LINEARLY DEPENDENT POWERS OF BINARY QUADRATIC FORMS

BRUCE REZNICK

Abstract. Given an integer $d \geq 2$, what is the least r so that there is a set of binary quadratic forms $\{f_1, \dots, f_r\}$ for which $\{f_j^d\}$ is non-trivially linearly dependent? We show that if $r \leq 4$, then $d \leq 5$, and for $d \geq 4$, construct such a set with $r = \lfloor d/2 \rfloor + 2$. Many explicit examples are given, along with techniques for producing others.

1. Introduction

For a fixed positive integer k, let $H_k(\mathbb{C}^2)$ denote the $(k+1)$-dimensional vector space of binary forms of degree k with complex coefficients. We say that two such forms are *distinct* if they are not proportional, and we say that a set $\mathcal{F} = \{f_1, \dots, f_r\} \subset H_k(\mathbb{C}^2)$ is *honest* if its elements are pairwise distinct. For $d \in \mathbb{N}$, let $\mathcal{F}^d = \{f_1^d, \dots, f_r^d\}$; if $\mathcal{F}$ is honest, then so is $\mathcal{F}^d$.

When $k = 1$, there is a simple classical criterion for the linear dependence of $\mathcal{F}^d$; see, e.g. [16, Thm.4.2].

Theorem 1.1. *If $\mathcal{F} = \{f_1, \dots, f_r\} \subset H_1(\mathbb{C}^2)$ is honest, then $\mathcal{F}^d = \{f_1^d, \dots, f_r^d\}$ is linearly independent if and only if $r \leq d + 1$.*

A version of this criterion is generally true for $k \geq 2$; see, e.g. [17, Thm.1.8]. (The proofs of these theorems are given at the start of section two.)

Theorem 1.2. *If $\mathcal{F} = \{f_1, \dots, f_r\} \subset H_k(\mathbb{C}^2)$, then it is generally true that $\mathcal{F}^d$ is linearly independent if and only if $r \leq kd + 1$.*

But there are singular cases, and these will be the focus of this paper. It is easy to find smaller values of r for which $\mathcal{F}^d$ is linearly dependent; for example, the Pythagorean parameterization gives three quadratics whose squares are dependent:

$$(x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2. \tag{1.1}$$

There are other ways of finding small dependent sets: let $\{g_j(x, y)\}$ be an honest set of $d+2$ linear forms, then both $\{g_j(x^k, y^k)\}$ and $\{\ell(x, y)^{k-1} g_j(x, y)\}$ (for a fixed linear form ℓ) will be dependent sets in $H_k(\mathbb{C}^2)$.

Date: March 27, 2019.

2000 *Mathematics Subject Classification.* Primary: 11E76, 11P05, 14M99; Secondary: 11D25, 11D41.

The author was supported by Simons Collaboration Grant 280987.

Given $r, d \in \mathbb{N}$, we say that an honest set of forms $\{f_1, \dots, f_r\} \subseteq H_k(\mathbb{C}^2)$ is a $\mathcal{W}_k(r,d)$*-set* if $\{f_j^d\}$ is linearly dependent. For example, (1.1) presents the $\mathcal{W}_2(3,2)$ set $\{x^2 - y^2, 2xy, x^2 + y^2\}$. Let $\Phi_k(d)$ denote the smallest r for which a $\mathcal{W}_k(r,d)$ set exists; clearly, $\Phi_k(d) \ge 3$. Theorem 1.1 implies that $\Phi_1(d) = d + 2$.

Our goal in this paper is two-fold. First, we give upper and lower bounds for $\Phi_k(d)$ for $k \ge 2$. Second, we describe all $\mathcal{W}_2(\Phi_2(d), d)$ sets for $d \le 5$. In (5) and (6) below, we use a peculiar-looking function. If $e \mid d$, let

$$\Theta_e(d) := 1 + \min_{t \in \mathbb{N}} (t \cdot \tfrac{d}{e} + \lfloor \tfrac{e}{t} \rfloor).$$

We summarize our main results.

Theorem 1.3 (Main Theorem)**.**

(1) $\Phi_{k+1}(d) \le \Phi_k(d)$.
(2) $\Phi_k(3) = 3$.
(3) *(Liouville)* $\Phi_k(d) \ge 4$ *for* $d \ge 3$ *and all* k.
(4) *(Hayman)* $\Phi_k(d) > 1 + \sqrt{d+1}$ *for* $d \ge 3$ *and all* k.
(5) *(Molluzzo-Newman-Slater)* $\Phi_d(d) \le \Theta_d(d) = 1 + \lfloor \sqrt{4d+1} \rfloor$.
(6) *If* $e \mid d$*, then* $\Phi_e(d) \le \min\{\Theta_k(d) : k \ge e,\ k \mid d\}$.
(7) $\Phi_k(d) = 4$ *for* $d = 3, 4, 5$ *and* $k \ge 2$.
(8) $\Phi_2(d) \ge 5$ *for* $d \ge 6$.
(9) $\Phi_2(d) = 5$ *for* $d = 6, 7$.
(10) $\Phi_2(14) \le 6$.
(11) $\Phi_2(d) \le \lfloor d/2 \rfloor + 2$ *for* $d \ge 4$.

All new parts of the Main Theorem except (8) and (11) have short proofs; these are given in section two. Examples give upper bounds for $\Phi_k(d)$; lower bounds are harder to find. The anomalous value in (10) for $d = 14$ is difficult to explain, and prevents us from conjecturing (11) as the exact value. This problem has been studied in [8] and [14] without the degree condition on the summands. The recent [13] contains a generalization of this question, replacing f_i^d with $\prod_j f_{ij}^{a_j}$ for fixed tuples (a_j).

If $\mathcal{F}$ is a $\mathcal{W}_k(r,d)$ set, then there is an obvious way to transform the linear dependence of the d-th powers into a more natural expression for any $m, 1 \le m \le r-1$:

$$\sum_{j=1}^{r} \lambda_j f_j^d = 0 \quad (\lambda_j \ne 0) \qquad \implies p = \sum_{j=1}^{m} \tilde{f}_j^d = \sum_{j=m+1}^{r} \tilde{f}_j^d, \tag{1.2}$$

where $\tilde{f}_j = (\pm\lambda_j)^{1/d} f_j$, for some p. In particular, a $\mathcal{W}_k(2m,d)$ set addresses the classical question of parameterizing two equal sums of m d-th powers. In this case, we say that (1.2) gives *two representations* of p as a sum of m d-th powers.

If $\alpha x + \beta y$ and $\gamma x + \delta y$ are distinct, then the map $M := (x,y) \mapsto (\alpha x + \beta y, \gamma x + \delta y)$ is an invertible change of variables (or *linear change* for short); let $(f \circ M)(x,y)$ denote $f(\alpha x + \beta y, \gamma x + \delta y)$. (This is a *scaling* if $\beta = \gamma = 0$.) If all members of $\mathcal{F}$ are subject to the same linear change, then the linear dependence of their d-th powers is unaltered. Any $\mathcal{W}_k(r,d)$ set can have its elements permuted and multiplied by various non-zero constants without essentially affecting the nature of the dependence.

So, suppose $\mathcal{F}$ is a $\mathcal{W}_k(r,d)$ set and

$$\sum_{j=1}^{r} \lambda_j f_j^d = 0. \tag{1.3}$$

If $\pi \in S_r$ is a permutation of $\{1,\dots r\}$, $c = (c_1,\dots,c_r) \in (\mathbb{C}\setminus\{0\})^r$, M is a linear change, and $g_j = c_j(f_{\pi(j)} \circ M)$, $1 \le j \le r$, then (1.3) is equivalent to

$$\sum_{j=1}^{r} (\lambda_{\pi(j)} \cdot c_j^{-d}) g_j^d = 0. \tag{1.4}$$

In this situation, we say that $\mathcal{F} = \{f_j\}$ and $\mathcal{G} = \{g_j\}$ (and the corresponding identities (1.3), (1.4)) are *cousins*. It is easy to show cousinhood by exhibiting M, π and c. Proving that $\mathcal{F}$ and $\mathcal{G}$ are *not* cousins may require *ad hoc* arguments.

We aim to present identities as symmetrically as possible, often guided by an old idea of Felix Klein. Associate to each non-zero linear form $\ell(x,y) = sx - ty$ the image of $t/s \in \mathbb{C}^*$ on the unit sphere S^2 under the Riemann map. (Assign $\ell(x,y) = y$ to ∞ and $(0,0,1)$.) Then associate to the binary form $\phi(x,y) = \prod_{j=1}^{k}(s_j x - t_j y)$ the image under the Riemann map of $\{t_j/s_j\}$, and call it the *Klein set of ϕ*. Given (1.3), we shall be interested in the Klein set of $\prod_{j=1}^{r} f_j$. In (1.1), the Klein set of $(x^2-y^2)(2xy)(x^2+y^2)$ is the regular octahedron with vertices $\{\pm e_k\}$.

Under the linear change M: $(x,y) \mapsto (\alpha x + \beta y, \gamma x + \delta y)$, the root $t/s \mapsto T(t/s)$, where T is the Möbius transformation $T(z) = \frac{\delta z - \beta}{-\gamma z + \alpha}$. Every rotation of the sphere corresponds to a Möbius transformation of the complex plane, and so a rotation of the Klein set can be effected by imposing a linear change on the forms. (Unfortunately, not every Möbius transformation gives a rotation.) It often happens that $p = \sum f_j^d$ and $p = p \circ M$, but $\sum (f_j \circ M)^d$ gives a different representation for p.

A trivial remark is surprisingly useful:

$$p = f_1^d + f_2^d = f_3^d + f_4^d \implies q = f_1^d - f_3^d = f_4^d - f_2^d$$

for suitable forms p, q; we call this a *flip*. For $k = 2$ and $d = 3, 4$, it can happen that q has a third representation as $q = f_5^d + f_6^d$, but that no such new expression exists for p. If $f_1^d + f_2^d = f_3^d + f_4^d$ and $g_1^d + g_2^d = g_3^d + g_4^d = g_5^d + g_6^d$, then $\mathcal{F} = \{f_1,\dots,f_4\}$ is a cousin of $\mathcal{G} = \{g_1,\dots,g_4\}$ and we say that $\mathcal{F}$ is a *sub-cousin* of $\mathcal{G}' = \{g_1,\dots,g_6\}$.

We now present some examples of small dependent sets of d-th powers. For integer $m \in \mathbb{N}$, let $\zeta_m = e^{\frac{2\pi i}{m}}$ be a primitive m-th root of unity, with the usual conventions that $\omega = \zeta_3$ and $i = \zeta_4$. A few interesting Klein sets will be noted.

The cubic identity with the simplest coefficients is probably

$$(x^2 + xy - y^2)^3 + (x^2 - xy - y^2)^3 = 2(x^2)^3 + 2(-y^2)^3 = 2x^6 - 2y^6. \tag{1.5}$$

The right-hand side of (1.5) is unchanged by the scalings $y \to \omega y$ and $y \to \omega^2 y$, so (1.5) shows that $2x^6 - 2y^6$ is a sum of two cubes in four different ways. Under the linear change $(x,y) \mapsto (\alpha+\beta, \alpha-\beta)$, (1.5) is due to Gérardin see [3, p.562] in 1910; in its present form, it was noted by Elkies in [1, p.542].

Here are two very simple quartic identities. The first generalizes to higher even degree; see (2.6), and the second is in $\mathbb{Z}[x,y]$:

$$(x^2+y^2)^4 + (\omega x^2 + \omega^2 y^2)^4 + (\omega^2 x^2 + \omega y^2)^4 = 18(xy)^4. \tag{1.6}$$

$$(x^2+2xy)^4 + (2xy+y^2)^4 + (x^2-y^2)^4 = 2(x^2+xy+y^2)^4. \tag{1.7}$$

These are cousins. Upon making the linear change $(x,y) \mapsto (i(x-\omega y), (x-\omega^2 y))$ and division by $\sqrt{-3}$, (1.6) becomes (1.7) up to a permutation of terms. The Klein set of (1.6) is a regular hexagon at the equator plus the poles.

A remarkable identity for $d=5$ was discovered independently by A. H. Desboves in 1880 (see [2], [3, p.684]) and N. Elkies in 1995 (see [1, p.542]):

$$\sum_{k=0}^{3} (-1)^k (i^k x^2 + \sqrt{-2}\ xy + i^{-k} y^2)^5 = 0. \tag{1.8}$$

The Klein set of (1.8) is the cube with vertices $\{(\pm\frac{\sqrt{2}}{\sqrt{3}}, 0, \pm\frac{1}{\sqrt{3}}), (0, \pm\frac{\sqrt{2}}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})\}$.

The next two examples appear to be new in detail, but are in the spirit of [15, §4]; the third explicitly appears there as (4.15); each is derived in section two:

$$\sum_{k=0}^{3} i^k (x^2 + i^k y^2)^6 = 80(xy)^6, \tag{1.9}$$

$$\sum_{k=0}^{3} \left(i^{-k} x^2 + \sqrt{-6/5}\ xy + i^k y^2\right)^7 = 26\sqrt{3} \cdot (-\sqrt{8/5}\ xy)^7, \tag{1.10}$$

$$\sum_{j=0}^{4} (\zeta_5^j x^2 + ixy + \zeta_5^{-j} y^2)^{14} = 5^7 (xy)^{14}. \tag{1.11}$$

The Klein set of (1.11) is the regular icosahedron, oriented so the vertices are the two poles plus two parallel regular pentagons at latitude $z = \pm\frac{1}{\sqrt{5}}$.

The second main focus of this paper is the characterization of $\mathcal{W}_2(\Phi_2(d), d)$ sets for $d = 3, 4, 5$. The characterization of $\mathcal{W}_k(3,2)$ is classical, and can be proved by emulating the standard analysis of $a^2 + b^2 = c^2$ over $\mathbb{N}$.

Theorem 1.4. *If $p, q, r \in \mathbb{C}[x_1, \dots, x_n]$, $n \ge 1$ and $p^2 + q^2 = r^2$, then then there exist $f, g, h \in \mathbb{C}[x_1, \dots, x_n]$ so that $p = f(g^2 - h^2), q = f(2gh), r = f(g^2 + h^2)$.*

The proof of the following theorem will be found in the companion paper [18].

Theorem 1.5. *Every $\mathcal{W}_2(4,3)$ set is a sub-cousin of a member of the $\mathcal{W}_2(6,3)$ family given below, for some $\alpha \neq 0, \pm 1$:*

$$\begin{aligned}
&(\alpha x^2 - xy + \alpha y^2)^3 + \alpha(-x^2 + \alpha xy - y^2)^3 \\
&= (\omega^2 \alpha x^2 - xy + \omega \alpha y^2)^3 + \alpha(-\omega^2 x^2 + \alpha xy - \omega y^2)^3 \\
&= (\omega \alpha x^2 - xy + \omega^2 \alpha y^2)^3 + \alpha(-\omega x^2 + \alpha xy - \omega^2 y^2)^3 \\
&= (\alpha^2 - 1)(\alpha x^3 + y^3)(x^3 + \alpha y^3).
\end{aligned} \tag{1.12}$$

If the first two lines of (1.12) are read as $f_1^3 + f_2^3 = f_3^3 + f_4^3$, then $f_1^3 - f_4^3 = f_3^3 - f_2^3$ also has a third representation as a sum of two cubes, but $f_1^3 - f_3^3 = f_4^3 - f_2^3$ does not.

(Put $(\alpha, x, y) \mapsto (i, \zeta_8^3 x, \zeta_8^5 y)$ in the first line of (1.12) to get (1.5).) After the linear change: $(x, y) \mapsto (ix + \sqrt{3}\, y, ix - \sqrt{3}\, y)$, (1.12) becomes

$$\begin{aligned}
&((1-2\alpha)x^2 + 3(1+2\alpha)y^2)^3 + \alpha((2-\alpha)x^2 - 3(2+\alpha)y^2)^3 \\
&= ((1+\alpha)x^2 + 6\alpha xy + 3(1-\alpha)y^2)^3 + \alpha(-(1+\alpha)x^2 - 6xy + 3(1-\alpha)y^2)^3 \\
&= ((1+\alpha)x^2 - 6\alpha xy + 3(1-\alpha)y^2)^3 + \alpha(-(1+\alpha)x^2 + 6xy + 3(1-\alpha)y^2)^3.
\end{aligned} \tag{1.13}$$

If $\alpha \in \mathbb{Q}$, then all forms in (1.13) are in $\mathbb{Q}[x, y]$, and if α is a rational cube, then (1.13) gives solutions to $f_1^3 + f_2^3 = f_3^3 + f_4^3$ in $\mathbb{Q}[x, y]$. Historically, these were used to parameterize solutions to the Diophantine equations $a^3 + b^3 = c^3 + d^3$ over $\mathbb{N}$.

Theorem 1.6. *Every $\mathcal{W}_2(4,4)$ is a cousin of (1.6) or a sub-cousin of (1.14):*

$$\begin{aligned}
&(x^2 + \sqrt{3}\, xy - y^2)^4 - (x^2 - \sqrt{3}\, xy - y^2)^4 \\
&= (\omega^2 x^2 + \sqrt{3}\, xy - \omega y^2)^4 - (\omega^2 x^2 - \sqrt{3}\, xy - \omega y^2)^4 \\
&= (\omega\, x^2 + \sqrt{3}\, xy - \omega^2 y^2)^4 - (\omega\, x^2 - \sqrt{3}\, xy - \omega^2 y^2)^4 \\
&= 8\sqrt{3}\, xy\, (x^6 - y^6).
\end{aligned} \tag{1.14}$$

In an earlier version of this work (see e.g. [15, (3.9)]), the identity

$$\begin{aligned}
&(\sqrt{3}\, x^2 + \sqrt{2}\, xy - \sqrt{3}\, y^2)^4 + (\sqrt{3}\, x^2 - \sqrt{2}\, xy - \sqrt{3}\, y^2)^4 \\
&= (\sqrt{3}\, x^2 + i\sqrt{2}\, xy + \sqrt{3}\, y^2)^4 + (\sqrt{3}\, x^2 - i\sqrt{2}\, xy + \sqrt{3}\, y^2)^4 \\
&= 18x^8 - 28x^4y^4 + 18y^8.
\end{aligned} \tag{1.15}$$

was given as an alternative in Theorem 1.6; (1.15) turns out to be a sub-cousin of (1.14), see Theorem 3.4. When scaled, (1.15) appears in Desboves [2, p.243]. The set in (1.6) is not a sub-cousin of (1.14): three of the quadratics in (1.6) are linearly dependent, and no three quadratics in (1.14) are dependent.

The situation for quintics is simpler.

Theorem 1.7. *Every $\mathcal{W}_2(4,5)$ set is a cousin of (1.8).*

Here is an outline of the rest of the paper. In section two, we prove Theorems 1.1 and 1.2 and Theorem 1.3 except (8). We also recall "synching" from [15] as a tool for finding "good" $\mathcal{W}_k(r,d)'s$ – the idea was inspired by a formula of Molluzzo [12] – and use it to prove several parts of Theorem 1.3.

In section three, we recall two results familiar to 19th century algebraists: a specialization of Sylvester's algorithm for determining the sums of two d-th powers of linear forms and a result on the simultaneous diagonalization of quadratic forms. We use these to lay out our strategy for proving Theorem 1.3(8). Suppose

$$p(x,y) = f_1^d(x,y) + f_2^d(x,y) = f_3^d(x,y) + f_4^d(x,y)$$

for an honest set $\{f_1, f_2, f_3, f_4\}$ of quadratics. There is a linear change which simultaneously diagonalizes f_1 and f_2 (making p even), but neither f_3 nor f_4 is even. We then make a systematic study of non-even $\{f_3, f_4\}$ for which $p = f_3^d + f_4^d$ is even, and check back to see whether p can be written as $f_1^d + f_2^d$. For $d \geq 3$, a shorter method can be used to prove Theorem 1.5; see the companion paper [18].

Section four is devoted to implementing in detail the strategy outlined above; this simultaneously proves Theorems 1.6 and 1.7, as well as Theorem 1.3(8). The proofs of Theorems 4.1 and 4.3 contain a great deal of "equation wrangling"; however, the reader should know that this has been greatly condensed from earlier drafts.

In section five, we do a brief review of the literature in the subject and derive the examples for for $d \leq 5$ via *a priori* constructions. We also discuss how Newton's theorem on symmetric forms helps explain (1.11), similar to the argument for (1.8) given in [15]. Corollaries 5.2 and 5.3 present the classification of forms which can be written as a sum of two d-th powers of quadratic forms and, for $d \geq 4$, those which have more than one representation. We suggest some further areas of exploration and finish with Conjecture 5.4 about the true growth of $\Phi_k(d)$.

The author has been working on this project for a very long time; online seminar notes [19] are dated 2000. He wishes to thank Andrew Bremner, Noam Elkies, Jordan Ellenberg, Andrew Granville, Samuel Lundqvist, Cordian Rainer and Boris Shapiro for encouraging conversations and useful emails, even if after all this time, they don't remember what they said. Many thanks to Becky Burner of the Illinois Mathematics Library for finding an online copy of [2].

2. Some proofs, and synching

We begin with proofs of Theorems 1.1 and 1.2.

Proof of Theorem 1.1. If $r > d + 1 = \dim(H_d(\mathbb{C}^2))$, then $\mathcal{F}^d$ is dependent. Suppose $r \leq d+1$ and let $f_i(x,y) = \alpha_i x + \beta_i y$. Define (if necessary) distinct f_j for $r+1 \leq j \leq d+1$ by $(\alpha_j, \beta_j) = (1, m_j)$, where $m_j \alpha_i \neq \beta_i, 1 \leq i \leq r$, and express $\{f_1^d, \dots, f_{d+1}^d\}$ in terms of the basis $\{\binom{d}{v} x^{d-v} y^v\}$. The resulting $(d+1) \times (d+1)$ matrix, $[\alpha_i^{d-v} \beta_i^v]$, is Vandermonde, with determinant $\prod_{1 \leq i < j \leq d+1} (\alpha_i \beta_j - \alpha_j \beta_i) \neq 0$, since $\mathcal{F}$ is honest. $\square$

Proof of Theorem 1.2. Again, if $r > kd+1$, then $\mathcal{F}^d$ is linearly dependent by dimension. Suppose $f_j(x,y) = \sum_{\ell=0}^{k}\binom{k}{\ell}\alpha_{\ell,j}x^{k-\ell}y^\ell$. If $r < kd+1$, again add pairwise distinct elements and assume that $r = kd+1$. Express $\{f_j^d\}$ in terms of the monomial basis $\{\binom{kd}{v}x^{kd-v}y^v\}$, obtaining a square matrix of order $kd+1$ whose entries are polynomials in the variables $\{\alpha_{\ell,j}\}$, and whose determinant is a polynomial $P(\{\alpha_{\ell,j}\})$. If we specialize to $f_j(x,y) = (x+jy)^k$, $1 \le j \le kd+1$, then $\alpha_{\ell,j} = j^\ell$, and $\mathcal{F}^d = \mathcal{G}^{kd}$ for $\mathcal{G} = \{x+jy\}$. By Theorem 1.1, $\mathcal{G}^{kd}$ is linearly independent, hence $P(\{j^\ell\}) \ne 0$, and so P is not identically zero. That is, $\mathcal{F}^d$, generally, is linearly independent. □

We defer the proofs of Theorem 1.3(5), (6) and (11) until we have defined synching; (8) will require sections three and four.

Partial Proof of Theorem 1.3.

(1) If $g_j(x,y) = xf_j(x,y)$, then $\sum \lambda_j f_j^d = 0 \implies \sum \lambda_j g_j^d = 0$.

(2) This follows from (1.1) and (1).

(3) As noted in (1.2), the existence of a $\mathcal{W}_k(3,d)$ set for $d \ge 3$ would imply the existence of a nontrivial identity

$$f_1^d(x,y) + f_2^d(x,y) = f_3^d(x,y).$$

After a linear change, we may assume that $f_j(x,y)$ is not a multiple of y^k. Let $p_j(t) = f_j(t,1)$. Then $p_1^d(t) + p_2^d(t) = p_3^d(t)$, where the p_j's are non-constant polynomials. In 1879, Liouville proved that the Fermat equation $X^d + Y^d = Z^d$ has no non-constant solutions over $\mathbb{C}[t]$ for $d \ge 3$. (See [20, pp.263–265] for a proof.)

(4) More generally, the elements of any $\mathcal{W}_k(r,d)$ set can be scaled as in (1.2) so that $\sum_{j=1}^{r-1} f_j^d(x,y) = f_r^d(x,y)$. Once again, by letting $p_j(t) = f_j(t,1)$ and $q_j(t) = f_j(t)/f_r(t)$ we obtain a set of $r-1$ rational functions so that $\sum_{j=1}^{r-1} q_j^d(t) = 1$. A 1984 theorem of Hayman [10] says that if $\{\phi_j\}$, $1 \le j \le r-1$, are $r-1$ holomorphic functions in n complex variables, no two of which are proportional, and $\sum_{j=1}^{r-1}\phi_j^d = 1$, then $d < (r-1)^2 - 1$, so $r > 1 + \sqrt{d+1}$. This was culmination of the work of Green [6] and others; see [8, pp.438-440] for a clear exposition and history.

(7) The equality for $k = 2$ follows from combining (3) with the equations (1.5), (1.6) and (1.8); for $k \ge 3$, apply (1).

(9) Subject to the as-yet unproved (8), this follows from (1.9) and (1.10).

(10) This follows from (1.11). □

Recall that for an integer $m \ge 2$ and for $s \in \mathbb{Z}$,

$$\frac{1}{m}\sum_{j=0}^{m-1}\zeta_m^{sj} = \begin{cases} 0 & \text{if } m \nmid s, \\ 1 & \text{if } m \mid s. \end{cases} \tag{2.1}$$

Synching was introduced in [15, §4] and is a generalization of the familiar formulas in which $\frac{1}{2}(f(x,y) \pm f(x,-y))$ give the even and odd parts of f.

Theorem 2.1. *Suppose* $p(x,y) = \sum_{i=0}^{k} a_i x^{k-i} y^i \in H_k(\mathbb{C}^2)$. *Then*

$$\frac{1}{m}\sum_{j=0}^{m-1} \zeta_m^{-rj} p(x, \zeta_m^j y) = \sum_{\substack{i \equiv r \pmod m, \\ 0 \le i \le k}} a_i x^{k-i} y^i. \tag{2.2}$$

Proof. We expand the left-hand side of (2.2), switch the order of summation:

$$\frac{1}{m}\sum_{j=0}^{m-1} \zeta_m^{-rj} p(x, \zeta_m^j\ y) = \sum_{i=0}^{k} \left(\frac{1}{m} \sum_{j=0}^{m-1} \zeta_m^{-rj}\zeta_m^{ij} \right) a_i x^{k-i} y^i,$$

and then apply (2.1) to the inner sum of $\zeta_m^{(i-r)j}$. □

In our applications, $p = f^d$; for example, if $p(x,y) = (x+\alpha y)^d$, then:

$$\frac{1}{m}\sum_{j=0}^{m-1} \zeta_m^{-rj} (x + \zeta_m^j \alpha y)^d = \sum_{-\frac{r}{m} \le i \le \frac{d-r}{m}} \binom{d}{r+im} \alpha^{r+im} x^{d-r-im} y^{r+im}. \tag{2.3}$$

Proof of Theorem 1.3(5), (6). We generalize an identity found in Molluzzo's thesis [12] (with $\ell = d$) and discussed in [14, p.485]; it follows from (2.3) with $r = 0$ that

$$\sum_{j=0}^{m-1} (x^\ell + \zeta_m^j y^\ell)^d = m \sum_{i=0}^{\lfloor d/m \rfloor} \binom{d}{im} x^{\ell d - im\ell} y^{im\ell}. \tag{2.4}$$

Suppose now that $d = ee'$, $\ell = e$ and $m = te'$ is a multiple of e'. Then the left-hand side of (2.4) is a sum of m d-th powers, and since $d \mid im\ell = itd$, the right-hand side is a sum of $1 + \lfloor d/m \rfloor$ d-th powers. Thus the total number of summands is $1 + t \cdot \frac{d}{e} + \lfloor \frac{e}{t} \rfloor$. We choose t to minimize this sum, obtaining $\Theta_e(d)$.

Newman and Slater took $d = e$, so $e' = 1$ ([14, p.485]); the minimum in $\Theta_d(d)$ is found by choosing $m \in \{\lfloor \sqrt{d} \rfloor, 1 + \lfloor \sqrt{d} \rfloor\}$, giving $\Phi_d(d) = 1 + \lfloor \sqrt{4d+1} \rfloor$.

If $e < d$, then $\Theta_e(d)$ is generally larger than $\Theta_d(d)$, since some m's are skipped in computing the minimum; however, $\Theta_e(d)$ need not be monotone in e, so Theorem 1.3(1) need not be be implemented. □

The first instance of non-monotonicity in $\Theta_e(d)$ occurs at $d = 72$; in general, $\Theta_{8n}(72n^2) = \Theta_{9n}(72n^2) = 1 + 17n$, but $\Theta_{12n}(72n^2) = 1 + 18n$. This suggests interesting questions in combinatorial number theory which we hope to pursue elsewhere.

When d is even, a specialization of (2.3) can be made more symmetric:

Corollary 2.2.

$$\frac{1}{s+1} \cdot \sum_{j=0}^{s} (\zeta_{2s+2}^{-j} x + \zeta_{2s+2}^{j} y)^{2s} = \binom{2s}{s} x^s y^s. \tag{2.5}$$

Proof. Set $r = s$, $d = 2s$ and $m = s+1$ in (2.3). Since $|\frac{r}{m}| = |\frac{d-r}{m}| < 1$, the summation on the right-hand side has a single term, $i = 0$, and (2.3) becomes

$$\frac{1}{s+1} \cdot \sum_{j=0}^{s} \zeta_{s+1}^{-sj}(x + \zeta_{s+1}^{j} y)^{2s} = \binom{2s}{s} x^s y^s;$$

(2.5) follows from $\zeta_{s+1}^{-sj}(x + \zeta_{s+1}^{j} y)^{2s} = \zeta_{2s+2}^{-2sj}(x + \zeta_{2s+2}^{2j} y)^{2s} = (\zeta_{2s+2}^{-j} x + \zeta_{2s+2}^{j} y)^{2s}$. □

Proof of Theorem 1.3(11) for even d. Take $(x, y) \mapsto (x^2, y^2)$ in (2.5), to obtain

$$\sum_{j=0}^{s} (\zeta_{2s+2}^{-j} x^2 + \zeta_{2s+2}^{j} y^2)^{2s} = (s+1)\binom{2s}{s}(xy)^{2s}, \tag{2.6}$$

a linear dependence among $s+2$ $2s$-th powers of an honest set of quadratic forms. □

If $s = 2v$, we have $(\zeta_{4v+2}^{-j}, \zeta_{4v+2}^{-j}) = ((-\zeta_{2v+1}^{v})^j, (-\zeta_{2v+1}^{v+1})^j)$, so

$$\sum_{j=0}^{2v} ((\zeta_{2v+1}^{v})^j x^2 + (\zeta_{2v+1}^{v+1})^j y^2)^{4v} = (2v+1)\binom{4v}{2v}(xy)^{4v}. \tag{2.7}$$

When $s = 1$, we have $\zeta_2 = -1$ and (2.7) reduces to (1.1); when $s = 2$ and 3, (2.7) becomes (1.6) and (1.9). Taking $(x, y) \mapsto (e^{-i\theta}(x + iy), e^{i\theta}(x - iy))$ in (2.5) (see [16, (5.8)], which is incorrect – unfortunately missing the factor of 2^{-2s}) gives

$$\frac{1}{s+1}\sum_{j=0}^{s}\Big(\cos(\tfrac{j\pi}{s+1} + \theta)x + \sin(\tfrac{j\pi}{s+1} + \theta)y\Big)^{2s} = \frac{1}{2^{2s}}\binom{2s}{s}(x^2 + y^2)^s, \quad \theta \in \mathbb{C}. \tag{2.8}$$

With $\theta \in \mathbb{R}$, (2.8) was a 19th century quadrature formula; see the discussion after [16, Cor.5.6] for details. Taking $\theta \in \mathbb{R}$ and $(x, y) \mapsto (x^2 - y^2, 2xy)$, so that $x^2 + y^2 \mapsto (x^2 + y^2)^2$ in (2.8), gives a nice family of $\mathcal{W}_2(s+2, 2s)$ cousins in $\mathbb{R}[x, y]$.

There doesn't seem to be such a simple proof of Theorem 1.3(11) for odd d, and we need to introduce powers of trinomials as summands. More generally, it is useful to present two quadratic cases, which are corollaries of Theorem 2.1; note that

$$\zeta_m^{-rj}(\zeta_m^{-j} x^2 + \alpha xy + \zeta_m^{j} y^2)^d = \zeta_m^{-(r+d)j}(x^2 + \alpha\zeta_m^{j} xy + \zeta_m^{2j} y^2)^d,$$

gives (2.9) the shape of Theorem 2.1 for $p(x, y) = (x^2 + \alpha xy + y^2)^d$.

Corollary 2.3. *Suppose $d, m \in \mathbb{N}, v \in \mathbb{Z}$ and $\alpha \in \mathbb{C}$. Let*

$$\Psi(v, m, d; \alpha) := \frac{1}{m}\sum_{j=0}^{m-1} \zeta_m^{-vj}(\zeta_m^{-j} x^2 + \alpha xy + \zeta_m^{j} y^2)^d. \tag{2.9}$$

(i) If $m > d$, then

$$\Psi(0, m, d; \alpha) = \left(\sum_{r=0}^{\lfloor d/2 \rfloor} \frac{d!}{(r!)^2(d-2r)!}\alpha^{d-2r}\right) x^d y^d. \tag{2.10}$$

(ii) If $2m > d \geq m$, *then*

$$\Psi(0,m,d;\alpha) = \left(\sum_{r=0}^{\lfloor d/2\rfloor} \frac{d!}{(r!)^2(d-2r)!}\alpha^{d-2r}\right) x^d y^d \tag{2.11}$$
$$+ \left(\sum_{r=0}^{\lfloor (d-m)/2\rfloor} \frac{d!}{r!(r+m)!(d-m-2r)!}\alpha^{d-m-2r}\right)(x^{d+m}y^{d-m} + x^{d-m}y^{d+m}).$$

Proof. By the trinomial theorem,

$$(x^2+\alpha xy+y^2)^d = \sum_{r+s+t=d} \frac{d!}{r!s!t!}\alpha^s x^{2r+s}y^{s+2t};$$

note that $(2r+s, s+2t) = (2d-i, i) \iff r-t = d-i$; all sums can only be taken over $r,s,t \geq 0$. In each case, m is relatively large compared to d and very few terms will be nonzero. In (i), $x^{2d-i}y^i$ appears when $i \equiv d \pmod m$. Since $d < m$, this only occurs when $i = d$, so $r = t$ and the coefficient of $x^d y^d$ is found by summing $\frac{d!}{r!s!t!}\alpha^s$ over the set $\{(r,s,t) = (r, d-2r, r)\}$. Similarly, in (ii), $v = 0$ and $2m > d$, so we have three cases: $r - t \in \{-m, 0, m\}$, and the terms sum as indicated. □

We use (2.11) when $d - m \geq 2$ by choosing $\alpha = \alpha_0$ to be a non-zero root of the polynomial coefficient of $(x^{d+m}y^{d-m} + x^{d-m}y^{d+m})$, so that the terms on both sides of the expression are d-th powers. In general, the Klein set of $\Psi(v,m,d;\alpha)$ will consist of two parallel regular m-gons, whose altitude and relative orientation depends on α. If $(xy)^d$ appears in the identity, then the two poles are added.

Proof of Theorem 1.3(11) for odd d. Suppose $d = 2s+1 \geq 5$. We have

$$\Psi(0, s+1, 2s+1;\alpha) = \sum_{j=0}^{s}(\zeta_{s+1}^{-j}x^2 + \alpha xy + \zeta_{s+1}^{j}y^2)^{2s+1} = \tag{2.12}$$
$$A_s(\alpha)x^{3s+2}y^s + B_s(\alpha)x^{2s+1}y^{2s+1} + A_s(\alpha)x^s y^{3s+2};$$
$$A_s(\alpha) = \binom{2s+1}{s}\alpha^s + (2s+1)\binom{2s}{s-2}\alpha^{s-2} + \dots.$$

Let $\alpha = \alpha_0$ be a non-zero root of $A_s(\alpha)$; this exists because $s \geq 2$, so (2.12) becomes

$$\Psi(0, s+1, 2s+1;\alpha_0) = B(\alpha_0)(xy)^{2s+1},$$

which is a sum of $s+1$ $(2s+1)$-st powers equal to another $(2s+1)$-st power. □

Alternate Proof of Theorem 1.3(11) for $d = 2s, s \geq 3$. Suppose $s \geq 3$. Then

$$\Psi(0, s+1, 2s;\alpha) = \tilde{A}_s(\alpha)(x^{3s+1}y^{s-1} + x^{s-1}y^{3s+1}) + \tilde{B}_s(\alpha)x^{2s}y^{2s}; \tag{2.13}$$
$$\tilde{A}_s(\alpha) = \binom{2s}{s-1}\alpha^{s-1} + (2s)\binom{2s-1}{s-3}\alpha^{s-3} + \dots$$

Again, choose $\alpha = \alpha_0$ to be a non-zero root of $\tilde{A}_s$. □

By looking at the pattern of linear dependence among the elements, it is not hard to show that the families in (2.6) and (2.13) are not cousins.

Here are other synching examples; (2.10) requires $m > d$. We have $\Psi(0,4,3;\alpha) = (\alpha^3+6\alpha)x^3y^3$, so $\Psi(0,4,3,\sqrt{-6})$ gives a $\mathcal{W}_2(4,3)$ set. In (ii) we need $d \in [m+2, 2m)$. For $m = 3$, this implies that $d = 5$, and we obtain a variant of [15, (4.12)]:

$$\begin{aligned} 3\Psi(0,3,5;\alpha) &= \sum_{j=0}^{2}(\omega^k x^2 + \alpha xy + \omega^{-k}y^2)^5 \\ &= 15(1+2\alpha^2)(x^8y^2+x^2y^8) + 3\alpha(\alpha^4+20\alpha^2+30)x^5y^5 \implies \\ \Psi\left(0,3,5;\sqrt{-1/2}\right) &= \left(\sqrt{-9/2}\;xy\right)^5. \end{aligned} \tag{2.14}$$

The linear change $(x,y) \mapsto (\sqrt{-2}\;x - (1+\sqrt{3})y, -(1+\sqrt{3})x + \sqrt{-2}\;y)$, applied to (2.14), gives $3(1+\sqrt{3})$ times a flip of (1.8). The Klein set here is again a cube, rotated so the vertices are the two poles and antipodal equilateral triangles at $z = \pm\frac{1}{3}$.

For $m = 4$, the possibilities are $d = 6, 7$; we have

$$4\Psi(0,4,6;\sqrt{-2/5}) = \sum_{k=0}^{3}(i^{-k}x^2 + \sqrt{-2/5}\;xy + i^k y^2)^6 = 11\cdot(\sqrt{-8/5}\;xy)^6;$$

$\Psi(0,4,7;\sqrt{-6/5})$ is just (1.10).

Two other examples show the range of Corollary 2.3. First,

$$4\Psi(2,4,4;\alpha) = \sum_{j=0}^{3}(-1)^k(i^{-k}x^2 + \alpha xy + i^k y^2)^4 = 8(2+3\alpha^2)(x^6y^2 + x^2y^6).$$

On taking $\alpha = \alpha_0 = \sqrt{-2/3}$, transposing two terms to get two equal sums of two fourth powers, and after multiplying through by $\sqrt{3}$, we obtain (1.15). For $d = 5$, we may recover (1.8) as $4\Psi(2,4,5,\sqrt{-2})$ from

$$4\Psi(2,4,5;\alpha) = \sum_{j=0}^{3}(-1)^k(i^{-k}x^2 + \alpha xy + i^k y^2)^5 = 40\alpha(2+\alpha^2)(x^7y^3 + x^3y^7).$$

An unusual phenomenon occurs with $\Psi(0,5,14;\alpha)$: by the general method,

$$\Psi(0,5,14;\alpha) = A(\alpha)(x^{24}y^4 + x^4y^{24}) + B(\alpha)(x^{19}y^9 + x^9y^{19}) + C(\alpha)x^{14}y^{14}.$$

It turns out that $A(\alpha)$ and $B(\alpha)$ have a common factor $1+\alpha^2$. Upon setting $\alpha = i$, we obtain (1.11). A computer search has not found other examples of this phenomenon. As noted earlier, the Klein form of (1.11) is an icosahedron, but an icosahedron can be rotated so that its vertices lie in four horizontal equilateral triangles. This suggests that (1.8) should be the cousin of a union of two $\Psi(v,3,14;\alpha)$'s. Indeed, with $\phi = \frac{1+\sqrt{5}}{2}$ as usual,

$$\sum_{k=0}^{2}(\omega^k x^2 + \phi^2 xy - \omega^{-k}y^2)^{14} + \sum_{k=0}^{2}(\omega^k \phi x^2 - \phi^{-1}xy - \omega^{-k}\phi y^2)^{14} = 0. \tag{2.15}$$

The Schönemann coefficients of the icosahedron, $\{(\phi^2+1)^{-1/2}\cdot(\pm\phi,\pm 1,0)\}$ and their cyclic images, lead to yet another cousin of (1.8):

$$\begin{aligned}(x^2 + 2\phi xy - y^2)^{14} + (x^2 - 2\phi xy - y^2)^{14} + ((\phi+i)(x^2 - \tfrac{1-2i}{\sqrt{5}}y^2))^{14} \\ +((\phi-i)(x^2 - \tfrac{1+2i}{\sqrt{5}}y^2))^{14} = (\phi x^2 + 2ixy + \phi y^2)^{14} + (\phi x^2 - 2ixy + \phi y^2)^{14}.\end{aligned} \tag{2.16}$$

The corresponding quadratics for a dodecahedron, alas, give a $\mathcal{W}_2(10,14)$ set.

There is no reason for synching to be limited to trinomials. Here is an example of a $\mathcal{W}_4(4,3)$ family of linearly independent elements:

$$\sum_{k=0}^{3}(-1)^k(x^4 + i^k\sqrt{6}\,x^3y - 6i^{2k}x^2y^2 - \sqrt{6}\,i^{3k}xy^3 + y^4)^3 = 0; \tag{2.17}$$

the quartics are linearly independent.

Finally, we compare Theorem 1.3(5), (6) and (11). The bound in (11) is linear in d and weaker than (5). This leads to the natural question: what is the smallest d so that $k \ge 2$ and $\Phi_{k+1}(d) < \Phi_k(d)$? Taking Theorem 1.3(7), (10) and (11), into account, we must have $d \ge 6$, and the smallest d for which (5) or (6) beats the bound for $k=2$ in (11) is $d = 15$: $1 + \lfloor\sqrt{61}\rfloor = 8 < 9 = 2 + \lfloor 15/2 \rfloor$.

3. Overview of $\mathcal{W}_2(4,d)$ sets and tools.

In order to prove Theorem 1.3(8), we need an abbreviated version of Sylvester's algorithmic theorem from 1851 on the representation of forms as a sum of powers of linear forms. We refer the reader to [16, Thm.2.1] for the general theorem and proof.

Theorem 3.1 (After Sylvester)**.** *Suppose $d \ge 3$ and*

$$p(x,y) = \sum_{j=0}^{d}\binom{d}{j}a_j x^{2d-2j}y^{2j}, \qquad q(x,y) = \sum_{j=0}^{d}\binom{d}{j}a_j x^{d-j}y^j. \tag{3.1}$$

Then p is a sum of d-th powers of two honest even quadratic forms if and only if there exists a non-square quadratic form $h(u,v) = c_0u^2 + c_1uv + c_2v^2 \ne 0$ so that

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ \vdots & \vdots & \vdots \\ a_{d-2} & a_{d-1} & a_d \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3.2}$$

Sketch of Proof. A comparison of the coefficients of monomials in p and q shows that

$$p(x,y) = (\alpha_1 x^2 + \beta_1 y^2)^d + (\alpha_2 x^2 + \beta_2 y^2)^d \iff$$
$$q(x,y) = (\alpha_1 x + \beta_1 y)^d + (\alpha_2 x + \beta_2 y)^d.$$

Assuming $\alpha_j \neq 0$, $q(x,y) = (\alpha_1 x+\beta_1 y)^d + (\alpha_2 x+\beta_2 y)^d$ implies that $a_j = \lambda_1\gamma_1^j + \lambda_2\gamma_2^j$, where $\lambda_i = \alpha_i^d$ and $\gamma_i = \beta_i/\alpha_i$, so (a_j) satisfies the linear recurrence given by (3.2) with $c_0 = \gamma_1\gamma_2, c_1 = -(\gamma_1+\gamma_2), c_2 = 1$; $h(u,v) = (\gamma_1 u - v)(\gamma_2 u - v)$. Conversely, any solution (a_j) to this recurrence has the indicated shape. If $\alpha_2 = 0$, then $\alpha_1 \neq 0$ by honesty; $a_j = \lambda_1\gamma_1^j$ for $j \le d-1$ and (3.2) holds with $h(u,v) = u(\gamma_1 u - v)$. □

The matrix in (3.2) is called the *2-Sylvester matrix* for p (or q). A necessary condition for p to be a sum of two d-th powers is that the 2-Sylvester matrix of p (with $d-2$ rows) has rank ≤ 2. As d increases, this becomes increasingly harder.

We also need a special case of a classical result about simultaneous diagonalization; there doesn't seem to be an easy-to-find modern proof.

Theorem 3.2 (Diagonalization). *If f_1 and f_2 are relatively prime binary quadratic forms, then there is a linear change M so that $f_1 \circ M$ and $f_2 \circ M$ are both even.*

Proof. Suppose without loss of generality that $rank(f_1) \ge rank(f_2)$. If $rank(f_1) = 1$, then $(f_1, f_2) = (\ell_1^2, \ell_2^2)$ and a linear change takes $(\ell_1, \ell_2) \mapsto (x, y)$. Otherwise, there exists M_1 so that $(f_1 \circ M_1)(x,y) = x^2 + y^2$ and $(f_2 \circ M_1)(x,y) = ax^2 + bxy + cy^2$. Since these are relatively prime, $a \pm ib - c \neq 0$.

Drop "M_1", and observe that for any $z \in \mathbb{C}$, f_1 is fixed by any orthogonal linear change $M_z : (x,y) \mapsto ((\cos z)x + (\sin z)y, -(\sin z)x + (\cos z)y)$, under which the coefficient of xy in $f_2 \circ M_z$ is $(a-c)\sin 2z + b\cos 2z$. If $a = c$, let $z = \frac{\pi}{4}$. Otherwise, choose z so that $\tan 2z = -\frac{b}{a-c}$; this is possible, since the range of $tan(z)$ is $\mathbb{C} \setminus \{\pm i\}$. The coefficient of xy in $f_2 \circ M_z$ vanishes, so $f_1 \circ M_z, f_2 \circ M_z$ are both even. □

Suppose $d \ge 3$ and we have a $\mathcal{W}_2(4,d)$ set, flipped and normalized so that

$$p(x,y) = f_1^d(x,y) + f_2^d(x,y) = f_3^d(x,y) + f_4^d(x,y), \tag{3.3}$$

for an honest set $\{f_1, f_2, f_3, f_4\}$ of binary quadratic forms.

Theorem 3.3. *If* (3.3) *holds, then there exists a linear change after which both f_1 and f_2 are even, so p is even. We have* $\gcd(f_1, f_2) = \gcd(f_3, f_4) = 1$*, but it is not true that f_3, f_4 are both even.*

Proof. If $\gcd(f_1, f_2) = \ell$ for a linear form ℓ, so that $f_1 = \ell\ell_1$ and $f_2 = \ell\ell_2$, then

$$\ell^d \mid f_3^d + f_4^d = \prod_{k=0}^{d-1}(f_3 + \zeta_d^k f_4).$$

Since $d \ge 3$, ℓ must divide at least two different quadratic factors on the right, say $\ell \mid f_3 + \zeta_d^{k_1} f_4, f_3 + \zeta_d^{k_2} f_4$ for $k_1 \neq k_2$. This implies that $\ell \mid f_3, f_4$ and $f_3 = \ell\ell_3$ and

$f_4 = \ell\ell_4$ for linear ℓ_3, ℓ_4. hence we can factor ℓ^d from (3.3) to obtain $\ell_1^d + \ell_2^d = \ell_3^d + \ell_4^d$, which contradicts Theorem 1.1, since $d \ge 3$. Similarly, $\gcd(f_3, f_4) = 1$.

Thus f_1 and f_2 are relatively prime, and by Theorem 3.2, we may simultaneously diagonalize them, after which (dropping M),

$$p(x,y) = (\alpha_1 x^2 + \beta_1 y^2)^d + (\alpha_2 x^2 + \beta_2 y^2)^d = f_3^d(x,y) + f_4^d(x,y).$$

Suppose $f_3(x,y) = \alpha_3 x^2 + \beta_3 y^2$ and $f_4(x,y) = \alpha_4 x^2 + \beta_4 y^2$ are both even. Then

$$\begin{aligned} &(\alpha_1 x^2 + \beta_1 y^2)^d + (\alpha_2 x^2 + \beta_2 y^2)^d = (\alpha_3 x^2 + \beta_3 y^2)^d + (\alpha_4 x^2 + \beta_4 y^2)^d \\ &\implies (\alpha_1 x + \beta_1 y)^d + (\alpha_2 x + \beta_2 y)^d = (\alpha_3 x + \beta_3 y)^d + (\alpha_4 x + \beta_4 y)^d. \end{aligned} \tag{3.4}$$

Since $\{f_j\}$ is honest, (3.4) violates Theorem 1.1, so f_3 and f_4 are not both even. □

Here then is our strategy. We seek to find all pairs $\{f_3, f_4\}$ which are not both even but for which $f_3^d + f_4^d$ *is* even. Then, from among those, we need to find those which can *also* be written as a sum of two d-th powers of even quadratic forms.

How can it happen that $f_3^d + f_4^d$ is even when at least one of $\{f_3, f_4\}$ is not even? Two cases come readily to mind:

$$(ax^2 + bxy + cy^2)^d + (ax^2 - bxy + cy^2)^d, \tag{3.5}$$

and, if d is even,

$$(ax^2 + cy^2)^d + b(xy)^d. \tag{3.6}$$

We call (3.5) and (3.6) the *tame* cases; otherwise $\{f_3, f_4\}$ are in the *wild* case. There is an important practical distinction. The tame expressions are formally symmetric under $y \mapsto -y$, but wild expressions are not. Thus, any wild (3.3) implies the existence of a *third* representation for p a sum of two d-th powers.

The case $d = 3$ is best handled by other techniques and is covered in the companion paper [18]. In preparation for implementing this strategy, we calculate the tame and wild cases which might occur from the list of $\mathcal{W}_2(4,d)$ sets for $d \ge 4$ in Theorems 1.6 and 1.7. Each identity (3.3) has two flips: $f_1^d - f_3^d = f_4^d - f_2^d$ and $f_1^d - f_4^d = f_3^d - f_2^d$, and since either side can be diagonalized, there are potentially six cases. (If there are three equal sums, there are potentially fifteen cases.) Fortunately, symmetry reduces the number of cases substantially.

Theorem 3.4.

(i) The diagonalizations of (1.6) *are, up to scaling,*

$$\begin{aligned} (x^2 + y^2)^4 - 18(xy)^4 &= -(\omega x^2 + \omega^2 y^2)^4 - (\omega^2 x^2 + \omega y^2)^4 \\ &= x^8 + 4x^6y^2 - 12x^2y^2 + 4x^2y^6 + y^8, \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} &-(2x^2 + 2y^2)^4 + 18(x^2 - y^2)^4 \\ &= (x^2 + 2\sqrt{-3}\ xy + y^2)^4 + (x^2 - 2\sqrt{-3}\ xy + y^2)^4 \\ &= 2(x^8 - 68x^6y^2 + 6x^4y^4 - 68x^2y^6 + y^8). \end{aligned} \tag{3.8}$$

(ii) The diagonalizations of (1.14) *are, up to scaling,*

$$
\begin{aligned}
&(\alpha x^2 - \beta y^2)^4 - (\beta x^2 - \alpha y^2)^4 \\
&= (\omega x^2 - \sqrt{3}\ xy - \omega^2 y^2)^4 - (\omega^2 x^2 - \sqrt{3}\ xy - \omega y^2)^4 \\
&= (\omega x^2 + \sqrt{3}\ xy - \omega^2 y^2)^4 - (\omega^2 x^2 + \sqrt{3}\ xy - \omega y^2)^4 \\
&= \sqrt{-3}\left(x^8 - 14x^6y^2 + 14x^2y^6 - y^8\right),
\end{aligned}
\tag{3.9}
$$

where $\alpha = \frac{2+\sqrt{-3}}{2}, \beta = \frac{2-\sqrt{-3}}{2}$*; and*

$$
\begin{aligned}
&((1+\sqrt{-6})x^2 + (1-\sqrt{-6})y^2)^4 + ((1-\sqrt{-6})x^2 + (1+\sqrt{-6})y^2)^4 \\
&= (x^2 + 2\sqrt{-6}\ xy + y^2)^4 + (x^2 - 2\sqrt{-6}\ xy + y^2)^4 \\
&= 2(x^8 - 140x^6y^2 + 294x^4y^4 - 140x^2y^6 + y^8).
\end{aligned}
\tag{3.10}
$$

(iii) The diagonalization of (1.8) *is, up to scaling,*

$$
\begin{aligned}
&((1-\sqrt{-2})x^2 + (1+\sqrt{-2})y^2)^5 + ((1+\sqrt{-2})x^2 + (1-\sqrt{-2})y^2)^5 \\
&= (x^2 - 2\sqrt{-2}\ xy + y^2)^5 + (x^2 + 2\sqrt{-2}\ xy + y^2)^5 = \\
&= 2(x^{10} - 75x^8y^2 + 90x^6y^4 + 90x^4y^6 - 75x^2y^8 + y^{10}).
\end{aligned}
\tag{3.11}
$$

Proof. (i) First, in (1.6), the summands on the left are cyclically permuted by $(x, y) \mapsto (\omega x, \omega^2 y)$, so there is only one choice up to scaling. One is already diagonalized as in (3.7). To diagonalize the left-hand side in (3.7), take $(x, y) \mapsto (x + y, x - y)$ and multiply through by -1, to obtain (3.8).

(ii) It is convenient to name the forms from (1.14) in (3.12) . Let

$$
\begin{aligned}
f_{1,1}(x,y) &= x^2 + \sqrt{3}\ xy - y^2, \quad & f_{1,2}(x,y) &= x^2 - \sqrt{3}\ xy - y^2, \\
f_{1,3}(x,y) &= f_{1,1}(\omega^2 x, \omega y), \quad & f_{1,4}(x,y) &= f_{1,2}(\omega^2 x, \omega y), \\
f_{1,5}(x,y) &= f_{1,1}(\omega x, \omega^2 y), \quad & f_{1,6}(x,y) &= f_{1,2}(\omega x, \omega^2 y);
\end{aligned}
$$
$$
f_{1,1}^4 - f_{1,2}^4 = f_{1,3}^4 - f_{1,4}^4 = f_{1,5}^4 - f_{1,6}^4 = 8\sqrt{3}\ xy\ (x^6 - y^6).
\tag{3.12}
$$

Let M_1 denote the linear change $(x, y) \mapsto (\omega^2 x, \omega y)$, so that M_1 cycles $f_{1,1} \mapsto f_{1,3} \mapsto f_{1,5} \mapsto f_{1,1}$ and $f_{1,2} \mapsto f_{1,4} \mapsto f_{1,6} \mapsto f_{1,2}$. Let M_2 denote the linear change $(x, y) \mapsto \frac{1}{\sqrt{2}}(x + iy, ix + y)$, which has two nice properties. First, M_2 cycles $f_{1,3} \mapsto f_{1,5} \mapsto f_{1,6} \mapsto f_{1,4} \mapsto f_{1,3}$, but it also takes $(f_{1,1}, f_{1,2}) \mapsto (\alpha x^2 - \beta y^2, \beta x^2 - \alpha y^2)$. On the Riemann sphere, M_1 induces a $\frac{2\pi}{3}$ rotation on the axis of the poles. and M_2 induces the rotation taking $(a, b, c) \mapsto (a, c, -b)$.

By repeatedly using M_1 and M_2, the fifteen pairs $\{f_{1,i}, f_{1,j}\}$ which might be simultaneously diagonalized given the identity $f_{1,3}^4 - f_{1,4}^4 = f_{1,5}^4 - f_{1,6}^4$, reduce to two cases, after linear changes. We have already seen one: M_2 diagonalizes (1.14) into (3.9).

For the other, note that

$$f_{1,4}^4(x,y) + f_{1,5}^4(x,y) = f_{1,3}^4(x,y) + f_{1,6}^4(x,y) \\ = -(x^8 + 14x^6y^2 + 42x^4y^4 + 14x^2y^6 + y^8). \tag{3.13}$$

An appeal to Theorem 3.1 shows that the octic in (3.13) is *not* a sum of two fourth powers of even quadratic forms. Under the linear change M_3, which takes $(x,y) \mapsto (x - (\sqrt{2}-1)y, i(\sqrt{2}-1)x + iy)$ and division by $\sqrt{2}-2$, (3.13) becomes (3.10).

(iii) We name the quadratics from (1.8) in (3.14). Let M_4 be the scaling $(x,y) \mapsto (\zeta_8 x, \zeta_8^3 y)$, which takes $(x^2, xy, y^2) \mapsto (ix^2, -xy, -iy^2)$, so that

$$f_{2,1}(x,y) = x^2 + \sqrt{-2}\, xy + y^2,\ f_{2,2} = f_{2,1} \circ M_4,\ f_{2,3} = f_{2,2} \circ M_4, \\ f_{2,4} = f_{2,3} \circ M_4; \qquad f_{2,1}^5 + f_{2,2}^5 + f_{2,3}^5 + f_{2,4}^5 = 0. \tag{3.14}$$

Thus M_4 cycles $f_{2,1} \mapsto f_{2,2} \mapsto f_{2,3} \mapsto f_{2,4} \mapsto f_{2,1}$. The symmetry of the Klein set for $\{f_{2,j}\}$ (the cube) suggests that we let M_5 be the linear change $(x,y) \mapsto \frac{1}{\sqrt{2}} \cdot (-x + \zeta_8^5 y, \zeta_8^3 x + y)$. Then M_5 fixes $f_{2,1}$ and $f_{2,4}$ and permutes $f_{2,2}$ and $f_{2,3}$.

Thus M_4 maps the flip $f_{2,1}^5 + f_{2,2}^5 = -f_{2,3}^5 - f_{2,4}^5$ into $f_{2,2}^5 + f_{2,3}^5 = -f_{2,4}^5 - f_{2,1}^5$ and M_5 maps it into $f_{2,1}^5 + f_{2,3}^5 = -f_{2,2}^5 - f_{2,4}^5$, so, up to cousin, we need only consider one flip. The easiest one to deal with is $f_{2,1}^5 + f_{2,3}^5 = -f_{2,2}^5 - f_{2,4}^5$. This is

$$(x^2 + \sqrt{-2}\, xy + y^2)^5 + (-x^2 + \sqrt{-2}\, xy - y^2)^5 \\ = -(ix^2 - \sqrt{-2}\, xy - iy^2)^5 - (-ix^2 - \sqrt{-2}\, xy + iy^2)^5 \\ = 2\sqrt{-2}\, xy(5x^8 - 6x^4y^4 + 5y^8). \tag{3.15}$$

Upon taking $(x,y) \mapsto (x+iy, x-iy)$, and dividing by $\sqrt{-2}$, (3.15) becomes (3.11). And under the linear change, $(x,y) \mapsto \frac{1}{\sqrt{2}}(x+iy, x-iy)$, (1.15) also becomes (3.11). The Klein set of the summands in (3.11) is a rotated cube lying in the planes $y = \pm\sqrt{1/3}$, so that the edge $(0, \pm\sqrt{1/3}, \sqrt{2/3})$ lies on top. □

4. Finishing the proof

We first make a simplifying observation in the tame case. If (f_3, f_4) is given in (3.5) or (3.6) and $a = 0$ (or $c = 0$), then f_3 and f_4 have a common factor of y (or x), violating Theorem 3.3. Similarly, we may assume that $b \neq 0$. Thus, after scaling, we may assume that (3.5) and (3.6) take the shape

$$(x^2 + bxy + y^2)^d + (x^2 - bxy + y^2)^d, \qquad b \neq 0; \tag{4.1}$$

$$(x^2 + y^2)^{2e} + b\binom{2e}{e}(xy)^{2e}, \qquad b \neq 0. \tag{4.2}$$

Theorem 4.1. *The only $\mathcal{W}_2(4,d)$ sets which come from a tame representation for $d \geq 4$ are given in Theorem 3.4 by* (3.7), (3.8), (3.10), *and* (3.11). *These sets are all cousins or sub-cousins of the families in Theorems 1.6, 1.7.*

Proof. We analyze (4.2) first. The 2-Sylvester matrix of $(x^2+y^2)^4 + 6b(xy)^4$ is

$$\begin{pmatrix} 1 & 1 & 1+b \\ 1 & 1+b & 1 \\ 1+b & 1 & 1 \end{pmatrix}, \tag{4.3}$$

which has rank 2 only if $-b^2(b+3) = 0$; if $b = -3$, we obtain (3.7).

If $d = 2s \ge 6$ and $p_{2s,b}(x,y) = (x^2+y^2)^{2s} + b\binom{2s}{s}(xy)^{2s}$, then the $(2s-1)\times 3$ 2-Sylvester matrix consists of (4.3), with $s-2$ rows of $(1,1,1)$ appended both at the top and the bottom. Such a matrix has rank 2 only if $b = 0$.

For (4.1), we first observe that

$$(x^2+bxy+y^2)^d + (x^2-bxy+y^2)^d = 2\sum_{0\le i\le d/2}\binom{d}{2i}(x^2+y^2)^{d-2i}(xy)^{2i}. \tag{4.4}$$

Suppose $d = 4$. Then the sum in (4.4) becomes

$$2x^8 + (8+12b^2)x^6y^2 + (12+24b^2+2b^4)x^4y^4 + (8+12b^2)x^2y^6 + 2y^8.$$

Apply Theorem 3.1: the 2-Sylvester matrix has discriminant $-\frac{b^8}{27}(12+b^2)(24+b^2)$, and has rank 2 only if $b^2 \in \{-12,-24\}$. These cases are presented in (3.8) and (3.10), and are a cousin of (1.6) and a sub-cousin of (1.14), respectively.

Suppose $d = 5$. Then applying Theorem 3.1 to (4.4) gives a 4×3 matrix; computing the 3×3 minors shows that the matrix has rank 2 only when $b = 0$ or $b^2 = -8$. Taking $b = \sqrt{-8}$, we obtain (3.11), which is a cousin of (1.8).

Now suppose $d \ge 6$; (4.4) gives

$$\begin{gathered} a_0 = a_d = 2, \quad a_1 = a_{d-1} = 2 + b^2(d-1), \\ a_2 = a_{d-2} = 2 + b^2(d-2)(12+(d-3)b^2)/6, \\ a_3 = a_{d-3} = 2 + b^2(d-3)(180 + b^2(30d-120) + b^4(d^2-9d+20))/60. \end{gathered}$$

The submatrix of the 2-Sylvester matrix consisting of the first and last two rows is

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_3 & a_2 & a_1 \\ a_2 & a_1 & a_0 \end{pmatrix}.$$

The 1,2,4 minor of this sub-matrix is $-\frac{b^8}{9(d-1)}\binom{d+1}{5}(12+b^2(d-3))(24+b^2(2d-7))$. If $b^2 = -\frac{12}{d-3}$, then the 1,2,3 minor becomes $\frac{55296\ d^2(d+1)(d-4)}{25(d-3)^5} \ne 0$. However, if $b^2 = -\frac{24}{2d-7}$, then all four minors vanish. (Note that $d = 4,5$ then give $b^2 = -24, b^2 = -8$, which we have already seen.) We re-compute the a_k's for $b^2 = -\frac{24}{2d-7}$, and find that the first three rows of the 2-Sylvester matrix give

$$\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix} = -\frac{3538944(d-5)(d-4)d(1+d)(2d-1)^2}{175(2d-7)^6} \neq 0.$$

Thus, no tame representations exist when $d \geq 6$. □

Suppose now that we have a wild representation

$$\begin{aligned} p(x,y) &= (a_1x^2 + b_1xy + c_1y^2)^d + (a_2x^2 + b_2xy + c_2y^2)^d \\ &= \sum_{i=0}^{2d} s_i(a_1, b_1, c_1, a_2, b_2, c_2; d)x^{2d-i}y^i, \end{aligned} \tag{4.5}$$

where $d \geq 4$, $s_{2j+1}(a_1, b_1, c_1, a_2, b_2, c_2; d) = 0$ for $0 \leq j \leq d-1$, $(b_1, b_2) \neq (0,0)$ and (4.5) is not in the form (3.5) or (3.6).

Lemma 4.2. *Suppose* $p \neq 0$ *and* (4.5) *holds. Then, after a scaling of* x *and* y,

$$p(x,y) = p_{\lambda,\alpha,\beta}(x,y) := (x^2 - \lambda\alpha xy + y^2)^d + \lambda(x^2 + \alpha xy + \beta y^2)^d, \tag{4.6}$$

where $\alpha\lambda \neq 0$, $\beta^{d-1} = 1$ *and* $\lambda^2 \neq 1$.

Proof. First suppose $b_1 = 0$ in (4.5). Then $s_1 = da_2^{d-1}b_2$ and $s_{2d-1} = db_2c_2^{d-1}$. Since $(b_1, b_2) \neq (0,0)$, we have $a_2 = c_2 = 0$ and $p(x,y) = (a_1x^2 + c_1y^2)^d + (b_2xy)^d$ is even, so d is even and we have (3.6). A similar argument lets us conclude that $b_2 \neq 0$.

Suppose now that $a_1 = 0$. Then $s_1 = da_2^{d-1}b_2 = 0$, and $b_2 \neq 0$ implies $a_2 = 0$. It then follows that y divides both f_3 and f_4, contradicting Theorem 3.3. Thus $a_1 \neq 0$, and by similar arguments, we have $a_2c_1c_3 \neq 0$. That is, we may assume that all the coefficients in (4.5) are non-zero.

We now scale x and y so that $a_1 = c_1 = 1$ and let $\lambda = a_2^d$, so that, after renaming,

$$p(x,y) = (x^2 + \alpha_1 xy + y^2)^d + \lambda(x^2 + \alpha_2 xy + \beta y^2)^d, \tag{4.7}$$

where all parameters are non-zero. Returning to the computation,

$$s_1 = d(\alpha_1 + \lambda\alpha_2) = 0, \quad s_{2d-1} = d(\alpha_1 + \lambda\alpha_2\beta^{d-1}) = 0.$$

It follows that $\alpha_1 = -\lambda\alpha_2$, and since $\lambda\alpha_2 \neq 0$, it also follows that $\beta^{d-1} = 1$. We now write $\alpha = \alpha_2$, so that $\alpha_1 = -\lambda\alpha$, and (4.7) becomes (4.6). Finally, if $\lambda^2 = 1$, then either $\lambda = 1$ (and (4.6) reduces to (3.5)), or $\lambda = -1$ (and (4.6) implies $p = 0$). □

Theorem 4.3. *For* $d \geq 4$, *the only* $\mathcal{W}_2(4,d)$ *set which comes from a wild representation is found in* (3.10), *and is a sub-cousin of* (1.14).

Proof. In view of Lemma 4.2, we simplify our notation: let

$$p_{\lambda,\alpha,\beta}(x,y) = \sum_{i=0}^{2d} a_i(\lambda, \alpha, \beta; d)x^{2d-i}y^i. \tag{4.8}$$

Since $p_{\lambda,\alpha,\beta}(x,y)$ is even, so is $p_{\lambda,\alpha,\beta}(y,x)$, as is their difference. For this reason, write

$$\begin{aligned}\lambda^{-1}(p_{\lambda,\alpha,\beta}(x,y) - p_{\lambda,\alpha,\beta}(y,x)) &= (x^2+\alpha xy+\beta y^2)^d - (\beta x^2+\alpha xy+y^2)^d \\ &= \sum_{i=0}^{2d} b_i(\alpha,\beta,d)x^{2d-i}y^i.\end{aligned} \tag{4.9}$$

We need to find the conditions under which $a_{2j+1}(\lambda,\alpha,\beta;d) = 0$ for $1 \le 2j+1 \le 2d-1$. Since $\lambda b_i(\alpha,\beta) = a_i(\lambda,\alpha,\beta;d) - a_{2d-i}(\lambda,\alpha,\beta;d)$ and $\lambda \ne 0$, it suffices to consider $a_{2j+1}(\lambda,\alpha,\beta;d) = b_{2j+1}(\alpha,\beta,d) = 0$ for $1 \le 2j+1 \le d$.

It follows from the definition and $\beta^{d-1} = 1$ that

$$p_{\lambda,\alpha,\beta}(x,y) = p_{\lambda,-\alpha,\beta}(x,-y), \quad p_{\lambda,\alpha,\beta}(x,y) = p_{\lambda\beta,\alpha/\beta,1/\beta}(y,x), \tag{4.10}$$

so that, up to linear change, if $\alpha^2 = \kappa$ is known, then choosing $\alpha = \pm\sqrt{\kappa}$ gives two equations that are cousins. Also, any solution for a particular value $\beta = \beta_0$ will be a cousin of a solution in which $\beta = \beta_0^{-1}$. This reduces the number of choices to check.

We now have

$$\begin{aligned}a_1(\lambda,\alpha,\beta) &= -d\alpha\lambda + d\alpha\lambda = 0, \quad b_1(\alpha,\beta) = d\alpha(\beta^{d-1}-1) = 0, \\ a_3(\lambda,\alpha,\beta) &= \frac{\lambda\alpha d(d-1)}{6}\cdot\left((d-2)\alpha^2(1-\lambda^2) + 6(\beta-1)\right), \\ b_3(\alpha,\beta) &= \frac{\alpha d(d-1)}{6}\cdot(1-\beta^{d-3})(6\beta+\alpha^2(d-2)).\end{aligned}$$

Now we claim that $\beta \ne 1$ and either

$$\beta = -1, \quad \alpha^2 = \frac{12}{(d-2)(1-\lambda^2)} \qquad \text{(and } d \text{ is odd);} \tag{4.11}$$

or

$$\beta = \frac{1}{\lambda^2}, \quad \alpha^2 = -\frac{6}{\lambda^2(d-2)}. \tag{4.12}$$

Indeed, since $\alpha(1-\lambda^2) \ne 0$, the equation $a_3 = 0$ implies that $\beta \ne 1$ and

$$\alpha^2 = \frac{6(1-\beta)}{(d-2)(1-\lambda^2)}. \tag{4.13}$$

The equation $b_3 = 0$ implies that $(1-\beta^{d-3})(6\beta+\alpha^2(d-2)) = 0$. If $\beta^{d-3} = 1$, then $\beta^{d-1} = 1$ implies $\beta^2 = 1$, and $\beta = 1$ is ruled out, so $\beta = -1$ and d is odd and (4.13) implies (4.11). Otherwise, we have by (4.13),

$$0 = 6\beta + \alpha^2(d-2) = 6\beta + \frac{6(1-\beta)}{(1-\lambda^2)} = \frac{6(1-\beta\lambda^2)}{1-\lambda^2},$$

so $1 = \beta\lambda^2$ and by (4.13),

$$\alpha^2 = \frac{6(1-\lambda^{-2})}{(d-2)(1-\lambda^2)} = -\frac{6}{\lambda^2(d-2)};$$

this is summarized as (4.12).

If $d = 4$, then only (4.12) can apply. Since $\beta^3 = 1$, $\beta \neq 1$ and $\omega \cdot \omega^2 = 1$, we can use (4.10) to assume that $\beta = \omega^2$. It follows from (4.12) that

$$\omega^2 = \frac{1}{\lambda^2}, \quad \alpha^2 = -\frac{3}{\lambda^2} \implies \lambda = \pm\omega^2,\ \alpha^2 = -3\omega^2.$$

By (4.10), it suffices to take $\alpha = \sqrt{-3}\,\omega$, but there are two values for λ: $\lambda = \pm\omega^2$. There are two wild cases: since $\lambda\alpha = \pm\sqrt{-3}$ and $(\omega^2)^4 = \omega^2$, these are

$$\begin{aligned} p_{4,\pm}(x,y) &:= (x^2 \mp \sqrt{-3}\ xy + y^2)^4 \pm \omega^2(x^2 + \sqrt{-3}\ \omega xy + \omega^2 y^2)^4 \\ &= (x^2 \mp \sqrt{-3}\ xy + y^2)^4 \pm (\omega^2 x^2 + \sqrt{-3}\ xy + \omega y^2)^4. \end{aligned} \tag{4.14}$$

We scale the two cases of (4.14) to make them easier to work with. First

$$\begin{aligned} \omega^2 p_{4,+}(x,\omega i y) := q_1(x,y) &= -x^8 - 14x^6y^2 - 42x^4y^4 - 14x^2y^6 - y^8 \\ &= (\omega^2x^2 - \sqrt{3}\ xy - \omega y^2)^4 + (\omega x^2 + \sqrt{3}\ xy - \omega^2 y^2)^4. \end{aligned} \tag{4.15}$$

The second line in (4.15) is $f_{1,4}^4 + f_{1,5}^4$, which gives a new representation after $y \mapsto -y$, namely, $f_{1,3}^4 + f_{1,6}^4$; c.f. (3.13). However, the 2-Sylvester matrix of q_1 has rank 3, so this case does not fall under Theorem 3.3.

For the other case, we have

$$\begin{aligned} -\omega^2 p_{4,-}(x,\omega i y) &:= q_2(x,y) = \\ -(\omega^2x^2 - \sqrt{3}\ xy - \omega y^2)^4 &+ (\omega x^2 - \sqrt{3}\ xy - \omega^2 y^2)^4 \\ &= \sqrt{-3}\ (x^8 - 14x^6y^2 + 14x^2y^6 - y^8). \end{aligned} \tag{4.16}$$

The 2-Sylvester matrix of q_2 has rank 2, so it has a representation as a sum of two fourth powers. Indeed, (4.16) is embedded in (3.9), with two other representations of q_2: one from taking $y \mapsto -y$ in (4.16), and the other by applying Theorem 3.1.

Now suppose $d \geq 5$; more equations need to be satisfied. If (4.11) holds, then

$$a_5 = -\frac{8\sqrt{3}\lambda(1+\lambda^2)(d+1)d(d-1)(d-3)}{5((d-2)(1-\lambda^2))^{3/2}} = 0,$$

so $\lambda^2 = -1$, and (4.11) becomes

$$\beta = -1, \quad \lambda^2 = -1, \quad \alpha^2 = \frac{6}{d-2}. \tag{4.17}$$

If (4.12) holds, then

$$a_5 = -\frac{\sqrt{6}(\lambda^4 - 1)(2d+1)d(d-1)(d-4)}{10\lambda^4(d-2)^{3/2}}. \tag{4.18}$$

Since $\lambda^2 \neq 1$, (4.18) implies $\lambda^2 = -1$, and simplification yields (4.17) again. Observe that $\lambda = \pm i$ implies that $d \equiv 1 \pmod 4$.

If $d = 5$, then $\beta = -1$, $\lambda^2 = -1$ and $\alpha^2 = 2$. We choose $\alpha = \sqrt{2}$ and obtain two solutions, for $\lambda = i$ and $\lambda = -i$, which we rewrite in terms of the $f_{2,j}$'s, upon noting that $\pm i = (\pm i)^5$:

$$\begin{aligned} p_{5,+}(x,y) &= (x^2 - i\sqrt{2}xy + y^2)^5 + i(x^2 + \sqrt{2}xy - y^2)^5 = -f_{2,3}^5 - f_{2,4}^5 \\ &= (1+i)(x^{10} + 15ix^8y^2 - 30x^6y^4 + 30ix^4y^6 - 15x^2y^8 - iy^{10}) \\ p_{5,-}(x,y) &= (x^2 + i\sqrt{2}xy + y^2)^5 - i(x^2 + \sqrt{2}xy - y^2)^5 = f_{2,1}^5 + f_{2,4}^5 \\ &= (1-i)(x^{10} - 15ix^8y^2 - 30x^6y^4 - 30ix^4y^6 - 15x^2y^8 + iy^{10}) \end{aligned} \tag{4.19}$$

The expressions in (4.19) are close cousins; in fact, $p_{5,-}(x,y) = -ip_{5,+}(x,iy)$. Theorem 3.1 shows that neither has a representation as a sum of two even 5th powers; however, $p_{5,-}(x,y) + ip_{5,+}(x,iy) = 0$ is a cousin of (1.8).

Suppose now that $d \ge 6$; since $d \equiv 1 \pmod 4$, we have $d \ge 9$. It turns out that $b_5 = 0$ under the conditions of (4.17), but

$$a_7\left(\pm i, \sqrt{\tfrac{6}{d-2}}, -1, d\right) = \pm\frac{8i\sqrt{2}(2d-1)(d^3-d)(d-3)(d-5)}{35\sqrt{3}(d-2)^{5/2}} = 0 \tag{4.20}$$

is clearly impossible for $d \ge 9$, so we are finally done with the wild case. □

Proof of Theorems 1.3(8), 1.6 and 1.7. Combine Theorems 3.3, 4.1, and 4.3. □

5. Final remarks

5.1. **Derivations and historical examples.** It is foolhardy for a living author to claim priority for any polynomial identity which is verifiable by hand and so might well have been given as a school algebra assignment. We have given previous attributions when we could find them; the pre-1920 literature was scoured by Dickson in [3], but with Diophantine equations over $\mathbb{N}$ in mind: the coverage of parameterizations over $\mathbb{C}$ must be regarded as incomplete. For example, the 1880 paper [2] by Desboves includes both (1.15) and (1.8), and Dickson only cites the latter, perhaps because there were no real quintic parameterizations.

Any four binary quadratic forms are linearly dependent, so any $\mathcal{W}_2(4,d)$ satisfies both $f_1^d + f_2^d = f_3^d + f_4^d$ and $c_1f_1 + c_2f_2 + c_3f_3 + c_4f_4 = 0$ for suitable c_i. It is remarkable that one can find the $\mathcal{W}_2(4,d)$'s for $d = 4, 5$ by guessing a simple choice of c_i's.

For example, Desboves [2, p.241] found his version of (1.8) by assuming $f_1 + f_2 = f_3 + f_4$ and $f_1^5 + f_2^5 = f_3^5 + f_4^5$ and parameterizing, to get

$$0 = (f+g)^5 + (f-g)^5 - ((f+h)^5 + (f-h)^5) = 10f(g^2 - h^2)(2f^2 + g^2 + h^2).$$

He then set $\{f, g, h\} = \{2xy, x^2 - 2y^2, i(x^2 + 2y^2)\}$ via Theorem 1.4 and by scaling via $y \mapsto \sqrt{-1/2}y$, this becomes essentially (1.8). Similarly, after noting that

$$(f+g)^4 + (f-g)^4 - ((f+h)^4 + (f-h)^4) = 2(g^2 - h^2)(6f^2 + g^2 + h^2),$$

Desboves solved $6f^2 + g^2 + h^2 = 0$ and derived a cousin of (1.15).

One might also guess $f_1+f_2+f_3=0$; an old observation (at least back to Proth in 1878 [3, p.657]) notes that

$$f_1^4+f_2^4+(-f_1-f_2)^4=2(f_1^2+f_1f_2+f_2^2)^2, \tag{5.1}$$

so if $f_1^2+f_1f_2+f_2^2=g^2$, we obtain a $\mathcal{W}_2(4,4)$. Take $f_1=x^2+y^2$ and $f_2=\omega x^2+\omega^2y^2$; this implies $-(f_1+f_2)=\omega^2x^2+\omega y^2$ and $f_1^2+f_1f_2+f_2^2=3x^2y^2$; hence (1.6).

In 1904, Ferrari (see [3, p.654]) gave the ostensibly ternary identity:

$$\begin{aligned}(a-b)^4(a+b+2c)^4+(b+c)^4(b-c-2a)^4+(c+a)^4(c-a+2b)^4\\ =2(a^2+b^2+c^2-ab+ac+bc)^4\end{aligned} \tag{5.2}$$

Let $x=a-b$ and $y=b+c$, so that $x+y=a+c$. Then (5.2) becomes (1.7):

$$x^4(x+2y)^4+y^4(-2x-y)^4+(x+y)^4(y-x)^4=2(x^2+xy+y^2)^4.$$

One can derive (1.14) by guessing $(a+d)^4-(a-d)^4=(b+d)^4-(b-d)^4=(c+d)^4-(c-d)^4$ for quadratics a,b,c,d with a,b,c distinct and $d\neq 0$. Then routine computations lead to $a+b+c=0$ and $d^2=-(a^2+ab+b^2)$. Now set $a=x^2+y^2, b=\omega x^2+\omega^2y^2, c=\omega^2x^2+\omega y^2$, with $d^2=-(a^2+ab+b^2)=-3x^2y^2$, and take $y\mapsto iy$ to get (1.14).

We derived (1.8) in [15, pp.119-120] using Newton's Theorem on symmetric polynomials. Every symmetric quaternary quintic polynomial p is contained in the ideal $\mathcal{I}=(t_1+t_2+t_3+t_4, t_1^2+t_2^2+t_3^2+t_4^2)$. In particular, $t_1^5+t_2^5+t_3^5+t_4^5\in\mathcal{I}$, so

$$f_1+f_2+f_3+f_4=0,\ f_1^2+f_2^2+f_3^2+f_4^2=0 \implies f_1^5+f_2^5+f_3^5+f_4^5=0.$$

Upon setting $f_4=-f_1-f_2-f_3$, the equation $f_1^2+f_2^2+f_3^2+(-f_1-f_2-f_3)^2=0$ can be analyzed as in Theorem 1.4 to obtain (1.8).

We present a similar *ad hoc, post hoc* derivation for (1.11).

Theorem 5.1. *Suppose $S(t_1,\dots,t_6)$ is a symmetric polynomial of degree 7. Then*

$$S\in\mathcal{I}:=\left(\sum_{k=1}^6 t_k, \sum_{k=1}^6 t_k^2, \sum_{k=1}^6 t_k^4\right).$$

Proof. Let e_k denote the k-th elementary symmetric polynomial. We have $\sum_{k=1}^6 t_k^2=e_1^2-e_2$ and $\sum_{k=1}^6 t_k^4=e_1^4-4e_1^2e_2+2e_2^2+4e_1e_3-4e_4$. Thus, $\mathcal{I}=(e_1,e_2,e_4)$. By Newton's Theorem, S is a linear combination of monomials in the e_k's: $e_1^{a_1}e_2^{a_2}e_3^{a_3}e_4^{a_4}e_5^{a_5}e_6^{a_6}$, where $\sum ka_k=7$. But 7 cannot be written as a non-negative linear combination of 3, 5 and 6, so each monomial in any such expression must contain one of $\{e_1,e_2,e_4\}$. □

Observe now that if we define $h_j=(\zeta_5^{j-1}x^2+ixy+\zeta_5^{-(j-1)}y^2)^2$ for $1\le j\le 5$ and $h_6=-5x^2y^2$, then a synching computation shows that $\sum_{j=1}^6 h_j=\sum_{j=1}^6 h_j^2=\sum_{j=1}^6 h_j^4=0$. Theorem 5.1 implies that $\sum_{j=1}^6 h_j^7=0$; that is, (1.11). The mystery now is *why* these particular squares work.

Jordan Ellenberg has suggested the following explanation to the author: The surface cut out by $\sum_{j=1}^6 X_j=\sum_{j=1}^6 X_j^2=\sum_{j=1}^6 X_j^4$ is a Hilbert modular surface (see [4, Lemma 2.1]). He adds [5]: "Dollars to donuts the nice low-degree rational curve you find on this surface arises as a modular curve on this modular surface, parametrizing abelian surfaces isogenous to a product of elliptic curves".

5.2. **Representations as a sum of at most two d-th powers of quadratic forms.** Which forms $p \in H_{2d}(\mathbb{C}^2)$ can be written as a sum of two d-th powers of linear forms, and in how many ways? Let $A_{d,2} = \{(\alpha_1 x + \beta_1 y)^d + (\alpha_2 x + \beta_2 y)^d\}$. It is tautological to say that $p \in A_{d,2}$ if and only if there is a linear change taking p into x^d or $x^d + y^d$. (A practical test is given by Theorem 3.1.)

Corollary 5.2. *If $p \in H_{2d}(\mathbb{C}^2)$ is not a d-th power, then p is a sum of two d-th powers of quadratic forms if and only if either (i) $p = \ell^d q$, where $q \in A_{d,2}$, or (ii) after a linear change in p, $p(x,y) = q(x^2,y^2)$, where $q \in A_{d,2}$.*

Proof. Sufficiency is clear. Conversely, suppose $p = f_1^d + f_2^d$ and $\{f_1, f_2\}$ is honest. As in Theorem 3.2, there are two cases. If $\gcd(f_1, f_2) = \ell$ for a linear form ℓ, then $f_j = \ell\ell_j$, giving case (i). Otherwise, we make a linear change which simultaneously diagonalizes f_1, f_2, giving case (ii). □

If p is a sum of two d-th powers in more than one way, then the two representations together give a $\mathcal{W}_2(d,4)$. The question is not interesting for $d = 2$, since $p = f^2 + g^2 \iff p = (f+ig)(f-ig)$, so two representations as a sum of two squares amount to two different factorizations into equal degrees. The situation for $d = 3$ is discussed in detail in [18]; by Theorem 1.3(8), it suffices now to consider $d = 4, 5$.

If p itself is a d-th power, then by Theorem 1.3(3), it does not have another representation as a sum of two d-th powers. In view of Theorems 1.6, 1.7, 3.4, we have an immediate corollary. We choose even representatives (from Theorem 3.3) and they also happen to be symmetric (we have taken $y \mapsto \zeta_{16} y$ in (3.9).)

Corollary 5.3.

(i) The form $p \in H_8(\mathbb{C}^2)$ has exactly two different representations as a sum of two fourth powers of binary forms if and only if, after a linear change, it is $x^8 + 4x^6y^2 - 12x^4y^4 + 4x^2y^6 + y^8$, $x^8 - 68x^6y^2 + 6x^4y^4 - 68x^2y^6 + y^8$, or $x^8 - 140x^6y^2 + 294x^4y^4 - 140x^2y^4 + y^8$.

(ii) The form $p \in H_8(\mathbb{C}^2)$ has three different representations as a sum of two fourth powers of binary forms if and only if, after a linear change, it is $x^8 - 7\sqrt{2}(1+i)x^6y^2 - 7\sqrt{2}(1+i)x^2y^6 + y^8$.

(iii) The form $p \in H_{10}(\mathbb{C}^2)$ has two different representations as a sum of two fifth powers of binary forms if and only if, after a linear change, it is $x^{10} - 75x^8y^2 + 90x^6y^4 + 90x^4y^6 - 75x^2y^8 + y^{10}$.

5.3. **Open questions.** We have already noted that there exists $k \ge 2$ and $d \ge 6$ so that $\Phi_k(d) > \Phi_{k+1}(d)$. Gundersen in [7] found three meromorphic (not rational) functions $g_j(t)$ so that $g_1^6 + g_2^6 + g_3^6 = 1$. It is unknown whether this can be achieved with rational functions. If so, a $\mathcal{W}_k(4,6)$ set would exist for some $k > 2$.

In case $m = rs$, an m-synching on m can be viewed as r coordinated s-synchings. We have not found a useful instance in this when $r = s = 2$, although (2.15) shows what can happen with $(r, s) = (2, 3)$. We hope that improvements on the bounds may come from careful investigations in this direction.

Another natural question is to restrict our attention to forms with coefficients in a fixed subfield of $\mathbb{C}$, such as $\mathbb{Q}$ or $\mathbb{R}$. Real forms with even degree also lead to a discussion of "signatures". From the Diophantine point of view, the equations $A^4 + B^4 + C^4 = D^4$ and $A^4 + B^4 = C^4 + D^4$ are completely different questions. In this point of view, the real equation (1.7) is "(3,1)". In 1772, Euler gave a famous (2,2) "septic" example of a $\mathcal{W}_7(4, 4)$ set (see [3, pp.644-646], [9, (13.7.11)], [11]). So far as we have been able to determine, there are no known real solutions of this kind of smaller degree, nor proofs that they cannot exist.

Theorem 1.4 shows that (1.1) is "universal" in presenting all $\mathcal{W}_k(3, 2)$ sets; that is, projectively, all families come from the substitution $(x, y) \mapsto (g, h)$. Are the solutions given in Theorems 1.5, 1.6 and 1.7 also universal in this sense? The answers are "no" for $d = 3, 4$. These families are all linearly dependent. For $d = 3$, the family in (2.17) is linearly independent, as are the parameterizations of the Euler-Binet solutions to $x^3 + y^3 = u^3 + v^3$ (see e.g [9, (13.7.8)]), when viewed as elements of $\mathbb{C}[a, b, \lambda]$. For $d = 4$, it can be checked that the Euler septics are also linearly independent. The case $d = 5$ is open. Can the sets $\mathcal{W}_k(4, d)$ themselves be parameterized for $k \geq 3$?

Finally, we note that the intricate calculations of section three and four suggest that new methods will be needed to study $\mathcal{W}_k(r, d)$ for $r > 4$ or $k > 2$. Nevertheless, we make the following conjecture, based on Theorem 1.3:

Conjecture 5.4. *There is a small constant M so that, for all k and d,*

$$\left| \Phi_k(d) - \min_{1 \leq i \leq k} \left(\frac{d}{i} + i \right) \right| < M.$$

References

[1] H. Darmon and A. Granville, *On the equations $z^m = F(x, y)$ and $Ax^p + By^q = Cz^r$*, Bull. London Math. Soc. **27** (1995), 513–543.

[2] A. Desboves, *Sur la résolution en nombres entier ou complexe de l'équation $U^n \pm V^n = S^n + W^n$*, Assoc. Franç. **9**, (1880), 239–243, online access located at https://babel.hathitrust.org/cgi/pt?id=nyp.33433012523159;view=1up;seq=337.

[3] L. E. Dickson, *History of the Theory of Numbers, Vol. II, Diophantine Analysis*, Chelsea, New York, 1966.

[4] J. Ellenberg, *Serre's Conjecture over F_9*, Ann. of Math. (3) **161** (2005), no. 3, 1111–1142, MR2180399 (2006g:11102).

[5] J. Ellenberg, email to the author, August 4, 2012.

[6] M. L. Green, *Some Picard theorems for holomorphic maps to algebraic varieties*, Amer. J. Math **97** (1975), 43–75, MR0367302 (51 #3544).

[7] G. G. Gundersen, *Meromorphic solutions of $f^6 + g^6 + h^6 = 1$*, Analysis (Munich) **18** (1998), 285–290, MR1660942 (99g:30040).

[8] G. G. Gundersen and W. K. Hayman, *The strength of Cartan's version of Nevanlinna theory*, Bull. London Math. Soc. **36** (2004), 433–454, MR2069006 (2005i:30045)

[9] G. H. Hardy and E. M. Wright, *An introduction to the Theory of Numbers*, Fifth edition, Oxford University Press, Oxford, 1979, MR0568909 (81i:10002).
[10] W. K. Hayman, *Warings Problem für analytische Funktionen*, Bayer. Akad. Wiss. Math.-Natur. Kl. Sitzungsber 1984, **13** (1985), MR0803374 (87a:11101)
[11] L. J. Lander, *Geometric aspects of Diophantine equations involving equal sums of like powers*, Amer. Math. Monthly, **75** (1968), 1061–1073, MR0237428 (38 #5710)
[12] J. Molluzzo, *Monotonicity of quadrature formulas and polynomial representation*, Ph.D. dissertation, Yeshiva University, 1972, MR2622862.
[13] G. Nenashev, B. Shapiro, M. Shapiro, *Secant degeneracy index of the standard strata in the space of binary forms*, Arnold Math. J. **3** (2017), 499–510, MR3766073.
[14] D. J. Newman and M. Slater, *Waring's problem for the ring of polynomials*, J. Number Theory **11** (1979), 477–487, MR0544895 (80m:10016)
[15] B. Reznick, *Patterns of dependence among powers of polynomials*, Algorithmic and quantitative real algebraic geometry (Piscataway, NJ, 2001), 101–121, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., **60**, Amer. Math. Soc., Providence, RI, 2003, MR1995017 (2004e:11036).
[16] B. Reznick, *On the length of forms*, Quadratic and higher degree forms, 207–232, Dev. Math., **31**, Springer, New York, 2013, MR3156559
[17] B. Reznick, *Some new canonical forms for polynomials*, Pac. J. Math, **266** (2013), 185–220, http://arxiv.org/abs/1203.5722, MR3105781.
[18] B. Reznick, *Binary sextic forms as a sum of two cubes*, in preparation.
[19] B. Reznick, https://faculty.math.illinois.edu/∼ reznick/92100notes.pdf
[20] P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, New York-Heidelberg, 1979, MR0551363 (81f:10023).

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
E-mail address: `reznick@math.uiuc.edu`