DENSITY OF SOLUTIONS FOR SYSTEMS OF FORMS

AMICHAI LAMPERT

ABSTRACT. Let K be a field of characteristic zero over which every diagonal form in sufficiently many variables admits a nontrivial solution. For example, K may be a totally imaginary number field or a finite extension of a p-adic field. Suppose $f_1, \ldots, f_s$ are forms of degree d over K. Bik, Draisma and Snowden recently proved that there exists a constant $B = B(d, s, K)$ such that the rational solutions to the system of equations $f_1 = \ldots = f_s = 0$ are Zariski dense, as long as the Birch rank of $f_1, \ldots, f_s$ is greater than B. We establish an effective bound for this constant. Combined with results of Skinner, we obtain as a corollary a sufficient condition for the integer zeros of a system of forms over a number field to satisfy the Hardy-Littlewood asymptotic with a positive leading term. This corollary generalizes a seminal result proved by Schmidt for the rational numbers.

1. INTRODUCTION

We are interested in studying solutions to systems of forms (i.e. homogeneous polynomials) over a field K. An especially simple class of forms is the diagonal ones.

Definition 1.1. The quantity $\phi_d = \phi_d(K)$ is defined to be the minimal integer such that the following holds: Whenever $n > \phi_d$ and $a_1, \ldots, a_n \in K$ there is some $0 \neq x \in K^n$ satisfying $a_1 x_1^d + \ldots + a_n x_n^d = 0$. If there is no such threshold, then $\phi_d = \infty$.

For example, $\phi_1 = 1$ over any field and $\phi_2(\mathbf{Q}) = \infty$. From now on we assume that $\phi_2, \ldots, \phi_d$ are finite. Consider a collection of forms (i.e. homogeneous polynomials) $f_1, \ldots, f_s \in K[x_1, \ldots, x_n]$ of common degree d.

1.1. **Existence of solutions.** In a landmark 1945 paper [5], Brauer proved:

Theorem 1.2 (Brauer). *There exists a constant $V_{d,s} = V(d, s, \phi_2, \ldots, \phi_d)$ such that if $n > V_{d,s}$ then there is some $0 \neq x \in K^n$ with $f_1(x) = \ldots = f_s(x) = 0$.*

The proof of theorem 1.2 proceeded via an elegant diagonalization argument, but resulted in an extremely large value for $V_{d,s}$. It was long thought impossible to obtain a reasonable bound by similar methods. However, in 1983 Leep and Schmidt [8] succeeded in that task, via carefully setting up an efficient inductive scheme. Their bound was further improved by Wooley [12].

Theorem 1.3 (Wooley). *Theorem 1.2 holds with*

$$V_{d,s} = 2s^{2^{d-1}} \phi_d^{2^{d-2}} \prod_{i=2}^{d-1} (\phi_i + 1)^{2^{i-2}}.$$

1.2. **Density of solutions.** Suppose now that K is infinite. Bik, Draisma and Snowden recently [3] proved a certain parallel of theorem 1.2 in this setting. To state it requires the following definition.

Definition 1.4. The *Birch rank* of f is

$$\text{Brk}(f) = \min(n, \text{codim}_{\mathbf{A}^n}(x : \nabla f(x) = 0)).$$

The minimum only plays a role when f is linear. More generally, the collective Birch rank of $f_1, \ldots, f_s$ is

$$\text{Brk}(f_1, \ldots, f_s) = \min\{\text{Brk}(a_1 f_1 + \ldots + a_s f_s) : 0 \neq a \in \overline{K}^s\}.^{[1]}$$

Bik, Draisma and Snowden's result is as follows.[2]

Theorem 1.5. *(Bik, Draisma and Snowden) There exists a constant $B_{d,s} = B(d, s, \varphi_2, \ldots, \varphi_d)$ such that if $\text{Brk}(f_1, \ldots, f_s) > B_{d,s}$ then the K-points of $Z(f_1, \ldots, f_s)$ are Zariski dense.*

As in theorem 1.2, the constant $B_{d,s}$ that can be extracted from their argument is extremely large. Our main result is an effective version of theorem 1.5.

Assume henceforth that the characteristic of K is either zero or greater than d.

Theorem 1.6. *Theorem 1.5 holds for $B_{d,s} = s^{2^{d-1}} C_d$, where*

$$C_d = 2^{(2\varphi)^{d-2}} \prod_{k=2}^{d} (\phi_k + 11)^{2^{d-2}(\varphi^{d-k-1} + 2^{k-d}) + 4}$$

and $\varphi = \frac{1+\sqrt{5}}{2} \approx 1.62$ is the golden ratio.

1.3. **Corollaries.** Bounds for ϕ_d were obtained by Skinner [11] when K is a finite extension of $\mathbf{Q}_p$. Plugging these bounds into theorem 1.6 yields the following corollary.

Corollary 1.7. *If K is a finite extension of $\mathbf{Q}_p$ and*

$$\text{Brk}(\underline{f}) > s^{2^{d-1}} (4d)^{(2\varphi)^{d-1} + 2^d + 8d} \cdot 2^{(2\varphi)^{d-2}}$$

then the K-points of $Z(\underline{f})$ are Zariski dense.

This is of particular interest for the Hardy-Littlewood circle method.

Definition 1.8. Let K be a number field. Following Schmidt [9], we say that:
   (1) $\underline{f}$ is a *Hardy-Littlewood system* (HLS for short) when the distribution of integer zeros of $\underline{f}$ satisfies the asymptotic formula predicted by the Hardy-Littlewood circle method and
   (2) $\underline{f}$ is a *proper Hardy-Littlewood system* (PHLS for short) when it is a HLS with a positive coefficient for the leading term.

---

[1]This is not the standard definition, but it differs from it by at most $s - 1$.

[2]Their result is phrased in terms of the strength of $f_1, \ldots, f_s$, but is equivalent to theorem 1.5 as stated. This is because Birch rank and strength are both bounded in terms of each other, see [4].

Example 1.9. Let $K = \mathbf{Q}$. A HLS satisfies

$$\#\{x \in [-P, P]^n \cap \mathbf{Z}^n : f_i(x) = 0 \text{ for all } 1 \le i \le s\} = \mathfrak{J}\mathfrak{S} P^{n-ds} + o(P^{n-ds}),$$

where $\mathfrak{J}, \mathfrak{S}$ are the *singular integral* and *singular series* respectively. A PHLS is one for which also $\mathfrak{J}\mathfrak{S} > 0$.

Skinner [10] established a sufficient criterion for $\underline{f}$ to be a PHLS. Together with corollary 1.7, this yields:

Corollary 1.10. *Suppose either that d is odd or else that K is totally imaginary. Then $\underline{f}$ is a PHLS as soon as*

$$\mathrm{Brk}(\underline{f}) > s^{2^{d-1}} (4d)^{(2\varphi)^{d-1} + 2^d + 8d} \cdot 2^{(2\varphi)^{d-2}}.$$

We are not aware of the existence of such a bound in the literature. For $K = \mathbf{Q}$, Schmidt used other methods to prove a result [9, supplement to theorem II] which implies that any system of forms of odd degree d satisfying $\mathrm{Brk}(\underline{f}) \ge C_d \cdot s^{2^{d-1}+1}$ is a PHLS. In fact, together with the recent results of Baily and the author [2], corollary 1.10 can also be used to give a new proof of Schmidt's result, albeit with a weaker constant C'_d.

Remark 1.11. Skinner's result was generalized by Lee to the setting of function fields [7]. Theorem 1.6 implies the analog of corollary 1.10 in that context as well.

1.4. **Structure of the paper.** The next section begins by proving corollaries 1.7 and 1.10, after which we state our three main technical propositions. We end section 2 by briefly discussing some key ideas in our proof. In section 3, theorem 1.6 is deduced from the three main propositions. Section 4 collects various geometric results required for the proofs of the main propositions, which are then carried out in sections 5-7.

1.5. **Acknowledgements.** The author thanks Tamar Ziegler for introducing him to Schmidt's result [9] and related work. Many thanks are due to Andrew Snowden for useful discussions about the proof of theorem 1.5. Trevor Wooley graciously shared his knowledge of the state of the art for various problems regarding forms in many variables. Last but not at all least, thank you to Noa for her unflagging encouragement and support throughout this project.

2. Proof of corollaries and statement of propositions

We begin this section by proving the corollaries of theorem 1.6.

2.1. **Proof of corollaries 1.7 and 1.10.**

Proof of corollary 1.7. Skinner proved [11] $\phi_k \le 8k^2$. This implies

$$\prod_{k=2}^{d} (\phi_k + 11)^{2^{d-2}(2^{k-d} + \varphi^{d-k-1}) + 4} \le (4d)^{2^{d-1} \sum_{k=2}^{d} (2^{k-d} + \varphi^{d-k-1}) + 8d}$$

$$\le (4d)^{(2\varphi)^{d-1} + 2^d + 8d}.$$

Plugging this into the bound of theorem 1.6 completes the proof. $\qquad\square$

Corollary 1.10 is a consequence of corollary 1.7 combined with the following result of Skinner [10].

Theorem 2.1 (Skinner). *Let K be a number field. If $\mathrm{Brk}(\underline{f}) > d2^d s^2$ then $\underline{f}$ is a HLS.[3] Moreover, if $Z(\underline{f})$ contains a non-singular point over every finite and infinite place of K, then $\underline{f}$ is a PHLS.*

Proof of corollary 1.10. The assumption that either d is odd or K is totally imaginary ensures the existence of a smooth point at every infinite place. By corollary 1.7, the assumption on $\mathrm{Brk}(\underline{f})$ ensures the existence of a smooth point at every finite place. $\square$

2.2. **Three propositions.** We begin with a more general version of theorem 1.6. Consider a system of forms $\underline{f} = (f_{i,j})_{i \in [d], j \in [s_i]}$, where each $f_{i,j}$ has degree i. The definition of Birch rank generalizes as follows.

Definition 2.2. The Birch rank of $\underline{f}$ is
$$\mathrm{Brk}(\underline{f}) = \min_{i \in [d]} \mathrm{Brk}(f_{i,1}, \ldots, f_{i,s_i}).$$

Theorem 1.6 is a special case of the following result.

Theorem 2.3. *Let $s = \max_{i \in [d]} s_i^{2^{i-d}}$. If*
$$\mathrm{Brk}(\underline{f}) > s^{2^{d-1}} 2^{(2\varphi)^{d-2}} \prod_{k=2}^{d} (\phi_k + 11)^{2^{d-2}(\varphi^{d-k-1} + 2^{k-d}) + 4}$$
then the K-points of $Z(\underline{f})$ are Zariski dense.

The proof of theorem 2.3 relies on three main propositions. Before we can state them, we need to set up some definitions.

Definition 2.4. Define $B = B(s_d, \ldots, s_1)$ to be the minimal integer such that for any collection of forms $\underline{f} = (f_{i,j})_{i \in [d], j \in [s_i]}$ with $\mathrm{Brk}(\underline{f}) > B$, the K-points of $Z(\underline{f})$ are Zariski dense. We sometimes write $B_d(s_d, \ldots, s_1)$ if the degree is otherwise unclear. For example, $B_1(s) = B(s) = s$.

There is a special case of this which will be particularly important for us. A *non-degenerate system of diagonal forms* $\underline{f} = (f_1, \ldots, f_{d+1})$ is a collection where:

- Each $f_i \in K[x_1, \ldots, x_n]$ is a diagonal form with nonzero coefficients,
- The degree of f_i equals i for $1 \le i \le d$ and $\deg(f_{d+1}) > d$.

Definition 2.5. Define β_d to be the minimal integer such that the following holds. Whenever $n > \beta_d$ and $\underline{f} = (f_1, \ldots, f_{d+1})$ is a non-degenerate system of diagonal forms then there exists $x \in \bar{K}^n$ with $f_1(x) = \ldots = f_d(x) = 0$ and $f_{d+1}(x) \ne 0$.

Our first proposition is a bound for β_d, which may be of independent interest.

---

[3]The bound we cite differs from the one in [10] to account for our different definition of Birch rank. This has no effect on corollary 1.10.

Proposition 2.6 (Bound for non-degenerate diagonal systems)**.** *There exists a sequence of even integers n_d such that for $d \geq 1$ we have*

$$\beta_d + 1 \leq n_d \leq 2^{1+\varphi^{d-1}} \prod_{k=2}^{d} (\phi_k + 11)^{\varphi^{d-k}}.$$

To state our next propositions, recall that the Taylor expansion of a form f of degree d is given by

$$f(x_1 + \ldots + x_m) = \sum_{|e|=d} f^e(x_1, \ldots, x_m),$$

where each f^e is multi-homogeneous of multi-degree e. Write $\left(\!\!\binom{m}{d}\!\!\right) = \binom{m+d-1}{d}$ for the number of terms in the above sum. For a collection of forms $\underline{f} = (f_{i,j})_{i\in[d], j\in[s_i]}$, write $T_m(\underline{f}) = (f_{i,j}^e)_{i\in[d], j\in[s_i], |e|=i}$ for the forms on V^m given by the Taylor expansion. These are also the coefficients of the polynomials obtained by restricting $\underline{f}$ to $\mathrm{span}(x_1, \ldots, x_m)$. We abbreviate $Z_m(\underline{f}) := Z(T_m(\underline{f}))$.

Assume $s_d > 0$ and write $\underline{f} = \underline{g} \sqcup \{f\}$, where $f = f_{d,s_d}$. Recall the following useful definition from [3].

Definition 2.7. Subspaces $U_1, \ldots, U_m$ are f-orthogonal if

$$f(u_1 + \ldots + u_m) = f(u_1) + \ldots + f(u_m)$$

for any tuple of vectors $u_i \in U_i$. Equivalently, $f^e(u_1, \ldots, u_m) = 0$ for any exponent e with multiple nonzero entries.

We consider the algebraic set

$$D_m = ((w_1, w_2, w_3, v_1, \ldots, v_{m-3}) \in Z_m(\underline{g}) : \mathrm{span}(w_1, w_2, w_3), v_1, \ldots, v_{m-3} \text{ are } f\text{-orthogonal}).$$

Definition 2.8. Let $\delta = \delta_m(s_d, \ldots, s_1)$ be the minimal number such that whenever $\mathrm{Brk}(\underline{f}) > \delta$ the K-points of D_m are Zariski dense.

Our second main proposition bounds $B(s_d, \ldots, s_1)$ in terms of $\delta_m(s_d, \ldots, s_1)$.

Proposition 2.9 (Diagonalization implies density)**.** *For $d \geq 2$ define*

$$m = m_d = (\phi_d + 9) \left\lceil \frac{\beta_{d-2} + 1}{2} \right\rceil.$$

Then

$$B(s_d, \ldots, s_1) \leq \max\left(\delta_m(s_d, \ldots, s_1), 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 1\right).$$

Conversely, $\delta_m(s_d, \ldots, s_1)$ can be bounded in terms of $B(s_d^*, \ldots, s_1^*)$, where $(s_d^*, \ldots, s_1^*)$ is a lexicographically smaller sequence. This is the content of our third main proposition.

Proposition 2.10 (Efficient diagonalization)**.** *For any m we have*

$$\delta_m(s_d, \ldots, s_1) \leq 2 \sum_{i=1}^{d} s_i m^i + B(s_d^*, \ldots, s_1^*),$$

where $s_d^* = s_d - 1$ and $s_i^* = s_i + m s_{i+1} + \ldots + s_d m^{d-i}$ for $1 \leq i \leq d - 1$.

Taken together, these three propositions allow us to prove theorem 2.3 by induction on the sequence $(s_d, \ldots, s_1)$. This analysis is carried out in the next section.

2.3. **Proof idea.** The proof of theorem 2.3 requires synthesizing the techniques used in the proofs of theorems 1.3 and 1.5. There are several technical challenges that arise, but we do not dwell on these here. Instead we emphasize a subtle but significant change in perspective compared to the proof of theorem 1.5 – this is the use of Birch rank rather than strength as our fundamental quantity.

Although Birch rank and strength are bounded in terms of each other (see [4]), there are many simple examples (such as $x^2 + y^2$) where they differ by a factor of 2. If the right hand side of proposition 2.10 were so much as doubled, then our inductive argument would no longer result in polynomial bounds for theorem 2.3. Thus, a key reason that the proof techniques of theorem 1.3 may be adapted to the setting of Birch rank is the *additive behavior* it exhibits – in the sense that proposition 2.10 holds.

Another important property of Birch rank (which also holds for strength) is that it is preserved under Taylor expansion. This is the content of lemma 4.2 and it is the reason we need to assume that the characteristic is either zero or greater than d. Since our argument proceeds by restricting forms to subspaces with desirable properties, this property plays a major role in our proofs.

3. Proof of theorem 2.3

In this section we deduce the main theorem from propositions 2.9, 2.10 and 2.6. We begin with some simple observations.

Lemma 3.1. *The function $B(s_d, \ldots, s_1)$ is monotonically increasing. Furthermore,*

$$B(s_d, \ldots, s_1) \le 2s_1 + B(s_d, \ldots, s_2, 0)$$
$$\delta_m(s_d, \ldots, s_1) \le 2s_1 + \delta_m(s_d, \ldots, s_2, 0).$$

Proof. The first claim is that

$$B(s_d, \ldots, s_i + 1, \ldots, s_1) \ge B(s_d, \ldots, s_i, \ldots, s_1).$$

Denote the expression on the left-hand side by B and suppose $\underline{f} = (f_{i,j})_{i \in [d], j \in [s_i]}$ is a collection of forms with $\mathrm{Brk}(\underline{f}) > B$. We need to show that K-points are dense in $Z(\underline{f})$. Let $W \cong V$ and let g be a smooth K-form of degree i on W. Consider the collection of forms $\underline{f}' = (f_{i,j} \circ \pi_1)_{i \in [d], j \in [s_i]} \cup \{g \circ \pi_2\}$ on $V \times W$, where $\pi_1 : V \times W \to V$ and $\pi_2 : V \times W \to W$ are the projections. It's easily verified that $\mathrm{Brk}(\underline{f}') = \mathrm{Brk}(\underline{f}) > B$ so that K-points are dense in $Z(\underline{f}')$. Since $Z(\underline{f}') = Z(\underline{f}) \times Z(g)$, we deduce that the K-points of $Z(\underline{f})$ are dense, as desired.

Now, supposing $\mathrm{Brk}(\underline{f}) > B(s_d, \ldots, s_2, 0) + 2s_1$, we claim that the K-points of $Z(\underline{f})$ are dense. Let $U \subset V$ be the K-subspace of codimension s_1 where the linear forms in $\underline{f}$ vanish and consider $\underline{g} = (g_{i,j})_{2 \le i \le d, j \in [s_i]}$, where $g_{i,j} = f \upharpoonright_U$. By lemma 4.4, we have $\mathrm{Brk}(\underline{g}) > B(s_d, \ldots, s_2, 0)$, so that the K-points of $Z(\underline{g}) = Z(\underline{f})$ are dense, as claimed. The proof of the corresponding inequality for δ_m is identical. $\square$

Iterating propositions 2.9 and 2.10 allows us to bound B_d in terms of B_{d-1} and m_d.

Lemma 3.2 (Degree-lowering). *Let* $m = m_d$ *be as in proposition 2.9. Then*

$$B_d(s_d, \ldots, s_1) \le 2 \sum_{i=1}^{d} s_i(ms_d)^i + B_{d-1}(r_{d-1}, \ldots, r_1),$$

where $r_i = s_i + (ms_d)s_{i+1} + \ldots + (ms_d)^{d-i}s_d$.

Proof. We begin by noting that

$$2 \sum_{i=1}^{d} s_i m^i \ge 2 \sum_{i=1}^{d} s_i \left(\binom{m}{i} \right) + d' - 1.$$

Combining propositions 2.9 and 2.10 and using the above observation gives

$$B_d(s_d, s_{d-1}, \ldots, s_2, s_1) \le 2 \sum_{i=1}^{d} s_i m^i + B(s_d^*, \ldots, s_2^*, s_1^*),$$

where $s_d^* = s_d - 1$ and $s_i^* = \sum_{k \ge i} s_i m^{k-i}$ for $1 \le i \le d - 1$. Iterating this, let $s_i^{(j)}$ be the number of degree i forms after j iterations. Thus,

$$s_i^{(0)} = s_i \ , \ s_d^{(j)} = s_d - j$$
$$s_i^{(j+1)} = \sum_{k \ge i} s_k^{(j)} m^{k-i}, \quad 1 \le i \le d - 1.$$

After s_d iterations we end up with

$$(1) \qquad B_d(s_d, s_{d-1}, \ldots, s_2, s_1) \le 2 \sum_{j=0}^{s_d-1} \sum_{i=1}^{d} s_i^{(j)} m^i + B_{d-1}(s_{d-1}^{(s_d)}, \ldots, s_2^{(s_d)}, s_1^{(s_d)}).$$

Note that

$$(2) \qquad\qquad s_i^{(j)} \le \sum_{\ell \ge i} s_\ell (jm)^{\ell - i}.$$

This is proved by downward induction on i, then upward on j. The base cases $i = d$ and $j = 0$ are clear. Supposing inequality (2) holds for $i' > i$ and for $s_i^{(j)}$, we have

$$s_i^{(j+1)} = \sum_{k \ge i} s_k^{(j)} m^{k-i} \le \sum_{k \ge i} m^{k-i} \sum_{\ell \ge k} s_\ell (jm)^{\ell-k}$$

$$= \sum_{\ell \ge i} s_\ell m^{\ell-i} \sum_{k=i}^{\ell} j^{\ell-k} \le \sum_{\ell \ge i} s_\ell m^{\ell-i}(j+1)^{\ell-i},$$

which proves the inductive step.

Therefore,

$$(3) \quad \sum_{j=0}^{s_d-1} \sum_{i=1}^{d} s_i^{(j)} m^i \le \sum_{j=0}^{s_d-1} \sum_{i=1}^{d} m^i \sum_{\ell \ge i} s_\ell (jm)^{\ell-i} = \sum_{\ell=1}^{d} s_\ell m^\ell \sum_{j=0}^{s_d-1} \sum_{i=1}^{\ell} j^{\ell-i}$$

$$\le \sum_{\ell=1}^{d} s_\ell m^\ell \sum_{j=0}^{s_d-1} (j+1)^{\ell-1} \le \sum_{\ell=1}^{d} s_\ell (ms_d)^\ell$$

The proof is completed by plugging inequalities (2) and (3) into (1) and using the monotonicity property from lemma 3.1.   □

We now specialize this bound.

Lemma 3.3. *We have*
$$B(s, s^2) \le 2s^2(m_2 + 1)^2$$
and for every $d \ge 3$,
$$B_d(s, s^2, \dots, s^{2^{d-1}}) \le ds^{2^{d-1}}(m_2 + 1)^2(m_3 + 1)^3 \prod_{i=4}^{d}(m_i + 1)^{2^{i-2}}.$$

Proof. The proof is by induction on d. When proving the claim for d we write $m = m_d$. For $d = 2$ we combine propositions 2.9, 2.10 and lemma 3.1 to get
$$B(s, 0) \le 2sm^2 + B(s - 1, ms) \le 2s(m^2 + m) + B(s - 1, 0)$$
$$\le \dots \le 2s^2(m^2 + m).$$
By lemma 3.1,
$$B(s, s^2) \le 2s^2 + B(s, 0) \le 2s^2(m + 1)^2.$$
For $d = 3$, we apply lemmas 3.2 and 3.1 to get
$$B_3(s, s^2, s^4) \le 2(m^2 s^4 + m^3 s^4 + s^4) + B_2(s^2 + ms^2, ms^3 + m^2 s^3)$$
$$\le 2s^4(m + 1)^3 + B_2(s^2(m + 1), s^4(m + 1)^2)$$
$$\le 2s^4(m_3 + 1)^3 + 2s^4(m_3 + 1)^2(m_2 + 1)^2 \le 3s^4(m_2 + 1)^2(m_3 + 1)^3,$$
where the penultimate inequality used the bound for $d = 2$.

For $d \ge 4$, plugging in $s_i = s^{2^{d-i}}$ to lemma 3.2 yields
$$B_d(s, s^2, \dots, s^{2^{d-2}}, 0) \le 2\sum_{i=2}^{d} s^{2^{d-i}}(ms)^i + B_{d-1}(r_{d-1}, \dots, r_2, r_1)$$
$$\le 2s^{2^{d-1}} \cdot (m + 1)^d + B_{d-1}(r_{d-1}, \dots, r_2, r_1)$$
For $1 \le i \le d - 1$ note that
$$r_i = s^{2^{d-i}} + (sm)s^{2^{d-i-1}} + \dots + (sm)^{d-i}s$$
$$\le s^{2^{d-i}}(m + 1)^{d-i} \le [s^2(m + 1)]^{2^{d-1-i}}.$$
By inductive hypothesis and lemma 3.1,
$$B_d(s, s^2, \dots, s^{2^{d-1}}) \le s^{2^{d-1}} \cdot [2(m_d + 1)^d + 2]$$
$$+ (s^2(m_d + 1))^{2^{d-2}}(d - 1)(m_2 + 1)^2(m_3 + 1)^3 \prod_{i=4}^{d-1}(m_i + 1)^{2^{i-2}}$$
$$\le ds^{2^{d-1}}(m_2 + 1)^2(m_3 + 1)^3 \prod_{i=4}^{d}(m_i + 1)^{2^{i-2}},$$
which completes the induction step.   □

We are now ready to prove our main result.

Proof of theorem 2.3. By lemma 3.1, it's enough to prove the bound for $B(s, s^2, \ldots, s^{2^{d-1}})$. Thanks to lemma 3.3, all that remains is to bound $(m_2+1)^2(m_3+1)^3 \prod_{i=4}^d (m_i+1)^{2^{i-2}}$. This will be accomplished using proposition 2.6.

By definition, $m_2 + 1 \le \phi_2 + 11$. By proposition 2.6 we have $m_3 + 1 \le 2(\phi_3 + 11)$ and in general

$$\prod_{i=4}^d (m_i+1)^{2^{i-2}} \le 2^{\sum_{i=4}^d 2^{i-2}\varphi^{i-3}} \prod_{i=4}^d (\phi_i+11)^{2^{i-2}} \prod_{k=2}^{i-2}(\phi_k+11)^{2^{i-2}\varphi^{i-2-k}}$$

$$\le 2^{(2\varphi)^{d-2}} \prod_{k=2}^d (\phi_k+11)^{2^{k-2}+\sum_{i=k+2}^d 2^{i-2}\varphi^{i-2-k}}$$

$$\le 2^{(2\varphi)^{d-2}} \prod_{k=2}^d (\phi_k+11)^{2^{d-2}(2^{k-d}+\varphi^{d-k-1})}.$$

$\square$

4. Some geometric results

We now collect some preliminary results that will be needed in the course of the proofs of propositions 2.9 and 2.10.

4.1. **Taylor expansion and Birch rank.** We begin by recording several elementary identities for the Taylor expansion.

Lemma 4.1. *The terms of the Taylor expansion satisfy the following formulas:*

(1) $f^{(1,\ldots,1)}(x_1, \ldots, x_d) = f^{(1,\ldots,1)}(x_{\sigma(1)}, \ldots, x_{\sigma(d)})$ *for any permutation* $\sigma \in S_d$,

(2) $f^{(1,d-1)}(h, x) = \nabla f(x) \cdot h$,

(3) $e_1! \ldots e_m! f^e(x_1, \ldots, x_m) = f^{(1,\ldots,1)}(x_1, \ldots, x_1, \ldots, x_m, \ldots, x_m)$, *where each* x_i *appears* e_i *times and*

(4) *for any scalars* $\lambda_1, \ldots, \lambda_m$,

$$\frac{\partial f^e}{\partial x_k}(\lambda_1 x, \ldots, \lambda_m x) = \binom{d-1}{e_1, \ldots, e_k-1, \ldots, e_m} \cdot \lambda_1^{e_1} \ldots \lambda_k^{e_k-1} \ldots \lambda_m^{e_m} \nabla f(x).$$

Proof. For the first property, note that for any permutation σ we have

$$f(x_1 + \ldots + x_d) = f(x_{\sigma(1)} + \ldots + x_{\sigma(d)}).$$

Comparing the multi-homogeneous terms of multi-degree $(1, \ldots, 1)$ on both sides proves the claim.

For the second formula, first note that $f^{(0,d)}(h, x) = f(x)$ follows from setting $h = 0$. Therefore, $f(x + th) = f(x) + t f^{(1,d-1)}(h, x) + O(t^2)$, proving the claim.

To prove the third formula, compute the coefficient of $\prod_{i=1}^m \prod_{j=1}^{e_i} t_{i,j}$ in the expression $f\left(\sum_{i=1}^m \sum_{j=1}^{e_i} t_{i,j} x_i\right)$ in two different ways. On the one hand, using the d-term Taylor expansion, this equals $f^{(1,\ldots,1)}(x_1, \ldots, x_1, \ldots, x_m, \ldots, x_m)$. On the other hand, using the m-term Taylor expansion, this equals $e_1! \ldots e_m! f^e(x_1, \ldots, x_m)$.

For the fourth identity, both sides are multi-homogeneous of the same multi-degree, so we may assume without loss of generality that $\lambda_1 = \ldots = \lambda_m = 1$ and that $k = 1$. Using formulas (1-3) and the multi-linearity of $f^{(1,\ldots,1)}$, we get

$$
\begin{aligned}
e_1! \ldots e_m! f^e(x + th, x, \ldots, x) &= f^{(1,\ldots,1)}(x + th, \ldots, x + th, x, \ldots, x) \\
&= e_1! \ldots e_m! f^e(x, \ldots, x) + t e_1 \cdot f^{(1,\ldots,1)}(h, x, \ldots, x) + O(t^2) \\
&= e_1! \ldots e_m! f^e(x, \ldots, x) + t e_1 (d-1)! f^{(1,d-1)}(h, x) + O(t^2) \\
&= e_1! \ldots e_m! f^e(x, \ldots, x) + t e_1 (d-1)! \nabla f(x) \cdot h + O(t^2),
\end{aligned}
$$

which proves the claim. $\square$

This last formula has a very useful consequence.

Lemma 4.2. *For any $m \geq 1$, $\mathrm{Brk}(T_m(\underline{f})) \geq \mathrm{Brk}(\underline{f})$.*

Proof. The statement is about each degree separately, so we may assume without loss of generality that $\underline{f} = (f_1, \ldots, f_s)$ is composed entirely of forms of degree d. For $d = 1$ the claim is immediate so assume also that $d > 1$. Let $g = \sum_{j \in [s], |e| = d} a_{j,e} f_j^e$ be a nontrivial linear combination and $S(g) = (\underline{x} : \nabla g(\underline{x}) = 0)$. For $\lambda = (\lambda_1, \ldots, \lambda_m) \neq 0$, consider the diagonal embedding $V \cong V_\lambda \subset V^m$ given by the image of $x \mapsto (\lambda_1 x, \ldots, \lambda_m x)$.

The intersection $S(g) \cap V_\lambda$ is nonempty because it contains 0, and we have

$$
(4) \qquad\qquad \mathrm{codim}_{V_\lambda}(S(g) \cap V_\lambda) \leq \mathrm{codim}_{V^m} S(g) = \mathrm{Brk}(g).
$$

By assumption, there is some e_0 for which a_{j,e_0} are not all zero. Let $k \in [m]$ be such that $(e_0)_k > 0$. By lemma 4.1,

$$
\frac{\partial g}{\partial x_k}(\lambda_1 x, \ldots, \lambda_m x) = \sum_{j \in [s]} \nabla f_j(x) \sum_{|e|=d, e_k > 0} a_{j,e} \binom{d-1}{e_1, \ldots, e_k - 1, \ldots, e_m} \cdot \lambda_1^{e_1} \ldots \lambda_k^{e_k - 1} \ldots \lambda_m^{e_m}.
$$

Choose λ such that the coefficient of some $\nabla f_j(x)$ above is nonzero and let

$$
f = \sum_{j \in [s]} f_j \sum_{|e|=d, e_k > 0} a_{j,e} \binom{d-1}{e_1, \ldots, e_k - 1, \ldots, e_m} \cdot \lambda_1^{e_1} \ldots \lambda_k^{e_k - 1} \ldots \lambda_m^{e_m}
$$

be the corresponding nontrivial linear combination of the f_j. We have

$$
S(f) \supset (x \in V : (\lambda_1 x, \ldots, \lambda_m x) \in S(g)),
$$

so that

$$
\mathrm{Brk}(\underline{f}) \leq \mathrm{Brk}(f) = \mathrm{codim}_V S(f) \leq \mathrm{codim}_{V_\lambda}(S(g) \cap V_\lambda).
$$

Combining this with inequality (4) gives $\mathrm{Brk}(\underline{f}) \leq \mathrm{Brk}(g)$ as desired. $\square$

4.2. Birch rank, restriction to subspaces and singularities.

We have already mentioned that our arguments will often rely on restricting forms to subspaces. An issue that arises is that the Birch rank of the restricted form may be smaller than that of the original form.

Consider, for example, the quadratic form $f = x^2 - y^2$ and the hyperplane $U = \{(x, x) : x \in K\}$. We have $\mathrm{Brk}(f) = 2$ but $\mathrm{Brk}(f \restriction_U) = 0$, so the Birch rank decreases by 2. Fortunately, this is the largest possible drop in Birch rank upon restriction to

a hyperplane – as we will soon see. This is a consequence of the following result of Ananyan and Hochster [1, Theorem 2.4].

Lemma 4.3. *Let M be a $h \times n$ matrix such that for each $i \in [h]$ the i-th row consists of forms of degree d_i in $K[x_1, \ldots, x_n]$ and the d_i are mutually distinct integers. If for each $i \in [h]$ we have*

$$\operatorname{codim}_{\mathbf{A}^n}(x : M_{i,j}(x) = 0 \ \forall j \in [n]) \ge c$$

then

$$\operatorname{codim}_{\mathbf{A}^n}(x : \operatorname{rk}(M(x)) < h) \ge c - h + 1.$$

We can now easily bound the decrease in Birch rank of the restriction of a form to a subspace.

Lemma 4.4. *Let $\underline{f}$ be a collection of forms of degree two or higher and $U \subset K^n$ a subspace of codimension c. Then $\operatorname{Brk}(\underline{f} \restriction_U) \ge \operatorname{Brk}(\underline{f}) - 2c$.*

Proof. It is enough to prove this in the case of a single form f and $c = 1$. In this case, $U = (x \in K^n : a_1 x_1 + \ldots + a_n x_n = 0)$ for some $0 \ne a \in K^n$. We have

$$\operatorname{Brk}(f \restriction_U) = \operatorname{codim}_U(x \in U : \nabla f(x) \in \operatorname{span}(a))$$
$$\ge \operatorname{codim}_{\mathbf{A}^n}(x \in \mathbf{A}^n : (\nabla f(x), a) \text{ are linearly dependent}) - 1 \ge \operatorname{Brk}(f) - 2,$$

where the final inequality follows from lemma 4.3. $\square$

Lemma 4.3 also implies a close connection between $\operatorname{Brk}(\underline{f})$ and the codimension of

$$S(\underline{f}) = (x \in \mathbf{A}^n : (\nabla f_{i,j}(x))_{i \in [d], j \in [s_i]} \text{ are linearly dependent}).$$

Lemma 4.5. *We have*

$$\operatorname{codim}_{\mathbf{A}^n} S(\underline{f}) \ge \operatorname{Brk}(\underline{f}) - \sum_{i=1}^{d} s_i - d' + 2,$$

where $d' = |\{i : s_i \ne 0\}|$.

Proof. Let $s = \sum_{i=1}^{d} s_i$. Write $S = \cup_{a \in \mathbf{P}^{s-1}} S_a$, where

$$S_a = \left(x : \sum_{i=1}^{d} \sum_{j=1}^{s_i} a_{i,j} \nabla f_{i,j}(x) = 0\right).$$

It's enough to show that for any fixed $a \in \mathbf{P}^{s-1}$ we have $\operatorname{codim}_{\mathbf{A}^n} S_a \ge \operatorname{Brk}(\underline{f}) - d' + 1$. Fixing a, let $I = \{i \in [d] : (a_{i,j})_{j \in [s_i]} \ne (0, \ldots, 0)\}$. For each $i \in I$, let $f_i = \sum_{j \in [s_i]} a_i f_{i,j}$. We now have the inclusion

$$S_a \subset (x \in \mathbf{A}^n : (\nabla f_i(x))_{i \in I} \text{ are linearly dependent}).$$

Applying lemma 4.3 yields

$$\operatorname{codim}_{A^n} S_a \ge \operatorname{codim}_{\mathbf{A}^n}(x \in \mathbf{A}^n : (\nabla f_i(x))_{i \in I} \text{ are linearly dependent})$$
$$\ge \operatorname{Brk}(\underline{f}) - |I| + 1 \ge \operatorname{Brk}(\underline{f}) - d' + 1.$$

$\square$

There is an important criterion relating singularities to irreducibility.

Lemma 4.6. *Let $f_1, \ldots, f_s$ be a sequence of forms of positive (possibly different) degrees such that $Z = Z(f_1, \ldots, f_s)$ is a complete intersection of codimension s. If*

$$\operatorname{codim}_Z(x \in Z : (\nabla f_1(x), \ldots, \nabla f_s(x) \text{ are linearly dependent}) > 1,$$

then the ideal $(f_1, \ldots, f_s)$ is prime. In particular, $Z(f_1, \ldots, f_s)$ is irreducible.

Proof. This is a well known fact that follows from [6, Theorem 18.15b]. Note that the hypothesis of that theorem is satisfied by [6, Proposition 18.13] and that $Z(f_1, \ldots, f_s)$ is connected since it's a cone. Therefore, if its coordinate ring is a product of domains then it must in fact be a domain. $\square$

As a consequence of the results of this section, we have:

Lemma 4.7. *If $\operatorname{Brk}(\underline{f}) > 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 1$ and $\underline{g} \subset T_m(\underline{f})$ consists of t equations then $Z(\underline{g})$ is an irreducible complete intersection of codimension t.*

Proof. It follows from lemmas 4.2 and 4.5 that

$$(5) \qquad \operatorname{codim}_{V^m} S(\underline{g}) \geq \operatorname{Brk}(\underline{f}) - \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) - d' + 2 > \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + 1.$$

The right hand side is larger than t, so in particular the zero locus $Z = Z(\underline{g})$ is a complete intersection of codimension t. Inequality (5) also implies that we have $\operatorname{codim}_Z Z \cap S(\underline{g}) > 1$, so Z is irreducible by lemma 4.6. $\square$

5. Proof of proposition 2.10

The proof will rely on the following geometric result.

Lemma 5.1. *Let $m > 1$ and suppose $\operatorname{Brk}(\underline{f}) > 2 \sum_{i=1}^{d} s_i m^i + c$. Let h be a polynomial which doesn't vanish identically on D_m. Then for a dense open set of $\underline{x} \in D_{m-1}$,*

(1) The polynomial $h(\underline{x}, \cdot)$ doesn't vanish identically on the fiber $D_m(\underline{x})$ and
(2) We have $\operatorname{Brk}\left((f_{i,j}^e(\underline{x}, \cdot))_{e_m > 0}\right) > c$.

Proof. Note that

$$2 \sum_{i=1}^{d} s_i m^i \geq 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 1,$$

so by lemma 4.7 both D_{m-1} and D_m are irreducible. It's therefore enough to show that each of the conditions holds on a nonempty open subset of D_{m-1}.

For the first condition, irreducibility of D_m implies that the proper subset $D' = (h = 0) \cap D_m$ has dimension $\dim D' < \dim D_m$. Therefore, for a dense open subset of $\underline{x} \in D_{m-1}$ the fiber satisfies $\dim D'(\underline{x}) < \dim D_m(\underline{x})$ as desired.

For the second condition, again by irreducibility it's enough to show this for each fixed $e_m = k > 0$. Let

$$S = \left(\underline{x} \in V^{m-1} : \operatorname{Brk}\left((f_{i,j}^e(\underline{x}, \cdot))_{e_m = k}\right) \leq c\right)$$

It's certainly enough to prove that

$$(6) \qquad \operatorname{codim}_{V^{m-1}} S > \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) > \operatorname{codim}_{V^{m-1}} D_{m-1}.$$

To prove this, consider some nontrivial linear combination $g = \sum_{e_m=k} a_{i,j}^e f_{i,j}^e(x_1, \ldots, x_m)$ and let

$$S_a = \left(\underline{x} \in V^{m-1} : \operatorname{Brk}(g(\underline{x}, \cdot) \le c)\right), \ \ Y_a = \left(\underline{x} \in V^m : \frac{\partial g}{\partial x_m}(\underline{x}) = 0\right).$$

For $0 \ne \lambda \in K^m$, we have

$$\operatorname{codim}_{V^m} Y_a \ge \operatorname{codim}_V \left(x \in V : (\lambda_1 x, \ldots, \lambda_m x) \in Y_a\right).$$

By lemma 4.1, the expression on the right hand side is

$$\operatorname{codim}_V \left(x \in V : \sum_{i \ge k, j \in [s_i]} \nabla f_{i,j}(x) \sum_{e_m=k} a_{i,j}^e \cdot \binom{i}{e_1, \ldots, e_{m-1}, k-1} \cdot \lambda_1^{e_1} \ldots \lambda_{m-1}^{e_{m-1}} \lambda_m^{k-1} = 0 \right).$$

Choosing λ such that the coefficients of $\nabla f_{i,j}(x)$ above are not all zero and applying lemma 4.5, we obtain

$$\operatorname{codim}_{V^m} Y_a \ge \operatorname{codim}_V S(\underline{f}) > \operatorname{Brk}(\underline{f}) - \sum_{i=1}^{d} s_i - d' \ge 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + c.$$

This gives

$$\operatorname{codim}_{V^{m-1}} S_a = \operatorname{codim}_{V^{m-1}} \left(\underline{x} : \operatorname{codim}_V Y_a(\underline{x}) \le c\right) > 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right).$$

Taking the union over $a \in \mathbf{P}^{\sigma-1}$ for $\sigma = \sum_{i=k}^{d} s_i \left(\!\!\binom{m}{i-k}\!\!\right)$ yields

$$\operatorname{codim}_{V^{m-1}} S > \min_{a \in \mathbf{P}^{\sigma-1}} \operatorname{codim}_{V^{m-1}} S_a - \sigma > \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right),$$

proving inequality (6). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can now prove the proposition.

Proof of proposition 2.10. Recall that our goal is to prove that

$$\delta_m(s_d, \ldots, s_2, s_1) \le 2 \sum_{i=1}^{d} s_i m^i + B(s_d^*, \ldots, s_1^*),$$

where $s_d^* = s_d - 1$ and $s_i^* = s_i + m s_{i+1} + \ldots + s_d m^{d-i}$ for $1 \le i \le d-1$. The proof is by induction on m. In the base case $m = 1$ we have $D_1 = Z(g)$ so by lemma 3.1,

$$\delta_1(s_d, \ldots, s_1) = B_d(s_d - 1, s_{d-1}, \ldots, s_1) \le B_d(s_d^*, \ldots, s_1^*).$$

Suppose now that the claim holds for $m-1$ and we wish to prove it for m. We assume that $\operatorname{Brk}(\underline{f}) > 2 \sum_{i=2}^{d} s_i m^i + B(s_d^*, \ldots, s_1^*)$ and that h is a polynomial not vanishing identically on D_m, and need to show that $(h \ne 0) \cap D_m$ contains a K-point. Let $U \subset D_{m-1}$ be the dense open subset guaranteed by lemma 5.1, with $c = B(s_d^*, \ldots, s_1^*)$.

By the inductive hypothesis, we may choose a K-point $\underline{x} \in U$. The fiber $D_m(\underline{x})$ is then defined by at most s_i^*-equations of degree i, and these equations are a subset of $(f_{i,j}^e(\underline{x},\cdot))_{e_m>0}$ so their Birch rank is greater than $B(s_d^*, \ldots, s_1^*)$. It follows that we may choose a K-point $y \in D_m(\underline{x}) \cap (h(\underline{x},\cdot) \neq 0)$, as desired. $\square$

6. Proof of proposition 2.6

The following result [3, Proposition 4.4] will play a key role in proving propositions 2.6 and 2.9.

Lemma 6.1. *Suppose that*

$$f(x,y,z,t_1,\ldots,t_m) = \alpha xy^{d-1} + \beta y^d + \gamma z^d + g(t_1,\ldots,t_m)$$

where $\alpha, \beta, \gamma \in K$, $\alpha\gamma \neq 0$ and g is some form. Then the K-points of $Z(f)$ are Zariski dense.

Definition 6.2. We say that f as above is *good*.

Proposition 2.6 is a direct consequence of the following lemma.

Lemma 6.3. *Define a sequence by $n_0 = 2, n_1 = 4$, and $n_d = \frac{(\phi_d+11)n_{d-1}n_{d-2}}{2}$ for $d \geq 2$. Then for all $d \geq 0$, n_d is an even integer and $\beta_d + 1 \leq n_d$.*

Proof. The statement about n_d being an even integer is immediate. We prove the inequality by induction on d. The case $d = 0$ is immediate. For $d = 1$, we have $\beta_1 \leq 2$ since for any $d \geq 2, n > 2$ and $a_1, \ldots, a_n \neq 0$, the form $a_1x_1^d + \ldots + a_nx_n^d$ is irreducible.

Suppose that that $\beta_{d'}+1 \leq n_{d'}$ for $d' < d$ and let $\underline{f} = (f_1, \ldots, f_{d+1})$ be a nondegenerate system of diagonal forms with $n \geq n_d$. By lemma 6.1, to prove the existence of a point $0 \neq x \in K^n$ with $f_{d+1}(x) \neq 0$ and $f_i(x) = 0$ for $i = 1, \ldots, d$, it's sufficient to construct a K-subspace L such that the following properties hold:

(1) The forms $f_1, \ldots, f_{d-1}$ vanish on L,
(2) The restriction $f_d \restriction_L$ is good and
(3) We have $\text{codim}_L Z(f_d, f_{d+1}) \cap L = 2$.

First, let $m = 3n_{d-1}$ and consider the subspace $K^m \cong L_2 \subset K^n$ spanned by the final m standard basis vectors. Write $L_2 = L_2^{(1)} \oplus L_2^{(2)} \oplus L_2^{(3)}$, where the $L_2^{(j)}$ are spanned by disjoint collections of n_{d-1} standard basis vectors. By the inductive hypothesis, there exist $w^{(j)} \in L_2^{(j)}$ with $f_1(w^{(j)}) = \ldots = f_{d-1}(w^{(j)}) = 0$ and $f_{d+1}(w^{(j)}) \neq 0$ for $j = 1, 2, 3$.

Next, let $m' = \frac{(\phi_d+1)n_{d-1}n_{d-2}}{2}+2n_{d-1}$ and consider the subspace $K^{m'} \cong U \subset K^n$ spanned by the first m' standard basis vectors. Note that $U \cap L_2 = \{0\}$ by the hypothesis on the size of n. We write $U = U^{(1)} \oplus \ldots \oplus U^{(m'/n_{d-1})}$, where the $U^{(j)}$ are spanned by disjoint collections of n_{d-1} standard basis vectors. By the inductive hypothesis, for each j there is a K-point $u^{(j)} \in U^{(j)}$ with $f_1(u^{(j)}) = \ldots = f_{d-1}(u^{(j)}) = 0$ and $f_d(u^{(j)}) \neq 0$. Thus far we have found linearly independent vectors $u^{(j)}, w^{(j)}$ such that $\text{span}(u^{(j)}, w^{(j)}) \subset Z(f_1, \ldots, f_{d-1})$, $f_d(u^{(j)}) \neq 0$ and $f_{d+1}(w^{(j)}) \neq 0$. By construction, $u^{(j)}, w^{(j)}$ are both f_d and f_{d+1} orthogonal.

Now, let $m'' = \frac{(\phi_d+1)n_{d-2}}{2}$, and $g(x_1, \ldots, x_{m''}) = f_d(x_1u^{(1)} + \ldots + x_{m''}u^{(m'')})$. By construction, g is a diagonal form with nonzero coefficients. Decompose $K^{m''} =

$V^{(1)} \oplus \ldots \oplus V^{(n_{d-2}/2)}$ where the $V^{(j)}$ are spanned by disjoint collections of $\phi_d + 1$ standard basis vectors. For each j there exists $0 \neq x^{(j)} \in V^{(j)}$ with $g(x^{(j)}) = 0$. Since g has nonzero coefficients, each $x^{(j)}$ must have at least two nonzero entries. The vector $x = x^{(1)} + \ldots + x^{(n_{d-2}/2)}$ satisfies $g(x) = 0$ and has at least n_{d-2} nonzero coefficients. For $1 \le e \le d - 1$, consider the corresponding term of the Taylor expansion of $g(x + y)$,

$$g^{(d-e,e)}(x,y) = \sum_{i=1}^{m''} f(u^{(j)}) \binom{d}{e} x_i^{d-e} y_i^e.$$

This is a system of diagonal forms in y with at least n_{d-2} nonzero coefficients, so by the inductive hypothesis we can find $y \in K^{m''}$ with $g^{(d-e,e)}(x,y) = 0$ for $1 \le e < d - 1$ and $g^{(1,d-1)}(x,y) \neq 0$. This last condition implies in particular that x, y are linearly independent.

Consider the subspace

$$L = \mathrm{span}\Big(\sum_{i=1}^{m''} x_i u^{(i)}, \sum_{i=1}^{m''} y_i u^{(i)}, u^{(m''+1)}, u^{(m''+2)}, w^{(1)}, w^{(2)}, w^{(3)}\Big).$$

By construction, the seven vectors spanning L are linearly independent. We claim that L satisfies properties (1-3). Properties (1) and (2) are straightforward to verify. For property (3), note that $\mathrm{Brk}(f_d \restriction_L, f_{d+1} \restriction_L) \ge 3$, so $\mathrm{codim}_L(L \cap S(f_d, f_{d+1})) \ge 2$ by lemma 4.3. It follows that f_d, f_{d+1} define a complete intersection in L. $\qquad\square$

Proof of proposition 2.6. We prove this by induction on d, where the base cases $d = 0, 1$ are immediate. For $d \ge 2$, applying lemma 6.3 and invoking the inductive hypothesis yields

$$n_d \le (\phi_d + 11) 2^{1+\varphi^{d-2}+\varphi^{d-3}} \prod_{k=2}^{d-1} (\phi_k + 11)^{\varphi^{d-1-k}+\varphi^{d-2-k}}$$

$$= 2^{1+\varphi^{d-1}} \prod_{k=2}^{d} (\phi_k + 11)^{\varphi^{d-k}}.$$

$$\square$$

7. Proof of proposition 2.9

The proof will follow a similar strategy to that of proposition 2.6, but will require some additional steps to deal with the fact that our forms are not diagonal. We begin by introducing the auxiliary variety

$$A_m = ((a_1, a_2, a_3, \underline{w}, \underline{v}) : (\underline{w}, \underline{v}) \in D_m \, , \, f(a_1 w_1 + a_2 w_2 + a_3 w_3) = 0).$$

Lemma 7.1. *If* $\mathrm{Brk}(\underline{f}) > 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 1$ *then* A_m *is irreducible.*

Proof. By lemma 4.7, the varieties $Z_m(\underline{g})$ and $Z(\underline{g}, f)$ are irreducible complete intersections. It follows that $f'(\underline{a}, \underline{w}) = f(a_1 w_1 + a_2 w_2 + a_3 w_3)$ doesn't vanish identically on $Z = \mathbf{A}^3 \times D_m$, so A_m is a complete intersection. Let $S = S(f', T_m(\underline{g}))$. To prove the claim, it's sufficient by lemma 4.6 to show that

(7) $\dim Z \cap S < \dim A_m - 1.$

To see this, decompose $S = \bigcup_{a \in \mathbf{A}^3} S_a$ where $S_a = S \cap (\{a\} \times V^m)$. For $a = 0$, we have

$$\mathrm{codim} Z \cap S_0 \geq \mathrm{codim} Z \cap (\{0\} \times V^m) = \mathrm{codim} Z + 3 = \mathrm{codim} A_m + 2.$$

For any $a \neq 0$, let

$$f_a(\underline{w}, \underline{v}) = f(a_1 w_1 + a_2 w_2 + a_3 w_3) = \sum_{e_1 + e_2 + e_3 = d} a_1^{e_1} a_2^{e_2} a_3^{e_3} f^{(e_1, e_2, e_3)}(\underline{w}).$$

If $(a, \underline{w}, \underline{v}) \in S$ then $(\underline{w}, \underline{v}) \in S(f_a, T_m(\underline{g})) \subset S(T_m(\underline{f}))$. By lemma 4.5 and 4.2 we get

$$\dim S_a \leq \dim(S(T_m(\underline{f}))) \leq m \dim V - \mathrm{Brk}(\underline{f}) + \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 2$$

$$< m \dim V - \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) - 1.$$

Taking the union over $0 \neq a \in A^3$, and using the bound for $a = 0$, we get

$$\dim Z \cap S < m \dim V - \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + 2 \leq \dim A_m - 1,$$

proving inequality (7). $\qquad \square$

The next lemma allows us to diagonalize while simultaneously keeping track of a Zariski open condition.

Lemma 7.2. *Suppose that* $\mathrm{Brk}(\underline{f}) > 2 \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) + d' - 1$ *and that h is a polynomial which doesn't vanish identically on $Z(\underline{f})$. Then there exists a dense open subset of* $(\underline{w}, \underline{v}) \in D_m$ *such that*
- *The vectors $w_1, w_2, w_3, v_1, \ldots, v_{m-3}$ are linearly independent,*
- *the polynomial h doesn't vanish identically on* $\mathrm{span}(w_1, w_2, w_3) \cap (f = 0)$ *and*
- *for each $i \in [m-3]$, $f(v_i) \neq 0$.*

Proof. By lemma 4.7 the variety D_m is irreducible so it's enough to show that each condition separately holds for a nonempty open subset.

The first condition holds because the codimension of the variety of linearly dependent tuples in V^m is

$$\dim V - (m-1) > \mathrm{Brk}(\underline{f}) - m > \sum_{i=1}^{d} s_i \left(\!\!\binom{m}{i}\!\!\right) > \mathrm{codim}_{V^m} D_m.$$

For the second condition, note that our hypothesis implies that

$$A' = A_m \cap (h(a_1 w_1 + a_2 w_2 + a_3 w_3) = 0)$$

is a proper subset of A_m. By lemma 7.1 this implies $\dim A' < \dim A_m$, and therefore the fiber satisfies $\dim A'(\underline{w}, \underline{v}) < \dim A_m(\underline{w}, \underline{v})$ for a dense open subset of D_m.

For the third condition, it is again enough to prove that it holds on a dense open subset for each fixed i. Indeed, by lemma 4.7, for each i the set $D_m \cap (f(v_i) = 0) \subset D_m$ has lower dimension. $\qquad \square$

Proof of proposition 2.9. Suppose $\mathrm{Brk}(\underline{f}) > \max(\delta_m(s_d, \ldots, s_1), 2\sum_{i=1}^d \left(\!\!\binom{m}{i}\!\!\right) + d' - 1)$ and let h be a polynomial which doesn't vanish identically on $Z(\underline{f})$. We wish to show that $Z(\underline{f}) \cap (h \neq 0)$ contains a K-point. By lemma 6.1, it's enough to find a K-subspace $L \subset K^n$ such that:

(1) The forms in $\underline{g}$ vanish on L,
(2) The restriction $f \restriction_L$ is good and
(3) the polynomial h doesn't vanish identically on $Z(f) \cap L$.

Let $U \subset D_m$ be the dense open subset of the previous lemma and let $(\underline{w}, \underline{v}) \in U$ be a K-point. This exists because $\mathrm{Brk}(\underline{f}) > \delta_m$. Let

$$f_0(x_1, \ldots, x_{m-4}) = f(x_1 v_1 + \ldots + x_{m-4} v_{m-4})$$

be the nondegenerate diagonal form given by the restriction of f and decompose $K^{m-4} = V^{(1)} \oplus \ldots \oplus V^{(\lceil \frac{\beta_{d-1}+1}{2} \rceil)}$, where the $V^{(j)}$ are spanned by disjoint collections of $\phi_d + 1$ standard basis vectors. By definition, for each j there exists $0 \neq x^{(j)} \in V^{(j)}$ with $f_0(x^{(j)}) = 0$. Each $x^{(j)}$ must have at least two nonzero entries, so the vector $x = x^{(1)} + \ldots + x^{(\lceil \frac{\beta_{d-1}+1)}{2} \rceil)}$ satisfies $f_0(x) = 0$ and has at least $\beta_{d-1} + 1$ nonzero entries.

For $1 \leq e \leq d - 1$, consider the terms of the Taylor expansion of $f_0(x + y)$,

$$f_0^{(d-e,e)}(x, y) = \sum_{i=1}^{m-4} f(v_j)\binom{d}{e}x_i^{d-e}y_i^e.$$

This is a system of diagonal forms in y with at least $\beta_{d-1}+1$ nonzero coefficients, so we can find $y \in K^{m-4}$ with $f_0^{(d-e,e)}(x, y) = 0$ for $1 \leq e < d-1$ and $f_0^{(1,d-1)}(x, y) \neq 0$. This last condition implies in particular that x, y are linearly independent. By construction, the subspace

$$L = \mathrm{span}\left(\sum_{i=1}^{m-4} x_i v_i, \sum_{i=1}^{m-4} y_i v_i, v_{m-3}, w_1, w_2, w_3\right)$$

satisfies conditions (1-3) above. This completes the proof. $\qquad\square$

REFERENCES

[1] Tigran Ananyan and Melvin Hochster. Small subalgebras of polynomial rings and Stillman's conjecture. *J. Amer. Math. Soc.*, 33(1):291–309, 2020.

[2] Benjamin Baily and Amichai Lampert. Strength, partition rank and algebraic closure. *arXiv e-prints*, page arXiv:2410.00248, September 2024.

[3] Arthur Bik, Jan Draisma, and Andrew Snowden. Two improvements in Brauer's theorem on forms. *arXiv e-prints*, page arXiv:2401.02067, January 2024.

[4] Arthur Bik, Jan Draisma, and Andrew Snowden. The geometry of polynomial representations in positive characteristic. *Math. Z.*, 310(1):Paper No. 13, 39, 2025.

[5] Richard Brauer. A note on systems of homogeneous algebraic equations. *Bull. Amer. Math. Soc.*, 51:749–755, 1945.

[6] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

[7] Siu-lun Alan Lee. Birch's theorem in function fields. *arXiv e-prints*, page arXiv:1109.4953, September 2011.

[8] D. B. Leep and W. M. Schmidt. Systems of homogeneous equations. *Invent. Math.*, 71(3):539–549, 1983.

[9] Wolfgang M. Schmidt. The density of integer points on homogeneous varieties. *Acta Math.*, 154(3-4):243–296, 1985.

[10] C. M. Skinner. Forms over number fields and weak approximation. *Compositio Math.*, 106(1):11–29, 1997.

[11] Christopher Skinner. Solvability of systems of diagonal equations over p-adic local fields. *Proc. Lond. Math. Soc. (3)*, 122(2):207–228, 2021.

[12] Trevor D. Wooley. On the local solubility of Diophantine systems. *Compositio Math.*, 111(2):149–165, 1998.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI
Email address: amichai@umich.edu