Tschirnhaus transformations after Hilbert

Jesse Wolfson*

May 31, 2019

Abstract

Let RD(n) denote the minimum d for which there exists a formula for the roots of the general degree n polynomial using only algebraic functions of d or fewer variables. In 1927, Hilbert sketched how the 27 lines on a cubic surface could be used to construct a 4-variable formula for the general degree 9 polynomial (implying RD(9) ≤ 4). In this paper, we turn Hilbert's sketch into a general method and use this to obtain best-to-date upper bounds on RD(n) for all n, improving earlier results of Hamilton, Sylvester, Segre and Brauer.

1 Introduction

Consider the problem of finding the roots of a polynomial

$$z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

in terms of the coefficients $a_1, \ldots, a_n$. A priori, the assignment

$$(a_1, \ldots, a_n) \mapsto \{z \mid z^n + a_1 z^{n-1} + \cdots + a_n = 0\}$$

is an algebraic function of n (complex) variables, and it is natural to ask whether there exists a formula using only algebraic functions of d or fewer variables. Call the minimum such d the *resolvent degree* and denote this by RD(n) (see Section 4 for a precise definition, and [FW18] for a detailed treatment). At present, no nontrivial lower bounds for RD(n) are known. The best general upper bounds in the literature are due to Brauer [Br75], who uses methods dating to Tschirnhaus [Ts1683] to prove that RD(n) $\leq n - r$ for $n \geq (r-1)! + 1$. As Brauer remarks, his bounds are not optimal.

In this paper we take a different approach to bounding RD(n), inspired by a geometric argument of Hilbert. In [Hi27] (for a contemporary translation, see [Hi27t]), Hilbert sketches how the 27 lines on a cubic surface can be used to produce a 4-variable formula for the general degree 9 polynomial, i.e. RD(9) ≤ 4. We turn Hilbert's sketch into a general method, whereby lines on cubic surfaces are replaced by r-planes on degree d hypersurfaces in $\mathbb{P}^m$ for appropriate choices of r,d and m. As a consequence, we are able to significantly improve Brauer's bounds and obtain the best current upper bounds on RD(n) for all n. While the exact statement takes some time to set up[1], the main result can be quickly stated as follows:

*Supported in part by NSF Grant DMS-1811846.

[1] See Theorem 5.6 and the discussion preceding it for a precise statement; see Table 1 for explicit bounds.

Theorem 1.1. *There exist monotone increasing functions* $\mathrm{F}, \varphi \colon \mathbb{N} \longrightarrow \mathbb{N}$ *so that:*

1. $\mathrm{RD}(n) \leq n - r$ *for* $n \geq \mathrm{F}(r)$.
2. *For all* $d \geq 0$, *if* $r \geq \varphi(d)$ *then* $\frac{(r-1)!+1}{\mathrm{F}(r)} \geq d!$.

(The first statement appears as Theorem 5.6 below, the second as Theorem 5.8.) In particular, in the limit, the ratio of Brauer's bounds to ours grows arbitrarily.

Besides the general interest in obtaining simpler formulas for polynomials, we hope this paper spurs work on two questions. For the first, we quote Dixmier [Di93, p. 90][2]:

> "Every reduction of RD(n) would be serious progress. In particular, it is time to know if RD(6) = 1 or RD(6) = 2." (Dixmier, 1993)

While the present methods cannot touch Hilbert's Sextic Conjecture (RD(6) = 2), they do contribute to Dixmier's call to lower the possible values of RD(n). They also contribute to a problem first posed (as far as we are aware) by Segre [Se51, III.5]:

Problem 1.2. Understand the large n behavior of RD(n).

As a clearer understanding of Segre's problem comes into view, we look forward to seeing the present bounds lowered in turn.

Historical Background.

> "The theory has been a plant of slow growth."
>
> (Sylvester and Hammond, 1887[3])

Tschirnhaus [Ts1683] introduced his transformation[4] to show that $\mathrm{RD}(n) \leq n-3$, improving upon the linear change of variables used by the Babylonians to set the first coefficient of the general polynomial to 0. A century later, Bring [Br1786] improved this for $n = 5$ to show that RD(5) = 1. Hamilton [Ha1836] was the first to show that

$$\lim_{n \longrightarrow \infty} n - \mathrm{RD}(n) = \infty.$$

More precisely, he showed the existence a monotone increasing function $H \colon \mathbb{N} \longrightarrow \mathbb{N}$, such that $n - \mathrm{RD}(n) \geq r$ for $n \geq H(r)$.[5] Hamilton computed the initial values of H (for $r \leq 7$). Five decades later, Sylvester [Sy1887] extended Hamilton's computations to give:

r	4	5	6	7	8	9
$H(r)$	5	11	47	923	409,619	83,763,206,255

Sylvester then sharpened Hamilton's bounds slightly (see [Sy1887, p. 485])[6], and Sylvester and Hammond [SH1887], [SH1888] gave a generating function for H.

---

[2] n.b. Dixmier writes "$s(n)$" for our RD(n).

[3] [SH1887, p. 286]

[4] We review Tschirnhaus' method in detail in Section 3.

[5] The numbers $H(r)$ are listed as the "Hamilton numbers" in the *Online Encyclopedia of Integer Sequences*.

[6] Writing H' for Sylvester's sharpening, the initial values are $H'(4) = 5$, $H'(5) = 10$, $H'(6) = 44$, $H'(7) = 905$.

Hilbert [Hi27] had a beautiful geometric idea to lower RD(9) to at most 4 (see also [Wi27]), but progress on the general problem stalled after Hilbert, and by the mid-20th century, Hamilton's work appears to have been forgotten. Segre [Se45], building on Hilbert, provided the first rigorous proof that $\mathrm{RD}(n) \leq n-5$ for $n \geq 9$, and proved that for $n \geq 157$, $\mathrm{RD}(n) \leq n-6$. Segre conjectured that in the limit

$$\lim_{n \longrightarrow \infty} n - \mathrm{RD}(n) = \infty.$$

(i.e. precisely what Hamilton had showed over a century earlier). Brauer [Br45] and Segre each reproved this statement, but without giving effective bounds *à la* Hamilton. In 1975, Brauer [Br75] proved that for $n \geq (r-1)! + 1$, $\mathrm{RD}(n) \leq n - r$. This improves Hamilton's bounds for $r >> 0$, and provided the best upper bound, of which we are aware, prior to this paper.

Remarks on the Proof. Dixmier [Di93, S8] observed that the argument in [Hi27] is incomplete. In particular, Hilbert takes for granted that a family of cubic surfaces arising in his argument for $n = 9$ is sufficiently generic (more precisely, that the generic member of the family admits a "pentahedral form"). Letting $\mathcal{H}_{3,3}$ denote the parameter space of cubic surfaces and $\mathcal{M}_{3,3}$ the (coarse) moduli space of smooth cubic surfaces, Hilbert equivalently assumes that a given map

$$X \longrightarrow \mathcal{H}_{3,3}$$

intersects the locus where the rational map

$$\mathcal{H}_{3,3} \dashrightarrow \mathcal{M}_{3,3}$$

is well-defined. The principal geometric contribution of this paper is to show that for all n, the family of "Tschirnhaus hypersurfaces" needed for Hilbert's argument and its generalization to arbitrary degrees are generically smooth; see Theorem 2.11.

Beyond this, we need two fundamental post-Hilbert advances to convert Hilbert's sketch into a general method. The first is Merkurjev and Suslin's theorem on Severi-Brauer varieties [MS83, Theorem 16.1], which allows us to trivialize the Severi-Brauer varieties which arise in Hilbert's argument by adjoining radicals.[7] The second is a theorem of Hochster–Laksov [HL87] which allowed Waldron [Wa08, Theorem 1.6] (see also [St17, Theorem 1.2]) to show that *every* degree d hypersurface in $\mathbb{P}^N$ contains an r-plane when an appropriate dimension count is non-negative.

Outline of the Paper. In Section 2 we introduce the Tschirnhaus complete intersections and study their geometry. In Section 3, we recall the geometric perspective on Tschirnhaus transformations, and connect this to the Tschirnhaus complete intersections. In Section 4, we develop the necessary results about the resolvent degree of a dominant map needed to implement Hilbert's idea for general degrees n. This extends the treatment of resolvent degree of generically finite dominant maps in [FW18]. In Section 5, we prove the upper bounds for $\mathrm{RD}(n)$ and compare them to Brauer's. In Appendix A, we give explicit values for the function $\mathrm{F}(r)$ discussed above.

[7] Neither Hilbert nor Dixmier comment on this gap in Hilbert's argument.

Acknowledgements. First, I thank Benson Farb, who was closely involved with the ideas that led to this paper, but who declined to be listed as a coauthor. Next, my sincere thanks to Sebastian Hensel who translated Hilbert's 1927 paper into English at Benson's and my request. This paper takes place in the context of ongoing joint work with Benson Farb and Mark Kisin, and their influence permeates the perspective here. Beyond this, I thank Jordan Ellenberg, Vlad Matei, Curt McMullen, Madhav Nori, Zinovy Reichstein, Daniil Rudenko and David Smyth for helpful conversations.

2 Tschirnhaus Complete Intersections

Fix $n \geq 0$. In this section, we work over $\mathbb{Z}$ unless otherwise specified, so that, e.g. $\mathbb{A}^n := Spec(\mathbb{Z}[a_1, \ldots, a_n])$. For ease of reading, we adopt the following notation.

Notation 2.1. Denote

$$\mathbf{a} := (a_1, \ldots, a_n) \in \mathbb{A}^n. \qquad |\kappa| := \sum_i k_i$$
$$\mathbf{b} := [b_1 : \cdots : b_{n-1}] \in \mathbb{P}^{n-2} \qquad ||\kappa|| := \sum_i i \cdot k_i$$
$$\kappa := (k_1, \ldots, k_{n-1}) \in \mathbb{N}^n \qquad \mathbf{b}^\kappa := \prod_{i=1}^{n-1} b_i^{k_i}$$

For $|\kappa| = i$, recall the multinomial coefficients

$$\binom{i}{\kappa} := \binom{i}{k_0, \ldots, k_{n-1}} := \frac{i!}{k_1! \cdots k_m!}.$$

We inductively define polynomials in the a_i by

$$p_0 := n, \tag{2.1}$$

while, for $k \leq n$

$$p_k := ka_k + \sum_{i=1}^{k-1} a_{k-i} p_i, \tag{2.2}$$

and for $k > n$

$$p_k := -\sum_{i=k-n}^{k-1} a_{k-i} p_i. \tag{2.3}$$

Remark 2.2. To interpret the polynomials p_i, let σ_i denote the i^{th} elementary symmetric polynomial in formal variables $x_1, \ldots, x_n$. If we write $a_i = (-1)^i \sigma_i$, then Newton's Identities give

$$p_i = \sum_{j=1}^{n} x_j^i.$$

Definition 2.3. For $i, n \geq 1$, let the $T_i^n \subset \mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}_{\mathbf{b}}^{n-1}$ be the scheme defined by the vanishing of the polynomial

$$\sum_{\kappa \ s.t. \ |\kappa|=i} \binom{i}{\kappa} p_{||\kappa||} \mathbf{b}^{\kappa}. \tag{2.4}$$

Note that this polynomial is homogeneous of degree i in the $\mathbf{b}$-coordinates. Projecting onto the first factor gives a family of degree i hypersurfaces in $\mathbb{P}^{n-1}$

$$T_i^n \longrightarrow \mathbb{A}_{\mathbf{a}}^n$$

We refer to this family as the n^{th} *Tschirnhaus hypersurface of degree* i.

Definition 2.4. Fix $n \geq 1$. For $1 \leq i_1 < \ldots < i_k$, define the n^{th} *Tschirnhaus complete intersection* $T_{i_1 \cdots i_k}^n$ *(of multi-degree* $i_1 \cdots i_k$*)* to be the scheme defined by the vanishing of the polynomials (2.4) for $i = i_1, \ldots, i_k$. Equivalently, define

$$T_{i_1 \cdots i_k}^n := T_{i_1}^n \times_{\mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}_{\mathbf{b}}^{n-1}} \cdots \times_{\mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}_{\mathbf{b}}^{n-1}} T_{i_k}^n \longrightarrow \mathbb{A}_{\mathbf{a}}^n.$$

Define the n^{th} *reduced Tschirnhaus complete intersection* $T_{i_1 \cdots i_k}^{n'}$ *(of multi-degree* $i_1 \cdots i_k$*)* by

$$T_{i_1 \cdots i_k}^{n'} := T_{i_1 \cdots i_k}^n \cap \{b_0 = 0\} \subset \mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}_{\mathbf{b}}^{n-1}.$$

Example 2.5. The hyperplane $T_1(\mathbf{a}) \subset \mathbb{P}_{\mathbf{b}}^{n-1}$ is given by the equation

$$nb_0 + \sum_{i=1}^{n-1} p_i b_i = 0$$

Over $\mathbb{Z}[1/n]$, we have an isomorphism

$$\mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}^{n-2} \xrightarrow{\cong} T_1$$
$$(\mathbf{a}, [b_1 : \cdots : b_{n-1}]) \mapsto (\mathbf{a}, [-\frac{1}{n} \sum_{i=1}^{n-1} p_i b_i : b_1 : \cdots : b_{n-1}]).$$

Likewise, the hyperplane $T_1'(\mathbf{a}) \subset \mathbb{P}_{\mathbf{b}}^{n-2}$ is given by the equation

$$\sum_{i=1}^{n-1} p_i b_i = 0$$

Over the locus $\{p_i \neq 0\} \subset \mathbb{A}_{\mathbf{a}}^n$ for $1 \leq i < n$, we have an isomorphism

$$\mathbb{A}_{\mathbf{a}}^n \times \mathbb{P}^{n-3} \xrightarrow{\cong} T_1$$
$$(\mathbf{a}, [b_1 : \cdots : b_{\hat{i}} : \cdots : b_{n-2}]) \mapsto (\mathbf{a}, [b_1 : \cdots : b_{i-1} : \frac{-1}{p_i} \sum_{j \neq i} p_j b_j : b_{i+1} : \cdots b_{n-2}]).$$

As a warm-up to Theorem 2.11 below, we prove the following.

Lemma 2.6. *The families of quadrics $T_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ and $T'_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ are generically smooth.*

Remark 2.7. The statement of the lemma for T_{12} (and most likely for T'_{12}) is classical, and follows from the fact that the discriminant of the quadratic form defining $T_{12}(\mathbf{a})$ is equal to $\frac{1}{n}$ times the discriminant of the polynomial $x^n + a_1x^{n-1} + \cdots + a_n$ (see, e.g. [Sy1887, p. 468-469]). We give a different proof in order to warm-up for Theorem 2.11.

Proof of Lemma 2.6. The quadric $T_{12}(\mathbf{a}) \subset \mathbb{P}^{n-2}_{\mathbf{b}}$ is given, in coordinates $[b_1 : \cdots : b_{n-1}]$ by the equation

$$-\frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n-1} p_i b_i\right)^2 + \sum_{1\le i<j\le n-1} p_{i+j} b_i b_j + \sum_{i=1}^{n-1} p_{2i} b_i^2 = 0.$$

We now specialize to the radical pencil $x^n + a = 0$, i.e. $\mathbf{a} = (0, \ldots, a)$. Then $T_{12}(a) := T_{12}(0, \ldots, a)$ is given by the equation

$$\begin{cases} -2na\left(\sum_{i=1}^{\frac{n-1}{2}} b_i b_{n-i}\right) & n \text{ odd} \\ -na\left(b_{\frac{n}{2}}^2 + 2\sum_{i=1}^{\frac{n}{2}-1} b_i b_{n-i}\right) & n \text{ even} \end{cases} \tag{2.5}$$

The partial derivatives of $T_{12}(a)$ are given by

$$\partial_{b_j} T_{12}(a) = -2nab_{n-j}.$$

We see that these vanish simultaneously if and only if $b_j = 0$ for all j, i.e. $T_{12}(a)$ is smooth over $\mathbb{Z}[1/n]$ so long as $a \neq 0$ (and thus $T_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ is generically smooth).

We now prove $T'_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ is generically smooth. Using (2.2), the hyperplane $T'_1(a)$ is given by

$$(n-1)ab_{n-1} = 0.$$

Over $\mathbb{Z}[1/(n-1)]$, and $a \neq 0$, we can therefore use the coordinates

$$[b_1 : \cdots : b_{n-2}]$$

on $T'_1(a)$. We now slightly modify Notation 2.1 by setting

$$\mathbf{b}' := [b_1 : \cdots : b_{n-2}] \in \mathbb{P}^{n-3}$$
$$\kappa' := (k_1, \ldots, k_{n-2}) \in \mathbb{N}^{n-2}$$

and carrying over the rest of the notation *mutatis mutandis.*

In these coordinates and this notation, we have

$$T'_{12}(a) = \begin{cases} -2(n-1)a\left(\sum_{i=1}^{\frac{n}{2}-1} b_i b_{n-1-i}\right) & n \text{ even} \\ -(n-1)a(b^2_{\frac{n-1}{2}} + \left(\sum_{i=1}^{\frac{n}{2}-1} b_i b_{n-1-i}\right) & n \text{ odd} \end{cases}$$

The partial derivatives of $T'_{12}(a)$ are given by

$$\partial_{b_j} T'_{12}(a) = -2(n-1)ab_{n-1-j}.$$

We see that these vanish simultaneously if and only if $b_j = 0$ for all j, i.e. $T'_{12}(a)$ is smooth over $\mathbb{Z}[1/(n-1)]$ so long as $a \neq 0$ (and thus $T'_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ is generically smooth). □

Tschirnhaus hypersurfaces as spaces of maps. In Section 3, we explain the origin of the Tschirnhaus complete intersections in the classical study of formulas for the general degree n polynomial (beginning with [Ts1683]). For the moment, we just observe that several varieties of classical interest are closely related to T_i^n for small i, n.

Let $\mathbf{x} := (x_1, \ldots, x_n)$ be coordinates on affine n-space, denoted $\mathbb{A}^n_{\mathbf{x}}$. Let $\sigma_i(\mathbf{x})$ denote the i^{th} elementary symmetric function on the x_i, and consider the map

$$\begin{aligned} q\colon \mathbb{A}^n_{\mathbf{x}} &\longrightarrow \mathbb{A}^n_{\mathbf{a}} \\ \mathbf{x} &\mapsto (-\sigma_1(\mathbf{x}), \ldots, (-1)^n\sigma_n(\mathbf{x})). \end{aligned}$$

By Newton's Theorem, this map realizes $\mathbb{A}^n_{\mathbf{a}}$ as the quotient of $\mathbb{A}^n_{\mathbf{x}}$ by the permutation action of the symmetric group S_n on $\mathbb{A}^n_{\mathbf{x}}$. As remarked above, Newton's Identities imply that

$$p_i(q(\mathbf{x})) = \sum_{j=1}^n x_j^i.$$

Let $\tilde{\mathbf{b}} := (b_0, \ldots, b_{n-1})$ viewed as affine coordinates on $\mathbb{A}^n_{\tilde{\mathbf{b}}}$. The relative affine cone on the pullback $T_i^n|_{\mathbb{A}^n_{\mathbf{x}}} \longrightarrow \mathbb{A}^n_{\mathbf{x}}$ is given by

$$\widetilde{T_i^n} := \left\{ (\mathbf{x}, \tilde{\mathbf{b}}) \in \mathbb{A}^n_{\mathbf{x}} \times \mathbb{A}^n_{\tilde{\mathbf{b}}} \mid \sum_{\kappa \ s.t.\ |\kappa|=i} \binom{i}{\kappa} \left(\sum_{j=1}^n x_j^{||\kappa||} \right) \tilde{\mathbf{b}}^\kappa = 0 \right\}.$$

Consider the map

$$\begin{aligned} \mathrm{ev}\colon \mathbb{A}^n_{\mathbf{x}} \times \mathbb{A}^n_{\tilde{\mathbf{b}}} &\longrightarrow \mathbb{A}^n_{\mathbf{x}} \\ (\mathbf{x}, \tilde{\mathbf{b}}) &\mapsto (\sum_{j=0}^{n-1} b_j x_1^j, \ldots, \sum_{j=0}^{n-1} b_j x_n^j). \end{aligned}$$

Lemma 2.8. *In the notation above,*

$$\widetilde{T_i^n} = \mathrm{ev}^{-1}(\{\mathbf{x} \in \mathbb{A}^n_{\mathbf{x}} \mid \sum_j x_j^i = 0.\}).$$

Proof. We prove this by explicit computation. For $i \geq 0$, write

$$p_i(\mathbf{x}) := \sum_{\ell=1}^n x_\ell^i.$$

In particular, $p_0(x_1, \ldots, x_n) = n$. Let $\mathrm{ev}(\mathbf{x})_\ell := \sum_{j=0}^{n-1} b_j x_\ell^j$. By the Multinomial Theorem,

$$\begin{aligned} p_i(\mathrm{ev}(\mathbf{x})) = \sum_\ell \mathrm{ev}(\mathbf{x})_\ell^i &= \sum_\ell \left(\sum_{j=0}^{n-1} b_j x_\ell^j \right)^i \\ &= \sum_\ell \left(\sum_{\kappa \ s.t.\ |\kappa|=i} \binom{i}{\kappa} \mathbf{b}^\kappa x_\ell^{||\kappa||} \right) \\ &= \sum_{\kappa \ s.t. |\kappa|=i} \binom{i}{\kappa} p_{||\kappa||} \mathbf{b}^\kappa \end{aligned} \tag{2.6}$$

where, in the final line, we use Newton's Identities to identify the power sums with the polynomials $p_{||\kappa||}$ in the a_i defined in Equations 2.1-2.3.

Setting the form (2.6) to 0, we obtain the hypersurface $\widetilde{T_i^n}$ as claimed. $\square$

Example 2.9. Let $S \subset \mathbb{P}^4$ be the Clebsch diagonal surface, i.e. the complete intersection

$$S := \{[x_1 : \cdots : x_5] \in \mathbb{P}^4 \mid \sum_{i=1}^{5} x_i = \sum_{i=1}^{5} x_i^3 = 0\}.$$

Let $\widetilde{S} \subset \mathbb{A}^5_{\mathbf{x}}$ be the affine cone over S. Then

$$\widetilde{T_{13}^5} = \mathrm{ev}^{-1}(\widetilde{S}).$$

As observed by Klein [Kl1884, Part II, Ch. 2], $\widetilde{T_{13}^5}$ can be understood as a space of S_5-equivariant maps of $\mathbb{A}^5_{\mathbf{x}} \longrightarrow \widetilde{S}$.

Example 2.10. Let $F \subset \mathbb{P}^6$ be the symmetric Fano sextic 3-fold, i.e. the complete intersection

$$F := \{[x_1 : \cdots : x_7] \in \mathbb{P}^6 \mid \sum_{i=1}^{7} x_i = \sum_{i=1}^{7} x_i^2 = \sum_{i=1}^{7} x_i^3 = 0\}.$$

Let $\widetilde{F} \subset \mathbb{A}^7_{\mathbf{x}}$ be the affine cone over F. Then

$$\widetilde{T_{123}^7} = \mathrm{ev}^{-1}(\widetilde{F}).$$

Though not remarked upon in [Be12], the symmetric Fano sextic arises as the "root space" of the normal form for the general degree 7 polynomial considered by Hilbert in his 13th problem [Hi1900]. The scheme $\widetilde{T_{123}^7}$ can be understood as a space of S_7-equivariant maps of $\mathbb{A}^7_{\mathbf{x}} \longrightarrow \widetilde{F}$, equivalently of ways of converting the general degree 7 polynomial into Hilbert's normal form.

Geometry of Tschirnhaus Complete Intersections

We can now state our main geometric theorem.

Theorem 2.11. *Let p be a prime. Let $i = p^r + 1 < n$ for some prime power p^r with $r > 0$.*

1. *If $p \nmid n$, the family of Tschirnhaus complete intersections*

$$T_{12i} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is generically smooth (i.e. there is a Zariski open $U \subset \mathbb{A}^n_{\mathbf{a}}$ such that for all $\mathbf{a} \in U$, $T_{12i}(\mathbf{a})$ is a smooth complete intersection).

2. *If $p \mid n$, the family of reduced Tschirnhaus complete intersections*

$$T'_{12i} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is generically smooth.

Deferring the proof for a moment, let K be a field of characteristic 0, now and throughout this paper.

Convention 2.12. *By a* K-variety, *we will always mean a reduced, separated* K*-scheme of finite type. By a* map *we will always mean a regular map of* K*-varieties.*

We now record a Bertini theorem for isotropic subspaces of smooth quadrics, which we prove below.

Proposition 2.13 (Bertini for isotropics)**.** *Let* K *be algebraically closed. Let* X *be a* K*-variety. Let* $Q \subset \mathbb{P}^n_X$ *be a smooth family of quadrics over* X*. For* $k \leq \lfloor \frac{n-1}{2} \rfloor$*, let* $Gr(k, Q) \longrightarrow X$ *denote the relative Grassmannian of* k*-dimensional isotropic subspaces in* Q*, and let* $\mathcal{L} \longrightarrow Gr(k, Q)$ *denote the tautological bundle. Let* $Y \subset \mathbb{P}^n_X$ *be a smooth family of varieties over* X *such that the family* $Q \times_{\mathbb{P}^n_X} Y \longrightarrow X$ *is smooth over some dense open* $V \subset X$*. Then there exists a dense open* $U \subset Gr(k, Q)|_V$ *such that the family* $\mathcal{L}|_U \times_{\mathbb{P}^n_X} Y|_V \longrightarrow X$ *is smooth.*

Combining Theorem 2.11, Lemma 2.6 and Proposition 2.13, we obtain the following.

Corollary 2.14. *Let* $Gr(T_{12}) \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ *denote the relative Grassmannian of maximal isotropics in the family of quadrics* $T_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$*, and let* $\mathcal{L} \longrightarrow Gr(T_{12})$ *denote the tautological bundle (with similar notation for the analogous objects for* T'_{12}*). Let* p *be a prime and let* $i = p^r + 1$ *for some* $r > 0$*.*

1. *If* $p \nmid n$*, there exists a dense open* $V \subset Gr(T_{12})$ *such that*

$$\mathcal{L}|_V \times_{\mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-1}_{\mathbf{b}}} T_{12i} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is smooth (i.e. for the generic polynomial, the intersection of $T_{12i}(\mathbf{a})$ *with a maximal isotropic in* $T_{12}(\mathbf{a})$ *is smooth).*

2. *If* $p \mid n$*, there exists a dense open* $V \subset Gr(T'_{12})$ *such that*

$$\mathcal{L}|_V \times_{\mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-2}_{\mathbf{b}'}} T'_{12i} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is smooth.

Proof. Note that to prove the existence of an open dense V, it suffices to restrict all of the schemes over $\mathbb{Z}$ above to a geometric generic point $Spec(K) \longrightarrow Spec(\mathbb{Z})$. The result now follows immediately from Theorem 2.11, Lemma 2.6 and Proposition 2.13. ☐

Remark 2.15. Corollary 2.14 (for the case $p = 2, i = 3, n = 9$) fills the gap in Hilbert's argument remarked upon by Dixmier [Di93, S8].

Proof of Proposition 2.13. We adapt Kleiman's proof of the classical Bertini theorem [Kl74]. Consider the canonical map

$$\mathrm{pr}_2 \colon \mathcal{L} \longrightarrow Q$$

(coming from the construction of $\mathcal{L}$ as an incidence variety $\mathcal{L} \subset Gr(k, Q) \times_X Q$). Observe that this map is smooth: indeed, the relative group scheme $\mathrm{O}(Q)$ acts transitively over X

on both $\mathcal{L}$ and Q (i.e. it acts transitively on fibers over X) and the map $\mathcal{L} \longrightarrow Q$ is an $\mathrm{O}(Q)$-equivariant fiber bundle, with fiber at $v \in Q$ given by $\mathrm{Stab}_{\mathrm{O}(Q)}(v)/\,\mathrm{Stab}_{\mathrm{O}(Q)}(L, v)$, (n.b. the stabilizer of an isotropic point v is a maximal parabolic, and the stabilizer of the flag $v \in L$ is a sub-parabolic).

Let $V \subset X$ be a dense open such that $Q \times_{\mathbb{P}^n_X} Y \longrightarrow X$ is smooth over V. Shrinking V as necessary, we can assume without loss of generality that V is a smooth variety over K (note that we are using characteristic 0 here), and thus $(Q \times_{\mathbb{P}^n_X} Y)|_V$ is also a smooth K-variety. Now consider the fiber product

$$\begin{array}{ccc} (\mathcal{L} \times_{\mathbb{P}^n_X} Y)|_V & \xrightarrow{f} & (Q \times_{\mathbb{P}^n_X} Y)|_V \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle \iota} \\ \mathcal{L}|_V & \xrightarrow{\mathrm{pr}_2} & Q|_V \\ {\scriptstyle \pi}\downarrow & & \\ Gr(k, Q)|_V & & \end{array}$$

The map f is smooth because pr_2 is smooth. Because $(Q \times_{\mathbb{P}^n_X} Y)|_V$ is a smooth K-variety, the K-variety $(\mathcal{L} \times_{\mathbb{P}^n_X} Y)|_V$ is smooth. We therefore have a dominant map of smooth K-varieties

$$q = \pi \circ g \colon (\mathcal{L} \times_{\mathbb{P}^n_X} Y)|_V \longrightarrow Gr(k, Q)|_V.$$

By generic smoothness (e.g. [Ha77, Corollary III.10.7]), there exists a nonempty open subset $U \subset Gr(k, Q)|_V$ such that $q \colon (\mathcal{L}|_U \times_{\mathbb{P}^n_X} Y|_V) \longrightarrow U$ is smooth, and thus the composite $(\mathcal{L}|_U \times_{\mathbb{P}^n_X} Y|_V) \longrightarrow U \longrightarrow V$ is smooth as well. □

We now prove Theorem 2.11.

Proof of Theorem 2.11. We prove the two cases separately, via parallel arguments.

Case 1: $p \nmid n$. The complete intersection $T_{12i}(\mathbf{a})$ is smooth if and only if the $3 \times n$ matrix

$$\begin{pmatrix} \partial_{b_1} T_1(\mathbf{a}) & \cdots & \partial_{b_{n-1}} T_1(\mathbf{a}) \\ \partial_{b_1} T_2(\mathbf{a}) & \cdots & \partial_{b_{n-1}} T_2(\mathbf{a}) \\ \partial_{b_1} T_i(\mathbf{a}) & \cdots & \partial_{b_{n-1}} T_i(\mathbf{a}) \end{pmatrix}$$

has full rank for all $\mathbf{b} \in T_{12i}(\mathbf{a})$. Choosing coordinates on T_1, we can equivalently check whether the $2 \times (n-1)$ matrix given by the partials of T_{12} and T_{1i} has rank 2 for all $\mathbf{b} \in T_{12i}(\mathbf{a})$. To show generic smoothness, it suffices to find a single $\mathbf{a}$ for which this holds. Further, because the matrix above is defined over $\mathbb{Z}$, to show it is nonsingular in characteristic 0, it suffices to find a prime p for which its reduction mod p is nonsingular.

We specialize to the locus of radical polynomials, i.e. those of the form

$$p(x) = x^n + a$$

i.e. $\mathbf{a} = (0, \ldots, a)$. It suffices to show there exists a such that $T_{12i}(a) := T_{12i}(0, \ldots, a)$ is smooth. Note that, restricting to $x^n + a$, the hyperplane $T_1(a)$ is given by

$$nb_0 = 0.$$

We can therefore use the coordinates

$$[b_1 : \cdots : b_{n-1}]$$

on $T_1(a)$ as above. As in (2.5), the form $T_{12}(a)$ is given in these coordinates by

$$T_{12}(a) = \begin{cases} -2na\left(\sum_{i=1}^{\frac{n-1}{2}} b_i b_{n-i}\right) & n \text{ odd} \\ -na\left(b_{\frac{n}{2}}^2 + 2\sum_{i=1}^{\frac{n}{2}-1} b_i b_{n-i}\right) & n \text{ even} \end{cases}$$

and the partial derivatives are given by

$$\partial_{b_j} T_{12}(a) = -2nab_{n-j}.$$

Similarly, the form $T_{1i}(a)$ is given by

$$T_{1i}(a) = n \cdot \left(\sum_{\ell=1}^{i-1} (-1)^\ell a^\ell \left(\sum_{\kappa \ s.t. |\kappa|=i, ||\kappa||=\ell n} \binom{i}{\kappa} \mathbf{b}^\kappa\right)\right)$$

The partial derivatives of $T_{1i}(a)$ are given by

$$\partial_{b_j} T_{1i}(a) = in \cdot \left(\sum_{\ell=1}^{i-1} (-1)^\ell a^\ell \left(\sum_{\kappa \ s.t. \ |\kappa|=i-1, ||\kappa||+j=\ell n} \binom{i-1}{\kappa} \mathbf{b}^\kappa\right)\right).$$

Define

$$\begin{aligned} T_{j,12}(a) &:= ab_{n-j} \\ T_{j,1i}(a) &:= \sum_{\ell=1}^{i-1} (-1)^\ell a^\ell \left(\sum_{\kappa \ s.t. \ |\kappa|=i-1, ||\kappa||+j=\ell n} \binom{i-1}{\kappa} \mathbf{b}^\kappa\right). \end{aligned}$$

Then, in characteristic 0, the matrix

$$\begin{pmatrix} \partial_{b_1} T_{12}(a) & \cdots & \partial_{b_{n-1}} T_{12}(a) \\ \partial_{b_1} T_{1i}(a) & \cdots & \partial_{b_{n-1}} T_{1i}(a) \end{pmatrix}$$

is singular if and only if the matrix

$$\begin{pmatrix} T_{1,12}(a) & \cdots & T_{n-1,12}(a) \\ T_{1,1i}(a) & \cdots & T_{n-1,1i}(a) \end{pmatrix}$$

is singular. Because this matrix is defined over $\mathbb{Z}[a]$, to show that it is generically nonsingular in characteristic 0, we can reduce mod p and find some $a \in \overline{\mathbb{F}_p}$ for which it is nonsingular.

Let $\overline{T_{j,12}}(a)$ and $\overline{T_{j,1i}}(a)$ denote the reduction of the above forms mod p.

Recall that multinomial coefficients can be written as a product of binomial coefficients

$$\binom{\ell}{k_1, \ldots, k_m} = \binom{k_1 + k_2}{k_2} \cdots \binom{k_1 + \ldots + k_{m-1}}{k_{m-1}} \binom{\ell}{k_m}.$$

Further, recall that Legendre's formula implies that a prime p divides all binomial coefficients $\binom{\ell}{k}$ for $1 \le k \le \ell - 1$ if and only if $\ell = p^r$. We conclude that p divides the multinomial coefficients $\{\binom{\ell}{k_1,\ldots,k_m} \mid k_j < \ell \text{ for all } j\}$ if and only if $\ell = p^r$.

Therefore, reducing the forms $T_{j,1i}(a)$ mod p, and using $i-1 = p^r$, Legendre's formula implies that

$$T_{j,1i}(a) = \sum_{\ell=1}^{i-1} (-1)^\ell a^\ell \left(\sum_{1 \le \nu \le n-1, p^r\nu + j = \ell n} b_\nu^{p^r} \right) \tag{2.7}$$

Now, because $p \nmid n$, $p^r \in (\mathbb{Z}/n\mathbb{Z})^\times$. Therefore, multiplication by p^{-r} determines a permutation of $\{1, \ldots, n-1\} = \mathbb{Z}/n\mathbb{Z} - \{0\}$, which we denote by

$$\nu(j) := p^{-r} \cdot j \in \mathbb{Z}/n\mathbb{Z} - \{0\}.$$

In this notation, we have

$$\begin{aligned} \overline{T_{j,12}}(a) &= ab_{-j} \\ \overline{T_{j,1i}}(a) &= (-a)^{\frac{p^r\nu(-j)+j}{n}} b_{\nu(-j)}^{p^r} \end{aligned}$$

where $\pm j$ and $\nu(\pm j)$ denote the corresponding elements of $\{1, \ldots, n-1\}$. Now, multiplication by p^{-r} on $\mathbb{Z}/n\mathbb{Z} - \{0\}$ generates a cyclic group, and so a partition of $\{1, \ldots, n-1\}$ into m orbits $\mathcal{O}_\alpha$ of size s_α. Let j_α denote the least element of the orbit $\mathcal{O}_\alpha$. Reorder the columns of the matrix we are considering so that it is of the form

$$M := \begin{pmatrix} M_1 & \cdots & M_m \end{pmatrix} \tag{2.8}$$

where each M_α denotes the $2 \times s_\alpha$ matrix

$$M_\alpha := \begin{pmatrix} ab_{j_\alpha} & ab_{\nu(j_\alpha)} & \cdots & ab_{\nu^{s_\alpha-1}(j_\alpha)} \\ (-a)^{\frac{p^r\nu(j_\alpha)+n-j_\alpha}{n}} b_{\nu(j_\alpha)}^{p^r} & (-a)^{\frac{p^r\nu^2(j_\alpha)+n-\nu(j_\alpha)}{n}} b_{\nu^2(j_\alpha)}^{p^r} & \cdots & (-a)^{\frac{p^r j_\alpha + n - \nu^{s_\alpha-1}(j_\alpha)}{n}} b_{j_\alpha}^{p^r} \end{pmatrix}$$

It suffices to prove that for any M_α, there exists $a \in \overline{\mathbb{F}}_p^\times$ such that $M_\alpha(a)$ is nonsingular for all $\mathbf{b} \in \mathbb{P}^{n-2}$.

By construction, for each j, all monomials containing b_j appear in precisely one M_α. For ease of notation, denote

$$\epsilon_\alpha(t) := \frac{p^r\nu^t(j_\alpha) + n - \nu^{t-1}(j_\alpha)}{n}$$

so that

$$M_\alpha := \begin{pmatrix} ab_{j_\alpha} & ab_{\nu(j_\alpha)} & \cdots & ab_{\nu^{s_\alpha-1}(j_\alpha)} \\ (-a)^{\epsilon_\alpha(1)} b_{\nu(j_\alpha)}^{p^r} & (-a)^{\epsilon_\alpha(2)} b_{\nu^2(j_\alpha)}^{p^r} & \cdots & (-a)^{\epsilon_\alpha(s_\alpha)} b_{j_\alpha}^{p^r} \end{pmatrix}$$

Now the matrix (2.8) is singular at $\mathbf{b} \in \mathbb{P}^{n-2}$ and $a \in \overline{\mathbb{F}}_p$ if and only if its two rows are linearly dependent. Equivalently, there exists $\lambda \in \overline{\mathbb{F}}_p^\times$ such that for all α and $0 \le t \le s_\alpha - 1$

$$ab_{\nu^t(j_\alpha)} = \lambda(-a)^{\epsilon_\alpha(t+1)} b_{\nu^{t+1}(j_\alpha)}^{p^r}. \tag{2.9}$$

Restrict to $a \in \overline{\mathbb{F}}_p^\times$. Then, by induction on t, we obtain that for all $j \in \mathcal{O}_\alpha$

$$b_j = (-\lambda)^{\sum_{t=1}^{s_\alpha} p^{(t-1)r}} (-a)^{\sum_{t=1}^{s_\alpha} p^{(t-1)r}(\epsilon_\alpha(t)-1)} b_j^{p^{s_\alpha r}}.$$

Therefore, for any $b_j \neq 0$ for $j \in \mathcal{O}_\alpha$ (and such a j and α must exist since $\mathbf{b} \in \mathbb{P}^{n-2}$), we have

$$\begin{aligned} b_j^{p^{s_\alpha r}-1} &= (-\lambda)^{-\sum_{t=1}^{s_\alpha} p^{(t-1)r}} (-a)^{-\sum_{t=1}^{s_\alpha} p^{(t-1)r}(\epsilon_\alpha(t)-1)} \\ &=: c_\alpha(a) \end{aligned}$$

But, if $j = \nu^t(j_\alpha)$, then by Equation (2.9),

$$c_\alpha(a) = b_j^{p^{s_\alpha r}-1} = (-\lambda(-a)^{\epsilon_\alpha(t+1)-1})^{p^{s_\alpha r}-1} c_\alpha(a)^{p^r}.$$

Expanding the definition of $c_\alpha(a)$ in terms of λ and a, we obtain

$$\begin{aligned} &(-\lambda)^{-\sum_{t=1}^{s_\alpha} p^{(t-1)r}} (-a)^{-\sum_{t=1}^{s_\alpha} p^{(t-1)r}(\epsilon_\alpha(t)-1)} \\ &\qquad = (-\lambda)^{p^{s_\alpha r}-1-\sum_{t=1}^{s_\alpha} p^{tr}} (-a)^{(p^{s_\alpha r}-1)(\epsilon_\alpha(t+1)-1)-\sum_{t=1}^{s_\alpha} p^{tr}(\epsilon_\alpha(t)-1)} \\ &\qquad = (-\lambda)^{-\sum_{t=1}^{s_\alpha} p^{(t-1)r}} (-a)^{(p^{s_\alpha r}-1)(\epsilon_\alpha(t+1)-1)-\sum_{t=1}^{s_\alpha} p^{tr}(\epsilon_\alpha(t)-1)}. \end{aligned}$$

Therefore, for all $0 \leq t \leq s_\alpha$

$$1 = (-a)^{(p^{s_\alpha r}-1)(\epsilon_\alpha(t+1)-1)-\sum_{t=1}^{s_\alpha} p^{(t-1)r}(p^r-1)(\epsilon_\alpha(t)-1)}$$

In particular,

$$a^{2(p^{s_\alpha r}-1)(\epsilon_\alpha(t+1)-1)-\sum_{t=1}^{s_\alpha} p^{(t-1)r}(p^r-1)(\epsilon_\alpha(t)-1)} = 1. \tag{2.10}$$

But, $s_\alpha, \epsilon_\alpha(t), p, r \in \mathbb{N}$ are fixed once and for all by our choice of p and n. In particular, there exists $N \in \mathbb{N}$ such that

$$N > \max_\alpha |2(p^{s_\alpha r}-1)(\epsilon_\alpha(t+1)-1) - \sum_{t=1}^{s_\alpha} p^{(t-1)r}(p^r-1)(\epsilon_\alpha(t)-1)|.$$

But, then for any primitive N^{th} root of unity $a \in \overline{\mathbb{F}}_p$, Equation 2.10 is never satisfied. Therefore, the matrix $M(a) = (M_1(a) \cdots M_m(a))$ of (2.8) has full rank for all $\mathbf{b} \in \mathbb{P}^{n-2}$ as claimed.

Case 2: $p \mid n$**.** This case is similar. We specialize to the pencil $x^n + ax = 0$, i.e. $\mathbf{a} = (0, \ldots, a, 0)$. It suffices to show there exists a such that $T'_{12i}(a) := T'_{12i}(0, \ldots, a, 0)$ is smooth.

As noted in the proof of Lemma 2.6, over $\mathbb{Z}[1/(n-1)]$, and $a \neq 0$, we can use the coordinates

$$[b_1 : \cdots : b_{n-2}]$$

on $T'_1(a)$. We adopt the same modification of Notation 2.1 as in the proof of Lemma 2.6, setting

$$\begin{aligned} \mathbf{b}' &:= [b_1 : \cdots : b_{n-2}] \in \mathbb{P}^{n-3} \\ \kappa' &:= (k_1, \ldots, k_{n-2}) \in \mathbb{N}^{n-2} \end{aligned}$$

and carrying over the rest of the notation *mutatis mutandis*.

In these coordinates and this notation, the partial derivatives of $T'_{12}(a)$ are given by

$$\partial_{b_j} T_{12}(a) = -2(n-1)ab_{n-1-j}$$

(as noted in the proof of Lemma 2.6). Similarly, we have

$$\begin{aligned} T'_{1i}(a) =& (n-1)\cdot\left(\sum_{\ell=1}^{i-1}(-a)^\ell \sum_{\kappa' \ s.t.\ |\kappa'|=i, ||\kappa'||=\ell(n-1)} \binom{i}{\kappa'} \mathbf{b}'^{\kappa'}\right) \\ \partial_{b_j} T'_{1i}(a) =& i(n-1)\cdot\left(\sum_{\ell=1}^{i-1}(-a)^\ell \sum_{\kappa' \ s.t.\ |\kappa'|=i-1, ||\kappa'||+j=\ell(n-1)} \binom{i-1}{\kappa'} \mathbf{b}'^{\kappa'}\right) \end{aligned}$$

Define

$$\begin{aligned} T'_{j,12}(a) &:= ab_{n-1-j} \\ T'_{j,1i}(a) &:= \sum_{\ell=1}^{i-1}(-a)^\ell \sum_{\kappa' \ s.t.\ |\kappa'|=i-1, ||\kappa'||+j=\ell(n-1)} \binom{i-1}{\kappa'} \mathbf{b}'^{\kappa'} \end{aligned}$$

Just as in Case 1, the matrix

$$\begin{pmatrix} \partial_{b_1} T_{12}(a) & \cdots & \partial_{b_{n-2}} T_{12}(a) \\ \partial_{b_1} T_{1i}(a) & \cdots & \partial_{b_{n-2}} T_{1i}(a) \end{pmatrix}$$

is everywhere nonsingular over $\mathbb{Z}[1/(n-1)]$ for some a if and only if the matrix

$$\begin{pmatrix} T'_{1,12}(a) & \cdots & T'_{n-2,12}(a) \\ T'_{1,1i}(a) & \cdots & T'_{n-2,1i}(a) \end{pmatrix}$$

is everywhere nonsingular for some a. We now reduce this matrix mod p. Because $i = p^r + 1$, the mod p reduction of $T'_{j,1i}(a)$ is given by

$$\overline{T'_{j,1i}}(a) = \sum_{\ell=1}^{i-1}(-a)^\ell \sum_{\kappa' \ s.t.\ |\kappa'|=i-1, ||\kappa'||+j=\ell(n-1)} \binom{i-1}{\kappa'} \mathbf{b}'^{\kappa'}$$

In particular, because $i-1 = p^r$, and $p^r \in (\mathbb{Z}/(n-1)\mathbb{Z})^\times$, the same arguments as above allow us to define a permutation $\nu \circlearrowleft \{1, \ldots, n-2\} = (\mathbb{Z}/(n-1)\mathbb{Z}) - \{0\}$ by

$$\nu(j) = p^{-r} j \in (\mathbb{Z}/(n-1)\mathbb{Z}) - \{0\}.$$

Using ν, we have

$$\overline{T'_{j,1i}}(a) = (-a)^{\frac{p^r\nu(-j)+j}{n-1}} b_{\nu(-j)}^{p^r}$$

where, analogous to the above, we write $\pm j$, $\nu(\pm j)$ for the corresponding elements of $\{1, \ldots, n-2\}$. *Mutatis mutandis*, we now complete the argument by the same reasoning as for Case 1. $\square$

Remark 2.16. A similar argument shows that the Tschirnhaus hypersurface $T_i \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ itself is generically smooth for $i = p^r + 1$ and $r \geq 0$. More generally, we see no reason not to expect this, as well as Theorem 2.11, to hold without restriction on $i < n$. In principle, this comes down to checking whether an appropriate discriminant identically vanishes on T_i (resp. T_{12i}), i.e. checking a polynomial condition on the form defining T_i. However, this discriminant is a polynomial of degree $(n-1)(d-1)^{n-1}$ in the coefficients of the form, and the number of terms in this polynomial grows so quickly as to make direct computation impossible except for very small d and n.

3 Algebraic Functions and Tschirnhaus Transformations

In this section, we recall the theory of Tschirnhaus transformations of algebraic functions and relate this to the complete intersections studied above.

Let X be an irreducible K-variety. We write $K(X)$ for the rational functions on X. More generally, for a (not necessarily reducible) K-variety Y with irreducible components $\{Y_i\}$, let $K(Y) := \prod_i K(Y_i)$.

Recall that an *algebraic function* Φ on X is a finite (rational) correspondence $X \dashrightarrow^{1:n} \mathbb{A}^1$, i.e. Φ is given by a span

$$\begin{array}{ccc} E_\Phi & \xrightarrow{z} & \mathbb{A}^1 \\ {\scriptstyle\pi}\downarrow & & \\ X & & \end{array}$$

where π is a dominant, quasi-finite map and z is a regular function. We say Φ is *irreducible* if E_Φ is an irreducible K-variety and z is a primitive element of the finite field extension $K(E_\Phi)/K(X)$.

Let $\mathrm{Mon}(\Phi)$ denote the *monodromy group* of Φ, equivalently the Galois group of the normal closure of $K(X)(\Phi)/K(X)$. Let

$$m_\Phi(z) := z^n + a_1 z^{n-1} + \ldots + a_n$$

denote the minimal polynomial of z, where the $a_i \in K(X)$ (i.e. $m_\Phi(z)$ is the monic generator of the ideal of $K(X)[z]$ corresponding to the extension $K(X)(\Phi)$). A classical perspective describes Φ as the assignment

$$x \mapsto \{z \in \bar{K} \mid m_{\Phi(x)}(z) = z^n + a_1(x)z^{n-1} + \ldots + a_n(x) = 0\}. \tag{3.1}$$

For any field extension $K(X) \hookrightarrow L$, write

$$L(\Phi) := L \otimes_{K(X)} K(X)(\Phi).$$

Note that since $\{1, z, \ldots, z^{n-1}\}$ is a basis for $K(X)(\Phi)$ over $K(X)$, it is also a basis for $L(\Phi)$ over L. Given this, for each $w \in L(\Phi)$, there exists unique $b_0, \ldots, b_{n-1} \in L$ such that

$$w = \sum_{i=0}^{n-1} b_i z^i.$$

Moreover, $\tilde{\mathbf{b}} = (b_0, \ldots, b_{n-1}) \in L^n$ determines an L-linear transformation

$$T_{\tilde{\mathbf{b}}} : L(\Phi) \longrightarrow L(\Phi)$$

given by (extending L-linearly) the assignment $T_{\tilde{\mathbf{b}}}(z^j) := w^j$ for each $0 \leq j \leq n-1$. Note that $T_{\tilde{\mathbf{b}}}$ is an automorphism if and only if w is a primitive element of the extension $L(\Phi)/L$.

Definition 3.1. Let X be an irreducible K-variety. Let Φ be an irreducible algebraic function on X with primitive element $z \in K(X)(\Phi)$. A *Tschirnhaus transformation* T of Φ is a $\overline{K(X)}$-linear automorphism

$$T\colon \overline{K(X)}(\Phi) \longrightarrow \overline{K(X)}(\Phi).$$

of the form

$$z^j \mapsto w^j = \left(\sum_{i=0}^{n-1} b_i z^i\right)^j$$

for $b_0, \ldots, b_{n-1} \in \overline{K(X)}$. We say the transformation is *rational* over X if $b_0, \ldots, b_{n-1} \in K(X)$. More generally, we say it is *rational* over $L/K(X)$ if all $b_i \in L$.

Picking an integral model $Y \longrightarrow X$ for $K(X)(\tilde{\mathbf{b}})/K(X)$, (i.e. a map of K-varieties $Y \longrightarrow X$ and an isomorphism $K(Y) \cong K(X)(\tilde{\mathbf{b}})$ as extensions of $K(X)$), we denote by $T(\Phi)$ the algebraic function on Y determined by the primitive element $w \in K(Y)(\Phi)$.

Now let Φ be an algebraic function as above, and T a Tschirnhaus transformation of Φ. Let $w = T(z)$, and let the minimal polynomial of multiplication by w on $\overline{K(X)}(\Phi)$ be given by

$$m_{T(\Phi)}(w) := w^n + c_1 w^{n-1} + \ldots c_n$$

where $c_i \in L = K(Y)$. The algebraic function $T(\Phi)$ on Y is given by the assignment

$$y \mapsto \{z \in \overline{K(X)} \mid m_{T(\Phi)(y)}(z) = z^n + c_1(y) z^{n-1} + \ldots + c_n(y) = 0\}.$$

Recall that $\mathbb{A}^n_X := X \times_{Spec(K)} \mathbb{A}^n_K$, viewed as a scheme over X.

Lemma 3.2. *Let X be irreducible, and let Φ be an irreducible, generically n-valued algebraic function on X. Then there is an open subvariety*

$$\mathcal{T}_\Phi \subset \mathbb{A}^n_X,$$

such that for all finite extensions $L/K(X)$, $\mathcal{T}_\Phi(L)$ is the set of Tschirnhaus transformations of Φ which are rational over L. In particular, the map

$$\begin{array}{ccc} \mathcal{T}_\Phi & \hookrightarrow & \mathbb{A}^n_X \\ & \searrow & \downarrow \\ & & X \end{array}$$

is smooth. Equivalently the parameter space of Tschirnhaus transformations $\mathcal{T}_\Phi \longrightarrow X$ is smooth over X.

Proof. We begin by constructing the variety $\mathcal{T}_\Phi$. Denote the set of $\overline{K(X)}$-rational Tschirnhaus transformations of Φ by $\mathcal{T}_\Phi(\overline{K(X)})$. We will show that this embeds as an explicit Zariski open subset of $\overline{K(X)}^n = \mathbb{A}^n_X(\overline{K(X)})$, and that its complement is defined over $K(X)$; we thus conclude that $\mathcal{T}_\Phi(\overline{K(X)})$ is the set of geometric generic points of a variety $\mathcal{T}_\Phi \subset \mathbb{A}^n_X$.

Let $z \in K(X)(\Phi)$ be the primitive element determined by Φ. Given $\tilde{\mathbf{b}} \in \overline{K(X)}^n$, we have a $\overline{K(X)}$-linear endomorphism

$$T_{\tilde{\mathbf{b}}} \colon \overline{K(X)}(\Phi) \longrightarrow \overline{K(X)}(\Phi)$$

given by

$$z^j \mapsto \left(\sum_{i=0}^{n-1} b_i z^i\right)^j .$$

Moreover, the assignment $\tilde{\mathbf{b}} \mapsto T_{\tilde{\mathbf{b}}}$ defines a $\mathrm{Gal}(\overline{K(X)}/K(X))$-equivariant map

$$T \colon \mathbb{A}^n(\overline{K(X)}) \longrightarrow \mathrm{End}_{\overline{K(X)}}(\overline{K(X)}(\Phi)) \cong \mathbb{A}^{n^2}(\overline{K(X)}).$$

By definition, $\mathcal{T}_\Phi(\overline{K(X)})$ is in bijection with the set

$$\{\tilde{\mathbf{b}} \in \overline{K(X)}^n \mid T_{\tilde{\mathbf{b}}} \in Aut_{\overline{K(X)}}(\overline{K(X)}(\Phi))\}$$

i.e.

$$\mathcal{T}_\Phi(\overline{K(X)}) = T^{-1}(Aut_{\overline{K(X)}}(\overline{K(X)}(\Phi))).$$

Since $Aut_{\overline{K(X)}}(\overline{K(X)}(\Phi))$ is the pullback to $\overline{K(X)}$ of an open subscheme of $\mathbb{A}^{n^2}_{\mathbb{Z}}$ (i.e. the locus $\{\det \neq 0\}$)) and T is defined over $K(X)$, we conclude that $\mathcal{T}_\Phi(\overline{K(X)}) \subset \mathbb{A}^n(\overline{K(X)}$ is Zariski open and defined over $K(X)$ as claimed. The remaining claims follow by direct inspection. □

Corollary 3.3. *Let Φ be an irreducible n-valued algebraic function on X such that $K(X)(\Phi)/K(X)$ has no intermediate subfields. Let $\mathbb{A}^n_X$ be given coordinates $(b_0, \ldots, b_{n-1})$ as above, and let $\mathbb{A}^1_{X,0} \subset \mathbb{A}^n_X$ denote the b_0-axis. Then*

$$\mathcal{T}_\Phi = \mathbb{A}^n_X - \mathbb{A}^1_{X,0}.$$

Proof. Because $K(X)(\Phi)/K(X)$ has no intermediate subfields, $y \in K(X)(\Phi)$ is a primitive element if and only if $y \notin K(X)$, i.e. if and only if y is of the form $y = \sum_{i=0}^{n-1} b_i z^i$ with $b_i \neq 0$ for some $i > 0$. □

Example 3.4. Let $X = \mathbb{A}^n_{\mathbf{a}}$, viewed as the parameter space for monic, degree n polynomials (parametrized by their coefficients $\mathbf{a} := (a_1, \ldots, a_n)$). Let P_n be the general degree n polynomial, i.e.

$$m_{P_n}(z) = z^n + a_1 z^{n-1} + \ldots + a_n.$$

Then the degree n extension $K(\mathbb{A}^n_{\mathbf{a}})(P_n)/K(\mathbb{A}^n_{\mathbf{a}})$ has no intermediate subfields, because it corresponds to the maximal subgroup $S_{n-1} \subset S_n = \mathrm{Mon}(P_n)$. In particular, the space of Tschirnhaus transformations of the general degree n polynomial is given by

$$\begin{aligned} \mathcal{T}_{P_n} &= \mathbb{A}^n_X - \mathbb{A}^1_{X,0} \\ &:= \mathbb{A}^n_{\tilde{\mathbf{b}}} \times \mathbb{A}^n_{\mathbf{a}} - \mathbb{A}^1_{b_0} \times \mathbb{A}^n_a \\ &= (\mathbb{A}^n_{\tilde{\mathbf{b}}} - \mathbb{A}^1_{b_0}) \times \mathbb{A}^n_{\mathbf{a}}. \end{aligned}$$

Now let Φ be an irreducible algebraic function on X, and let T be a Tschirnhaus transformation of Φ as above, with minimal polynomial

$$m_{T(\Phi)}(y) := y^n + c_1 y^{n-1} + \dots c_n$$

Observe that the assignment

$$x \mapsto (c_1(x), \dots, c_n(x))$$

determines a rational map

$$X \dashrightarrow \mathbb{A}^n$$

which fits into a pullback square

$$\begin{array}{ccc} E_\Phi & \dashrightarrow & E_{P_n} \\ {\scriptstyle\pi}\downarrow & & \downarrow \\ X & \dashrightarrow & \mathbb{A}^n \end{array}$$

In particular, the Tschirnhaus transformation T transforms Φ into a function of $d = \dim(\mathrm{Image}(X \dashrightarrow \mathbb{A}^n))$ variables.

We now study loci of interest in the space of Tschirnhaus transformations. The basic observation (essentially going back to Tschirnhaus [Ts1683]) is as follows. First, the collection of n-valued algebraic functions on X is given by $\mathbb{A}^n_X$, where $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{A}^n_X$ corresponds to the function $\Phi_{\mathbf{a}}$ of (3.1), i.e. the function

$$x \mapsto \{z \in \bar{K} \mid m_{\Phi_{\mathbf{a}}(x)}(z) = z^n + a_1(x) z^{n-1} + \dots + a_n(x) = 0\}.$$

Next, the assignment $(\Phi_{\mathbf{a}}, \tilde{\mathbf{b}}) \mapsto T_{\tilde{\mathbf{b}}}(\Phi_{\mathbf{a}})$ determines an "evaluation" map

$$\begin{aligned} \mathbb{A}^n_{X,\mathbf{a}} \times \mathbb{A}^n_{X,\tilde{\mathbf{b}}} &\xrightarrow{\mathrm{ev}} \mathbb{A}^n_{X,\mathbf{a}} \\ (\mathbf{a}, \tilde{\mathbf{b}}) &\mapsto T_{\tilde{\mathbf{b}}}(\mathbf{a}) \end{aligned}$$

(where we write $(-)_{\mathbf{a}}$ and $(-)_{\tilde{\mathbf{b}}}$ to distinguish the different roles of the $\mathbf{a}$ and $\tilde{\mathbf{b}}$ coordinates). The coordinates of $T_{\tilde{\mathbf{b}}}(\mathbf{a})$ can be computed explicitly as follows. By definition, $\tilde{\mathbf{b}} \in \mathbb{A}^n_X$ corresponds to the assignment

$$z \mapsto \sum_{i=0}^{n-1} b_i z^i = y$$

for z a value of $\Phi_{\mathbf{a}}$. Passing to a Galois closure of $K(X)(\Phi)$, the transformation T maps the roots z_i of m_Φ to y_i given by

$$y_i = \sum_{j=0}^{n-1} b_j z_i^j.$$

In particular, the polynomial $m_{T(\Phi)}$ is given by

$$m_{T(\Phi)}(y) = \prod_{i=1}^{n}(y - y_i).$$

i.e. the coordinates of T_Φ are obtained (up to sign) by expanding the elementary symmetric polynomials in the y_i as polynomials in $\mathbf{b}$ with coefficients given by polynomials in the coordinates $\mathbf{a}$. In particular, the j^{th} coefficient is a homogeneous polynomial of total degree j in the coordinates $\tilde{\mathbf{b}}$.

As a result, every Zariski closed subscheme $Z \subset \mathbb{A}^n_{X,\mathbf{a}}$ determines a Zariski closed subscheme

$$\mathrm{ev}^{-1}(Z) \subset \mathbb{A}^n_{X,\mathbf{a}} \times \mathbb{A}^n_{X,\tilde{\mathbf{b}}},$$

Specializing to a particular algebraic function Φ, and its space of Tschirnhaus transformations $\mathcal{T}_\Phi \subset \mathbb{A}^n_{X,\tilde{\mathbf{b}}}$, we obtain a Zariski closed subscheme (concretely $\mathcal{T}_\Phi \cap \mathrm{ev}^{-1}(Z)$), which, by abuse of notation, we denote again by

$$\mathrm{ev}^{-1}(Z) \subset \mathcal{T}_\Phi.$$

By construction, this subscheme parametrizes Tschirnhaus transformations of Φ such that $T(\Phi)$ (or more precisely, the coefficients of its minimal polynomial) lie in $Z \subset \mathbb{A}^n_{X,\mathbf{a}}$.

We can now make contact with the Tschirnhaus complete intersections introduced in Section 2. For $1 \leq i_1 < \ldots < i_k$, define

$$Z_{i_1 \cdots i_k} := \{\mathbf{a} \in \mathbb{A}^n_{\mathbf{a}} \mid p_{i_1}(\mathbf{a}) = \cdots = p_{i_k}(\mathbf{a}) = 0\}$$

where the p_is are as in Section 2.

Definition 3.5. Let $n > 0$. For $1 \leq i_1 < \ldots < i_k$, define the *affine Tschirnhaus complete intersection* $\tilde{T}_{i_1 \cdots i_k}(P_n)$ to be

$$\tilde{T}_{i_1 \cdots i_k}(P_n) := \mathrm{ev}^{-1}(Z_{i_1 \cdots i_j}) \subset \mathcal{T}_{P_n} \subset (\mathbb{A}^n_{\tilde{\mathbf{b}}} - \mathbb{A}^1_{b_0=0}) \times \mathbb{A}^n_{\mathbf{a}}.$$

Projecting onto $\mathbb{A}^n_{\mathbf{a}}$ gives the family $\tilde{T}_{i_1 \cdots i_k}(P_n) \longrightarrow \mathbb{A}^n_{\mathbf{a}}$.

Similarly, define the *Tschirnhaus complete intersection*

$$T_{i_1 \cdots i_k}(P_n) \subset (\mathbb{P}^{n-1}_{\mathbf{b}} - \{[1 : 0 : \cdots : 0]\}) \times \mathbb{A}^n_{\mathbf{a}}$$

to be the (fiberwise) projectivization of the family $\tilde{T}_{i_1 \cdots i_k}(P_n) \longrightarrow \mathbb{A}^n_{\mathbf{a}}$.

Define the *reduced affine Tschirnhaus complete intersection* by

$$\tilde{T}'_{i_1 \cdots i_k}(P_n) := T_{i_1 \cdots i_k} \cap \{b_0 = 0\}.$$

Similarly, define the *reduced Tschirnhaus complete intersection*

$$T'_{i_1 \cdots i_k}(P_n) \subset \mathbb{P}^{n-2}_{\mathbf{b}'} \times \mathbb{A}^n_{\mathbf{a}}$$

to be the (fiberwise) projectivization of the family $\tilde{T}'_{i_1 \cdots i_k} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$.

Lemma 2.8 can now be equivalently restated as follows.

Lemma 3.6. *For all n and all $1 \leq i_1 \leq \cdots \leq i_k$, we have*

$$\mathcal{T}_{i_1 \cdots i_k}(P_n) = T^n_{i_1 \cdots i_k}$$

as subschemes of $\mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-1}_{\mathbf{b}}$, where the right hand side denotes the Tschirnhaus complete intersection of Definition 2.4.

Similarly, we have

$$\mathcal{T}'_{i_1 \cdots i_k}(P_n) = T^{n'}_{i_1 \cdots i_k}$$

as subschemes of $\mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-2}_{\mathbf{b}}$.

4 The Resolvent Degree of a Dominant Map

Recall the following (see [Br75, AS76, FW18]).[8]

Definition 4.1 (Resolvent degree). Let $Y \longrightarrow X$ be a generically finite dominant map of K-varieties. Its *resolvent degree* $\mathrm{RD}(Y \longrightarrow X)$ is the minimum d for which there exists a dense Zariski open $U \subset X$ and a tower of generically finite dominant maps

$$E_r \longrightarrow \cdots \longrightarrow E_1 \longrightarrow E_0 = U$$

such that $E_r \longrightarrow U$ factors through a dominant map $E_r \longrightarrow Y$ and such that for each $i \geq 0$, there exists a pullback diagram

$$\begin{array}{ccc} E_i & \longrightarrow & \tilde{Z}_i \\ \downarrow & & \downarrow \\ E_{i-1} & \longrightarrow & Z_i \end{array}$$

where $\tilde{Z}_i \longrightarrow Z_i$ is a generically finite dominant map with $\dim(Z_i) \leq d$.

Example 4.2. Consider the space $\mathbb{A}^n_{\mathbf{a}}$ of monic degree n-polynomials. This has a canonical n-sheeted branched cover $E_{P_n} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ where E_{P_n} is the space of monic degree n polynomials with a choice of root, and the map forgets the root. By definition

$$\mathrm{RD}(n) := \mathrm{RD}(E_{P_n} \longrightarrow \mathbb{A}^n_{\mathbf{a}}).$$

We now extend the notion of resolvent degree to general dominant maps. We adopt the following convention to avoid irrelevant pathologies.

Convention 4.3. *By a* dominant *map, we mean a map $Y \longrightarrow X$ that is both dominant, and is such that every irreducible component of Y maps dominantly onto some irreducible component of X.*

[8] With the exception of Proposition 4.11, the results of this section are valid in arbitrary characteristic.

Definition 4.4 (**Rational multi-section).** Let $Y \xrightarrow{\pi} X$ be a dominant map of K-varieties. A *rational multi-section* is a subvariety $U \subset Y$ such that the restriction $\pi|_U : U \longrightarrow X$ is a generically finite dominant map.

Lemma 4.5. *Every dominant map $Y \longrightarrow X$ admits a dense set of rational multi-sections, i.e. the closure of their union is all of Y.*

Proof. First assume that X is irreducible. Let $\overline{K(X)}$ be an algebraic closure of the rational functions of X. Then every point of $Y(\overline{K(X)})$ is a germ of a rational multi-section, and the closure of the union of all of these contains the generic fiber of $Y \longrightarrow X$; in particular it is dense. For the general case, the argument above exhibits a dense set of rational multi-sections over each irreducible component. Their union gives a dense set of rational multi-sections of $Y \longrightarrow X$. □

It will be useful to extend the definition of resolvent degree from generically finite dominant rational maps to all dominant rational maps.

Definition 4.6 (**Resolvent degree of a dominant map).** Let $Y \xrightarrow{\pi} X$ be a dominant map of K-varieties. The *resolvent degree of the dominant map*, $\mathrm{RD}(Y \longrightarrow X)$ is defined to be the minimum d for which there exists a dense set of rational multi-sections $\{U_\alpha \subset Y\}$ with $\mathrm{RD}(U_\alpha \longrightarrow X) \leq d$ for all α.

We will need a few basic facts about the resolvent degree of a dominant map.

Lemma 4.7. *Let $Y \longrightarrow X$ be a dominant map of K-varieties.*

1. $\mathrm{RD}(Y \longrightarrow X) \leq \dim(X)$.
2. *Let $Z \longrightarrow X$ be any dominant map of K-varieties. Then*
$$\mathrm{RD}(Y \times_X Z \longrightarrow Z) \leq \mathrm{RD}(Y \longrightarrow X).$$
3. *If $Y \longrightarrow X$ is birationally equivalent to $W \longrightarrow Z$, then*
$$\mathrm{RD}(Y \longrightarrow X) = \mathrm{RD}(W \longrightarrow Z).$$
4. *If $X = \bigcup X_i$ is a union of irreducible components, write $\{Y_{i,j}\}$ for the set of irreducible components of Y which dominate X_i. Then*
$$\mathrm{RD}(Y \longrightarrow X) = \max_{i,j}\{\mathrm{RD}(Y_{i,j} \longrightarrow X_i)\}.$$

Proof. These follow immediately from the definition and the analogous properties for resolvent degree of generically finite dominant maps (cf. [FW18, Lemmas 2.7, 2.8]). □

Lemma 4.8. *Let $Y \longrightarrow X$ be a surjective map (on geometric points). Let $Z \longrightarrow X$ be any map. Then*
$$\mathrm{RD}(Y|_Z \longrightarrow Z) \leq \dim(X).$$

Proof. Let $W \subset X$ be the Zariski closure of the image of $Z \longrightarrow X$. By construction, the map $Z \longrightarrow W$ is dominant. The surjectivity of $Y \longrightarrow X$ implies that the restriction

$$Y|_W \longrightarrow W$$

is dominant. Therefore, by Lemma 4.7,

$$\begin{aligned}\mathrm{RD}(Y|_Z \longrightarrow Z) &\leq \mathrm{RD}(Y|_W \longrightarrow W)\\ &\leq \dim(W)\\ &\leq \dim(X).\end{aligned}$$

□

Lemma 4.9. *Let $Y \longrightarrow X$ be a generically finite dominant map. Then Definition 4.6 specializes to Definition 4.1 for $Y \longrightarrow X$, i.e. they give equivalent notions of resolvent degree.*

Proof. By Lemma 4.7 4 and [FW18, Lemma 2.8], it suffices to prove this when Y is irreducible. In this case, any rational multi-section $U \subset Y$ of $Y \longrightarrow X$ must be dense in Y. In particular, it must be birational to Y. From the birational invariance of RD for generically finite dominant maps, we conclude that $\mathrm{RD}(U \longrightarrow X) = \mathrm{RD}(Y \longrightarrow X)$ (as generically finite dominant maps). The lemma follows. □

Lemma 4.10. *Let $Z \xrightarrow{\pi_1} Y \xrightarrow{\pi_2} X$ be a pair of dominant maps of K-varieties. Then*

$$\mathrm{RD}(Z \longrightarrow X) \geq \mathrm{RD}(Y \longrightarrow X)$$

and

$$\mathrm{RD}(Z \longrightarrow X) \leq \max\{\mathrm{RD}(Z \longrightarrow Y), \mathrm{RD}(Y \longrightarrow X)\}.$$

with equality when either $Z \longrightarrow Y$ or $Y \longrightarrow X$ is generically finite.

Proof. For the first inequality, let $\{U_\alpha \subset Z\}$ be a dense set of rational multi-sections of $Z \longrightarrow X$ with $\mathrm{RD}(U_\alpha \longrightarrow X) \leq d$ for all α. Then, shrinking each U_α as necessary (e.g. restricting to the preimage in U of an affine open in Y), its (scheme theoretic) image $V_\alpha := \mathrm{Image}(U_\alpha \longrightarrow Y)$ is a subscheme of Y, and thus a rational multi-section of $Y \longrightarrow X$. Since $Z \longrightarrow Y$ is dominant, that $\{U_\alpha \subset Z\}$ is dense implies that $\{V_\alpha \subset Y\}$ is dense. By [FW18, Lemma 2.9], we conclude that $\mathrm{RD}(U_\alpha \longrightarrow X) \geq \mathrm{RD}(V_\alpha \longrightarrow X)$. Minimizing over all $\{U_\alpha \subset Z\}$, we conclude that

$$\mathrm{RD}(Z \longrightarrow X) \geq \mathrm{RD}(Y \longrightarrow X).$$

For the second inequality, let $\{U_\alpha \subset Z\}$ be a dense set of rational multi-sections for $Z \longrightarrow Y$ and $\{V_\beta \subset Y\}$ a dense set of rational multi-sections for $Y \longrightarrow X$. Then

$$\{W_{\alpha,\beta} := U_\alpha \times_Y V_\beta \subset Z\}$$

is a dense set of rational multi-sections for $Z \longrightarrow X$. By [FW18, Lemmas 2.7, 2.9],

$$\mathrm{RD}(W_{\alpha,\beta} \longrightarrow X) \leq \max\{\mathrm{RD}(U_\alpha \longrightarrow Y), \mathrm{RD}(V_\beta \longrightarrow X)\}.$$

Minimizing over all such collections $\{U_\alpha\}, \{V_\beta\}$, we conclude

$$\mathrm{RD}(Z \longrightarrow X) \leq \max\{\mathrm{RD}(Z \longrightarrow Y), \mathrm{RD}(Y \longrightarrow X)\}.$$

To show the equalities when $\dim(Y) = \dim(X)$ or $\dim(Z) = \dim(Y)$, it suffices, by Lemma 4.74, to prove the case when X and Y are irreducible. Under this assumption, if $\dim(X) = \dim(Y)$ or if $\dim(Z) = \dim(Y)$, then any rational multi-section U for $Z \longrightarrow Y$ is a rational multi-section for $Z \longrightarrow X$ and vice versa. In particular,

$$\mathrm{RD}(U \longrightarrow Y) \leq \mathrm{RD}(U \longrightarrow X)$$

and taking the minimum over dense subsets of such, we see that $\mathrm{RD}(Z \longrightarrow Y) \leq \mathrm{RD}(Z \longrightarrow X)$. The equality

$$\mathrm{RD}(Z \longrightarrow X) = \max\{\mathrm{RD}(Z \longrightarrow Y), \mathrm{RD}(Y \longrightarrow X)\}$$

follows from what we have shown above. □

Special cases of the following are implicit in [Se45, Br45, Br75].

Proposition 4.11. *Let $Y \longrightarrow X$ be a dominant map of K-varieties. Let $S \longrightarrow X$ be a Severi-Brauer variety and let $\overline{K(X)}$ be an algebraic closure of $K(X)$. Suppose that there exists an embedding over X*

$$Y \hookrightarrow S$$

such that the closure of the geometric generic fiber $Y|_{\overline{K(X)}}$ in $S|_{\overline{K(X)}} \cong \mathbb{P}^n_{\overline{K(X)}}$ has degree d. Then

$$\mathrm{RD}(Y \longrightarrow X) \leq \mathrm{RD}(d) < d.$$

Proof. By the Merkurjev-Suslin theorem [MS83, Theorem 16.1], using that K is a field of characteristic 0, there exists a solvable étale map $E \longrightarrow X$ such that $S|_E \cong \mathbb{P}^n_E$. By Galois descent, we conclude that the embedding $Y \hookrightarrow S$ pulls back to an embedding

$$Y|_E \hookrightarrow S|_E \cong \mathbb{P}^n_E$$

whose closure is a degree d subvariety. Points of $Y|_E$ are thus of height at most d over $K(E)$ (and the general point is of height d). Therefore, by [FW18, Lemma 2.11], $Y|_E$ admits a dense set of rational multi-sections $\{U_\alpha \subset Y|_E\}$ with $\mathrm{RD}(U_\alpha \longrightarrow E) \leq \mathrm{RD}(d)$. The images of these rational multi-sections in Y, $\{V_\alpha \subset Y\}$ are thus a dense set of rational multi-sections, and by [FW18, Lemma 2.8], we have

$$\begin{aligned}\mathrm{RD}(V_\alpha \longrightarrow X) &\leq \mathrm{RD}(U_\alpha \longrightarrow E \longrightarrow X)\\ &= \max\{\mathrm{RD}(U_\alpha \longrightarrow E), \mathrm{RD}(E \longrightarrow X)\}\\ &\leq \max\{\mathrm{RD}(d), 1\} = \mathrm{RD}(d) < d.\end{aligned}$$

□

Now let X be a variety, and let $\mathbb{A}^n_{X,\mathbf{a}}$ be the parameter space for n-valued algebraic functions on X as in Section 3. Observe that the action of $\mathbb{G}_m$ on algebraic functions by rescaling their values corresponds to a weighted action $\mathbb{G}_m \circlearrowleft \mathbb{A}^n_{X,\mathbf{a}}$ where

$$\lambda \cdot (a_1, \ldots, a_n) = (\lambda a_1, \ldots, \lambda^n a_n).$$

Moreover, if $Z \subset \mathbb{A}^n_{X,\mathbf{a}}$ is weighted homogeneous with respect to this action, then $\mathrm{ev}^{-1}(Z) \subset \mathcal{T}_\Phi$ is homogeneous (with respect to the diagonal action of $\mathbb{G}_m$ on $\mathbb{A}^n_{X,\mathbf{b}}$).

Lemma 4.12. *Let X be an irreducible K-variety. Let Φ be an algebraic function on X. Let $Z \subset \mathbb{A}^n_{X,\mathbf{a}}$ be a Zariski closed subvariety which is weighted homogeneous (relative to the above action). Let*

$$U \subset \mathrm{ev}^{-1}(Z) \subset \mathcal{T}_\Phi$$

be any rational multi-section for $\mathrm{ev}^{-1}(Z) \longrightarrow X$. Then

$$\mathrm{RD}(\Phi) \leq \max\{\mathrm{RD}(U \longrightarrow X), \dim(Z) - 1\}.$$

Proof. The multi-section $U \longrightarrow \mathrm{ev}^{-1}(Z)$ determines a Tschirnhaus transformation T of $\Phi|_U$ which is rational over $K(U)$. By the observations above, we have a pullback square

$$\begin{array}{ccc} (E_\Phi)|_U & \dashrightarrow & (E_{P_n})|_Z \\ \downarrow & & \downarrow \\ U & \dashrightarrow & Z \end{array}$$

Since Z is weighted homogeneous, we can projectivize $(E_{P_n})_Z \longrightarrow Z$ to obtain a pullback square

$$\begin{array}{ccc} (E_\Phi)|_U & \dashrightarrow & \mathbb{P}(E_{P_n})|_{\mathbb{P}(Z)} \\ \downarrow & & \downarrow \\ U & \dashrightarrow & \mathbb{P}(Z) \end{array}$$

where $\mathbb{P}(Z) \subset \mathbb{P}(\mathbb{A}^n_{\mathbf{a}})$ and $\mathbb{P}(\mathbb{A}^n_{\mathbf{a}})$ now denotes the weighted projective space. The result now follows by applying Lemmas 4.7 and 4.10. $\square$

5 Hilbert's Formula for the Degree 9 and New General Upper Bounds

We now apply the results of the previous sections to complete and extend Hilbert's argument from [Hi27]. We work throughout this section over an algebraically closed field K of characteristic 0.

Let $\mathcal{H}_{d,N}$ denote the *parameter space* of degree d hypersurfaces in $\mathbb{P}^N$, i.e. $\mathcal{H}_{d,N} \cong \mathbb{P}^{\binom{N+d}{d}-1}$. Let $\mathcal{M}_{d,N}$ denote the coarse moduli space of smooth hypersurfaces, i.e

$$\mathcal{M}_{d,N} = (\mathcal{H}_{d,N} - \Sigma)/\,\mathrm{PGL}_{N+1}$$

where Σ denotes the locus of singular hypersurfaces. Let $\mathcal{H}^r_{d,N}$ denote the space of such hypersurfaces with a choice of r-plane on them, i.e. $\mathcal{H}^r_{d,N}$ is the incidence variety

$$\mathcal{H}^r_{d,N} := \{(X, L) \in \mathcal{H}_{d,N} \times Gr(r+1, N+1) \mid L \subset X\}.$$

Similarly to above, let $\mathcal{M}^r_{d,N}$ denote the moduli of smooth degree d hypersurfaces equipped with an incident r-plane, i.e.

$$\mathcal{M}^r_{d,N} = (\mathcal{H}^r_{d,N} - \tilde{\Sigma}) / \operatorname{PGL}_{N+1},$$

where $\tilde{\Sigma} \subset \mathcal{H}^r_{d,N}$ denotes the locus where the hypersurface is singular.

We will need the following theorem of Waldron [Wa08, Theorem 1.6] (see also [St17, Theorem 1.2]).

Theorem 5.1 (Waldron)**.** *Let $d \geq 3$. The map*

$$\mathcal{H}^r_{d,N} \longrightarrow \mathcal{H}_{d,N}$$

is surjective for r, N such that

$$(r+1)(N-r) - \binom{d+r}{r} \geq 0.$$

Motivated by this theorem, we introduce the following notation:

Notation 5.2. Given $(d, k) \in \mathbb{N}_{\geq 3} \times \mathbb{N}$, define

$$\psi(d, k)_0 = k.$$

For $0 \leq i < d-2$, define

$$\psi(d, k)_{i+1} = \lceil \psi(d, k)_i + \binom{\psi(d, k)_i + d - i}{\psi(d, k)_i} / (\psi(d, k)_i + 1) \rceil.$$

Finally, define

$$\psi(d, k)_{d-1} = 2\psi(d, k)_{d-2} + 1.$$

By Waldron's Theorem, for all $0 \leq i < d-2$, the map

$$\mathcal{H}^{\psi(d,k)_i}_{d-i, \psi(d,k)_{i+1}} \longrightarrow \mathcal{H}_{d-i, \psi(d,k)_{i+1}}$$

is surjective. Similarly, by the classical theory of quadratic forms, the locus of smooth quadrics is contained in the image of the map

$$\mathcal{H}^{\psi(d,k)_{d-2}}_{2, \psi(d,k)_{d-1}} \longrightarrow \mathcal{H}_{2, \psi(d,k)_{d-1}}$$

In words, the integers $\psi(d, k)_i$ are defined so that every smooth quadric in a $\mathbb{P}^{\psi(d,k)_{d-1}}$ contains a $\psi(d, k)_{d-2}$ plane, every cubic hypersurface in this $\psi(d, k)_{d-2}$ plane contains a $\psi(d, k)_{d-3}$ plane, every quartic in this $\psi(d, k)_{d-3}$ plane contains a $\psi(d, k)_{d-4}$ plane, and on down until we arrive at a $\psi(d, k)_1$ plane such that every degree d hypersurface inside it contains a k-plane.

Lemma 5.3. *For all $d \geq 3$ and all $k \geq 1$,*

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) \geq \max\{\dim(\mathcal{H}_{d-i,\psi(d,k)_{i+1}})\}_{i=0}^{d-3}$$

and

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) + d + k + 1 \geq \psi(d,k)_{d-1} + 2.$$

Proof. For each i,

$$\dim(\mathcal{H}_{d-i,\psi(d,k)_{i+1}}) = \binom{d-i+\psi(d,k)_{i+1}}{d-i} - 1$$

From the definition of the $\psi(d,k)_i$s, we conclude for all i that

$$\dim(\mathcal{H}_{d-i,\psi(d,k)_{i+1}}) \geq \dim(\mathcal{H}_{d-i+1,\psi(d,k)_i})$$

and thus

$$\dim(\mathcal{H}_{4,\psi(d,k)_{d-3}}) = \max\{\dim(\mathcal{H}_{d-i,\psi(d,k)_{i+1}})\}_{i=0}^{d-4}.$$

Similarly,

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) = \binom{3+\psi(d,k)_{d-2}}{3} - (\psi(d,k)_{d-2}+1)^2.$$

From the definition, this is a monotone increasing degree 6 polynomial in $\psi(d,k)_{d-3}$, while $\dim(\mathcal{H}_{4,\psi(d,k)_{d-3}})$ is a monotone increasing quartic. Therefore, the inequality

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) \geq \dim(\mathcal{H}_{4,\psi(d,k)_{d-3}})$$

for all (d,k) follows from the inequality for $(d,k) = (4,1)$ (for which the left hand side is 220 while the right hand side is 70).

Finally, from the definition,

$$\psi(d,k)_{d-1} + 2 = 2\psi(d,k)_{d-2} + 3.$$

By the same reasoning as above, the inequality

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) + d + k + 1 \geq \psi(d,k)_{d-1} + 2$$

for all $(d,k) \in \mathbb{N}_{\geq 3} \times \mathbb{N}_{>0}$ follows from the inequality for $(d,k) = (3,1)$ (in which case both left and right hand side are 9). □

The lemma implies that for $d \geq 3$, $\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})$ gives a coarse upper bound on the resolvent degree of the surjective maps

$$\mathcal{M}_{3,\psi(d,k)_{d-2}}^{\psi(d,k)_{d-3}} \longrightarrow \mathcal{M}_{3,\psi(d,k)_{d-2}}$$
$$\mathcal{H}_{d-i,\psi(d,k)_{i+1}}^{\psi(d,k)_i} \longrightarrow \mathcal{H}_{d-i,\psi(d,k)_{i+1}}.$$

This motivates the following definition.

Notation 5.4. Given $(d,k) \in \mathbb{N}_{\geq 3} \times \mathbb{N}_{>0}$, define

$$\Phi(d,k) := \max\{\frac{(d+k)!}{d!} + 1, \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) + d + k + 1\}$$

For $r \in \mathbb{N}_{\geq 5}$, define

$$\mathrm{F}(r) := 2\lfloor \frac{1}{2} \cdot \left(\min_{d+k+1=r} \Phi(d,k) \right) \rfloor + 1. \tag{5.1}$$

For $r \leq 4$, define $\mathrm{F}(r) = r + 1$.

Lemma 5.5. *For all $r \in \mathbb{N}$, $\mathrm{F}(r+1) > \mathrm{F}(r)$, i.e. F is monotone increasing.*

Proof. The maximum of two monotone increasing functions is monotone increasing, as is any linear combination with positive integer coefficients of the integer part of a monotone increasing function. □

We can now state our main theorem.

Theorem 5.6. *Let $\mathrm{F}\colon \mathbb{N} \longrightarrow \mathbb{N}$ be the monotone increasing function (5.1). For all $n \geq \mathrm{F}(r)$,*

$$\mathrm{RD}(n) \leq n - r.$$

Example 5.7. Observe that

$$\begin{aligned} \mathrm{F}(5) = \Phi(3,1) &= \max\{\frac{4!}{3!} + 1, \dim(\mathcal{M}_{3,3}) + 5\} \\ &= \max\{5, 9\} = 9. \end{aligned}$$

The theorem thus asserts that for $n \geq 9$, $\mathrm{RD}(n) \leq n - 5$, as first stated by Hilbert.

We can compare the upper bounds of Theorem 5.6 to Brauer's bounds as follows. Both the previous theorem and Brauer's theorem prove the existence, for each r, of an explicit cut-off (for n) after which $\mathrm{RD}(n) \leq n - r$. More precisely, define

$$B(r) := (r-1) + 1!$$

Brauer proved [Br75, Theorem 1] that for $n \geq B(r)$,

$$\mathrm{RD}(n) \leq n - r.$$

The cut-off functions $B(r)$ and $\mathrm{F}(r)$ are related as follows.

Theorem 5.8. *There exists a monotone indecreasing function $\varphi\colon \mathbb{N} \longrightarrow \mathbb{N}$, such that for $r \geq \varphi(d)$,*

$$B(r)/\,\mathrm{F}(r) \geq d!$$

In particular, in the limit, the ratio of Brauer's cut-off to that of Theorem 5.6 grows arbitrarily.

Remark 5.9.

1. In Appendix A, we give explicit computations of F(r) for r up to 15 (at which point F(r) is approximately 3.6 billion).
2. We do not expect that the upper bounds of Theorem 5.6 are themselves sharp for two reasons: first, we expect that further optimizizations to the present method should be possible; and second, we have not made contact in this paper with the methods introduced by Sylvester and Hammond [Sy1887, SH1887, SH1888] in their study of Hamilton's work [Ha1836].

It remains to prove Theorems 5.6 and 5.8.

5.1 Proof of Theorem 5.6

Our proof follows the strategy outlined by Hilbert [Hi27]. We recall a classical lemma on quadrics.

Lemma 5.10. *Let K be a field of characteristic 0, let $K \subset \overline{K}$ be an algebraic closure, and let $K^{2\text{-}solv} \subset \overline{K}$ denote the fixed field of the 2-Sylow in the profinite group* $\mathrm{Gal}(\overline{K}/K)$*. For any smooth quadric Q over K, with maximal isotropic Grassmannian $Gr(Q)$, the inclusion*

$$Gr(Q)(K^{2\text{-}solv}) \subset Gr(Q)(\overline{K})$$

is Zariski dense. Moreover, for any $x \in Gr(Q)(K^{2\text{-}solv})$, the associated Severi-Brauer variety over $K^{2\text{-}solv}$ is trivial.

Proof. The proof is classical, and goes back at least to work of Sylvester. Recall that by completing the squares, every nonsingular, definite quadratic form Q over K admits a K-rational change of coordinates to one of the form

$$Q'(x_1, \ldots, x_n) = a_1 x_1^2 + \cdots + a_n x_n^2 \tag{5.2}$$

for $a_i \in K^{\times}$. For example, see [Fo36] for explicit formulas for the a_i in terms of minors of the matrix associated to the quadratic form (n.b. Fort states the results for real definite forms, but the method holds over any base field).

Let $L = K(\sqrt{a_1}, \ldots, \sqrt{a_2}) \subset K^{2\text{-solv}}$. The L-rational change of coordinates

$$x_i =: \frac{y_i}{\sqrt{a_i}}$$

converts the above quadratic form (5.2) to

$$Q''(y_1, \ldots, y_n) = y_1^2 + \cdots + y_n^2.$$

Finally, let $L' = L(\sqrt{-1}) \subset K^{2\text{-solv}}$. Then the quadratic form Q'' vanishes identically on the linear subspace Λ defined by

$$y_{2i-1} = \sqrt{-1} y_{2i}$$

for $i = 1, \ldots, \frac{n}{2}$. Counting the dimension, Λ is a maximal isotropic, i.e.

$$\Lambda \in Gr(Q)(L') \subset Gr(Q)(K^{2\text{-solv}}).$$

Using that $Gr(Q)$ is a homogeneous space for the algebraic group $O(Q)$, and that K (and thus L') is an infinite field, we conclude that the $O(Q)(L')$ orbit of Λ is dense in $Gr(Q)(\bar{K})$ as claimed. Finally, because Λ has an L' point (e.g. for n even $[y_1 : \cdots : y_n] = [\sqrt{-1} : 1 : \cdots : \sqrt{-1} : 1]$, with the analogous formula if n is odd), the Severi-Brauer variety associated to Λ over L' splits completely. We conclude the same for every point in the $\mathrm{O}(Q)(L')$ orbit of Λ. □

Corollary 5.11. *Let X be a variety over a field K of characteristic 0. For any generically smooth family of quadrics $Q \longrightarrow X$, the solvable multi-sections of $Gr(Q) \longrightarrow X$ are Zariski dense in $Gr(Q)(K(\bar{X}))$.*

Proof of Theorem 5.6. Because F is a monotone increasing function (by Lemma 5.5), if $n \geq \mathrm{F}(r)$, then $n-1 \geq \mathrm{F}(r-1)$. We can therefore induct on r.

The base cases $r \leq 4$ are classical: for $n \leq 4$, solutions in radicals imply $\mathrm{RD}(n) = 1$. That $\mathrm{RD}(n) \leq n-4$ for $n \geq 5$ follows from Bring [Br1786] and [Ha1836].

For the induction step, assume that we have shown that for all $s < r$, $n \geq \mathrm{F}(s)$ implies that $\mathrm{RD}(n) \leq n - s$. Let $n \geq \mathrm{F}(r)$. Note that if $\min_{d+k+1=r} \Phi(d,k)$ is odd, then the definition of F implies that

$$\mathrm{F}(r) = \min_{d+k+1=r} \Phi(d,k).$$

Conversely, if $\min_{d+k+1=r} \Phi(d,k)$ is even, then

$$\mathrm{F}(r) = \min_{d+k+1=r} \Phi(d,k) + 1.$$

Consequently, if n is odd, then

$$n \geq \min_{d+k+1=r} \Phi(d,k),$$

while if n is even

$$n \geq \min_{d+k+1=r} \Phi(d,k) + 1.$$

Let (d,k) be such that

$$\Phi(d,k) = \min_{d'+k'+1=r} \Phi(d',k').$$

If n is odd (and thus $n \geq \Phi(d,k)$), we will explicitly construct a rational multi-section

$$U \longrightarrow T_{1\cdots d+k}$$

for $T_{1\cdots d+k} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ with

$$\mathrm{RD}(U \longrightarrow \mathbb{A}^n_{\mathbf{a}}) \leq \max\{\mathrm{RD}(\frac{(d+k)!}{d!}), \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})\}.$$

If n is even (and thus $n \geq \Phi(d,k)+1$), *mutatis mutandis* the same argument will produce a rational multi-section

$$U \longrightarrow T'_{1\cdots d+k}$$

with $\mathrm{RD}(U \longrightarrow \mathbb{A}^n_{\mathbf{a}}) \leq\leq \max\{\mathrm{RD}(\frac{(d+k)!}{d!}), \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})\}$.

Case 1: n odd. Let $U_1 = \mathbb{A}^n_{\mathbf{a}}$. By Lemma 2.6, the family $T_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ is generically smooth. By Corollary 2.14, there exists a dense open $V \subset Gr(T_{12})$, such that

$$\mathcal{L}|_V \times_{\mathbb{P}^n_{\mathbb{A}^n_{\mathbf{a}}}} T_{123} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is smooth (i.e. for the generic polynomial, the intersection of $T_{123}(\mathbf{a})$ with a generic maximal isotropic in $T_{12}(\mathbf{a})$ is smooth).

By Corollary 5.11,

$$\mathrm{RD}(V \longrightarrow \mathbb{A}^n_{\mathbf{a}}) = 1$$

More precisely, there exists a multi-section $U_2 \subset V$ such that $U_2 \longrightarrow U_1$ is a solvable cover of its image, and such that

$$\mathcal{L}|_{U_2} \cong \mathbb{P}^{\frac{n-3}{2}}_{U_2}.$$

Now, by Lemma 5.3 and our assumption on n,

$$\begin{aligned} n \geq \Phi(d,k) &\geq \psi(d,k)_{d-1} + 2 \\ &= 2\psi(d,k)_{d-2} + 3. \end{aligned}$$

Therefore,

$$\frac{n-3}{2} \geq \psi(d,k)_{d-2}$$

If $\frac{n-3}{2} = \psi(d,k)_{d-2}$, then we obtain a map

$$\begin{aligned} U_2 &\longrightarrow \mathcal{M}_{3,\psi(d,k)_{d-2}} \\ x &\mapsto \mathcal{L}|_x \times_{\mathbb{P}^{n-2}} T_{123}|_x. \end{aligned}$$

If $\frac{n-3}{2} > \psi(d,k)_{d-2}$, by the Bertini Theorem for isotropics (Proposition 2.13), there exists a dense open

$$V' \subset Gr(\psi(d,k)_{d-2}, \mathcal{L}|_{U_2})$$

such that the family of cubic hypersurfaces in $\mathbb{P}^{\psi(d,k)_{d-2}}$ given by

$$V' \times_{\mathcal{L}|_{U_2}} (T_{123} \times_{\mathbb{P}^{n-2}_{U_1}} \mathcal{L}|_{U_2}) \longrightarrow U_2$$

is generically smooth. Because rational points are dense in Grassmannians, perhaps after shrinking U_2, we obtain a section $U_2 \longrightarrow V'$. As above, we again obtain a map

$$U_2 \xrightarrow{\cap T_{123}} \mathcal{M}_{3,\psi(d,k)_{d-2}}.$$

Note that, from the construction above, $\mathrm{RD}(U_2 \longrightarrow U_1) = 1$.

By Waldron's Theorem (Theorem 5.1) and the definition of the numbers $\psi(d,k)_i$, the map

$$\mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}} \longrightarrow \mathcal{M}_{3,\psi(d,k)_{d-2}}$$

is surjective. Therefore, the map

$$\mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}}|_{U_2} \longrightarrow U_2$$

is surjective, and by Lemma 4.8,

$$\mathrm{RD}(\mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}}|_{U_2} \longrightarrow U_2) \leq \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}).$$

Let $U' \subset \mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}}|_{U_2}$ be any rational multi-section such that

$$\mathrm{RD}(U' \longrightarrow U_2) = \mathrm{RD}(\mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}}|_{U_2} \longrightarrow U_2).$$

Let $\bar{\mathcal{L}} \longrightarrow \mathcal{M}^{\psi(d,k)_{d-3}}_{3,\psi(d,k)_{d-2}}$ denote the tautological $\psi(d,k)_{d-3}$-plane bundle. By the Merkurjev-Suslin Theorem [MS83, Theorem 16.1], there exists a solvable étale map $U_3 \longrightarrow U'$ such that

$$\bar{\mathcal{L}}|_{U_3} \cong \mathbb{P}^{\psi(d,k)_{d-3}}_{U_3}.$$

By Lemma 4.10 and the construction above,

$$\mathrm{RD}(U_3 \longrightarrow U_2) = \max\{\mathrm{RD}(U' \longrightarrow U_2), 1\} \leq \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}).$$

Further, intersecting with the Tschirnhaus hypersurface T_4, we obtain a map

$$\begin{aligned} U_3 &\xrightarrow{\cap T_4} \mathcal{H}_{4,\psi(d,k)_{d-3}} \\ x &\mapsto (T_{123}|_x \times_{U_3} \bar{\mathcal{L}}|_{U_3}) \times_{\mathbb{P}^{n-1}_{U_3}} T_4|_{U_3}. \end{aligned}$$

By induction, we now construct, for each $4 \leq i \leq d$, a quasi-finite dominant map

$$U_i \longrightarrow U_{i-1}$$

such that

1. $\mathrm{RD}(U_i \longrightarrow U_{i-1}) \leq \dim(\mathcal{H}_{i,\psi(d,k)_{d-i+1}})$,
2. we have a commuting diagram

$$\begin{array}{ccc} U_i & \longrightarrow & \mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}} \\ \downarrow & & \downarrow \\ U_{i-1} & \xrightarrow{\cap T_i} & \mathcal{H}_{i,\psi(d,k)_{d-i+1}} \end{array}$$

   with a trivialization

$$\mathcal{L}|_{U_i} \cong \mathbb{P}^{\psi(d,k)_{d-i}}_{U_i},$$

   where $\mathcal{L} \longrightarrow \mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}$ denotes the tautological $\psi(d,k)_{d-i}$-plane bundle;

3. and the assignment

$$x \mapsto (T_{1\cdots i}|_x \times_{U_i} \mathcal{L}_{i,\psi(d,k)_{d-i+1}}|_{U_i}) \times_{\mathbb{P}^{n-1}_{U_i}} T_{i+1}|_x$$

   defines a map

$$U_i \xrightarrow{\cap T_{i+1}} \mathcal{H}_{i+1,\psi(d,k)_{d-i}}.$$

The construction proceeds along the same lines as the construction of U_3 above. Given U_{i-1} with the map

$$U_{i-1} \xrightarrow{\cap T_i} \mathcal{H}_{i,\psi(d,k)_{d-i+1}},$$

by the definition of the $\psi(d,k)_j$s and Waldron's Theorem (Theorem 5.1), the map

$$\mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}} \longrightarrow \mathcal{H}_{i,\psi(d,k)_{d-i+1}}$$

is surjective. Therefore, the map

$$\mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}|_{U_{i-1}} \longrightarrow U_{i-1}$$

is surjective, and by Lemma 4.8,

$$\mathrm{RD}(\mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}|_{U_{i-1}} \longrightarrow U_{i-1}) \leq \dim(\mathcal{H}_{i,\psi(d,k)_{d-i+1}}).$$

Let $U' \subset \mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}|_{U_{i-1}}$ be any rational multi-section such that

$$\mathrm{RD}(U' \longrightarrow U_{i-1}) = \mathrm{RD}(\mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}|_{U_{i-1}} \longrightarrow U_{i-1}).$$

Let $\mathcal{L} \longrightarrow \mathcal{H}^{\psi(d,k)_{d-i}}_{i,\psi(d,k)_{d-i+1}}$ denote the tautological $\psi(d,k)_{d-i}$-plane bundle. By the Merkurjev-Suslin Theorem [MS83, Theorem 16.1], there exists a solvable étale map $U_i \longrightarrow U'$ such that

$$\mathcal{L}|_{U_i} \cong \mathbb{P}^{\psi(d,k)_{d-i}}_{U_i}.$$

By Lemma 4.10 and the construction above,

$$\mathrm{RD}(U_i \longrightarrow U_{i-1}) = \max\{\mathrm{RD}(U' \longrightarrow U_{i-1}), 1\} \leq \dim(\mathcal{H}_{i,\psi(d,k)_{d-i+1}}).$$

Finally, to complete the induction step, we observe that, by intersecting with the Tschirnhaus hypersurface T_{i+1}, we obtain a map

$$U_i \xrightarrow{\cap T_{i+1}} \mathcal{H}_{i+1,\psi(d,k)_{d-i}}$$
$$x \mapsto (T_{1\cdots i}|_x \times_{U_i} \mathcal{L}|_{U_i}) \times_{\mathbb{P}^{n-1}_{U_i}} T_{i+1}|_{U_i}.$$

This completes the induction step. We have thus constructed a tower of maps

$$U_d \longrightarrow \cdots U_4 \longrightarrow U_3 \longrightarrow U_2 \longrightarrow U_1 = \mathbb{A}^n_{\mathbf{a}}.$$

Further, from the inductive construction and Lemmas 4.10 and 5.3, we have

$$\mathrm{RD}(U_d \longrightarrow \mathbb{A}^n_{\mathbf{a}}) \leq \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}).$$

Further, let $\mathcal{L} \longrightarrow \mathcal{H}_{d,\psi(d,k)_1}$ denote the tautological k-plane bundle (n.b. $k = \psi(d,k)_0$). Then, by construction, we have an isomorphism

$$\mathcal{L}|_{U_d} \cong \mathbb{P}^k_{U_d}.$$

For $i_1 < \ldots < i_k$, and N, let

$$\mathcal{H}_{i_1 \cdots i_k, N}$$

denote the parameter space of complete intersections of degree $(i_1, \ldots, i_k)$. Let

$$\mathcal{I} \longrightarrow \mathcal{H}_{i_1 \cdots i_k, N}$$

denote the tautological family of complete intersections. By Proposition 4.11,

$$\mathrm{RD}(\mathcal{I} \longrightarrow \mathcal{H}_{i_1 \cdots i_k, N}) \leq \mathrm{RD}(i_1 \cdots i_k).$$

By our inductive construction, we have a map

$$\begin{aligned} U_d &\xrightarrow{\cap T_{(d+1)\cdots(d+k)}} \mathcal{H}_{(d+1)\cdots(d+k),k} \\ x &\mapsto (T_{1\cdots d}|_x \times_{U_d} \mathcal{L}|_{U_d}) \times_{\mathbb{P}^{n-1}_{U_d}} T_{(d+1)\cdots(d+k)}|_{U_d}. \end{aligned}$$

Because, $\mathcal{I} \longrightarrow \mathcal{H}_{(d+1)\cdots(d+k),k}$ is surjective, by Lemma 4.8,

$$\mathrm{RD}(\mathcal{I}|_{U_d} \longrightarrow U_d) \leq \mathrm{RD}(\frac{(d+k)!}{d!}).$$

Let $U_{d+1} \subset \mathcal{I}|_{U_d}$ be a rational multi-section of $\mathcal{I}|_{U_d} \longrightarrow U_d$ such that

$$\mathrm{RD}(U_{d+1} \longrightarrow U_d) \leq \mathrm{RD}(\frac{(d+k)!}{d!}).$$

Then, by construction, U_{d+1} carries a canonical map

$$U_{d+1} \longrightarrow T_{1\cdots(d+k)}$$

making it a rational multi-section of the Tschirnhaus complete intersection. Further, by the above construction and Lemma 4.10,

$$\mathrm{RD}(U_{d+1} \longrightarrow \mathbb{A}^n_{\mathbf{a}}) \leq \max\{\mathrm{RD}(\frac{(d+k)!}{d!}), \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})\}.$$

By assumption, $n \geq \mathrm{F}(r) = \Phi(d,k) \geq \frac{(d+k)!}{d!} + 1$. Lemma 5.5 thus implies that $\frac{(d+k)!}{d!} \geq \mathrm{F}(r-1)$. Therefore, by the inductive hypothesis,

$$\mathrm{RD}(\frac{(d+k)!}{d!}) \leq \frac{(d+k)!}{d!} - (r-1).$$

Moreover, from the definition of $\Phi(d,k)$, $n \geq \Phi(d,k)$ implies that $n \geq \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})+r$.

By Lemma 4.12, we therefore conclude that

$$\begin{aligned}\mathrm{RD}(n) &\leq \max\{\mathrm{RD}(U_{d+1} \longrightarrow \mathbb{A}^n_{\mathbf{a}}, \dim(\mathrm{ev}(T_{1\cdots(d+k)})) - 1\}\\ &\leq \max\{\frac{(d+k)!}{d!} - (r-1), \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}), n-r.\}\\ &= n-r.\end{aligned}$$

Case 2: n **even.** Let $U_1 = \mathbb{A}^n_{\mathbf{a}}$. By Lemma 2.6, the family $T'_{12} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$ is generically smooth. By Corollary 2.14, there exists a dense open $V \subset Gr(T'_{12})$, such that

$$\mathcal{L}|_V \times_{\mathbb{P}^n_{\mathbb{A}^n_{\mathbf{a}}}} T'_{123} \longrightarrow \mathbb{A}^n_{\mathbf{a}}$$

is smooth (i.e. for the generic polynomial, the intersection of $T'_{123}(\mathbf{a})$ with a generic maximal isotropic in $T'_{12}(\mathbf{a})$ is smooth).

By Corollary 5.11,

$$\mathrm{RD}(V \longrightarrow \mathbb{A}^n_{\mathbf{a}}) = 1$$

More precisely, there exists a multi-section $U_2 \subset V$ such that $U_2 \longrightarrow U_1$ is a solvable cover of its image, and such that

$$\mathcal{L}|_{U_2} \cong \mathbb{P}^{\frac{n}{2}-2}_{U_2}.$$

Now, by Lemma 5.3 and our assumption on n

$$\begin{aligned}n-1 \geq \Phi(d,k) &\geq \psi(d,k)_{d-1} + 2\\ &= 2\psi(d,k)_{d-2} + 3.\end{aligned}$$

Therefore,

$$\frac{n}{2} - 2 \geq \psi(d,k)_{d-2}$$

If $\frac{n}{2} - 2 = \psi(d,k)_{d-2}$, then we obtain a map

$$\begin{aligned}U_2 &\longrightarrow \mathcal{M}_{3,\psi(d,k)_{d-2}}\\ x &\mapsto \mathcal{L}|_x \times_{\mathbb{P}^{n-2}} T'_{123}|_x.\end{aligned}$$

If $\frac{n}{2} - 2 > \psi(d,k)_{d-2}$, by the Bertini Theorem for isotropics (Proposition 2.13), there exists a dense open

$$V' \subset Gr(\psi(d,k)_{d-2}, \mathcal{L}|_{U_2})$$

such that the family of cubic hypersurfaces in $\mathbb{P}^{\psi(d,k)_{d-2}}$ given by

$$V' \times_{\mathcal{L}|_{U_2}} (T'_{123} \times_{\mathbb{P}^{n-2}_{U_1}} \mathcal{L}|_{U_2}) \longrightarrow U_2$$

is generically smooth. Because rational points are dense in Grassmannians, perhaps after shrinking U_2, we obtain a section $U_2 \longrightarrow V'$. As above, we again obtain a map

$$U_2 \xrightarrow{\cap T'_{123}} \mathcal{M}_{3,\psi(d,k)_{d-2}}.$$

Note that, from the construction above, $\mathrm{RD}(U_2 \longrightarrow U_1) = 1$. The remainder of the proof now proceeds exactly as in the case of n odd. $\square$

5.2 Proof of Theorem 5.8

Proof of Theorem 5.8. We deduce the theorem from the following:

Claim 1. There exists a monotone increasing function $\rho\colon \mathbb{N} \longrightarrow \mathbb{N}$ such that

1. for $k \geq \rho(d)$,

$$\begin{aligned} \frac{(d+k)!}{d!} + 1 &= \Phi(d,k) \\ &\leq \Phi(d-1,k+1) \end{aligned}$$

   (i.e. both conditions hold for $k \geq \rho(d)$);

2. for all $k < \rho(d)$,

$$\Phi(d,k) > \Phi(d-1,k+1).$$

   (i.e. $\rho(d)$ is the least integer such that the inequality holds).

Granting the claim, let $\varphi(d) := \rho(d+1) + d + 2$. Then for $r \geq \varphi(d)$, we have

$$\begin{aligned} k &:= (r-1) - (d+1) \\ &\geq \varphi(d) - (d+2) \\ &\geq \rho(d+1) \end{aligned}$$

As a result,

$$\mathrm{F}(r) = \frac{(r-1)!}{(d+1)!} + 1$$

and therefore,

$$\begin{aligned} B(r)/\,\mathrm{F}(r) &= \frac{(r-1)!+1}{(r-1)!/(d+1)!+1} \\ &\geq d! \end{aligned}$$

We now prove Claim 1 by asymptotic estimates; more precisely, we show that for each d, $\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})$ grows polynomially in k, while $\frac{(d+k)!}{d!}$ grows superexponentially. Precise formulas for the function ρ require a more detailed analysis.

Continuing to follow Notation 5.2, we claim the following:

Claim 2. Fix d. Then as a function of k,

$$\mathcal{O}((d+k)!) \geq \max\{\mathcal{O}(\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})), \mathcal{O}(\dim(\mathcal{M}_{3,\psi(d-1,k+1)_{d-3}}))\},$$

where $\mathcal{O}(f)$ denotes the asymptotic growth of a function f.

Granting the claim, we see that for $k >> d$,

$$\Phi(d,k) = \frac{(d+k)!}{d!} < \frac{(d+k)!}{(d-1)!} = \Phi(d-1,k+1).$$

Note that by definition,

$$\Phi(d,k) = \max\{\frac{(d+k)!}{d!} + 1, \dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) + d + k + 1\}$$

Therefore Claim 1 follows from Claim 2. To prove Claim 2, recall Stirling's formula (cf. [Ro55])

$$\sqrt{2\pi} m^{m+\frac{1}{2}} e^{\frac{1}{12m+1} - m} \leq m! \leq \sqrt{2\pi} m^{m+\frac{1}{2}} e^{\frac{1}{12m} - m}$$

This implies that

$$\mathcal{O}(\ln((d+k)!) = \mathcal{O}((d+k+\frac{1}{2})\ln(d+k)).$$

It suffices to prove that

$$\max\{\mathcal{O}(\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})), \mathcal{O}(\dim(\mathcal{M}_{3,\psi(d-1,k+1)_{d-3}}))\} = \mathcal{O}(k^{\alpha_d})$$

for some α_d, as then

$$\begin{aligned} \max\{\mathcal{O}(\ln(\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}))), \mathcal{O}(\ln(\dim(\mathcal{M}_{3,\psi(d-1,k+1)_{d-3}})))\} &= \mathcal{O}(\alpha_d \cdot \ln(k)) \\ &\leq \mathcal{O}((d+k+\frac{1}{2})\ln(d+k)) \\ &= \mathcal{O}(\ln(d+k)!). \end{aligned}$$

Recall that $\psi(d,k)_0 = k$ and for $i > 0$,

$$\psi(d,k)_i = \lceil \psi(d,k)_{i-1} + \binom{\psi(d,k)_{i-1} + d - (i-1)}{\psi(d,k)_{i-1}} / (\psi(d,k)_{i-1} + 1) \rceil.$$

Therefore

$$\psi(d,k)_i \sim \frac{(d-i+1+\psi(d,k)_{i-1}) \cdots (\psi(d,k)_{i-1}+2)}{(d-i+1)!} \sim (\psi(d,k)_{i-1})^{d-i}.$$

Because $\psi(d,k)_1 \sim k^{d-1}$, by induction, we obtain

$$\begin{aligned} \dim(\mathcal{H}_{d-i,\psi(d,k)_{i+1}}) &= \binom{d-i+\psi(d,k)_{i+1}}{\psi(d,k)_{i+1}} - 1 \\ &\sim \psi(d,k)_{i+1}^{d-i} \\ &\sim k^{(d-i)\frac{(d-1)!}{(d-i-2)!}}. \end{aligned}$$

Similarly,

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) \sim k^{3(d-1)!}.$$

By the same argument,

$$\dim(\mathcal{M}_{3,\psi(d-1,k+1)_{d-3}}) \sim (k+1)^{3(d-2)!} \sim k^{3(d-2)!},$$

and, thus, as functions of k,

$$\mathcal{O}((d+k)!) \geq \max\{\mathcal{O}(\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}})), \mathcal{O}(\dim(\mathcal{M}_{3,\psi(d-1,k+1)_{d-3}}))\}$$

as claimed. □

A Explicit Bounds

Table 1: Upper Bounds on RD(n)

r	F(r)	$B(r)$	$B(r)/$ F(r)	Subspace on Cubic	Subspace on Quartic
2	3				
3	4				
4	5				
5	9	25	2.78	1	
6	41	121	2.95	2	
7	121	721	5.96	3	
8	841	5041	5.99	4	
9	6721	40321	5.99	5	
10	60481	362881	5.99	6	
11	604801	3628801	5.99	7	
12	6652801	39916801	5.99	8	
13	78485043	$12!+1$	6.10	63	8
14	320082459	13!+1	19.45	81	9
15	3632428801	14!+1	24	101	10

References

[AS76] V. Arnold and G. Shimura, Superpositions of algebraic functions, *Proc. Symposia in Pure Math.* vol. 28 (1976), AMS, Providence, 45–46.

[Be12] A. Beauville, Non-rationality of the symmetric sextic Fano threefold, *Geometry and arithmetic*, 57–60, EMS Series of Congress Reports, Eur. Math. Soc., Zrich, 2012.

[Br45] R. Brauer, A note on systems of homogeneous algebraic equations, *Bull.* AMS, vol. 51 (1945), 749–755.

[Br75] R. Brauer, On the resolvent problem, *Ann. Mat. Pura Appl.* (4) 102 (1975), 45–55.

[Br1786] E. Bring, Meletemata quædam Mathematica circa Transformationem Æquationum Algebraicarum ("Some Selected Mathematics on the Transformation of Algebraic Equations"), Lund, 1786.

[Di93] J. Dixmier, Histoire de 13e problème de Hilbert, *Cahiers du sminaire dhistoire des mathématiques*, 2e série, tome 3 (1993), 85–94.

[FW18] B. Farb and J. Wolfson, Resolvent degree, Hilbert's 13th problem and Geometry, arXiv:1803.04063.

[Fo36] T. Fort, Formulas for Reducing a Quadratic Form to a Sum of Squares, *American Mathematical Monthly*, vol. 43, no. 8, 1936, 477–481.

[Ha1836] W. Hamilton, Inquiry into the validity of a method recently proposed by George B. Jerrard, esq., for transforming and resolving equations of elevated degrees, *Report of the Sixth Meeting of the British Association for the Advancement of Science* (1836), Bristol, 295-348.

[Ha77] R. Hartshorne, Algebraic Geometry, GTM 52, Springer-Verlag, New York, NY, 1977, xvi+496 pp.

[Hi1900] D. Hilbert, Mathematical Problems, from *Proceedings of the* 1900 ICM, English translation reprinted in *Bull.* AMS, Vol. 37, No. 4 (2000), 407–436.

[Hi27] D. Hilbert, Über die Gleichung neunten Grades, *Math. Ann.* vol. 97 (1927), no. 1, 243–250.

[Hi27t] D. Hilbert, On the equation of ninth degree, *Math. Ann.* vol. 97 (1927), no. 1, 243–250. English translation by S. Hensel, 2017, available at https://www.mathematik.uni-muenchen.de/ hensel/papers/hilbert2.pdf.

[HL87] M. Hochster and D. Laksov, The linear syzygies of generic forms, *Comm. Algebra* vol. 15 (1987), no. 1-2, 227–239.

[Kl74] S. Kleiman, The transversality of a general translate, *Compositio Math.* vol. 28 (1974), no. 3, 287–297.

[Kl1884] F. Klein, Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen vom fünften Grade, Teubner, Leipzig, 1884. English translation: Lectures on the icosahedron and solution of equations of the fifth degree, translated by G. G. Morrice, 2nd and rev. edition, New York, Dover Publications, 1956.

[MS83] A. Merkurjev and A. Suslin, K-Cohomology of Severi-Brauer Varieties and the Norm Residue Homomorphism, *Math. USSR Izv.* vol. 21 (1983), no. 2, 307–340.

[Ro55] H. Robbins, A Remark on Stirling's Formula, *The American Mathematical Monthly*, vol. 62 (1955) no. 1: pp. 26-29.

[Se45] B. Segre, The Algebraic Equations of Degrees 5, 9, 157, ..., and the Arithmetic Upon an Algebraic Variety, *Ann. of Math.* (2), vol. 46 (1945), 287–301.

[Se51] B. Segre, *Arithmetical Questions on Algebraic Varieties*, University of London, Athlone Press, London, 1951.

[St17] J.M. Starr, Veronese varieties contained in hypersurfaces, arXiv:1703.03294.

[Sy1887] J. Sylvester, On the so-called Tschirnhausen transformation, *J. Reine Angew. Math.* vol. 100 (1887), 465–486.

[SH1887] J. Sylvester and J. Hammond, On Hamilton's Numbers, *Phil. Trans. R. Soc. London A* vol. 178 (1887), 285–312.

[SH1888] J. Sylvester and J. Hammond, On Hamilton's Numbers II, *Phil. Trans. R. Soc. London A* vol. 179 (1888), 65–71.

[Ts1683] E.W. von Tschirnhaus, Methodus auferendi omnes terminos intermedios ex data aeqvatione (Method of eliminating all intermediate terms from a given equation), *Acta Eruditorum* (1683), 204–207.

[Wa08] A. Waldron, Fano Varieties of Low-Degree Smooth Hypersurfaces and Unirationality, Bachelor thesis, Harvard University, Cambridge, Massachusetts, 2008.

[Wi27] A. Wiman, Über die Anwendung der Tschirnhausen-Transformation auf die Reduktion algebraischer Gleichungen, *Nova acta regiae societatis scientiarum Uppsaliensis*, vol. 16 (1927), 3–8.

Dept. of Mathematics, University of California, Irvine
E-mail: wolfson@uci.edu