L'Enseignement Mathématique (2) 66 (2020), 489–540 DOI 10.4171/LEM/66-3/4-9

Tschirnhaus transformations after Hilbert

Jesse Wolfson

Abstract. In this paper, we use enumerative geometry to simplify the formula for the roots of the general one-variable polynomial of degree n, for all n. More precisely, let $\mathrm{RD}(n)$ 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 $\mathrm{RD}(9) \leq 4$). In this paper, we turn Hilbert's sketch into a general method. We show this method produces best-to-date upper bounds on $\mathrm{RD}(n)$ for all n, improving earlier results of Hamilton, Sylvester, Segre and Brauer.

Mathematics Subject Classification (2020). Primary: 14G25; Secondary: 11C08, 12E12.

Keywords. Resolvent degree, polynomials, lines on cubic surfaces.

1. Introduction

The goal of this paper is to use enumerative geometry to produce simplest-to-date formulas for the roots of the general one-variable polynomial of degree n, for all n. 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 $\mathrm{RD}(n)$ (see Section 4 for a precise definition, and [FW] for a detailed treatment). At present, no nontrivial lower bounds for $\mathrm{RD}(n)$ are known. The best general upper bounds in the literature are due to Brauer [Bra2], who uses methods dating to Tschirnhaus [Tsch] to prove

that $\mathrm{RD}(n) \leq n - r$ for $n \geq (r-1)! + 1$. As Brauer remarks, his bounds are not optimal for small r.[1]

In this paper we take a different approach to bounding $\mathrm{RD}(n)$, inspired by a geometric argument of Hilbert. In [Hil2], 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., $\mathrm{RD}(9) \leq 4$. 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. This defines an explicit increasing function $\mathrm{F}\colon \mathbb{N} \to \mathbb{N}$ (Definition 5.4) for which we prove the following:

Theorem 1.1. *Let* $\mathrm{F}\colon \mathbb{N} \to \mathbb{N}$ *be the function defined in Definition 5.4.*

(1) *For all* r *and all* $n \geq \mathrm{F}(r)$, $\mathrm{RD}(n) \leq n - r$.

(2) *For all* r, $n = \mathrm{F}(r)$ *is the least value for which we know* $\mathrm{RD}(n) \leq n - r$ *to hold.*[2] *In particular, the initial values are given by*

r	1	2	3	4	5	6	7
$\mathrm{F}(r)$	2	3	4	5	9	41	121

(3) *Writing* $B(r) = (r-1)! + 1$ *for Brauer's bound, then*

$$\lim_{r\to\infty} B(r)/\mathrm{F}(r) = \infty.$$

The first statement appears as Theorem 5.6 below, while the last two appear as Theorem 5.8.

Remark 1.2. (1) The construction of F, the proof that $\mathrm{F}(5) = 9$ and that this implies $\mathrm{RD}(9) \leq 4$ marks the first rigorous construction of the 4-variable formula for the general degree 9 sketched by Hilbert in [Hil2].[3]

(2) The first improvement over prior bounds occurs at $\mathrm{F}(6) = 41$. Previously, Sylvester proved [Syl, p. 485] that for $n \geq 44$, $\mathrm{RD}(n) \leq n - 6$.

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 [Dix, p. 90][4]:

[1] Brauer's first improvement over prior bounds occurs for $r = 7$.

[2] i.e., $n = \mathrm{F}(r)$ is the least value for which $\mathrm{RD}(n) \leq n - r$ is currently proven to hold in any of the literature of which we are aware. Note that G. Chebotarev [Cheb] claimed to have extended an argument of Wiman [Wim] to conclude $\mathrm{RD}(n) \leq n - 6$ for $n \geq 21$. His proof has gaps similar to those observed by Dixmier [Dix] in the arguments of Hilbert and Wiman, namely he takes for granted that certain forms are generic, when they are not.

[3] Rigorous 4-variable formulas have been previously constructed by Segre [Seg1] and Dixmier [Dix].

[4] n.b. Dixmier writes "$s(n)$" for our $\mathrm{RD}(n)$.

> "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 [Seg2, III.5]:

Problem 1.3. 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.

Remarks on the Proof. Given a polynomial

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_n = \prod_{i=1}^{n} (z - z_i),$$

a *Tschirnhaus transformation* is a "change of variables"

$$y = \sum_{j=0}^{n-1} b_j z^j.$$

This gives a new polynomial

$$q(y) = \prod_i (y - \sum_{j=0}^{n-1} b_j z_i^j) = y^n + c_1 y^{n-1} + \cdots + c_n,$$

and we can ask for Tschirnhaus transformations which normalize the resulting polynomial so that, e.g.

$$c_1 = \cdots = c_k = 0. \tag{1.1}$$

The space of all $(b_0, \ldots, b_{n-1})$ such that the conditions (1.1) are satisfied forms an affine cone, and the projectivization gives a complete intersection

$$T^n_{1\cdots k} \subset \mathbb{P}^{n-1};$$

when the superscript n is clear from context, we suppress it and write $T_{1\cdots k}$. If we can find a point of $T_{1\cdots k}$ over a convenient extension of $\mathbb{C}(a_1, \ldots, a_n)$, e.g. one defined using only algebraic functions of at most d variables, then we can write a formula for the general degree n polynomial using only functions of at most d variables and the algebraic function

$$(c_{k+1}, \ldots, c_n) \mapsto \{y \mid y^n + c_{k+1} y^{n-k-1} + \cdots + c_n = 0\},$$

This, together with a final rational change of coordinates, gives an upper bound

$$\mathrm{RD}(n) \leq \max\{d, n-k-1\}.$$

In [Hil2], Hilbert sketched how to use the 27 lines on a smooth cubic surface to find points on T_{1234} for $n = 9$: Here, $T_1 \subset \mathbb{P}^8$ is a hyperplane, and thus T_{12} is a quadric 6-fold in $T_1 \cong \mathbb{P}^7$. Over a solvable extension $L/\mathbb{C}(a_1, \ldots, a_9)$, every smooth quadric contains a 3-plane P in $\mathbb{P}^7$. The intersection of this 3-plane P with T_{123} is a cubic surface, and this gives a map from $\mathrm{Spec}(L)$ to the moduli of cubic surfaces. Since every smooth cubic surface has 27 lines, and the moduli space of cubic surfaces is 4-dimensional, the algebraic function which assigns a line to a cubic surface is a function of at most 4-variables. Given a line on our cubic surface $P \cap T_{123}$, we can then intersect it with T_{1234} to get a quartic polynomial in one variable, and by adjoining radicals, we can find a point on $T_{1234}(L')$, where $L'/\mathbb{C}(a_1, \ldots, a_9)$ is defined using algebraic functions of at most $d = 4$ variables.

As Dixmier observed [Dix, §8], the argument above is incomplete. In particular, Hilbert takes for granted that the family of cubic surfaces $P \cap T_{123}$ is sufficiently generic. 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 essentially assumes that the above map

$$\mathrm{Spec}(L) \to \mathcal{H}_{3,3}$$

lands in the locus where the rational map

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

is well-defined.[5] 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) is generically smooth; see Theorem 2.12.

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 [MS, Theorem 16.1], which allows us to trivialize the Severi–Brauer varieties which arise in Hilbert's argument by adjoining radicals.[6] The second is a theorem of Hochster–Laksov [HL] which allowed Waldron [Wal, Theorem 1.6] (see also [Sta, 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. Given these, we can generalize Hilbert's sketch to explicitly construct the function F and obtain the bounds on $\mathrm{RD}(n)$ stated above.

[5] n.b. Hilbert actually assumes that the generic member of the family $P \cap T_{123}$ admits a "pentahedral form", but one can weaken this as above without any loss in the argument.

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

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 [FW]. 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. In Appendix B, we review the history of the search for simple formulas for the general degree n polynomial and the summarize the major prior work to date.

Conventions. Throughout the paper, by a *variety* over a field K or over $\mathbb{Z}$, we mean a reduced, separated, not-necessarily irreducible K or $\mathbb{Z}$-scheme. For maps of varieties $X \to Z$ and $Y \to Z$, we will use the notation $Y|_X$ to denote the fiber product $X \times_Z Y$.

2. Tschirnhaus complete intersections

Given a polynomial

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_n = \prod (z - x_i),$$

a *Tschirnhaus transformation* is a "change of variables"

$$y = \sum_{j=0}^{n-1} b_j x^j.$$

This gives a new polynomial

$$q(z) = z^n + c_1 z^{n-1} + \cdots + c_n = \prod_i (z - y_i).$$

We are interested in Tschirnhaus transformations such that $q(z)$ is "better normalized" than $p(z)$, e.g. in the sense that for some i,

$$\sum_j y_j^i = 0,$$

or more generally such that

$$\sum_j y_j^{i_1} = \cdots = \sum_j y_j^{i_k} = 0$$

for some $i_1, \ldots, i_k$. In this section, we study the collection of all $\mathbf{b} = (b_0, \ldots, b_{n-1})$ such that the above normalizations hold. These are affine varieties which we denote $\widetilde{T^n_{i_1 \cdots i_k}}$, and we refer to their projectivizations $T^n_{i_1 \cdots i_k}$ as *Tschirnhaus complete intersections*.

In this section, we introduce the varieties $T^n_{i_1 \cdots i_k}$ as objects of interest in their own right, i.e., via explicit equations. We relate them to classical examples of interest, and study their geometry. In Section 3, we review the classical subject of Tschirnhaus transformations for algebraic functions, and we identify the varieties $T^n_{i_1 \cdots i_k}$ considered here with the spaces of "normalized changes of variables" described above.

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

Notation 2.1. Denote

$$\begin{array}{ll} \mathbf{a} := (a_1, \ldots, a_n) \in \mathbb{A}^n. & |\kappa| := \sum_i k_i \\ \mathbf{b} := [b_0 : \cdots : b_{n-1}] \in \mathbb{P}^{n-1} & ||\kappa|| := \sum_i i \cdot k_i \\ \kappa := (k_0, \ldots, k_{n-1}) \in \mathbb{N}^n & \mathbf{b}^\kappa := \prod_i b_i^{k_i} \end{array}$$

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

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

We also introduce two variants of the above.

Notation 2.2.

$$\begin{aligned} {}'\mathbf{b} &:= [b_1 : \cdots : b_{n-1}] \in \mathbb{P}^{n-2} \\ {}'\kappa &:= (k_1, \ldots, k_{n-1}) \in \mathbb{N}^{n-1} \\ {}'\mathbf{b}' &:= [b_1 : \cdots : b_{n-2}] \in \mathbb{P}^{n-3} \\ {}'\kappa' &:= (k_1, \ldots, k_{n-2}) \in \mathbb{N}^{n-2} \end{aligned}$$

Mutatis mutandis, we will also write $|'\kappa|$, $||'\kappa'||$, $\binom{i}{'\kappa}$, etc. Note that the meaning of $||-||$ depends on whether the first coordinate is the zeroth coordinate or the first coordinate. Our notation indicates that any tuple without a $'$ preceding its label starts with a zeroth coordinate, while any tuple with a $'$ preceding its label starts with a first coordinate.

We now inductively define polynomials in the a_i by

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

while, for $0 < k \leq n$

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

and for $k > n$

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

Remark 2.3. 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 2.4. 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 variety 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 $\mathbb{P}^{n-1}$

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

We refer to this family as the n^{th} *Tschirnhaus hypersurface of degree* i. When the superscript n is clear from context, we will suppress it for ease of reading.

Definition 2.5. 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 variety defined by the vanishing of the polynomials (2.4) for $i = i_1, \ldots, i_k$. Equivalently, define

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

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 2.6. The hyperplane $T_1(\mathbf{a}) \subset \mathbb{P}^{n-1}_{\mathbf{b}}$ 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

$$\begin{aligned} \mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-2} &\to^{\cong} T_1 \\ \left(\mathbf{a}, [b_1 : \cdots : b_{n-1}]\right) &\mapsto \left(\mathbf{a}, \left[-\frac{1}{n}\sum_{i=1}^{n-1} p_i b_i : b_1 : \cdots : b_{n-1}\right]\right). \end{aligned}$$

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

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

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

$$\begin{aligned} \mathbb{A}^n_{\mathbf{a}} \times \mathbb{P}^{n-3} &\to^{\cong} T_1 \\ \left(\mathbf{a}, [b_1 : \cdots : \hat{b_i} : \cdots : b_{n-2}]\right) &\mapsto \left(\mathbf{a}, \left[b_1 : \cdots : b_{i-1} : \frac{-1}{p_i}\sum_{j\neq i} p_j b_j : b_{i+1} : \cdots b_{n-2}\right]\right). \end{aligned}$$

As a warm-up to Theorem 2.12 below, we prove the following.

Lemma 2.7. *The families of quadrics $T_{12} \to \mathbb{A}^n_{\mathbf{a}}$ and $T'_{12} \to \mathbb{A}^n_{\mathbf{a}}$ are generically smooth.*

Remark 2.8. 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 $x^n + a_1 x^{n-1} + \cdots + a_n$ (see, e.g., [Syl, p. 468–469]). We give a different proof in order to warm-up for Theorem 2.12.

Proof of Lemma 2.7. 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

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

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

$$\begin{cases} -2na\left(\sum_{i=1}^{\frac{n-1}{2}} b_i b_{n-i}\right) = 0 & n \quad \text{odd}, \\ -na\left(b_{\frac{n}{2}}^2 + 2\sum_{i=1}^{\frac{n}{2}-1} b_i b_{n-i}\right) = 0 & n \quad \text{even}. \end{cases} \tag{2.5}$$

The partial derivatives of the defining polynomial 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 $\mathbb{Z}[1/2n]$ so long as $a \neq 0$ (and thus $T_{12} \to \mathbb{A}^n_{\mathbf{a}}$ is generically smooth).

We now prove $T'_{12} \to \mathbb{A}^n_{\mathbf{a}}$ 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 $T'_1(a)$. In these coordinates, and abusing notation by writing the same symbol for a hypersurface and its defining polynomial, we have

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

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 $\mathbb{Z}[1/2(n-1)]$ so long as $a \neq 0$ (and thus $T'_{12} \to \mathbb{A}^n_{\mathbf{a}}$ 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 [Tsch]). 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

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

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

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

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 $T_i|_{\mathbb{A}^n_{\mathbf{x}}} \to \mathbb{A}^n_{\mathbf{x}}$ is given by

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

Consider the map

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

Lemma 2.9. *In the notation above,*

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

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 $\mathrm{ev}(\mathbf{x}, \tilde{\mathbf{b}})_\ell := \sum_{j=0}^{n-1} b_j x_\ell^j$. By the Multinomial Theorem,

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

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. □

Example 2.10. Let $S \subset \mathbb{P}^4$ be the Clebsch diagonal surface, i.e., the complete intersection

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

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

$$\widetilde{T^5_{13}|_{\mathbb{A}^5_{\mathbf{x}}}} = \mathrm{ev}^{-1}(\widetilde{S}).$$

As observed by Klein [Klei3, Part II, Ch. 2], $\widetilde{T^5_{13}|_{\mathbb{A}^5_{\mathbf{x}}}}$ can be understood as a space of S_5-equivariant maps of $\mathbb{A}^5_{\mathbf{x}} \to \widetilde{S}$.

Example 2.11. Let $F \subset \mathbb{P}^6$ be the symmetric Fano sextic 3-fold as in [Bea], i.e., the complete intersection

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

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

$$\widetilde{T^7_{123}|_{\mathbb{A}^7_{\mathbf{x}}}} = \mathrm{ev}^{-1}(\widetilde{F}).$$

Though not remarked upon in [Bea], 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 [Hil1]:

$$z^7 + az^3 + bz^2 + cz + 1 = 0.$$

The variety $\widetilde{T^7_{123}|_{\mathbb{A}^7_{\mathbf{x}}}}$ can be understood as a space of S_7-equivariant maps of $\mathbb{A}^7_{\mathbf{x}} \to \widetilde{F}$, 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 2.12. *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*

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

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*

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

is generically smooth.

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

We now record a special case of Kleiman's Bertini Theorem [Kle]; for ease of reading, we include the proof 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* $Gr(k, Q) \to X$ *denote the relative Grassmannian of* k*-dimensional isotropic subspaces in* Q, *and let* $\mathcal{L} \to Gr(k, Q)$ *denote the tautological bundle. Let* $Y \subset \mathbb{P}^n_X$ *be a smooth family of varieties over* X *such that the family* $Q \times_{\mathbb{P}^n_X} Y \to X$ *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* $\mathcal{L}|_U \times_{\mathbb{P}^n_X} Y|_V \to X$ *is smooth.*

Combining Theorem 2.12, Lemma 2.7 and Proposition 2.13, we obtain the following.

Corollary 2.14. *Let* $Gr(T_{12}) \to \mathbb{A}^n_{\mathbf{a}}$ *denote the relative Grassmannian of maximal isotropics in the family of quadrics* $T_{12} \to \mathbb{A}^n_{\mathbf{a}}$, *and let* $\mathcal{L} \to Gr(T_{12})$ *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*

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

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*

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

is smooth.

Proof. Note that to prove the existence of an open dense V, it suffices to restrict all of the varieties over $\mathbb{Z}$ above to a geometric generic point $\mathrm{Spec}(K) \to \mathrm{Spec}(\mathbb{Z})$. The result now follows immediately from Theorem 2.12, Lemma 2.7 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 [Dix, §8].

Proof of Proposition 2.13. We recall Kleiman's proof [Kle]. Consider the canonical map

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

(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 $\mathcal{L} \to Q$ 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 $Q \times_{\mathbb{P}^n_X} Y \to X$ 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

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

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

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

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

We now prove Theorem 2.12.

Proof of Theorem 2.12. We prove the two cases separately, via parallel arguments. As in the proof of Lemma 2.7, if it will not cause confusion, we will abuse notation by writing the same symbol to denote a complete intersection and its defining polynomials.

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., $\mathbf{a} = (0, \ldots, 0, a)$. 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

$$n b_0 = 0.$$

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

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

and the partial derivatives are given by

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

Similarly, using Notation 2.2, 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

$$T_{j,12}(a) := ab_{n-j}$$
$$T_{j,1i}(a) := \sum_{\ell=1}^{i-1}(-1)^{\ell}a^{\ell}\left(\sum_{'\kappa \ s.t. \ |'\kappa|=i-1,||'\kappa||+j=\ell n}\binom{i-1}{'\kappa}{}'\mathbf{b}^{'\kappa}\right).$$

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 Legendre's formula implies that a prime p divides all 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 .

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

(n.b. as we remark just below, only one term in the above sum is nonzero). 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

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

In this notation, we have

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

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 $\mathcal{O}_\alpha$. For ease of notation, denote

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

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

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

Note that, by construction, for each j, all monomials containing b_j appear in precisely one M_α.

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

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

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

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

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

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

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., $\mathbf{a} = (0, \ldots, 0, a, 0)$. 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 2.7, 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 follow Notation 2.2. In these coordinates and this notation, the partial derivatives of $T'_{12}(a)$ are given by

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

(as noted in the proof of Lemma 2.7). Similarly, we have

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

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

Define

$$T'_{j,12}(a) := a b_{n-1-j}$$

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

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 in characteristic 0 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

$$\nu(j) = p^{-r} j \in \big(\mathbb{Z}/(n-1)\mathbb{Z}\big) - \{0\} = \{1, \ldots, n-2\}.$$

Using ν, we have

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

Mutatis mutandis, we now complete the argument by the same reasoning as for Case 1. □

Remark 2.16. A similar argument shows that the Tschirnhaus hypersurface $T_i \to \mathbb{A}^n_{\mathbf{a}}$ 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 2.12, 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 Tschirnhaus 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 $K(E_\Phi)/K(X)$. As a bridge to the classical literature, we will also denote $K(E_\Phi)$ as $K(X)(\Phi)$ to emphasize that $K(E_\Phi)$ is obtained from the field $K(X)$ by adjoining the values of Φ.

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

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

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 exist 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

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

given by (extending L-linearly) the assignment $T_{\tilde{\mathbf{b}}}(z^j) := w^j$ for each $0 \le j \le n-1$. 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

$$T : \overline{K(X)}(\Phi) \to \overline{K(X)}(\Phi).$$

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 $Y \to X$ for $K(X)(\tilde{\mathbf{b}})/K(X)$, (i.e., a map of K-varieties $Y \to X$ 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

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

Recall that $\mathbb{A}^n_X := X \times_{\mathrm{Spec}(K)} \mathbb{A}^n_K$, viewed as a variety 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 $\mathcal{T}_\Phi \to X$ 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

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

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

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

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

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

i.e.,

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

Since $Aut_{\overline{K(X)}}(\overline{K(X)}(\Phi))$ is the pullback to $\overline{K(X)}$ of an open subvariety 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

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

Observe that the assignment

$$x \mapsto \big(c_1(x), \ldots, c_n(x)\big)$$

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 [Tsch]) is as follows. First, the collection of n-valued algebraic functions on X is given by $\mathbb{A}^n_X$, where $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{A}^n_X$ corresponds to the function $\Phi_{\mathbf{a}}$ of (3.1), i.e., the function

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

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

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

(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 subvariety $Z \subset \mathbb{A}^n_{X,\mathbf{a}}$ determines a Zariski closed subvariety

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

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 subvariety (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 subvariety 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

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

where the p_i s 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 $\tilde{T}_{i_1 \cdots i_k}(P_n) \to \mathbb{A}^n_{\mathbf{a}}$.

Similarly, define the *Tschirnhaus complete intersection*

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

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

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 $\tilde{T}'_{i_1\cdots i_k} \to \mathbb{A}^n_{\mathbf{a}}$.

Lemma 2.9 can now be equivalently restated as follows.

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

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

as subvarieties 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 2.5.

Similarly, we have

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

as subvarieties 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 [Bra2, AS, FW]).

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

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

such that $E_r \to U$ factors through a dominant map $E_r \to Y$ and such that for each $i \ge 0$, 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 $\tilde{Z}_i \to Z_i$ is a generically finite dominant map with $\dim(Z_i) \le d$.

Example 4.2. Consider the space $\mathbb{A}^n_{\mathbf{a}}$ of monic degree n-polynomials. This has a canonical n-sheeted branched cover $E_{P_n} \to \mathbb{A}^n_{\mathbf{a}}$ 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

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

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

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

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

Lemma 4.5. *Every dominant map $Y \to X$ 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, by Hilbert's Nullstellensatz, the closure of the union of all of these contains the generic fiber of $Y \to X$; 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 $Y \to X$. □

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

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

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

(1) $\mathrm{RD}(Y \to X) \leq \dim(X)$.

(2) *Let $Z \to X$ be any dominant map of K-varieties. Then*

$$\mathrm{RD}(Y \times_X Z \to Z) \leq \mathrm{RD}(Y \to X).$$

(3) *If $Y \to X$ is birationally equivalent to $W \to Z$, then*

$$\mathrm{RD}(Y \to X) = \mathrm{RD}(W \to Z).$$

(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*

$$\mathrm{RD}(Y \to X) = \max_{i,j}\{\mathrm{RD}(Y_{i,j} \to X_i)\}.$$

Proof. These follow immediately from the definition and the analogous properties for resolvent degree of generically finite dominant maps (cf. [FW, Lemmas 2.5, 2.6]). □

Lemma 4.8. *Let $Y \to X$ be a surjective map (on geometric points). Let $Z \to X$ be any map. Then*

$$\mathrm{RD}(Y|_Z \to Z) \leq \dim(X).$$

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

$$Y|_W \to W$$

is dominant. Therefore, by Lemma 4.7,

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

□

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

Proof. By Lemma 4.7 4 and [FW, Lemma 2.6], it suffices to prove this when Y is irreducible. In this case, any rational multi-section $U \subset Y$ of $Y \to X$ 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 $\mathrm{RD}(U \to X) = \mathrm{RD}(Y \to X)$ (as generically finite dominant maps). The lemma follows. □

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

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

and

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

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

Proof. For the first inequality, let $\{U_\alpha \subset Z\}$ be a dense set of rational multi-sections of $Z \to X$ with $\mathrm{RD}(U_\alpha \to X) \leq d$ 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 $V_\alpha := \mathrm{Image}(U_\alpha \to Y)$ is a subscheme of Y, and thus a rational multi-section of $Y \to X$. Since $Z \to Y$ is dominant, that $\{U_\alpha \subset Z\}$ is dense implies that $\{V_\alpha \subset Y\}$ is dense. By [FW, Lemma 2.7], we conclude that $\mathrm{RD}(U_\alpha \to X) \geq \mathrm{RD}(V_\alpha \to X)$. Minimizing over all $\{U_\alpha \subset Z\}$, we conclude that

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

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

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

is a dense set of rational multi-sections for $Z \to X$. By [FW, Lemmas 2.5, 2.7],

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

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

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

To show the equalities when $\dim(Y) = \dim(X)$ or $\dim(Z) = \dim(Y)$, it suffices, by Lemma 4.7(4), 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 $Z \to Y$ is a rational multi-section for $Z \to X$ and vice versa. In particular,

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

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

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

follows from what we have shown above. □

Special cases of the following are implicit in [Segl, Bral, Bra2].

Proposition 4.11. *Let $Y \to X$ be a dominant map of K-varieties. Let $S \to X$ be a map such that the generic fiber is a Severi–Brauer variety over $K(X)$, 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

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

Proof. By the Merkurjev–Suslin theorem [MS, Theorem 16.1], using that K is a field of characteristic 0, there exists a solvable extension $L./K(X)$ such that $S|_{\mathrm{Spec}(L)} \cong \mathbb{P}^n_L$. Because we are in characteristic 0, the extension $L/K(X)$ is separable, so picking a primitive element z and writing $L \cong K(X)(z)$, we can, by clearing denominators in the minimal polynomial for z over $K(X)$ and using that the discriminant of this minimal polynomial is not identically 0, realize L as $K(E)$ for $E \subset \mathbb{A}^1_X$ a locally closed subvariety such that the projection $E \to X$ is solvable and étale. Shrinking E as needed, we can extend the above isomorphism $S|_{\mathrm{Spec}(L)} \cong \mathbb{P}^n_L$ to an isomorphism $S|_E \cong \mathbb{P}^n_E$. 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 degree at most d over $K(E)$ (and the generic point is of degree d). Therefore, by [FW, Lemma 2.9], $Y|_E$ admits a dense set of rational multi-sections $\{U_\alpha \subset Y|_E\}$ with $\mathrm{RD}(U_\alpha \to E) \leq \mathrm{RD}(d)$. The images of these rational multi-sections in Y, $\{V_\alpha \subset Y\}$ are thus a dense set of rational multi-sections, and by [FW, Lemma 2.6], we have

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

□

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 $\mathrm{ev}^{-1}(Z) \to X$. Then

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

Proof. The multi-section $U \to \mathrm{ev}^{-1}(Z)$ 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 $(E_{P_n})_Z \to Z$ 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. □

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 [Hil2]. 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

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

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 [Wal, Theorem 1.6] (see also [Sta, Theorem 1.2]).

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

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

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

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

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

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

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

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

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 $d \geq 2$ and all $k \geq 1$,*

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

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

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

Similarly,

$$\dim(\mathcal{M}_{3,\psi(d,k)_{d-2}}) = \max\left\{0, \binom{3+\psi(d,k)_{d-2}}{3} - (\psi(d,k)_{d-2}+1)^2\right\}.$$

From the definition, this is a maximum of a ceiling function of a monotone increasing degree 6 polynomial in $\psi(d,k)_{d-3}$, all of whose derivatives are monotone increasing in the domain $\psi(d,k)_{d-3} \geq 1$, while $\dim(\mathcal{H}_{4,\psi(d,k)_{d-3}})$ is a monotone increasing quartic, all of whose derivatives are monotone increasing in the same domain. 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 equality for $(d,k) = (3,1)$ and direct inspection of the higher derivatives of the sextic and quartic polynomials in the interval $\psi(d,k)_{d-3} \geq 1$ (for which both left and right hand side equal 4; note that the inequality is vacuously true for $(d,k) = (2,1)$).

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 $(d,k) \in \mathbb{N}_{\geq 2} \times \mathbb{N}_{>0}$ follows from the inequality for $(d,k) = (2,1)$ (in which case the left hand side is 8 and the right hand side is 4). □

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

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

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

This motivates the following definition.

Definition 5.4. Given $(d,k) \in \mathbb{N}_{\geq 2} \times \mathbb{N}_{>0}$, define

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

For $r \in \mathbb{N}_{\geq 4}$, define

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

For $r \leq 3$, 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 first main theorem.

Theorem 5.6. *Let $\mathrm{F}\colon \mathbb{N} \to \mathbb{N}$ 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

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

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 [Bra2, 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. *Let $B(r)$ and $\mathrm{F}(r)$ be as above. There exists a monotone increasing function $\varphi\colon \mathbb{N} \to \mathbb{N}$, such that $\varphi(2) = 5$, and such that for $r \geq \varphi(d)$,*

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

In particular, $\mathrm{F}(r) \leq B(r)$ for all r and

$$\lim_{r\to\infty} B(r)/\mathrm{F}(r) = \infty.$$

Remark 5.9. (1) As remarked above, Brauer's bound $B(r)$ gives the best prior general bound once $r \geq 7$; in this range, Theorem 5.8 shows that F is the best current bound. For $r = 6$, Sylvester [Syl] proved that the bound $n = 44$ is sufficient, while for $r = 5$, Segre and Dixmier proved that $n = 9$ suffices. In Appendix A, we give explicit computations of $\mathrm{F}(r)$ for r up to 15 (at which point $\mathrm{F}(r)$ is approximately 3.6 billion). In particular, we see that $\mathrm{F}(5) = 9$ recovers the Hilbert-Wiman-Segre-Dixmier bound, and $\mathrm{F}(6) = 41$ improves Sylvester.

(2) We do not expect that the upper bounds of Theorem 5.6 are themselves sharp for two reasons: first, we expect that further optimizations 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 [Syl, SH1, SH2] in their study of Hamilton's work [Ham].

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 [Hil2]. 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 quadratic closure of 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

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

for $a_i \in K^\times$. For example, see [For] 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 $L = K(\sqrt{a_1}, \dots, \sqrt{a_n}) \subset K^{2\text{-solv}}$. The L-rational change of coordinates

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

converts the above quadratic form (5.2) to

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

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 $i = 1, \dots, \lfloor \frac{n}{2} \rfloor$. 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* $Q \to X$*, the solvable multi-sections of* $Gr(Q) \to X$ *are Zariski dense in* $Gr(Q)(K(X))$.

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.

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 [Bri] and Hamilton [Ham]. We reprove this Bring-Hamilton bound as the base of our induction, in order to show the uniform general method; simple modifications of the below can be used to re-derive the bound $\mathrm{F}(r)$ for $r \leq 3$.

For $n \geq 5$ we have a generically smooth family of quadrics $T_{12} \to \mathbb{A}^n_{\mathbf{a}}$ (by Lemma 2.7) of dimension at least 2. By Lemma 5.10, there exists a solvable branched cover

$$U_1 \to \mathbb{A}^n_{\mathbf{a}}$$

with a map over $\mathbb{A}^n_{\mathbf{a}}$ to the relative Grassmannian of maximal isotropics $Gr(T_{12})$, i.e., there exists a linear embedding

$$\mathcal{L}\colon U_1 \times \mathbb{P}^{\lfloor \frac{n-3}{2} \rfloor} \to T_{12}|_U.$$

Because $n \geq 5$, the dimension of the linear subspaces is at least 1. We can therefore intersect with $T_3|_{U_1}$ to get a rational map

$$\begin{aligned} U_1 &\dashrightarrow \mathbb{A}^3 \\ u &\mapsto \mathcal{L}(u) \cap T_3. \end{aligned}$$

Adjoining the solution of this family of cubics, we get a solvable branched cover

$$U_2 \to U_1$$

and a map $U_2 \to T_{123}$. By Lemma 4.12, we conclude that

$$\begin{aligned} \mathrm{RD}(n) &\leq \max\bigl\{\mathrm{RD}(U_2 \to \mathbb{A}^n_{\mathbf{a}}), \dim(\mathbb{A}^{n-3}_{a_1=a_2=a_3=0}) - 1\bigr\} \\ &= \max\{1, n-4\} = n-4. \end{aligned}$$

For the induction step, let $r \geq 5$ and 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

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

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

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

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

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

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

Case 1: *n* odd. Let $U_1 = \mathbb{A}^n_{\mathbf{a}}$. By Lemma 2.7, the family $T_{12} \to \mathbb{A}^n_{\mathbf{a}}$ is generically smooth. By Corollary 2.14, there exists a dense open $V \subset Gr(T_{12})$, such that

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

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,

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

More precisely, there exists a multi-section $U_2 \subset V$ such that $U_2 \to U_1$ 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

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

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

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

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

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

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

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

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

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

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

is surjective. Therefore, the map

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

is surjective, and by Lemma 4.8,

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

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

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

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

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

By Lemma 4.10 and the construction above,

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

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

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

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

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

such that

(1) $\mathrm{RD}(U_i \to U_{i-1}) \leq \dim(\mathcal{H}_{i,\psi(d,k)_{d-i+1}})$,

(2) we have a commuting diagram

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

with a trivialization

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

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

(3) and the assignment

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

defines a map

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

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

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

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

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

is surjective. Therefore, the map

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

is surjective, and by Lemma 4.8,

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

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

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

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

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

By Lemma 4.10 and the construction above,

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

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

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

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

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

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

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

Now let $\mathcal{L} \to \mathcal{H}_{d,\psi(d,k)_1}$ 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

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

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

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

By our inductive construction, we have a map

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

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

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

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

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

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

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

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

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

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,

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

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

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

Case 2: *n* even. Let $U_1 = \mathbb{A}^n_{\mathbf{a}}$. By Lemma 2.7, the family $T'_{12} \to \mathbb{A}^n_{\mathbf{a}}$ is generically smooth. By Corollary 2.14, there exists a dense open $V \subset Gr(T'_{12})$, such that

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

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,

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

More precisely, there exists a multi-section $U_2 \subset V$ such that $U_2 \to U_1$ 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

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

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

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

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

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

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

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

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

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 $\rho\colon \mathbb{N} \to \mathbb{N}$ 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 $k < \rho(d)$, either

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

or

$$\frac{(d+k)!}{d!} + 1 \neq \Phi(d,k)$$

(i.e., $\rho(d)$ is the least integer such that both conditions hold).

Granting the claim, let $\varphi(d) := \rho(d+1) + d + 2$. From Definition 5.4, we see that $\rho(3) = 2$, and thus $\varphi(2) = 6$. However, $\mathrm{F}(5) = 9$ while $B(5) = 25$, so we can modify φ by setting $\varphi(2) := 5$ as claimed. Moreover, 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,

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

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,

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

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

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

This implies that

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

It suffices to prove that

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

for some α_d, as then

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

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

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

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,

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

as claimed. □

A. Explicit bounds

TABLE 1
Upper Bounds on $\mathrm{RD}(n)$. In the rightmost column above, k is the dimension of the linear subspace on the degree d hypersurface that we use to construct the necessary Tschirnhaus transformation, e.g., for $r = 5$, $(d,k) = (3,1)$ and we are using a line on a cubic surface *à la* Hilbert to prove $\mathrm{F}(5) = 9$.

r	$\mathrm{F}(r)$	Best Prior Bound $B'(r)$	Source of $B'(r)$	$B'(r)/\mathrm{F}(r)$	(d,k)
2	3	3	Babylonians	1	
3	4	4	Ferrari	1	
4	5	5	Bring [Bri]	1	(2,1)
5	9	9	Segre [Segl]	1	(3,1)
6	41	44	Sylvester [Syl]	1.07	(3,2)
7	121	721	Brauer [Bra2]	5.95	(3,3)
8	841	5041	"	5.99	(3,4)
9	6721	40321	"	5.99	(3,5)
10	60481	362881	"	5.99	(3,6)
11	604801	3628801	"	5.99	(3,7)
12	6652801	39916801	"	5.99	(3,8)
13	78485043	$12!+1$	"	6.10	(4,8)
14	320082459	13!+1	"	19.45	(4,9)
15	3632428801	14!+1	"	24	(4,10)

B. Historical background

The theory has been a plant of slow growth.
Sylvester and Hammond, 1887[7]

Tschirnhaus [Tsch] introduced his transformation 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 [Bri] improved this for $n = 5$ to show that $\mathrm{RD}(5) = 1$. Hamilton [Ham] was the first to show that

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

More precisely, he showed the existence a monotone increasing function $H : \mathbb{N} \to \mathbb{N}$, such that $n - \mathrm{RD}(n) \geq r$ for $n \geq H(r)$.[8] Hamilton computed the initial

[7] [SHI, p. 286]
[8] The numbers $H(r)$ are listed as the "Hamilton numbers" in the *Online Encyclopedia of Integer Sequences*.

values of H (for $r \leq 7$). Five decades later, Sylvester [Syl] 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 [Syl, p. 485])[9], and Sylvester and Hammond [SH1], [SH2] gave a generating function for H.

Preceding Sylvester (and apparently unbeknownst to him at the time of [SH1]), Klein [Klei1] initiated a new approach to solving polynomials, linking it with group theory, representation theory, projective geometry, classical invariant theory, and the theory of elliptic and automorphic functions. Fundamental to Klein's vision was the goal of reducing a given algebraic function to a simplest possible "normal form", with the ideal being a normal form given by the action of the monodromy group of the function on a projective space of minimal dimension.[10] For $n = 5, 6, 7$, this program allowed Klein [Klei3, Klei3, Klei6] to reproduce the Bring/Hamilton bounds of $\mathrm{RD}(n) \leq n-4$ with substantial simplifications in both the algebra of the formulas and the geometry of the normal forms involved. Klein also popularized the problem of finding simplest solutions of polynomials [Klei7, Second Part, Ch. II], was the first, or among the first, to explicitly consider the problem of lower bounds for RD [Klei5, Klei6], and worked, over a 50 year span, to anchor this problem firmly within the central mathematical concerns of his time (see also [Klei4, Klei2], and more generally [Klei8, Fri]).

In his 1900 address at the Universal Exposition in Paris, Hilbert [Hil1, Problem 13] explicitly posed the problem of the non-existence of 2-variable formulas for the general degree 7 polynomial. Hilbert's address cements two decisive shifts for the problem: first, he explicitly called attention to the question of lower bounds on resolvent degree, made conjectures as to lower bounds, and advocated for this as the fundamental problem. Second, Hilbert built upon Enriques' 1897 ICM address [Enr] by generalizing the problem to encompass formulas using analytic functions and even continuous ones; he then proved by a dimension count that the general three variable analytic function does not admit a formula in analytic functions of two or fewer variables. Hilbert returned to this problem at the end of his career in [Hil2], where he explicitly conjectured that $\mathrm{RD}(6) = 2$, $\mathrm{RD}(7) = 3$, $\mathrm{RD}(8) = 4$, and then sketched a beautiful geometric idea to lower $\mathrm{RD}(9)$ to at most 4. Shortly after, Wiman [Wim] sketched another approach to showing $\mathrm{RD}(n) \leq n-5$ for $n \geq 9$. As Dixmier observed [Dix], there are gaps in both

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

[10] As Wiman proved, this program cannot produce a solution in $\mathrm{RD}(n)$ variables for the general degree n polynomial once n is at least 8.

Hilbert and Wiman's proofs due to their assuming certain forms are sufficiently generic.

Progress on the general problem of bounding $\mathrm{RD}(n)$ stalled after Hilbert. N. Chebotarev highlighted this and related questions in his 1932 ICM address [Che3], and in several papers in the 1930s and 1940s [Che1, Che2, Che4, Che5]. However, by the mid-20th century, much of the 19th century work appears to have been forgotten. Segre [Seg1], building on Hilbert, provided the first rigorous proof that $\mathrm{RD}(n) \leq n-5$ for $n \geq 9$, and proved that for $n \geq 157$, $\mathrm{RD}(n) \leq n-6$ (n.b. Hamilton proved this for $n \geq 47$, while Sylvester proved it for $n \geq 44$). G. Chebotarev (N.'s son) worked to extend Wiman's methods to show $\mathrm{RD}(n) \leq n-6$ for $n \geq 21$ [Cheb], but his proof is incomplete.[11] Segre (*loc. cit.*) conjectured that in the limit

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

(i.e., precisely what Hamilton had showed over a century earlier). Brauer [Bra1] and Segre each reproved this statement, but without giving effective bounds *à la* Hamilton (see also [Seg2]).

In 1957, Arnold (then 19 years old) published a theorem which he described as a "complete solution of the 13th problem of Hilbert" [Arn1]. A strengthening of Arnold's theorem, published soon after by Kolmogorov [Kol], states that for any continuous map $f : [0,1]^n \to \mathbb{R}$, there exist continuous functions $g_j, \varphi_{ij} : [0,1] \to \mathbb{R}$ such that

$$f(t_1, \ldots, t_n) = \sum_{j=1}^{2n-1} g_j \left(\sum_{i=1}^{k} \varphi_{ij}(t_i) \right)$$

To apply this to Hilbert's problem, one must interpret Hilbert as having asked for an obstruction to expressing a single-valued branch of the general degree 7 polynomial as a composition of (single-valued) continuous functions of two or fewer variables. Following Arnold and Kolmogorov, work on the problem in all of its forms largely collapsed, this despite Arnold's efforts over a four decade span [Arn2, Arn3, Arn4, AS, Arn5] to call attention to and solve Hilbert's (still open!) thirteenth problem.[12]

[11] As remarked above, Chebotarev's argument has the same gap that Dixmier [Dix] observed in Hilbert and Wiman, namely certain non-generic forms are assumed to be generic.

[12] See also [Arn6, Problems 1972-27, 1976-34, 1979-10, 1980-10, 1985-18]

In 1971, Khovanskii [Kho] showed that if one prohibited the use of division in a formula (i.e., one only allowed "entire" algebraic functions), then the quintic was not solvable in 1-variable functions.[13] Khovanskii emphasized that, more than anything else, this result shows the importance of division.[14]

In 1975, Brauer [Bra2] gave the first rigorous definition of resolvent degree in the literature (followed soon after by Arnold and Shimura [AS]). Brauer then proved that for $n \geq (r-1)! + 1$, $\mathrm{RD}(n) \leq n - r$. This improves Sylvester and Hamilton's bounds for $r \geq 7$, and for such r provides the best upper bound, of which we are aware, prior to this paper.

While not strictly on $\mathrm{RD}(n)$, McMullen's work on iterative algorithms [McM] and his iterative solution of the quintic with Doyle [DM] represent one of the major outgrowths of Arnold's efforts to obstruct solutions of polynomials. More recently, Buhler–Reichstein's formalization of the Kronecker–Klein resolvent problem [BR1, BR2], and the broader theory of essential dimension that this given rise to, provides the closest contemporary body of work (see, e.g., [Rei], [Mer], [FKW1]).

The interested reader can find other discussions of the history of the problem in Sylvester and Hammond [SH1], in Klein [Klei9], or more recently in the surveys by Dixmier [Dix] and Vitushkin [Vit]. For a contemporary treatment of resolvent degree and its relation to classical problems see also [FW, FKW2].

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. I thank Curt McMullen for helpful conversations and for extensive helpful comments and suggestions on a draft. I thank Aaron Landesman and Igor Dolgachev for helpful comments on a draft. I thank Jordan Ellenberg, Vlad Matei, Madhav Nori, Zinovy Reichstein, Daniil Rudenko and David Smyth for helpful conversations. Last, I thank the referee for many helpful comments. Supported in part by NSF Grant DMS-1811846.

[13] A late paper of Abhyankar [Abh], apparently unaware of Khovanskii's result, proves the analogous theorem for the sextic.

[14] Lin has also extensively investigated what one can say for the general degree n polynomial if one rules out division and possibly imposes further restrictions, see the papers [Lin1, Lin2, Lin3].

References

[Abh] S. Abhyankar, Hilbert's thirteenth problem. Alev, J. (ed.) et al., *Algèbre non commutative, groupes quantiques et invariants.* Septième contact Franco–Belge, Reims, France, June 26-–30, 1995. Société Mathématique de France. Sémin. Congr. 2 (1995), pp. 1–11.

[Arn1] V. Arnold, On continuous functions of three variables. *Dokl. Akad. Nauk SSSR* **114** (1957), 679–681.

[Arn2] —— On some topological invariants of algebraic functions. *Trans. Moscow Math. Soc.* **21** (1970), 30–52. Zbl 0225.14005

[Arn3] —— Cohomology classes of algebraic functions invariant under Tschirnhausen transformations. *Funct. Anal. Appl.* **4** (1970), 74–75. Zbl 0232.14006

[Arn4] —— Topological invariants of algebraic functions, II. *Funct. Anal. Appl.* **4** (1970), 91–98. Zbl 0239.14012 MR 0276244

[Arn5] —— From Hilbert's superposition problem to dynamical systems. *The Arnoldfest. Proceedings of a Conference in Honour of V. I. Arnold for his 60th Birthday, Toronto, Canada, June 15–21, 1997*, AMS, Fields Inst. Commun. **24** (1999), 1–18. Zbl 0967.01012 MR 1733564

[Arn6] —— *Arnold's Problems*. Springer, 2000. Zbl 1052.00003 MR 2078115

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

[Bea] A. Beauville, Non-rationality of the symmetric sextic Fano threefold. *Geometry and Arithmetic*, 57–60, EMS Series of Congress Reports, Eur. Math. Soc., Zürich, 2012. Zbl 1317.14033 MR 2987652

[Bra1] R. Brauer, A note on systems of homogeneous algebraic equations, *Bull.* AMS **51** (1945), 749–755. Zbl 0063.00599 MR 0013127

[Bra2] —— On the resolvent problem. *Ann. Mat. Pura Appl.* (*4*) **102** (1975), 45–55. Zbl 0299.12105 MR 0371854

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

[BR1] J. Buhler and Z. Reichstein, On the essential dimension of a finite group. *Compositio Math.* **106** (1997), 159–179. Zbl 0905.12003 MR 1457337

[BR2] —— On Tschirnhaus transformations. *Topics in Number Theory* (*University Park, PA, 1997*), 127–142, Math. Appl., 467, Kluwer Acad. Publ., Dordrecht. Zbl 0937.12001 MR 1691314

[Che1] N. Chebotarev, Über ein algebraisches Problem von Herrn Hilbert. I. *Math. Ann.* 104 (1931), 459–471. Zbl 0001.05103

[Che2] —— Über ein algebraisches Problem von Herrn Hilbert. II. *Math. Ann.* 105 (1931), 240–255. Zbl 0002.18101

[Che3] —— Die Probleme der modernen Galoisschen Theorie. *Proceedings of the International Congress of Mathematicians*, 1932. Zbl 0006.28904

[Che4] —— Über das Klein–Hilbertsche Resolventenproblem. *Bull. Soc. Phys.-Math. Kazan, III.* Ser. **6** (1934), 5–22. Zbl 0010.29002

[Che5] —— The problem of resolvents and critical manifolds. *Izvestia Akad. Nauk SSSR* **7** (1943), 123—146. Zbl 0060.04901 MR 0011068

[Cheb] G. Chebotarev, On the problem of resolvents. *Uchenye Zapiski Kazanskogo Universiteta* **114** (1954) Book 2, 189–193. MR 0077496

[Dix] J. Dixmier, Histoire de 13e problème de Hilbert. *Cahiers du séminaire d'histoire des mathématiques*, 2e série, tome **3** (1993), 85–94. Zbl 0795.01012 MR 1240756

[DM] P. Doyle and C. McMullen, Solving the quintic by iteration. *Acta Math.* **163** (1989), 151–180. Zbl 0705.65036 MR 1032073

[Enr] F. Enriques, Sur les problèmes qui se rapportent à la résolution des équations algébriques renfermant plusieurs inconnues. *Proceedings of the International Congress of Mathematicians*, 1897. JFM 29.0073.03

[FKW1] B. Farb, M. Kisin and J. Wolfson, Essential dimension of congruence covers. arXiv:1901.09013

[FKW2] —— Modular functions and resolvent problems (with an appendix by N. Harman). arXiv:1912.12536

[FW] B. Farb and J. Wolfson, Resolvent degree, Hilbert's 13th problem and Geometry. *L'Enseignement Mathématique* **65** (2019), 303–376. MR 4113045

[For] T. Fort, Formulas for reducing a quadratic form to a sum of squares. *American Mathematical Monthly* **43** (1936), 477–481. JFM 62.1082.01

[Fri] R. Fricke, *Lehrbuch der Algebra*, vol. 2. Vieweg und Sohn, Braunschweig, 1926. JFM 52.0095.01

[Ham] 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.

[Har] R. Hartshorne, *Algebraic Geometry*, GTM 52, Springer-Verlag, New York, NY, 1977. Zbl 0531.14001 MR 0463157

[Hil1] D. Hilbert, Mathematical problems. From *Proceedings of the 1900 ICM*, English translation reprinted in *Bull. AMS* **37** (2000), 407–436. Zbl 0979.01028 MR 1779412

[Hil2] —— Über die Gleichung neunten Grades. *Math. Ann.* **97** (1927), 243–250. (English translation: On the equation of JFM 52.0103.02) MR 1512361

[HL] M. Hochster and D. Laksov, The linear syzygies of generic forms. *Comm. Algebra* **15** (1987), 227–239. Zbl 0619.13007 MR 0876979

[Kho] A. Khovanskii, The representability of algebroidal functions by superpositions of analytic functions and algebroidal functions of one variable. *Funct. Anal. Appl.* **4** (1970), 152–156. Translation from *Funkts. Anal. Prilozh.* **4** (1970), 74–79. MR 0267139

[Kle] S. Kleiman, The transversality of a general translate. *Compositio Math.* **28** (1974), 287–297. Zbl 0288.14014 MR 0360616

[Klei1] F. Klein, Ueber eine geometrische Repräsentation der Resolventen algebraischer Gleichungen. *Math. Ann.* **4** (1871), 346–358. JFM 03.0060.02 MR 1509753

[Klei2] —— Ueber die Auflösung gewisser Gleichungen vom siebenten und achten Grade. *Math. Ann.* 15 (1879), JFM 11.0074.03 MR 1510011

[Klei3] —— *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.

[Klei3] —— Zur Theorie der allgemeinen Gleichungen sechsten und siebenten Grades. *Math. Ann.* **28** (1887), 499–532. JFM 19.0082.03 MR 1510392

[Klei4] —— Sur la resolution, par les fonctions hyperelliptiques de l'equation du vingt-septieme degre, de laquelle depend la determination des vingt-sept droites d'une surface cubique. *Journal de Mathématiques pures et appliquées* (*4*) **4** (1888), 169–176. Zbl 20.0816.01

[Klei5] —— *Lectures on Mathematics*, MacMillan and Co., 1894.

[Klei6] —— Über die Auflösung der allgemeinen Gleichungen fünften und sechsten Grades. *Journal für die reine und angewandte Mathematik* **129** (1905), 150–174. English translation: About the solution of the general equations of fifth and sixth degree, translation by A. Sutherland, 2019, arXiv:1911.02358 JFM 36.0132.03

[Klei7] —— *Elementarmathematik vom höheren Standpunkte aus. Teil I: Arithmetik, Algebra, Analysis*. Teubner. (1908). English translation: *Elementary Mathematics from a Higher Standpoint*, Vol. 1: *Arithmetic, Algebra, Analysis*, translated from the 3rd German ed. by E. R. Hedrick and C. A. Noble, New York, Dover Publications (1953).

[Klei8] —— *Gesammelte Mathematische Abhandlungen*, vol. 2, pp. 255–504, 1922.

[Klei9] —— *History of the Development of Mathematics in the 19th Century*. Springer, 1926.

[Kol] A. Kolmogorov, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Dokl. Akad. Nauk SSSR* **114** (1957), 953–956. Zbl 0090.27103 MR 0111809

[Lin1] V. Lin, On superpositions of algebraic functions. *Funct. Anal. Appl.* 6 (1973), 240–241; translation from *Funkts. Anal. Prilozh.* **6** (1972), 77–78. Zbl 0272.12103

[Lin2] —— Superpositions of algebraic functions. *Funct. Anal. Appl.* **10** (1976), 32–38; translation from *Funkts. Anal. Prilozh.* **10** (1976), 37–45. Zbl 0346.32009 MR 0460329

[Lin3] —— Around the 13th Hilbert problem for algebraic functions. Teicher, Mina (ed.), *Proceedings of the Hirzebruch 65 Conference on Algebraic Geometry, Bar-Ilan University, Ramat Gan, Israel, May 2–7, 1993*. Bar-Ilan University. *Isr. Math. Conf. Proc.* **9** (1996), 307–327. Zbl 0846.32014 MR 1360511

[McM] C. MCMULLEN, Braiding of the attractor and the failure of iterative algorithms. *Invent. Math.* **91** (1988), 259–272. Zbl 0654.58023 MR 0922801

[Mer] A. MERKURJEV, Essential dimension. *Bull. AMS* **54** (2017), 635–661. MR 3683628

[MS] A. MERKURJEV and A. SUSLIN, K-cohomology of Severi–Brauer varieties and the norm residue homomorphism. *Math. USSR Izv.* **21** (1983), 307–340.

[Rei] Z. REICHSTEIN, Essential dimension. *Proceedings of the International Congress of Mathematicians*, Hyderabad, India, 2010. Zbl 1232.14030 MR 2827790

[Rob] H. ROBBINS, A remark on Stirling's formula. *The American Mathematical Monthly* **62** (1955), 26–29. Zbl 0068.05404 MR 0069328

[Seg1] B. SEGRE, The algebraic equations of degrees $5, 9, 157, \ldots$, and the arithmetic upon an algebraic variety. *Ann. of Math.* (*2*) **46** (1945), 287–301. Zbl 0061.01807 MR 0013126

[Seg2] —— *Arithmetical Questions on Algebraic Varieties*. University of London, Athlone Press, London, 1951. Zbl 0042.15204 MR 0043498

[Sta] J. STARR, Veronese varieties contained in hypersurfaces. arXiv:1703.03294

[Syl] J. SYLVESTER, On the so-called Tschirnhausen transformation. *J. Reine Angew. Math.* **100** (1887), 465–486. JFM 19.0079.02 MR 1580113

[SH1] J. SYLVESTER and J. HAMMOND, On Hamilton's numbers. *Phil. Trans. R. Soc. London A* **178** (1887), 285–312. JFM 19.0080.02

[SH2] —— On Hamilton's numbers II. *Phil. Trans. R. Soc. London A* **179** (1888), 65–71. JFM 20.0081.01

[Tsch] E. 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.

[Vit] A. VITUSHKIN, On Hilbert's thirteenth problem and related questions. *Russian Math. Surveys* **59** (2004), 11–25.

[Wal] A. WALDRON, Fano varieties of low-degree smooth hypersurfaces and unirationality. Bachelor thesis, Harvard University, Cambridge, Massachusetts, 2008.

[Wim] A. WIMAN, Über die Anwendung der Tschirnhausen-Transformation auf die Reduktion algebraischer Gleichungen. *Nova acta regiae societatis scientiarum Uppsaliensis*, **16** (1927), 3–8. JFM 55.0073.04

(*Reçu le 5 juin 2020*)

Jesse WOLFSON, Department of Mathematics,
University of California, Irvine,
340 Rowland Hall, Irvine, CA 92697-3875, USA

e-mail: wolfson@uci.edu