



Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra



Duality of multiple root loci

Hwangrae Lee^{a,*}, Bernd Sturmfels^b^a Pohang University of Science and Technology, South Korea^b University of California, Berkeley, USA

ARTICLE INFO

Article history:

Received 2 August 2015

Available online 9 October 2015

Communicated by Seth Sullivant

Keywords:

Multiple root

Dual variety

Conormal variety

Real rank

ED degree

ABSTRACT

The multiple root loci among univariate polynomials of degree n are indexed by partitions of n . We study these loci and their conormal varieties. The projectively dual varieties are joins of such loci where the partitions are hooks. Our emphasis lies on equations and parametrizations that are useful for Euclidean distance optimization. We compute the ED degrees for hooks. Among the dual hypersurfaces are those that demarcate the set of binary forms whose real rank equals the generic complex rank.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Univariate polynomials of degree n correspond to points in a projective space \mathbb{P}^n . The *multiple root locus* Δ_λ associated with a partition $\lambda = (\lambda_1, \dots, \lambda_d)$ is the subvariety of \mathbb{P}^n given by polynomials that have d distinct roots with multiplicities $\lambda_1, \dots, \lambda_d$. The dimension of Δ_λ is d . The singular locus of Δ_λ is the union of certain codimension one subloci Δ_μ , as described in [15, §3]. The degree of Δ_λ was determined by Hilbert in [11].

* Corresponding author.

E-mail addresses: meso@postech.ac.kr (H. Lee), bernd@berkeley.edu (B. Sturmfels).

He showed that

$$\deg(\Delta_\lambda) = \frac{d!}{m_1!m_2!\cdots m_p!} \cdot \lambda_1\lambda_2\cdots\lambda_d, \quad (1)$$

where m_j denotes the number of parts λ_i in the partition λ that are equal to the integer j .

The multiple root loci Δ_λ have been studied in a wide range of contexts and under various different names: coincident root loci [5,9,15], pejorative manifolds [16,23], strata of the discriminant [11,14,18], λ -Chow varieties [19], factorization manifolds [24, Definition 5.2.4], etc. Our motivation arose from the desire to understand the geometry of a model selection problem considered at the interface of symbolic computation [13] and numerical analysis [23]: given a univariate polynomial h , identify a low-dimensional Δ_λ such that h is close to Δ_λ .

Finding a point in Δ_λ that is closest to a given h is a problem of polynomial optimization [4]. We here characterize the geometric duality that underlies this optimization problem, in the sense of [4, Chapter 5]. The key player is the dual variety $(\Delta_\lambda)^\vee$. This variety lives in the dual projective space \mathbb{P}^n and it parametrizes all hyperplanes that are tangent to Δ_λ .

The duals to multiple root loci were studied by Oeding in [19]. He shows that $(\Delta_\lambda)^\vee$ is a hypersurface if and only if $m_1 = 0$, i.e. all parts of λ satisfy $\lambda_i \geq 2$. In [19, Theorem 5.3] an explicit formula is given for the degree of the polynomial that cuts out this hypersurface:

$$\deg((\Delta_\lambda)^\vee) = \frac{(d+1)!}{m_2!\cdots m_p!} \cdot (\lambda_1 - 1)(\lambda_2 - 1)\cdots(\lambda_d - 1). \quad (2)$$

For the application to optimization in [4, Theorem 5.23], this is the number of complex critical points one encounters when minimizing a general linear function over an affine chart of Δ_λ . For instance, consider $n = 5$ and $\lambda = (3, 2)$. Following Example 4.2 and [6, §3], the polynomial for $(\Delta_\lambda)^\vee$ is the *apple invariant* of degree 12. So, optimizing a linear function over quintics with a triple root and a double root leads to solving an equation of degree 12.

The present paper is a continuation of the studies by Hilbert [11] and Oeding [19]. It is organized as follows. In Section 2 we set up notation and basics. In Theorem 2.5 we parametrize the conormal variety Con_λ that links Δ_λ and $(\Delta_\lambda)^\vee$. Theorem 2.9 offers a parametrization for the projective duals of multiple root loci. These results are derived from the apolarity theory for binary forms, as described in the book by Iarrobino and Kanev [12].

In Section 3 we study the multidegree of the conormal variety Con_λ , and we summarize what is known about the ideals of Δ_λ and $(\Delta_\lambda)^\vee$. Table 1 offers a census of small

instances up to $n = 7$. We also examine duality for hooks, and inclusions among the dual strata.

Section 4 offers an application to tensor rank over \mathbb{R} . Theorem 4.1 characterizes the algebraic boundary of the set of binary forms whose real rank equals the generic complex rank. When n is even then this involves the use of Chow forms [10] and Hurwitz forms [21].

In Section 5 we turn to Euclidean distance (ED) optimization. Theorem 5.1 determines the ED degree of Δ_λ for hook shapes λ , both for invariant coordinates and for generic coordinates. This generalizes known results for the rational normal curve $\Delta_{(n)}$, in particular the tight connection to eigenvectors of symmetric tensors; cf. [8, Ex. 5.2 and Cor. 8.7].

Section 6 discusses implications for the ED optimization problem (21). The focus lies on locating all the critical points in Δ_λ and in $(\Delta_\lambda)^\vee$, and on certifying the global optimum. We work this out for several examples, including one from tensor decomposition over \mathbb{R} .

2. Duality for binary forms

We fix a field K of characteristic zero and we represent univariate polynomials by binary forms. Let $V = K[x, y]_n$ denote the space of binary forms of degree n and $V^\vee = K[u, v]_n$ its dual vector space. For $f \in V$ and $g \in V^\vee$, the pairing is defined by using partial derivatives:

$$\langle f, g \rangle = \frac{1}{n!} \cdot g \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) f(x, y) = \frac{1}{n!} \cdot f \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right) g(u, v). \quad (3)$$

This is the scalar in K obtained by interpreting one polynomial as a differential operator and applying it to the other polynomial. Introducing coordinates on V and V^\vee , we have:

$$\text{If } f = \sum_{i=0}^n \binom{n}{i} a_i x^i y^{n-i} \text{ and } g = \sum_{i=0}^n \binom{n}{i} b_i u^i v^{n-i} \text{ then } \langle f, g \rangle = \sum_{i=0}^n \binom{n}{i} a_i b_i. \quad (4)$$

We regard f and g as elements in the projective spaces $\mathbb{P}(V)$ and $\mathbb{P}(V^\vee)$. Both of these spaces are identified with \mathbb{P}^n using homogeneous coordinates $(a_0 : \cdots : a_n)$ and $(b_0 : \cdots : b_n)$.

A partition λ of n is represented alternatively as a list of integers $(\lambda_1, \dots, \lambda_d)$ satisfying $\lambda_1 \geq \cdots \geq \lambda_d \geq 1$ and $\sum_{i=1}^d \lambda_i = n$, or as a multiset $\{1^{m_1}, \dots, p^{m_p}\}$ satisfying $\sum_{j=1}^p j m_j = n$. So, $m_j = \#\{i : \lambda_i = j\}$, and d is the number of parts of λ , and p is the largest part of λ .

The multiple root variety $\Delta_\lambda \subset \mathbb{P}^n$ comprises all binary forms f of degree n that have m_j roots of multiplicity j . Equivalently, writing ℓ_i for linear forms, Δ_λ is the image of

$$(\mathbb{P}^1)^d \rightarrow \mathbb{P}^n, (\ell_1, \ell_2, \dots, \ell_d) \mapsto f = \ell_1^{\lambda_1} \ell_2^{\lambda_2} \cdots \ell_d^{\lambda_d}.$$

The variety Δ_λ has dimension d and degree as in (1). Among its smooth points are those given by d distinct linear forms ℓ_i . We determine the tangent space of Δ_λ at such a point.

Lemma 2.1. *The tangent space of Δ_λ at a general smooth point $f = \prod_{i=1}^d \ell_i^{\lambda_i}$ equals*

$$T_f \Delta_\lambda = \left\{ h(x, y) \cdot \prod_{i=1}^d \ell_i^{\lambda_i-1}(x, y) \mid h \in \mathbb{P}(K[x, y]_d) \right\} \simeq \mathbb{P}^d.$$

Proof. Write $\ell_i = a_i x + b_i y$ for some indeterminate $(a_i : b_i) \in \mathbb{P}^1$. The tangent space $T_f \Delta_\lambda$ is spanned by the binary forms $\frac{\partial}{\partial a_i} f = x f / \ell_i$ and $\frac{\partial}{\partial b_i} f = y f / \ell_i$ for $i = 1, 2, \dots, d$. These generators have the required form if we take $h(x, y)$ to be $x \cdot (\prod_{j=1}^d \ell_j) / \ell_i$ or $y \cdot (\prod_{j=1}^d \ell_j) / \ell_i$. For the converse we note that $K[x, y]_{d-1}$ is spanned by $\{(\prod_{j=1}^d \ell_j) / \ell_i \mid i = 1, \dots, d\}$. Indeed, this is a Lagrange basis, since the ℓ_i are pairwise linearly independent. Hence, by multiplying these polynomials with x and y , we obtain a spanning set for the vector space $K[x, y]_d$. \square

Lemma 2.2. *Let f be as above and $g \in \mathbb{P}(V^\vee)$. Then $g \perp T_f \Delta_\lambda$ if and only if*

$$\left[\prod_{i=1}^d \ell_i^{\lambda_i-1} \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right) \right] \text{ annihilates } g(u, v). \quad (5)$$

Proof. Let $\tilde{g}(u, v)$ be the binary form of degree d obtained from $g(u, v)$ by applying the operator on the left of (5). Then \tilde{g} is zero if and only if $\tilde{g}(u, v)$ is annihilated by $h(\frac{\partial}{\partial u}, \frac{\partial}{\partial v})$ for all $h \in K[x, y]_d$ if and only if $g(u, v)$ is annihilated by $(h \cdot \prod_{i=1}^d \ell_i^{\lambda_i-1})(\frac{\partial}{\partial u}, \frac{\partial}{\partial v})$ for all $h \in K[x, y]_d$. By Lemma 2.1, this means that $g(u, v)$ is orthogonal to the space $T_f \Delta_\lambda$. \square

The *conormal variety* of Δ_λ , here denoted Con_λ , is the Zariski closure of the set

$$\{(f, g) \mid f \text{ is a smooth point of } \Delta_\lambda \text{ and } g \perp T_f \Delta_\lambda\} \quad \text{in } \mathbb{P}(V) \times \mathbb{P}(V^\vee) = \mathbb{P}^n \times \mathbb{P}^n.$$

General results on projective duality ensure that Con_λ is an irreducible variety of dimension $n - 1$ in $\mathbb{P}^n \times \mathbb{P}^n$. The *dual variety* $(\Delta_\lambda)^\vee$ is the image of the conormal variety Con_λ under projection onto the second factor $\mathbb{P}(V^\vee)$. This is an irreducible variety of dimension $\leq n - 1$. The Biduality Theorem [10] states that the conormal variety of $(\Delta_\lambda)^\vee$ coincides with Con_λ , and hence $((\Delta_\lambda)^\vee)^\vee = \Delta_\lambda$. Lemma 2.2 implies the following description of the dual variety.

Corollary 2.3. *The points on the variety $(\Delta_\lambda)^\vee$ are binary forms $g(u, v)$ that are annihilated by some order $n - d$ operator of the form $\prod_{i=1}^d \ell_i^{\lambda_i-1}(\frac{\partial}{\partial u}, \frac{\partial}{\partial v})$ where $\ell_1, \dots, \ell_d \in K[x, y]_1$.*

At this point, let us pause to illustrate the concepts seen above with a small example.

Example 2.4. Let $n = 3$ and take $\lambda = (3)$, the partition with a single part. The conormal variety $\text{Con}_{(3)}$ is the surface in $\mathbb{P}^2 \times \mathbb{P}^2$ whose defining homogeneous prime ideal equals

$$\langle a_2^2 - a_1a_3, a_1a_2 - a_0a_3, a_1^2 - a_0a_2, 3b_1^2b_2^2 - 4b_0b_2^3 - 4b_1^3b_3 + 6b_0b_1b_2b_3 - b_0^2b_3^2, \\ a_0b_0 + 2a_1b_1 + a_2b_2, a_0b_1 + 2a_1b_2 + a_2b_3, a_1b_0 + 2a_2b_1 + a_3b_2, a_1b_1 + 2a_2b_2 + a_3b_3 \rangle. \quad (6)$$

The multiple root locus $\Delta_{(3)}$ is the twisted cubic curve in \mathbb{P}^3 , consisting of cubes of linear forms, and defined by the first three quadrics in (6). Its dual is $(\Delta_{(3)})^\vee = \Delta_{(2,1)}$. It consists of binary cubics with a double root, and it is the discriminant surface in \mathbb{P}^3 defined by the quartic in (6). We have $\text{Con}_{(3)} = \text{Con}_{(2,1)}$, after swapping the a -variables with the b -variables.

We illustrate Corollary 2.3 for the point $((1 : 1 : 1 : 1), (1 : 2 : -5 : 8))$ in $\text{Con}_{(3)}$. This represents the pair (f, g) where $f = (x + y)^3$ and $g = (u - v)^2(u + 8v)$. Then $g(u, v)$ lies in $(\Delta_{(3)})^\vee$ because $(\frac{\partial}{\partial u} + \frac{\partial}{\partial v})^2 g(u, v) = 0$, and $f(x, y)$ lies in $(\Delta_{(2,1)})^\vee$ because $(\frac{\partial}{\partial x} - \frac{\partial}{\partial y})f(x, y) = 0$. Note that the degree formulas (1) and (2) evaluate to 3 and 4 for $\lambda = (3)$. \diamond

The main result in this section is a parametric representation of the conormal variety.

Theorem 2.5. *The conormal variety Con_λ is the set of pairs $(f, g) \in \mathbb{P}^n \times \mathbb{P}^n$ of the form*

$$f(x, y) = \prod_{i=1}^d (t_i x - s_i y)^{\lambda_i} \quad \text{and} \quad g(u, v) = \sum_{\substack{i=1 \\ \lambda_i \neq 1}}^d (s_i u + t_i v)^{n-\lambda_i+2} \cdot g_i(u, v), \quad (7)$$

where $(s_i : t_i)$ runs over \mathbb{P}^1 and g_i runs over binary forms of degree $\lambda_i - 2$. This parametrization is a finite-to-one map $(\mathbb{P}^1)^{m_1} \times \mathcal{J}_\lambda \dashrightarrow \text{Con}_\lambda \subset \mathbb{P}^n \times \mathbb{P}^n$ whose degree is $m_1!m_2! \cdots m_p!$.

In order for this theorem to make sense, we need to define the parameter space. We set

$$\mathcal{J}_\lambda = \text{Join}(\mathbb{P}^1 \times \mathbb{P}^{\lambda_1-2}, \mathbb{P}^1 \times \mathbb{P}^{\lambda_2-2}, \dots, \mathbb{P}^1 \times \mathbb{P}^{\lambda_d-2}).$$

This is the *free join* (or *abstract join*) of the $d - m_1$ varieties in the argument. Here m_1 is subtracted from d because the j -th argument is empty and gets deleted when $\lambda_j = 1$. Note that \mathcal{J}_λ is the projective toric variety whose associated polytope is the free join of the product of simplices $\sigma_1 \times \sigma_{\lambda_j-2}$ for $j = 1, 2, \dots, d$. The dimension of this join equals

$$\dim(\mathcal{J}_\lambda) = (d - m_1 - 1) + \sum_{i=1}^d \max(0, 1 + (\lambda_i - 2)) = n - m_1 - 1.$$

The point of the toric variety \mathcal{J}_λ is to ensure that (7) gives a well-defined rational map. We will see in Example 2.10 that it is not a morphism. But that is not a problem since we can replace both the parameter space and image by their affine cones. The formulas in (7) are homogeneous. We use them in the next section to carry out the computations for Table 1.

By projection onto the second factor, Theorem 2.5 immediately yields a parametrization $(\mathbb{P}^1)^{m_1} \times \mathcal{J}_\lambda \dashrightarrow (\Delta_\lambda)^\vee \subset \mathbb{P}^n$ of the dual variety. Indeed, $(\Delta_\lambda)^\vee$ consists of all polynomials $g(u, v)$ of the form in (7). This can be rephrased in the language of projective geometry:

Corollary 2.6. *The dual variety $(\Delta_\lambda)^\vee$ indexed by λ is the join of $d - m_1$ multiple root loci determined by hook shapes, namely the loci $\Delta_{\{1^{\lambda_i-2}, n-\lambda_i+2\}}$, where $i = 1, \dots, d$ with $\lambda_i \geq 2$.*

Recall that a partition λ is a *hook* if at most one part is not 1. This is special for us:

Corollary 2.7. *The dual variety $(\Delta_\lambda)^\vee$ is also a multiple root locus Δ_μ if and only if the partition λ is a hook. In that case $\lambda = \{1^{n-a}, a\}$ and $\mu = \{1^{a-2}, n-a+2\}$ for some $a \in \{2, \dots, n\}$. In particular, Δ_λ is self-dual whenever n is even and $\lambda = \{1^{n/2-1}, n/2+1\}$.*

We easily derive the theorem from results that are well-known in commutative algebra.

Proof of Theorem 2.5. We use apolarity as presented in the book of Iarrobino and Kanev [12]. Let f and g be binary forms, possibly of different degrees. We say that f is *apolar* to g if $g(x, y)$ is annihilated by the operator $f(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$. Our Lemma 2.2 says that the tangent space $T_f \Delta_\lambda$ consists of all binary forms g of degree n such that $\ell_1^{\lambda_1-1} \ell_2^{\lambda_2-1} \dots \ell_d^{\lambda_d-1}$ is apolar to g . By [12, Lemma 1.31], this condition is equivalent to g having a *generalized additive decomposition* of the form $g = \ell_1^{n-(\lambda_1-1)+1} g_1 + \dots + \ell_d^{n-(\lambda_d-1)+1} g_d$. Here, only terms with $\lambda_i \geq 2$ may appear. This is precisely the representation on the right of (7), and we conclude that the proposed parametric representation of the conormal variety Con_λ is correct.

The parametrization is finite-to-one because $\dim(\text{Con}_\lambda) = n - 1$, which is the dimension of the parameter space $(\mathbb{P}^1)^{m_1} \times \mathcal{J}_\lambda$. Consider the fiber over a general point (f, g) in Con_λ . The first entry f is the image of precisely $m_1! m_2! \dots m_p!$ points in $(\mathbb{P}^1)^d$. For each such point $((s_1 : t_1), \dots, (s_d : t_d))$, the second entry gives a homogeneous system of linear equations:

$$\text{const.} \cdot g(u, v) = \sum_{\substack{i=1 \\ \lambda_i \neq 1}}^d (s_i u + t_i v)^{n-\lambda_i+2} \cdot g_i(u, v). \quad (8)$$

The unknowns are the coefficients of g_1, \dots, g_d . The solution set is a linear subspace of \mathcal{J}_λ . It is non-empty since (f, g) was assumed to lie in Con_λ . Since the parametrization is finite-

to-one, the solution space of the linear system must be a point. We conclude that the parametrization of the conormal variety Con_λ given in (7) has degree $m_1!m_2!\cdots m_p!$. \square

At this point it is important to note that our approach in this section is quite classical. All the ingredients are known and have been published elsewhere. What is new is their arrangement and interpretation. For instance, the parametrization (7) appears in [12] but the connection to dual and conormal varieties was not made explicit. Likewise, our Lemma 2.1 is among the statements in [14, Theorem 7.1]. The dual variety does make an appearance in [14, Section 6] but it was not seen that the dual of a multiple root locus for a hook is again such a locus. By putting all the known puzzle pieces together, we can go further in this paper. For instance, Katz states an inequality in [14, Corollary 6.4]. He conjectures in his introduction [14, p. 220] that equality holds. The following result proves Katz' conjecture.

Corollary 2.8. *The variety of binary forms of degree n with a root of multiplicity a satisfies*

$$\deg((\Delta_{\{1^{n-a}, a\}})^\vee) = \deg(\Delta_{\{1^{n-a+1}, a-1\}}) = (a-1)(n-a+2).$$

Proof. The dual on the left is $\Delta_{\{1^{a-2}, n-a+2\}}$ by Corollary 2.7. Using Hilbert's formula (1), we see that both varieties have the same degree, namely $(a-1)(n-a+2)$. \square

We close this section by recording the dimension of the dual variety, and by pointing out that our parametrization is in fact always *identifiable*, i.e. birational modulo permutations.

Theorem 2.9. *For any partition λ of n , the dual variety $(\Delta_\lambda)^\vee$ has dimension $n - m_1 - 1$. The generalized additive decomposition in (7) represents a finite-to-one parametrization $\mathcal{J}_\lambda \dashrightarrow (\Delta_\lambda)^\vee \subset \mathbb{P}^n$. This rational map has degree $m_2!\cdots m_p!$ and it is given by*

$$\left(((t_i : s_i))_{\lambda_i \neq 1}, (g_1, \dots, g_d) \right) \mapsto g(u, v) = \sum_{\substack{i=1 \\ \lambda_i \neq 1}}^d (s_i u + t_i v)^{n-\lambda_i+2} \cdot g_i(u, v). \quad (9)$$

Proof. It was shown in [14, Corollary 7.3] that $\dim((\Delta_\lambda)^\vee) = n - m_1 - 1$. This is also the dimension of the parameter space \mathcal{J}_λ . Theorem 2.5 implies that the map (9) is finite-to-one.

Fix a generic point $g(u, v)$ in the dual variety $(\Delta_\lambda)^\vee$. Consider any smooth point $f(x, y)$ of the primal Δ_λ at which g is tangent. The unordered set of linear forms ℓ_i that appear with multiplicity ≥ 2 in $f(x, y)$ can be recovered uniquely by Lemma 2.2. Permuting linear forms that appear with the same multiplicity accounts for the size $m_2!\cdots m_p!$ of the fiber when writing out the multipliers $(s_i u + t_i v)^{n-\lambda_i+2}$ in (9). At this point, we still need to recover the forms (g_1, \dots, g_d) , but this can be done uniquely by

the same argument as in the proof of [Theorem 2.5](#). We conclude that the degree of the map [\(9\)](#) equals $m_2! \cdots m_p!$ \square

Example 2.10. The rational parametrization in [Theorem 2.9](#) can have base points, so it is generally not a morphism. Let $\lambda = (3, 2)$, so $(\Delta_\lambda)^\vee$ is the hypersurface in \mathbb{P}^5 referred to as *little apple* in [Example 4.2](#). The toric fourfold \mathcal{J}_λ is the free join of $\mathbb{P}^1 \times \mathbb{P}^1$ and $\mathbb{P}^1 \times \mathbb{P}^0$. In these products, we fix the points $((s_1 : t_1), g_1)$ and $((s_2 : t_2), g_2)$, where $s_1 = s_2 = t_1 = t_2 = 1$, $g_1 = u + v$, and $g_2 = -1$. We also fix the scalar 1 for a point on the line that joins them. These choices specify a point in \mathcal{J}_λ . Plugging into the formula [\(9\)](#), we obtain

$$g(u, v) = (s_1 u + t_1 v)^4 \cdot g_1 + (s_2 u + t_2 v)^5 \cdot g_2 = (u + v)^4 \cdot (u + v) + (u + v)^5 \cdot (-1) = 0.$$

Hence our point in \mathcal{J}_λ is a base point of the parametrization [\(9\)](#) for $\lambda = (3, 2)$. \diamond

3. Equations, multidegree, and more

Given any parametrically represented variety, it is natural to ask for its *implicitization*. This concerns finding the ideals of polynomials that vanish on Con_λ , on Δ_λ , and on $(\Delta_\lambda)^\vee$. For the multiple root loci, this is a well-studied problem [\[5, 15, 22\]](#). Before reviewing what is known, we give the relevant definitions and we present a census of our varieties for $n \leq 7$.

We write $I(\text{Con}_\lambda)$ for the ideal of the conormal variety in the \mathbb{N}^2 -graded polynomial ring

$$\mathbb{Q}[\mathbf{a}, \mathbf{b}] = \mathbb{Q}[a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_n], \quad \text{with} \quad \deg(a_i) = (1, 0), \deg(b_i) = (0, 1).$$

The ideal $I(\text{Con}_\lambda)$ is bihomogeneous and prime. According to textbook definition [\[17, §8.5\]](#), the *multidegree* of the \mathbb{N}^2 -graded algebra $\mathbb{Q}[\mathbf{a}, \mathbf{b}]/I(\text{Con}_\lambda)$ is a binary form of degree $n + 1$:

$$C_\lambda(t_1, t_2) = \delta_1 t_1 t_2^n + \delta_2 t_1^2 t_2^{n-1} + \cdots + \delta_n t_1^n t_2. \quad (10)$$

The coefficients δ_i are nonnegative integers, known among geometers as the *polar classes* of the variety Δ_λ . Geometrically, δ_i is the number of intersection points in $(L_i \times M_i) \cap \text{Con}_\lambda$ where L_i is a generic plane of dimension i in $\mathbb{P}(V)$ and M_i is a generic plane of codimension $i - 1$ in $\mathbb{P}(V^\vee)$. In particular, we have $\delta_i = 0$ for $i < \text{codim}(\Delta_\lambda)$ and for $i > \dim((\Delta_\lambda)^\vee) + 1$.

The ideal of the multiple root locus Δ_λ is obtained by eliminating the variables b_0, \dots, b_n , and the ideal of its dual $(\Delta_\lambda)^\vee$ is obtained by eliminating the variables a_0, \dots, a_n . In symbols,

$$I(\Delta_\lambda) = I(\text{Con}_\lambda) \cap \mathbb{Q}[\mathbf{a}] \quad \text{and} \quad I((\Delta_\lambda)^\vee) = I(\text{Con}_\lambda) \cap \mathbb{Q}[\mathbf{b}]. \quad (11)$$

Table 1Multiple root loci, their duals, and their conormal varieties for $n \leq 7$.

λ	Equations of Δ_λ	Multidegree	Equations of $(\Delta_\lambda)^\vee$	Hooks	ED-degrees
2	2	2, 2	2	2	2, 4
3	2^3	0, 3, 4	4	21	3, 7
21	4	4, 3, 0	2^3	3	3, 7
4	2^6	0, 0, 4, 6	6	211	4, 10
31	$2^1, 3^1$	0, 6, 6, 0	$2^1, 3^1$	31	4, 12
211	6	6, 4, 0, 0	2^6	4	4, 10
22	3^7	0, 4, 6, 3	3	4, 4	7, 13
5	2^{10}	0, 0, 0, 5, 8	8	2111	5, 13
41	2^3	0, 0, 8, 9, 0	4^6	311	5, 17
311	4^6	0, 9, 8, 0, 0	2^3	41	5, 17
2111	8	8, 5, 0, 0, 0	2^{10}	5	5, 13
221	5^{10}	0, 12, 16, 6, 0	3^4	5, 5	16, 34
32	4^{28}	0, 0, 12, 21, 12	12	41, 5	21, 45
6	2^{15}	0, 0, 0, 0, 6, 10	10	21111	6, 16
51	2^6	0, 0, 0, 10, 12, 0	$4^1, 5^3, 6^1$	3111	6, 22
411	$2^1, 3^3, 4^1$	0, 0, 12, 12, 0, 0	$2^1, 3^3, 4^1$	411	6, 24
3111	$4^1, 5^3, 6^1$	0, 12, 10, 0, 0, 0	2^6	51	6, 22
21111	10	10, 6, 0, 0, 0, 0	2^{15}	6	6, 16
2211	7^{13}	0, 24, 30, 10, 0, 0	3^{10}	6, 6	28, 64
222	4^{45}	0, 0, 8, 16, 12, 4	4	6, 6, 6	20, 40
33	3^{29}	0, 0, 0, 9, 18, 12	12	51, 51	19, 39
321	$4^1, 5^3, 6^{31}$	0, 0, 36, 56, 24, 0	$4^1, 6^1$	51, 6	44, 116
42	$2^1, 3^3, 4^{31}$	0, 0, 0, 16, 30, 18	18	411, 6	26, 64
7	2^{21}	0, 0, 0, 0, 0, 7, 12	12	211111	7, 19
61	2^{10}	0, 0, 0, 0, 12, 15, 0	6^{10}	31111	7, 27
511	$2^3, 3^4$	0, 0, 0, 15, 16, 0, 0	4^{20}	4111	7, 31
4111	4^{20}	0, 0, 16, 15, 0, 0, 0	$2^3, 3^4$	511	7, 31
31111	6^{10}	0, 15, 12, 0, 0, 0, 0	2^{10}	61	7, 27
211111	12	12, 7, 0, 0, 0, 0, 0	2^{21}	7	7, 19
22111	9^{16}	0, 40, 48, 15, 0, 0, 0	3^{20}	7, 7	43, 103
2221	6^{78}	0, 0, 32, 60, 40, 10, 0	4^6	7, 7, 7	62, 142
3211	$6^{10}, 8^{38}$	0, 0, 72, 105, 40, 0, 0	$4^5, 6^{10}$	61, 7	73, 217
322	6^{364}	0, 0, 0, 36, 80, 66, 24	24	61, 7, 7	94, 206
331	3^{10}	0, 0, 0, 27, 48, 24, 0	7^8	61, 61	39, 99
421	$4^{20}, 6^{42}$	0, 0, 0, 48, 80, 36, 0	$8^1, 10^3, 11^2, 12^{49}$	511, 7	52, 164
43	$3^{10}, 4^{66}$	0, 0, 0, 0, 24, 51, 36	36	511, 61	51, 111
52	$2^3, 3^4, 4^{38}$	0, 0, 0, 0, 20, 39, 24	24	4111, 7	31, 83

These elimination ideals are \mathbb{N} -graded and prime. Their degrees and codimensions in the respective \mathbb{P}^n can be read off from the first term and the last term of the multidegree:

$$\mathcal{C}_\lambda(t_1, t_2) = \deg(\Delta_\lambda) \cdot t_1^{\text{codim}(\Delta_\lambda)} t_2^{\dim(\Delta_\lambda)+1} + \cdots + \deg((\Delta_\lambda)^\vee) \cdot t_1^{\dim((\Delta_\lambda)^\vee)+1} t_2^{\text{codim}((\Delta_\lambda)^\vee)}.$$

We computed the objects in (10) and (11) for all partitions up to $n = 7$. Our results are listed in Table 1. Each row corresponds to one partition λ . The second column lists the degrees of the minimal generators of $I(\Delta_\lambda)$. The third column lists the polar classes $\delta_1, \delta_2, \dots, \delta_n$. The leftmost nonzero entry in that list is the degree of Δ_λ and the rightmost nonzero entry is the degree of $(\Delta_\lambda)^\vee$. The codimension of the variety in question is the number of consecutive left (resp. right) zeros plus one. The fourth

column lists the degrees of the minimal generators of $I((\Delta_\lambda)^\vee)$. The fifth column lists the corresponding collection of hooks $\{1^{\lambda_i-2}, n - \lambda_i + 2\}$ that make up the dual variety in the join construction of [Corollary 2.6](#).

Finally, the last column refers to the Euclidean distance degrees in two coordinate systems. This will be explained in [Section 5](#). The second one is always $C_\lambda(1, 1) = \delta_1 + \delta_2 + \cdots + \delta_n$.

We now discuss some general facts that a reader might discover by looking at [Table 1](#). Whenever all hooks dual to λ are identical then $(\Delta_\lambda)^\vee$ is a secant variety of that hook variety.

Proposition 3.1. *Suppose $n \geq k(n - a + 2)$. The k -th secant variety of the variety of binary forms of degree n with a root of multiplicity a is projectively dual to the multiple root locus with partition $\lambda = \{1^{n-k(n-a+2)}, (n-a+2)^k\}$. In symbols,*

$$\sigma_k(\Delta_{\{1^{n-a}, a\}}) = (\Delta_\lambda)^\vee. \quad (12)$$

This secant variety is non-defective: it has the expected dimension $k(n - a + 1) + k - 1$.

Proof. This is the special case of [Corollary 2.6](#) where λ has the shape of a rectangle together with some singleton blocks, say $\lambda = \{1^{n-ku}, u^k\}$, and we set $a = n - u + 2$. The statement concerning the dimension follows from the first sentence in [Theorem 2.9](#). \square

The best known special case of this statement arises when $a = n$, or $u = 2$. The dual of $\Delta_{\{1^{n-2k}, 2^k\}}$ is the k -th secant variety of the rational normal curve $\Delta_{(n)}$. Its ideal is generated in degree $k+1$, namely by the minors of a Hankel matrix. We see this in [Table 1](#) for $\lambda = 21, 211, 22, 2111, 221, \dots$. Also of interest is the case of a rectangular partition, when the secant variety [\(12\)](#) is a hypersurface. Oeding's formula [\(2\)](#) and [Proposition 3.1](#) imply

Corollary 3.2. *The secant variety referred to in [Proposition 3.1](#), under the same hypothesis, is a hypersurface if and only if $n = k(n - a + 2)$. The degree of that hypersurface is*

$$\deg(\sigma_k(\Delta_{\{1^{n-a}, a\}})) = (k+1)(n-a+1)^k.$$

Example 3.3. Let $k = 2$, $n = 6$, $a = 5$, so $\lambda = (3, 3)$. The variety [\(12\)](#) is a hypersurface of degree 12 in \mathbb{P}^6 . It consists of all binary sextics $f = \ell_1^5 \ell_2 + \ell_3^5 \ell_4$ where the ℓ_i are linear forms.

Consider also the case $n = 12$. Here the construction gives four interesting hypersurfaces:

k	a	λ	forms	degree
2	8	(6, 6)	$\ell_1^8 g_1 + \ell_2^8 g_2$	75
3	10	(4, 4, 4)	$\ell_1^{10} g_1 + \ell_2^{10} g_2 + \ell_3^{10} g_3$	108
4	11	(3, 3, 3, 3)	$\ell_1^{11} g_1 + \ell_2^{11} g_2 + \ell_3^{11} g_3 + \ell_4^{11} g_4$	80
6	12	(2, 2, 2, 2, 2, 2)	$\ell_1^{12} + \ell_2^{12} + \cdots + \ell_6^{12}$	7

Of course, the last hypersurface is defined by the determinant of a 7×7 Hankel matrix. \diamond

The containment relation among multiple root loci is the order on partitions by refinement. Indeed, $\Delta_\lambda \subset \Delta_\mu$ if and only if μ *refines* λ , i.e. every part of λ is a sum of parts of μ . Our next result characterizes the containment relation among their dual varieties. Note that

$$(\Delta_\lambda)^\vee \subset (\Delta_\mu)^\vee \quad (13)$$

cannot hold unless λ has more 1's than μ , for dimension reasons. Given a partition $\lambda = \{1^{m_1}, 2^{m_2}, \dots, p^{m_p}\}$, we write λ' for the partition $\{1^{m_2}, \dots, (p-1)^{m_p}\}$. Note that if λ is a partition of n then λ' is a partition of $|\lambda'| = n - d$, where $d = \sum m_i$ is the number of parts.

Proposition 3.4. *Given two partitions λ and μ of n , the inclusion (13) holds if and only if $|\lambda'| \leq |\mu'|$ and, by adding to the parts, λ' can be transformed to a partition λ'' refined by μ' .*

Proof. This can be seen from Corollary 2.3. Let $\ell^{\lambda'}$ be the differential operator that is used in that corollary to characterize $(\Delta_{\lambda'})^\vee$. Here ℓ represents an arbitrary collection of linear forms. Our refinement condition means that every form g annihilated by $\ell^{\lambda'}$ is also annihilated by $\ell^{\mu'}$. This condition is equivalent to (13). More precisely, write m_a for the number of parts in λ' , and m_b for that in μ' . If such a λ'' exists, then we claim that

$$\bigcup_{\ell_{m_a}} \text{Ann}((\ell_{m_a})^{\lambda'}) \subset \bigcup_{\ell_{m_a}} \text{Ann}((\ell_{m_a})^{\lambda''}) \subset \bigcup_{\ell_{m_b}} \text{Ann}((\ell_{m_b})^{\mu'}).$$

Here, ℓ_k is a k -vector of linear forms and the exponent is written in multi-index notation. The first inclusion holds since each $(\ell_{m_a})^{\lambda''}$ can be seen as a multiple of $(\ell_{m_a})^{\lambda'}$, and the second inclusion holds since being annihilated by $(\ell_{m_a})^{\lambda''}$ can be seen as a special case of being annihilated by $(\ell_{m_b})^{\mu'}$, where ℓ_{m_b} is obtained by duplicating some parts of ℓ_{m_a} . \square

Example 3.5. We can check the inclusion (13) for various cases in Table 1. The inclusion holds for $\lambda = 2211$ and $\mu = 321$. Indeed, $\lambda' = 11$, and $\mu' = 21$ refines $\lambda'' = 21$. Likewise, if $\lambda = 321$ and $\mu = 222$ then $\mu' = 111$ refines $\lambda' = \lambda'' = 21$. This explains the inclusions $(\Delta_{2211})^\vee \subset (\Delta_{321})^\vee \subset (\Delta_{222})^\vee \subset \mathbb{P}^6$. It follows that the unique quartic polynomial 4^1 that vanishes on $(\Delta_{321})^\vee$ must be the 4×4 Hankel determinant whose hypersurface is $(\Delta_{222})^\vee$. \diamond

We now discuss some of the entries in the column *Equations of Δ_λ* in Table 1, and thereby review the literature on the ideals $I(\Delta_\lambda)$. For $\lambda = (n)$, we see the $\binom{n}{2}$ quadrics

that define the rational number curve, and for $\lambda = \{1^{n-2}, 2\}$ we get the discriminant of degree $2n - 2$. If $a \geq \lfloor n/2 \rfloor + 2$ then $I(\Delta_{\{1^{n-a}, a\}})$ is generated in degree ≤ 4 , as shown by Weyman [22]. The case $a = \lfloor n/2 \rfloor + 1$ is of special interest: here $\Delta_{\{1^{n-a}, a\}}$ is the *nullcone*, i.e. the variety defined set-theoretically by all SL_2 -invariants of binary n -ics. Note that the nullcone is self-dual when n is even (by Corollary 2.7), and its ideal is generated by quartics when n is odd.

Chipalkatti [5, Conjecture 6.1] had conjectured that $I(\Delta_\lambda)$ is generated in degree ≤ 4 whenever λ has only $d = 2$ parts, and this was proved by Abdesselam and Chipalkatti in [2, Proposition 20]. For further details on their approach see [1, Section 7]. The papers [1, 2, 5, 22] stress the fact that all our ideals are invariant under the action of SL_2 , and hence each of their graded components are direct sums of irreducible SL_2 -modules. Chipalkatti describes the minimal free resolutions of $I(\Delta_\lambda)$ in terms of SL_2 -modules for $\lambda = (3, 2)$ in [5, §4.1] and for $\lambda = (3, 3)$ in [5, §4.2]. He discusses the 364 sextics for $\lambda = (3, 3, 2)$ in [5, §3.1].

The ideal $I((\Delta_\lambda)^\vee)$ of the dual variety has been studied only in two cases, namely when λ is a hook (by self-duality) and when λ has only parts 1 and 2. In the latter case, it is generated by minors of a Hankel matrix. In all other cases, little seems to be known. Of course, $I((\Delta_\lambda)^\vee)$ will often be principal (when $m_1 = 0$), and we seek to find the generator of such an ideal explicitly. This is accomplished for some interesting cases in Section 4.

4. Real rank boundaries

In this section we present an application of our duality theory to the study of real ranks of binary forms [3, 6]. Let $f \in \mathbb{R}[x, y]_n$ be a general binary form of degree n with real coefficients. The complex rank of f equals $r = \lceil (n+1)/2 \rceil$. This means that f is a sum of r powers of linear forms over \mathbb{C} , but not fewer. We consider the set \mathcal{R}_n of all real binary forms f that admit a rank r decomposition also over \mathbb{R} . In other words, \mathcal{R}_n is the set of all forms $f \in \mathbb{R}[x, y]_n$ such that both the real rank of f and the complex rank of f are equal to r .

The set \mathcal{R}_n is a full-dimensional semi-algebraic set inside $\mathbb{R}[x, y]_n$. Its *topological boundary* $\partial\mathcal{R}_n$ is the set-theoretic difference of the closure of \mathcal{R}_n minus the interior of the closure of \mathcal{R}_n . Thus, if $f \in \partial\mathcal{R}_n$ then every open neighborhood of f contains a generic form of real rank equal to r and also a generic form of real rank bigger than r . The topological boundary $\partial\mathcal{R}_n$ is a semi-algebraic subset of pure codimension one inside the real affine space $\mathbb{R}[x, y]_n$.

We define the *real rank boundary*, denoted $\partial_{\mathrm{alg}}(\mathcal{R}_n)$, to be the Zariski closure of the topological boundary $\partial\mathcal{R}_n$. We view $\partial_{\mathrm{alg}}(\mathcal{R}_n)$ as a closed subvariety in the complex projective space $\mathbb{P}(\mathbb{C}[x, y]_n) = \mathbb{P}^n$. It is pure codimension one, so it is defined by a unique (up to scaling) squarefree polynomial in the coordinates a_0, a_1, \dots, a_n on \mathbb{P}^n . We compute that polynomial:

Theorem 4.1. *Let $n \geq 5$. If $n = 2k - 1$ is odd then the real rank boundary $\partial_{\text{alg}}(\mathcal{R}_n)$ is an irreducible hypersurface of degree $2k(k-1)$ in \mathbb{P}^n . This hypersurface is the dual $(\Delta_\lambda)^\vee$ of the multiple root locus Δ_λ where $\lambda = \{2^{k-2}, 3\}$. Its defining polynomial is the discriminant of*

$$q(u, v) = \det \begin{pmatrix} u^k & u^{k-1}v & \cdots & uv^{k-1} & v^k \\ a_0 & a_1 & \cdots & a_{k-1} & a_k \\ a_1 & a_2 & \cdots & a_k & a_{k+1} \\ a_2 & a_3 & \cdots & a_{k+1} & a_{k+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k-1} & a_k & \cdots & a_{n-1} & a_n \end{pmatrix}. \quad (14)$$

If $n = 2k$ is even then the real rank boundary $\partial_{\text{alg}}(\mathcal{R}_n)$ is the union of the two irreducible hypersurfaces $(\Delta_\lambda)^\vee$ where λ is $\{2^{k-3}, 3^2\}$ or $\{2^{k-2}, 4\}$. Their degrees are $2k(k-1)(k-2)$ and $3k(k-1)$. These arise as irreducible factors of the Hurwitz form of the discriminant $\Delta_{\{1^{k-1}, 2\}}$ when evaluated at the line in $\mathbb{P}(\mathbb{R}[x, y]_{k+1})$ that is the kernel of the $k \times (k+2)$ -matrix

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_k & a_{k+1} \\ a_1 & a_2 & \cdots & a_{k+1} & a_{k+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k-1} & a_k & \cdots & a_{n-1} & a_n \end{pmatrix}. \quad (15)$$

A third irreducible factor appearing in this specialized Hurwitz form is the Hankel determinant $(\Delta_{\{2^k\}})^\vee$, which has degree $k+1$, but this is not part of the real rank boundary $\partial_{\text{alg}}(\mathcal{R}_n)$.

For the relevant background on Chow forms and Hurwitz forms we refer to [10] and [21]. Before presenting the proof of Theorem 4.1, here is an illustration of the first three cases.

Example 4.2 (*Little apple and big apple*). Let $n = 5$, so $k = 3$ and $\lambda = (3, 2)$. In this first case, Theorem 4.1 was proved by Comon and Ottaviani in [6, §5]. The polynomial defining the hypersurface $(\Delta_\lambda)^\vee \subset \mathbb{P}^5$ is their *apple invariant* I_{12} . It has 228 terms of degree 12.

The next odd case is $n = 7$, so $k = 4$ and $\lambda = (3, 2, 2)$. The real rank boundary $(\Delta_\lambda)^\vee$ is a hypersurface of degree 24 in \mathbb{P}^7 , namely the join of the surface $\Delta_{(6,1)}$ and the threefold $\sigma_2(\Delta_{(7)}) = \sigma_2(\nu_7(\mathbb{P}^1))$. Its defining polynomial is computed via (14). It has 38082 terms. \diamond

Example 4.3 (*Chow and Hurwitz forms*). Let $n = 6$. We consider the discriminant $\Delta_{(2,1,1)}$ of binary forms of degree $k+1 = 4$. This discriminant is a threefold of degree 6 in \mathbb{P}^4 . Its singular locus consists of the surfaces $\Delta_{(3,1)}$ and $\Delta_{(2,2)}$. The Hurwitz form of $\Delta_{(2,1,1)}$ is a polynomial of degree $30 = 3 \cdot 6 + 2 \cdot 4 + 4$ in the ten Plücker coordinates on the

Grassmannian of lines in \mathbb{P}^4 . It is reducible because $\Delta_{(2,1,1)}$ is singular in codimension 1. It factors as

$$\text{Hur}(\Delta_{(2,1,1)}) = \text{Chow}(\Delta_{(3,1)})^3 \cdot \text{Chow}(\Delta_{(2,2)})^2 \cdot \text{Tan}. \quad (16)$$

Here $\text{Chow}(X)$ denotes the *Chow form* of a surface X in \mathbb{P}^4 , i.e. the polynomial in Plücker coordinates that vanishes when the line intersects X . The degree of the Chow form equals the degree of X , so it is 6 and 4 in our two cases. The last factor Tan is the condition for the line to be tangent at a smooth point of $\Delta_{(2,1,1)}$. This is a quartic in the Plücker coordinates.

We now take our line in \mathbb{P}^4 to be the kernel of the 3×5 matrix in (15). More precisely, in the formula (16) we substitute the 10 maximal minors of (15) for the 10 Plücker coordinates. Here signs have to be taken into consideration carefully. Each maximal minor is a cubic, so the degrees above have to be tripled. Equation (16) now has degree $90 = 3 \cdot 18 + 2 \cdot 12 + 3 \cdot 4$.

The polynomial obtained from $\text{Chow}(\Delta_{(3,1)})$ is $(\Delta_{(4,2)})^\vee$. It has 3140 terms of degree 18. The polynomial obtained from $\text{Chow}(\Delta_{(2,2)})$ equals $(\Delta_{(3,3)})^\vee$. It has 560 terms of degree 12. The degree 12 polynomial obtained from Tan is the third power of the Hankel determinant

$$(\Delta_{(2,2,2)})^\vee = \det \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 & a_4 \\ a_2 & a_3 & a_4 & a_5 \\ a_3 & a_4 & a_5 & a_6 \end{pmatrix}. \quad (17)$$

To illustrate the statement of Theorem 4.1, we present two explicit forms that lie in $\partial\mathcal{R}_6$.

The binary form $f = y^6 + 15x^4y^2$ lies in $\partial\mathcal{R}_6 \cap (\Delta_{(4,2)})^\vee$ but not in $(\Delta_{(3,3)})^\vee \cup (\Delta_{(2,2,2)})^\vee$. Its line of apolar quartics is $L_f = \{s \cdot uv^3 + t \cdot (u^4 - v^4) \mid (s : t) \in \mathbb{P}^1\}$. This has discriminant $(27s^4 + 256t^4)t^2$, with only real root at $(s : t) = (1 : 0)$. That quartic lies in $\kappa = \Delta_{(3,1)}$, and it is a limit of quartics with four real roots, and also a limit of quartics with two real roots. From this we can construct generic sextics $f_{\pm\epsilon}$ close to f whose real ranks are 4 and 5.

The binary form $f = y^6 + 5x^2y^4 - 5x^4y^2 - x^6$ lies in $\partial\mathcal{R}_6 \cap (\Delta_{(3,3)})^\vee$ but not in $(\Delta_{(4,2)})^\vee \cup (\Delta_{(2,2,2)})^\vee$. Its line of apolar quartics is $L_f = \{s \cdot (u - v)^2(u + v)^2 + t \cdot uv(u^2 + v^2) \mid (s : t) \in \mathbb{P}^1\}$. This has discriminant $(16s^2 + t^2)^2t^2$, with only real root $(s : t) = (1 : 0)$. That quartic lies in $\nu = \Delta_{(2,2)}$. We can construct nearby generic sextics $f_{\pm\epsilon}$ whose real ranks are 4 and 5.

The proof below will explain the relevance of L_f , κ , and ν for $\partial\mathcal{R}_6$. It will also show why the Hankel determinant $(\Delta_{(2,2,2)})^\vee$ meets the boundary of \mathcal{R}_6 only in lower dimension. \diamond

Remark 4.4. It is instructive to see the factorization (16) using *iterative discriminants*. Let $A(x)$ and $B(x)$ be univariate quartics, and consider $A(x) + tB(x)$ where t is an unknown. Then

$$\text{discr}_t(\text{discr}_x(A(x) + tB(x))) = ([6, 6]_{A,B})^3 \cdot ([4, 4]_{A,B})^2 \cdot ([4, 4]_{A,B}),$$

where $[e, e]_{A,B}$ stands for a big expression that has degree e in the coefficients of A and of B .

Proof of Theorem 4.1. We write f for a generic form in $\mathbb{R}[x, y]_n$. We first consider the odd case $n = 2k - 1$. The apolar ideal, consisting of all forms in $\mathbb{R}[u, v]$ such that $g(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ annihilates $f(x, y)$, is generated by forms $q(u, v)$ of degree k and $r(u, v)$ of degree $k + 1$. Furthermore, the dual form $q(u, v)$ has the Hankel determinantal representation in (14).

Since f is generic, there exists a unique decomposition $f(x, y) = \sum_{i=1}^k (s_i x + t_i y)^n$. The points $(s_i : t_i)$ are the complex roots of q . Hence f lies in \mathcal{R}_n precisely when q is real-rooted. By *real-rooted* we mean that q is square-free and its roots are all real.

Suppose now that the form f moves and passes through the boundary of \mathcal{R}_n . Then two real roots of $q(u, v)$ merge and become a double root. At this point, the discriminant of $q(u, v)$ vanishes. Corollary 2.3 implies that this discriminant equals $(\Delta_\lambda)^\vee$ where $\lambda = \{2^{k-2}, 3\}$.

Next consider the even case $n = 2k$. The apolar ideal of f is generated by two forms of degree $k + 1$. Let L_f be the line in \mathbb{P}^{k+1} spanned by these two forms. The points q on L_f correspond to the distinct decompositions $f(x, y) = \sum_{i=1}^{k+1} (s_i x + t_i y)^n$. Namely, the $(s_i : t_i)$ are the roots of q . This means that f lies in \mathcal{R}_n if and only if some q in L_f is real-rooted.

The set of all real-rooted forms is a connected full-dimensional semi-algebraic subset of $\mathbb{P}_{\mathbb{R}}^{k+1}$, and its algebraic boundary is the discriminant $\delta = \Delta_{\{1^{k-1}, 2\}}$. Note that δ is a hypersurface of degree $2k$. Its singular locus consists of the codimension two loci $\kappa = \Delta_{\{1^{k-2}, 3\}}$ and $\nu = \Delta_{\{1^{k-3}, 2, 2\}}$. Their degrees are $\deg(\kappa) = 3(k - 1)$ and $\deg(\nu) = 2(k - 1)(k - 2)$.

Suppose that the form f moves along a general curve. Consider its image under the map $f \mapsto L_f$ into the Grassmannian of lines in $\mathbb{P}_{\mathbb{R}}^{k+1}$, here denoted Gr . As f crosses the boundary of \mathcal{R}_n , the line L_f transitions from intersecting to not intersecting the subset of real-rooted forms. At the transition point, the form f is in $\partial_{\text{alg}}(\mathcal{R}_n)$. One of the following three scenarios might happen: (i) L_f intersects κ , or (ii) L_f intersects ν , or (iii) L_f is tangent to δ at a smooth point. Each of these conditions defines an irreducible hypersurface in the Grassmannian Gr . Their equations are the irreducible factors in the Hurwitz form $\text{Hur}(\delta)$.

The Hurwitz form of the discriminant factors as follows in the coordinate ring of Gr :

$$\text{Hur}(\delta) = \text{Chow}(\kappa)^3 \cdot \text{Chow}(\nu)^2 \cdot \text{Tan}(\delta). \quad (18)$$

The exponents 3 and 2 arise from the classical Plücker formula for the dual of a plane curve. Now, since δ is a hypersurface of degree $2k$, its Hurwitz form has degree $2k(2k - 1)$. Hence

$$2k(2k - 1) = 3\deg(\kappa) + 2\deg(\nu) + \deg(\text{Tan}(\delta)) = 9(k - 1) + 4(k - 1)(k - 2) + \deg(\text{Tan}(\delta)).$$

This uses the fact that the degree of a Chow form in Plücker coordinates equals the degree of its variety. We conclude that $\text{Tan}(\delta)$ is a polynomial of degree $k+1$ in Plücker coordinates.

The map $\mathbb{P}^n \dashrightarrow \text{Gr}$, $f \mapsto L_f$ is birational. Indeed, every generic line has the form L_f for a unique sextic f that is recovered by solving the differential equations represented by L_f . Moreover, the base locus of this inverse map is precisely the tangential hypersurface $\text{Tan}(\delta)$.

We now examine what happens to the three irreducible factors in (18) when pulled back under the map $\mathbb{P}^n \dashrightarrow \text{Gr}$. Algebraically, the line L_f is given as the kernel of the matrix (15). So, to compute the pullback of (18), we need to replace the Plücker coordinates by the (appropriately signed and scaled) maximal minors of (15). These minors are polynomials of degree k , so we must multiply the above degrees by k . This gives the degrees $3k(k-1)$, $2k(k-1)(k-2)$ and $k(k+1)$ for the pullbacks of the three irreducible factors in (18).

The first two pullbacks are irreducible as polynomials in $\mathbb{R}[\mathbf{a}] = \mathbb{R}[a_0, \dots, a_n]$. They are

$$\text{Chow}(\kappa)(L_f) = (\Delta_{\{2^{k-2}, 4\}})^\vee \quad \text{and} \quad \text{Chow}(\nu)(L_f) = (\Delta_{\{2^{k-3}, 3^2\}})^\vee. \quad (19)$$

As before, this follows from the description of the dual hypersurfaces in Corollary 2.3.

The third factor in (18) becomes a reducible polynomial in $\mathbb{R}[\mathbf{a}]$. Namely, we have

$$\text{Tan}(L_f) = ((\Delta_{\{2^k\}})^\vee)^k. \quad (20)$$

In words: the pullback of Tan is the k th power of the Hankel determinant of order $k+1$.

This can be seen as follows. A general point f on the Hankel hypersurface satisfies $\text{rank}_{\mathbb{C}}(f) = k$. It is annihilated by some binary form $g(u, v)$ of degree k , and the line L_f is spanned by $u \cdot g(u, v)$ and $v \cdot g(u, v)$. This means that L_f is tangent to δ at k points over \mathbb{C} , given by the k roots of $g(u, v)$. We see that the pullback $\text{Tan}(L_f)$ contains $(\Delta_{\{2^k\}})^\vee$ set-theoretically. Both are irreducible varieties, and hence they are equal as sets in \mathbb{P}^n . By comparing degrees, we conclude that (20) holds.

Our argument also shows the following fact: if a line of the form L_f is tangent to the discriminant δ at one smooth point then it is tangent to δ at a scheme of length k .

We have proved that $\partial_{\text{alg}}(\mathcal{R}_n)$ has at most three irreducible factors. However, the correct number is two. The two hypersurfaces that appear in $\partial_{\text{alg}}(\mathcal{R}_n)$ are those in (19). This was shown already for $n = 6$. Towards the end of Example 4.3, we exhibited one binary form f for each of these two boundary strata. In general, a construction can be made as follows.

We consider lines in the space \mathbb{P}^{k+1} of binary forms of degree $k+1$ that look like

$$L_\kappa = \{s \cdot uv^3 f(u, v) + t \cdot (u^2 + v^2)g(u, v) \mid (s : t) \in \mathbb{P}^1\} \quad \text{and}$$

$$L_\nu = \{s \cdot (u - v)^2(u + v)^2 f(u, v) + t \cdot (u^2 + v^2)g(u, v) \mid (s : t) \in \mathbb{P}^1\},$$

where f and g are generic real-rooted binary forms of degrees $k-3$ and $k-1$ respectively. The line L_κ meets the discriminant δ only in the cusp locus, and the line L_ν meets δ only in the node locus. The binary form has $k-1$ distinct real roots at these intersection points. Now, for suitable choices of f and g , the discriminant of the pencil of binary forms has no real roots other than $(s:t) = (1:0)$. We fix such f and g , and we set $L = L_\kappa$ or $L = L_\nu$.

The pencil L consists of binary forms of degree $k+1$ that have precisely $k-1$ real roots. Moreover, L is not contained in the hypersurface Tan . Recall that the inverse to $\mathbb{P}^n \dashrightarrow \text{Gr}$ is well-defined in a neighborhood \mathcal{U} of L . Hence, for each $U \in \mathcal{U}$ there exists a unique form f_U of degree n such that $L_{f_U} = U$. By construction, there exist generic points $U_{\pm\epsilon}$ in \mathcal{U} such that $U_{+\epsilon}$ contains a real-rooted form, and $U_{-\epsilon}$ contains no real-rooted form. Then $f_{+\epsilon} = f_{U_{+\epsilon}}$ has real rank $k+1$ while $f_{-\epsilon} = f_{U_{-\epsilon}}$ has real rank $\geq k+2$. We conclude that f is in $\partial\mathcal{R}_n$.

It remains to be seen that the Hankel determinant $\Delta_{\{2k\}}$ is not a factor of the real rank boundary $\partial_{\text{alg}}(\mathcal{R}_n)$. The proof is by contradiction. Suppose that $f \in \partial\mathcal{R}_n$ has $\text{rank}_{\mathbb{C}}(f) = k$, and write $f = \sum_{i=1}^k \ell_i^n$. The ℓ_i are linear forms over \mathbb{C} , and these are unique. We can approximate f by a sequence of forms in \mathcal{R}_n . Each is a sum of $k+1$ powers of *real* linear forms. A convergence argument shows that ℓ_1, \dots, ℓ_k must have real coefficients as well.

Fix a generic real linear form ℓ_0 and consider $f + \epsilon\ell_0^n$ and $f - \epsilon\ell_0^n$ where $\epsilon > 0$ is small. The square Hankel matrices corresponding to these two forms are invertible. Their determinants have opposite signs. This can be seen from the Matrix Determinant Lemma. The two forms lie on different sides of the Hankel hypersurface. Both have real rank $k+1$, and hence both lie in \mathcal{R}_n . This contradicts our assumption that f is in the topological boundary of \mathcal{R}_n . \square

Remark 4.5. The real projective space $\mathbb{P}_{\mathbb{R}}^n$ of binary forms of degree n is stratified by real rank. Let $r = \lceil (n+1)/2 \rceil$. Blekherman [3] showed that each integer in $\{r, r+1, \dots, n-1, n\}$ arises as the real rank of some open stratum. It would be very interesting to determine the algebraic boundary that separates real rank i from real rank $i+1$, for any $i \in \{r, \dots, n-1\}$. Currently, only the two extreme cases are known. The algebraic boundary for $i = n-1$ is the discriminant $\Delta_{\{1^{n-2}, 2\}}$, by [6, Prop. 3.1], and the case $i = r$ is resolved by Theorem 4.1.

5. Euclidean distance degrees

This section is concerned with the following optimization problem over the real field $K = \mathbb{R}$. Let $h \in \mathbb{R}[x, y]_n$ be a fixed binary form. We seek $f \in \Delta_\lambda$ that is closest to h . In symbols:

$$\text{minimize } \langle f - h, f - h \rangle \quad \text{subject to } f \in \Delta_\lambda. \quad (21)$$

We call this the *special ED problem* when the inner product is as in (3) and (4). Alternatively, we may also use a positive definite quadratic form that is generic. The resulting scenario (21) is the *generic ED problem* for Δ_λ . For instance, if we replace (4) with $\langle f, g \rangle = \sum_{i=0}^n \gamma_i a_i b_i$, where $\gamma_0, \gamma_1, \dots, \gamma_n$ are random positive reals, then this leads to a generic ED problem.

The map that takes the given input h to the optimal solution $f^* = f^*(h)$ of the problem (21) is an algebraic function. The degree of that algebraic function is known as the *ED degree* of the variety Δ_λ . For an introduction to this topic and its basic results see [8]. A follow-up study, aimed at varieties that admit a determinantal representation, was undertaken in [20].

In the algebraic approach to solving (21) one writes the critical equations using Lagrange multipliers and one removes the singular locus. We shall see concrete examples for Δ_λ shortly. The ED degree is the number of complex solutions to these equations for generic data h . Of course, the optimal point f^* is real and it is among these complex critical points.

Varieties such as Δ_λ come with a natural invariant coordinate system, and our special inner product (3)–(4) gives the standard Euclidean distance for that coordinate system. The corresponding ED degree is the *special ED degree* of Δ_λ . By contrast, the *generic ED degree* of Δ_λ is usually larger. This is the degree for the generic ED problem described above.

In the last column of Table 1 we list the special ED degree and the generic ED degree. The first pair in each box, corresponding to the rational normal curve $\Delta_{(n)}$, equals $n, 3n - 2$. These numbers were derived, for arbitrary n , in [8, Example 5.12] and [8, Corollary 8.7]. Both notions of ED degrees are preserved under duality [8, Theorem 5.2], and the generic ED degree coincides with the sum of the polar classes, $C_\lambda(1, 1) = \sum \delta_i$, by [8, Theorem 5.4].

Some of the ED degrees in Table 1 appear in [20]. For instance, the ED degrees 7, 13 for $\lambda = (2, 2)$ appear in [20, Example 1.1]. When λ is dual to a collection of hooks of the form n, n or n, n, n , the variety $(\Delta_\lambda)^\vee$ is given by Hankel matrices of rank 2 or 3. The corresponding ED degrees in Table 1 are found in [20, Table 4]. Therein, the left table for $\Lambda = \Omega_n$ gives generic ED degree, while the right table for $\Lambda = \Theta_n$ gives special ED degree.

From Table 1 we can guess the two ED degrees for the variety of binary forms of degree n that have a root of multiplicity a . Our next goal is to prove that this guess is correct.

Theorem 5.1. *For any hook $\lambda = \{1^{n-a}, a\}$, the special ED degree of the variety Δ_λ is always n , independently of a , whereas the generic ED degree of Δ_λ equals $(2a - 1)n - 2(a - 1)^2$.*

Proof. We begin with the generic ED degree. It is the sum of the polar classes δ_i . The codimension $a - 1$ of $\Delta_{\{1^{n-a}, a\}}$ equals the dimension of its dual $(\Delta_{\{1^{n-a}, a\}})^\vee =$

$\Delta_{\{1^{a-2}, n-a+2\}}$. This means that precisely two polar classes are nonzero, and the generic ED degree equals

$$\delta_{a-1} + \delta_a = \deg(\Delta_{\{1^{n-a}, a\}}) + \deg(\Delta_{\{1^{a-2}, n-a+2\}}) = a(n-a+1) + (n-a+2)(a-1), \quad (22)$$

by Hilbert's formula (1). This expression equals $(2a-1)n - 2(a-1)^2$, as desired.

Next consider the special ED degree. Our goal is to compute the complex critical points f^* of the optimization problem (21) where h is a fixed generic binary form in $V = \mathbb{R}[x, y]_n$. We now identify V with its dual space V^\vee by way of the distinguished inner product (3), and we regard the conormal variety Con_λ as an affine cone of dimension $n+1$ in $V \times V = \mathbb{R}^{2n+2}$.

By ED duality [8, §5], our problem is equivalent to solving the system of linear equations

$$f + g = h \quad \text{for } (f, g) \in \text{Con}_\lambda. \quad (23)$$

These $n+1$ inhomogeneous linear equations have finitely many complex solutions (f^*, g^*) on the $(n+1)$ -dimensional affine variety Con_λ . Their number is the special ED degree.

We now write the equations (23) using the parametrization of Con_λ given in Theorem 2.5:

$$h(x, y) = (tx - sy)^a \cdot g_1(x, y) + (sx + ty)^{n-a+2} \cdot g_2(x, y). \quad (24)$$

The two summand are unknown points in $\Delta_{\{1^{n-a}, a\}}$ and in $\Delta_{\{1^{a-2}, n-a+2\}}$. Both are now regarded as affine varieties in $V = \mathbb{C}^{n+1}$. The forms g_1 and g_2 are unknown and they have degrees $n-a$ and $a-2$ respectively. Furthermore, $(s : t)$ is an unknown point in \mathbb{P}^1 . Hence there are $n+1$ unknown parameters in total, to match the $n+1$ given coefficients of $h(x, y)$. Thus (24) is a square polynomial system in \mathbb{C}^{n+1} , and we need to count its solutions.

Any representation (24) over \mathbb{C} of the given binary form $h(x, y)$ will be called an *orthogonal decomposition of type a* . Thus, our proof reduces to establishing the following assertion: a general binary form of degree n has precisely n orthogonal decompositions of type a .

In what follows we describe the binary form whose zeros are the points $(s : t)$ that can occur in (24). In other words, we eliminate the unknown binary forms g_1 and g_2 from (24). Once $(s : t)$ is known, the coefficients of g_1 and g_2 can be recovered by solving a linear system of equations. So, it suffices to identify that binary form and to show it has degree n .

We define an endomorphism $L^{(k)}$ of the vector space $V = \mathbb{R}[x, y]_n$ as follows:

$$(L^{(k)}(f))(x, y) = \frac{(n-k)!}{n!} \sum_{i=0}^k (-1)^i \binom{k}{i} x^{k-i} y^i \frac{\partial^k}{\partial x^i \partial y^{k-i}} f(x, y).$$

This linear differential operator generalizes the SO_2 -invariant vector field

$$L^{(1)} = \frac{1}{n} \left[x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right]$$

and the second order operator

$$L^{(2)} = \frac{1}{n(n-1)} \left[x^2 \frac{\partial^2}{\partial y^2} - 2xy \frac{\partial^2}{\partial x \partial y} + y^2 \frac{\partial^2}{\partial x^2} \right].$$

The linear map $L^{(k)} : V \rightarrow V$ can be written explicitly in terms of coordinates as follows. If $f = \sum_{i=0}^n \binom{n}{i} a_i x^i y^{n-i}$ then the coefficients of $L^{(k)}(f) = \sum_{j=0}^n \binom{n}{j} b_j x^j y^{n-j}$ are

$$\binom{n}{j} b_j = \sum_{i=\max\{0, k-j\}}^{n-j} (-1)^i \binom{k}{i} \binom{n-k}{i+j-k} a_{2i+j-k}. \quad (25)$$

If we apply the n -th order operator then this amounts to a rotation by 90 degrees:

$$(L^{(n)}(f))(x, y) = f(-y, x). \quad (26)$$

Now, the relevance of the differential operator $L^{(k)}$ for our proof is as follows:

Lemma 5.2. *Let h be a binary form of degree n . Suppose that $(s : t) \in \mathbb{P}^1$ occurs in an orthogonal decomposition (24) of type a . Then $(s : t)$ is a root of the binary form $L^{(a-1)}(h)$.*

Lemma 5.2 will be proved further below. We first derive the theorem from the lemma.

The linear map $L^{(a-1)}$ is not zero. Hence, for generic h , the binary form $L^{(a-1)}(h)$ is nonzero and has degree n . Such a binary form has at most n distinct roots. Therefore, by Lemma 5.2, we know that the special ED degree is at most n . What we must prove is that the special ED degree is at least n . We do this by exhibiting, for every n and a , one particular binary form of degree n that has n distinct orthogonal decompositions of type a .

We begin with the observation that the following two binary forms have real coefficients:

$$\begin{aligned} h_n(x, y) &= \frac{1}{2} ((x + \sqrt{-1} \cdot y)^n + (x - \sqrt{-1} \cdot y)^n), \\ k_n(x, y) &= \frac{1}{2\sqrt{-1}} ((x + \sqrt{-1} \cdot y)^n - (x - \sqrt{-1} \cdot y)^n). \end{aligned} \quad (27)$$

These two forms are invariant under rotation by $2\pi/n$, i.e. they are fixed by the endomorphism $\rho : V \rightarrow V$ that maps x to $(\cos \frac{2\pi}{n})x - (\sin \frac{2\pi}{n})y$ and y to $(\sin \frac{2\pi}{n})x + (\cos \frac{2\pi}{n})y$.

Case 1. Suppose that n is odd. For our special binary form we take

$$h_n(x, y) = x^n - \binom{n}{2}x^{n-2}y^2 + \binom{n}{4}x^{n-4}y^4 - \cdots \pm \binom{n}{n-1}xy^{n-1}.$$

Note that all the exponents of x are odd while that of y are even. For odd a we have

$$h = x^a g_1 + y^{n-a+2} g_2, \quad (28)$$

for suitable forms g_1 and g_2 depending on a . Similarly, for even a we have

$$h = y^a g_1 + x^{n-a+2} g_2.$$

For each decomposition, we obtain $n-1$ others by acting with the rotations $\rho, \rho^2, \dots, \rho^{n-1}$.

Case 2. Suppose that n even and a is even. We also take

$$h_n(x, y) = x^n - \binom{n}{2}x^{n-2}y^2 + \binom{n}{4}x^{n-4}y^4 - \cdots \pm y^n.$$

Then (28) also holds. Now, h_n is fixed by $\rho^{n/2}$, so applying the rotations ρ^i gives only $n/2$ distinct orthogonal decompositions of type a . However, we have $h_n(x, y) = \pm h_n(y, x)$, so by permuting the two variables we get additional decompositions. These are not equal to any of the previous ones. In total, this yields $\frac{n}{2} \cdot 2 = n$ distinct decompositions (24) for h_n .

Case 3. Suppose that n even and a is odd. In that case we take

$$k_n(x, y) = \binom{n}{1}x^{n-1}y - \binom{n}{3}x^{n-3}y^3 + \binom{n}{5}x^{n-5}y^5 - \cdots \pm \binom{n}{n-1}xy^{n-1},$$

and the argument is the same as in Case 2. This completes the proof of [Theorem 5.1](#). \square

We now turn to [Lemma 5.2](#). Note that this result is familiar in the case $a = 2$, when (21) asks for the best rank 1 approximation of a symmetric $2 \times 2 \times \cdots \times 2$ tensor h . Finding that approximation amounts to computing the *eigenvectors* of h ; see [8, [Corollary 8.7](#)]. However, by definition, the eigenvectors of h are the roots of $L^{(1)}(h) = (1/n) \cdot \det \begin{pmatrix} x & y \\ \partial h / \partial x & \partial h / \partial y \end{pmatrix}$.

Proof of Lemma 5.2. We claim that the operator $L^{(k)}$ satisfies the following identity

$$L^{(k)}(f(x, y)) = (-1)^{n-k} \cdot L^{(n-k)}(f(-y, x)) \quad \text{for } k = 0, 1, \dots, n. \quad (29)$$

The special case $k = n$ is (26). One can show that (29) holds by a direct computation, using the formula (25). Equivalently, we check that it holds for monomials and extend by linearity.

Suppose now that $(s : t) \in \mathbb{P}^1$ occurs in an orthogonal decomposition (24) of type a . We apply the differential operator $L^{(a-1)}$ to both sides of that equation. This implies

$$L^{(a-1)}(h(x, y)) = L^{(a-1)}((tx - sy)^a g_1(x, y)) + L^{(a-1)}((sx + ty)^{n-a+2} g_2(x, y)).$$

Applying (29) to the summand on the right, we conclude

$$\begin{aligned} L^{(a-1)}(h(x, y)) &= L^{(a-1)}((tx - sy)^a g_1(x, y)) \\ &\quad + (-1)^{n-a+1} L^{(n-a+1)}((tx - sy)^{n-a+2} g_2(-y, x)). \end{aligned}$$

In both summands, a k -th order differential operator $L^{(k)}$ is applied to a binary form that has $(s : t)$ as a root of multiplicity at least $k + 1$. Both of the resulting binary forms still have $(s : t)$ among their roots. This shows that $L^{(a-1)}(h(x, y))$ vanishes at $(s : t)$. \square

The difference between the two ED degrees in Theorem 5.1 is $2(n - a + 1)(a - 1)$. We shall explain this number and how to think about the operator $L^{(a-1)}$. If we fix values for the parameters s and t , then the equation (24) translates into an inhomogeneous linear system of equations whose unknowns are the coefficients of g_1 and g_2 . We have $n + 1$ linear equations in $n = (n - a + 1) + (a - 1)$ unknowns. There is no solution for generic s and t .

We write our inhomogeneous linear system as an $(n + 1) \times (n + 1)$ -matrix $\mathcal{M}_{n,a}$. The first row consists of the coefficients of h . The next $n - a + 1$ rows contain the monomials of degree a in (s, t) , and the last $a - 1$ rows contain the monomials of degree $n - a + 2$ in (s, t) . We seek row vectors of the form $(-1, g_1, g_2)$ that lie in the left kernel of $\mathcal{M}_{n,a}$. The matrix has a banded structure, like Sylvester's matrix for the resultant. For instance, for $\lambda = (3, 1)$,

$$\mathcal{M}_{4,3} = \begin{pmatrix} h_0 & h_1 & h_2 & h_3 & h_4 \\ -s^3 & 3s^2t & -3st^2 & t^3 & 0 \\ 0 & -s^3 & 3s^2t & -3st^2 & t^3 \\ t^3 & 3st^2 & 3s^2t & s^3 & 0 \\ 0 & t^3 & 3st^2 & 3s^2t & s^3 \end{pmatrix}.$$

In order for our linear equations to have a solution, the determinant of $\mathcal{M}_{n,a}$ must be zero.

The degree of $\det(\mathcal{M}_{n,a})$ in (s, t) is precisely the generic ED degree. This is best seen from (22). If we replace (3) by a generic inner product then the rows for g_1 change by a linear transformation, whilst the rows for g_2 change by the inverse linear transformation. Thus, for the generic ED problem, the critical points are precisely the roots of the determinant.

However, for the special ED problem, our determinant admits the following factorization:

$$\det(\mathcal{M}_{n,a}) = \pm(L^{(a-1)}(h))(s, t) \cdot (s^2 + t^2)^{(n-a+1)(a-1)}. \quad (30)$$

The binary form $s^2 + t^2$ vanishes if and only if the last n rows of $\mathcal{M}_{n,a}$ are linearly dependent. The degree of the extraneous factor is $2(n-a+1)(a-1)$, and subtracting that from the generic ED degree gives n . The remaining factor is the remarkable binary form $L^{(a-1)}(h)$.

We close this section with a conjecture, namely that the two ED degrees of Δ_λ always have the same parity. According to Table 1, this holds for all partitions with $n \leq 7$. For hooks, it is proved by Theorem 5.1, and the underlying reason is seen clearly in (30). In general, the extraneous components should come from isotropic quadrics like $s^2 + t^2$, and parallelities like $sv - tu$. These quadrics suggest that the difference in ED degrees is even for all λ . This conjecture is related to [20, Eqn. (3.5)]. At present we do not know how to prove it.

6. ED duality in action

We now illustrate how our results can be applied to find exact solutions to the optimization problem (21). Following [8,20], our approach is to compute all critical points and then select the best real critical point. By ED duality [8, §5], the critical points are found by solving linear equations on the conormal variety. Given h , we need to compute all decompositions

$$h(x, y) = f(x, y) + g(x, y) \quad \text{where } (f, g) \in \text{Con}_\lambda. \quad (31)$$

If h is generic then the number of such decompositions is the special ED degree. The binary forms f that arise in the decompositions (31) are precisely the critical points on Δ_λ for the Euclidean distance to h , and similarly the forms g are the critical points on its dual $(\Delta_\lambda)^\vee$. The proximity of the solution is reversed under duality because $\|h - f\|^2 + \|h - g\|^2 = \|h\|^2$. This follows from $h = f + g$ and $\langle f, g \rangle = 0$. In particular, if f is the closest point to h among those on Δ_λ , then it is paired with the farthest critical point g on $(\Delta_\lambda)^\vee$, and vice versa.

In what follows we illustrate how one might solve (31) and hence (21) in practice. We begin with $n = 5$ and $\lambda = (3, 1, 1)$. For our given data point we take the binary quintic

$$h(x, y) = \frac{1}{2}(x + \sqrt{-1} \cdot y)^5 + \frac{1}{2}(x - \sqrt{-1} \cdot y)^5 + y^5 = x^5 - 10x^3y^2 + 5xy^4 + y^5. \quad (32)$$

This is a slight variant of (27). The primal problem is to find the closest quintic $f \in \Delta_{(3,1,1)}$ with a triple root, and the dual problem is to find the closest quintic $g \in \Delta_{(4,1)}$ with a quadruple root. By Theorem 5.1, the equation (31) has five solutions on Con_λ . They are

$ g ^2$	$f = h - g$	$g = h - f$	$ f ^2$
3.02	$(1.471x^2 - 5.582xy + 3.585y^2)(x + 0.785y)^3$	$(-1.238x - 0.735y)(0.785x - y)^4$	13.98
4.92	$(-0.263x^2 - 0.211xy - 0.338y^2)(x - 3.132y)^3$	$(0.0131x - 0.403y)(3.132x + y)^4$	12.08
5.09	$(-0.263x^2 + 0.167xy + 0.0346y^2)(x + 3.020y)^3$	$(0.0152x + 0.468y)(3.020x - y)^4$	11.91
6	$(x^2 - 10y^2)x^3$	$(5x + y)y^4$	11
8.12	$(1.473x^2 + 5.397xy + 2.867y^2)(x - 0.6732y)^3$	$(-2.301x + 1.875y)(0.6732x + y)^4$	8.88

The upper left quintic f in $\Delta_{(3,1,1)}$ is closest to h , at distance $3.0215666805997121633^{1/2}$, with triple root $(-0.78519451639408253233 : 1)$. The lower right quintic g in $\Delta_{(4,1)}$ is closest to h , at distance $8.880827761458859783^{1/2}$, with quadruple root $(-1 : 0.67321557299682647408)$. We could have guessed the decomposition $h = f + g$ in the fourth row from the input (32). This one is indeed a critical point, but it is neither primal optimal nor dual optimal.

Our five solutions to (31) were found by using the matrix that was introduced in Section 5:

$$\mathcal{M}_{5,3}(s, t) = \begin{pmatrix} 1 & 5 & 0 & -10 & 0 & 1 \\ -s^3 & 3s^2t & -3st^2 & t^3 & 0 & 0 \\ 0 & -s^3 & 3s^2t & -3st^2 & t^3 & 0 \\ 0 & 0 & -s^3 & 3s^2t & -3st^2 & t^3 \\ t^4 & 4st^3 & 6s^2t^2 & 4s^3t & s^4 & 0 \\ 0 & t^4 & 4st^3 & 6s^2t^2 & 4s^3t & s^4 \end{pmatrix}.$$

The triple (resp. quadruple) roots $(s : t) \in \mathbb{P}^1$ of f (resp. g) are the roots of the binary form

$$\det(\mathcal{M}_{5,3}(s, t)) = (s^2 + t^2)^6 \cdot (s^5 - 10s^3t^2 - s^2t^3 + 5st^4) = -(s^2 + t^2)^6 \cdot (L^{(2)}(h))(s, t).$$

We compute the five real roots numerically, and at each of them we compute the left kernel of $\mathcal{M}_{5,3}(s, t)$. The result of that computation is precisely the list of five pairs (f, g) above.

This method scales well for hooks $\lambda = \{1^{n-a}, a\}$. Here, the special ED degree is always n , and Lemma 5.2 furnishes the minimal polynomial for the desired a -fold root of $f \in \Delta_\lambda$. The matrix $\mathcal{M}_{n,a}(s, t)$ represents our task as a *homogeneous polynomial eigenvalue problem*, and this makes it amenable to well-developed methods of numerical linear algebra; see e.g. [7].

For an illustration we fix $n = 15$ and $a = 6$. The data point is the binary form

$$h(x, y) = \sum_{i=0}^{15} \binom{15}{i} u_i x^i y^{15-i}, \quad \text{with randomly chosen coefficient vector}$$
$$(u_0, \dots, u_{15}) = (20, -17, 3, 16, 12, 14, -16, -5, 7, 8, -13, 5, -13, -16, 7, -11).$$

Among forms f of degree 15 with a root of multiplicity 6, the following is the closest to h :

$$\sum_{i=0}^{15} \binom{15}{i} v_i x^i y^{15-i} = 10^{-6} (x - 8.70886y)^6 (131903x^9 + 11375.9x^8y - \cdots - 552.901xy^8 + 45.8419y^9),$$

$$(v_0, v_1, \dots, v_{15}) = (20, -17, 3, 16, 12, 14, -16, -5, 7.04, 8.19, -12.17, 8.09, -3.78, 1.42, -0.46, 0.13).$$

All 15 critical points for this optimization problem are real because $L^{(5)}(h)$ is real-rooted for our h . Its 15 roots have the form $(s : 1)$ where $s \in \mathbb{R}$. They appear in the first column of

Root s of $L^{(5)}(h)$	Distance ²	Local
8.70886	86 791	min
3.70567	111 796	min
2.19850	163 470	max
0.05736	476 068	?
-0.38870	550 056	?
-3.49092	564 363	min
-5.71229	565 936	min
-0.22118	657 621	?
1.25359	723 240	max
0.25811	727 831	?
0.48187	774 941	?
0.80694	934 884	max
-0.68808	1 058 800	max
-1.67383	1 150 260	max
-1.06515	1 256 200	max

By computing the left kernel of $\mathcal{M}_{15,6}(s, t)$ at each root, we find the 15 decompositions (31). The squared norms $\|g\|^2$ of the dual solutions $g = h - f$ are listed in the second column. So, the first row gives the optimal solution for $\lambda = \{1^9, 6\}$, while the last row gives the optimal solution for the dual problem $\mu = \{1^4, 11\}$. Local optima are indicated in the third column. There are four local optima on Δ_λ (marked with “min”) and six local optima on Δ_μ (marked with “max”). These were certified using the signature of the Hessian of the distance function. The local nature of the other five points cannot be decided with the second-order criterion because their Hessians are singular, in both the primal and the dual formulation.

Our optimization problem is more challenging when λ is not a hook. A small interesting case is the partition $\lambda = (3, 2)$. Let us see what happens here if we take the same input h as in (32). We seek a binary form $f \in \Delta_{(3,2)}$ with a triple root and a double root that is closest to h . The desired decomposition (31) on the conormal variety $\text{Con}_{(3,2)}$ now takes the form

$$h(x, y) = \alpha(tx - sy)^3(vx - uy)^2 + (\beta x + \gamma y)(sx + ty)^4 + \delta(ux + vy)^5.$$

This means that the vector $(-1, \alpha, \beta, \gamma, \delta)$ lies in the left kernel of the 5×6 -matrix

$$\begin{pmatrix} 1 & 5 & 0 & -10 & 0 & 1 \\ -s^3u^2 & 2s^3uv + 3s^2tu^2 & -6s^2tuv - s^3v^2 - 3st^2u^2 & 3s^2tv^2 + 6st^2uv + t^3u^2 & -3st^2v^2 - 2t^3uv & t^3v^2 \\ 0 & t^4 & 4st^3 & 6s^2t^2 & 4s^3t & s^4 \\ t^4 & 4st^3 & 6s^2t^2 & 4s^3t & s^4 & 0 \\ v^5 & 5uv^4 & 10u^2v^3 & 10u^3v^2 & 5u^4v & u^5 \end{pmatrix}.$$

All 5×5 minors of this matrix must be 0. However, the ideal of 5×5 minors has some extraneous associated primes that must be removed. This is analogous to the factor s^2+t^2 in the hook case, but more complicated, so we do not pursue that primary decomposition.

Instead we simply work directly with the squared Euclidean distance function

$$D = \| h - \alpha(tx - sy)^3(vx - uy)^2 \|^2,$$

and we solve the following system of polynomial equations in five unknowns:

$$\frac{\partial D}{\partial \alpha} = \frac{\partial D}{\partial s} = \frac{\partial D}{\partial t} = \frac{\partial D}{\partial u} = \frac{\partial D}{\partial v} = 0 \quad \text{and} \quad sv - ut \neq 0.$$

This has 20 complex solutions, while the ED degree of $\Delta_{(3,2)}$ is 21. One checks that h lies in the ED discriminant [8, §7]. Only 4 of the 20 solutions are real. Setting $t = v = 1$, they are

α	s	u	distance ²
1.817238	0.673272	−1.316853	7.724678
−0.266252	−3.020572	0.274673	8.643701
−0.265424	3.131909	−0.387044	12.017703
1.815280	−0.785143	1.428712	13.105926

The worst critical point is given in the bottom row. The corresponding quintic is

$$f = 1.81528x^5 - 0.911262x^4y - 5.15521x^3y^2 + 0.0137459x^2y^3 + 4.34203xy^4 + 1.79341y^5.$$

ED duality gives us the optimal solution $g = h - f$ for the dual problem on $(\Delta_{(3,2)})^\vee$, namely

$$\begin{aligned} g &= -0.81528x^5 + 0.911262x^4y - 4.84479x^3y^2 - 0.0137459x^2y^3 \\ &\quad + 0.657973xy^4 - 0.793414y^5 \\ &= (\beta x + \gamma y)(-0.785143x + y)^4 + \delta(1.428712x + y)^5 \end{aligned}$$

for suitable real constants β, γ , and δ , which can be found by solving a linear system.

The form g is very interesting for the application described in Section 4. Recall that the generic tensor rank for binary quintics is 3. The unique rank 3 decomposition of the given quintic h is complex. It is shown on the left in (32). One might therefore ask for a best approximation to h that has real rank 3. This question is not well-posed because the set \mathcal{R}_5 from Section 4 is not closed. Instead one should ask for the closest binary quintic in the closure of \mathcal{R}_5 , i.e. among those whose *border rank* equals 3. That optimal quintic must be our g , because it is a critical point on real rank boundary $\partial_{\text{alg}}(\mathcal{R}_5)$, which equals the little apple hypersurface $(\Delta_{(3,2)})^\vee$. The fact that g has border rank 3 is verified by the representation

$$g = \lim_{\epsilon \rightarrow 0} \left(\frac{1}{5\epsilon} ((-0.785143 + \epsilon\beta)x + (1 + \epsilon\gamma)y)^5 - \frac{1}{5\epsilon} (-0.785143x + y)^5 + \delta(1.428712x + y)^5 \right).$$

This computation offers a concrete illustration of [Theorem 4.1](#) and [Example 4.2](#).

Acknowledgments

We thank Michael Burr for a conversation in March 2015 that ignited our interest in the ED problem for multiple root loci. We are grateful to Greg Blekherman for his help in getting [Theorem 4.1](#) into final shape. This project was carried out in the summer of 2015, when both authors visited the National Institute of Mathematical Sciences, Daejeon, Korea. Hwangrae Lee was supported by the National Research Foundation of Korea (MSIP 2011-0030044). Bernd Sturmfels was supported by the National Science Foundation (DMS-1419018).

References

- [1] A. Abdesselam, J. Chipalkatti, Brill–Gordan loci, transvectants and an analogue of the Foulkes conjecture, *Adv. Math.* 208 (2007) 491–520.
- [2] A. Abdesselam, J. Chipalkatti, The bipartite Brill–Gordan locus and angular momentum, *Transform. Groups* 11 (2006) 341–370.
- [3] G. Blekherman, Typical real ranks of binary forms, *Found. Comput. Math.* 15 (2015) 793–798.
- [4] G. Blekherman, P. Parrilo, R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Ser. Optim., vol. 13, 2012.
- [5] J. Chipalkatti, On equations defining coincident root loci, *J. Algebra* 267 (2003) 246–271.
- [6] P. Comon, G. Ottaviani, On the typical rank of real binary forms, *Linear Multilinear Algebra* 60 (2012) 657–667.
- [7] J.-P. Dedieu, F. Tisseur, Perturbation theory for homogeneous polynomial eigenvalue problems, *Linear Algebra Appl.* 358 (2003) 71–94.
- [8] J. Draisma, E. Horobeț, G. Ottaviani, B. Sturmfels, R. Thomas, The Euclidean distance degree of an algebraic variety, *Found. Comput. Math.* (2015) 1–51.
- [9] L.M. Fehér, A. Némethi, R. Rimányi, Coincident root loci of binary forms, *Michigan Math. J.* 54 (2006) 375–392.
- [10] I.M. Gel’fand, M.M. Kapranov, A.V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.
- [11] D. Hilbert, Singularitäten der Diskriminantenfläche, *Math. Ann.* 30 (1887) 437–441.
- [12] A. Iarrobino, V. Kanev, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, *Lecture Notes in Math.*, vol. 1721, Springer-Verlag, Berlin, 1999.
- [13] E. Kaltofen, J. May, Z. Yang, L. Zhi, Approximate factorization of multivariate polynomials using singular value decomposition, *J. Symbolic Comput.* 43 (2008) 359–376.
- [14] G. Katz, How tangents solve algebraic equations, or a remarkable geometry of discriminant varieties, *Expo. Math.* 21 (2003) 219–261.
- [15] S. Kurmann, Some remarks on equations defining coincident root loci, *J. Algebra* 352 (2012) 223–231.
- [16] J. McNamee, *Numerical Methods for Roots of Polynomials, Part I*, *Stud. Comput. Math.*, vol. 14, Elsevier B.V., Amsterdam, 2007.
- [17] E. Miller, B. Sturmfels, *Combinatorial Commutative Algebra*, *Grad. Texts in Math.*, Springer Verlag, New York, 2004.
- [18] F. Napolitano, Topology of complements of strata of the discriminant of polynomials, *C. R. Acad. Sci. Paris, Sér. I* 327 (1998) 665–670.

- [19] L. Oeding, Hyperdeterminants of polynomials, *Adv. Math.* 231 (2012) 1308–1326.
- [20] G. Ottaviani, P.-J. Spaenlehauer, B. Sturmfels, Exact solutions in structured low-rank approximation, *SIAM J. Matrix Anal. Appl.* 35 (2014) 1521–1542.
- [21] B. Sturmfels, The Hurwitz form of a projective variety, [arXiv:1410.6703](https://arxiv.org/abs/1410.6703).
- [22] J. Weyman, The equations of strata for binary forms, *J. Algebra* 122 (1989) 244–249.
- [23] Z. Zeng, Computing multiple roots of inexact polynomials, *Math. Comp.* 74 (2005) 869–903.
- [24] Z. Zeng, Regularization and matrix computation in numerical polynomial algebra, in: *Approximate Commutative Algebra*, in: *Texts Monogr. Symbol. Comput.*, Springer, Vienna, 2009, pp. 125–162.