Journal of Combinatorial Theory, Series A 117 (2010) 248–271

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series A

www.elsevier.com/locate/jcta

Gale duality bounds for roots of polynomials with nonnegative coefficients

Julian Pfeifle [1]

Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, C/ Jordi Girona, 31, Barcelona, Spain

ARTICLE INFO

Article history:
Received 13 February 2008
Available online 11 November 2009

Keywords:
Gale diagrams
Location of roots
Nonreal roots of polynomials
Ehrhart polynomials
Chromatic polynomials
Linear relations among coefficients of a polynomial

ABSTRACT

We bound the location of roots of polynomials that have nonnegative coefficients with respect to a fixed but arbitrary basis of the vector space of polynomials of degree at most d. For this, we interpret the basis polynomials as vector fields in the real plane, and at each point in the plane analyze the combinatorics of the Gale dual vector configuration. This approach permits us to incorporate arbitrary linear equations and inequalities among the coefficients in a unified manner to obtain more precise bounds on the location of roots. We apply our technique to bound the location of roots of Ehrhart and chromatic polynomials. Finally, we give an explanation for the clustering seen in plots of roots of random polynomials.

1. Introduction

The Ehrhart polynomial of a d-dimensional lattice polytope Q is a real polynomial of degree d, which has the following two representations:

$$i_Q = i_Q(z) = \sum_{j=0}^{d} c_j z^j = \sum_{i=0}^{d} a_i \binom{z+d-i}{d}.$$

Here we chose the letter z for the independent variable in order to emphasize that we think of i_Q as a polynomial defined over the complex numbers. The coefficients c_0, c_{d-1} and c_d in the first representation are positive, while the others generally can vanish or take on either sign. In contrast,

E-mail address: julian.pfeifle@upc.edu.

[1] Partially supported by MEC grants MTM2008-03020 and MTM2009-07242 of the Spanish Ministry of Science, and SGR grant 2009-SGR-1040 of the Generalitat de Catalunya.

0097-3165/$ – see front matter
doi:10.1016/j.jcta.2009.10.009

a famous theorem of Stanley [11] asserts that all coefficients a_i of i_Q in the latter representation are nonnegative, $a_i \geqslant 0$ for $0 \leqslant i \leqslant d$.

Such nonnegativity information is also available for other combinatorially defined polynomials, a case in point being the chromatic polynomial of a graph (cf. Proposition 6.1 below). An early example of how combinatorial information might be gleaned from studying roots of such polynomials is the *Birkhoff–Lewis Conjecture*, which asserts that no chromatic polynomial has a root in the real interval $[4, \infty)$. Somewhat ironically, even though it was formulated as a new inroad towards settling the Four Color Conjecture (which it implies), the latter is now a theorem, while the former is still open. Nevertheless, since at least 1965 [7], the complex roots of chromatic polynomials have received close scrutiny. A well-known recent result by Sokal [10] states that their complex roots are dense in the entire complex plane, if one allows arbitrarily large graphs. He was motivated by applications in physics to the *Potts model partition function*.

Coming back to Ehrhart polynomials, first bounds obtained in [1] on the location of the roots of i_Q for fixed d were substantially improved by Braun [3] and Braun and Develin [4]. All of these papers use the nonnegativity of the a_i's, but Braun's crucial new insight is to think of the value $i_Q(z)$ at each $z \in \mathbb{C}$ as a linear combination with nonnegative coefficients of the $d+1$ complex numbers $b_i = b_i(z) = \binom{z+d-i}{d}$. In particular, for z_0 to be a zero of i_Q, there must be a nonnegative linear combination of the $b_i(z_0)$ that sums to zero.

In this paper, we extend and generalize Braun's bounds on the location of roots for the binomial coefficient basis. We propose a unified approach using *Gale duality* to bound the location of roots, that

- works in exactly the same way for *all* bases of the vector space P_d of polynomials of degree at most d (Theorem 3.3), and
- allows one to incorporate arbitrary additional linear equations and inequalities between the coefficients a_i beyond mere nonnegativity (Theorem 5.1). This is applied in Section 6 to the case of Ehrhart and chromatic polynomials (Figs. 6 and 7).

We apply our approach in Section 3 to explicitly bound the location of the roots of polynomials with nonnegative coefficients with respect to four common bases of P_d; the detailed treatment of the binomial coefficient basis comprises Section 4. Throughout, we focus on bounding the location of the nonreal roots, as the case of real roots is much more straightforward (Observation 3.4).

In Section 7, we use our method to explain why the roots of "random" polynomials with nonnegative coefficients (for a suitable meaning of "random") tend to clump together, by tracing this behavior back to properties of the basis polynomials (Figs. 9 and 10).

1.1. Sketch of the method

Let $B = \{b_0, \dots, b_d\}$ be any basis of P_d, the $(d+1)$-dimensional vector space of real polynomials of degree at most d in one variable.

- We regard B as a collection of *vector fields*: for each complex number $z \in \mathbb{C}$, the basis elements $b_0(z), \dots, b_d(z)$ define a configuration $\mathcal{B}(z) = (w_0(z), \dots, w_d(z))$ of real vectors $w_j(z) = (\operatorname{Re} b_j(z), \operatorname{Im} b_j(z))^T$ in the plane $\mathbb{R}^2$. This point of view converts the algebraic problem of bounding the location of roots of a polynomial into a combinatorial problem concerning the discrete geometry of vector configurations.
- We analyze the combinatorics of $\mathcal{B}(z)$ in terms of the Gale dual configuration $\mathcal{B}^*(z)$. In particular, there exists a polynomial $f = \sum_{i=0}^{d} a_i b_i(z)$ with nonnegative coefficients $a_i \geqslant 0$ and a root at $z = z_0$ whenever the vector configuration $\mathcal{B}(z_0)$ has a nonnegative circuit, and this occurs whenever $\mathcal{B}^*(z_0)$ has a nonnegative cocircuit.
The important point is that we obtain a semi-explicit expression for $\mathcal{B}^*$ for *any* basis of P_d, not just the binomial coefficient basis. In fact, for the power basis $b_i = z^i$, the rising and falling

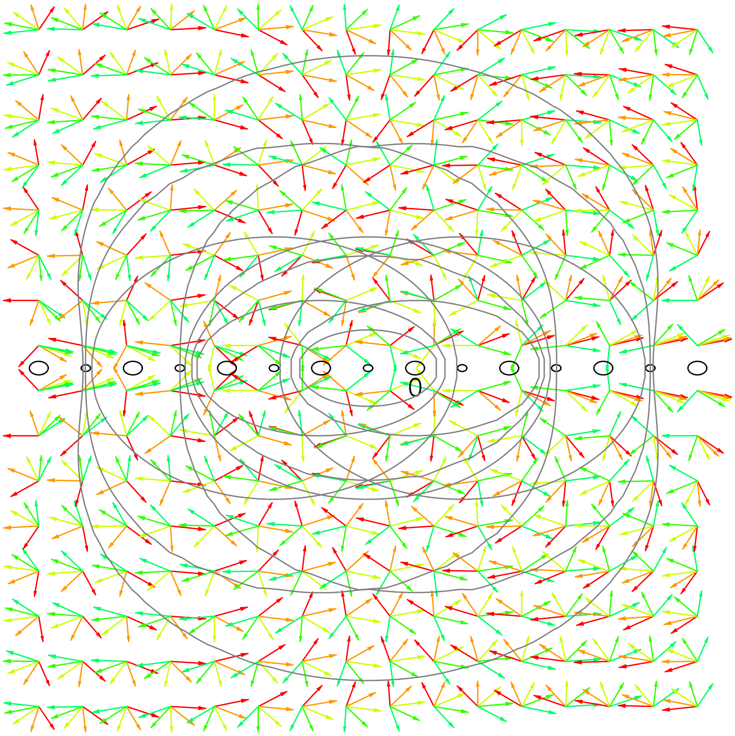

Fig. 1. The values of $\{\binom{z+d-i}{d}\colon\ 0 \leqslant i \leqslant d\}$ at different points in the complex plane, for $d = 6$. All vectors are normalized to the same length. In gray, the locus of points where two vectors become collinear. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

factorial bases $b_i = z^{\bar{i}}, z^{\underline{i}}$, and the binomial coefficient basis $b_i = \binom{z+d-i}{d}$ we can make the Gale dual completely explicit.

- In concrete situations one often has more information about f. Gale duality naturally allows to incorporate any linear equations and inequalities on the coefficients, and in some cases this leads to additional restrictions on the location of roots.
- As an illustration, we show how the inequality $a_d \leqslant a_0 + a_1$ that is valid for Ehrhart polynomials further constrains the location of the roots of i_Q. We also study the case of chromatic polynomials, for which Brenti [5] has shown the nonnegativity of the coefficients with respect to the binomial coefficient basis.
- Braun and Develin [4] derive an implicit equation for a curve $\mathcal{C}$ bounding the possible locations of roots of $f = \sum_{i=0}^{d} a_i \binom{z+d-i}{d}$, and our method gives an explicit equation for a real algebraic curve whose outermost oval is precisely $\mathcal{C}$.

It is instructive to visualize the vector fields $w_0, \dots, w_d$ for the binomial coefficient basis, i.e., when $b_j(z) = \binom{z+d-j}{d} = R_j(z) + iI_j(z)$; recall that $w_j(z) = (R_j(z), I_j(z))^T$.

From Fig. 1, it appears that at points far away from the origin the vectors w_i are all "acute", i.e., contained in a half-plane (that varies from point to point), while closer to the origin they positively span the entire space. If true in general, this would imply that *far away from the origin, f cannot have any roots.*

The detailed analysis (and proof) of this observation will take up the bulk of the paper, Sections 2–6, and in this special case may be summarized as follows:

Theorem 4.11. *Let d be a positive integer and $\mathcal{Z}_d$ the set of complex, nonreal numbers that are zeros of non-identically vanishing polynomials of the form*

$$f(z) = \sum_{j=0}^{d} a_j \binom{z+d-j}{d},$$

with $a_j \geqslant 0$ for $j = 0, \dots, d$. Then $\mathcal{Z}_d$ is the set of nonreal points in the region bounded by the outermost oval of the real algebraic curve of degree $d-1$ in the complex plane with equation

$$\frac{\binom{z}{d}\overline{\binom{z+d}{d}} - \overline{\binom{z}{d}}\binom{z+d}{d}}{z-\bar{z}} = 0,$$

where $\bar{\cdot}$ denotes complex conjugation. This bound is tight, in the sense that any point inside $\mathcal{Z}_d$ is a root of some such f. Moreover, there is an explicit representation of this equation as the determinant of a tridiagonal matrix; *see* (3) *and Proposition* 2.4.

The real roots of any such f all lie in the real interval $[-d, d-1]$.

From contemplating Fig. 1, a naive strategy for bounding the locations of the roots comes to mind: First, try to prove that for "far away" z the $w_i(z)$ positively span a convex pointed 2-dimensional cone τ. Then determine the generators $w_k(z)$, $w_l(z)$ of its facets, and the locus $\mathcal{C}$ of all $z \in \mathbb{C}$ for which these facet vectors "tip over", i.e., become collinear. By continuity, for z_0 inside $\mathcal{C}$ the origin is a nonnegative linear combination of the w_i, and thus z_0 is a possible root.

The alert reader will perhaps have lost track of even the *number* of holes in this argument! As a sample, it is a priori not clear (but true, at least for the binomial coefficient basis) that the $w_i(z)$ in fact span a pointed cone for *all* z of large enough absolute value. It is even less clear (but true in this case) that the vectors spanning facets of τ far away from the origin will still define facets just before τ ceases to be convex closer to the origin. Furthermore, the locus $\mathcal{C}$ might (and does) have multiple components, suggesting that one has to exercise more care when talking about points z_0 "inside" $\mathcal{C}$.

However, the real problem with this approach lies with the fact that the locus of collinearity of $w_i(z)$ and $w_j(z)$ is the vanishing locus of the determinant $\Delta_{ij} = \begin{vmatrix} R_i & R_j \\ I_i & I_j \end{vmatrix}$, and evaluating this polynomial explicitly quickly becomes a daunting task; moreover, it is not at all clear how the knowledge of Δ_{ij} for any particular basis would help for other bases of P_d.

We now present our method that overcomes all these obstacles.

2. Gale duality

2.1. Overview

Consider a polynomial $f = \sum_{i=0}^{d} a_i b_i$ of degree d, expanded with respect to a basis $B = \{b_0, \dots, b_d\}$ of P_d, the vector space of all polynomials in one complex variable of degree at most d. For the moment, we will focus on the complex, nonreal roots of f. To find these, rewrite the real and complex parts of the condition $f(z) = 0$ in the form

$$\begin{pmatrix} R_0 & R_1 & \cdots & R_d \\ I_0 & I_1 & \cdots & I_d \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix} = 0, \tag{1}$$

where $R_j = R_j(x, y)$ and $I_j = I_j(x, y)$ stand for the real and imaginary parts of the polynomial $b_j(x + iy)$.

As suggested in the Introduction, we now regard each basis element b_i not as a complex polynomial, but as a real vector $w_i(x, y) = (R_i, I_i)^T \in \mathbb{R}^2$. Then there exists some polynomial f with a root at $z = x + iy$ if and only if there exist real coefficients $a_0, \ldots, a_d$ with

$$\sum_{i=0}^{d} a_i w_i(x, y) = 0.$$

If we impose the additional restriction that the a_i be nonnegative but not all zero, this is only possible if the positive span of the w_i includes the origin. Among all such linear combinations summing to zero, we now consider only *support-minimal* ones, i.e., those with the minimum number of nonzero coefficients a_i. In oriented matroid terminology, the ordered collection σ of signs of the coefficients of such a support-minimal linear combination is called a *circuit* of the (full-dimensional) vector configuration $W = (w_0, \ldots, w_d) \subset \mathbb{R}^2$. To proceed, we regard W as a $(2 \times (d + 1))$-matrix. A *Gale dual* vector configuration $\bar{W} = (\bar{w}_0, \ldots, \bar{w}_d) \subset \mathbb{R}^{d-1}$ of W is the ordered set of rows of any matrix, also called $\bar{W}$, whose columns form a basis for the (row) kernel of the matrix W, so that $W\bar{W} = 0$ [12]. Gale duality states that the collection of signs σ is a *cocircuit* of $\bar{W}$. This means that there exists a linear form g on $\mathbb{R}^{d-1}$ with $(\operatorname{sign} g(\bar{w}_i)\colon\ i = 0, \ldots, d) = \sigma$.

Clearly, any circuit of W has either two or three nonzero entries (unless it is the zero circuit, which we exclude from the discussion). Because z_0 is a root of f if and only if there exists a nonnegative circuit of $W(z_0) = (w_0(z_0), \ldots, w_d(z_0)) \subset \mathbb{R}^2$, by Gale duality this happens if and only if there exists a cocircuit of $\bar{W}(z_0) = (\bar{w}_0(z_0), \ldots, \bar{w}_d(z_0)) \subset \mathbb{R}^{d-1}$, i.e., if and only if there exists a linear form on $\mathbb{R}^{d-1}$ that vanishes on all of the $\bar{w}_i$ except for either two or three of them, and on those evaluates to the same sign. Geometrically, there must exist a linear hyperplane in $\mathbb{R}^{d-1}$ that contains all vectors $\bar{w}_i$ except for two or three, and has those on the same side.

Thus, we have traded the search for the locus of two collinear vectors among the $w_i \in \mathbb{R}^2$ (a problem involving only two pieces of input data) for the task of finding a Gale dual $\bar{W}$ in the much higher-dimensional space $\mathbb{R}^{d-1}$, and hyperplanes passing through almost all of the $\bar{w}_i$ — a problem involving almost the entire input!

That this is not crazy, but instead effective, is explained by the fact that passing to the higher-dimensional representation is possible in great generality, and moreover greatly simplifies the structure of the problem; see Proposition 2.1 below.

2.2. Implementation

Let $B = \{b_i\colon\ 0 \leqslant i \leqslant d\}$ be any basis of P_d.

2.2.1. The Gale dual

Form the matrix

$$W = W(x, y) = \begin{pmatrix} R_0 & R_1 & \cdots & R_d \\ I_0 & I_1 & \cdots & I_d \end{pmatrix},$$

where $R_j = R_j(x, y)$ and $I_j = I_j(x, y)$ denote the real and imaginary parts of the complex polynomial $b_j = b_j(x + iy)$. The rank of W is 2, so any Gale dual matrix $\bar{W}$ to W has size $(d + 1) \times (d - 1)$. The following proposition gives an explicit representative for $\bar{W}$ involving polynomials p_k, q_k, r_k that depend on the basis B. For four especially relevant bases, we will make the Gale dual $\bar{W}$ completely explicit. These bases are:

- The *power basis*, where $b_i = z^i$;
- the *falling factorial basis*, where $b_i = z^{\underline{i}} = z(z - 1) \cdots (z - i + 1)$;
- the *rising factorial basis*, where $b_i = z^{\overline{i}} = z(z + 1) \cdots (z + i - 1)$; and
- the *binomial coefficient basis*, where $b_i = \binom{z+d-i}{d}$.

Here $z^0 = z^{\underline{0}} = z^{\overline{0}} = 1$.

Proposition 2.1. *A Gale dual matrix to W may be chosen to have exactly three nonzero diagonals*

$$\bar{W} = \bar{W}(x, y) = \begin{pmatrix} p_0 & 0 & 0 & \cdots & 0 \\ -q_0 & p_1 & 0 & \cdots & 0 \\ r_0 & -q_1 & \ddots & & \vdots \\ 0 & r_1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & p_{d-2} \\ & & & r_{d-3} & -q_{d-2} \\ 0 & \cdots & & 0 & r_{d-2} \end{pmatrix}. \tag{2}$$

Moreover, its entries may be chosen to lie in $\mathbb{R}[x, y]$. For the four bases considered, we may choose the following explicit values:

b_i	p_k	q_k	r_k
z^i	$x^2 + y^2$	$2x$	1
$z^{\underline{i}}$	$(x-k)^2 + y^2$	$2(x-k) - 1$	1
$z^{\overline{i}}$	$(x+k)^2 + y^2$	$2(x+k) + 1$	1
$\binom{z+d-i}{d}$	$(x-k)^2 + y^2$	$p_k + r_k - d(d-1)$	p_{k+1-d}

Note that in the last row, $q_k = 2(x - (k - \frac{d-1}{2}))^2 + 2y^2 - \frac{d^2-1}{2}$.

Proof. We first prove that the matrix $\bar{W}$ can be chosen to have the displayed triple band structure regardless of the basis B chosen for P_d. For this, define the rational functions $g_k = \frac{b_{k+1}}{b_k} \in \mathbb{R}(z)$ for $0 \leqslant k \leqslant d-1$; specific values for g_k become apparent from the relations $z^{k+1} = z \cdot z^k$, $z^{\underline{k+1}} = (z-k)z^{\underline{k}}$, $z^{\overline{k+1}} = (z+k)z^{\overline{k}}$ and $\binom{z+d-k-1}{d} = \frac{z-k}{z+d-k}\binom{z+d-k}{d}$.

The triple (p_k, q_k, r_k) lists nontrivial coefficients of a real syzygy

$$p_k b_k + q_k b_{k+1} + r_k b_{k+2} = b_k(p_k + g_k q_k + g_k g_{k+1} r_k) = 0$$

whenever

$$\begin{pmatrix} 1 & \operatorname{Re} g_k & \operatorname{Re} g_k g_{k+1} \\ 0 & \operatorname{Im} g_k & \operatorname{Im} g_k g_{k+1} \end{pmatrix} \begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

But the displayed matrix with entries in $\mathbb{R}(x, y)$, call it M, obviously has rank at least 1, and rank 2 whenever $\operatorname{Im} g_k(x+iy) \neq 0$, so that such triples certainly exist. Moreover, by multiplying with a common denominator we may assume $p_k, q_k, r_k \in \mathbb{R}[x, y]$, and so the relations $p_k b_k + q_k b_{k+1} + r_k b_{k+2} = 0$ imply that $\bar{W}$ is in fact a Gale dual of W. The concrete syzygies listed above arise by choosing explicit bases for $\ker M$. □

Remark 2.2. Another interesting case is that of polynomials with symmetric coefficients. For instance, if $f = \sum_{i=0}^{d} a_i \binom{z+d-i}{d}$ and $a_i = a_{d-i}$, we may expand f in the basis $B = \{\binom{z+d-i}{d} + \binom{z+i}{d}\colon 0 \leqslant i \leqslant \lfloor \frac{d}{2} \rfloor\}$ of the vector space of polynomials with symmetric coefficients in the binomial coefficient basis. However, the coefficients of syzygies of these b_k do not appear to be as simple as the ones listed in Proposition 2.1. For example, a typical coefficient (namely, q_1 for $d = 8$) reads

$$-8\big(\big((x+\alpha_1)^2 + y^2\big)\big((x+\alpha_2)^2 + y^2\big) + \gamma_1\big)\big(\big((x+\beta_1)^2 + y^2\big)\big((x+\beta_2)^2 + y^2\big) + \gamma_2\big) + \gamma,$$

where α_1, α_2 are the roots of $\alpha^2 - \alpha + \rho_1 = 0$ (so that $\alpha_1 + \alpha_2 = 1$), β_1, β_2 are the roots of $\beta^2 - \beta + \rho_2 = 0$, ρ_1, ρ_2 are the roots of $\rho^2 - \frac{29}{2}\rho - \frac{231}{8} = 0$, $\gamma_1 + \gamma_2 = \frac{135}{4}$, $\gamma_1 = \frac{135}{8}(1 - \frac{61}{\sqrt{649}})$, and $\gamma = \frac{874800}{649}$. We will not pursue this basis further in this paper.

2.2.2. The determinants

Recall that two vectors $w_j(z)$, $w_k(z)$ become collinear at some point $z \in \mathbb{C}$ whenever there exists a circuit of the vector configuration $W(z)$ with exactly two nonzero entries. By Gale duality, this means that the Gale dual vector configuration $\bar{W}(z)$ has a cocircuit with support 2, i.e., the determinant of the matrix obtained by deleting two rows from $\bar{W}$ vanishes. Our approach rests on the fact that we can give fairly explicit expressions for these determinants for the four bases considered here.

Lemma 2.3. *Let $\bar{W}_{(j,k)} = \bar{W}_{(j,k)}(x, y)$ denote the square matrix obtained by deleting rows j and k from $\bar{W}$, where $0 \leqslant j < k \leqslant d$ (so that we number the rows from 0 to d). Then*

$$\det \bar{W}_{(j,k)} = p_0 \cdots p_{j-1} D_{j,k} r_{k-1} \cdots r_{d-2},$$

where $D_{j,k} = D_{j,k}(x, y)$ is the determinant of the tridiagonal matrix

$$\begin{pmatrix} -q_j & p_{j+1} & 0 & \cdots & 0 \\ r_j & -q_{j+1} & \ddots & & \vdots \\ 0 & r_{j+1} & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & p_{k-2} \\ & & & r_{k-3} & -q_{k-2} \end{pmatrix}. \tag{3}$$

Here $D_{j,j+1} := 1$, and the leading resp. trailing products are 1 if $j = 0$ resp. $k = d$. In particular, $D_{j,j+2} = -q_j$.

Proof. The matrix $\bar{W}_{(j,k)}$ decomposes into three blocks, whose determinants yield the stated formula, and two additional elements r_{j-1} and p_{k-1} that do not contribute to $\det \bar{W}_{(j,k)}$. $\square$

Proposition 2.4. *Set $z = x + iy$ and $\bar{z} = x - iy$. Then*

$$D_{j,k}(x, y) = \frac{(-1)^{k-j-1}}{z - \bar{z}} \bigl(f_{j,k}(z) - f_{j,k}(\bar{z}) \bigr),$$

where the polynomials $f_{j,k}(z)$ are given in the following table:

b_i	$f_{j,k}(z)$
z^i	z^{k-j}
$z^{\underline{i}}$	$(z-k+1)\cdots(z-j)$
$z^{\overline{i}}$	$(z+k-1)\cdots(z+j)$
$\binom{z+d-i}{d}$	$\frac{1}{d}(z-k+1)\cdots(z-j)(\bar{z}+d-k+1)\cdots(\bar{z}+d-j)$

The $D_{j,k}$ are real polynomials with even degrees in y.

Proof. It is well known that the determinant D_n of an $n \times n$ tridiagonal matrix $A = (a_{ij})$ satisfies the three-term recursion relation $D_n = a_{nn} D_{n-1} - a_{n,n-1} a_{n-1,n} D_{n-2}$. Solving this recursion for the matrix from Lemma 2.3 with the values from Proposition 2.1 and the boundary conditions $D_{j,j+1} = 1$ and $D_{j,j+2} = -q_j$ yields the stated expressions. $\square$

2.3. The real case

Up to now, we have only considered complex, nonreal roots of f. The case of real roots is much simpler, and the machinery used for complex roots specializes in a straightforward way to the real case. If we regard both f and the b_i as polynomials in one real variable, the matrix $W = W^{\mathbb{R}}$ reduces to the single row $W^{\mathbb{R}} = (b_0, \ldots, b_d)$.

Proposition 2.5. *A basis for the kernel of $W^{\mathbb{R}}$ is given by the columns of the matrix*

$$\bar{W}^{\mathbb{R}} = \begin{pmatrix} -b_1 & 0 & 0 & \cdots & 0 \\ b_0 & -b_2 & 0 & \cdots & 0 \\ 0 & b_1 & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & -b_d \\ 0 & \cdots & & 0 & b_{d-1} \end{pmatrix}$$

of size $(d+1) \times d$. The determinant of the matrix obtained by deleting row j from W is

$$\det \bar{W}_j = (-1)^j b_j \Pi \quad \text{for } j = 0, 1, \ldots, d,$$

with $\Pi = b_1 b_2 \cdots b_{d-1}$.

3. Bounding the location of roots

We first treat the case of complex, nonreal roots. For each ordered triple of indices i, j, k with $0 \leqslant i < j < k \leqslant d$, denote by $H_{i,j,k}$ the hyperplane in $\mathbb{R}^{d-1}$ spanned by the rows of the matrix $\bar{W}_{(i,j,k)}$, obtained by deleting the rows $\bar{w}_i, \bar{w}_j, \bar{w}_k$ from $\bar{W}$.

Definition 3.1. $\mathcal{S}_{i,j,k}$ is the set of all $z = x + iy \in \mathbb{C}$ such that $H_{i,j,k} = H_{i,j,k}(x, y)$ induces a nonnegative cocircuit, i.e., the vectors $\bar{w}_i = \bar{w}_i(x, y)$, $\bar{w}_j = \bar{w}_j(x, y)$, $\bar{w}_k = \bar{w}_k(x, y)$ all (weakly) lie on the same side of $H_{i,j,k}(x, y)$.

The sets $\mathcal{S}_{i,j,k}$ are crucial for our purposes for the following reason: If $z \in \mathcal{S}_{i,j,k}$, then the corresponding Gale primal vectors w_i, w_j, w_k form a nonnegative circuit, and thus yield a nonnegative combination of all w's that sums to zero; in other words, there exists some polynomial f with nonnegative coefficients in the chosen basis B that has a zero at z. On the other hand, if $z \notin \mathcal{S}_{i,j,k}$, we can only conclude that the three particular Gale primal vectors w_i, w_j, w_k do *not* form a circuit, and so are not responsible for the possible zero z of f.

Proposition 3.2. *For $0 \leqslant i < j < k \leqslant d$ and $z \in \mathbb{C} \setminus \mathbb{R}$, let $\sigma_{i,j,k}(z)$ be the set of signs*

$$\begin{aligned} &\bigl\{(-1)^i \operatorname{sign} \det \bar{W}_{(j,k)}(x, y), (-1)^{j-1} \operatorname{sign} \det \bar{W}_{(i,k)}(x, y), (-1)^{k-2} \operatorname{sign} \det \bar{W}_{(i,j)}(x, y)\bigr\} \\ &\quad = \bigl\{(-1)^i \operatorname{sign} D_{j,k}(x, y), (-1)^{j-1} \operatorname{sign} D_{i,k}(x, y), (-1)^{k-2} \operatorname{sign} D_{i,j}(x, y)\bigr\}. \end{aligned}$$

Then each $\mathcal{S}_{i,j,k} \subset \mathbb{R}^2$ is a semialgebraic set defined as the locus of all (x, y) such that

$$\{\pm 1\} \not\subset \sigma_{i,j,k}(x + iy).$$

Proof. We obtain a linear form $\varphi_{i,j,k}$ on $\mathbb{R}^{d-1}$ whose vanishing locus is the hyperplane $H_{i,j,k}$ by adding a first row of variables $x_1, \ldots, x_{d-1}$ to $\bar{W}_{(i,j,k)}$ and expanding the determinant of that square matrix along the first row. The value of $\varphi_{i,j,k}$ on $\bar{w}_i$, say, is given by the sign $(-1)^i$ of the permutation that interchanges rows 0 and i in the matrix $\bar{W}_{(j,k)}$, times $\det \bar{W}_{(j,k)}$. For $H_{i,j,k}$ to define a (positive or negative) cocircuit, the signs obtained in this way for $\bar{w}_i$, $\bar{w}_j$ and $\bar{w}_k$ must agree. Finally, by Lemma 2.3 the signs of $\det \bar{W}_{(j,k)}$ and $D_{j,k}$ agree except perhaps on the real axis (on the vanishing locus of the p_k's and r_k's), and we may assume that $j - i \geqslant 2$ and $k - j \geqslant 2$. □

In summary:

Theorem 3.3. *Let f be a polynomial of degree d with nonnegative coefficients with respect to some basis of the vector space P_d. Then the set of nonreal roots of f is contained in the union of the semialgebraic sets $\mathcal{S}_{i,j,k}$, for $0 \leqslant i < j < k \leqslant d$. Put differently, if*

$$\{-1, 1\} \subseteq \left\{(-1)^i D_{j,k}(z_0), (-1)^{j+1} D_{i,k}(z_0), (-1)^k D_{i,j}(z_0)\right\}$$

for each triple (i, j, k) with $0 \leqslant i < j < k \leqslant d$, then z_0 is not a root of f.

After a short discussion of the real case, we will apply this result to our four representative bases. We only discuss the power basis and binomial coefficient basis in any detail, as the procedure for the rising and falling factorial bases is almost exactly the same.

3.1. The real case

A real number $x \in \mathbb{R}$ is a real root of some polynomial f with nonnegative coefficients with respect to a fixed basis if and only if either x is a root of some basis polynomial, or two basis polynomials differ in sign when evaluated at x. In other words:

Observation 3.4. *Let f be a polynomial of degree d with nonnegative coefficients with respect to some basis $\{b_0, \dots, b_d\}$ of P_d. Then the locus of possible real roots of f is the set of $x \in \mathbb{R}$ for which $b_i(x)b_j(x) \leqslant 0$ for some $i \neq j$.*

For the sake of completeness, and in response to the query of one of the referees, we briefly rederive this result using our framework of Gale transforms.

Proof of Observation 3.4. In complete analogy to the complex case, denote for $0 \leqslant i < j \leqslant d$ by $\bar{W}^{\mathbb{R}}_{(i,j)}$ the matrix obtained by deleting rows w_i and w_j from $\bar{W}^{\mathbb{R}}$, by $H_{i,j}$ the hyperplane in $\mathbb{R}^d$ spanned by the rows of $\bar{W}^{\mathbb{R}}_{(i,j)}$, and by $\mathcal{S}_{i,j} \subseteq \mathbb{R}$ the set of all $x \in \mathbb{R}$ such that $H_{i,j}(x)$ induces a positive cocircuit, i.e., the vectors $\bar{w}_i = \bar{w}_i(x)$ and $\bar{w}_j = \bar{w}_j(x)$ lie on the same side of $H_{i,j}$.

To find these cocircuits explicitly, build a linear form $\varphi_{i,j}$ on $\mathbb{R}^d$ that defines $H_{i,j}$ by adding a first row $(x_1, \dots, x_d)$ of variables to $\bar{W}^{\mathbb{R}}_{i,j}$ and expanding the determinant of that square matrix along the first row. Just as in the proof of Proposition 3.2, $\varphi_{i,j}(\bar{w}_i) = (-1)^i \det \bar{W}^{\mathbb{R}}_i = b_i\Pi$ and $\varphi_{i,j}(\bar{w}_j) = (-1)^{j-1} \det \bar{W}^{\mathbb{R}}_j = -b_j\Pi$ by Proposition 2.5. In consequence, $\mathcal{S}_{i,j}$ is the locus of points $x \in \mathbb{R}$ such that $b_i(x)$ and $b_j(x)$ differ in sign. This finishes the proof. □

3.2. The power basis

For $b_i = z^i$, we set $n = k - j - 1$, write D_n for $D_{j,k}$, and substitute $z = re^{i\theta}$ into D_n:

$$D_n = (-1)^n \frac{z^{n+1} - \bar{z}^{n+1}}{z - \bar{z}} = (-1)^n r^n e^{-in\theta} \frac{e^{2i(n+1)\theta} - 1}{e^{2i\theta} - 1}.$$

This vanishes iff $\theta = \pi l/(n+1)$ for integer l with $1 \leqslant l \leqslant 2n+1$ and $l \neq n+1$. The zero locus of D_n thus consists of n lines through the origin, the ones closest to the x-axis having angles $\theta = \pm\frac{\pi}{n+1}$. We conclude that D_n has the same sign throughout the entire open sector $Z_{n+1} = \{z \in \mathbb{C}\colon -\frac{\pi}{n+1} < \arg z < \frac{\pi}{n+1}\}$. By substituting a positive, real value of z into $D_n = (-1)^n \sum_{j=0}^n z^j \bar{z}^{n-j}$, we determine this sign to be $(-1)^n = (-1)^{k-j-1}$.

For $z \in Z_d$ and $0 \leqslant i < j < k \leqslant d$ the set of signs of the polynomials in Proposition 3.2 is

$$\sigma_{i,j,k}(z) = \left\{(-1)^{i+j+k+1}, (-1)^{i+j+k}, (-1)^{i+j+k-1}\right\} = \{\pm 1\}.$$

This implies $\mathcal{S}_{i,j,k} = \emptyset$, and thus Theorem 3.3 recovers the classical result that a polynomial of degree d with positive coefficients in the power basis has no zeros in Z_d; of course, this includes the case of real roots.

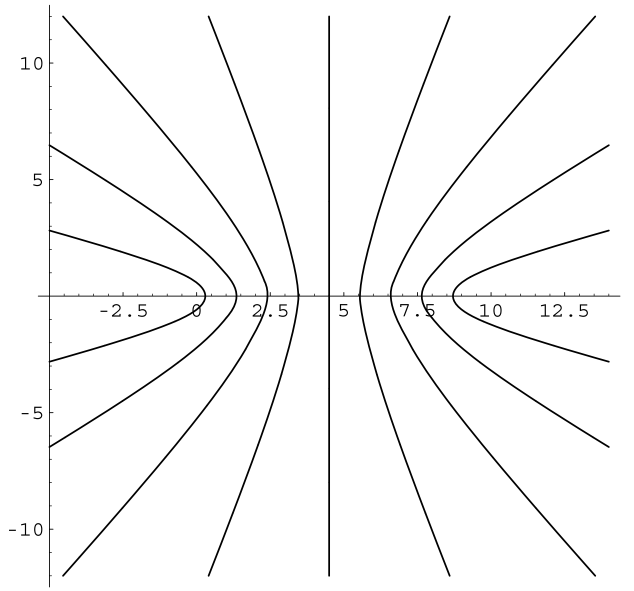

Fig. 2. The locus $\mathcal{D}_{0,10}$ in the case of the rising factorial basis.

3.3. *Rising and falling factorial basis*

In both cases, the polynomials $f_{j,k}$ from Proposition 2.4 have the form $f_{j,k}(z) = \prod_{i=1}^{n+1}(z - a_i)$, with $n = k - j - 1$ and $a_i = j - 1 + i$ for the falling powers and $a_i = -(j - 1 + i)$ in the case of the rising powers. The transform $z \mapsto z \pm \frac{j+k-1}{2}$ remedies this asymmetry, where we choose the '−' sign for $b_i = z^{\underline{i}}$ and the '+' sign for $b_i = z^{\bar{i}}$. The a_i then become integers or half-integers in the range $\pm\frac{n}{2}$.

Using the same type of analysis as will be detailed in Section 4 for the binomial coefficient basis, one can prove that the zero locus $\mathcal{D}_{j,k} = \{z \in \mathbb{C}\colon\ D_{j,k}(z) = 0\}$ is smooth everywhere, that one component intersects the real axis between each pair of adjacent a_i's, and that far away from the origin $\mathcal{D}_{j,k}$ approaches the arrangement of lines through the origin with slopes $\pm\frac{1}{n+1}, \dots, \pm\frac{n}{n+1}$; cf. Fig. 2. We will not enter into the details here, but instead treat the remaining basis in a separate section.

4. The binomial coefficient basis

We first get the real case out of the way: The basis polynomials all have the same sign outside the closed interval $[-d, d-1]$, and at each point inside this interval there are two basis polynomials that evaluate to opposite signs. By Observation 3.4, $[-d, d-1]$ is exactly the set of possible real roots.

For the nonreal roots, as before we pass to an adapted coordinate system with respect to which the vanishing locus of $D_{j,k}$ is centro-symmetric, by replacing

$$z \mapsto z' + (k + j - d - 1)/2 \tag{4}$$

in $d \cdot f_{j,k}(z)$. Writing again z for z' yields

$$df_{j,k}(z) = \prod_{i=j}^{k-1}\left(z + \frac{k + j - d - 1 - 2i}{2}\right)\prod_{i=j}^{k-1}\left(\bar{z} + \frac{k + j + d - 1 - 2i}{2}\right).$$

Next, we replace i by $i + j$ in the first product and by $k - 1 - i$ in the second, to obtain

$$df_{j,k}(z) = \prod_{i=1}^{k-j}\left(z - i - \frac{\Delta}{2}\right)\left(\bar{z} + i + \frac{\Delta}{2}\right),$$

where $\Delta = d - 1 - k + j$. Introducing $a_i = i + \frac{\Delta}{2}$ we obtain

$$D_{j,k}(z) = D_n(z) = \frac{(-1)^{n+1}}{d(\bar{z} - z)}\left(\prod_{i=1}^{n+1}(z - a_i)(\bar{z} + a_i) - \prod_{i=1}^{n+1}(\bar{z} - a_i)(z + a_i)\right), \tag{5}$$

where we have set $n = k - j - 1$ (so that $\Delta = d - 2 - n$), in accordance with the fact that the degree of $D_{j,k}(z)$ in z is n.

Before examining the zero locus of $D_{j,k}(z)$, we pause to calculate the leading coefficient. This result will be used in Section 6.1.

Lemma 4.1. *The leading coefficient of $D_{j,k}(z)$ is*

$$[2n]D_{j,k}(z) = (z\bar{z})^{k-j-1}(-1)^{k-j-1}(k - j) = r^{2n}(-1)^n(n + 1), \tag{6}$$

where $z = re^{i\phi}$. It is invariant under substitutions of the form $z \mapsto z + z_0$, and the sign of $D_n(z)$ outside the outermost component of $\mathcal{D}_n$ is $(-1)^n$, and $+1$ inside the innermost one.

Proof. See Appendix A. □

We now treat the zero locus of D_n. First, whenever $D_n(z) = 0$,

$$\sum_{i=1}^{n+1}\arg(z - a_i) + \sum_{i=1}^{n+1}\arg(\bar{z} + a_i) - \sum_{i=1}^{n+1}\arg(\bar{z} - a_i) - \sum_{i=1}^{n+1}\arg(z + a_i) = 2l\pi$$

for some integer l. Because $\arg(z \pm a_i) = -\arg(\bar{z} \pm a_i)$, this relation reads

$$\sum_{i=1}^{n+1}\arg(z - a_i) - \sum_{i=1}^{n+1}\arg(z + a_i) = \sum_{i=1}^{n+1}\alpha_i = l\pi, \tag{7}$$

where α_i is the angle under which the segment $[-a_i, a_i]$ appears as seen from z (cf. Fig. 3).

We may assume without loss of generality that z lies in the upper half-plane, and therefore that $\arg(z - a_i) > \arg(z + a_i) > 0$, which implies $l \geqslant 1$. On the other hand, the maximal value $(n+1)\pi$ of (7) is achieved for real z between $-a_1$ and a_1, so that $l \leqslant n$ for nonreal z. From this, we can draw several conclusions, which we detail in Section 4.1. The reader may want to just skim this material, and otherwise skip ahead to Section 4.2, where we apply it to conclude that the root locus is bounded.

4.1. Limiting behavior and global geometric properties of $D_{j,k}$

Proposition 4.2. *When d becomes large with respect to n, the zero locus of $D_{j,k}$ approaches a union of circles passing through $\pm d/2$ and symmetric about the imaginary axis. For $l = 1, \ldots, k - j - 1$, these circles have center*

$$z_l = -\frac{d + 1 - k - j}{2} - i\frac{d}{2}\cot\frac{l\pi}{k - j}$$

and radius

$$r_l = \frac{d}{2\sin\frac{l\pi}{k-j}}.$$

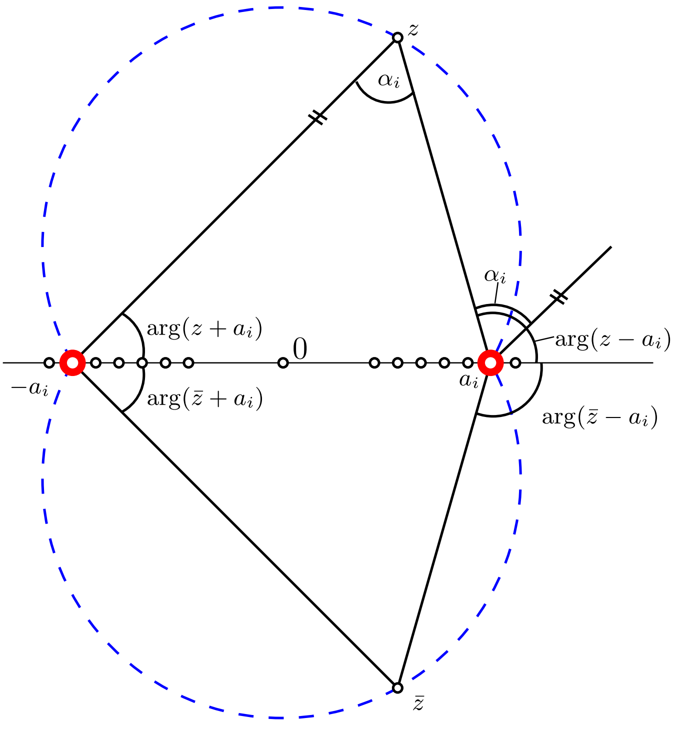

Fig. 3. The segments $[-a_i, a_i]$ as seen from z.

Proof. For d large with respect to n, the points $\pm a_i$ fuse to $\pm a = \pm\frac{d}{2}$, so that (7) reads

$$\alpha := \arg(z - a) - \arg(z + a) = \frac{l\pi}{n+1}.$$

By elementary geometry, the locus of these points is a union of two circular arcs with the specified equations. □

Example 4.3. For $n = 1$, we obtain $a_1 = \frac{d-1}{2}$ and $a_2 = \frac{d+1}{2}$. For large d, they approach $a = \frac{d}{2}$ and Eq. (7) says $\alpha = \frac{\pi}{2}$. In that limit, $\mathcal{D}_1$ thus approaches the circumference with center 0 and radius $\frac{d}{2}$. For smaller values of d, directly evaluating Eq. (5) yields

$$D_1(z) = z\bar{z} - a_1 a_2,$$

which describes a circumference of center 0 and radius $\sqrt{a_1 a_2} = \frac{1}{2}\sqrt{d^2 - 1} < \frac{d}{2}$.

Proposition 4.4. *The plane algebraic curve $\mathcal{D}_{j,k}$ with equation $D_{j,k}(z) = 0$ is smooth. The only points where it has horizontal tangent vectors lie on the y-axis.*

This is proved in Appendix A.

Proposition 4.5. *All algebraic curves $\mathcal{D}_{j,k}$ consist of $n = k - j - 1$ nested ovals. The i-th oval intersects the real axis inside the union of open intervals $\pm(a_i, a_{i+1})$, for $i = 1, \dots, n$.*

Proof. Let $\phi \in S^1 \setminus S^0$ be a nonreal unit vector and ρ the ray through the origin and ϕ. At each point p of $\mathcal{D}_{j,k} \cap \rho$, the angle sum $\sum_{i=1}^{n+1} \alpha_i$ takes on some value $l\pi$ among the discrete set

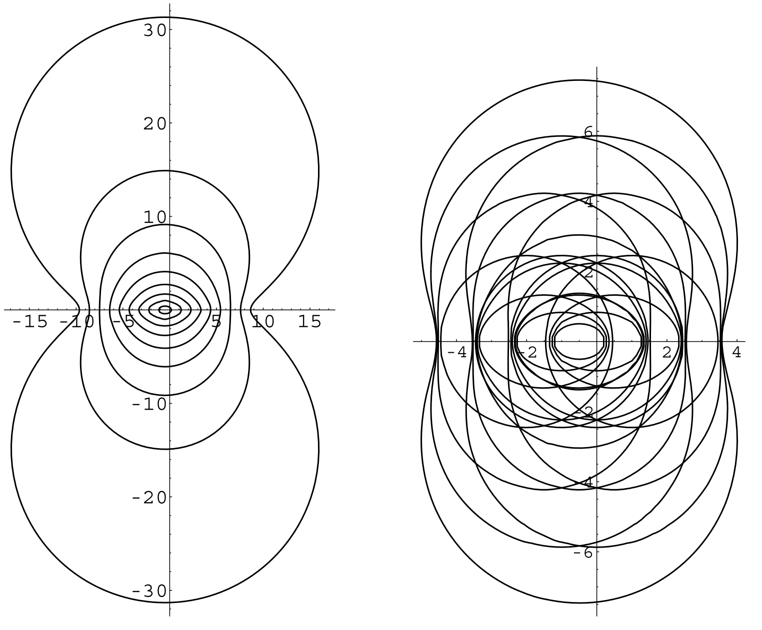

Fig. 4. *Left*: The curve $\mathcal{D}_{0,10}$ for $d = 10$. *Right*: The curves $\mathcal{D}_{i,j}$ with $0 \leqslant i < j \leqslant 5$.

$\{\pi, \dots, n\pi\}$, and therefore this value remains constant on the entire connected component to which p belongs. The argument extends to the real axis by smoothness of $\mathcal{D}_{j,k}$.

For the second statement, observe that the value of $\alpha_j = \arg(z - a_j) - \arg(z + a_j)$ increases by almost π as z travels from $a_j + \varepsilon + i\delta$ to $a_j - \varepsilon + i\delta$, for $0 < \delta \ll \varepsilon \ll 1$. $\square$

Example 4.6. For $D_{0,10} = D_9$ and $d = 10$, we obtain the picture of Fig. 4.

To continue, we introduce some useful notation. By (5), the formula for $D_{j,k}(z)$ involves the points $a_{j,k;i} = i + \frac{1}{2}(d - 1 - k + j)$ for $1 \leqslant i \leqslant k - j$, so that

$$(a_{j,k;1}, \dots, a_{j,k;k-j}) = \left(\frac{d}{2} - \frac{k-j-1}{2}, \dots, \frac{d}{2} + \frac{k-j-1}{2}\right).$$

We write $\alpha(\pm a_{j,k;i}; z)$ for the angle under which $z \in \mathbb{C}$ sees the segment $[-a_{j,k;i}, a_{j,k;i}]$, and $A(j,k;z) = \sum_{i=1}^{k-j} \alpha(\pm a_{j,k;i}; z)$ for the corresponding angle sum. Moreover, let

$$\mathcal{D}_{j,k;l} = \left\{(x, y) \in \mathbb{R}^2\colon\ A(j,k;x+iy) = l\pi\right\} \quad \text{for } l = 1, \dots, k-j-1,$$

be the l-th oval of $\mathcal{D}_{j,k}$, and $\operatorname{cl}\mathcal{D}_{j,k;l}$ the closure of the region in $\mathbb{R}^2$ bounded by $\mathcal{D}_{j,k;l}$.

Remark 4.7. The arrangement of ovals $\{\mathcal{D}_{j,k;l}\colon\ 0 \leqslant j < k \leqslant d,\ 1 \leqslant l \leqslant k-j\}$ has several interesting combinatorial properties, which we will not pursue in this paper. Here we would only like to point out the triple points where components of $\mathcal{D}_{j,r}$, $\mathcal{D}_{r,k}$ and $\mathcal{D}_{j,k}$ intersect.

Proposition 4.8. *Let $0 \leqslant j \leqslant j' < k' \leqslant k \leqslant d$ and $1 \leqslant l \leqslant k - j - 1$ be integers.*

Then $\mathcal{D}_{j',k'} \subset \operatorname{cl}\mathcal{D}_{j,k;1} \smallsetminus \operatorname{cl}\mathcal{D}_{j,k;k-j-1}$. In particular, all components of all curves $\mathcal{D}_{j,k}$ are contained in the topological closure of $\operatorname{cl}\mathcal{D}_{0,d;1} \smallsetminus \operatorname{cl}\mathcal{D}_{0,d;d-1}$. Moreover, for all integers δ_1, δ_2 with $0 \leqslant \delta_1 \leqslant j$, $0 \leqslant \delta_2 \leqslant d-k$ and (where appropriate) $1 + \delta_1 + \delta_2 \leqslant l \leqslant d - \delta_1 - \delta_2$,

$$\operatorname{cl}\mathcal{D}_{j-\delta_1,k+\delta_2;l+\delta_1+\delta_2} \subseteq \operatorname{cl}\mathcal{D}_{j,k;l} \subseteq \operatorname{cl}\mathcal{D}_{j+\delta_1,k-\delta_2;l-\delta_1-\delta_2}, \tag{8}$$

$$\operatorname{cl}\mathcal{D}_{j+\delta_1,k-\delta_2;l} \subseteq \operatorname{cl}\mathcal{D}_{j,k;l} \subseteq \operatorname{cl}\mathcal{D}_{j-\delta_1,k+\delta_2;l}. \tag{9}$$

Proof. We first show that $\mathcal{D}_{j',k';l} \subset \operatorname{cl}\mathcal{D}_{j,k;1}$ for all l with $1 \leqslant l \leqslant k'-j'-1$. The first set consists of all points $z \in \mathbb{C}$ such that $\sum_{i=1}^{k'-j'} \alpha(\pm a_{j',k';i}; z) = l\pi$ in the centro-symmetric coordinates. Undoing the coordinate change (4) yields

$$\mathcal{D}_{j',k';l} = \{z \in \mathbb{C}\colon \alpha(\pm(d-k'+1); z) + \cdots + \alpha(\pm(d-j'); z) = l\pi\},$$
$$\operatorname{cl}\mathcal{D}_{j,k;1} = \{z \in \mathbb{C}\colon \alpha(\pm(d-k+1); z) + \cdots + \alpha(\pm(d-j); z); \geqslant \pi\}.$$

Now the required inclusion is clear, because the first set of points of which the viewing angle is taken is a subset of the second one. It remains to prove that $\mathcal{D}_{j,k;k-j-1} \subset \operatorname{cl}\mathcal{D}_{j',k';l}$ for all $1 \leqslant l \leqslant k'-j'-1$; proving the extremal case $l = k'-j'-1$ is sufficient. Thus, we are required to show that $\sum_{m=d-k+1}^{d-j} \alpha(\pm m; z) = (k-j-1)\pi$ implies $\sum_{m=d-k'+1}^{d-j'} \alpha(\pm m; z) \geqslant (k'-j'-1)\pi$. But this is true because the first sum has $k-j$ summands, the second $k'-j'$ summands, and removing each of the $(k-j)-(k'-j') = (k-k')+(j'-j) \geqslant 0$ pairs of points from the points corresponding to the first summand decreases the total viewing angle by at most π.

Similarly, the first inclusion of (8) follows because $\sum_{m=d-k+1-\delta_2}^{d-j+1+\delta_1} \alpha(\pm m; z) \geqslant (l+\delta_1+\delta_2)\pi$ implies $\sum_{m=d-k+1}^{d-j} \alpha(\pm m; z) \geqslant l\pi$, by removing $\delta_1+\delta_2$ pairs of points, and the second one from an appropriate change of variables. Relations (9) are proved in exactly the same way. □

Corollary 4.9. *For $z \notin \operatorname{cl}\mathcal{D}_{0,d;1}$, the facets of the cone $\tau(z) = \mathbb{R}_{\geqslant 0}\langle w_0(z), \ldots, w_d(z)\rangle$ are the rays spanned by $w_0(z)$ and $w_d(z)$, and $w_1(z), \ldots, w_{d-1}(z)$ appear in cyclic order inside $\tau(z)$.*

Proof. The argument of $w_j(z) = \binom{z+d-j}{d}$ is $(\sum_{i=1}^{d} \arg(z+i-j)) \bmod 2\pi$, so that the difference of the arguments of $w_j(z)$ and $w_{j+1}(z)$ equals $\beta_j := \arg(z+d-j) - \arg(z-j) \bmod 2\pi$; cf. [4]. If we choose z to have the form $z = N + i\varepsilon$, with $N \gg \varepsilon > 0$, it is not necessary to reduce β_j modulo 2π, and $0 < \beta_0 < \beta_1 < \cdots < \beta_d$; we may even achieve $\beta_d < \frac{\pi}{d}$, so that the total angle subtended by the $w_i(z)$ is strictly less than π, and $w_0(z)$ and $w_d(z)$ span the facets of $\tau(z)$. Now note that two vectors $w_i(z), w_j(z)$ become collinear iff there is a (not necessarily positive or negative) circuit involving the two, iff there is such a cocircuit involving $\bar{w}_i(z), \bar{w}_j(z)$, iff $\mathcal{D}_{j,k}(z) = 0$. An invocation of Proposition 4.8 finishes the proof. □

We close with a lemma regarding the relative orientations of w_j, w_k on $\mathcal{D}_{j,k}$.

Lemma 4.10. *Let $z \in \mathcal{D}_{j,k;l}$, and regard $w_i(z) = \binom{z+d-i}{d}$ as a vector in $\mathbb{R}^2$. Then $w_j(z)$ and $w_k(z)$ point in the same direction iff l is even, and in opposite directions iff l is odd*:

$$\operatorname{sign}(w_j(z) \cdot w_k(z)) = (-1)^l \quad \textit{for } z \in \mathcal{D}_{j,k;l}.$$

4.2. Conclusion: The root locus is bounded

Theorem 4.11. *Let $f = \sum_{j=0}^{d} a_j \binom{z+d-j}{d}$ be a polynomial of degree d with nonnegative coefficients $a_j \geqslant 0$ with respect to the binomial coefficient basis. Then all nonreal roots of f are contained in the region $\operatorname{cl}\mathcal{D}_{0,d;1}$ bounded by the outermost oval of the algebraic curve with equation $D_{0,d}(z) = 0$, and any point inside $\operatorname{cl}\mathcal{D}_{0,d;1}$ arises as a root of some such f.*

The real roots of f all lie in the real interval $[-d, d-1]$.

Proof. The first statement can be proved by a short calculation involving Lemma 4.1 and the general tool of Theorem 3.3. However, we have accumulated enough information about the special curves $\mathcal{D}_{j,k}$

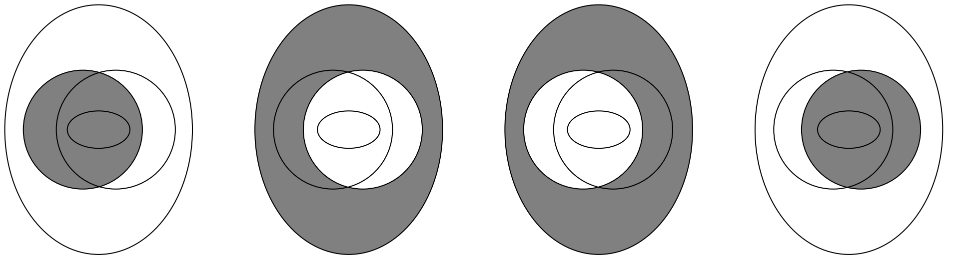

Fig. 5. From left to right, the semialgebraic sets $\mathcal{S}_{012}$, $\mathcal{S}_{013}$, $\mathcal{S}_{023}$, $\mathcal{S}_{123}$ (shaded). Their union equals the entire interior of the bounding curve C, which by Theorem 4.11 is precisely the locus of possible nonreal roots.

arising for the binomial coefficient basis to give a direct proof: By Corollary 4.9, the vectors $w_0 = w_0(z), \dots, w_d = w_d(z)$ are positively spanning for $z \notin \operatorname{cl} \mathcal{D}_{0,d;1}$.

Next, suppose that $z \in \operatorname{cl} \mathcal{D}_{0,d;1}$ falls inside the region $S_k := \operatorname{cl} \mathcal{D}_{0,k;1} \setminus \operatorname{cl} \mathcal{D}_{0,k-1;1}$ for some $k \in \mathbb{N}$ with $2 \leqslant k \leqslant d$. Such a k exists, because $\operatorname{cl} \mathcal{D}_{0,k-1;1} \subset \operatorname{cl} \mathcal{D}_{0,k;1}$ by (9), and $\operatorname{cl} \mathcal{D}_{0,1;1} = \emptyset$. We claim that in this situation, the vectors w_0, w_{k-1} and w_k are positively spanning. Indeed, the locus of points in the complex plane where the combinatorics of this subconfiguration changes is exactly $\mathcal{D}_{0,k-1} \cup \mathcal{D}_{0,k}$, because $\mathcal{D}_{k,k-1} = \emptyset$. Moreover, $\mathcal{D}_{0,k;l} \subset \operatorname{cl} \mathcal{D}_{0,k-1;1}$ for $l \geqslant 2$ by (8), so the boundary of the region S_k is $\mathcal{D}_{0,k-1;1} \cup \mathcal{D}_{0,k;1}$, and the property of the three vectors being spanning or not remains constant inside S_k. Since outside of $\operatorname{cl} \mathcal{D}_{0,k;1}$, these vectors are *not* positively spanning by Corollary 4.9, but this changes when crossing ∂S_k, the second statement follows.

Finally, the case of real roots was dealt with at the beginning of the present Section 4. $\square$

Example 4.12. Let $d = 3$. Then

$$\bar{W} = \begin{pmatrix} p_0 & 0 \\ -q_0 & p_1 \\ r_0 & -q_1 \\ 0 & r_1 \end{pmatrix},$$

$q_0 = 2(x+1)^2 + 2y^2 - 4$, $q_1 = 2x^2 + 2y^2 - 4$, $p_0 = x^2 + y^2$, $p_1 = (x-1)^2 + y^2$, $r_0 = (x+2)^2 + y^2$ and $r_1 = (x+1)^2 + y^2$. Furthermore, $D_{0,2} = -q_0 r_1$, $D_{0,3} = q_0 q_1 - p_1 r_0$, $D_{1,3} = -p_0 q_1$, and $D_{i,j} \geqslant 0$ otherwise. Now

$$\mathcal{S}_{012} = \{z\colon\ D_{1,2} \geqslant 0,\ D_{0,2},\ D_{0,1} \geqslant 0\}, \qquad \mathcal{S}_{013} = \{z\colon\ D_{1,3},\ D_{0,3},\ -D_{0,1} \leqslant 0\},$$

$$\mathcal{S}_{023} = \{z\colon\ D_{2,3} \geqslant 0,\ -D_{0,3},\ -D_{0,2}\}, \qquad \mathcal{S}_{123} = \{z\colon\ -D_{2,3} \leqslant 0,\ -D_{1,3},\ -D_{1,2} \leqslant 0\},$$

so that by Fig. 5 and Theorem 4.11 all nonreal roots of polynomials of degree 3 with nonnegative coefficients in the binomial coefficient basis lie in the union of these regions.

5. Incorporating additional linear constraints

5.1. Linear inequalities

Suppose we not only know that the coefficients a_i of a polynomial $f = \sum_{i=0}^{d} a_i b_i$ with respect to some basis $B = \{b_i\colon\ i = 0, \dots, d\}$ are nonnegative, but also that they satisfy a linear inequality $\sum_{i=0}^{d} \lambda_i a_i \leqslant 0$; the '$\geqslant 0$' case is of course accounted for by reversing the signs of the λ_i. We use a slack variable $s \geqslant 0$ to rewrite our inequality as

$$\sum_{i=0}^{d} \lambda_i a_i + s = 0.$$

To incorporate this into our Gale dual matrices W and $\bar{W}$, we introduce the vector $\tilde{a} = (a_0, \dots, a_d, s)^T$. The analogue $W\tilde{a} = 0$ of (1) is

$$\begin{pmatrix} R_0 & R_1 & \cdots & R_d & 0 \\ I_0 & I_1 & \cdots & I_d & 0 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_d & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \\ s \end{pmatrix} = 0,$$

and we name the columns of this new W by $w_0, \dots, w_{d+1}$. We obtain a Gale dual $\tilde{W} = \tilde{W}(z)$ of W by appending the row vector

$$\bar{w}_{d+1} = (\omega_0, \dots, \omega_{d-2}) = (-\lambda_i p_i + \lambda_{i+1} q_i - \lambda_{i+2} r_i \colon 0 \leqslant i \leqslant d-2)$$

to the matrix $\bar{W}$ from (2). For the polynomial $f(z) = \sum_{i=0}^{d} a_i b_i(z)$ with $a_i \geqslant 0$ and $\sum_{i=0}^{d} \lambda_i a_i \leqslant 0$ to have a zero at $z = z_0$, the vector $\tilde{a}$ must lie in the column space of $\tilde{W}(z_0)$ and have nonnegative entries; equivalently, there must exist a vector $\mu = (\mu_0, \dots, \mu_{d-2})^T$ with $\tilde{W}(z_0)\mu = \tilde{a}$. Geometrically, we think of μ as the normal vector of a linear hyperplane that leaves all vectors $\bar{w}_i$ (weakly) on one side. In particular, if the linear inequality is strict (so that $s > 0$), then we are only interested in linear hyperplanes that do not contain $\bar{w}_{d+1}$.

In general, $m \geqslant 1$ independent linear inequalities yield a $((d+1+m) \times (d-1))$-matrix $\tilde{W}$. Consider the configuration of $d+1+m$ vectors in $\mathbb{R}^{d-1}$ spanned by the rows of $\tilde{W}$. Each $(d-2)$-tuple of vectors among these spans a linear hyperplane, and we would like to know when the $m+3$ remaining vectors all lie on the same side of it. As before, we treat strict inequalities by only considering those linear hyperplanes that do not contain any of the m "new" vectors $\bar{w}_j$, and to simplify the discussion we will focus on these.

We thus fix an ordered subset $J = \{j_1, \dots, j_{m+3}\} = \{j_1 < j_2 < j_3\} \cup \{d+1, \dots, d+m\}$ of $\{0, \dots, d+m\}$; this set will index the rows of $\tilde{W}$ *not* on a linear hyperplane. Next, we calculate a linear form $\varphi_{\bar{J}}$ whose vanishing locus is the hyperplane spanned by the $d-2$ vectors not indexed by J: it is the determinant of the matrix obtained by deleting from $\tilde{W}$ all rows indexed by J, and adding a first row of variables. The sign $\sigma_{\bar{J},i}(z)$ of $\varphi_{\bar{J}}(w_{j_i})$ at a point $z \in \mathbb{C}$ is then obtained by plugging the coordinates of $w_{j_i} = w_{j_i}(z)$ into these variables, i.e., by *not* deleting the row with index j_i, but instead permuting it to the first row and then taking the sign of the determinant of the resulting matrix. More precisely, if we denote by $\tilde{W}_K$ the matrix obtained from $\tilde{W}$ by deleting the rows indexed by $K \subset \{0, \dots, d+m\}$, then

$$\sigma_{\bar{J},i}(z) = (-1)^{j_i+i+1} \operatorname{sign}\det \tilde{W}_{J \setminus \{j_i\}}(z), \quad \text{for } i = 1, \dots, m+3. \tag{10}$$

Writing $\sigma(\bar{J}, z) = \{\sigma_{\bar{J},1}(z), \dots, \sigma_{\bar{J},m+3}(z)\}$, we can summarize our discussion as follows:

Theorem 5.1. *Assume that the coefficients of f satisfy $m \geqslant 1$ strict linear inequalities, indexed from $d+1$ to $d+m$. Let*

$$\mathcal{S}(J) = \big\{z \in \mathbb{C} \colon \sigma(\bar{J}, z) = \{-1, 0\} \text{ or } \sigma(\bar{J}, z) = \{0, +1\}\big\}.$$

Then the set of roots of f is contained in the union $\bigcup_J \mathcal{S}(J)$, where J runs through all sets of the form $\{j_1, j_2, j_3\} \cup \{d+1, \dots, d+m\}$ with $0 \leqslant j_1 < j_2 < j_3 \leqslant d$; put differently, if $\{-1, 1\} \subseteq \sigma(\bar{J}, z)$ for each such J, then z_0 is not a root of f.

In the case $m = 1$ and $J = \{j, k, l, d+1\}$, we obtain from (10) that

$$\sigma(\bar{J}, z) = \big\{(-1)^j \operatorname{sign} D_{k,l}, (-1)^{k+1} \operatorname{sign} D_{j,l}, (-1)^l \operatorname{sign} D_{j,k}, (-1)^d \operatorname{sign}\det \tilde{W}_{\{j,k,l\}}\big\}. \tag{11}$$

Expanding the last determinant along its last row yields

$$\det \tilde{W}_{\{j,k,l\}} = (-1)^d \sum_{c=0}^{d-2} (-1)^c \omega_i [\bar{W}]_{\{j,k,l\};c}, \tag{12}$$

where $[\bar{W}]_{\{j,k,l\};c}$ stands for the minor of $\bar{W}$ obtained by deleting rows j, k, l and column c. This formula can be evaluated as follows:

Lemma 5.2. *Let $m = 1$, $0 \leqslant j < k < l \leqslant d$, and $0 \leqslant c \leqslant d-2$. Then*

$$[\bar{W}]_{\{j,k,l\};c} = \begin{cases} p_0 \cdots p_{j-1} D_{j,c+1} p_{c+1} \cdots p_{k-1} D_{k,l} r_{l-1} \cdots r_{d-2} & \text{if } 0 \leqslant c \leqslant k-1, \\ p_0 \cdots p_{j-1} D_{j,k} r_{k-1} \cdots r_{c-1} D_{c+1,l} r_{l-1} \cdots r_{d-2} & \text{if } k-1 \leqslant c \leqslant d-2. \end{cases}$$

Here we follow the convention that $p_a \cdots p_b = r_a \cdots r_b = 1$ if $a > b$, but $D_{a,b} = 0$ for $a \geqslant b$. In particular, $[\bar{W}]_{\{j,k,l\};c} = 0$ for $0 \leqslant c \leqslant j-1$ and $l-1 \leqslant c \leqslant d-2$.

Proof. In each case, $\bar{W}_{K;c}$ decomposes into square blocks on the diagonal whose determinants yield the stated expressions. The elements outside these blocks do not contribute to $[\bar{W}]_{K;c}$, because the determinant of a block matrix of the form $\left(\begin{smallmatrix} A & 0 \\ C & D \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} A & B \\ 0 & D \end{smallmatrix}\right)$ is $\det A \det D$. □

To recapitulate, additional linear inequalities can only restrict further the location of possible roots of f. If the vectors $w_i(z_0), w_j(z_0), w_k(z_0) \in \mathcal{B}$ do not witness a possible root of f, in other words $\{\pm 1\} \subseteq \{(-1)^j \operatorname{sign} D_{k,l}(z_0), (-1)^{k+1} \operatorname{sign} D_{j,l}(z_0), (-1)^l \operatorname{sign} D_{j,k}(z_0)\}$, nothing changes after incorporating the additional sign $(-1)^d \operatorname{sign} \det \tilde{W}_{\{j,k,l\}}(z_0)$: the vectors $w_i(z_0), w_j(z_0), w_k(z_0), w_l(z_0)$ do still not witness a root of f at z_0. If, on the other hand, the new sign is different from the old ones, there is "one reason less" for z_0 to be a root.

5.2. Linear equations

If the coefficients of f satisfy m independent linear *equations* of the form $\sum_{i=0}^d \lambda_i a_i = 0$ (corresponding to the case $s = 0$), the $d-1-m$ columns of the new Gale dual $\tilde{W}$ will of course be linear combinations of the columns of the old one, but in general we will not be able to give an explicit expression for them. We therefore only treat some special cases that arise in the context of Ehrhart and chromatic polynomials, and defer further discussion to Section 6.2.

6. Applications

6.1. Ehrhart polynomials

From [1], we know that the following inequalities hold for the coefficients of i_Q in the binomial basis:

$$a_d + a_{d-1} + \cdots + a_{d-s} \leqslant a_0 + \cdots + a_s + a_{s+1} \quad \text{for all } 0 \leqslant s \leqslant \lfloor (d-1)/2 \rfloor.$$

For $s = 0$, the inequality reads $a_d \leqslant a_0 + a_1$, and $\bar{w}_{d+2}$ is

$$(p_0 - q_0, p_1, 0, \ldots, 0, -r_{d-2}).$$

Eq. (12) and Lemma 5.2 thus specialize as follows:

$$\det \tilde{W}_{\{0,k,d\}} = \begin{cases} (-1)^d (p_0 - q_0) D_{1,d} - (-1)^d p_1 r_0 D_{2,d} - r_0 \cdots r_{d-2} & \text{if } k = 1, \\ (-1)^d p_0 \cdots p_{k-1} D_{k,d} - D_{0,k} r_{k-1} \cdots r_{d-2} & \text{if } 2 \leqslant k \leqslant d-1, \end{cases}$$

$$\det \tilde{W}_{\{1,k,l\}} = (-1)^d q_0 p_0 \cdots p_{k-1} D_{k,l} r_{l-1} \cdots r_{d-2} - p_0 D_{1,k} r_{k-1} \cdots r_{d-2} [\![l = d]\!].$$

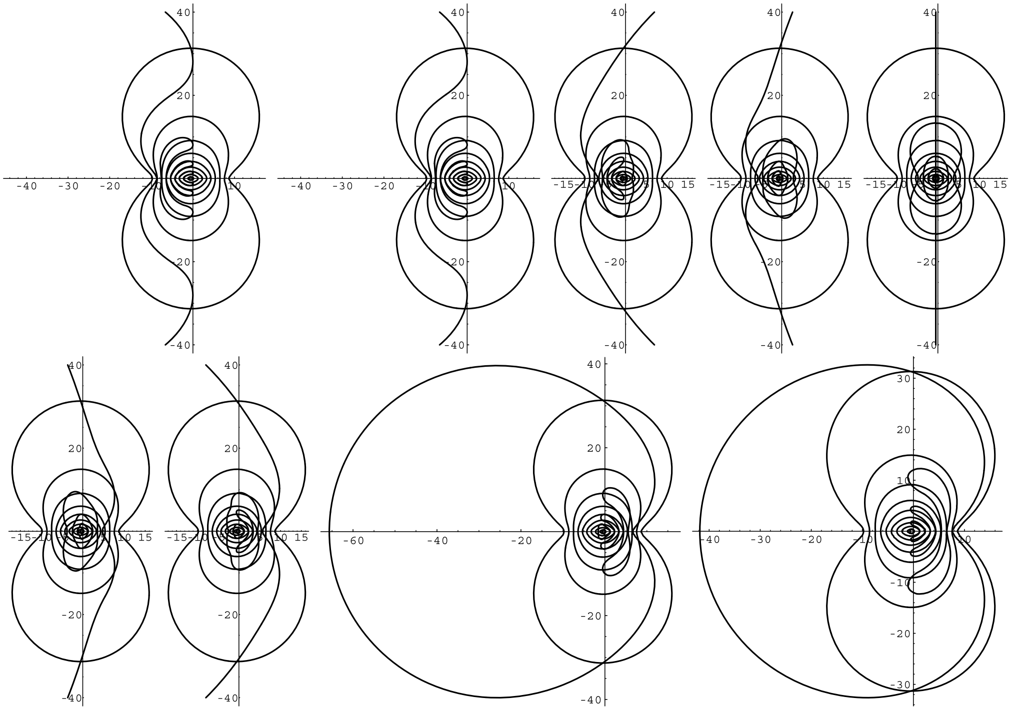

Fig. 6. From top left to bottom right, the vanishing loci of $D_{0,d}$ and $\tilde{W}_{0,k,d}$, for $d = 10$, $j = 0$, $l = d$, and $1 \leqslant k \leqslant d - 1$. In the first four pictures, the outermost oval is only partially shown, but in fact curves around to the right and intersects the real axis at a point with large positive coordinate. Thus, the points with positive real part just inside $\mathcal{D}_{0,d;1}$ lie inside the outermost component of the zero locus of $\det \tilde{W}_{\{0,k,d\}}$ for $0 < k < d/2$, but outside all components of $\det \tilde{W}_{\{0,k,d\}}$ for $d/2 < k < d$.

(Here we have used Iverson's notation: $[\![l = d]\!]$ evaluates to 1 if $l = d$, and to 0 otherwise.) Explicit calculation using Lemma 4.1 yields that the coefficient of the leading term r^{2d-2} in $(-1)^d \det \tilde{W}_{\{0,k,d\}}$ is $2(-1)^{d+1}$ for $k = 1$ and $(-1)^{d+k+1}(d - 2k)$ for $2 \leqslant k \leqslant d - 1$. Thus, the sign of this coefficient is

$$\operatorname{sign}\bigl(\bigl[r^{2d-2}\bigr](-1)^d \det \tilde{W}_{\{0,k,d\}}\bigr) = \begin{cases} (-1)^{d+k+1} & \text{for } 0 < k < d/2, \\ (-1)^{d+k} & \text{for } d/2 < k < d. \end{cases} \tag{13}$$

We examine the effect that this has on $\sigma(\bar{J}, z)$. If $z \in \mathbb{C}$ does not lie in $\operatorname{cl} D_{0,d;1}$, the first three entries of (11) already yield two different signs, no matter what sign the last determinant takes. Now let z lie inside $\operatorname{cl} D_{0,d;1}$, but outside the union of all $\operatorname{cl} D_{i,j;1}$ with $(i, j) \neq (0, d)$. If $\{j, k, l\}$ does not contain $\{0, d\}$, the first three signs of $\sigma(\bar{J}, z)$ in (11) will again contain two different ones. The interesting situation is thus $J = \{0, k, d\}$, in which case $\sigma(\bar{J}, z) = \{(-1)^{k+d+1}, (-1)^d \operatorname{sign}\det \tilde{W}_{\{0,k,d\}}(z)\}$. Combining this with (13), we see that these signs are different, i.e., z's "last opportunity" J also does not make it an Ehrhart zero, if z lies inside the outermost component of the zero locus of $\det \tilde{W}_{\{0,k,d\}}$ for $0 < k < d/2$, but outside all components of $\det \tilde{W}_{\{0,k,d\}}$ for $d/2 < k < d$. Fig. 6 shows that this actually occurs.

6.2. Chromatic polynomials

Let G be a graph on d vertices. The value of the chromatic polynomial $P(G, t)$ of G at $z = t_0$ counts the number of colorings of G with z_0 colors. The chromatic number $\chi(G)$ is the first positive integer that is not a zero of $P(G, t)$.

Proposition 6.1. *Let G be an undirected graph on d vertices with m edges, κ connected components, chromatic number $\chi = \chi(G)$, and ω acyclic orientations. Let $P(G,z) = \sum_{i=0}^{d} a_i b_i$ be the chromatic polynomial of G expressed in the basis $B = \{b_0, \dots, b_d\}$ of P_d.*

(a) *Let $b_i = (-1)^{d-i} z^i$, so that B is the* alternating power basis. *Then $a_i \geqslant 0$ for $i = 0, \dots, d$, $a_i = 0$ for $i = 0, \dots, \kappa - 1$, $a_\kappa > 0$, $a_{d-1} = m$ and $a_d = 1$* [8, Theorem 2.7].
(b) *Let $b_i = z^{\underline{i}}$, so that B is the* falling factorial basis. *Then $a_i \geqslant 0$ for $i = 0, \dots, d$* [8, Theorem 2.1].
(c) *Let $b_i = (-1)^{d-i} z^{\overline{i}}$, so that B is the* alternating rising factorial basis. *Then $a_i \geqslant 0$ for $i = 0, \dots, d$* [6, Proposition 2.1].
(d) *Let $b_i = \binom{z+d-i}{i}$, so that B is the* binomial coefficient basis. *Then $a_i \geqslant 0$ for $i = 0, \dots, d$, $\sum_{i=0}^{d} a_i = d!$, $a_i = 0$ precisely for $0 \leqslant i \leqslant \chi - 1$, and $a_d = \omega$* [5, Proposition 4.5], [6].

The roots of chromatic polynomials simultaneously satisfy all restrictions implied by these non-negativity conditions. Here we only treat two of these in any detail.

6.2.1. The alternating power basis

To evaluate these conditions for the alternating power basis, only slight modifications from the power basis case are needed. First, $q_k = -2x$ instead of $q_k = 2x$ in Proposition 2.1, and so

$$D_{j,k}(z) = -\frac{z^{n+1} - \bar{z}^{n+1}}{\bar{z} - z}$$

with $n = k - j - 1$. Next, the relations $a_0 = \dots = a_{\kappa-1} = 0$ say that effectively,

$$\begin{pmatrix} R_\kappa & R_{\kappa+1} & \cdots & R_d \\ I_\kappa & I_{\kappa+1} & \cdots & I_d \end{pmatrix} \begin{pmatrix} a_\kappa \\ \vdots \\ a_d \end{pmatrix} = 0,$$

so that $\bar{W}$ starts out with the column $(p_\kappa, -q_\kappa, r_\kappa, 0, \dots, 0)^T$. But in the present case of the alternating power basis, none of p_k, q_k, r_k actually depends on k. The matrix $\bar{W}$ thus stays the same, only the effective dimension has dropped to $d' = d - \kappa$. The discussion in Section 3.2 still applies, except that the excluded region for roots of $P(G,z)$ is now the opposite half-open sector, i.e., the cone τ bounded by the lines of angles $\pm(1 - \frac{1}{d-\kappa})\pi$.

We may incorporate the linear equation $ma_d - a_{d-1} = 0$ by appending the row vector $(0, \dots, 0, -1, m)$ of length $d - \kappa + 1$ to W, and replacing the last two columns of $\bar{W}$ by their linear combination $(0, \dots, 0, g, h, m, 1)^T$ with $g = (m - 2x)(x^2 + y^2)$ and $h = (m - 2x)2x + x^2 + y^2$. The rows of the resulting matrix $\bar{W}'$ represent $d' + 1$ vectors in $\mathbb{R}^{d'-2}$, so any linear hyperplane spanned by members of this set is defined by a linear form $\varphi_{i,j,k,l}$. The signs of the values of this linear form on the four row vectors $\bar{w}'_i, \bar{w}'_j, \bar{w}'_k, \bar{w}'_l$ are

$$\sigma_{i,j,k,l} = \big\{(-1)^i \operatorname{sign}\det \bar{W}'_{(j,k,l)}, (-1)^{j+1} \operatorname{sign}\det \bar{W}'_{(i,k,l)}, \\ (-1)^k \operatorname{sign}\det \bar{W}'_{(i,j,l)}, (-1)^{l+1} \operatorname{sign}\det \bar{W}'_{(i,j,k)}\big\},$$

where $\bar{W}'_{(i,j,k)}$, for instance, is obtained from $\bar{W}'$ by deleting rows i, j, k. The sets of signs

$$\sigma_{i,j,k,d-1} = \big\{(-1)^i \operatorname{sign} D_{j,k}, (-1)^{j+1} \operatorname{sign} D_{i,k}, (-1)^k \operatorname{sign} D_{i,j}, 0\big\}$$

tell us that any root allowed by the conditions $a_i \geqslant 0$ is also allowed under the additional restriction $ma_d = a_{d-1}$, so that the set of possible roots does not change under this restriction.

6.2.2. The binomial coefficient basis

The relations $a_0 = \dots = a_{\chi-1} = 0$ say that effectively,

$$\begin{pmatrix} R_\chi & R_{\chi+1} & \cdots & R_d \\ I_\chi & I_{\chi+1} & \cdots & I_d \end{pmatrix} \begin{pmatrix} a_\chi \\ \vdots \\ a_d \end{pmatrix} = 0,$$

so that $\bar{W}$ starts out with the column $(p_\chi, -q_\chi, r_\chi, 0, \dots, 0)^T$. The transformation $x \mapsto x + \chi$ maps $(p_{\chi+i}, -q_{\chi+i}, r_{\chi+i})$ to $(p_i, -q_i, r_i)$, so after this translation the effective dimension has dropped to $d' = d - \chi$.

The two affine linear relations $a_\chi + \dots + a_d = d!$ and $a_d = \omega$ of course do not individually influence the location of roots, but may be combined to the linear relation $\sum_{i=\chi}^{d} a_i - \frac{1}{\varepsilon} a_d = 0$ with $\varepsilon = \frac{\omega}{d!}$. A Gale dual compatible with this linear relation is the matrix $\tilde{W}$ of size $(d'+1) \times (d'-2)$ with columns

$$\tilde{W} = (v_1 - v_0, \dots, v_{d'-3} - v_0, \lambda v_{d'-2} - \mu v_0),$$

where v_i is the i-th column of $\bar{W}$, and the coefficients are $\lambda = \varepsilon d(d-1)$ and $\mu = \lambda - r_{d'-2}$. To calculate the sets $\sigma_{i,j,k,l}$ of signs, we must evaluate the determinant $[\tilde{W}]_K$ of the submatrix of $\tilde{W}$ obtained by deleting the three rows indexed by $K = \{i, j, k\}$, say. By multilinearity of the determinant, we obtain

$$\begin{aligned}
[\tilde{W}]_K &= \lambda \det(v_1 - v_0, \dots, v_{d'-3} - v_0, v_{d'-2}) - (-1)^{d'-3} \mu \det(v_0, v_1, \dots, v_{d'-3}) \\
&= \lambda \sum_{c=0}^{d'-3} (-1)^c [\bar{W}]_{K;c} + (-1)^{d'-2} \mu [\bar{W}]_{K;d'-2} \\
&= \frac{\omega}{(d-2)!} \sum_{c=0}^{d'-2} (-1)^c [\bar{W}]_{K;c} - (-1)^{d'-2} r_{d'-2} [\bar{W}]_{K;d'-2}.
\end{aligned}$$

This formula can be evaluated using Lemma 5.2. In Fig. 7 we show the zero loci of $[\tilde{W}]_K$ in the case $d = 4$ and $\omega = \frac{d!}{2}$.

7. Distribution of random roots

In closing, we explain a phenomenon encountered several times in the literature [1,4]: The roots of "randomly" generated polynomials with nonnegative coefficients tend to cluster together in several clumps, and usually lie well inside the region permitted by theory; cf. Fig. 8.

Our explanation is this: in these simulations, the coefficient vector $(a_0, \dots, a_d)$ is usually picked uniformly at random from some cube $[0, N]^{d+1}$ (except that sometimes the cases $a_0 = 0$ and $a_d = 0$ are excluded; we will gloss over this minor point). By linearity of expectation, the expected value $E(f(z_0))$ of $f(z_0) = \sum_{i=0}^{d} a_i b_i(z_0)$ at a point $z_0 \in \mathbb{C}$ is $\sum_{i=0}^{d} E(a_i) b_i(z_0) = \frac{N}{2} \sum_{i=0}^{d} b_i(z_0)$. Thus, as a first approximation, the closer the barycenter $\beta(z_0) = \sum_{i=0}^{d} b_i(z_0)$ is to zero, i.e., the smaller its absolute value $|\beta(z_0)|$, the more likely it is for z_0 to be a root of f! For example, in the case of the binomial coefficient basis,

$$\beta(z_0) = \sum_{i=0}^{d} \binom{z_0 + d - i}{d} = \sum_{i=0}^{d} \binom{z_0 + i}{d} = \binom{z_0 + d + 1}{d+1} - \binom{z_0}{d+1},$$

by an elementary identity for binomial coefficients. Fig. 9 shows the regions where $|\beta(z_0)|$ is small, together with the roots of several random polynomials. Note that $\beta(z_0)$ is the Ehrhart polynomial of the simplex $\operatorname{conv}\{e_1, \dots, e_d, -e_1 - \dots - e_d\}$ by [2, Proposition 1.3]; see also [9].

Fig. 10 shows the corresponding regions for the rising and falling factorial bases; in the case of the power basis ($b_i = z^i$), of course $\beta(z_0) = 0$ iff $z_0 \neq 1$ is a d-th root of unity.

Clearly, the predictive power of this simple model can be easily improved by considering additional parameters of the data; however, we will not do this here.

Acknowledgments

The author would like to thank Matthias Beck, Christian Haase and Günter M. Ziegler for helpful comments, and the two anonymous referees for their careful reading and their pertinent suggestions that have helped to improve the paper.

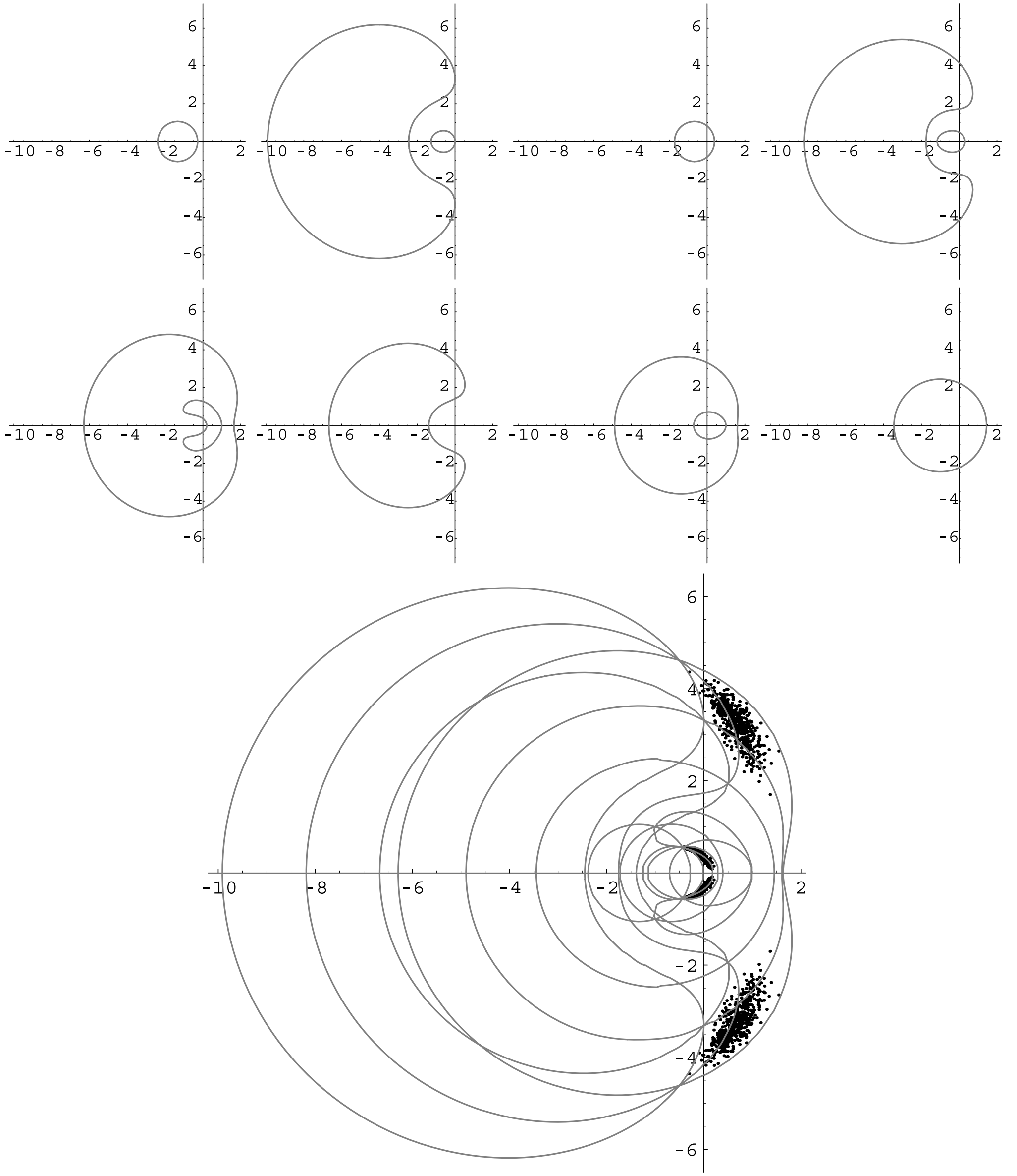

Fig. 7. From left to right and top to bottom, the zero loci of $[\tilde{W}]_K$ for $d = 4$ and $\varepsilon = \frac{1}{2}$. Here K runs through $\binom{\{0,\dots,d\}}{3}$ in lexicographic order, except that the zero loci corresponding to $K = \{0, 1, 2\}$ and $K = \{1, 2, 3\}$ are empty and not shown. The last figure combines all the zero loci with the roots of 500 random polynomials whose coefficients satisfy $\sum_{i=0}^{d} a_i - \frac{1}{\varepsilon} a_d = 0$.

Appendix A

Proof of Lemma 4.1. It suffices to do the calculation for D_n from Eq. (5). So let's expand the difference

$$\prod_{i=1}^{n+1} (z - a_i)(\bar{z} + a_i) - \prod_{i=1}^{n+1} (\bar{z} - a_i)(z + a_i), \tag{14}$$

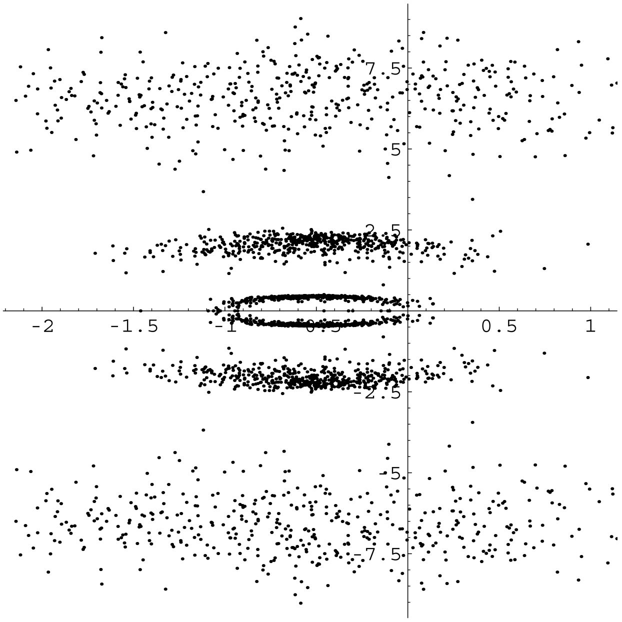

Fig. 8. The roots of 1000 random polynomials of degree $d = 6$ with nonnegative coefficients in the binomial coefficient basis and $a_0, a_d \neq 0$.

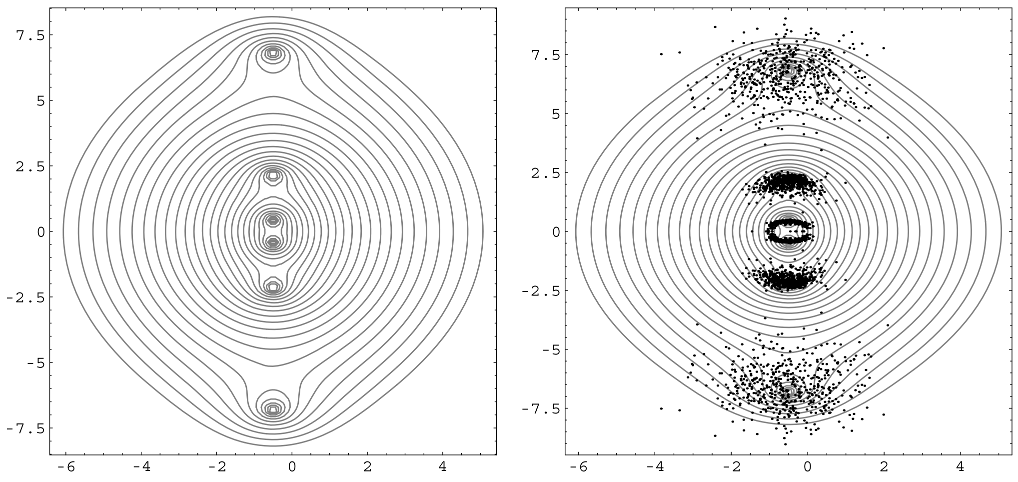

Fig. 9. *Left*: For $d = 6$, the contours $|\beta(z_0)| = c$ for $b_i = \binom{z+d-i}{d}$ and varying c; the innermost contours correspond to the smallest c. *Right*: additionally, the roots of 500 polynomials of degree d whose coefficients with respect to the b_i are chosen uniformly at random from $[0, d!]$, except that $a_0, a_d \neq 0$.

and pick out a term in the expansion with l 'z's and $m - l$ '$\bar{z}$'s. The coefficient of this term is a sum of terms of the form

$$(-1)^{m-l} a_{i_1} \cdots a_{i_{m-l}} a_{j_1} \cdots a_{j_l} - a_{i_1} \cdots a_{i_{m-l}} (-1)^l a_{i_{m-l}} a_{j_1} \cdots a_{j_l},$$

and each of these terms vanishes for m even. In particular, the term $z^{n+1}\bar{z}^{n+1}$ does not occur, which is also easy to see directly. The first nonzero term in (14) is then

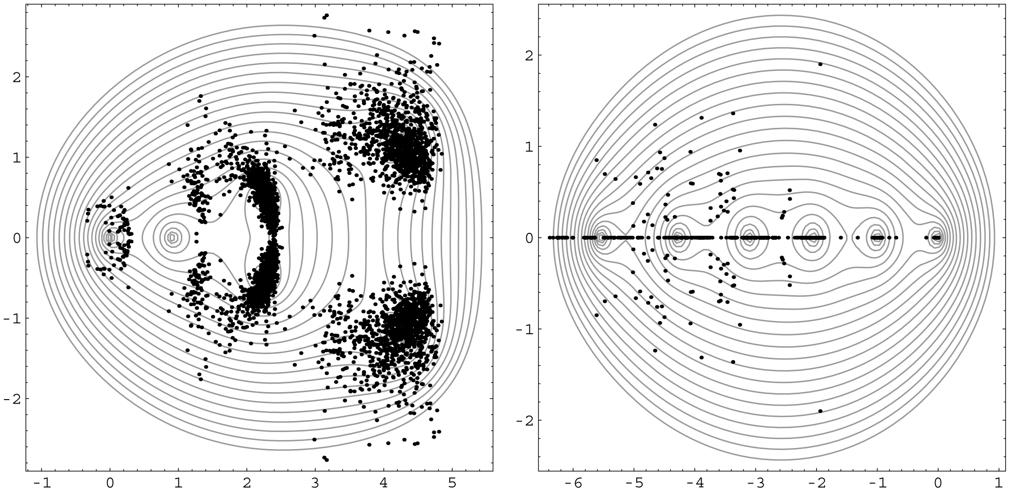

Fig. 10. Contours $|\beta(z_0)| = c$ for $d = 6$ and $b_i = z^{\underline{i}}$ (left), $b_i = z^{\overline{i}}$ (right), together with the roots of 1000, respectively 100 random polynomials with nonnegative coefficients with respect to these bases. For the rising factorial basis, the minima of $|\beta(z_0)|$ turn out to be real, so they only govern the distribution of the real zeros.

$$2z^n\bar{z}^{n+1}(-a_1 - \cdots - a_{n+1}) + 2z^{n+1}\bar{z}^n(a_1 + \cdots + a_{n+1}) = 2z^n\bar{z}^n(z - \bar{z})(a_1 + \cdots + a_{n+1}).$$

It is easy to work out $\sum_{i=1}^{n+1} a_i = d(n+1)/2$ for $a_i = i + \Delta/2$, and this finishes the proof. $\square$

Proof of Proposition 4.4. The curve $\mathcal{D} = \mathcal{D}_n$ is also described by the equation $g = (h_1 - h_2)/(\bar{z} - z)$, where $h_1 = \prod_{i=1}^{n+1}(z - a_i)(\bar{z} + a_i)$, $h_2 = \prod_{i=1}^{n+1}(z + a_i)(\bar{z} - a_i)$, and $a_i = i + \Delta/2$ with $\Delta = d - 2 - n$. Thus, $\mathcal{D}$ has a singular point if and only if the Jacobi matrix of g vanishes at some point of the locus $g = 0$. Using the chain rule and the relations $\partial z/\partial x = 1$, $\partial z/\partial y = i$, $\partial\bar{z}/\partial x = 1$, $\partial\bar{z}/\partial y = -i$, we calculate the partial derivatives of $g(z)$ with respect to x and y:

$$\frac{\partial g(z)}{\partial x} = \frac{h_{1,z} - h_{2,z} + h_{1,\bar{z}} - h_{2,\bar{z}}}{\bar{z} - z},$$
$$\frac{\partial g(z)}{\partial y} = 2i\frac{h_1 - h_2}{(\bar{z} - z)^2} + i\frac{h_{1,z} - h_{2,z} - h_{1,\bar{z}} + h_{2,\bar{z}}}{\bar{z} - z}.$$

Here $h_{j,z}$, $h_{j,\bar{z}}$ denote the partial derivatives of h_j with respect to z, $\bar{z}$; by explicit differentiation, $h_{1,z} = \sum_{i=1}^{n+1} h_1/(z - a_i)$ and $h_{2,z} = \sum_{i=1}^{n+1} h_2/(z + a_i)$.

To prove that $\mathcal{D}$ has no real singular points, we pick $z \in \mathcal{D}$ and calculate

$$\begin{aligned}
\frac{\partial g}{\partial x}(z) &= \frac{h_{1,z} + h_{1,\bar{z}}}{\bar{z} - z} - \frac{h_{2,z} + h_{2,\bar{z}}}{\bar{z} - z} \\
&= \frac{h_1}{\bar{z} - z}\sum_{i=1}^{n+1}\left(\frac{1}{z - a_i} + \frac{1}{\bar{z} + a_i} - \frac{1}{z + a_i} - \frac{1}{\bar{z} - a_i}\right) \\
&= \frac{h_1}{\bar{z} - z}\sum_{i=1}^{n+1}\frac{2a_i(\bar{z}^2 - z^2)}{(z^2 - a_i^2)(\bar{z}^2 - a_i^2)} \\
&= 4h_1\sum_{i=1}^{n+1}\frac{xa_i}{((x - a_i)^2 + y^2)((x + a_i)^2 + y^2)}.
\end{aligned}$$

For real nonzero $z \in \mathcal{D}$, this expression never vanishes. The same calculation already proves the second statement, because a tangent vector to the curve $g(x, y) = 0$ at a nonsingular point (x_0, y_0) is given by $\pm(-\frac{\partial g}{\partial y}(x_0, y_0), \frac{\partial g}{\partial x}(x_0, y_0))$.

We now examine a nonreal singular point z_0 of $\mathcal{D}$. Any such point must satisfy

$$\frac{h_1(z_0) - h_2(z_0)}{\bar{z}_0 - z_0} = 0 = h_{1,z}(z_0) - h_{2,z}(z_0).$$

The first equation tells us that $h_1(z_0) = h_2(z_0)$, so that $h_{1,z}(z_0) - h_{2,z}(z_0) = 0$ if and only if $h_1(z_0) = 0$ (which is incompatible with $z_0 \notin \mathbb{R}$ and $g(z_0) = 0$), or

$$0 = \sum_{i=1}^{n+1} \frac{1}{z_0 - a_i} - \sum_{i=1}^{n+1} \frac{1}{z_0 + a_i} = 2 \sum_{i=1}^{n+1} \frac{a_i}{z_0^2 - a_i^2}.$$

Writing $z_0^2 = x_0 + i y_0$ and separating the real and imaginary parts in the last expression yields

$$\sum_{i=1}^{n+1} \frac{a_i}{(x_0 - a_i)^2 + y_0^2} = \sum_{i=1}^{n+1} \frac{a_i^3}{(x_0 - a_i)^2 + y_0^2} = 0.$$

But the denominators of these expressions are positive (the a_i and the origin do not lie on $\mathcal{D}$), and $a_i > 0$ for $i = 1, \ldots, n+1$, so we conclude that $\mathcal{D}$ has no singular points. □

References

[1] M. Beck, J.A. De Loera, M. Develin, J. Pfeifle, R.P. Stanley, Coefficients and roots of Ehrhart polynomials, in: Integer Points in Polyhedra—Geometry, Number Theory, Algebra, Optimization, in: Contemp. Math., vol. 374, Amer. Math. Soc., Providence, RI, 2005, pp. 15–36.
[2] C. Bey, M. Henk, J.M. Wills, Notes on the roots of Ehrhart polynomials, Discrete Comput. Geom. 38 (2007) 81–98.
[3] B. Braun, Norm bounds for Ehrhart polynomial roots, arXiv:math.CO/0602464, 3 pp.
[4] B. Braun, M. Develin, Ehrhart polynomial roots and Stanley's non-negativity theorem, arXiv:math.CO/0610399, 11 pp.
[5] F. Brenti, Expansions of chromatic polynomials and log-concavity, Trans. Amer. Math. Soc. 332 (1992) 729–756.
[6] F. Brenti, G.F. Royle, D.G. Wagner, Location of zeros of chromatic and related polynomials of graphs, Canad. J. Math. 46 (1994) 55–80.
[7] D. Hall, J. Siry, B. Vanderslice, The chromatic polynomial of the truncated icosahedron, Proc. Amer. Math. Soc. 16 (1965) 620–628.
[8] R. Read, W. Tutte, Chromatic polynomials, in: Selected Topics in Graph Theory, vol. 3, Academic Press, New York, 1988, pp. 15–42.
[9] F. Rodriguez-Villegas, On the zeros of certain polynomials, Proc. Amer. Math. Soc. 130 (2002) 2251–2254.
[10] A.D. Sokal, Chromatic roots are dense in the whole complex plane, Combin. Probab. Comput. 13 (2004) 221–261.
[11] R.P. Stanley, Decompositions of rational convex polytopes, in: Combinatorial Mathematics, Optimal Designs and Their Applications, Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, CO, 1978, in: Ann. Discrete Math., vol. 6, 1980, pp. 333–342.
[12] G.M. Ziegler, Lectures on Polytopes, Grad. Texts in Math., vol. 152, Springer-Verlag, 1995.