Diagonal asymptotics for symmetric rational functions via ACSV[1]

Abstract: We consider asymptotics of power series coefficients of rational functions of the form $1/Q$ where Q is a symmetric multilinear polynomial. We review a number of such cases from the literature, chiefly concerned either with positivity of coefficients or diagonal asymptotics. We then analyze coefficient asymptotics using ACSV (Analytic Combinatorics in Several Variables) methods. While ACSV sometimes requires considerable overhead and geometric computation, in the case of symmetric multilinear rational functions there are some reductions that streamline the analysis. Our results include diagonal asymptotics across entire classes of functions, for example the general 3-variable case and the Gillis-Reznick-Zeilberger (GRZ) case [GRZ83], where the denominator in terms of elementary symmetric functions is $1 - e_1 + ce_d$ in any number d of variables. The ACSV analysis also explains a discontinuous drop in exponential growth rate for the GRZ class at the parameter value $c = (d-1)^{d-1}$, previously observed for $d = 4$ only by separately computing diagonal recurrences for critical and noncritical values of c. We prove this phenomenon for any even value of $d \geq 4$, show that the exponential drop occurs not only on the diagonal but in an open cone, and give the leading asymptotic behavior throughout the cone. This last result depends on identifying a chain of integration as homologous to a sum of signed chains that partially cancel (the so-called *lacuna* of [ABG70]).

Yuliy Baryshnikov, University of Illinois, Department of Mathematics, 273 Altgeld Hall 1409 W. Green Street (MC-382), Urbana, IL 61801, `ymb@illinois.edu`, partially supported by NSF grant DMS-1622370.

Stephen Melczer, University of Pennsylvania, Department of Mathematics, 209 South 33rd Street, Philadelphia, PA 19104, `smelczer@sas.upenn.edu`, partially supported by an NSERC postdoctoral fellowship.

Robin Pemantle, University of Pennsylvania, Department of Mathematics, 209 South 33rd Street, Philadelphia, PA 19104, `pemantle@math.upenn.edu`, partially supported by NSF grant DMS-1612674.

Armin Straub, University of South Alabama, Department of Mathematics and Statistics, 411 University Blvd N, MSPB 325. Mobile, AL 36688, `straub@southalabama.edu`, partially supported by a Simons Collaboration Grant.

Subject classification: 05A16; secondary 57Q99.

Keywords: Analytic combinatorics, Generating function, Coefficient, Lacuna, Positivity, Morse theory, D-finite, Smooth point

---

[1] All authors gratefully acknowledge the support and hospitality of the Erwin Schrödinger Institute

1 Introduction

We study the power series coefficients of rational functions of the form $F(x_1, \ldots, x_d) = 1/Q(x_1, \ldots, x_d)$ where Q is a symmetric multilinear function with $Q(\mathbf{0}) \neq 0$. Let

$$F(\mathbf{x}) = \frac{1}{Q(\mathbf{x})} = \sum_{\mathbf{r} \in \mathbb{Z}^d} a_{\mathbf{r}} \mathbf{x}^{\mathbf{r}},$$

converging in some polydisk $\mathcal{D} \subset \mathbb{C}^d$. Often one focuses on the diagonal coefficients $\delta_n := a_{n,\ldots,n}$, whose univariate generating function $\mathsf{diag}_F(z) := \sum_n \delta_n z^n$ satisfies a linear differential equation with polynomial coefficients, but may be transcendental. A number of questions are natural, including nonnegativity (are all coefficients nonnegative), eventual nonnegativity (all but finitely many coefficients nonnegative), diagonal extraction (computing diag_F from Q), diagonal asymptotics, multivariate asymptotics and phase transitions in the asymptotics of $\{a_{\mathbf{r}}\}$.

The positivity (nonnegativity) question is the most classical, dating back at least to Szegő's work in [Sze33]. The techniques, some of which are indicated in the next section, used in the literature are diverse and include integral methods and special functions, positivity preserving operators, combinatorial identities, computer algebra such as cylindrical algebraic decomposition, or determinantal methods. Contrasting to these methods are analytic combinatorial several-variable methods (ACSV) as developed in [PW13]. These are typically asymptotic, rather than exact, and therefore less useful for proving classical positivity statements, though they can be used to disprove them. Their chief advantages are their broad applicability and, increasingly, the level to which they have been automated. Our aim in this paper is to apply ACSV methods to a number of previously studied family of rational coefficient sequences, thereby extending what is known as well as illuminating the relative advantages of each method.

1.1 Previously studied instances

Let $\mathcal{M}_d$ denote the class of symmetric functions of d variables that are multilinear (degree 1 in each variable). This class of generating functions $F(\mathbf{x}) := 1/Q(\mathbf{x})$ where $Q \in \mathcal{M}_d$ includes a great number of previously studied cases, some of which we now review. Here and in the following, we use d for the number of variables and boldface $\mathbf{x}, \mathbf{y}, \mathbf{z}$, etc., for vectors of length d of integer, real or complex numbers. When d is small we use x, y, z, w for x_1, x_2, x_3, x_4. Let $e_k = e_{k,d}$ denote the k^{th} elementary symmetric function of d variables. An equivalent description of the class $\mathcal{M}_d$ is that it contains all linear combinations of $\{e_{k,d} : 0 \leq k \leq d\}$.

The Askey-Gasper rational function is

$$A(x, y, z) := \frac{1}{1 - x - y - z + 4xyz}, \tag{1.1}$$

which, in the previous notation, is $A(\mathbf{x}) = F(\mathbf{x})$ when $d = 3$ and $Q = 1 - e_1 + 4e_3$. Gillis, Reznick and Zeilberger [GRZ83] deduce positivity of A from positivity of a 4-variate extension due to Koornwinder [Koo78], for which they give a short elementary proof using a positivity preserving operation. Gillis, Reznick and Zeilberger also provide an elementary proof of the stronger result by Askey and Gasper [AG77] that A^β is positive for $\beta \geq (\sqrt{17} - 3)/2 \approx 0.56$, by deriving a recurrence relation for the coefficients that makes positivity apparent.

Specific functions in $\mathcal{M}_4$ that have shown up in the literature include the Szegő rational function

$$S(x,y,z,w) := \frac{1}{e_3(1-x,1-y,1-z,1-w)} \tag{1.2}$$

as well as the Lewy-Askey function

$$L(x,y,z,w) := \frac{1}{e_2(1-x,1-y,1-z,1-w)}, \tag{1.3}$$

which is a rescaled version of $1/Q(\mathbf{x})$ with $d = 4$ and $Q = 1 - e_1 + \frac{2}{3}e_2$. Szegő [Sze33] proved that (1.2) is positive. In fact, he showed that $e_{d-1,d}^{-\beta}(1-\mathbf{x})$ is nonnegative if $\beta \geq 1/2$. His proof relates the power series coefficients to integrals of products of Bessel functions and, among other ingredients, employs the Gegenbauer–Sonine addition theorem. Scott and Sokal [SS13] establish a vast and powerful generalization of this result by showing that, if T_G is the spanning-tree polynomial of a connected series-parallel graph, then $T_G^{-\beta}(1-\mathbf{x})$ is nonnegative if $\beta \geq 1/2$. In the simplest non-trivial case, if G is a d-cycle, then $T_G = e_{d-1,d}$, thus recovering Szegő's result. Relaxing the condition on β, Scott and Sokal further extend their results to spanning-tree polynomials of general connected graphs. They do so by realizing that Kirchhoff's matrix-tree theorem implies that these polynomials can be expressed as determinants, and by proving that determinants of this kind are nonnegative. As another consequence of this determinantal nonnegativity, Scott and Sokal conclude that (1.3) is nonnegative, thus answering a question originating with Lewy [AG72] (with positivity replaced by nonnegativity). Kauers and Zeilberger [KZ08] show that positivity of the Lewy-Askey rational function (1.3) would follow from positivity of the four variable function

$$K(x,y,z,w) := \frac{1}{1 - e_1 + 2e_3 + 4e_4}. \tag{1.4}$$

However, the conjectured positivity (or even nonnegativity) of (1.4) remains open.

As noted above, $e_{d-1,d}^{-\beta}(1-\mathbf{x})$ is nonnegative if $\beta \geq 1/2$. The asymptotics of $e_{k,d}^{-\beta}(1-\mathbf{x})$ are computed in [BP11] for $(k,d) = (2,3)$. They are asymptotically positive when $\beta > 1/2 = (d-k)/2$ in this case, and not when $\beta < 1/2$. A conjecture of Scott and Sokal that remains open in both directions is that, for general k and d, the condition $\beta \geq (d-k)/2$ is necessary and sufficient for nonnegativity of the coefficients of $e_{k,d}^{-\beta}(1-\mathbf{x})$.

Gillis, Reznick and Zeilberger [GRZ83] consider the family

$$F_{c,d}(x_1,\ldots,x_d) := \frac{1}{1 - e_1 + c\, e_d}\,. \tag{1.5}$$

of rational functions, where c is a real parameter. When $c < 0$, the coefficients are trivially positive, therefore it is usual to assume $c > 0$. Gillis, Reznick and Zeilberger show that $F_{c,3}$ has nonnegative coefficients if $c \leq 4$ (and this condition is shown to be necessary in [Str08]), but they conjecture that the threshold for $d \geq 4$ has a different form, namely that $F_{c,d}$ has nonnegative coefficients if and only if $c \leq d!$. It is claimed in [GRZ83], but the proof is omitted due to its length, that nonnegativity of $F_{d!,d}$ is implied by nonnegativity of the diagonal power series coefficients. In the cases $d = 4, 5, 6$, Kauers [Kau07] proved nonnegativity of these diagonal coefficients by applying cylindrical algebraic decomposition (CAD) to the respective recurrences. On the other hand, it is suggested in [SZ15] that the diagonal coefficients are eventually positive if $c < (d-1)^{d-1}$.

1.2 Previous questions and results on diagonals

The diagonal generating function diag_F and the sequence $\delta_n := a_{n,\ldots,n}$ it generates have received special attention. One reason is that the question of multivariate asymptotics in the diagonal direction is simply stated, whereas the question of asymptotics in all possible directions requires discussion of different possible phase regimes, a notion of uniformity over directions, degeneracies when the coordinates are not of comparable magnitudes, and so forth. Another reason is that there are effective methods for determining diag_F from Q, transferring the problem to the familiar univariate realm.

We briefly recall the theory of diagonal extraction. A d-variate power series F is said to be D-finite if the formal derivatives $\{\partial_{\mathbf{r}} F : \mathbf{r} \in (\mathbb{Z}^+)^d\}$ form a finite dimensional vector space over $\mathbb{C}[\mathbf{x}]$. In one variable, this is equivalent to F satisfying a linear differential equation with polynomial coefficients,

$$\sum_{i=0}^{k} q_i(z) \frac{d^i}{dz^i} F = 0\,.$$

Proposition 1.1 (D-finiteness of rational diagonals [Lip88])**.** *Let $F(z)$ be a D-finite power series. Then $\mathsf{diag}(z) := \sum_n \delta_n z^n$ is D-finite, where $\delta_n := a_{n,\ldots,n}$.* □

When F is a rational function and $d = 2$, it was known that diag is algebraic (and thus D-finite) at least as early as the late 1960's in a result of Furstenburg, recorded in [HK71], and in special cases by Pólya in the 1920's [Pól21]. In the rational function $F(x, y) = P(x, y)/Q(x, y)$ one substitutes $y = 1/x$ and computes a residue integral to extract the constant coefficient. The basis for Lipshitz' proof was the realization that the complex integration can be viewed as purely formal. With the advent of computer algebra this formal D-module computation was automated, with an early package in Macaulay and more widely used modern implementations in Magma, Mathematica and Maple. Due to advances in software and processor speed, these computations are often completable on functions arising in applications.

The following relationship between D-finiteness of a univariate function and the existence of a polynomial recursion satisfied by its coefficient sequence is the result of translating a formal differential equation into a relation among the coefficients.

Proposition 1.2. *The series $f(z) = \sum_{n\geq 0} a_n z^n$ is D-finite if and only if it is polynomially recursive, meaning that there is a $k > 0$ and there are polynomials $p_0, \ldots, p_k$, not all zero, such that for all but finitely many n,*

$$\sum_{i=0}^{k} p_i(n) f(n+i) = 0\,.$$

Let f be a D-finite power series in one variable. If f has positive finite radius of convergence and integer coefficients, then it is a so-called *G-function* and has well behaved asymptotics according to following result.

Proposition 1.3 (Asymptotics of G-Function Coefficients)**.** *Suppose f is D-finite with finite radius of convergence and integer coefficients annihilated by a minimal order linear differential operator $\mathcal{L}$ with polynomial coefficients. Then $\mathcal{L}$ has only regular singular points in the Frobenius sense. Consequently, the coefficients $\{a_n\}$ are given asymptotically by a formula*

$$a_n \sim \sum_{\alpha} C_\alpha n^{\beta_\alpha} \rho_\alpha^{-n} (\log n)^{k_\alpha} \tag{1.6}$$

where the sum is over quadruples $(C_\alpha, b_\alpha, \rho_\alpha, k_\alpha)$ as α ranges over a finite set A with the following properties. The base ρ_α is an algebraic number, a root of the leading polynomial coefficient of $\mathcal{L}$. The β_α are rational and for each value of ρ_α can be determined as roots of an explicit polynomial constructed from ρ_α and $\mathcal{L}$.

The log powers k_α are nonnegative integers, zero unless for fixed ρ_α there exist two values of β_α differing by an integer (including multiplicities in the construction of β_α). The C_α are not in general closed form analytic expressions, but may be determined rigorously to any desired accuracy.

PROOF: See discussion in [Mel17, page 37] for references to several published results that together establish this proposition. Determination of all rational and algebraic numbers other than C_α is known to be effective. $\square$

Because there are computational methods for the study of diagonals, it is of interest to reduce positivity questions to those involving only diagonals. For the Gillis-Reznick-Zeilberger class $F_{c,d}$, such a result is conjectured.

Conjecture 1.4 ([GRZ83])**.** *For $d \geq 4$, the following three statements are equivalent.*

(i) $c \leq d!$

(ii) The diagonal coefficients of $F_{c,d}$ are nonnegative

(iii) All coefficients of $F_{c,d}$ are nonnegative

To be precise, $(iii) \Rightarrow (ii) \Rightarrow (i)$ is trivial (look at δ_1); nonnegativity of all coefficients of $F_{c,d}$ holds for some interval $c \in [0, c_{\max}]$, therefore the conjecture comes down to nonnegativity of $F_{d!,d}$. A proof for $(iii) \Rightarrow (ii)$ in the case $c = d!$ is claimed in [GRZ83] but omitted from the paper due to length. This question is generalized in [SZ15] to all of $\mathcal{M}_d$.

Question 1 ([SZ15, Question 1.1 and following])**.** *For $Q \in \mathcal{M}_d$ and $F = 1/Q$, under what conditions does nonnegativity of the coefficients of diag_F imply nonnegativity of all coefficients of F?*

The authors of that paper detail a necessary condition and ask whether nonnegativity of the coefficients of $F(x_1, \ldots, x_{d-1}, 0)$ suffices. Restricting to $d = 3$, they conjecture both halves of a sharp criterion for nonegativity of F.

Conjecture 1.5 ([SZ15, Conjecture 3.2])**.** *When $d = 3$, nonnegativity of the diagonal together with nonnegativy of $F(x, y, 0)$ is enough to imply nonnegativity of F.*

Conjecture 1.6 ([SZ15, Conjecture 3.3])**.** *Let $F = 1/Q$ where $Q = 1 - e_2 + ae_2 + be_3$, which is, up to rescaling, the general element of $\mathcal{M}_3$. Then diag_F is nonnegative in the interesting case $a \geq 1$ (equivalent to nonnegativity of $F(x, y, 0)$) if and only if $b \leq -a^3$.*

1.3 Present results

In the present work we use ACSV to answer asymptotic versions of these questions. Aside from computing special cases, the main new results are (1) simplification for diagonals with symmetric denominators via the Grace-Walsh-Szegö Theorem (Lemma 2.3 below); (2) an easy further simplification for the Gillis-Reznick-Zeilberger class (Lemma 4.1 below); and (3) a topological computation to explain drop in magnitude of coefficients at critical parameter values (Theorem 5.1 below).

The first special case we look at is the diagonal of the general element of $\mathcal{M}_3$, corresponding to Conjecture 1.6.

Theorem 1.7. *Let* $Q = 1 - e_1 + ae_2 + be_3$*, let* $F = 1/Q = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}}$ *and let* $\delta_n = a_{n,\dots,n}$ *be the diagonal coefficients of* F*. If* $r_1(a)$ *denotes the upper branch of the discriminant of* $Q(x,x,x)$ *then* δ_n *is eventually positive when*

$$b < \begin{cases} -9a & a \leq -3 \\ r_1(a) & -3 \leq a \leq 1 \\ -a^3 & a \geq 1 \end{cases} \tag{1.7}$$

while, when the inequality is reversed, δ_n *attains an infinite number of positive and negative values.*

Theorem 1.7 is obtained by examining asymptotic regimes, captured in the following result.

Theorem 1.8. *Let* $Q, F,$ *and* δ_n *be as in Theorem 1.7. Assuming that* b *is not equal to the piecewise function in Equation (1.7),*

$$\delta_n = \sum_{x \in E} \left(\frac{x^{-3n}}{n} \cdot \left| \frac{1 - 2ax - bx^2}{1 - ax} \right| \cdot \frac{1}{2\sqrt{3}(1 - 2x + ax^2)} \right) \left(1 + O\left(\frac{1}{n}\right) \right), \tag{1.8}$$

where E *consists of the minimal modulus roots of the polynomial* $Q(x,x,x) = 1 - 3x + 3ax^2 + bx^3$*.*

The situation when equality does hold in Equation (1.7) is more delicate.

Proposition 1.9. *Suppose equality holds in Equation (1.7). The following table captures behaviour of the diagonal; refer to Section 2 for notational information related to our asymptotic results.*

	Minimal Points	*Eventual Sign*
$a < -3$	$\pm r$ *for some* $r > 0$	*Positive*
$a = -3$	$\rho_\pm = \pm(1/3, 1/3, 1/3)$ ρ_+ *a cone point,* ρ_- *a degenerate smooth point*	*Conjecturally Alternates*
$-3 < a < 1$	*cone point* (r,r,r) *with* $r = (1 - \sqrt{1-a})/a$	*Positive*
$a = 1$	$(1,1,1)$	*Sequence identically 1*
$a > 1$	$\rho_{0,1,2}$ *where* $\rho_k = e^{2k\pi i/3}(1/a, 1/a, 1/a)$ *all smooth, but* ρ_0 *degenerate*	*Conjecturally Positive*

When $a < -3$ *asymptotics are given by evaluating Equation (1.8) with* $E = \{\pm r\}$*. When* $-3 < a < 1$ *the results of Baryshnikov and Pemantle [BP11, Theorem 3.7] give asymptotics*

$$\delta_n = \left(\frac{\left((4-a)\sqrt{1-a} - 3a + 4\right)^n}{n} \right) \left(\frac{\sqrt{3}}{3\pi(\sqrt{1-a} - a + 1)} + O\left(\frac{1}{n}\right) \right).$$

Checking numerically this is off by some constant factor (depending on a*).* *When* $a = 1$ *the sequence is identically 1. When* $a = -3$ *or* $a > 1$*, asymptotics are given by a saddle-point integral with degenerate phase, which is beyond the scope of this article. The asymptotics appear to have the form* $Ca^{3n}/n^{2/3}$ *for constant* C*; these cases will be handled in a forthcoming extension of this article.*

Our second set of results concern the diagonal of the general element of GRZ rational function $F_{c,d}$. Let

$$c_* = c_*(d) := (d-1)^{d-1} \,. \tag{1.9}$$

The following corresponds to Conjecture 1.4.

Theorem 1.10. *Let $d \geq 4$. Then the diagonal coefficients of $F_{c,d}$ are eventually positive when $c < c_*$ and contain an infinite number of positive and negative values when $c > c_*$. When $c < c_*$, there is a conical neighborhood $\mathcal{N}$ of the diagonal such that $a_{\mathbf{r}} > 0$ for all but finitely many $\mathbf{r} \in \mathcal{N}$.*

Again, the result is obtained through an explicit asymptotic analysis.

Theorem 1.11. *Let δ_n be the diagonal coefficients of $F_{c,d}$. Then when $c \neq c_*$,*

$$\delta_n = \sum_{x \in E} \left(\frac{x^{-dn}}{n^{(d-1)/2}} \cdot \left(\frac{2\pi(1-(d-1)r)}{r^{(d-1)/2}} \right)^{(d-1)/2} \cdot \frac{1}{d^{1/2}(1-(d-1)r)} \right) \left(1 + O\left(\frac{1}{n}\right) \right),$$

where E consists of the minimal modulus roots of the polynomial $1/F_{c,d}(x,\ldots,x) = 1 - dx + cx^d$.

These theorems are proven in Section 4, using ACSV smooth point methods summarized in Section 2, however the case $c = c_*$ for the GRZ rational function requires the more delicate results of Section 5.

1.4 Exponential drop and further results

In the GRZ family, for even values of $d \geq 4$ the exponential growth rate of the coefficients drops at the special value $c = (d-1)^{d-1}$. This special value, and the corresponding drop in exponential growth, may be identified for each fixed d from the differential equation annihilating diag_F. For example, when $d = 4$ an annihilating differential equation for $F_{c,4}$ is computed by D-module integration in the Mathematica package of Koutschan [Kou10] producing the annihilating operator $\mathcal{L}$, of order 3 and maximum coefficient degree 6, such that $\mathcal{L}F_{c,4} = 0$ for all c:

$$\begin{aligned}
\mathcal{L} = {} & w^2(c^4w^4 + 4c^3w^3 + 6c^2w^2 + 4cw - 256w + 1)(3cw-1)^2\partial_w^3 \\
& - 3w(3cw-1)(6c^5w^5 + 15c^4w^4 + 8c^3w^3 - 6c^2w^2 - 384cw^2 - 6cw + 384w - 1)\partial_w^2 \\
& - (cw+1)(63c^5w^5 - 3c^4w^4 - 66c^3w^3 + 18c^2w^2 + 720cw^2 + 19cw - 816w + 1)\partial_w \\
& - 9c^6w^5 + 3c^5w^4 + 6c^4w^3 - 18c^3w^2 + 360c^2w^2 - 13c^2w + 384cw - c + 24.
\end{aligned} \tag{1.10}$$

When $c = 27$, all coefficients in (1.10) acquire enough zeros at $w = 1/81$ that the quantity $(81w-1)^4$ may be factored out of the entire operator, leaving the following operator of order 4 and maximum degree 2:

$$\begin{aligned}
\mathcal{L}_{27} := {} & w^2(81w^2 + 14w + 1)\,\partial_w^3 + 3w(162w^2 + 21w + 1)\,\partial_w^2 \\
& + (21w+1)(27w+1)\partial_w + 3(27w+1).
\end{aligned} \tag{1.11}$$

Asymptotics for δ_n may be extracted via the methodology described in Proposition 1.3. In the special case $d = 4, c = 27$, the recursion may be found on the OEIS (entry A125143) and identifies $\{\delta_n\}$ as the

Almkvist–Zudilin numbers[2] from [AvSZ11, sequence (4.12)(δ)]. The known asymptotic formula implies that $|\delta_n|^{1/n} \to 9$. However, as $c \neq 27$ approaches 27 from either side, we have

$$\lim_{c \to 27} \lim_{n \to \infty} |\delta_n|^{1/n} = 81;$$

in other words, the growth rate at $c = 27$ drops suddenly from 81 to 9. The occurrence of a phase change at $(d-1)^{d-1}$ for all d and drop in exponential rate for even $d \geq 4$ had not previously been proved. The special role of the case $c = (d-1)^{d-1}$ was observed in [SZ15, Example 4.4] and claimed to agree with intuition from hypergeometric functions. We verify this, first by identifying the singularity from an ACSV point of view and then by checking that this singularity indeed produces the observed dimension drop.

Theorem 1.12 (exponential growth approaching criticality)**.** *For all $d \geq 2$,*

$$\lim_{c \to c_*} \limsup_{n \to \infty} |\delta_n|^{1/(dn)} = d - 1$$

Theorem 1.13 (dimension drop at criticality)**.** *When $c = c_*$ and $d \geq 4$ is even,*

$$\limsup_{n \to \infty} |\delta_n|^{1/(dn)} < d - 1\,.$$

Theorem 1.13 is proved in Section 5.

2 ACSV

In this section we describe the basic setup for ACSV and state some existing results. Throughout this section let $F(\mathbf{z}) = P(\mathbf{z})/Q(\mathbf{z}) = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}}$ denote a rational series in d variables, with P and Q co-prime polynomials. Assume that F has a (finite) positive radius of convergence; that is, $Q(\mathbf{0}) \neq 0$ and P/Q is not a polynomial. Let $\mathcal{V} := \{\mathbf{z} \in \mathbb{C}^d : Q(\mathbf{z}) = 0\}$ denote the singular variety for F and let $\mathcal{M} = (\mathbb{C}^*)^d \setminus \mathcal{V}$ where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Coefficients $a_{\mathbf{r}}$ are extracted via the multivariable Cauchy formula

$$a_{\mathbf{r}} = \frac{1}{2\pi i} \int_{\mathbf{T}} \mathbf{z}^{-\mathbf{r}} F(\mathbf{z}) \frac{d\mathbf{z}}{\mathbf{z}} \tag{2.1}$$

where $d\mathbf{z}/\mathbf{z}$ denotes the holomorphic logarithmic volume form $(dz_1/z_1) \wedge \cdots \wedge (dz_d/z_d)$ and $\mathbf{T}$ denotes a small torus (a product of sufficiently small circles about the origin in each coordinate that the product of the corresponding disks is disjoint from $\mathcal{V}$). The fundamental insight of ACSV is that the integral depends only on the homology class of $\mathbf{T}$ in $H_d(\mathcal{M})$. Therefore, one tries to replace $\mathbf{T}$ by some homologous chain $\mathcal{C}$ over which the integral is easier, typically via some combination of residue reductions and saddle point estimates.

A *direction* of asymptotics is an element $\hat{\mathbf{r}} \in (\mathbb{RP}^d)^+$, that is, a projective vector in the positive orthant; if $\mathbf{r} \in \mathbb{R}^d$ we write $\hat{\mathbf{r}}$ to denote the projective equivalence class containing $\mathbf{r}$, and set $|\mathbf{r}| = |\mathbf{r}|_1 := r_1 + \cdots + r_d$. Given a Whitney stratification of $\mathcal{V}$ into smooth manifolds, the *critical set* $\mathsf{crit}(\hat{\mathbf{r}})$ for a direction $\hat{\mathbf{r}}$ is the set of $\mathbf{z} \in \mathcal{V}$ such that $\hat{\mathbf{r}}$ is orthogonal to the tangent space of the stratum of $\mathbf{z}$ in $\mathcal{V}$. If $\mathbf{z}$ is a smooth point of $\mathcal{V}$ and Q is square-free, this means $\hat{\mathbf{r}}$ should be parallel to the logarithmic gradient $(z_1 \partial Q/\partial z_1, \ldots, z_d \partial Q/\partial z_d)$. A *minimal* point for direction $\hat{\mathbf{r}}$ is a point $\mathbf{z} \in \mathsf{crit}(\hat{\mathbf{r}})$ such that the open polydisk $\mathcal{D}(\mathbf{z}) := \{\mathbf{w} : |w_j| < |z_j| \, \forall 1 \leq j \leq d\}$ does not intersect $\mathcal{V}$. The minimal point $\mathbf{z}$ is called *strictly minimal* if the closed polydisk $\overline{\mathcal{D}(\mathbf{z})}$ intersects $\mathcal{V}$ only at $\mathbf{z}$. Except in Section 5, all ACSV computations are based on the following result.

---

[2] That these are the diagonals of the rational function $F_{27,4}$ was observed in [Str14], where it is further conjectured that the coefficients of $F_{27,4}$ satisfy very strong congruences.

Theorem 2.1 (smooth point formula)**.** *Fix* $F = P/Q = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}}$*, a direction* $\hat{\mathbf{r}} \in (\mathbb{R}^+)^d$ *viewed also as an element of* $\mathbb{RP}^{d-1}$*, and a vector* $\beta \in \mathbb{R}^d$*. Let* $\mathbf{T}(\beta) = \{\mathbf{w} : |w_j| = \exp(\beta_j)\ \forall 1 \leq j \leq d\}$ *denote the torus of points with log modulus vector* β *and let* $c := -\hat{\mathbf{r}} \cdot \beta$ *denote the common height of points on the torus. Assume two hypotheses.*

1 **On the torus.** *The set* $E := \mathbf{T} \cap \mathsf{crit}(\hat{\mathbf{r}})$ *is finite and nonempty and contains only minimal smooth points.*

2 **Quadratic nondegeneracy**. *At each* $\mathbf{z} \in E$ *fix* k *such* $\partial Q/\partial z_k(\mathbf{z}) \neq 0$ *and let* $z_k = g(z_1, \ldots, \hat{z_k}, \ldots, z_d)$ *be a smooth local parametrization of* z_k *on* $\mathcal{V}$ *as a function of* $\{z_j : j \neq k\}$*. We suppose that the Hessian determinant* $\mathcal{H}_k$ *of second partial derivatives of* $g\left(w_1 e^{i\theta_1}, \ldots, w_d e^{i\theta_d}\right)$ *at the origin is non-zero for each* $\mathbf{w} \in E$*.*

Then there exists a closed neighborhood $\mathcal{N}$ *of* $\hat{\mathbf{r}}$ *on which all the above hypotheses hold and, uniformly on any such neighborhood,*

$$a_{\mathbf{r}} = (2\pi)^{(1-d)/2} \sum_{\mathbf{z} \in E} \det \mathcal{H}_{k(\mathbf{z})}^{-1/2} \frac{P(\mathbf{z})}{z_k (\partial Q/\partial z_k)(\mathbf{z})} r_k^{(1-d)/2} \mathbf{z}^{-\mathbf{r}} + O\left(r_k^{-d/2} \mathbf{z}^{-\mathbf{r}}\right) . \tag{2.2}$$

Remark. A number of other formulae for $a_{\mathbf{r}}$ are equivalent to this one and hold under the same hypotheses. An explicit formula for $\mathcal{H}_k$ in terms of partial derivatives of Q is given in [Mel17, Theorem 54]. The following coordinate-free formula for the constants involved in terms of the complexified Gaussian curvature $\mathcal{K}$ at a smooth point $\mathbf{z} \in \mathcal{V}$ is given in [PW13, (9.5.2)].

$$a_{\mathbf{r}} = (2\pi)^{(1-d)/2} \left[\sum_{\mathbf{z} \in E} \mathcal{K}(\mathbf{z})^{-1/2} \left| \nabla_{\log} Q(\mathbf{z}) \right|^{-1} P(\mathbf{z})\, |\mathbf{r}|^{(1-d)/2} \mathbf{z}^{-\mathbf{r}} \right] + O\left(|\mathbf{r}|^{-d/2} |\mathbf{z}|^{-\mathbf{r}} \right) \tag{2.3}$$

Proof: Assume first that $\log|\mathbf{w}|$ is the unique minimizer of $\mathbf{r} \cdot \mathbf{x}$ on the boundary of the log domain of convergence (this being a component of the complement of the amoeba). Under no assumptions on E or $\mathcal{K}$, Theorem 9.3.2 of [PW13] writes the multivariate Cauchy integral 2.1 as the integral of a residue form ω over an intersection cycle, $\mathcal{C}$. Taking into account that E is finite, and assuming an extra hypothesis that $\mathbf{r}$ is a *proper direction* (see [BP11, Definition 2.3]), Theorem 9.4.2 of [PW13] identifies $\mathcal{C}$ as a sum of quasi-local cycles near the points of E. For each such $\mathbf{z}$, if $\partial Q/\partial z_k$ and $\det \mathcal{H}_k$ do not vanish, Theorem 9.2.7 of [PW13] identifies the integral as the corresponding summand in (2.2). Nonvanishing of $\mathcal{H}_k$ is equivalent to nonvanishing of $\mathcal{K}$, leading to the coordinate-free formula (2.3), which may be found in [PW13, Theorem 9.3.7]. This proves the theorem under an extra hypothesis on the amoeba boundary.

To remove the hypothesis, consider the intersection cycle $\mathcal{C}$ obtained from expanding the small torus $\mathbf{T}$ to a torus $\mathbf{T}(\beta + \varepsilon)$. The construction in [PW13, Section A4] gives a compact $(d-1)$-chain representing a relative cycle in $H_{d-1}(\mathcal{V}^{c+\varepsilon}, \mathcal{V}^{c-\varepsilon})$, that is a chain of maximum height $c+\varepsilon$ with maximum boundary height $c - \varepsilon$. Applying the gradient flow on $\mathcal{V}$ for arbitrarily small time, we arrive again at a chain satisfying the conclusions of [PW13, Theorem 9.4.2]. Because the deformed chain has nonvanising boundary, one must add a term for the chain swept out by the deformation applied to this boundary, but this chain lies entirely in $\mathcal{V}^{c-\varepsilon}$, which is enough to complete the proof. □

Corollary 2.2. *Assume the hypotheses of Theorem 2.1, and fix a vector* $\mathbf{v}$ *in direction* $\hat{\mathbf{r}}$*.*

(i) If $E = \{\mathbf{z}\}$ *for some* $\mathbf{z}$ *in the positive real orthant in* $\mathbb{C}^d$ *and the leading constant of Equation* (2.2) *is positive, then there exists a neighbourhood of* $\hat{\mathbf{r}}$ *such that all but finitely many coefficients* $\{a_{\mathbf{r}} : \hat{\mathbf{r}} \in \mathcal{N}\}$ *are positive.*

(ii) If $E = \{\mathbf{z}\}$ for some $\mathbf{z}$ such that $\mathbf{z}^{\mathbf{v}} := \prod_{j=1}^{d} v_j$ is positive real and the leading constant of Equation (2.2) is positive, then all but finitely many coefficients $a_{n\mathbf{v}}$ are positive.

(iii) If E does not contain a point $\mathbf{z}$ with $\mathbf{z}^{\mathbf{v}}$ positive real and the sum in Equation (2.2) is not identically zero, then infinitely many coefficients $a_{n\mathbf{v}}$ are positive and infinitely many $a_{n\mathbf{v}}$ are negative.

Remark. When E contains a point in the positive real orthant but it is not a singleton, the corollary does not provide information as to eventual positivity.

Proof: Conclusions (i) and (ii) follow immediately from (2.2) because the sum is a single positive term.

For conclusion (iii), grouping the elements of E by conjugate pairs we note that up to scaling by $\mathbf{z}^{n\mathbf{v}} n^{d/2}$ the asymptotic leading term of $a_{n\mathbf{v}}$ has the form

$$l_n = \sum_{i=1}^{|E|} a_i \cos(2\pi\theta_i n + \beta_i),$$

where each θ_i, a_i, β_i is real, and $\theta_i \in (0,1)$. If r_n is any sequence satisfying a linear recurrence relation with constant coefficients, and $r_n = O(1/n)$, then Bell and Gerhold [BG07, Section 3] show that $l_n > r_n$ infinitely often. Since the modulus of the error term in Equation (2.2) can be bounded by a linear recurrence sequence with growth $O(1/n)$, we see that $a_{n\mathbf{v}}$ is positive infinitely often. Repeating the argument with $-l_n$ shows that $a_{n\mathbf{v}}$ is negative infinitely often. $\square$

Any computer algebra system can compute the set of smooth critical points in $\mathsf{crit}(\hat{\mathbf{r}})$ by solving the $d-1$ equations $\nabla_{\log}(\mathbf{z}) \parallel \hat{\mathbf{r}}$ together with the equation $Q(\mathbf{z}) = 0$. Identifying which points in crit are minimal is more difficult, although still effective [MS16]. For our cases, we can use results about symmetric functions to help with the computations. For any polynomial Q in d variables, let δ^Q denote the codiagonal: the univariate polynomial defined by $\delta^Q(x) = Q(x, \ldots, x)$.

Lemma 2.3 (polynomials in $\mathcal{M}_d$ have diagonal minimal points)**.** *Let $F = 1/Q$ with $Q \in \mathcal{M}_d$. Let x be a zero of δ^Q of minimal modulus. Then $\mathbf{x} := (x, \ldots, x)$ is a minimal point for F in $\mathsf{crit}(1, \ldots, 1)$.*

This follows directly from the classical Grace-Walsh-Szegő Theorem, of which we now sketch a modern proof.

Proof: Let $\alpha_1, \ldots, \alpha_k$ be the roots of δ^Q, where $k \le d$ is the common degree of Q and δ^Q and $|\alpha_1|$ is minimal among $\{|\alpha_j| : j \le k\}$. For any $\varepsilon > 0$, the polynomial

$$M(\mathbf{x}) := \prod_{j=1}^{k} (x_j - \alpha_j)$$

has no zeros in the polydisk $\mathcal{D}$ centered at the origin whose radii are $\alpha_1 - \varepsilon$. The symmetrization of M (see [BB09]) is defined to be the multilinear symmetric function m such that $m(x, \ldots, x) = M(x, \ldots, x)$. In our case $M = \delta^Q$, and it immediately follows that $m = Q$. By the Borcea-Brändén symmetrization lemma (see [BB09, Theorem 2.1]), the polynomial Q has no zeros in the polydisk $\mathcal{D}$. We conclude that the zero $\mathbf{x}$ of Q is a minimal point of F. $\square$

3 Symmetric multilinear functions of three variables

In this section we determine the diagonal asymptotics for general $Q = 1 - e_1 + ae_2 + be_3 \in \mathcal{M}_3$. Taking the coefficient of e_1 to be 1 loses no generality because of the rescaling $x_j \to \lambda x_j$ which preserves $\mathcal{M}_d$ and affects coefficient asymptotics in a trivial way. In order to use Theorem 2.1, we begin by identifying minimal points. Lemma 2.3 dictates that our search should be on the diagonal.

To that end, let $\delta^Q(x) = Q(x,x,x) = 1 - 3x + 3ax^2 + bx^3$. The discriminant of δ^Q is a positive real multiple of $p(a,b) := 4a^3 - 3a^2 + 6ab + b^2 - 4b = (a - 1 + 3(b-1))^2 - 4(b-1)^3$, and the zero set of δ^Q is obtained from that of the cubic $4b^3 = -a^2$ by centering at $(1,-1)$ and shearing via $(a,b) \mapsto (a+3b, b)$. The discriminant $p(a,b)$ vanishes along the red curve (solid and dashed) in Figure 1. Let $r_1(a)$ and $r_2(a)$ denote respectively the upper and lower branches of the solution to $p(a,b) = 0$.

Lemma 3.1. *Let p be a minimal modulus root of δ^Q. Then any critical point of $1/Q$ on the torus $T(p,p,p)$ has the form (q,q,q) where $\delta^Q(q) = 0$.*

Proof: A Gröbner basis computation shows that at any critical point at least two coordinates are equal, so we may assume without loss of generality that $x = y$. The only critical point possibly not on the diagonal is the point $\left(\frac{1}{a}, \frac{1}{a}, \frac{a(1-a)}{a^2+b}\right)$, when $a^2 + b \neq 0$. When the moduli of these coordinates are equal, $\frac{1}{a} = \pm\frac{a(1-a)}{a^2+b}$, meaning $a^3 + b = 0$ or $a^3 - 2a^2 - b = 0$. Another Gröbner computation verifies that there is no critical point off the diagonal when $a^3 + b = 0$. Similar manipulations show that when $a^3 - 2a^2 - b = 0$ then the only critical points off the diagonal occur at the permutations of $(1/a, 1/a, -1/a)$. As shown in Proposition 3.2, when $a \le 1$ and $a^3 - 2a^2 - b = 0$ then the minimal root of $\delta^Q(x)$ is positive and real, and it can be checked that the only time the positive root of $\delta^Q(x)$ has modulus $1/|a|$ is the trivial case $(a,b) = (1,-1)$. Furthermore, when $a^3 - 2a^2 - b = 0$ and $a > 1$ the modulus of the product of the roots of $\delta^Q(x)$ equals $\frac{1}{a^2(a-2)}$ and the minimal roots of $\delta^Q(x)$ are a pair of complex conjugates. If this pair has modulus $1/a$, then the real root of $\delta^Q(x)$ is $\pm\frac{1}{a^4(a-2)}$, but $\delta^Q\left(\pm\frac{1}{a^4(a-2)}\right) \neq 0$ for $a > 1$. □

Determining asymptotics is thus a matter of determining the minimal modulus roots of $\delta^Q(x)$.

Proposition 3.2. *The function δ^Q has a minimal positive real zero if and only if*

$$b \le \begin{cases} -9a & a \le -3 \\ r_1(a) & -3 \le a \le 1 \\ -a^3 & a \ge 1 \end{cases}$$

This corresponds to the set of points lying on and below the solid curve in Figure 1.

Proof: The real curve $p(a,b) = 0$ divides the plane into two regions. We first examine what happens in the region R to its right, which includes all points with $a > 1$ and some points with $a \le 1$.

Case 1: $a > 1$. The polynomial δ^Q has one real root and two complex roots. The real root is negative if and only if $b > 0$ and infinite when $b = 0$. We claim that the real root, r, has the least modulus of the three precisely when $b < -a^3$. First we check that the moduli of all roots agree when $b = -a^3$. The product of all three roots is $-1/b$ so, because $r^3 b < 0$, the moduli of the three are equal exactly when the real root, r, satisfies $r^3 b + 1 = 0$. Checking when δ^Q and $x^3 b + 1$ have a common root, we find this happens precisely

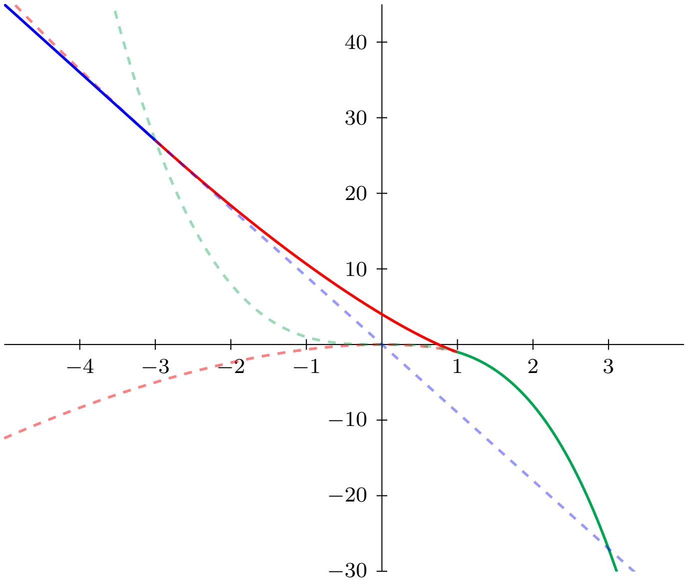

Figure 1: The three regimes defined by Proposition 3.2, made up of the curves $b = -9a$, $p(a,b) = 0$, and $b = -a^3$. Dashed lines represent the curves where they do not determine positivity of coefficients; note smoothness in the transitions between regimes.

when $b = -a^3$. By continuity of the roots with respect to the coefficients of δ^Q, because the real root has large modulus when b is near zero, the real root has larger modulus than the complex conjugate roots when $0 < b < \infty$ and when $0 > b > -a^3$. Conversely, when $b < -a^3$, the modulus of the real root is either always or never less than the modulus of the complex roots, and it is easy to check that the real root always has minimal modulus. This proves the claim, hence the proposition in the case $a > 1$.

Case $1'$*:* $a = 1$. Follows from continuity of the roots and by examining the trivial special case $(a,b) = (1,-1)$.

Case 2: $-3 < a < 1$. Again, when $b > r_1(a)$, two zeros are complex and one is real. The real root is negative except when $b < 0$ which can only happen for $3/4 \leq a \leq 1$. There $r_1(a) > -a^3$, so the real root is greater than the modulus of the complex roots and we conclude that for $b > r_1(a)$ there is no minimal positive real root.

Next, when $b = r_1(a)$, there is a new doubled real root, the limit of a complex conjugate pair of roots as $b \downarrow r_1(a)$. If $3/4 \leq a < 1$ then all roots are real and positive. If $a < 3/4$ then by a Gröbner basis computation, on $p(a,b) = 0$, the doubled root $y(a)$ is a solution to $ay^2 - 2y + 1 = 0$, while the single root satisfies $(4a-3)y^2 - 2y + 1 = 0$. On the upper branch $b = r_1(a)$, the doubled root $(1 - \sqrt{1-a})/a$ is positive and of smaller magnitude than the single root $-(1 + 2\sqrt{1-a})/(3-4a)$. This uses the fact that $a > -3$; at this value the single and doubled roots switch which has greater magnitude, though neither changes sign.

Decreasing b from $r_1(a)$, the minimal positive real root remains minimal and positive until either it crosses in magnitude with another root or it becomes negative. We check for a doubled root, a root and its negative, or a zero root. The cubic δ^Q never has a zero root. We already know a doubled root occurs precisely at $b = r_2(a)$. Two distinct real roots of equal magnitude can occur when a, b, x are such that $f_{ab}(x) = f_{ab}(-x) = 0$; a Gröbner basis for $f_{ab}(x), f_{ab}(-x)$ is $\{b + 9a, 1 + 3ax^2\}$, whence the equal magnitude roots can occur for

$b = -9a$. It suffices therefore to check what happens when b descends past $-9a$, then past $r_2(a)$. When $b = -9a$, the three roots are $1/3$ and $\pm 1/\sqrt{-3a}$. As long as $a > -9$, the root at $1/3$ is minimal on both sides of the crossing. When $b = r_2(a)$, the single real root $(2\sqrt{1-a}-1)/(3-4a)$ is positive and less than the doubled root $(\sqrt{1+a}+1)/a$, therefore after b decreases further, the single root is the only real root and is minimal. This completes the case when $-3 < a < 1$.

Case 2′: $a = -3$. Again, the result follows by continuity.

Case 3: $a < -3$. Observe that $b = -9a$ is tangent to $b = r_1(a)$, coinciding precisely at $a = -3$. The only thing different from Case 2 is that as b decreases below r_1, the two new positive real roots are no longer smaller than the magnitude of the negative real root. Instead, this does not happen until b crosses below $-9a$. After that, we see as before that the minimal positive real root cannot disappear. This finishes the case $a < -3$, the last case of the proof. $\square$

PROOF OF THEOREMS 1.7 AND 1.8: Suppose b is greater than the piecewise expression in the proposition; then δ^Q has no minimal positive zero, so the product of the three coordinates of the minimal points determined above do not lie in the positive orthant. By part (iii) of Corollary 2.2, the diagonal coefficients are not eventually positive. Asymptotics of δ_n are determined by Theorem 2.1, and when b is less than the piecewise expression it can be verified that the dominant term is positive.

4 The Gillis-Reznick-Zeilberger classes

Throughout this section, let $F_{c,d} = 1/Q_{c,d} = 1/(1 - e_1 + ce_d)$ and recall that $c_* = (d-1)^{d-1}$. Lemma 2.3 implies that for $Q \in \mathcal{M}_d$, in the diagonal direction, one may find diagonal minimal points. For $\mathcal{F}_{c,d}$, things are even simpler: all critical points for diagonal asymptotics are diagonal points.

Lemma 4.1. *Let $F = 1/Q$ with $Q \in F_{c,d}$. If $\mathbf{z} \in \mathsf{crit}(1,\ldots,1)$ then $z_i = z_j$ for all $1 \le i,j \le d$.*

PROOF: From $Q = 1 - e_1 + ce_d$ we see that $(\nabla_{\log} Q)_j = -z_j - ce_d$ and hence that $(\nabla_{\log} Q)_i = (\nabla_{\log} Q)_j$ if and only if $z_i = z_j$. $\square$

Proposition 4.2 (Smoothness of $F_{c,d}$ for $c \neq c_*$). *Let $F_{c,d} = 1/Q_{c,d} = 1/(1 - e_1 + ce_d)$. If $c \neq c_*$ then $\mathcal{V}$ is smooth. If $c = c_*$ then $\mathcal{V}$ fails to be smooth at the single point $\mathbf{z}_* = (1/d,\ldots,1/d)$. When $c = c_*$, the singularity at $\mathbf{z}_*$ has tangent cone e_2.*

PROOF: Checking smoothness of $F_{c,d}$ we observe that for d fixed and c and $x_1,\ldots,x_d$ variable, vanishing of the gradient of $F_{c,d}$ with respect to the x variables implies $x_j = ce_d$ for all j. This common value, x, cannot be zero, hence $x_j \equiv x$ and $c = x^{1-d}$. Vanishing of $F_{c,d}$ then implies $1 - dx + x$, hence $x = 1/(d-1)$ and $c = c_*$. This proves the first two statements. Setting $c = c_*$ and $x_j = 1/(d-1) + y_j$ centers $F_{c_*,d}$ at the singularity and produces a leading term of $(d-1)e_2(\mathbf{y})$, proving the third statement. $\square$

4.1 Proof of Theorems 1.10 and 1.11 in the case $c < c_*$

When $c \le 0$, the denominator of $F_{c,d}$ is one minus the sum of positive monomials, which leaves no doubt as to positivity. Assume, therefore, that $0 < c < c_*$. Apply Lemma 2.3 to see that if x is a minimum modulus zero of $\delta^Q := Q_{c,d}(x, \dots, x)$ then $(x, \dots, x)$ is a minimal point for $F_{c,d}$ in the diagonal direction. Apply Lemma 4.1 to conclude that the set E in Theorem 2.1 of minimal critical points on $\mathbf{T}(|x|, \dots, |x|)$ consists only of points $(y, \dots, y)$ such that y is a root of δ^Q. By part (i) of Corollary 2.2, it suffices to check that $\delta^Q = 1 - dx + cx^d$ has a unique minimal modulus root ρ and that $\rho \in \mathbb{R}^+$. Thus, the conclusion follows from the following proposition.

Proposition 4.3. *For $c \in (0, c_*)$, the polynomial $\delta^Q = 1 - dx + cx^d$ has a root $\rho \in \left[\frac{1}{d}, \frac{1}{d-1}\right]$ which is the unique root of δ^Q of modulus less than $1/(d-1)$.*

Proof: Checking signs we find that $\delta^Q(1/d) = cd^{-d} > 0$ while $\delta^Q(1/(d-1)) = -(d-1)^{-1} + c(d-1)^{-d} < -(d-1)^{-1} + c_*(d-1)^{-d} = 0$, therefore there is at least one root, call it ρ, of δ^Q in the interval $[1/d, 1/(d-1)]$. On the other hand, when $|z| = 1/(d-1)$, we see that $|dz| \geq |1 + cz^d|$ and therefore, by applying Rouché's theorem to the functions $-dz$ and $1 + cz^d$, we see that δ^Q has as many zeros on $|z| < 1/(d-1)$ as does $-dz$: precisely one root, ρ. $\square$

4.2 Proof of Theorems 1.10 and 1.11 in the case $c > c_*$

Again, by Lemmas 2.3 and 4.1, we may apply part (iii) of Corollary 2.2 to the set E of points $(y, \dots, y)$ for all minimal modulus roots y of δ^Q. The result then reduces to the following proposition.

Proposition 4.4. *For $c > c_*$, the set of minimal modulus roots of the polynomial $\delta^Q = 1 - dx + cx^d$ contains no point whose d^{th} power is real and positive.*

Proof: First, if z^d is real then the imaginary part of $\delta^Q(z)$ is equal to the imaginary part of $-dz$, hence any root z of δ^Q with z^d real is itself real.

Next we check that δ^Q has no positive real roots. Differentiating $\delta^Q(x)$ with respect to x gives the increasing function $d(-1 + cz^{d-1})$ with a unique zero at $c^{-1/(d-1)}$. This gives the location of the minimum of δ^Q on $\mathbb{R}^+$, where the function value is $1 - dc^{-1/(d-1)} + c^{1-d/(d-1)} = 1 - (d-1)/c^{1/(d-1)}$ which is positive because $c > (d-1)^{d-1}$.

If d is even, δ^Q clearly has no negative real roots, hence no real roots at all, finishing the proof in this case. If d is odd δ^Q will have a negative real root u, however because d is odd, the product of the coordinates of $(u, \dots, u)$ is $u^d < 0$. $\square$

We conjecture that the roots of minimal modulus when $c > c_*$ are always a complex conjugate pair, however this determination does not affect our positivity results.

4.3 Proof of Theorem 1.12

When $c < c_*$ we have seen there is a single real minimal point $(\rho_c, \dots, \rho_c)$ in the diagonal direction and that $\rho_c \uparrow 1/(d-1)$ as $c \uparrow c_*^-$. The limit from below in Theorem 1.12 then follows directly from Theorem 1.11.

For the limit from from above, it suffices to show that in the diagonal direction, for c sufficiently close to c_* and greater, E consists of a single diagonal complex conjugate pair $(\zeta_c, \dots, \zeta_c)$ and $(\overline{\zeta_c}, \dots, \overline{\zeta_c})$, and that

$\overline{\zeta_c} \to 1/(d-1)$ as $c_* \downarrow c$. First, we check that at $c = c_*$ the unique minimum modulus root of δ^Q is the doubled root at $1/(d-1)$. For $c = c_*$, the first and third terms of $\delta^Q = 1 - dz + c_* z^d$ have modulus 1 and $1/(d-1)$ when $|z| = 1/(d-1)$, respectively, summing to the modulus of the middle term; therefore if $\delta^Q(z) = 0$ and $|z| = 1/(d-1)$ then the third term is positive real. But then the second term must be positive real too, hence the unique solution of modulus at most $1/(d-1)$ is $z = 1/(d-1)$. A quick computation shows the multiplicity to be precisely 2. We know that for $c > c_*$ there are no real roots. Therefore, as c increases from c_*, the minimum modulus doubled root splits into two conjugate roots, which, in a neighborhood of c_*, are still the only minimum modulus roots. $\square$

5 Lacuna computations

Theorem 1.13 follows immediately from Theorem 5.1 below, with the following specifications: $d \geq 4$ is any even number, $c = c_*$, $P = 1$, $Q = Q_{c,d}$, $\mathbf{z}_* = (1/d, \dots, 1/d)$, $\hat{\mathbf{r}} = (1, \dots, 1)$, B is the component of the complement of the amoeba of Q containing $(a, \dots, a)$ for $a < -\log d$, $\mathbf{x}_* = (-\log d, \dots, -\log d)$, $\mathbf{y}_* = \mathbf{0}$ and $\mathcal{N}$ taken to be the diagonal. Proposition 4.2 guarantees the correct shape for the tangent cone to Q at $\mathbf{z}_*$.

Theorem 5.1. *Suppose $F = P/Q^k$ with P a holomorphic function and Q a real Laurent polynomial. Fix $\hat{\mathbf{r}} \in \mathbb{RP}^d$, let B be a component of the complement of the amoeba of Q, let $\sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}}$ be the Laurent expansion for F convergent for $\mathbf{z} = \exp(\mathbf{x} + i\mathbf{y})$ and $\mathbf{x} \in B$. Let $\mathbf{x}_* \in \partial B$ be the a maximizing point for $\mathbf{r} \cdot \mathbf{x}$ on ∂B. Assume that $\mathcal{V}$ has a unique singularity $\mathbf{z}_* = \exp(\mathbf{x}_* + i\mathbf{y}_*)$, and that the tangent cone of Q at $\mathbf{z}$ transforms by a real linear map to $z_d^2 - \sum_{j=1}^{d-1} z_j^2$. Let $\mathcal{N}$ be any closed cone such that $\mathbf{x}_*$ maximizes $\mathbf{r} \cdot \mathbf{x}$ for all $\mathbf{r} \in \mathcal{N}$.*

If $d > 2k$ is even then there is an $\varepsilon > 0$ and a chain Γ contained in the set $\mathcal{V}_\varepsilon := \{\mathbf{z} \in \mathcal{V} : |\mathbf{z}^{-\mathbf{r}}| \leq \exp(-\mathbf{r} \cdot \mathbf{x}_* - \varepsilon|\mathbf{r}|)$ such that

$$a_{\mathbf{r}} = \int_\Gamma \mathbf{z}^{-\mathbf{r}} \frac{P}{Q^k} \frac{d\mathbf{z}}{\mathbf{z}} \,. \tag{5.1}$$

In other words, the chain of integration can be slipped below the height of the singular point.

The outline of the proof is as follows. Let $\Phi = \Phi_{\mathbf{r}} := -\hat{\mathbf{r}} \cdot \log \mathbf{z}$ denote the phase function from the integral (2.1), whose real part h must be minimized in order to apply saddle point techniques.

Step 1: For $c \in \mathbb{C}$, let the subscript c denote replacement of Q by $Q - c$. Thus, $Q_c = Q - c$, $\mathcal{V}_c := \{\mathbf{z} : Q_c(\mathbf{z}) = 0\}$, and so forth. For small positive real c, the variety $\mathcal{V}_c$ is smooth and has two critical points $\mathbf{z}_c^+$ and $\mathbf{z}_c^-$ near c, with $h(\mathbf{z}_c^+) > h(\mathbf{z}_*) > h(\mathbf{z}_c^-)$ and $\mathbf{z}_c^\pm \to \mathbf{z}_*$ as $c \downarrow 0$. Let B_c denote the component of the amoeba boundary for the power series expansion of P/Q_c.

Near $\mathbf{z}_c^+$ the boundary of B_c is mapped by the log Gauss map diffeomorphically to a set D_c. Let $\mathcal{S}^\pm$ denote respectively the stable manifold of points that flow into $\mathbf{z}_c^\pm$ under the downhill gradient flow under h on $\mathcal{V}_c$.

Lemma 5.2. *The following three sets are the same in a ball around $\mathbf{z}_c^+$:*

(i) The stable manifold of all points flowing into $\mathbf{z}_c^+$ under the downward gradient flow for h;

(ii) the real variety $\mathcal{V}_c^{\mathbb{R}} := \mathcal{V}_c \cap \mathbb{R}^d$;

(iii) the pre-image under the log Gauss map $\mathbf{z} \mapsto \nabla_{\log} Q(\mathbf{z})$ of ∂B_c.

Step 2: let $\mathbf{T}_t$ denote the centered torus of polyradius $(1+t)|\mathbf{z}_*|$. Fix $\varepsilon > 0$ such that for all $t \leq \varepsilon$ and $c \leq \varepsilon$, the torus $\mathbf{T}_t$ is disjoint from an ε neighborhood of $\mathcal{V}_c$. Let Γ_c denote the intersection of the homotopy

$$H_\varepsilon := \bigcup_{t \in [-\varepsilon, \varepsilon]} \mathbf{T}_t$$

with $\mathcal{V}_c$. We perform the usual Thom intersection on the homotopy H and the pair $(\mathbb{C}^d, \mathcal{V}_c)$ to see that $\mathbf{T}_{-\varepsilon}$ is homologous in $H_d(\mathcal{M})$ to a compact chain Γ_c supported in the union of $\mathbf{T}_\varepsilon$ and a neighborhood of $\mathbf{z}_c^+$; the part in the neighborhood of $\mathbf{z}_c^+$ is in fact equal to the product of a $(d-1)$ sphere with the intersection chain $\mathcal{I}$ which is the intersection of $\mathcal{V}_c$ and the image of H. We observe (see Step 1) that Γ_c intersects $\mathcal{S}^+$ at the unique point $\mathbf{z}_c$.

Step 3: For $c \leq \varepsilon$, mapping by log in the ε-neighborhood of $\mathbf{z}_*$ causes the tori $\mathbf{T}_t$ become imaginary fibers: sets of the form $v + iB$ for $v \in \mathbb{R}^d, B \subseteq \mathbb{R}^d$. There is an affine change of coordinates $\mathbf{z} \mapsto \mathbf{w}$, independent of c, such that in the new coordinates, Q is given locally by

$$Q(\mathbf{w}) = q(\mathbf{w}) + r(\mathbf{w})$$

where

$$q(\mathbf{w}) = w_1^2 - 1 - \sum_{k \geq 2} w_k^2 \tag{5.2}$$

with r holomorphic and vanishing to order 3 at the origin and $h(\mathbf{w}) = \Re\{w_0\}$. Rescaling by $c^{1/2}$ gives coordinates $\mathbf{Z} = c^{1/2}\mathbf{w}$ in which Q is given locally by

$$Q(\mathbf{Z}) = q(\mathbf{Z}) + c^{1/2} r(\mathbf{Z}, c)$$

with r continuous in c. In these new coordinates, the homotopy $\mathbf{T}$ becomes

$$\mathbf{T}_{\varepsilon t} = t\mathbf{e}_1 + iB_t \tag{5.3}$$

for $-1 \leq t \leq 1$, some neighborhood B_t of the origin in $\mathbb{R}^d$, and $-\mathbf{e}_1$ denoting the first basis vector, oriented to point inside each B_c at $\mathbf{z}_c$. We will henceforth refer to zero set of q as **the quadric**. The two critical points for h on the quadric are $\pm\mathbf{e}_1$,

Step 4: For small c's, all the subvarieties

$$\mathcal{V}_c^{\mathbf{Z}} = \{q(\mathbf{Z}) + c^{1/2} r'(\mathbf{Z}, c)\}$$

are diffeomorphic between themselves and to the limiting variety. Intersections of slightly perturbed manifolds define homologous chains, so that we can work directly with the quadric. The downward gradient flow on the quadric admits a simple description. The real part of the quadric is a two-sheeted hyperboloid, the upper sheet H^+ passing through $\mathbf{e}_1$ and the lower sheet H^- passing through $-\mathbf{e}_1$.

Step 5: Let $z_k = x_k + iy_k$ and let S denote the $(d-1)$-dimensional sphere $x_1^2 + \sum_{k=2}^d y_k^2 = 1$. The north pole and south pole of S are $\pm\mathbf{e}_1$ respectively.

Lemma 5.3. *At the upper critical point $\mathbf{e}_1$, the stable manifold (the points whose gradient flow approach $\mathbf{e}_1$) is H^+, while the unstable manifold (the points whose upward gradient flow approach $\mathbf{e}_1$) is the sphere, S. For the lower critical point $-\mathbf{e}_1$ it is the opposite: the stable manifold is S and the unstable manifold is H^-.*

Let us see what happens when we apply the gradient flow to any chain, stopped whenever height $-1-\varepsilon$ is reached. Most trajectories will flow to height $-1-\varepsilon$ in finite time, with two exceptions. Points in the upper branch will approach $\mathbf{e}_1$ and points in S will flow, at any finite time, to a lower point of S, approaching but never hitting $-\mathbf{e}_1$. Any chain disjoint from H^+ and from S will hit height $-1-\varepsilon$.

Step 6: In the $\mathbf{Z}$ coordinates, the homotopy $iB_t - \mathbf{e}_1$ avoids the rescaled variety, given locally by

$$\mathcal{V}_1 := \{\mathbf{Z} : Z_0^2 = 1 + \sum_{k=2}^d Z_k^2\}\,. \tag{5.4}$$

The homotopy from $iB_{-\varepsilon} - a\mathbf{e}_1$ to $iB_\varepsilon + a\mathbf{e}_1$ cuts out the intersection chain Γ_1. We compute this by setting $Z_k = x_k + iy_k$. The range of the homotopy is the set where the real part satisfies $-a \le x_1 \le a$ and $x_k = 0$ for $k \ge 2$. Accordlingly, $H \cap \mathcal{V}_1$ is described by the following pair of equations, together with the inequality on x_1.

$$x_1^2 - y_1^2 \;=\; 1 - \sum_{k=2}^d y_k^2 \tag{5.5}$$

$$x_1 y_1 \;=\; 0 \tag{5.6}$$

Clearly this is the union of two sets, one obtained by solving (5.5) when $y_1 = 0$ and the other obtained by solving (5.5) when $x_1 = 0$. Denoting these by Γ_s and Γ_h respectively, we see that Γ_s is the sphere S, while Γ_h is the single-sheeted hyperboloid $\{-y_1^2 = 1 - \sum_{k=2}^d y_k^2\}$ a hypersurface in the purely imaginary space. Each of these is smooth, however they intersect at a $(d-2)$-sphere, the equator of the imaginary unit sphere, everywhere on which Γ_1 fails to be smooth.

Step 7: Remarkably, the orientations are discontinuous!

Lemma 5.4. *When d is even, the orientations of Γ_s and Γ_h change signs when crossing their mutual intersection.*

Step 8: Define the chain Γ_s' to be the sphere S with continuous orientation agreeing with Γ_s in the northern hemisphere (hence disagreeing in the southern hemisphere).

Lemma 5.5. *The chain $\Gamma_s + \Gamma_h$ is homotopic to a chain $\Gamma_s' + \Gamma_h'$, where Γ_h' is the result of applying the gradient flow to Γ_h.*

PROOF: Under gradient flow, Γ_s is an invariant set, while the equator of Γ_s flows down to the south pole. All other points flow below height -1. At large finite time, this shows that $\Gamma_s + \Gamma_h$ is homotopic to S with orientation reversed in a small neighborhood of the south pole, plus a chain mostly below height -1 that intersects S at the orientation reversing latitude. Taking the limit, one obtains $\Gamma_s' + \mathcal{D}$ for some chain $\mathcal{D}$ with maximum height -1 obtained uniquely at the south pole. □

Step 9: We denote by $\mathcal{C} \times S^1$ a chain in the quadric with this product structure. The original chain of integration $\mathbf{T}$ in (2.1) is homologous in $H_d(\mathcal{M})$ to a compact chain $\mathcal{C} = \Gamma_s' \times S^1 + \Gamma_h' \times S^1$.

$$a_{\mathbf{r},c} = \int_{\mathbf{T}} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^n} \frac{d\mathbf{z}}{\mathbf{z}} = \int_{\Gamma_s' \times S^1} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^k} \frac{d\mathbf{z}}{\mathbf{z}} + \int_{\Gamma_h' \times S^1} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^k} \frac{d\mathbf{z}}{\mathbf{z}}$$

whence,

$$a_{\mathbf{r}} = \lim_{c \downarrow 0} a_{\mathbf{r},c} = \lim_{c \downarrow 0} \int_{\Gamma_s' \times S^1} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^k} \frac{d\mathbf{z}}{\mathbf{z}} + \lim_{c \downarrow 0} \int_{\Gamma_h' \times S^1} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^k} \frac{d\mathbf{z}}{\mathbf{z}}\,. \tag{5.7}$$

Step 10: We evaluate the integral over Γ'_s via residues. Denoting the integrand as the d-form $\omega := \mathbf{z}^{-\mathbf{r}}(P/Q_c^k)d\mathbf{z}/\mathbf{z}$, the residue in the $\mathbf{Z}$-coordinates at $\mathcal{V}_c$, straightened so that Q_c reduces to the leading quadratic term, is the $(d-1)$-form given by

$$\operatorname{Res}\omega = \frac{R(\mathbf{Z})}{Z_1 Q^{k-1}} dZ_1 \wedge \cdots \wedge dZ_d \,.$$

Here, R is some locally holomorphic function. This form is homogeneous of degree $d-2k$. Changing back to unrescaled coordinates $\mathbf{z} = c^{1/2}\mathbf{Z}$ scales the integral by $c^{(d-2)/2}$ sending it to zero as $c \downarrow 0$ as long as $d > 2k$. In other words,

$$\lim_{c\downarrow 0} \int_{\Gamma_s \times S^1} \mathbf{z}^{-\mathbf{r}} \frac{P}{Q_c^k} \frac{d\mathbf{z}}{\mathbf{z}} = 0 \,.$$

Step 11: It remains to see why Γ'_h is supported strictly below the lower critical height. Working back in the $\mathbf{Z}$-coordinates, we claim that, locally, Γ'_h is null homologous. In fact, the computation at (10) and following of the lacuna writeup shows that locally, Γ_h may be deformed into two copies of the lower sheet of the real hyperboloid, oppositely oriented. This is only local because the coordinate change making $\mathcal{V}_c$ into an exact quadric happens only in a (unrescaled) neighborhood of $\mathbf{z}_c^-$. We see that such a neighborhood is excluded from the support of Γ'_h. Thus Γ'_h is supported on a a set at height $-1-\varepsilon$, proving the theorem with $\Gamma = \Gamma'_h$.

□

References

[ABG70] M. Atiyah, R. Bott, and L. Gårding. Lacunas for hyperbolic differential operators with constant coefficients, I. *Acta Mathematica*, 124:109–189, 1970.

[AG72] R. Askey and Gasper G. Certain rational functions whose power series have positive coefficients. *Amer. Math. Monthly*, 79:327–341, 1972.

[AG77] R. Askey and Gasper G. Convolution structures for Laguerre polynomials. *J. D'Analyse Math.*, 31:48–68, 1977.

[AvSZ11] G. Almkvist, D. van Straten, and W. Zudilin. Generalizations of Clausen's formula and algebraic transformations of Calabi–Yau differential equations. *Proc. Edin. Math. Soc.*, 54:273–295, 2011.

[BB09] J. Borcea and P. Brändén. The Lee-Yang and Pólya-Schur programs, II: Theory of stable polynomials and applications. *Comm. Pure Appl. Math.*, 62:1595–1631, 2009.

[BG07] J. Bell and S. Gerhold. On the positivity set of a linear recurrence sequence. *Israel J. Math.*, 157:333–345, 2007.

[BP11] Y. Baryshnikov and R. Pemantle. Asymptotics of multivariate sequences, part III: quadratic points. *Adv. Math.*, 228:3127–3206, 2011.

[GRZ83] J. Gillis, B. Reznick, and D. Zeilberger. On elementary methods in positivity theory. *SIAM J. Math. Anal.*, 14:396–398, 1983.

[HK71] M. Hautus and D. Klarner. The diagonal of a double power series. *Duke Math. J.*, 23:613–628, 1971.

[Kau07] M. Kauers. Computer algebra and power series with positive coefficients. page ??? 2007.

[Koo78] T. Koornwinder. Positivity proofs for linearization and connection coefficients of orthogonal polynomials satisfying an addition formula. *J. London Math. Soc.*, 18:101–114, 1978.

[Kou10] C. Koutschan. HolonomicFunctions (User's Guide). Technical report, no. 10-01 in RISC Report Series, University of Linz, Austria, January 2010.

[KZ08] M. Kauers and D. Zeilberger. Experiments with a positivity-preserving operator. *Exper. Math.*, 17:341–345, 2008.

[Lip88] L. Lipshitz. The diagonal of a D-finite power series is D-finite. *J. Algebra*, 113(2):373–378, 1988.

[Mel17] S. Melczer. *Analytic combinatorics in several variables: effective asymptotics and lattice path enumeration.* PhD thesis, University of Waterloo, 2017.

[MS16] S. Melczer and B. Salvy. Symbolic-numeric tools for analytic combinatorics in several variables. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 333–340, New York, NY, USA, 2016. ACM.

[Pól21] G. Pólya. Sur les séries entières, dont la somme est une fonction algébrique. *L'Enseignement Mathématique*, 22:38–47, 1921.

[PW13] R. Pemantle and M. Wilson. *Analytic Combinatorics in Several Variables*, volume 340 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, New York, 2013.

[SS13] A. Scott and A. Sokal. Complete monotonicity for inverse powers of some combinatorially defined poynomials. *Acta Math.*, 213:323–392, 2013.

[Str08] A. Straub. Positivity of Szegö's rational function. *Adv. Appl. Math.*, 41(2):255–264, 2008.

[Str14] A. Straub. Multivariate Apéry numbers and supercongruences of rational functions. *Algebra and Number Theory*, 8:1985–2008, 2014.

[SZ15] A. Straub and W. Zudilin. Positivity of rational functions and their diagonals. *J. Approx. Theory*, 195:57–69, 2015.

[Sze33] G. Szegö. Über gewisse Potenzreihen mit lauter positiven Koeffizienten. *Math. Zeit.*, 37:674–688, 1933.