Optimal Bounds for the Number of Pieces of Real Near-Circuit Hypersurfaces

Weixun Deng
Department of Mathematics
Texas A&M University
College Station, TX, USA
deng15521037237@tamu.edu

J. Maurice Rojas
Department of Mathematics
Texas A&M University
College Station, Texas, USA
rojas@tamu.edu

Cordelia Russell
Department of Mathematics
University of New Mexico
Albuquerque, New Mexico, USA
cordelia03@gmail.com

Abstract

Suppose f is a polynomial in n variables with real coefficients and exactly $n+k$ monomial terms. Estimating the number of connected components of the zero set of f in the positive orthant is a fundamental problem in real algebraic geometry, and tight estimates have applications in computational complexity and topology. We prove that, for generic coefficients and exponent vectors, the number of connected components is at most 3 when $k=3$, settling an open question from Fewnomial Theory. Our results also extend to exponential sums with real exponents. A key contribution is a deeper analysis of the underlying $\mathcal{A}$-discriminant contours, which should be useful for other quantitative geometric problems.

CCS Concepts

• **Mathematics of computing** → **Computations on polynomials**; • **Theory of computation** → **Computational geometry**.

Keywords

Exponential Sum, Morse Theory, Connected Component, Discriminant, Fewnomial

ACM Reference Format:

Weixun Deng, J. Maurice Rojas, and Cordelia Russell. 2025. Optimal Bounds for the Number of Pieces of Real Near-Circuit Hypersurfaces. In *International Symposium on Symbolic and Algebraic Computation (ISSAC '25), July 28–August 01, 2025, Guanajuato, Mexico.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3747199.3747576

1 Introduction

Estimating the number of connected components (a.k.a. *pieces*) of the real zero set of a polynomial is a fundamental problem with numerous applications. For univariate polynomials with exactly t monomial terms, Descartes' (17th century) Rule of Signs reveals that the maximal number of isolated positive roots is $t-1$. In the 1980s, Khovanskii [14] extended this to higher dimensions by developing *Fewnomial Theory*. One of his results gave, for any n-variate polynomial f with real coefficients and exactly $n+k$ monomial terms, an upper bound of $2^{O((n+k-1)^2)}$ for the number of pieces of the zero set, $Z_+(f)$, of f in the positive orthant $\mathbb{R}^n_+$.[1]

Although Khovanskii's fewnomial bounds were eventually found to be far from optimal, significant improvements took decades to find: After improvements by Li, Rojas, Wang [15], and Perrucci [18] in the early 2000s, Bihan and Sottile sharpened the last bound to $2^{O(k^2+n+k\log n)}$ [3, 5].

Let $\mathcal{A}$ denote the *support* (i.e., set of exponent vectors) of f. If $\mathcal{A}$ does *not* lie in an affine hyperplane then we call f an *honest* n-variate $(n+k)$-nomial. We make the natural assumption that our polynomials (and exponential sums) are honestly n-variate in this sense, for otherwise a simple change of variables can be used to reduce the number of variables and increase k while preserving $n+k$. Refining a 2016 $\mathcal{A}$*-discriminant* approach from [8], Bihan, Humbert, and Tavenas [2] proved an even sharper upper bound of the form $2^{O(k^2+k\log n)}$ around 2022: Their best upper bound for the honest n-variate $(n+3)$-nomial case was $\left\lfloor \frac{n-1}{2} \right\rfloor + 3$. Our main result improves this bound to an optimal *constant*: 3.

Remark 1.1. *The maximal numbers of pieces in the honest n-variate $(n+k)$-nomial case, for $k\in\{1,2\}$, are respectively* 1 *and* 2. *(In these cases, the real zero set is always smooth when $k=1$ and, for $k=2$,* neither *singularities nor non-generic supports increases the maximal number of pieces.) It took about* 18 *years from the publication of Khovanskii's book on fewnomials [14] until the 2011 habilitation thesis of Bihan for the $k=2$ case to be settled.*

The cases $k\in\{1,2,3\}$ are respectively known as the simplex, circuit, *and* near-circuit *cases. The maximal number of pieces for honest n-variate $(n+4)$-nomials remains unknown.* ⋄

If no $n+1$ distinct points of $\mathcal{A}$ lie in the same affine $(n-1)$-plane, then we say that $\mathcal{A}$ is *generic*.

Theorem 1.2. *Let f be a real n-variate $(n+3)$-nomial with generic support and smooth $Z_+(f)$. Then $Z_+(f)$ has at most* 3 *pieces.*

We prove Theorem 1.2 in Section 4. The case of non-generic supports is more subtle in the near-circuit case than for the simplex and circuit cases. (See Example 1.4 below and the discussion just before Example 1.4.) So we leave this case, as well as the case of singular $Z_+(f)$, for the journal version of this paper, which is currently in progress.

For any $x=(x_1,\ldots,x_n)\in\mathbb{C}^n$ let $\mathrm{Log}\,|x|:=(\log|x_1|,\ldots,\log|x_n|)$, where we use base-$e$ logarithm throughout. The following example shows that our bound is in fact optimal.

[1]Since real (not just integral) powers of positive numbers are well-defined, we allow, like Khovanskii did, *real* exponents for our $(n+k)$-nomials throughout. Also, in our use of O and Ω notation, all our constants are effective and absolute, i.e., they can be made explict, and there is *no* dependence on any further parameters.

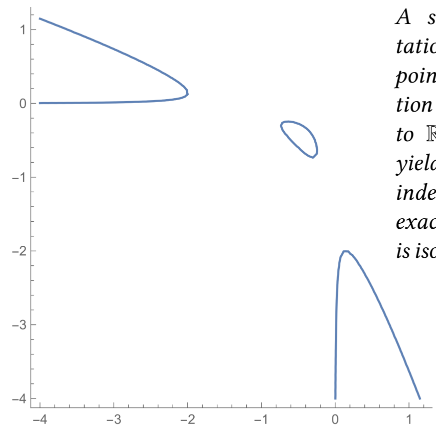
ISSAC '25, Guanajuato, Mexico

ACM ISBN 979-8-4007-2075-8/25/07
https://doi.org/10.1145/3747199.3747576

Example 1.3. *[17, Ex. 1.8] Consider*

$$f(x_1,x_2):=1-x_1-x_2+\tfrac{6}{5}x_1x_2^4+\tfrac{6}{5}x_1^4x_2$$

which has $\mathrm{Log}\,|Z_+(f)|$ *plotted below:*



A standard computation of the critical points of the projection mapping $Z_+(f)$ to $\mathbb{R}_+\times\{0\}$ easily yields a proof that we indeed have 3 pieces, exactly one of which is isotopic to a circle. ⋄

Example 1.3 is bivariate and has generic support. One might think that a simpler example like

$$f(x_1,x_2):=(x_1x_2-1)(x_1x_2-2)(x_1x_2-3)(x_1x_2-4)$$

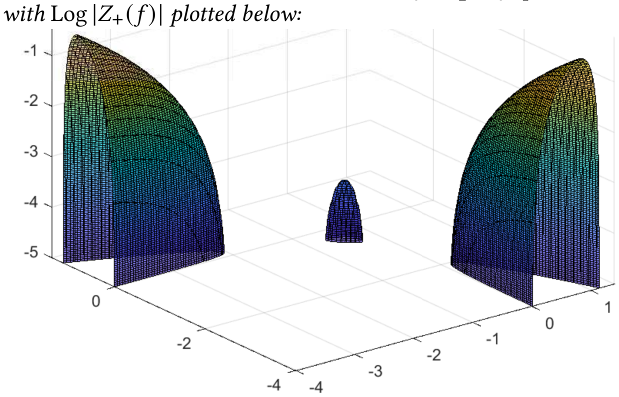
(with 5 terms in its monomial term expansion and $Z_+(f)$ having 4 pieces) contradicts Theorem 1.2. However, this simpler f has support lying in a line, making it a *dis*honest bivariate 5-nomial with (extremely) non-generic support.

One can in fact prove that generic support is a necessary condition for a bivariate pentanomial f to have $Z_+(f)$ with 3 pieces (see the upcoming journal version of our paper). However, the situation changes when f has 3 or more variables.

Example 1.4. *Consider the trivariate 6-nomial*

$$f(x_1,x_2,x_3):=1-x_1-x_2+x_3+\tfrac{6}{5}x_1x_2^4+\tfrac{6}{5}x_1^4x_2$$

with $\mathrm{Log}\,|Z_+(f)|$ *plotted below:*



This f has non-generic support, since its support has 6 points, but 5 of them lie in a 2-plane. However, this f is at least honestly trivariate. Note also that while $Z_+(f)$ here is smooth and has 3 pieces, all its pieces are non-compact since the x_1x_2-plane is not part of $Z_+(f)$. ⋄

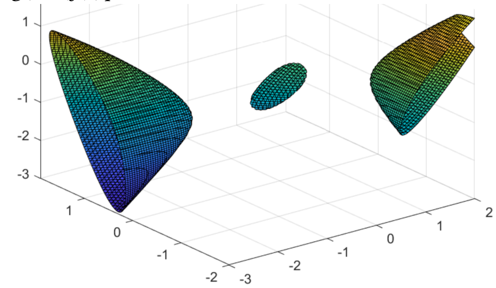
Remark 1.5. *A basic consequence of our framework is that $Z_+(f)$ is smooth for any choice of coefficients outside of an analytic hypersurface in $\mathbb{R}^{n+k}$. So, on a practical level, non-generic supports and singular $Z_+(f)$ are rare.* ⋄

We can also find trivariate f with *generic* near-circuit support and $Z_+(f)$ having 3 pieces:

Example 1.6. *Consider the trivariate 6-nomial*

$$f(x_1,x_2,x_3):=x_3^2-x_1x_3^2-x_2x_3^2+x_3^3+e^{-2.2}x_1x_2^4x_3+e^{-2.7}x_1^4x_2$$

with $\mathrm{Log}\,|Z_+(f)|$ *plotted below:*



(Note that $e^{-2.2}=0.110803\ldots$ and $e^{-2.7}=0.067205\ldots$) The support $\mathcal{A}$ here is generic: Every point of $\mathcal{A}$ is a vertex on the convex hull, P, of $\mathcal{A}$; and P has positive volume and exactly 8 facets. ⋄

We conjecture that n-variate f with *generic* near-circuit support and $Z_+(f)$ having 3 pieces exist for each $n\geq 4$.

1.1 High-Level Summary of Proof of Main Result

To achieve our new optimal bound of 3 pieces we closely analyze the structure of the underlying $\mathcal{A}$-*discriminant* (see Subsections 2.3 and 2.4). The key is that once a near-circuit $\mathcal{A}$ and a choice ε of coefficient signs is fixed, the resulting $(n+3)$-dimensional family of polynomials can be parametrized via the pieces of the complement of a (possibly singular) analytic curve $\Gamma_\varepsilon\subset\mathbb{R}^2$ called a *(signed) reduced contour*. The number of cusps on Γ_ε determines, in part, the number of possible isotopy types for $Z_+(f)$, building on earlier work of Rojas and Rusek [19] (see Lemma 2.4 and Theorem 3.2 below). When the number of cusps is two or more we need to work harder to determine how the isotopy type of $Z_+(f)$ changes as we vary the coefficients of f and sometimes cross over Γ_ε.

The influence of cusps in reduced contours on the possible isotopy types of $Z_+(f)$ is a new tool that we hope will be used more broadly in real algebraic geometry.

2 Notation and Background

For any two vectors $v=(v_1,\ldots,v_n), w=(w_1,\ldots,w_n)\in\mathbb{R}^n$ we set $v\cdot w:=v_1w_1+\cdots+v_nw_n$, $\mathbb{R}^*:=\mathbb{R}\setminus\{0\}$, and let $M^\top$ denote the transpose of a matrix M. For any twice differentiable function $h:\mathbb{C}^n\longrightarrow\mathbb{R}$ we let let $Z_\mathbb{R}(h)$ denote the zero set of h in $\mathbb{R}^n$ and $\mathrm{Hess}(f)$ the $(n\times n)$ *Hessian matrix* $\left[\frac{\partial^2 h}{\partial x_i\partial x_j}\right]$.

2.1 Deforming Zero Sets of Exponential Sums

Suppose $\mathcal{A}=\{\alpha_1,\ldots,\alpha_{n+k}\}\subset\mathbb{R}^n$ has cardinality $n+k$. Technically, it will be easier to work with exponential sums instead of polynomials. So let $f_c(x):=\sum_{i=1}^{n+k}c_ie^{\alpha_i\cdot x}$, where $c=(c_1,\ldots,c_{n+k})\in(\mathbb{R}^*)^{n+k}$. Since log defines a homeomorphism from $\mathbb{R}$ to $\mathbb{R}_+$ our main problem is clearly equivalent to bounding the number of pieces of $Z_\mathbb{R}(f_c)$.

We call $\varepsilon=\mathrm{sign}(c)\in\{\pm\}^{n+k}$ a *sign distribution*, and $(\mathcal{A},\varepsilon)$ a *signed support*. We also say that f_c is *honestly n-variate* if and only if the dimension of $\mathrm{Conv}(\mathcal{A})$ is n. We write $\mathbb{R}^{n+k}_\varepsilon=\{c\in\mathbb{R}^{n+k}\mid \mathrm{sign}(c)=\varepsilon\}$ for the appropriate sub-orthant of $\mathbb{R}^{n+k}$.

Two subsets $Z_0, Z_1 \subseteq \mathbb{R}^n$ are *isotopic* (ambiently in $\mathbb{R}^n$) if and only if there is a continuous map $H\colon [0,1]\times\mathbb{R}^n \to \mathbb{R}^n$ such that (1) $H(t,\cdot)$ is a homeomorphism for all $t\in[0,1]$, (2) $H(0,\cdot)$ is the identity on $\mathbb{R}^n$, and (3) $H(1,Z_0)=Z_1$. Isotopy is an equivalence relation on subsets of $\mathbb{R}^n$ [12, Ch. 10.1]. So we can speak of isotopy *type*.

2.2 Signed $\mathcal{A}$-discriminant Contours

We recall the notion of $\mathcal{A}$*-discriminant* from [9], but extended to real exponents as in [1, 4, 19]. A point $x\in\mathbb{R}^n$ is a *singular zero* of f_c if and only if $f_c(x)=\frac{\partial f_c(x)}{\partial x_1}=\cdots=\frac{\partial f_c(x)}{\partial x_n}=0$. We denote the set of singular zeros of f_c by $\mathrm{Sing}(f_c)$. For a fixed signed support $(\mathcal{A},\varepsilon)$, we define the *signed generalized* $\mathcal{A}$*-discriminant variety*, $\Xi_{\mathcal{A},\varepsilon}$, as the (Euclidean) closure of $\left\{c\in\mathbb{R}^{n+k}_\varepsilon \mid \mathrm{Sing}(f_c)\neq\emptyset\right\}$.

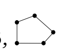
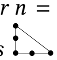
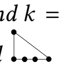
We recall a natural invariance property of the signed $\mathcal{A}$-discriminant.

Proposition 2.1. *[6, Prop. 2.3] Let f_c be an exponential sum with support $\mathcal{A}=\{\alpha_1,\ldots,\alpha_{n+k}\}\subset\mathbb{R}^n$. For an invertible matrix $M\in\mathbb{R}^{n\times n}$ and $v\in\mathbb{R}^n$ consider the exponential sum $g_c(x)=\sum_{i=1}^{n+k} c_i e^{(M\alpha_i+v)\cdot x}$. Then:*

(i) If $\det(M)>0$ then $Z_{\mathbb{R}}(f_c)$ and $Z_{\mathbb{R}}(g_c)$ are isotopic.
(ii) $\mathrm{Sing}(f_c)=M^\top\,\mathrm{Sing}(g_c)$.
(iii) For all $x\in\mathrm{Sing}(g_c)$ the matrices $\mathrm{Hess}(f_c(M^\top x))$ and $\mathrm{Hess}(g_c(x))$ have the same number of positive, negative and zero eigenvalues. ■

Remark 2.2. *An immediate consequence of Proposition 2.1 is that one can transform any full-dimensional support $\mathcal{A}\subset\mathbb{R}^n$ to a support containing the standard basis vectors $e_1,\ldots,e_n\in\mathbb{R}^n$ and the zero vector* **O** *without changing the isotopy type of the underlying hypersurface.* ⋄

For any $\mathcal{A}=\{\alpha_1,\ldots,\alpha_{n+k}\}\subset\mathbb{R}^n$ we let $\widehat{\mathcal{A}}\in\mathbb{R}^{(n+1)\times(n+k)}$ denote the $(n+1)\times(n+k)$ matrix with ith column $[1,\alpha_i]^\top$. If $x\in\mathbb{R}^n$ is a singular zero of f_c then $[c_1e^{\alpha_1\cdot x},\ldots,c_{n+k}e^{\alpha_{n+k}\cdot x}]\in\ker(\widehat{\mathcal{A}})$ (defining $\ker(\widehat{\mathcal{A}})$ as the right nullspace of $\widehat{\mathcal{A}}$). Choose a basis of $\ker(\hat{\mathcal{A}})$ and write these vectors as columns of a matrix $B\in\mathbb{R}^{(n+k)\times(k-1)}$. Such a choice of B is called a *Gale dual matrix* of $\mathcal{A}$. We assume throughout that $\mathcal{A}$ does *not* lie in any affine hyperplane. This is equivalent to $\dim\ker(\widehat{\mathcal{A}})=k-1$.

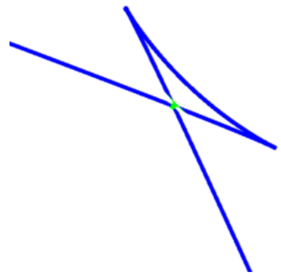
Example 2.3. *Observe that, for $n=2$ and $k=3$,*  *is generic (no 3 points on a line) whereas*  *and*  *are not. However, $\dim\ker(\widehat{\mathcal{A}})=2$ for each of these 3 supports.* ⋄

To understand how the number of pieces of $Z_{\mathbb{R}}(f_c)$ changes as we vary c, let us recall a variant of [19, Dfn. 2.5 & Thm. 3.8].

Lemma 2.4. *Following the notation above, assume $\mathcal{A}\subset\mathbb{R}^n$ is generic, has cardinality $n+k$ and Gale dual matrix B, and let $c,c'\in\mathbb{R}^{n+k}_\varepsilon$. We define a* (signed reduced) $\mathcal{A}$-discriminant contour *of $(\mathcal{A},\varepsilon)$ to be Γ_ε where $\Gamma_\varepsilon:=\Gamma_\varepsilon(\mathcal{A},B):=(\mathrm{Log}|\Xi_{\mathcal{A},\varepsilon}|)B$. If $(\mathrm{Log}|c|)B$ and $(\mathrm{Log}|c'|)B$ are in the same piece of $\mathbb{R}^{k-1}\setminus\Gamma_\varepsilon$, then the zero sets $Z_{\mathbb{R}}(f_c)$ and $Z_{\mathbb{R}}(f_{c'})$ are ambiently isotopic in $\mathbb{R}^n$.* ■

The lemma follows essentially from *Hardt's Triviality Theorem* [10] which, despite the title of [10], holds in the broader context of bounded sub-analytic sets. (Alternatively, one can avail to the *Regular Interval Theorem* [11, Thm. 2.2, Pg. 153].) We will see later, through Theorem 3.2, that as $(\mathrm{Log}\,|c|)B$ crosses the discriminant contour Γ_ε *transversally*, we gain or lose at most one piece for $Z_{\mathbb{R}}(f_c)$.

Example 2.5. *Taking $\mathcal{A}=\{(0,0),(1,0),(0,1),(4,1),(1,4)\}$ as in Example 1.3, just 10 choices of $\varepsilon\in\{\pm\}^5$ yield non-empty Γ_ε: $++--+$, $-+++-$, $-++-+$, $+-++-$, and $+--++$, and their respective negatives. The first 4 choices yield smooth Γ_ε (and thus just two reduced and unbounded chambers in each respective orthant $\mathbb{R}^5_\varepsilon$), while the last yields Γ_{+--++} having exactly 2 cusps:*



In particular, there are exactly 3 signed reduced chambers in this copy of $\mathbb{R}^2$, exactly one of which is bounded: The quasi-triangular reduced chamber to the left is the collection of all $B^\top\,\mathrm{Log}\,|c|$ such that f_c, with signed support $(\mathcal{A},+--++)$, yields $Z_{\mathbb{R}}(f_c)$ having exactly 3 pieces. ⋄

Pieces of the complement of $\Xi_{\mathcal{A},\varepsilon}$ and Γ_ε are respectively called *(signed) chambers* and *(signed) reduced chambers*. In particular, although reduced chambers are $(k-1)$-dimensional, they carry the central information needed to count the pieces of $Z_{\mathbb{R}}(f_c)$ from all $n+k$ coefficients of f_c: The definition of Γ_ε leverages the natural homogeneities of the underlying discriminant obtained by rescaling coefficients and equations.

The contour Γ_ε also admits a very special parametrization known as the *Horn-Kapranov Uniformization* [13]:

Theorem 2.6. *Let $(\mathcal{A},\varepsilon)$ be a signed (generic) support with Gale dual matrix B and set $\mathcal{H}_B:=\{\lambda\in\mathbb{R}^{k-1}\mid$ some coordinate of $B\lambda$ is $0\}$ and $I_{B,\varepsilon}:=\{\lambda\in\mathbb{R}^{k-1}\mid B\lambda\in\mathbb{R}^{n+k}_\varepsilon\}$. Also set $\xi_{B,\varepsilon}(\lambda):=(\mathrm{Log}|\lambda B^\top|)B$ for any $\lambda\in I_{B,\varepsilon}$. Then $\Gamma_\varepsilon=\xi_{B,\varepsilon}(I_{B,\varepsilon})$.* ■

Lemma 2.7. *Any $\lambda\in\mathbb{R}^{k-1}\setminus\mathcal{H}_B$ is normal to the point $\xi_{B,\varepsilon}(\lambda)$, i.e., $\sum_{i=1}^{n+k}\lambda_i\cdot\frac{\partial(\xi_{B,\varepsilon})_i}{\partial\lambda_j}=0$ for all j.* ■

The theorem and lemma were observed in the special case $\mathcal{A}\subset\mathbb{Z}^n$ in [13, Thm. 2.1]. The extension to $\mathcal{A}\subset\mathbb{R}^n$ is in fact an elementary calculation, upon observing that $[c_1e^{\alpha_1\cdot x},\ldots,c_{n+k}e^{\alpha_{n+k}\cdot x}]^\top\in\ker(\widehat{\mathcal{A}})$ if and only if $x\in\mathbb{R}^n$ is a singular point of f_c.

Modifying a Gale dual matrix using elementary column operations gives another choice of Gale dual matrix. Thus we may assume without loss of generality that the last row of the Gale dual matrix B has the form $B_{n+k}=[0,\ldots,0,-1]$. In which case, since $\xi_{B,\varepsilon}$ is homogeneous, we may replace $\mathbb{R}^{k-1}\setminus\mathcal{H}_B$ by the $(k-2)$−dimensional quasi-affine subspace $\{\lambda\in\mathbb{R}^{k-1}\setminus\mathcal{H}_B\mid\lambda_{k-1}=1\}$. In Section 3, we will prefer this latter choice and work with a refined version of $\xi_{B,\varepsilon}$ defined as $\bar{\xi}_{B,\varepsilon}(\mu):=(\mathrm{Log}\,|[\mu,1]B^\top|)B$, for any $\mu\in\mathbb{R}^{k-2}$ with $[\mu,1]B^\top\in\mathbb{R}^{n+k}_\varepsilon$.

2.3 Near-Circuit Exponential Sums

Now let us focus on the case when $k=3$. Then B will have 2 columns and $I_{B,\varepsilon}$ (the domain of the parametrization $\xi_{B,\varepsilon}$ of Γ_ε) will be an open sub-interval of $\mathbb{R}$ with endpoints $-b_{i,2}/b_{i,1}$ and $-b_{j,2}/b_{j,1}$ for some distinct $i,j\in\{1,\ldots,n+3\}$. By reordering $\alpha_1,\ldots,\alpha_{n+3}$ we may assume $i=n+2$ and $j=n+3$. By Proposition 2.1 and translation of x, we may then reduce f_c to:

$$\varepsilon_0+\varepsilon_1e^{x_1}+\cdots+\varepsilon_ne^{x_n}+\varepsilon_{n+1}e^{\beta\cdot x-c_1}+\varepsilon_{n+2}e^{\gamma\cdot x-c_2} \tag{2.1}$$

By column operations on B, we can then reduce the last two rows of B to $[-1, 0]$ and $[0, -1]$. Finally, since $\hat{\mathcal{A}}B = \mathbf{O}$, we then obtain that the Gale dual matrix B has the following form:

$$B = \begin{bmatrix} 1 - \sum_{i=1}^{n} \beta_i & \beta_1 & \cdots & \beta_n & -1 & 0 \\ 1 - \sum_{i=1}^{n} \gamma_i & \gamma_1 & \cdots & \gamma_n & 0 & -1 \end{bmatrix}^{\top} \tag{2.2}$$

Remark 2.8. *We may assume that $[\beta_1 : \cdots : \beta_n]$ and $[\gamma_1 : \cdots : \gamma_n]$ are distinct as points of $\mathbb{P}^{n-1}_{\mathbb{R}}$ for otherwise, by subtracting a suitable multiple of one column of B from another, we could get an affine relation involving just 3 points of $\mathcal{A}$. In other words, 3 points of $\mathcal{A}$ would be collinear, thus contradicting our earlier assumption that $\mathcal{A}$ was generic.* $\diamond$

Via Theorem 2.6 and (2.1) we may parametrize degenerate f_c via $(c_1, c_2) = \left(\sum_{i=1}^{n+3} b_{i,1} \log|b_{i,1}\lambda_1 + b_{i,2}\lambda_2|, \sum_{i=1}^{n+3} b_{i,2} \log|b_{i,1}\lambda_1 + b_{i,2}\lambda_2|\right)$. So we can compute partial derivatives with respect to λ_1:

$$\frac{\partial c_1}{\partial \lambda_1} = \frac{(1 - \sum_{i=1}^{n} \beta_i)^2}{(1 - \sum_{i=1}^{n} \beta_i)\lambda_1 + (1 - \sum_{i=1}^{n} \gamma_i)\lambda_2} + \sum_{i=1}^{n} \frac{\beta_i^2}{\beta_i \lambda_1 + \gamma_i \lambda_2} - \frac{1}{\lambda_1} \tag{2.3}$$

$$\frac{\partial c_2}{\partial \lambda_1} = \frac{1 - \sum_{i=1}^{n} \gamma_i}{\lambda_1 + \frac{1 - \sum_{i=1}^{n} \gamma_i}{1 - \sum_{i=1}^{n} \beta_i} \lambda_2} + \sum_{i=1}^{n} \frac{\gamma_i}{\lambda_1 + \frac{\gamma_i}{\beta_i} \lambda_2} \tag{2.4}$$

Note that, in the near-circuit case, Γ_ε is a piece-wise analytic curve in $\mathbb{R}^2$, provided $I_{B,\varepsilon}$ is non-empty. In which case, we say that Γ_ε *has a cusp at* $p = \xi_{B,\varepsilon}(\ell_1, \ell_2)$ if and only if $\frac{\partial c_1(\ell_1,\ell_2)}{\partial \lambda_1} = \frac{\partial c_2(\ell_1,\ell_2)}{\partial \lambda_1} = 0$. Since $\xi_{B,\varepsilon}$ parametrizes Γ_ε, the preceding condition is certainly necessary for the curve Γ_ε to have a cusp in the usual sense of singularity theory. Thanks to Lemma 2.7, this necessary condition is also sufficient. But we can characterize cusps even more easily:

Lemma 2.9. *Γ_ε has a cusp at $\xi_{B,\varepsilon}(\lambda_1, \lambda_2)$ if and only if $\frac{\partial c_1}{\partial \lambda_1} \frac{\partial c_2}{\partial \lambda_1} = 0$. Moreover, Γ_ε has at most n cusps.*

Proof: By Lemma 2.7, $\lambda_1 \frac{\partial c_1}{\partial \lambda_1} + \lambda_2 \frac{\partial c_2}{\partial \lambda_1} = 0$, one partial vanishing implies the other vanishes. Now just observe that the numerator of $\frac{\partial c_i}{\partial \lambda_1}$ is a homogeneous polynomial of degree n after clearing denominators. ∎

2.4 Reduced Chamber Structure

Following [19], in the near-circuit case $k=3$, we call the bounded (resp. unbounded) pieces of $\mathbb{R}^2 \setminus \Gamma_\varepsilon$ *signed reduced inner (resp. outer) chambers.*

Lemma 2.10. *[6, Proposition 4.6] If $\mathcal{A}$ is a near-circuit and Γ_ε has at most one cusp then the complement of the signed reduced $\mathcal{A}$-discriminant Γ_ε has at most two pieces, both of which are unbounded.* ∎

We want to find the chamber structure when Γ_ε has more cusps. Let us first recall *Cauchy's Mean Value Theorem.*

Lemma 2.11. *(Cauchy's Mean Value Theorem) If $f, g \colon [a, b] \to \mathbb{R}$ are each differentiable on the open interval (a, b) then there is $c \in (a, b)$ such that $(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$.* ∎

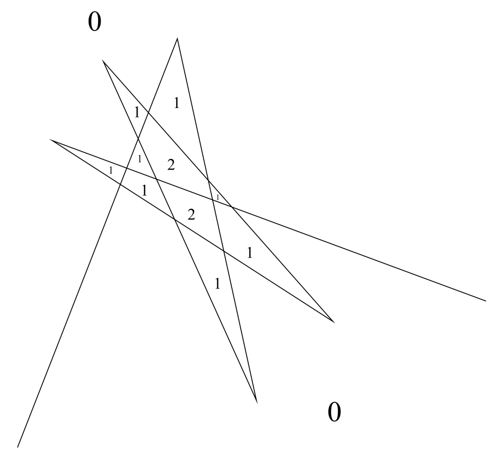
By Lemma 2.9, the signed reduced contour Γ_ε has only finitely many cusps, so Γ_ε is a finite union of smooth images of $[0, 1]$, each of which we will call a *curve segment*, with endpoints that are cusps.

Lemma 2.12. *Every pair of curve segments has relative interiors intersecting at most once.*

Proof: Suppose there are two distinct curve segments intersecting at points $p_1 \neq p_2$, each in the relative interior of both curve segments. Then by Lemma 2.11, in each curve segment, there exists a point such that the tangent line at this point is parallel to the line connecting p_1 and p_2. Therefore, we find two points in Γ_ε such that the tangent lines at these two points are parallel, which contradicts Lemma 2.7. ∎

We call two distinct reduced signed chambers S and S' (i.e., pieces of $\mathbb{R}^{k-1} \setminus \Gamma_\varepsilon$ for some fixed sign-vector ε) *adjacent* if and only if the intersection of the closures of S and S' with Γ_ε contains a relatively open (non-empty) subset of Γ_ε. Clearly, if there is a differentiable path from a point in S to a point in S', and this path intersects Γ_ε transversally, then S and S' must be adjacent. More generally, we define the *depth* of a reduced chamber S as the minimum number of transversal intersections with Γ_ε a differentiable path can have if the path connects S to some *outer* chamber S'. For example, the depth of an outer chamber is 0, and the depth of an inner chamber is 1 if and only if it is adjacent to an outer chamber.

Example 2.13.



The illustration above shows a piece-wise linear model of a hypothetical reduced contour with exactly 5 cusps. (A true contour would have an entire non-empty open interval-worth of slopes.) Two of the segments drawn are intended to be rays going off to infinity, and thus there are exactly 2 outer chambers and 10 inner chambers. Exactly 8 of the inner chambers have depth 1 while the remaining 2 have depth 2. $\diamond$

Lemma 2.14. *Suppose Γ_ε has exactly m cusps. Then each chamber has depth no more than $\lfloor m/2 \rfloor$.*

Proof: We have already built B so that it has bottom row $[0, -1]$. Suppose Γ_ε is parametrized by $\bar{\xi}_{\varepsilon,B}(t)$ where $t \in (t_0, t_\infty)$ and C is a chamber of $\mathbb{R}^2 \setminus \Gamma_\varepsilon$. Let s be a curve segment of Γ_ε such that $s = \bar{\xi}_{\varepsilon,B}((t_1, t_2))$, where $(t_1, t_2) \subset (t_0, t_\infty)$. Given $\delta \in \mathbb{R}$, we define the following path:

$$\gamma_{\delta,s}(t) = \begin{cases} \dfrac{d\bar{\xi}_{\varepsilon,B}}{dt}(t_1) \cdot t + \bar{\xi}_{\varepsilon,B}(t_1) + (0, \delta) & t_0 < t \leq t_1 \\ \bar{\xi}_{\varepsilon,B}(t) + (0, \delta) & t_1 < t < t_2 \\ \dfrac{d\bar{\xi}_{\varepsilon,B}}{dt}(t_2) \cdot t + \bar{\xi}_{\varepsilon,B}(t_2) + (0, \delta) & t_2 \leq t < t_\infty \end{cases} \tag{2.5}$$

Then $(\gamma_{\delta,s}(t))_1$ and $(\gamma_{\delta,s}(t))_2$ have continuous derivatives on (t_0, t_∞). We can choose a point $p \in C$ and a real number δ such that

$p \in \gamma_{\delta,s}(t)$ and $\gamma_{\delta,s}(t)$ intersects Γ_ε transversally. Let $p = \gamma_{\delta,s}(t_p)$. It suffices to show that the number of intersections between $\gamma_{\delta,s}(t)$ and Γ_ε is at most m. Consequently, one of the paths $\gamma_{\delta,s}((t_0, t_p))$ or $\gamma_{\delta,s}((t_p, t_\infty))$ must intersect Γ_ε at most $\lfloor m/2 \rfloor$ times.

In fact, when $\delta \neq 0$, we have $\gamma_{\delta,s} \bigcap s = \emptyset$. Suppose, to the contrary, there exists $t_a \in (t_1, t_2)$ and $t_b \in (t_0, t_\infty)$ such that $\bar{\xi}_{\varepsilon,B}(t_a) = \gamma_{\delta,s}(t_b)$. Then we have $\gamma_{\delta,s}(t_a) - \gamma_{\delta,s}(t_b) = (0, \delta)$. By Lemma 2.11, there exists $t_c \in (t_a, t_b)$ such that the tangent vector at $\gamma_{\delta,s}(t_c)$ is parallel to $(0, \delta)$. In other words, the normal vector at $\gamma_{\delta,s}(t_c)$ is $(1, 0)$. However, by Lemma 2.7, the normal vector at $\gamma_{\delta,s}(t_c)$ is $(t_c, 1)$, leading to a contradiction.

For the remaining m curve segments of Γ_ε, we show that $\gamma_{\delta,s}$ intersects each of them at most once. To see this, let us argue by contradiction: suppose $\gamma_{\delta,s}$ intersects a given segment at two distinct points p_1 and p_2. By Lemma 2.11, there exist a point on the curve segment and a point on $\gamma_{\delta,s}$ where the tangent lines are parallel to the line joining p_1 and p_2. However, there always exists a point in s where the tangent vector is identical to that of $\gamma_{\delta,s}$. This leads to a situation where two points on Γ_ε have the same tangent vectors, which yields a contradiction. ■

Corollary 2.15. *If Γ_ε has exactly 2 or 3 cusps then all inner chambers have depth 1.* ■

Lemma 2.16. *If Γ_ε has exactly 4 cusps, then the depth of any chamber is at most 2. Moreover, any chamber adjacent to a depth-2 chamber must have depth 1.*

Proof: The first statement follows directly from Lemma 2.14. For the second statement, observe that any chamber adjacent to a depth-2 chamber must have depth at least 1; otherwise, it would be an outer chamber, which cannot be adjacent to a depth-2 inner chamber.

Now suppose there exist two adjacent depth-2 inner chambers, denoted by C_1 and C_2, with a common boundary lying on a curve segment s. Let p be a point on this common boundary, and choose a curve segment s' distinct from s. Consider the curve $\gamma_{\delta,s'}(t)$ defined in (2.5), chosen so that $p \in \gamma_{\delta,s'}(t)$. By the proof of Lemma 2.14, $\gamma_{\delta,s'}(t)$ intersects $\Gamma_\varepsilon(\mathcal{A}, B)$ at most 4 times, dividing $\gamma_{\delta,s'}(t)$ into 5 segments. It follows that only one of these segments can have depth 2. However, $C_1 \cap \gamma_{\delta,s'}$ and $C_2 \cap \gamma_{\delta,s'}$ belong to different segments and both have depth 2, which leads to a contradiction. ■

3 Morse Theory and Hessians Along Contours

In this section, we discuss the number of pieces of $Z_\mathbb{R}(f_c)$ for c in each discriminant chamber. Theorem 2.4 established that if two exponential sums belong to the same chamber, then their zero sets contain the same number of pieces. In Theorem 3.2, we describe how the number of pieces varies between adjacent chambers.

Morse Theory plays a crucial role in proving these results. We begin by recalling the parts of Morse Theory we'll need (see, e.g., [16]). Recall that a critical point p for a twice-differentiable function on $\mathbb{R}^n$ is *non-degenerate* if and only if its Hessian at p is nonsingular.

Theorem 3.1. *(Morse Lemma) Suppose $M \subset \mathbb{R}^N$ is an n-dimensional C^∞ submanifold, $h : M \longrightarrow \mathbb{R}$ is C^∞, distinct critical points of h yield distinct critical values, and p is a non-degenerate critical point for h. Then there is a local coordinate system $(y_1, \ldots, y_n)$ in a neighborhood $U \subseteq M$ of p with $y_i(p) = 0$ for all i and such that the identity $h(y) = h(p) - y_1^2 + \cdots - y_s^2 + y_{s+1}^2 + \cdots + y_n^2$ holds on U, where s is the number of negative eigenvalues of Hessian of h at p. We call s the* index *of h at p.* ■

Now let us focus on when the number of pieces changes for c in two adjacent chambers. Let f_c be an exponential sum with $\alpha_1 = \mathbf{O}$ in its support. Then on $Z_\mathbb{R}(f_c)$ we always have $-c_1 = \sum_{i=2}^{n+k} c_i e^{\alpha_i \cdot x}$. By varying the value of c_1 while keeping $c_2, \ldots, c_{n+k}$ fixed, we can trace out an axis-parallel (closed) line segment $\bar{L} \subset \mathbb{R}_\varepsilon^{n+k}$. In particular, if $\bar{L}$ is sufficiently small, then $L = (\mathrm{Log}\,|\bar{L}|)B$ is a 1-dimensional analytic sub-manifold of $\mathbb{R}^{k-1}$ with exactly two boundary points (which we will call *endpoints*). So we obtain a 1-parameter family of zero sets $Z_\mathbb{R}(f_c)$ and, if L intersects Γ_ε transversally, then the corresponding exponential sums transition from one chamber to another. We can then examine the change in the number of pieces of $Z_\mathbb{R}(f_c)$ via the Morse Lemma.

Theorem 3.2. *Following the notation above, assume L connects two adjacent reduced chambers, intersecting Γ_ε only at $\ell^* := (\mathrm{Log}\,|c^*|)B$, and transversally so. (So ℓ^* is a smooth point of Γ_ε.) Also let the endpoints of L be $(\mathrm{Log}\,|c|)B$ and $(\mathrm{Log}\,|c'|)B$, assume c, c', c^* all have sign vector ε, and let N_c (resp. $N_{c'}$, N_{c^*}) denote the number of pieces of $Z_\mathbb{R}(f_c)$ (resp. $Z_\mathbb{R}(f_{c'})$, $Z_\mathbb{R}(f_{c^*})$). Then $Z_\mathbb{R}(f_{c^*})$ has a unique singular point x^*, and $N_c \neq N_{c'}$ only when the index (i.e., the number of negative eigenvalues of the Hessian) of f_{c^*} at the critical point x^* of $c_1(x_1, \ldots, x_n)$ is 0, 1, $n-1$, or n. In particular, $|N_c - N_{c'}| \leq 1$.*

Proof: Since ℓ^* is a smooth point of Γ_ε, c^* must be a smooth point of $\Xi_{\mathcal{A},\varepsilon}$. So then, by [9, Thm. 1.5, Pg. 16], $Z_\mathbb{R}(f_{c^*})$ has a unique singular point x^*. Also note that $\frac{\partial f_c}{\partial x_i} = \frac{\partial f_{c^*}}{\partial x_i} = -\frac{\partial c_1}{\partial x_i}$ for all i. Consider the function h that maps the analytic n-manifold $\{(x_1, \ldots, x_n, c_1) \in \mathbb{R}^{n+1} \mid f_c = 0\}$ to c_1. By [7, Thm. 3.5], the critical point (x^*, c_1^*) is non-degenerate. By Theorem 3.1, in a small neighborhood of the critical point (x^*, c_1^*), we can find a new chart $(y_1, \ldots, y_n)$ such that $c_1 - c_1^* = -y_1^2 - y_2^2 - \cdots - y_s^2 + y_{s+1}^2 + \cdots + y_n^2$, where s is the index of f_{c^*}.

In this quadratic form, as c_1 increases across c_1^*, the isotopy type of the level set of values of $(y_1, \ldots, y_n)$ changes as follows:

- If $s = 0$ or $s = n$, the set changes from an empty set to an $(n-1)$-sphere (or vice versa).
- If $s = 1$ or $s = n-1$, the set changes from a hyperboloid of two sheets to a hyperboloid of one sheet (or vice versa). Both of these cases change the number of pieces.
- If $2 \leq s \leq n-2$, the isotopy type of the set of $(y_1, \ldots, y_n)$ changes from $\mathbb{R}^{n-s} \times \mathbb{S}^{s-1}$ to $\mathbb{R}^s \times \mathbb{S}^{n-s-1}$, which does not change the number of pieces. ■

Thanks to Theorem 3.2, we can determine the number of pieces of $Z_\mathbb{R}(f_c)$, given the number of pieces of $Z_\mathbb{R}(f_{c'})$, by computing the index of the Hessian of an exponential sum corresponding to a point $c^* \in \Xi_{\mathcal{A},\varepsilon}$. From this point onward, we focus on honest n-variate exponential sums with exactly $n+3$ terms and generic supports. As noted earlier, we can reduce any near-circuit exponential sum to a special form so that the Gale dual matrix B has the form shown in (2.2). Let us now compute the Hessian of f_c when $\mathrm{Sing}\, f_c \neq \emptyset$.

Lemma 3.3. *Let f and B be defined as above. If $(c_1, c_2) \in \Gamma_\varepsilon$ with $(c_1, c_2) = \mathrm{Log}\,|(\lambda_1, \lambda_2) \cdot B^\top| \cdot B$, then f has a critical point $x^* = (x_1^*, x_2^*)$ with $x_i^* = \log|\beta_i \lambda_1 + \gamma_i \lambda_2| - \log|(1 - \sum_{i=1}^n \beta_i)\lambda_1 + (1 - \sum_{i=1}^n \gamma_i)\lambda_2|$*

for all i. In particular, the Hessian of f at x^* is

$$\frac{M(\beta)\lambda_1+M(\gamma)\lambda_2}{(1-\sum_{i=1}^n\beta_i)\lambda_1+(1-\sum_{i=1}^n\gamma_i)\lambda_2},\ \text{where } M(\beta)=\begin{bmatrix}\beta_1(\beta_1-1)&\beta_1\beta_2&\cdots&\beta_1\beta_n\\ \beta_1\beta_2&\beta_2(\beta_2-1)&\cdots&\beta_2\beta_n\\ \vdots&\vdots&\ddots&\vdots\\ \beta_1\beta_n&\beta_2\beta_n&\cdots&\beta_n(\beta_n-1)\end{bmatrix}.$$

Proof: It is not hard to check that $f(x^*)=\frac{\partial f}{\partial x_1}(x^*)=\cdots=\frac{\partial f}{\partial x_n}(x^*)=0$ by direct computation. For the Hessian, note that for all i we have $\varepsilon_i e^{x_i^*}=-\beta_i\varepsilon_{n+1}e^{\beta\cdot x^*-c_1}-\gamma_i\varepsilon_{n+2}e^{\gamma\cdot x^*-c_2}$ since $\frac{\partial f}{\partial x_i}(x^*)=0$. ■

As derived earlier, we can assume $\lambda_1=\mu$ and $\lambda_2=1$ since $[0,-1]$ is a row vector of B. Let us now find the characteristic polynomial of the Hessian obtained in Lemma 3.3.

Lemma 3.4. *The characteristic polynomial of $M(\beta)\mu+M(\gamma)$ with $\mu\neq 0$, is $p_\mu(\zeta)=\mu g(\zeta,\mu)\prod_{i=1}^n(\zeta+\beta_i\mu+\gamma_i)$, where $g(\zeta,\mu)$ is the determinant of*

$$\begin{bmatrix}\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & -(1-\sum\limits_{i=1}^n\beta_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta}\\ -(1-\sum\limits_{i=1}^n\beta_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta} & -(1-\sum\limits_{i=1}^n\beta_i)\mu-(1-\sum\limits_{i=1}^n\gamma_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i\mu+\gamma_i}{\beta_i\mu+\gamma_i+\zeta}\end{bmatrix}.$$

Proof: First note that $M(\beta)=\beta^\top\beta-\mathrm{Diag}(\beta_1,\ldots,\beta_n)$. Then

$$\begin{aligned}p_\mu(\zeta)&=\det(\zeta I_n-M(\beta)\mu-M(\gamma))\\ &=\det(\mathrm{Diag}(\zeta+\beta_1\mu+\gamma_1,\ldots,\zeta+\beta_n\mu+\gamma_n)-(\beta^\top\beta\mu+\gamma^\top\gamma))\\ &=\left(\prod_{i=1}^n(\zeta+\beta_i\mu+\gamma_i)\right)\cdot\det\left(I_n-\mathrm{Diag}\left((\zeta+\beta_1\mu+\gamma_1)^{-1},\ldots,(\zeta+\beta_n\mu+\gamma_n)^{-1}\right)\begin{bmatrix}\beta^\top&\gamma^\top\end{bmatrix}\begin{bmatrix}\beta\mu\\ \gamma\end{bmatrix}\right)\\ &=\prod_{i=1}^n(\zeta+\beta_i\mu+\gamma_i)\\ &\quad\times\det\left(I_2-\begin{bmatrix}\beta\mu\\ \gamma\end{bmatrix}\mathrm{Diag}\left((\zeta+\beta_1\mu+\gamma_1)^{-1},\ldots,(\zeta+\beta_n\mu+\gamma_n)^{-1}\right)\begin{bmatrix}\beta^\top&\gamma^\top\end{bmatrix}\right)\\ &=\mu\det\begin{bmatrix}\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & \sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i}\\ \sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i} & \sum\limits_{i=1}^n\frac{\gamma_i^2}{\zeta+\beta_i\mu+\gamma_i}-1\end{bmatrix}\prod_{i=1}^n(\zeta+\beta_i\mu+\gamma_i).\end{aligned}$$

Above, we used Sylvester's determinant identity, which states that $\det(I_m-AB)=\det(I_n-BA)$ if A and B are matrices of sizes $m\times n$ and $n\times m$, respectively.

Now note that we have the identities

$$(1-\sum_{i=1}^n\beta_i)+\mu\left(\sum_{i=1}^n\frac{\beta_i^2}{\beta_i\mu+\gamma_i+\zeta}-\frac{1}{\mu}\right)+\sum_{i=1}^n\frac{\beta_i\gamma_i}{\beta_i\mu+\gamma_i+\zeta}+\zeta\sum_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta}=0$$

and

$$(1-\sum_{i=1}^n\gamma_i)+\mu\left(\sum_{i=1}^n\frac{\beta_i\gamma}{\beta_i\mu+\gamma_i+\zeta}\right)+\sum_{i=1}^n\frac{\gamma_i^2}{\beta_i\mu+\gamma_i+\zeta}-1+\zeta\sum_{i=1}^n\frac{\gamma_i}{\beta_i\mu+\gamma_i+\zeta}=0.$$

Therefore $g(\zeta,\mu)$

$$\begin{aligned}&=\det\begin{bmatrix}\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & \mu\left(\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu}\right)+\sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i}\\ \sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i} & \mu\left(\sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i}\right)+\sum\limits_{i=1}^n\frac{\gamma_i^2}{\zeta+\beta_i\mu+\gamma_i}-1\end{bmatrix}\\ &=\det\begin{bmatrix}\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & -(1-\sum\limits_{i=1}^n\beta_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta}\\ \sum\limits_{i=1}^n\frac{\beta_i\gamma_i}{\zeta+\beta_i\mu+\gamma_i} & -(1-\sum\limits_{i=1}^n\gamma_i)-\zeta\sum\limits_{i=1}^n\frac{\gamma_i}{\beta_i\mu+\gamma_i+\zeta}\end{bmatrix}\\ &=\det\begin{bmatrix}\sum\limits_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & -(1-\sum\limits_{i=1}^n\beta_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta}\\ -(1-\sum\limits_{i=1}^n\beta_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i}{\beta_i\mu+\gamma_i+\zeta} & -(1-\sum\limits_{i=1}^n\beta_i)\mu-(1-\sum\limits_{i=1}^n\gamma_i)-\zeta\sum\limits_{i=1}^n\frac{\beta_i\mu+\gamma_i}{\beta_i\mu+\gamma_i+\zeta}\end{bmatrix}\ ■\end{aligned}$$

For each $\mu\in\mathbb{R}$ the parametrization $\bar{\xi}_{B,\varepsilon}$ provides a point $\bar{\xi}_{B,\varepsilon}(\mu)\in\Gamma_\varepsilon$ corresponding to an exponential sum, allowing us to compute its Hessian. Although we have derived the characteristic polynomial of the Hessian, determining the index directly remain a challenge. However, since μ varies along reduced contour Γ_ε, the index changes only at the cusps.

Lemma 3.5. *For $\mu\neq 0$, the index of the Hessian $(M(\beta)\mu+M(\gamma))$ changes only at the cusps of a discriminant contour, where μ satisfies the condition $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)=0$.*

Proof: The sign of the eigenvalues changes only when the matrix becomes singular, i.e., when 0 is an eigenvalue. From the characteristic polynomial (Lemma 3.4), we deduce that $p_\mu(0)$ is

$$\mu\prod_{i=1}^n(\beta_i\mu+\gamma_i)\det\begin{bmatrix}\sum_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu} & -(1-\sum_{i=1}^n\beta_i)\\ -(1-\sum_{i=1}^n\beta_i) & -(1-\sum_{i=1}^n\beta_i)\mu-(1-\sum_{i=1}^n\gamma_i)\end{bmatrix}.$$

Since $\mu\neq 0$, and $\beta_i\mu+\gamma_i\neq 0$ for all i, we have $p_\mu(0)=0$ if and only if

$$\left(\sum_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu}\right)\left(-(1-\sum_{i=1}^n\beta_i)\mu-(1-\sum_{i=1}^n\gamma_i)\right)-(1-\sum_{i=1}^n\beta_i)^2=0.$$

That is, $\sum_{i=1}^n\frac{\beta_i^2}{\zeta+\beta_i\mu+\gamma_i}-\frac{1}{\mu}+\frac{(1-\sum_{i=1}^n\beta_i)^2}{(1-\sum_{i=1}^n\beta_i)\mu+(1-\sum_{i=1}^n\gamma_i)}=0$, which corresponds exactly to $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)=0$. Therefore, this condition defines the cusps of a discriminant contour. ■

We also show that 0 is a simple root of the characteristic polynomial $p_\mu(\zeta)$ when μ satisfies $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)=0$. This implies that, at the cusps, at most one eigenvalue changes its sign.

Lemma 3.6. *Let $\mu_0\neq 0$ be a root of $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)$ with $\beta_i\mu_0+\gamma_i\neq 0$ for all i and $(1-\sum_{i=1}^n\beta_i)\mu_0+(1-\sum_{i=1}^n\gamma_i)\neq 0$. Then $\frac{\partial p_{\mu_0}}{\partial\zeta}(0)\neq 0$.*

Proof: $\zeta=0$ is a root of $p_{\mu_0}(\zeta)$ if and only if $(0,\mu_0)$ is a root of $g(\zeta,\mu)$ by our assumption. Also, we have that $\frac{\partial p_{\mu_0}}{\partial\zeta}(0)$ is $\mu_0\prod_{i=1}^n(\zeta+\beta_i\mu_0+\gamma_i)\frac{\partial g}{\partial\zeta}(0,\mu_0)$.

By direct computation, one can find the derivative of g with respect to ζ at the point $(0,\mu_0)$:

$$\frac{\partial g}{\partial\zeta}(0,\mu_0)=\left((1-\sum_{i=1}^n\beta_i)\mu_0+(1-\sum_{i=1}^n\gamma_i)\right)\cdot\sum_{i=1}^n\left(\frac{\beta_i}{\beta_i\mu_0+\gamma_i}-\frac{1-\sum_{j=1}^n\beta_j}{(1-\sum_{j=1}^n\beta_j)\mu_0+(1-\sum_{j=1}^n\gamma_j)}\right)^2.$$

By our assumption, $\frac{\partial g}{\partial\zeta}(0,\mu_0)=0$ only when $\frac{\beta_i}{\beta_i\mu_0+\gamma_i}-\frac{1-\sum_{j=1}^n\beta_j}{(1-\sum_{j=1}^n\beta_j)\mu_0+(1-\sum_{j=1}^n\gamma_j)}=0$ for all i, but this can't happen: If so, we would obtain $\frac{\beta_1}{\gamma_1}=\frac{\beta_2}{\gamma_2}=\cdots=\frac{\beta_n}{\gamma_n}$, which contradicts our earlier assumption of $\mathcal{A}$ being generic (see Remark 2.8). ■

In Lemma 3.6, we proved that only one eigenvalue may change its sign at a cusp. Setting $p_\mu(\zeta)=0$ then defines an implicit function $\zeta=\zeta(\mu)$ (from $\mathbb{R}^*$ to $\mathbb{R}$), with $\zeta(\mu_0)=0$, that is well-defined in a neighborhood of μ_0 by Lemma 3.6. Moreover, we can detect the sign change of this eigenvalue by analyzing the derivatives of c_1 with respect to λ_1. The following lemma provides further details:

Lemma 3.7. *Let $\mu_0\neq 0$ be a root of $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)$ as stated in Lemma 3.6, and let $\zeta(\mu)$ be the implicit function defined above. Then, for a sufficiently small neighborhood $\mu\in(\mu_0-\delta,\mu_0+\delta)$ of μ_0, we have $\mathrm{sign}(\zeta(\mu))=\mathrm{sign}\left(\frac{\partial c_1}{\partial\lambda_1}(\mu,1)\right)$.*

Proof: Suppose μ_0 is a root of $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)$ with multiplicity ℓ, which yields $\frac{\partial^{\ell+1}c_1}{\partial\lambda_1^{\ell+1}}(\mu_0,1)\neq 0$ and $\frac{\partial^j c_1}{\partial\lambda_1^j}(\mu_0,1)=0$ for all $1\leq j\leq\ell$.

Let $p(\zeta,\lambda)=\lambda\prod_{i=1}^{n}(\zeta+\beta_i\lambda+\gamma_i)g(\zeta,\lambda)$, where g is the same as in Lemma 3.5. Also, by the proof of Lemma 3.5, one can show that

$$g(0,\mu)=-\left((1-\textstyle\sum_{i=1}^{n}\beta_i)\mu+(1-\sum_{i=1}^{n}\gamma_i)\right)\frac{\partial c_1}{\partial\lambda_1}(\mu,1).$$

Then we have

$\frac{\partial^\ell g}{\partial\mu^\ell}(0,\mu_0)=-\left((1-\sum_{i=1}^{n}\beta_i)\mu_0+(1-\sum_{i=1}^{n}\gamma_i)\right)\frac{\partial^{\ell+1}c_1}{\partial\lambda_1^{\ell+1}}(\mu_0,1)$ and $\frac{\partial^j g}{\partial\mu^j}(\mu_0,1)=0$ for all $0\le j\le\ell-1$.

Therefore, on a small neighborhood of μ_0, we have the following (by induction on the order of the derivatives):

$\frac{\partial^j\zeta}{\partial\mu^j}(\mu_0)=-\frac{\partial^j p}{\partial\mu^j}(0,\mu_0)\Big/\frac{\partial p}{\partial\zeta}(0,\mu_0)=-\frac{\partial^j g}{\partial\mu^j}(0,\mu_0)\Big/\frac{\partial g}{\partial\zeta}(0,\mu_0)=0$ for all $0\le j\le\ell-1$. Also, $\frac{\partial^\ell\zeta}{\partial\mu^\ell}(\mu_0)=-\frac{\partial^\ell g}{\partial\mu^\ell}(0,\mu_0)\Big/\frac{\partial g}{\partial\zeta}(0,\mu_0)$

$$=\frac{\partial^{\ell+1}c_1}{\partial\lambda_1^{\ell+1}}(\mu_0,1)\Bigg/\sum_{i=1}^{n}\left(\frac{\beta_i}{\beta_i\mu_0+\gamma_i}-\frac{1-\sum_{j=1}^{n}\beta_j}{(1-\sum_{j=1}^{n}\beta_j)\mu_0+(1-\sum_{j=1}^{n}\gamma_j)}\right)^2.$$

Now let us consider the Taylor expansions of $\zeta(\mu)$ and $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)$ centered at μ_0 within a small neighborhood $(\mu_0-\delta,\mu_0+\delta)$. Based on the calculations above, the leading terms of these expansions are $\frac{\partial^\ell\zeta}{\partial\mu^\ell}(\mu_0)\frac{(\mu-\mu_0)^\ell}{\ell!}$ and $\frac{\partial^{\ell+1}c_1}{\partial\lambda_1^{\ell+1}}(\mu_0,1)\frac{(\mu-\mu_0)^\ell}{\ell!}$, respectively. These terms dominate the signs of $\zeta(\mu)$ and $\frac{\partial c_1}{\partial\lambda_1}(\mu,1)$ in the neighborhood $(\mu_0-\delta,\mu_0+\delta)$. Since $\frac{\partial^\ell\zeta}{\partial\mu^\ell}(\mu_0)$ and $\frac{\partial^{\ell+1}c_1}{\partial\lambda_1^{\ell+1}}(\mu_0,1)$ share the same sign within this neighborhood, the proof follows. ■

4 Reduced $\mathcal{A}$-discriminants with multiple cusps

We now study near-circuit exponential sums under the additional assumption that their signed contours have at least two cusps. To count the pieces of their real zero sets, we first consider the case where the contour contains at least five cusps. We show that the Hessian has enough positive and negative eigenvalues so that, via the Morse-Theoretic approach of Theorem 3.2, adjacent chambers always yield $Z_{\mathbb{R}}(f_c)$ with the same number of pieces. Subsequently, we examine the scenario where the contour has at most four cusps, similarly analyzing the change in the number of pieces as we traverse contours between chambers.

We begin by recalling a useful result of Bihan, Humbert, and Tavenas.

Lemma 4.1. *[2, Prop. 6.3] Suppose f_c is a near-circuit exponential sum, corresponding to a point in an outer chamber. Then $Z_{\mathbb{R}}(f_c)$ has at most 2 pieces.* ■

By combining Theorem 3.2 and Lemma 4.1, we obtain a rough bound on the number of pieces.

Corollary 4.2. *If f_c lies in a reduced signed chamber of depth d then $Z_{\mathbb{R}}(f_c)$ has at most $2+d$ pieces.* ■

Remark 4.3. *We'll abuse slightly by saying "f in a chamber" to mean f such that $(\operatorname{Log}|c|)B$ is in a chamber. Similarly, when we speak of the number of pieces in a chamber, we'll really mean the number of pieces of $Z_{\mathbb{R}}(f_c)$ for $(\operatorname{Log}|c|)B$ in a chamber.* ⋄

Using Theorem 3.2, we can more precisely count the pieces in the inner chambers by determining the index of f_c when f_c lies on a discriminant contour. While it may be challenging to directly compute the index of f_c at an arbitrary point of a discriminant contour, Lemma 3.5 and Lemma 3.7 allow us to determine the index of f_c as $\mu\to 0$, and track the sign changes of eigenvalues as μ changes.

Before explaining this, recall that we use the canonical form of f_c as given in (2.1) and the Gale dual matrix in (2.2). If f_c contains a singular zero, the Hessian of f_c is $M(\beta)\mu+M(\gamma)$, up to a scalar, as shown in Lemma 3.3. By taking the limit $\mu\to 0$, we obtain $M(\gamma)$, whose index can be determined.

Lemma 4.4. *The number of the positive (resp. negative) eigenvalues of $M(\gamma)$ (counted with multiplicity) is q (resp. $n-q$), where q is the number of the positive entries of $((\sum_{i=1}^{n}\gamma_i)-1,-\gamma_1,\ldots,-\gamma_n)$.*

Proof: By the proof of Lemma 3.4, as $\mu\to 0$, the characteristic polynomial is given by $p_0(\zeta)=\left(1-\sum_{i=1}^{n}\frac{\gamma_i^2}{\zeta+\gamma_i}\right)\prod_{i=1}^{n}(\zeta+\gamma_i)$.

We first assume that the γ_i are distinct. In this case, if $\zeta=-\gamma_j$ for some j, then $p_0(-\gamma_j)=\gamma_j^2\prod_{i\neq j}(\gamma_i-\gamma_j)$. Since $\gamma_j\neq 0$ (by the assumption of the Gale dual matrix B) and $\gamma_i\neq\gamma_j$, it follows that $-\gamma_j$ is not a root of $p_\mu(\zeta)$.

As $p_\mu(\zeta)$ is a polynomial of degree n, its n roots must all come from solving $\bar p_0(\zeta):=1-\sum_{i=1}^{n}\frac{\gamma_i^2}{\zeta+\gamma_i}=0$. Note that $\bar p$ has n poles at $-\gamma_1,\ldots,-\gamma_n$. Without loss of generality, assume that $-\gamma_1<-\gamma_2<\cdots<-\gamma_k<0<-\gamma_{k+1}<\cdots<-\gamma_n$ for some $k\in\{1,\ldots,n\}$. Moreover, we have $\bar p_0'(\zeta)=\sum_{i=1}^{n}\frac{\gamma_i^2}{(\zeta+\gamma_i)^2}>0$, showing $\bar p$ is strictly increasing in each interval between its poles.

At each pole $-\gamma_j$, the behavior of $\bar p$ is characterized as follows: $\lim_{\zeta\to-\gamma_j^-}\bar p_0(\zeta)=+\infty$ and $\lim_{\zeta\to-\gamma_j^+}\bar p_0(\zeta)=-\infty$. Thus $\bar p_0$ has the following types of roots:

- $k-1$ **negative roots**, respectively located in the open intervals $(-\gamma_1,-\gamma_2),\ldots,(-\gamma_{k-1},-\gamma_k)$;
- $n-k-1$ **positive roots**, respectively located in the intervals $(-\gamma_{k+1},-\gamma_{k+2}),\ldots,(-\gamma_{n-1},-\gamma_n)$;
- One **positive root** in $(-\gamma_n,+\infty)$, since $\lim_{\zeta\to\infty}\bar p_0(\zeta)=1$.

The sign of the root in $(-\gamma_k,-\gamma_{k+1})$ depends on the value of $\bar p_0(0)=1-\sum_{i=1}^{n}\gamma_i$ as follows:

- If $\bar p_0(0)>0$, the root in this interval is negative, so the number of positive (resp. negative) eigenvalues of $M(\gamma)$ is $n-k$ (resp. k). Since $q=n-k$, this completes the proof.
- Otherwise, the root is positive, so the number of positive (resp. negative) eigenvalues of $M(\gamma)$ is $n-k+1$ (resp. $k-1$). Since $q=n-k+1$, this also completes the proof.

Now we consider the case when the γ_i are not distinct. Suppose $\gamma_1,\ldots,\gamma_\ell$ are distinct with $1\le\ell\le n$, and $p_0(\zeta)$ is $\left(1-\sum_{i=1}^{\ell}\frac{k_i\gamma_i^2}{\zeta+\gamma_i}\right)\prod_{i=1}^{\ell}(\zeta+\gamma_i)^{k_i}$, where $k_i\ge 1$ is the multiplicity of γ_i.

Now, consider the polynomial $\hat p_0(\zeta):=p_0(\zeta)/\prod_{i=1}^{\ell}(\zeta+\gamma_i)^{k_i-1}$ $=\left(1-\sum_{i=1}^{\ell}\frac{k_i\gamma_i^2}{\zeta+\gamma_i}\right)\prod_{i=1}^{\ell}(\zeta+\gamma_i)$. The signs of the roots of $\hat p_0(\zeta)$ correspond to the case we discussed previously, where the γ_i (with $1\le i\le\ell$) are distinct.

Also, $p_\mu(\zeta)$ has roots $-\gamma_1,\ldots,-\gamma_\ell$, with respective multiplicities $k_1-1,\ldots,k_\ell-1$. Therefore, the number of positive (resp. negative) eigenvalues is still given by q (resp. $n-q$). ■

Consider now the case where the signed reduced contour Γ_ε is parameterized by $\mu\in(0,\infty)$, i.e., $\Gamma_\varepsilon=\{\xi_{\varepsilon,B}(\mu,1)\mid\mu\in(0,\infty)\}$. (The case where $\mu\in(-\infty,0)$ is similar.)

Lemma 4.5. *Suppose the signed reduced contour Γ_ε has no poles when $\mu > 0$. If Γ_ε has m cusps (counted with multiplicities, and $2 \le m \le n$), then $\left(1 - \sum_{i=1}^n \gamma_i, \gamma_1, \ldots, \gamma_n\right)$ contains $\lfloor m/2 \rfloor + 1$ positive entries and $\lfloor (m+1)/2 \rfloor$ negative entries.*

Proof: By Lemma 2.9, the reduced contour has a cusp if and only if $\frac{\partial c_2}{\partial \lambda_1} = 0$. Let μ_i $(1 \le i \le m)$ be the positive roots of $\frac{\partial c_2}{\partial \lambda_1}$. Also, $\frac{\partial c_2}{\partial \lambda_1}$ is a rational function whose numerator has degree n, which yields that $\frac{\partial c_2}{\partial \lambda_1}$ has at most $n - m$ negative roots.

Let b_{ij} (for $1 \le i \le n+3$ and $1 \le j \le 2$) be the entries of B. We call μ a *pole* of $\frac{\partial c_2}{\partial \lambda_1}$ if $b_{i,1}\mu + b_{i,2} = 0$ for some i. Then by our assumption, $\frac{\partial c_2}{\partial \lambda_1}$ has $n+1$ negative poles: $-\frac{b_{i,2}}{b_{i,1}}$ for $i \in \{1, \ldots, n+1\}$. These $n+1$ poles cut the negative axis $(-\infty, 0)$ into $n+2$ sub-intervals, or n sub-intervals excluding those with endpoints $-\infty$ or 0.

Case 1: m is odd. Since $\frac{\partial c_2}{\partial \lambda_1}$ has at most $n - m$ negative roots, there are at least m sub-intervals with no real roots of $\frac{\partial c_2}{\partial \lambda_1}$. Among these m sub-intervals, there are $(m+1)/2$ of them that are not adjacent when m is odd.

Suppose the endpoints of these $(m+1)/2$ intervals are $(-b_{i_1,2}/b_{i_1,1}, -b_{i_2,2}/b_{i_2,1}), \ldots, (-b_{i_m,2}/b_{i_m,1}, -b_{i_{m+1},2}/b_{i_{m+1},1})$. We claim that $\mathrm{sign}(b_{i_1,2}, b_{i_2,2}, \ldots, b_{i_{m+1},2})$ contains $(m+1)/2$ positive signs and $(m+1)/2$ negative signs. In fact, on each interval $(-b_{i_{2t-1},2}/b_{i_{2t-1},1}, -b_{i_{2t},2}/b_{i_{2t},1})$ (for $t \in \{1, \ldots, (m+1)/2\}$), we have that $\frac{\partial c_2}{\partial \lambda_1}$ has no root, and hence the sign does not change. We now show that $b_{i_{2t-1},2}$ and $b_{i_{2t},2}$ have different signs: Assume without loss of generality that $\frac{\partial c_2}{\partial \lambda_1} > 0$ on $(-b_{i_{2t-1},2}/b_{i_{2t-1},1}, -b_{i_{2t},2}/b_{i_{2t},1})$. Therefore,

$$\lim_{\mu \to (-b_{i_{2t-1},2}/b_{i_{2t-1},1})^+} \frac{\partial c_2}{\partial \lambda_1}(\mu, 1) = \lim_{\mu \to (-b_{i_{2t},2}/b_{i_{2t},1})^-} \frac{\partial c_2}{\partial \lambda_1}(\mu, 1) = +\infty.$$

By Equation (2.4), the expression of $\frac{\partial c_2}{\partial \lambda_1}(\mu, 1)$ is a sum of fractions. So the signs as $\mu \to (-b_{i_{2t-1},2}/b_{i_{2t-1},1})^+$ (resp. $\mu \to (-b_{i_{2t},2}/b_{i_{2t},1})^-$) only depend on the term $b_{i_{2t-1},2}/(\mu + \frac{b_{i_{2t-1},2}}{b_{i_{2t-1},1}})$ (resp. $b_{i_{2t},2}/(\mu + \frac{b_{i_{2t},2}}{b_{i_{2t},1}})$). Hence, $b_{i_{2t-1},2} > 0$ and $b_{i_{2t},2} < 0$. This concludes the proof.

Case 2: m is even. Let $(-b_{i_l,2}/b_{i_l,1}, 0)$ be the rightmost sub-interval. If there is a root of $\frac{\partial c_2}{\partial \lambda_1}$ on $(-b_{1,2}/b_{1,1}, 0)$, then by the proof above, there are $m+1$ sub-intervals not containing any real roots of $\frac{\partial c_2}{\partial \lambda_1}$, so there are $(m+2)/2$ of them that are not adjacent, which give $(m+2)/2$ positive signs and $(m+2)/2$ negative signs as in the proof above.

Now we just need to consider the case when $\frac{\partial c_2}{\partial \lambda_1}$ has no roots on $(-b_{i_l,2}/b_{i_l,1}, 0)$. Note that by (2.4), we have $\frac{\partial c_2}{\partial \lambda_1}(0, 1) = 1 > 0$. Hence $\frac{\partial c_2}{\partial \lambda_1} > 0$ on $(-b_{i_l,2}/b_{i_l,1}, 0)$. Thus $\lim_{\mu \to (-b_{i_l,2}/b_{i_l,1})^+} \frac{\partial c_2}{\partial \lambda_1}(\mu, 1) = +\infty$, which means $b_{i_l,2} > 0$. Similar to the prove above, precluding the intervals with endpoints $-\infty$ or 0, there are m sub-intervals with no real roots of $\frac{\partial c_2}{\partial \lambda_1}$. Among these m sub-intervals, there are $m/2$ sub-intervals not adjacent to each other and none is adjacent to $(-b_{i_l,2}/b_{i_l,1}, 0)$. Therefore, we have $m/2$ negative $b_{i,2}$ and $m/2$ positive $b_{i,2}$. In addition to $b_{i_l,2} > 0$, there are thus $\frac{m}{2} + 1$ positive entries and $m/2$ negative entries in $(b_{12}, b_{22}, \ldots, b_{n+1,2}) = (1 - \sum_{i=1}^n \gamma_i, \gamma_1, \ldots, \gamma_n)$. ■

Corollary 4.6. *Under the same assumptions as Lemma 4.5, if Γ_ε has m cusps (counted with multiplicities), then $M(\gamma)$ has at least $\lfloor (m+1)/2 \rfloor$ positive eigenvalues and $\lfloor m/2 \rfloor$ negative eigenvalues.*

Proof of Corollary 4.6: This follows directly from Lemma 4.4 and Lemma 4.5. ■

Proof of Theorem 1.2: By Lemma 2.1, we can suppose f is as in Equation (2.1): $\varepsilon_0 + \varepsilon_1 e^{x_1} + \cdots + \varepsilon_n e^{x_n} + \varepsilon_{n+1} e^{\beta \cdot x - c_1} + \varepsilon_{n+2} e^{\gamma \cdot x - c_2}$, and the Gale dual matrix B is as in Equation (2.2). Furthermore, we have that the signed reduced contour Γ_ε admits a parametrization of the form $(c_1, c_2) = (\mathrm{Log}\,|[\mu, 1]B^\top|)B$, where $\mu \in (-\infty, 0)$ or $\mu \in (0, +\infty)$. Let us discuss the cases of the number of the cusps (counting multiplicities) in Γ_ε.

- Γ_ε **has at least 6 cusps.** By Corollary 4.6, $M(\gamma)$ has at least 3 positive eigenvalues and 3 negative eigenvalues. Since $\frac{\partial c_1}{\partial \lambda_1}(\mu, 1)$ is a continuous function on $(-\infty, 0)$ or $(0, +\infty)$, by Lemma 3.7, there are at least 2 positive eigenvalues and 2 negative eigenvalues in the Hessian $M(\beta)\mu + M(\gamma)$ for any $\mu \in (-\infty, 0)$ or $\mu \in (0, +\infty)$. Therefore, by Theorem 3.2 and Lemma 4.1, the real zero set of f has at most 2 pieces.
- Γ_ε **has exactly 5 cusps.** We first show that this case only occurs when $\mu \in (-\infty, 0)$. Indeed, if $\mu \in (0, +\infty)$, then $\frac{\partial c_2}{\partial \lambda_1}(\mu, 1)$ has exactly 5 positive roots. Since $\frac{\partial c_2}{\partial \lambda_1}(\mu, 1)$ is a rational function and all the poles are negative, then it has the form $\frac{\partial c_2}{\partial \lambda_1}(\mu, 1) = \frac{(\mu - \mu_1) \cdots (\mu - \mu_5)(\mu + \mu_6) \cdots (\mu + \mu_\ell)}{(\mu + p_1) \cdots (\mu + p_{n+1})} h(\mu)$, where $\mu_i > 0$ and $p_i > 0$ for all i, and $h(\mu)$ is a monic polynomial with no real roots. Since $\lim_{\mu \to +\infty} h(\mu) = +\infty$ we have $h(0) > 0$. Thus we have $\frac{\partial c_2}{\partial \lambda_1}(0, 1) = -\frac{\mu_1 \cdots \mu_\ell}{p_1 \cdots p_{n+1}} h(0) < 0$, which contradicts the fact that $\frac{\partial c_2}{\partial \lambda_1}(0, 1) = 1 > 0$ by (2.4).
  By Corollary 4.6, $M(\gamma)$ has at least 3 positive eigenvalues and 2 negative eigenvalues. Also, by Lemma 3.7 and the fact that $\lim_{\mu \to 0^-} \frac{\partial c_1}{\partial \lambda_1}(\mu, 1) = +\infty$, the Hessian $M(\beta)\mu + M(\gamma)$ has at least 2 positive eigenvalues and 2 negative eigenvalues when $\mu \in (-\infty, 0)$. It follows that the real zero set of f has at most 2 pieces, by Theorem 3.2 and Lemma 4.1.
- Γ_ε **has exactly 4 cusps.** By Lemma 2.16, the depth of inner chambers in this case is at most two. We aim to show that the number of pieces in a depth-2 inner chamber matches that of an adjacent depth-1 chamber. Specifically, the boundary of a depth-2 inner chamber has at least 3 edges, each originating from a distinct curve segment. By Corollary 4.6 and Lemma 3.7, there are only two curve segments on which $M(\beta)\mu + M(\gamma)$ may have exactly 1 positive or 1 negative eigenvalue, across which the number of pieces could change. Thus, among the 3 edges on the boundary of a depth-2 inner chamber, there is an edge along which $M(\beta)\mu + M(\gamma)$ has at least 2 positive and 2 negative eigenvalues. By Theorem 3.2, this depth-2 inner chamber yields $Z_+(f)$ having the same number of pieces as those $Z_+(f)$ coming from some adjacent chamber across this edge, which, by Lemma 2.16, has depth 1. Hence, the $Z_+(f)$ coming from depth-2 inner chambers have the same number of pieces as those $Z_+(f)$ coming from some depth-1 chamber. Consequently, the real zero set of f has at most 3 pieces.
- Γ_ε **has exactly 2 or 3 cusps.** By Corollary 2.15 and 4.2, the real zero set of f has at most 3 pieces.
- Γ_ε **has exactly 0 or 1 cusps.** By Lemma 2.10 and 4.1, the real zero set of f has at most 2 pieces. ■

References

[1] Osbert Bastani, Christopher J. Hillar, Dimitar Popov, and J. Maurice Rojas. 2011. Randomization, sums of squares, near-circuits, and faster real root counting. In *Randomization, relaxation, and complexity in polynomial equation solving*. Contemp. Math., Vol. 556. Amer. Math. Soc., Providence, RI, 145–166. https://doi.org/10.1090/conm/556/11012

[2] F. Bihan, T. Humbert, and S. Tavenas. 2022. New bounds for the number of connected components of fewnomial hypersurfaces. *arXiv* 2208.04590 (2022).

[3] F. Bihan, J. M. Rojas, and Frank Sottile. 2008. On the sharpness of fewnomial bounds and the number of components of fewnomial hypersurfaces. *Algorithms in algebraic geometry* (2008), 15–20.

[4] Frédéric Bihan, J. Maurice Rojas, and Casey E. Stella. 2009. Faster real feasibility via circuit discriminants. In *ISSAC 2009—Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation*. ACM, New York, 39–46. https://doi.org/10.1145/1576702.1576711

[5] F. Bihan and Frank Sottile. 2009. Betti number bounds for fewnomial hypersurfaces via stratified morse theory. *Proc. Amer. Math. Soc.* 137(9) (2009), 2825–2833.

[6] W. Deng, J. M. Rojas, and M. L. Telek. 2025. Viro's patchworking and the signed reduced A-discriminant. *J. Symb. Comput.* in revision (2025).

[7] J. Forsgård. 2019. Defective dual varieties for real spectra. *J. Algebraic Comb.* 49 (2019), 49–67.

[8] J. Forsgård, M. Nisse, and J. M. Rojas. 2017. New Subexponential Fewnomial Hypersurface Bounds. *arXiv* 1710.00481 (2017).

[9] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. 1994. *Discriminants, Resultants, and Multidimensional Determinants.* Birkhäuser.

[10] Robert M. Hardt. 1980. Semi-algebraic local-triviality in semi-algebraic mappings. *Amer. J. Math.* 102, 2 (1980), 291–302. https://doi.org/10.2307/2374240

[11] Morris W. Hirsch. 1994. *Differential topology*. Graduate Texts in Mathematics, Vol. 33. Springer-Verlag, New York. x+222 pages. Corrected reprint of the 1976 original.

[12] S. Kalajdzievski. 2014. *An Illustrated Introduction to Topology and Homotopy.* CRC Press.

[13] M. Kapranov. 1991. A characterization of A-discriminantal hypersurfaces in terms of the logarithmic Gauss map. *Math. Ann.* 290 (1991), 277–285. https://api.semanticscholar.org/CorpusID:121876304

[14] A. G. Khovanskii. 1991. *Fewnomials.* American Mathematical Society, Providence, R.I.

[15] T.-Y. Li, J. M. Rojas, and X. Wang. 2003. Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces. *Discrete Comput. Geom.* 30 (2003), 379–414.

[16] J. Milnor. 1963. *Morse Theory.* Princeton University Press.

[17] Philippe Pébay, J. Maurice Rojas, and David C. Thompson. 2011. Optimizing n-variate $(n+k)$-nomials for small k. *Theoret. Comput. Sci.* 412, 16 (2011), 1457–1469. https://doi.org/10.1016/j.tcs.2010.11.053

[18] D. Perrucci. 2005. Some bounds for the number of components of real zero sets of sparse polynomials. *Discrete Comput. Geom.* 34 (2005), 475–495.

[19] J. M. Rojas and K. Rusek. 2017. A-Discriminants for Complex Exponents and Counting Real Isotopy Types. *arXiv* 1612.03458 (2017).