Quickly Computing Isotopy Type for Exponential Sums over Circuits (Extended Abstract)

Frédéric Bihan
Université Savoie Mont Blanc
Frederic.Bihan@univ-savoie.fr

Erika Croy
First Interstate Bank
erefsland@gmail.com

Weixun Deng
Texas A&M University
deng15521037237@tamu.edu

Kaitlyn Phillipson
University of Wisconsin, Madison
kphillipson@wisc.edu

Robert J. Rennie
Colorado College
robert.rennie.phd@gmail.com

J. Maurice Rojas
Texas A&M University
rojas@tamu.edu

Fewnomial Theory [Kho91] has established bounds on the number of connected components (a.k.a. *pieces*) of a broad class of real analytic sets as a function of a particular kind of input complexity, e.g., the number of distinct exponent vectors among a generating set for the underlying ideal. Here, we pursue the algorithmic side: We show how to efficiently compute the exact isotopy type of certain (possibly singular) real zero sets, instead of just estimating their number of pieces. While we focus on the *circuit* case, our results form the foundation for an approach to the general case that we will pursue later.

Assume $f \in \mathbb{R}\left[x_1^{\pm 1}, \ldots, x_n^{\pm 1}\right]$ can be written in the form $f(x) = \sum\limits_{i=1}^{n+k} c_i x^{a_i}$ where $a_i = (a_{1,i}, \ldots, a_{n,i})^\top$ ($(\cdot)^\top$ denoting transpose), the notation $x^{a_i} = x_1^{a_{i,1}} \cdots x_n^{a_{i,n}}$ is understood, and the $(n+1)\times(n+k)$ matrix $\widehat{\mathcal{A}} := \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_n \end{bmatrix}$ has rank $n+1$. (So $k \geq 1$ is forced.) We call such an f an *honestly n-variate* $(n+k)$-nomial (over $\mathbb{R}$), and we call $\mathrm{Supp}(f) := \{a_1, \ldots, a_{n+k}\}$ the *support* of f.

Writing $Z_+(f)$ for the zero set of f in the positive orthant $\mathbb{R}^n_+$, it has been known at least since the 1990s that the maximal number of pieces for $Z_+(f)$ is $2^{O(n^2)}$ — independent of the degree of f — provided $k \leq n$ and $Z_+(f)$ is smooth (see, e.g., [Kho91, Sec. 3.14, Cor. 5]). After various improvements and extensions, a *tight* upper bound of 2 was proved for the special case $k = 2$ [BRS09, Bih11]. Tight upper bounds still remain unknown even for $k = 3 \leq n$ and $n = 2 \leq k-2$. Theorem 1.2 below completes the case $k = 2$ by giving an efficient algorithm that completely determines the underlying isotopy type, in addition to counting the exact number of pieces.

Fewnomial Theory has always considered more than just algebraic sets. For example, assume $g(y) := \sum_{i=1}^{n+k} c_i e^{a_i \cdot y}$ where $a_i \in \mathbb{R}^n$, $y := (y_1, \ldots, y_n)$, $a_i \cdot y$ denotes the standard dot product, and (as before) the matrix $\widehat{\mathcal{A}}$ has rank $n+1$. We then call such a g an *honestly n-variate exponential* $(n+k)$-*sum (over $\mathbb{R}$)*, and call $\{a_1, \ldots, a_{n+k}\}$ the *spectrum* of g and each a_i a *frequency*. Writing $Z_\mathbb{R}(g)$ for the real zero set of g, it is clear that $Z_{\mathbb{R}}(g)$ and $Z_+(f)$ are homeomorphic (and even isotopic) when the spectrum of g lies in $\mathbb{Z}^n$. More to the point, the two upper bounds on the number of pieces of $Z_+(f)$ stated in the preceding paragraph *also* apply to $Z_{\mathbb{R}}(g)$.

We will use g to state our topological results and f (restricted to integer coefficients) to state our algorithmic results. Let us now recall the notions of isotopy and diffeotopy.

Definition 1.1 *Given any two subsets $X, Y \subseteq \mathbb{R}^n$, an* isotopy from X to Y (ambient in $\mathbb{R}^n$) *is a continuous map $I : [0,1] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ satisfying (1) $I(t,\cdot)$ is a homeomorphism for all $t \in [0,1]$, (2) $I(0,x) = x$ for all $x \in \mathbb{R}^n$, and (3) $I(1,X) = Y$. Similarly, a* (C^∞-) diffeotopy (ambient in $\mathbb{R}^n$) *is the corresponding notion where homeomorphism is replaced by (C^∞-) diffeomorphism.* $\diamond$

Our algorithmic results will be stated relative to the classical Turing model. In what follows, S^ℓ denotes the unit ℓ-sphere in $\mathbb{R}^{\ell+1}$ (so $S^0 = \{-1,1\}$) and, for $1 \leq \ell \leq m$, we set

$$C_{m,\ell} := Z_{\mathbb{R}}(x_1^2 + \cdots + x_\ell^2 - x_{\ell+1}^2 - \cdots - x_m^2) \subset \mathbb{R}^m.$$

Note that $C_{m,m}$ is a point (the origin, $\mathbf{O}$, in $\mathbb{R}^m$) and, for all $m \geq 2$, $C_{m,\ell}$ is a singular hypersurface.

Theorem 1.2 *Suppose g is an honestly n-variate exponential $(n+k)$-sum over $\mathbb{R}$, f is an honestly n-variate $(n+k)$-nomial with* $\mathrm{Supp}(f) \subset [-d,d]^n$, *and all the coefficients of f lie in $\{\pm 1, \ldots, \pm H\}$. Then:*

1. *If $k = 1$ then $Z_{\mathbb{R}}(g)$ is either empty or ambiently diffeotopic in $\mathbb{R}^n$ to $\mathbb{R}^{n-1} \times \{1\}$, with the latter occuring if and only if the coefficients of g are* not *all of the same sign.*

2. *If $k = 2$ then $Z_{\mathbb{R}}(g)$ is isotopic to one of the following types of set:*
   (a) the empty set
   (b) $S^\ell \times \mathbb{R}^{n-\ell-1}$ for some $\ell \in \{0, \ldots, n-1\}$
   (c) $C_{m,\ell} \times \mathbb{R}^{n-m}$ for some $m \in \{0, \ldots, n\}$ and $\ell \in \{\lceil m/2 \rceil, \ldots, m\}$.

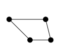
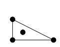
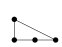
3. *If $k = 2$ then, in time $(\log(dH))^{O(n)}$, we can find $\varepsilon_0, \ldots, \varepsilon_n \in \{-1,0,1\}$ such that $q(x) := \varepsilon_0 + \varepsilon_1 x_1^2 + \cdots + \varepsilon_n x_n^2$ has $Z_{\mathbb{R}}(q)$ ambiently isotopic in $\mathbb{R}^n$ to $Z_+(f)$.*

4. *Fix any $\varepsilon > 0$. If there is an algorithm that can decide for all $n \in \mathbb{N}$ and $k \leq n^\varepsilon$ whether $Z_+(f)$ is non-empty, using a number bit operations polynomial in $n + \log(dH)$, then* $\mathbf{P} = \mathbf{NP}$.

While Assertions (1) and (4) were known earlier (see, e.g., [BRS09]), Assertions (2) and (3) are new. In particular, the best previous deterministic complexity bound in the setting of Assertion (3) appears to be $n^{O(n \log^3 n)} d^{O(n \log^2 n)}$ arithmetic operations [BR14].

The spectra we have focussed on so far have a particular combinatorial structure:

Definition 1.3 *We say that $\{a_1, \ldots, a_t\} \subset \mathbb{R}^n$, with cardinality $t \geq 3$, is a* non-degenerate circuit *if and only if $\{a_1, \ldots, a_t\}$ is affinely dependendent but every proper subset is affinely independent. Finally, we call $\{a_1, \ldots, a_t\}$ a* degenerate circuit *if and only if $\{a_1, \ldots, a_t\}$ has a proper subset B, and an index j, such that $\{a_1, \ldots, a_t\} \setminus \{a_j\}$ is affinely independent and B is a non-degenerate circuit.* $\diamond$

For instance, both  and  are circuits, but  is a degenerate circuit. For any degenerate circuit $\mathcal{A}$, the subset B named above is always unique.

When the spectrum of g is a fixed circuit and the signs of the coefficients are fixed, but the norms of the coefficients are allowed to vary, the number of possible isotopy types for $Z_{\mathbb{R}}(g)$ is actually very small: No more than 3. (See Corollary 1.4 below). A simple example is given by the family of positive zero sets $\{Z_+(1+x^3+y^2-cxy) \mid c>0\}$: The resulting isotopy types are empty, a point, and an oval, according as c is less than, equal to, or greater than $\frac{6}{2^{1/3}3^{1/2}}=2.749459...$

The total number of possible isotopy types, over all circuit spectra in $\mathbb{R}^n$ *and* real coefficients, is a bit larger: Quadratic in n.

Corollary 1.4 *Following the notation of Theorem 1.2, fix n. Then the number of possible isotopy types for $Z_+(g)$ is:*

1. *exactly 2, when $k=1$ (and* both *the coefficients and spectrum of g are allowed to vary).*
2. *at most 3, when $k=2$ and the spectrum of g and the signs of the coefficients are fixed, but the norms of the coefficients of g are allowed to vary.*
3. *exactly $\frac{1}{4}n^2+n+\varepsilon_n$, where ε_n is $\frac{7}{4}$ or 3, according as $n\geq 1$ is odd or even, when $k=2$ and* both *the spectrum and coefficients of g are allowed to vary.*

In particular, $Z_+(g)$ has at most 2 pieces, and has just 1 piece if there is a compact piece.

Our results yield the first upper bounds *polynomial* in n for the number of possible isotopy types of positive zero sets of honestly n-variate $(n+2)$-nomials. (Which isotopy types can actually occur depends on the oriented matroid structure of the spectrum.) More general results on homotopy types (e.g., [GVZ04, BV07]) only yield upper bounds exponential in n in our setting. We also note that, in the setting of Theorem 1.2, the number of distinct isotopy types remains the same if one restricts to spectra in $\mathbb{Z}^n$.

When the spectrum of g is fixed and the coefficients of g vary, the isotopy type of $Z_{\mathbb{R}}(g)$ depends on a generalization [BRS09] of the $\mathcal{A}$-*discriminant* [GKZ94] to real exponents.

Definition 1.5 *Given any non-degenerate circuit $\mathcal{A}=\{a_1,\ldots,a_{n+2}\}\subset\mathbb{R}^n$ of cardinality $n+2$, let $b=(b_1,\ldots,b_{n+2})^\top$ be any generator for the right nullspace of $\widehat{\mathcal{A}}$. We then call b* a circuit relation *for $\mathcal{A}$ and define the* generalized circuit discriminant *to be*

$$\Xi_{\mathcal{A}}(c_1,\ldots,c_{n+2}):=\prod_{i=1}^{n+2}\left(\frac{\varepsilon c_i}{b_i}\right)^{b_i}-1,$$

provided $\mathrm{sign}(b_1c_1)=\cdots=\mathrm{sign}(b_{n+2}c_{n+2})\in\{\pm 1\}$. $\diamond$

Note that by adding 1 and taking logs, it is clear that deciding the sign of $\Xi_{\mathcal{A}}$ is equivalent to deciding the sign of a linear combination of logarithms.

Determining isotopy types of positive zero sets of sparse polynomials was considered in a piecewise-linear (a.k.a. *tropical*) way starting with work of Viro in the 1980s [Vir84]. However, one difficulty with this approach is that combinatorics alone can not determine a unique piecewise linear hypersurface isotopic to $Z_+(f)$. This is where the sign of the discriminant enters and, in particular, the efficient approximation of linear forms in logarithms. Baker's Theorem [Bak77] gives explicit bounds on how well we need to approximate the underlying logarithms to decide sign and, through our techniques, his bound is just good enough to yield our new complexity bounds.

Another difficulty with approaching isotopy type by purely tropical means is that such a description becomes unwieldy in higher dimensions, and omits the singular case. This is where Morse Theory enters: In the setting of real exponential sums g with spectrum a circuit, we show how to efficiently find a quadratic polynomial whose real zero set is isotopic to $Z_{\mathbb{R}}(g)$ via a reasonable linear algebra calculation, after the underlying discriminant sign is computed.

So the key ingredients in the proof of our main results are Baker's Theorem, Morse Theory, and a recent metric estimate for tropical varieties from [EPR21]. The latter estimate enables us to extend Morse Theory (for compact manifolds) to our setting where we deal with non-compact stratified manifolds, and do so more succinctly than applying stratified Morse Theory.

References

[Bak77] Alan Baker, *"The Theory of Linear Forms in Logarithms,"* in Transcendence Theory: Advances and Applications: proceedings of a conference held at the University of Cambridge, Cambridge, Jan.–Feb., 1976, Academic Press, London, 1977.

[BR14] Saugata Basu and Marie-Françoise Roy, *"Divide and conquer roadmap for algebraic sets,"* Discrete and Computational Geometry, 52:278–343, 2014.

[BV07] Saugata Basu and Nicolai N. Vorobjov, *"On the Number of Homotopy Types of Fibres of a Definable Map,"* Journal of the London Mathematical Society, vol. 76, part 3, 2007, pp. 757–776.

[Bih11] Frederic Bihan, *"Topologie des variétés creuses,"* Habilitation à Diriger des Recherches, Universitde Savoie, France, 2011. Downloadable from `http://www.lama.univ-savoie.fr/ bihan/MyPapers/HDRfinal.pdf`

[BRS09] Frederic Bihan, J. Maurice Rojas, and Casey E. Stella, *"Faster real feasibility via circuit discriminants,"* proceedings of ISSAC 2009 (July 28-31, Seoul, Korea), pp. 39–46, ACM Press, 2009.

[EPR21] Alperen Ergür, Grigoris Paouris, and J. Maurice Rojas, *"Tropical Varieties for Exponential Sums,"* Mathematische Annalen (2020) 377:863-882.

[GVZ04] Andrei Gabrielov, Nikolai N. Vorobjov; and Thierry Zell, *"Betti Numbers of Semialgebraic and Sub-Pfaffian Sets,"* J. London Math. Soc. **69** (2004), pp. 27–43.

[GKZ94] Israel Moseyevitch Gel'fand, Misha M. Kapranov, and Andrei V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants,* Birkhäuser, Boston, 1994.

[Kho91] Askold Khovanski, *Fewnomials,* AMS Press, Providence, Rhode Island, 1991.

[Mil69] John Milnor, *Morse Theory,* Annals of Mathematics studies, no. 51, Princeton University Press, 1969.

[Vir84] Oleg Ya. Viro, *"Gluing of Plane Real Algebraic Curves and Constructions of Curves of Degrees* 6 *and* 7*,"* Topology (Leningrad, 1982), pp. 187–200, Lecture Notes in Math., 1060, Springer, Berlin, 1984.