EFFICIENT COMPUTATION OF REEB SPACES AND FIRST HOMOLOGY GROUPS

by

Sarah Percival

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Mathematics

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Saugata Basu, Chair

Department of Mathematics

Dr. Tamal Dey

Department of Computer Science

Dr. Andrei Gabrielov

Department of Mathematics

Dr. Elena Grigorescu

Department of Computer Science

Approved by:

Dr. Plamen Stefanov

To my family

ACKNOWLEDGMENTS

First and foremost, I want to thank my family for always believing in me. Without their support, this thesis would not be possible.

I would also like to thank my advisor, Professor Saugata Basu, for his guidance over the past five years, and thank Nathanael Cox and Negin Karisani for being great collaborators and friends.

Thank you to Tyler Billingsley, Frankie Chan, Kyle Dahlin, Vianney Filós-González, Joan Ponce, Van Vo, Ellen Weld, and Katy Yochman for making the math department infinitely more enjoyable.

Finally, I would like to thank Amy Brandfonbrener, Samira Fatemi, Lauren Lahey, Yvonne Muller, and Emily Rosenthal for reminding me there's a world outside of math.

TABLE OF CONTENTS

ABSTRACT 6

1 INTRODUCTION 7

1.1 Computational topology 7

1.2 Semi-algebraic geometry 7

1.2.1 Algorithmic semi-algebraic geometry 8

1.3 Computing A Basis for Higher Homology Groups 8

1.4 Reeb Spaces 10

2 PRELIMINARIES 12

2.1 Algorithmic Preliminaries 16

3 COMPUTATION OF A BASIS OF THE FIRST HOMOLOGY GROUP OF A SEMI-ALGEBRAIC SET 25

4 REEB SPACES OF SEMI-ALGEBRAIC MAPS 44

4.1 Reeb Spaces as Semi-algebraic Quotients 47

4.2 A Bound on the Topological Complexity of Reeb Spaces 50

4.2.1 Outline of the proof of Theorem 4.2.1 50

4.2.2 Technical necessities 52

4.2.3 Proof of Theorem 4.2.1 54

5 AN EFFICIENT ALGORITHM FOR THE COMPUTATION OF REEB SPACES FROM ROADMAPS 57

5.1 Example 58

5.2 Algorithm to Compute the Reeb Graph 59

REFERENCES 67

ABSTRACT

This thesis studies problems in computational topology through the lens of semi-algebraic geometry. We first give an algorithm for computing a semi-algebraic basis for the first homology group, $\mathrm{H}_1(S, \mathbb{F})$, with coefficients in a field $\mathbb{F}$, of any given semi-algebraic set $S \subset \mathrm{R}^k$ defined by a closed formula. The complexity of the algorithm is bounded singly exponentially. More precisely, if the given quantifier-free formula involves s polynomials whose degrees are bounded by d, the complexity of the algorithm is bounded by $(sd)^{k^{O(1)}}$. This algorithm generalizes well known algorithms having singly exponential complexity for computing a semi-algebraic basis of the zero-th homology group of semi-algebraic sets, which is equivalent to the problem of computing a set of points meeting every semi-algebraically connected component of the given semi-algebraic set at a unique point.

We then turn our attention to the Reeb graph, a tool from Morse theory which has recently found use in applied topology due to its ability to track the changes in connectivity of level sets of a function. The roadmap of a set, a construction that arises in semi-algebraic geometry, is a one-dimensional set that encodes information about the connected components of a set. In this thesis, we show that the Reeb graph and, more generally, the Reeb space, of a semi-algebraic set is homeomorphic to a semi-algebraic set, which opens up the algorithmic problem of computing a semi-algebraic description of the Reeb graph. We present an algorithm with singly-exponential complexity that realizes the Reeb graph of a function $f\colon X \to Y$ as a semi-algebraic quotient using the roadmap of X with respect to f.

1. INTRODUCTION

1.1 Computational topology

The field of applied topology was born out of a desire to better understand the underlying structure of real-world data by using ideas, such as homology, from the mathematical field of algebraic topology. As collections of data grow in size, it is increasingly important to have a way to efficiently analyze data. Applied topology addresses this problem by combining theoretical advances in computational topology with practical applications in topological data analysis. In particular, this thesis explores two problems in computational topology: efficient computation of Reeb spaces, which have gained recent attention in applied topology, and the classical topological problem of computing a basis for the first homology group of a space. In order to obtain meaningful results, we work within the category of semi-algebraic sets and maps over a real closed field R (see Chapter 2 for an overview of semi-algebraic geometry).

1.2 Semi-algebraic geometry

Because the goal of applied topology is to study real-world data, we choose to work within the framework of semi-algebraic geometry since the class of semi-algebraic sets is large enough that semi-algebraic sets can approximate any set that will occur in applications. We fix a real closed field R, and denote by $\mathrm{D} \subset \mathrm{R}$ a fixed ordered domain. We will denote by $\mathrm{C} = \mathrm{R}[i]$ the algebraic closure of R. For example, one can take $\mathrm{R} = \mathbb{R}$, and $\mathrm{D} = \mathbb{Z}$. Semi-algebraic sets are the subsets of R^n that are defined by a finite number of polynomial equations ($P = 0$) and inequalities ($P > 0$), and semi-algebraic maps are maps whose graphs are semi-algebraic sets. Given any finite family of polynomials $\mathcal{P} \in \mathrm{R}[X_1, \ldots, X_k]$, we call a Boolean formula without negations and with atoms $P\{\geq, \leq\}0$, $P \in \mathcal{P}$, a $\mathcal{P}$-*closed* formula, and we call the realization, $\mathcal{R}(\Phi, \mathrm{R}^k)$, a $\mathcal{P}$-*closed* semi-algebraic set. This framework allows us to make use of prior results in semi-algebraic geometry related to algorithmic complexity, as defined below.

1.2.1 Algorithmic semi-algebraic geometry

The field of semi-algebraic geometry arose from the problem of counting the real roots of a real univariate polynomial. Whether there exist real solutions to a finite set of polynomial equations and inequalities is decidable, and equivalent to the statement "the projection of a semi-algebraic set is semi-algebraic" [1]. There are several models of computation that can be considered when working with semi-algebraic sets (and also several notions of what constitutes an algorithm). If the real closed field $\mathrm{R} = \mathbb{R}$, and $\mathrm{D} = \mathbb{Z}$, one can consider these algorithmic problems in the classical Turing model and measure the bit complexity of the algorithms. In this paper, we will follow the book [1] and take a more general approach valid over arbitrary real closed fields. In the particular case, when $\mathrm{R} = \mathbb{R}$, our method will yield bit-complexity bounds.

It is an ongoing field of research to construct "efficient" algorithms, that is, algorithms with singly-exponential complexity, in order to answer mathematical questions. The complexity of an algorithm refers to the the the supremum over all inputs of the number of arithmetic operations performed by the algorithm, measured by the number of number of polynomials, their degrees, and the number of variables. Many algorithms in semi-algebraic geometry make use of the techniques of triangulation and cylindrical algebraic decomposition (CAD), both of which have doubly exponential complexity. In our paper, we bypass these techniques to construct efficient algorithms to compute topological invariants in special cases.

1.3 Computing A Basis for Higher Homology Groups

Given the relation between the upper bounds on topological complexity of a set and the complexity of algorithms computing their invariants, we can expect that there should exist algorithms for computing the Betti numbers of semi-algebraic sets with complexity bounded singly exponentially. Indeed, algorithms for computing the zero-th Betti number have been investigated in depth, and nearly optimal algorithms are known for this problem [2], [3], [4], [5]. We build on these results in Chapter 5, where present an efficient algorithm that takes as input a space X and returns a semi-algebraic basis for the first homology group, $\mathrm{H}_1(S)$.

To avoid computationally infeasible doubly-exponential complexity, our result relies on previous work on finding a basis for the 0-dimensional homology. That is, given a description of a semi-algebraic set, it is possible to compute a finite set of points that mean each connected component of the set exactly once. The solution to this problem is in two steps. The first, more simple step, is to compute a finite subset $\Gamma \subset S$ of "sample points" that meet each connected component of S. This is equivalent to finding a set Γ such that $\mathrm{H}_0(S, \Gamma) = 0$. The second step is to select a subset of the points in Γ so that there is only one point meeting each connected component of S. This is equivalent to finding a set $\{\mathbf{x}_i\} \subset \Gamma$ such that $\mathrm{H}_0(\{\mathbf{x}_i\}) \cong \mathrm{H}_0(S)$.

Here, we extend this result to H_1: given a semi-algebraic set S, we produce an algorithm that outputs semi-algebraic cycles that form a basis of $\mathrm{H}_1(S)$. As before, we first compute a one-dimensional subset Γ of S such that $\mathrm{H}_1(\Gamma) \twoheadrightarrow \mathrm{H}_1(S)$, or equivalently, $\mathrm{H}_1(S, \Gamma) = 0$. The cycles in Γ span $\mathrm{H}_1(S)$, but may not be a minimal spanning set, so we use a result in [6] to select cycles in Γ to form a basis of $\mathrm{H}_1(S)$. We construct an algorithm that proves the following theorem:

Theorem. *[7] There exists an algorithm that takes as input a finite set*

$$\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k],$$

and a $\mathcal{P}$-closed formula Φ, where $s = \mathrm{card}(\mathcal{P})$ and the maximum of the degrees of the polynomials in $\mathcal{P}$ and $f_1, \ldots, f_m$ is bounded by d,and outputs a finite set $\mathcal{Q} \subset \mathrm{D}[X_1, \ldots, X_k]$, as well as a finite tuple $(\Psi_j)_{j \in J}$, in which each Ψ_j is a $\mathcal{Q}$-formula, such that the realizations $\Gamma_j = \mathcal{R}(\Psi_j, \mathrm{R}^k)$ have the following properties:

1. *For each $j \in J$, $\Gamma_j \subset S$ and Γ_j is semi-algebraically homeomorphic to $\mathbf{S}^1$;*
2. *the inclusion map $\Gamma_j \hookrightarrow S$ induces an injective map $\mathbb{F} \cong \mathrm{H}_1(\Gamma_j) \to \mathrm{H}_1(S)$, whose image we denote by $[\Gamma_j]$;*
3. *the tuple $([\Gamma_j])_{j \in J}$ forms a basis of $\mathrm{H}_1(S)$.*

The complexity of each $\Gamma_j, j \in J$ (as a semi-algebraic subset of R^k defined over D), as well as the complexity of this algorithm, are bounded by $(sd)^{k^{O(1)}}$.

1.4 Reeb Spaces

The Reeb space of a continuous function $f\colon X \to Y$, where X and Y are topological spaces, is defined to be the set $X/\sim$, where $x \sim x' \in X$ if and only if $f(x) = f(x')$ and x and x' are in the same connected component of $f^{-1}(f(x)) = f^{-1}(f(x')) \in X$. To emphasize the dependence of the Reeb space on the choice of function $f\colon X \to Y$, we denote the Reeb space of a function by $\mathrm{Reeb}(f)$. When restricted to the special case where $f\colon X \to \mathbb{R}$, we obtain the Reeb graph of f. In order to ensure that the Reeb graph is a semi-algebraic set, we consider the case where X and Y are semi-algebraic sets and f is a proper semi-algebraic function and think of the Reeb space as a semi-algebraic set. Originally created as a tool in Morse theory [8], Reeb spaces have gained attention in applied topology [9], [10], [11] due to their ability to capture the underlying topological properties of a space. Burlet and de Rham first introduced the Reeb space in [12] as the *Stein factorization* of a map f, but their work was limited to bivariate, generic, smooth mappings. The authors of [13] defined the Reeb space of a multivariate piecewise linear mapping on a combinatorial manifold, and proved results regarding its local and global structure.

Reeb spaces, and in particular Reeb graphs, have previously been used to simplify and study data sets. The authors of [14] observe that, since shape data of natural objects is often available in cross-sections, Reeb graphs can be used to reconstruct the original three-dimensional surface. Furthermore, Mapper, an approximation of the Reeb graph introduced in [15], can be used to visualize the relationship between clusters of points in point-cloud data. The authors of [16] define the *interleaving distance* for Reeb spaces to show the convergence between the Reeb space and Mapper. Mapper has been used in a diverse range of applications, including the analysis breast cancer tumor expression [17]. Although there are fewer applications of higher dimensional Reeb spaces, the authors of [18] believe that Reeb spaces have shown promise in interpreting the output of standard statistical machine learning methods. These varied applications illustrate the need for an efficient method to compute Reeb spaces, which Chapter 3 provides.

While progress has been made in computing Reeb graphs [19] and Reeb spaces of certain maps [13], [20], little has been made in computing Reeb spaces of semi-algebraic maps. In

[21] it was shown that the Reeb space of a proper semi-algebraic map is homeomorphic to a semi-algebraic set, and furthermore that there is a singly exponential upper bound on the Betti numbers of the Reeb space of a semi-algebraic map $f\colon X \to Y$, where X is a closed and bounded semi-algebraic set, in terms of the number and the degrees of the polynomials defining X, Y and f.

Theorem. *[21] Let $S \subset \mathrm{R}^n$ be a bounded $\mathcal{P}$-closed semi-algebraic set, and $f = (f_1, \ldots, f_m) : S \to \mathrm{R}^m$ be a polynomial map. Suppose that $s = \mathrm{card}(\mathcal{P})$ and the maximum of the degrees of the polynomials in $\mathcal{P}$ and $f_1, \ldots, f_m$ is bounded by d. Then,*

$$b(\mathrm{Reeb}(f)) \leq (sd)^{(n+m)^{O(1)}}.$$

It is a meta theorem in algorithmic semi-algebraic geometry that upper bounds on topological complexity of objects are closely related to the worst-case complexity of algorithms computing the topological invariants of such objects. As a consequence of Theorem 4.2.1, we can expect to obtain an algorithm with singly-exponential complexity that computes the Reeb space of a semi-algebraic map. To that end, we will prove the following theorem:

Theorem. *There is an algorithm that takes as input a family of polynomials $\mathcal{P}$ and formulas describing a semi-algebraic set and polynomial map f and computes as output a semi-algebraic description of the Reeb graph with complexity $s^{k+1}d^{O(k^2)}$, where $s = \mathrm{card}(\mathcal{P})$ and the maximum of the degrees of the polynomials in $\mathcal{P}$.*

2. PRELIMINARIES

We begin by defining concept of *semi-algebraic sets*, which play a critical role throughout this thesis.

Definition 2.0.1 (Semi-algebraic Sets). *Let* R *be a real closed field. We define the* semi-algebraic sets *of* R^k *as the smallest family of sets in* R^k *that contains the sets of the form* $\{x \in \mathrm{R}^k \mid P(x) = 0\}$ *and* $\{x \in \mathrm{R}^k \mid P(x) > 0\}$ *for some polynomial* $P \in \mathrm{R}[X_1, \ldots, X_k]$ *and is closed under finite unions, finite intersections, and complementation.*

A map $f\colon X \to Y$ *between two semi-algebraic sets is is* semi-algebraic *if its graph is a semi-algebraic set.*

In this thesis, we assume that X and Y are semi-algebraic sets, and $f\colon X \to Y$ is a proper semi-algebraic map. We require f to be proper to ensure that $\mathrm{Reeb}(f)$ is a semi-algebraic set. In fact, we can assume without loss of generality that the map f is projection onto the first coordinate. To see that the choice of projection map does not reduce generality, consider a semi-algebraic set $S \in \mathrm{R}^k$ and let $f\colon S \to \mathrm{R}$ be any semi-algebraic map. Then we can add a new variable y and consider the set $S' = \{(x, y) \in \mathrm{R}^k \times \mathrm{R} \mid x \in S, y = f(x)\}$. Then $f(S)$ is the projection of S' onto R.

For computational efficiency, we represent points in R^k by univariate representations and an associated Thom encoding. To define these concepts, we must first define the notion of a *sign condition.*

Definition 2.0.2 (Sign Condition). *Let* $\mathcal{Q}$ *be a finite subset of* $\mathrm{R}[X_1, \ldots, X_k]$. *A* sign condition *on* $\mathcal{Q}$ *is an element of* $\{0, 1, -1\}^{\mathcal{Q}}$, *i.e. a mapping from* $\mathcal{Q}$ *to* $\{0, 1, -1\}$. *We say that* $\mathcal{Q}$ realizes *the sign condition* σ *if* $\bigwedge_{Q \in \mathcal{Q}} \mathbf{sign}(Q(x)) = \sigma(Q)$.

The realization of the sign condition σ *is*

$$\mathrm{Reali}(\sigma) = \{x \in \mathrm{R}^k \mid \bigwedge_{Q \in \mathcal{Q}} \mathbf{sign}(Q(x)) = \sigma(Q)\}$$

The sign condition σ *is* realizable *if* $\mathrm{Reali}(\sigma)$ *is non-empty.*

In particular, for our algorithms, we need to consider triangular Thom encodings. We first introduce some notation:

Notation 1. *Consider a finite subset* $\mathcal{P} \subset \mathrm{C}[X_1, \ldots, X_k]$, *where* C *is an algebraically closed field. We write the* set of zeros of $\mathcal{P}$ in C^k *as*

$$\mathrm{Zer}(\mathcal{P}, \mathrm{C}^k) = \{x \in \mathrm{C}^k \mid \bigwedge_{P \in \mathcal{P}} P(x) = 0\}$$

Notation 2. *Let* P *be a univariate polynomial of degree* p *in* $\mathrm{R}[X]$. *We denote by* $\mathrm{Der}(P)$ *the list* $P, P', \ldots, P^{(p)}$

Definition 2.0.3 (Thom Encoding)**.** *Let* $P \in \mathrm{R}[X]$ *and* $\sigma \in \{0, 1, -1\}^{\mathrm{Der}(P)}$. *The sign condition* σ *is a* Thom encoding *of* $x \in \mathrm{R}$ *if* $\sigma(P) = 0$ *and* $\mathrm{Reali}(\sigma) = \{x\}$.

Definition 2.0.4 (Triangular Thom Encoding and Associated Point)**.** *A* triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ *of size* i *is a tuple (triangular system) of polynomials,*

$$\mathbf{F} = (f_1, \ldots, f_i)$$

where $f_j \in \mathrm{R}[X_1, \ldots, X_j], 1 \leq j \leq i$, *and a tuple of Thom encodings* $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_i)$, *with* $\sigma_j \in \{0, 1, -1\}^{\mathrm{Der}_{X_j}(f_j)}$, *such that for each* $j, 1 \leq j \leq i$, *there exists* $t_j \in \mathrm{R}$, *such that* t_j *is a root of the polynomial* $f_j(t_1, \ldots, t_{j-1}, T_j)$ *with Thom encoding* σ_j. *We call* $(t_1, \ldots, t_i) \in \mathrm{R}^i$ *the* point associated to $\mathcal{T}$ *and denote* $\mathrm{ass}(\mathcal{T}) = (t_1, \ldots, t_i)$.

Given a triangular Thom encoding

$$\mathcal{T}^+ = ((f_1, \ldots, f_{i+1}), (\sigma_1, \ldots, \sigma_{i+1})),$$

with $\mathrm{ass}(\mathcal{T}^+) = (t_1, \ldots, t_{i+1})$, *we will sometimes call the pair* $\tau = (f_{i+1}, \sigma_{i+1})$ *a* Thom encoding over the triangular Thom encoding $\mathcal{T} = ((f_1, \ldots, f_i), (\sigma_1, \ldots, \sigma_i))$. *In this case we will denote* t_{i+1} *by* $\mathrm{ass}(\tau)$ *(generalizing Definition 2.0.3).*

We will call $\max_{1 \leq j \leq i} \deg(F_j)$ *the degree of the triangular Thom encoding* $\mathcal{T}$, *and denote it by* $\deg(\mathcal{T})$.

If $\tau = (f_{i+1}, \sigma_{i+1})$ *a is Thom encoding over a triangular Thom encoding*

$$\mathcal{T} = ((f_1, \ldots, f_i), (\sigma_1, \ldots, \sigma_i)),$$

we will call $\deg_{T_{i+1}}(f_{i+1})$*, the degree of* τ*, and denote it by* $\deg(\tau)$*.*

Finally, given a triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ *of size* i*, we denote by* $\theta(\mathcal{T})$*, the formula*

$$\bigwedge_{1 \le j \le i} \; \bigwedge_{0 \le h \le \deg_{X_j}(f_j)} \left(\mathbf{sign}(f_j^{(h)}) = \sigma_j(f_j^{(h)})\right).$$

Notation 3. *Given a triangular Thom encoding* $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ *of size* i *(following the same notation as in Definition 2.0.4 above), we will denote by* $\overline{\boldsymbol{\sigma}}$*, the closed formula obtained from* $\boldsymbol{\sigma}$ *by replacing each sign condition on the derivatives by the corresponding weak inequality (i.e. replacing* $f_j^{(h)} > 0$ *by* $f_j^{(h)} \geq 0$ *and* $f_j^{(h)} < 0$ *by* $f_j^{(h)} \leq 0$*).*

We will make use of the following consequence of Thom's lemma[1]:

Lemma 1. *Given a triangular Thom encoding* $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ *of size* i*,*

$$\operatorname{ass}(\mathcal{T}) = \mathcal{R}(\overline{\boldsymbol{\sigma}}, \mathrm{R}^i).$$

Proof. Follows directly from [1, Proposition 5.39] (Generalized Thom's Lemma). □

We now introduce the notion of a *univariate representation* to efficiently describe the coordinates of the solutions of a zero-dimensional polynomial system as rational functions of the roots of a univariate polynomial.

Definition 2.0.5 (Univariate Representation)**.** *Let* K *be a field. A* k-univariate representation *is a* $k+2$*-tuple of polynomials in* $\mathrm{K}[T]$*,*

$$u = (f(T), g(T)), \text{ with } g = (g_0(T), g_1(T), \ldots, g_k(T))$$

such that f *and* g_0 *are coprime.*

A real k-univariate representation *is a pair* u, σ *where* u *is a* k*-univariate representation and* σ *is the Thom encoding of a root of* $r, t_\sigma \in \mathrm{R}$*.*

A real univariate triangular representation $\mathcal{T}, \sigma, u$ *of level* $i-1$ *consists of*

- *a triangular Thom encoding* $\mathcal{T}, \sigma$ *specifying* $(z,t) \in \mathrm{R}^i$ *with* $z \in \mathrm{R}^{i-1}$
- *a parametrized univariate representation*

$$u(X_{<i}) = (\mathcal{T}_i(X_{<i}, T), g_0(X_{<i}, T), g_i(X_{<i}, T), \ldots, g_k(X_{<i}, T)$$

with parameters $X_{<i} = (X_1, \ldots, x_{i-1})$.

The point associated to $\mathcal{T}, \sigma, u$ *is*

$$\left(z, \frac{g_i(z,t)}{g_0(z,t)}, \ldots, \frac{g_k(z,t)}{g_0(z,t)}\right)$$

A curve segment representation u, ρ above $\mathcal{V}_1 \tau_1, \mathcal{V}_2, \tau_2$ *consists of*

- *a parametrized univariate representation with parameters* $(x_{\leq i})$

$$u = (f(X_{\leq i}, T), g_0(X_{\leq i}, T), g_{i+1}(X_{\leq i}, T), \ldots, g_k(X_{\leq i}, T),$$

- *a sign condition* ρ *on* $\mathrm{Der}(f)$ *such that for every* $v \in (a,b)$ *there exists a real root* $t(V)$ *of* $f(z,v,T)$ *with Thom encoding* ρ, σ *and* $g_0(z,v,t(v)) \neq 0$.

We need the following notation:

Notation 4. *For* R *a real closed field we denote by* $\mathrm{R}\langle\varepsilon\rangle$ *the real closed field of algebraic Puiseux series in* ε *with coefficients in* R*. As a real closed field* $\mathrm{R}\langle\varepsilon\rangle$ *is uniquely ordered, and this order extends the order on* R*. It is the unique order in which* $\varepsilon > 0$ *and* $\varepsilon < x$ *for every* $x \in \mathrm{R}, x > 0$*. In particular, the subring* $\mathrm{D}[\varepsilon] \subset \mathrm{R}\langle\varepsilon\rangle$ *is ordered by:*

$$\sum_{i \geq 0} a_i \varepsilon^i > 0 \textit{ if and only if } a_p > 0 \textit{ where } p = \min\{i \mid a_i \neq 0\}$$

Notation 5. *For* $P = a_p T^p + \cdots + a_q T^q, p \geq q \in \mathrm{D}[T]$*, we denote*

$$c'(P) = \left((p+1) \cdot \sum_i \frac{a_i^2}{a_q^2}\right)^{-1}.$$

In Chapter 4, we make use of a tool from semi-algebraic geometry called the *roadmap* [22], which provides a useful description of the connected components of a set.

Definition 2.0.6 (Roadmap)**.** *A* roadmap *for S is a semi-algebraic set M of dimension at most one contained in S which satisfies the following roadmap conditions:*

- *RM_1: For every semi-algebraically connected component D of S, $D \cap M$ is semi-algebraically connected.*

- *RM_2: For every $x \in R$ and for every semi-algebraically connected component D' of $S_x = \{y \in \mathrm{R}^{k-1} \mid (x, y) \in S\}$, $D' \cap M \neq \emptyset$.*

We conclude this section with the following defintion:

Definition 2.0.7 (Semi-algebraic Description)**.** *Given a pair (S, s) where S is a semi-algebraic set and $s \colon S \to \mathrm{R}$, a* semi-algebraic description *of (S, s) is a set D homeomorphic to S and a map $p \colon D \to \mathrm{R}$ such that the following diagram commutes:*

$$\begin{tikzcd} S \arrow[r, "\sim"] \arrow[d, "s"] & D \arrow[dl, "p"] \\ \mathrm{R} & \end{tikzcd}$$

2.1 Algorithmic Preliminaries

Before we introduce our algorithmic contributions, we must first discuss previous contributions to algorithmic semi-algebraic geometry upon which we build our algorithm. We rely the roadmap algorithm detailed in [1], reprinted here for convenience.

Algorithm 1 (Roadmap of a Semi-Algebraic Set)

Input:

(a) a polynomial $Q \in D[X_1, \ldots, X_k]$ such that $\mathrm{Zer}(Q, \mathrm{R}^k)$ is of real dimension k',

(b) a semi-algebraic subset S of $\mathrm{Zer}(Q, \mathrm{R}^k)$ described by a finite set $\mathcal{P} \subset R[X_1, \ldots, X_k]$.

Output: a roadmap for S.

Complexity: $s^{k'+1} d^{O(k^2)}$ where s is a bound on the number of elements of $\mathcal{P}$, and d is a bound on the degree of Q and of the polynomials in $\mathcal{P}$.

We also include a semi-algebraic generalization of the Curve Selection Algorithm in [1], which relies on the following algorithm, appearing (without a triangular Thom encoding in the input) in [6].

Algorithm 2 (Morse partition)

Input:

(a) $r \in \mathrm{D}, r > 0$;

(b) A triangular Thom encoding $\mathcal{T}$ of size $i, 0 \leq i \leq k$;

(c) A finite set $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{D}[X_1, \ldots, X_k]$;

(d) A $\mathcal{P}$-closed formula Φ.

Output:

A finite set of $\mathcal{F}$ of Thom encodings over $\mathcal{T}$, with associated points $t_1 < \cdots < t_N$, with $-r \leq t_1, t_N \leq r$, and such that for each $j, 1 \leq j \leq N-1$, and all $t \in [t_j, t_{j+1})$ the inclusion maps

$$S_{\{\mathrm{ass}(\mathcal{T})\} \times (\infty, t_i]} \hookrightarrow S_{\{\mathrm{ass}(\mathcal{T})\} \times (\infty, t]}$$

induce isomorphisms

$$\mathrm{H}_*(S_{\{\mathrm{ass}(\mathcal{T})\} \times (\infty, t_i]}) \to \mathrm{H}_*(S_{\{\mathrm{ass}(\mathcal{T})\} \times (\infty, t]}),$$

where

$$S = \mathcal{R}(\Phi, \mathrm{R}^k) \cap \mathrm{clos}(B_k(\mathbf{0}, r)).$$

Complexity: The complexity of the algorithm is bounded by

$$D^{O(i)}(sd)^{O(k)},$$

where $s = \mathrm{card}(\mathcal{P})$, $d = \max_{P \in \mathcal{P}} \deg(P)$, and $D = \deg(\mathcal{T})$.

Moreover, $\deg(\tau) \leq d^{O(k)}$ for each $\tau \in \mathcal{F}$, and $\mathrm{card}(\mathcal{F}) \leq (sd)^{O(k)}$.

We will need the following extra property of the output of Algorithm 2.

Proposition 2.1.1. *For each $j, 1 \leq j \leq N-1$, and all $t \in [t_j, t_{j+1}]$ the inclusion maps*

$$S_{\{\mathrm{ass}(\mathcal{T})\}\times\{t\}} \hookrightarrow S_{\{\mathrm{ass}(\mathcal{T})\}\times[t_j t_{j+1}]}$$

induce isomorphisms

$$\mathrm{H}_*(S_{\{\mathrm{ass}(\mathcal{T})\}\times\{t\}}) \to \mathrm{H}_*(S_{\{\mathrm{ass}(\mathcal{T})\}\times[t_j t_{j+1}]}),$$

where

$$S = \mathcal{R}(\Phi, \mathrm{R}^k) \cap \mathrm{clos}(B_k(\mathbf{0}, r)).$$

Proof. Let

$$\begin{aligned}
A_1 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,t]},\\
A_2 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times[t,\infty)},\\
B_1 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,t_{j+1}]},\\
B_2 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times[t_j,\infty)}.
\end{aligned}$$

Then, $A_h \subset B_h, h = 1, 2$ and

$$\begin{aligned}
A_1 \cap A_2 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times\{t\}},\\
B_1 \cap B_2 &= S_{\{\mathrm{ass}(\mathcal{T})\}\times[t_j t_{j+1}]},
\end{aligned}$$

and

$$A_1 \cup A_2 = B_1 \cup B_2 = S_{\mathrm{ass}(\mathcal{T})}.$$

Moreover, the properties of the output of Algorithm 2 imply that that the homomorphisms $\mathrm{H}_*(A_h) \to \mathrm{H}_*(B_h)$, $h = 1, 2$, induced by inclusions are isomorphisms. The Mayer-Vietoris exact sequence in homology then yields the following commutative diagram with exact rows and vertical homomorphisms induced by inclusion (where A_{12} (resp. B_{12}) denotes $A_1 \cap A_2$ (resp. $B_1 \cap B_2$), and A^{12} (resp. B^{12}) denotes $A_1 \cup A_2$ (resp. $B_1 \cup B_2$)):

$$\begin{array}{ccccccccc}
\mathrm{H}_{m+1}(A_1)\oplus \mathrm{H}_{m+1}(A_2) & \longrightarrow & \mathrm{H}_{m+1}(A^{12}) & \longrightarrow & \mathrm{H}_m(A_{12}) & \longrightarrow & \mathrm{H}_m(A_1)\oplus \mathrm{H}_m(A_2) & \longrightarrow & \mathrm{H}_m(A^{12}) \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\mathrm{H}_{m+1}(B_1)\oplus \mathrm{H}_{m+1}(B_2) & \longrightarrow & \mathrm{H}_{m+1}(B^{12}) & \longrightarrow & \mathrm{H}_m(B_{12}) & \longrightarrow & \mathrm{H}_m(B_1)\oplus \mathrm{H}_m(B_2) & \longrightarrow & \mathrm{H}_m(B^{12})
\end{array}$$

It follows from the fact that homomorphisms $\mathrm{H}_*(A_h) \to \mathrm{H}_*(B_h)$, $h = 1,2$ induced by inclusions are isomorphisms, and the fact that $A^{12} = B^{12}$, that all the vertical arrows other than the middle one are isomorphisms. Hence, by the "five-lemma" [23] the middle arrow is also an isomorphism, thus proving the proposition. □

Algorithm 4 requires the following construction construction as input:

Algorithm 3 (Big enough radius)

Input:

(a) A triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ of size $i, 0 \leq i \leq k$;

(b) a finite set $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$;

(c) a $\mathcal{P}$-closed formula Φ such that $\mathcal{R}(\Phi, \mathrm{R}^k)$ is bounded.

Output:

Elements $a, b \in \mathrm{D}[\mathrm{ass}(\mathcal{T})], a, b > 0$, such that the inclusion map

$$\mathcal{R}(\Phi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})} \cap \mathrm{clos}(\{\mathrm{ass}(\mathcal{T})\} \times B_{k-i}(\mathbf{0}, r)) \hookrightarrow \mathcal{R}(\Phi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})},$$

where $r = \frac{a}{b}$, induces an isomorphism

$$\mathrm{H}_*\left((\mathcal{R}(\Phi, \mathrm{R}^k) \cap \mathrm{clos}(\{\mathrm{ass}(\mathcal{T})\} \times B_{k-i}(\mathbf{0}, r)))_{\mathrm{ass}(\mathcal{T})}\right) \to \mathrm{H}_*\left(\mathcal{R}(\Phi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})}\right).$$

Procedure:

1: $P_1 \leftarrow Y - (X_{i+1}^2 + \cdots + X_k^2)$.

2: $P_2 \leftarrow (\varepsilon^2(X_{i+1}^2 + \cdots + X_k^2) - 1)$.

3: $\widetilde{\Phi} \leftarrow \Phi \wedge (P_1 = 0) \wedge (P_2 \leq 0)$.

4: $\widetilde{\mathrm{D}} \leftarrow \mathrm{D}[\varepsilon]$ where $D[\varepsilon] \subset \mathrm{R}\langle\varepsilon\rangle$ (cf. Definition 4).

5: Call Algorithm 2 (Morse partition) treating Y as the $(i+1)$-st coordinate, with computations occurring in the domain $\widetilde{\mathrm{D}}$, and with input $\mathcal{T}, \mathcal{P} \cup \{P_1, P_2\}, \widetilde{\Phi}$.

6: $\mathcal{Q} \in \mathrm{D}[\mathrm{ass}(\mathcal{T})][\varepsilon] \leftarrow$ the set of polynomials whose signs are determined during the call to Algorithm 2 (Morse partition) in the previous step.

7: $c = \frac{b}{a} \leftarrow \min_{Q \in \mathcal{Q}} c'(Q), a, b \in \mathrm{D}[\mathrm{ass}(\mathcal{T})]$ (cf. Notation 5).

8: $r \leftarrow \frac{a}{b}$.

9: Output a, b.

Complexity: The complexity of the algorithm is bounded by

$$D^{O(i)}(sd)^{O(k)},$$

where $s = \mathrm{card}(\mathcal{P})$, $d = \max_{P \in \mathcal{P}} \deg(P)$, and $D = \deg(\mathcal{T})$.

Proof of Correctness of Algorithm 3. Note that the formula $\widetilde{\Phi}$ defines a semi-algebraic subset

$$\widetilde{S} \subset \mathrm{R}\langle\epsilon\rangle^i \times \mathrm{clos}(B_{k-i}(\mathbf{0}, \frac{1}{\varepsilon})).$$

It follows from the conic structure theorem at infinity of semi-algebraic sets (see for example [1, Proposition 5.49]) and the Tarski-Seidenberg transfer principle (see for example [1, Theorem 2.98]) that $\mathrm{ext}(S, \mathrm{R}\langle\epsilon\rangle)_{\mathrm{ass}(\mathcal{T})}$ is semi-algebraically homeomorphic to $\widetilde{S}_{\mathrm{ass}(\mathcal{T})}$, and hence

$$\mathrm{H}_*(\mathrm{ext}(S, \mathrm{R}\langle\varepsilon\rangle)_{\mathrm{ass}(\mathcal{T})}) \cong \mathrm{H}_*(\widetilde{S}_{\mathrm{ass}(\mathcal{T})}). \tag{2.1}$$

It follows from the way r is computed in the algorithm in Line 7, and Lemma 10.7 in [1] that $c = \frac{1}{r}$ is strictly positive, and smaller than all roots in R of the polynomials in $\mathcal{Q}$ (defined in Line 6 of the algorithm). It now follows from the ordering of the ring $\mathrm{D}[\mathrm{ass}(\mathcal{T})][\varepsilon]$ (cf. Notation 4) that all the branchings in the algorithm, each of which depend on the determination of the sign of an element in $\mathrm{D}[\mathrm{ass}(\mathcal{T})][\varepsilon]$, remain the same if c is substituted for ε.

It now follows from the correctness of Algorithm 2 (Morse partition), that the inclusion

$$\widetilde{S}_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,r]} \hookrightarrow \widetilde{S}_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,\frac{1}{\varepsilon}]} = \widetilde{S}_{\mathrm{ass}(\mathcal{T})}$$

(with r as computed in the algorithm) induces an isomorphism

$$\mathrm{H}_*(\widetilde{S}_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,r]}) \to \mathrm{H}_*(\widetilde{S}_{\mathrm{ass}(\mathcal{T})}). \tag{2.2}$$

Moreover, for any $r' > 0$, $\widetilde{S}_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,r']}$ is semi-algebraically homeomorphic to $\mathrm{ext}(S, \mathrm{R}\langle\varepsilon\rangle) \cap \{\mathrm{ass}(\mathcal{T})\} \times \mathrm{clos}(B_{k-i}(\mathbf{0}, r'))$, and hence

$$\mathrm{H}_*(\widetilde{S}_{\{\mathrm{ass}(\mathcal{T})\}\times(-\infty,r]}) \cong \mathrm{H}_*(\mathrm{ext}(S, \mathrm{R}\langle\varepsilon\rangle) \cap \{\mathrm{ass}(\mathcal{T})\} \times \mathrm{clos}(B_{k-i}(\mathbf{0}, r))). \tag{2.3}$$

Finally, for any closed semi-algebraic set $X \subset \mathrm{R}^k$,

$$\mathrm{H}_*(X) \cong \mathrm{H}_*(\mathrm{ext}(X, \mathrm{R}\langle\varepsilon\rangle)). \tag{2.4}$$

The isomorphisms (2.1), (2.2), (2.3), and (2.4) imply that

$$\mathrm{H}_*(S_{\mathrm{ass}(\mathcal{T})}) \cong \mathrm{H}_*(S \cap \{\mathrm{ass}(\mathcal{T})\} \times \mathrm{clos}(B_{k-i}(\mathbf{0}, r))).$$

This proves the correctness of Algorithm 3. □

Complexity of Algorithm 3. The stated complexity follows from the complexity of Algorithm 2 (Morse partition). □

Algorithm 4 (Curve segments)

Input:

1. A triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ of size i with $0 \leq i \leq k-1$;
2. a finite set $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$;
3. a $\mathcal{P}$-closed formula Φ such that $\mathcal{R}(\Phi, \mathrm{R}^k)$ is bounded.

Output:

1. A finite tuple $\mathcal{F} = (\tau_1, \ldots, \tau_N)$ of Thom encodings over $\mathcal{T}$, with

$$t_1 = \mathrm{ass}(\tau_1) < \cdots < t_N = \mathrm{ass}(\tau_N);$$

2. for each $j, 1 \leq j \leq N-1$, an indexing set I_j, a finite tuple $\mathcal{C}_j = (\gamma_h)_{h \in I_j}$ of curve segment representations over $\mathcal{T}$, such that for each $h \in I_j$

$$\tau_1(\gamma_h) = \tau_j, \tau_2(\gamma_h) = \tau_{j+1}$$

(we will let $C_0 = \mathcal{C}_{N+1} = \emptyset$);

3. for each $j, 1 \leq j \leq N$ a finite set $\mathcal{U}_j$ of real univariate representations over $\mathcal{T}$, such that for each $u \in U_j$, the set of points $\{\mathrm{ass}(u) \mid u \in \mathcal{U}_j\}$ includes the set of end-points of the curve segment representations in $C_{j-1} \cup C_j$;

4. mappings $L_j, R_{j-1} : I_j \to \mathcal{U}_j$, defined by $\mathrm{ass}(L_j(h))$ is the left end-point of γ_h, and $\mathrm{ass}(R_j(\gamma_h))$ is the right end-point of γ_h.

Procedure:

1: Call Algorithm 3 with input $(\mathcal{T}, \mathcal{P}, \Phi)$ and compute $r > 0, r \in \mathrm{D}$.

2: $\mathcal{P} \leftarrow \mathcal{P} \cup \{X_1^2 + \cdots + X_k^2 - r^2\}$.

3: $\Phi \leftarrow \Phi \wedge (X_1^2 + \cdots + X_k^2 - r^2 \leq 0)$.

4: Call Algorithm 2 (Morse partition) with $(\mathcal{T}, r, \mathcal{P}, \Phi)$ as input and obtain a finite set $\mathcal{F}$ of Thom encodings over $\mathcal{T}$ as output.

5: Call Algorithm 16.26 (General Roadmap) on [1] with input $\mathcal{P}, \Phi$, performing all computations over the ring $\mathrm{D}[\mathrm{ass}(\mathcal{T})]$.

6: Retain from the output of the previous step, the set $\mathcal{C}$ of curve segment representations over $\mathcal{T}$ parametrized by X_{i+1}, and the set $\mathcal{U}$ of real univariate representations over $\mathcal{T}$.

7: For each pair $\gamma, \gamma' \in \mathcal{C}$, compute a description of $\mathrm{ass}(\gamma) \cap \mathrm{ass}(\gamma')$ using Algorithm 14.25 (Parametrized Sign Determination) in [1], and refine the set $\mathcal{C}$ to have the property that $\mathrm{ass}(\gamma) \cap \mathrm{ass}(\gamma') = \emptyset$, for all $\gamma, \gamma' \in \mathcal{C}\gamma \neq \gamma'$.

8: Augment the set $\mathcal{U}$ to also contain the set of real univariate representations over $\mathcal{T}$ whose associated points are the end points of the curve segments in $\mathcal{C}$.

9: Compute Thom encodings over $\mathcal{T}$ whose associated values are the X_{i+1}-coordinates of the asssociated points of $\mathcal{U}$, and add these to $\mathcal{F}$.

10: Using Algorithm 12.69 (Triangular Comparison of Roots) in [1] order the Thom encodings in $\mathcal{F}$, and let the associated values be $t_1 = \mathrm{ass}(\tau_1), \ldots, t_N = \mathrm{ass}(\tau_N)$.

$$\mathcal{F} \leftarrow (\tau_1, \ldots, \tau_N).$$

11: Further refine $\mathcal{C}$, such that for each $\gamma \in \mathcal{C}$, there exists $j, 1 \leq j < N$, such that

$$\tau_1(\gamma) = \tau_j, \tau_2(\gamma) = \tau_2.$$

12: Augment the set $\mathcal{U}$ to include the left and the right end points of each $\gamma \in \mathcal{C}$.

13: **for** each $j, 1 \leq j \leq N-1$ **do**

14: Let I_j denote a set indexing the set of curve segment representations $\gamma \in \mathcal{C}$, such that $\tau_1(\gamma) = \tau_j, \tau_2(\gamma) = \tau_{j+1}$. For $h \in I_j$, we denote the corresponding curve segment representation in $\mathcal{C}$ by γ_h.

15: $\mathcal{C}_j \leftarrow (\gamma_h)_{h \in I_j}$

16: $\mathcal{U}_j \leftarrow \{u \in \mathcal{U} \mid \pi_{i+1}(\mathrm{ass}(u)) = \mathrm{ass}(\tau_j)\}$.

17: Compute the maps $L_j : I_j \to \mathcal{U}_j, R_j : I_j \to \mathcal{U}_{j+1}$, such that $\mathrm{ass}(L_j(h))$ is the left end-point of γ_h, and $\mathrm{ass}(R_j(h))$ is the right endpoint of γ_h.

18: **end for**

19: Output $\Big(\mathcal{F}, (I_j, \mathcal{C}_j, \mathcal{U}_j, L_j, R_j)_{j \in [1,N]}\Big)$.

Complexity: The complexity of the algorithm is bounded by $D^{O(i)}(sd)^{k^{O(1)}}$, where $s = \mathrm{card}(\mathcal{P})$, $d = \max_{P \in \mathcal{P}} \deg(P)$, and $D = \deg(\mathcal{T})$.

The degrees of the curve segment representations in the various $\mathcal{C}_j$, and the degrees of the real univariate representations in $\mathcal{U}_j$ are both bounded by $(D, d^{O(k)})$. Finally the sum of the cardinalities,

$$\sum_{j=1}^{N} (\mathrm{card}(\mathcal{C}_j) + \mathrm{card}(\mathcal{U}_j))$$

is bounded by $(sd)^{k^{O(1)}}$.

Proposition 2.1.2. *The output of Algorithm 4 (Curve segments) satisfies the following:*

(a) For each j, $1 \leq j \leq N-1$, and all $x_{i+1} \in [\mathrm{ass}(\tau_j), \mathrm{ass}(\tau_{j+1}))$ the inclusion maps

$$S_{\{\mathrm{ass}(\mathcal{T})\} \times (-\infty, \mathrm{ass}(\tau_j)]} \hookrightarrow S_{\{\mathrm{ass}(\mathcal{T})\} \times (-\infty, x_{i+1}]}$$

are homological equivalences;

(b) for each $h \in I_j$, $\mathrm{ass}(\gamma_h) \subset S = \mathcal{R}(\Phi, \mathrm{R}^k)$;

(c) for each $x_{i+1} \in (\mathrm{ass}(\tau_j), \mathrm{ass}(\tau_{j+1}))$ and each semi-algebraically connected component C of $S_{\mathbf{y}}$, where $\mathbf{y} = (\mathrm{ass}(\mathcal{T}), x_{i+1}) \in \mathrm{R}^{i+1}$, there exists $h \in I_j$ such that $\mathrm{ass}(\gamma_h)_{\mathbf{y}} \in C$;

(d) if $h_1, h_2 \in I_j$ with $h_1 \neq h_2$, then $\mathrm{ass}(\gamma_{h_1}) \cap \mathrm{ass}(\gamma_{h_2}) = \emptyset$.

3. COMPUTATION OF A BASIS OF THE FIRST HOMOLOGY GROUP OF A SEMI-ALGEBRAIC SET

Given a quantifier-free formula Φ defining a semi-algebraic subset $S \subset \mathrm{R}^k$, the problem of computing a basis for $\mathrm{H}_0(S)$ has been well-studied, and very efficient algorithms for computing such a basis via roadmaps have been produced [24]–[28]. It is natural to ask if this result can be extended to higher homology groups. In this chapter, we present an efficient algorithm to compute a basis for $\mathrm{H}_1(S)$.

To achieve this result, we first construct a one-dimensional subset Γ of S such that there is a surjection $\mathrm{H}_1(\Gamma) \twoheadrightarrow \mathrm{H}_1(S)$. This Γ is semi-algebraically equivalent to the realization $|G|$ of a finite graph G that is singly exponential in size. It is then a relatively easy combinatorial task to choosing a basis of simple cycles, $\Gamma_1, \ldots, \Gamma_N$, for the cycle space of G. Letting $[\Gamma_i]$ denote the image of $\mathrm{H}_1(|\Gamma_i|)$ in $\mathrm{H}_1(S)$ under the homomorphism induced by the inclusion $|\Gamma_i| \hookrightarrow S$, the images $[\Gamma_1], \ldots, [\Gamma_N]$ span $\mathrm{H}_1(S)$ but are not necessarily linearly independent. To select a minimal spanning subset from amongst the $[\Gamma_1], \ldots, [\Gamma_N]$, we rely on an algorithm from [6] that replaces a given semi-algebraic set and a tuple of subsets by a simplicial complex and a tuple of corresponding subcomplexes, which are homologically ℓ-equivalent (cf. Definition 3.0.2) for any fixed ℓ, and which has singly exponentially bounded complexity [6] (cf. Algorithm 6 below).

The set Γ is very similar to the roadmap construction, however the roadmap requirements alone do not guarantee that $\mathrm{H}_1(\Gamma) \twoheadrightarrow \mathrm{H}_1(S)$. To see why the roadmap is insufficient for our purposes, consider the 2-torus, which has as its basis for H_1 a wedge of two copies of $\mathbf{S}^1$, shown below.

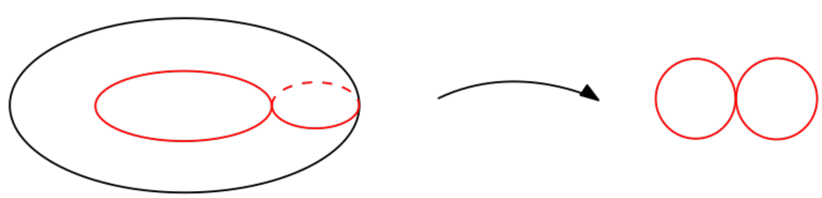


Figure 3.1. A basis for $\mathrm{H}_1(T^2)$

The outer equator E of a torus T, shown in 3.2a, is a roadmap for T. However, $H_1(D)$ does not surject onto $H_1(T)$. To remedy this problem, we present an algorithm (cf Algorithm 7) to construct a set Γ, shown in 3.2b, such that $H_1(\Gamma) \to H_1(S^1 \wedge S^1)$ is a surjection.

(a) The outer equator of T^2 (b) Our construction

Figure 3.2. Two roadmaps of T^2

Algorithm 7 takes as input a $\mathcal{P}$-closed formula Φ, and produces as output a description of a semi-algebraic subset $\Gamma \subset S = \mathcal{R}(\Phi)$, having dimension ≤ 1, and such that the homomorphism $i_{*,1} : \mathrm{H}_1(\Gamma) \to \mathrm{H}_1(S)$ induced by the inclusion $\Gamma \hookrightarrow S$ is surjective, and the homomorphism $i_{*,0} : \mathrm{H}_0(\Gamma) \to \mathrm{H}_0(S)$ is an isomorphism. More explicitly, we have the following theorem.

Theorem 3.0.1. *There exists an algorithm that takes as input a finite set*

$$\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k],$$

and a $\mathcal{P}$-closed formula Φ, where $s = \mathrm{card}(\mathcal{P})$ and the maximum of the degrees of the polynomials in $\mathcal{P}$ and $f_1, \ldots, f_m$ is bounded by d,and outputs a finite set $\mathcal{Q} \subset \mathrm{D}[X_1, \ldots, X_k]$, as well as a finite tuple $(\Psi_j)_{j \in J}$, in which each Ψ_j is a $\mathcal{Q}$-formula, such that the realizations $\Gamma_j = \mathcal{R}(\Psi_j, \mathrm{R}^k)$ have the following properties:

1. *For each $j \in J$, $\Gamma_j \subset S$ and Γ_j is semi-algebraically homeomorphic to $\mathbf{S}^1$;*
2. *the inclusion map $\Gamma_j \hookrightarrow S$ induces an injective map $\mathbb{F} \cong \mathrm{H}_1(\Gamma_j) \to \mathrm{H}_1(S)$, whose image we denote by $[\Gamma_j]$;*
3. *the tuple $([\Gamma_j])_{j \in J}$ forms a basis of $\mathrm{H}_1(S)$.*

The complexity of each $\Gamma_j, j \in J$ (as a semi-algebraic subset of R^k defined over D), as well as the complexity of this algorithm, are bounded by $(sd)^{k^{O(1)}}$.

We will prove Theorem 3.0.1 by describing an algorithm (cf. Algorithm 8 below) for computing a semi-algebraic basis of $\mathrm{H}_1(\mathcal{R}(\Phi, \mathrm{R}^k))$, for any given closed formula Φ. Theorem 3.0.1 will then follow from the proof of correctness this algorithm and the analysis of the complexity of the algorithm. Before presenting Algorithm 8, we need to address address the following technical issue: the output of Algorithm 7 (Computing one-dimensional subset) contains amongst other objects, a set of curve segment representations. In order to use Algorithm 6 which only accepts such descriptions by closed formulas in the input, we will need to convert these descriptions into closed formulas describing the closure of the associated curves

Note that the algorithmic problem of computing a *closed* formula describing the closure of a given semi-algebraic set described by a quantifier-free (but not necessarily closed-) formula is far from being easy, and no algorithm with singly exponential complexity is known for solving this problem in general. (A doubly exponential algorithm is known, using the notion of a stratifying family [1, Chapter 5]). However, the curve segment representations describing the associated curve have a special structure, namely, it is clear that given a curve segment representation γ, it is algorithmically quite simple to obtain a description of the the image, $\pi_{\{1,j\}}(\mathrm{ass}(\gamma))$, of the projection of $\mathrm{ass}(\gamma)$, to each of the coordinate subspaces spanned by $(X_1, X_j), 2 \leq j \leq k$. More precisely, suppose that γ is the curve segment representation with

$$u(\gamma) = ((f, g_0, \ldots, g_k), \sigma),$$

where $f, g_i \in \mathrm{D}[X_1, T]$, and $\sigma \in \{0, 1, -1\}^{\mathrm{Der}_T(f)}$. Then, $\mathrm{ass}(\gamma)$ is defined by

$$\mathrm{ass}(\gamma) = \left\{ \left(x_1, \frac{g_2(x_1, t(x_1))}{g_0(x_1, t(x_1))}, \ldots, \frac{g_k(x_1, t(x_1))}{g_0(x_1, t(x_1))} \right) \;\middle|\; \mathrm{ass}(\tau_1(\gamma)) < x_1 < \mathrm{ass}(\tau_2(\gamma)) \right\},$$

where for each $x_1, \mathrm{ass}(\tau_1(\gamma)) < x_1 < \mathrm{ass}(\tau_2(\gamma))$, and $t(x_1)$ is a root of $f(x_1, T)$ with Thom encoding σ.

Now for each j, $2 \leq j \leq k$, the projection of $\mathrm{ass}(\gamma)$ to the (X_1, X_j)-plane is described by

$$\pi_{\{1,j\}}(\mathrm{ass}(\gamma)) = \left\{ \left(x_1, \frac{g_j(x_1, t(x_1))}{g_0(x_1, t(x_1))} \right) \;\middle|\; \mathrm{ass}(\tau_1(\gamma)) < x_1 < \mathrm{ass}(\tau_2(\gamma)) \right\}.$$

Using an effective quantifier elimination (eliminating T), one can obtain from the above description a quantifier-free formula with free variables X_1, X_j whose realization is equal to $\pi_{\{1,j\}}(\mathrm{ass}(\gamma))$.

We will use the following claim.

Claim 1. *Suppose that γ is a curve segment representation and* $\mathrm{ass}(\gamma)$ *is bounded, then*

$$\mathrm{clos}(\mathrm{ass}(\gamma)) = \bigcap_{2 \le j \le k} \pi^{-1}_{\{1,j\}}(\mathrm{clos}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma)))).$$

Proof of Claim 1. First, suppose that $\mathbf{x} \in \mathrm{clos}(\mathrm{ass}(\gamma))$. Then we have that

$$\mathbf{x} \in \pi^{-1}_{\{1,j\}}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma)))$$

for all j. It is a general fact from topology that for any continuous function f and set A, $f(\mathrm{clos}(A)) \subset \mathrm{clos}(f(A))$. Hence,

$$\pi^{-1}_{\{1,j\}}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma))) \subset \pi^{-1}_{\{1,j\}}(\mathrm{clos}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma))))$$

for all j. Therefore $\mathbf{x} \in \bigcap_{2 \le j \le k} \pi^{-1}_{\{1,j\}}(\mathrm{clos}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma))))$.

Now, suppose that

$$\mathbf{x} = (x_1, \dots, x_k) \in \bigcap_{2 \le j \le k} \pi^{-1}_{\{1,j\}}(\mathrm{clos}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma))))$$

and let $\mathbf{x}_{1,j} = \pi_{\{1,j\}}(\mathbf{x})$ for $j = 2, \dots, k$.

Since, $\mathbf{x}_{1,j} \in \mathrm{clos}(\pi_{\{1,j\}}(\mathrm{ass}(\gamma)))$, using the semi-algebraic curve selection lemma (see for example [1, Theorem 3.19]) there exists $t_{j,0} > 0$, such that there exists a semi-algebraic curve, $\gamma_j = (\gamma_{1,j}, \gamma_{2,j}) : [0, t_{j,0}] : \mathrm{R}^2$, such that $\gamma_{1,j}, \gamma_{2,s}$ are continuous semi-algebraic functions, $\gamma_j(0) = \mathbf{x}_{1,j}$, and $\gamma((0, t_{j,0}]) \subset \pi_{\{1,j\}}(\mathrm{ass}(\gamma))$. Moreover, since $\pi_{\{1,j\}}(\mathrm{ass}(\gamma))$ is a curve parametrized by the X_1 coordinate, $\gamma_{1,j}$ is not a constant function, and without loss of generality (choosing $t_{j,0}$ smaller if necessary) we can assume that $\gamma_{1,j}$ is a strictly increasing function. For $j = 2, \ldots, k$, let $f_j : [x_1, \gamma_{1,j}(t_{j,0})] \to \mathrm{R}$ be defined by

$$f_j(X_1) = \gamma_{2,j}(\gamma_{1,j}^{-1}(X_1)).$$

Taking $x'_1 = \min_{2 \le j \le k} \gamma_{1,j}(t_{j,0})$, we obtain a semi-algebraic curve $\widetilde{\gamma} : [x_1, x'_1] \to \mathrm{ass}(\gamma)$, defined by

$$\widetilde{\gamma}(X_1) = (X_1, f_2(X_1), \ldots, f_k(X_1)).$$

It is easy to check that

$$\widetilde{\gamma}(x_1) = \mathbf{x},$$

and

$$\widetilde{\gamma} : (x_1, x'_1] \subset \mathrm{ass}(\gamma),$$

which proves that $\mathbf{x} \in \mathrm{clos}(\mathrm{ass}(\gamma))$. $\square$

Using Claim 1 we reduce the problem of computing a closed description of $\mathrm{clos}(\mathrm{ass}(\gamma))$ to the problem of computing the closures of $\pi_{\{1,j\}}(\mathrm{ass}(\gamma))$, $2 \le j \le k$, and each of the latter is a 2-dimensional problems which can be solved within our allowed complexity bound using the doubly exponential algorithm referred to previously.

Finally, as in the case of the other algorithms in this paper we include in the input a triangular Thom encoding $\mathcal{T}$ that fixes the first i-coordinates, and the curve segment representations in the input is over $\mathcal{T}$. The computations in the algorithm take place in the ring $\mathrm{D}[\mathrm{ass}(\mathcal{T})]$, and in the description given above, the first coordinate is replaced by the $(i+1)$-st coordinate.

Algorithm 5 (Conversion of curve segment representations to closed formulas)

Input:

1. A triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ of size $i, 0 \leq i \leq k$;
2. a curve segment representation γ over $\mathcal{T}$.

Output:

1. A finite set of polynomials $\mathcal{Q} \subset \mathrm{D}[\mathrm{ass}(\mathcal{T})][X_{i+1}, \ldots, X_k]$;
2. A $\mathcal{Q}$-closed formula Ψ such that

$$\mathcal{R}(\Psi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})} = \mathrm{clos}(\mathrm{ass}(\gamma)).$$

Procedure:

$u \leftarrow u(\gamma) = ((f, g_0, g_{i+2}, \ldots, g_k), \sigma)$.

for $j = i+2, \ldots, k$ **do**

Using Algorithm 14.5 (Quantifier Elimination) in [1, pp. 591] with the formula

$$(\exists T)(f = 0) \wedge \bigwedge_{1 \leq h \leq \deg_T(f)} (\mathbf{sign}(f^{(h)}) = \sigma(f^{(h)})) \wedge (X_j g_0 - g_j = 0),$$

as input, and obtain a $\widetilde{\mathcal{Q}}_j$-formula quantifier-free formula $\widetilde{\phi}_j$, for some $\widetilde{\mathcal{Q}}_j \subset \mathrm{D}[\mathrm{ass}(\mathcal{T})][X_{i+1}, X_j]$ describing such that $\mathcal{R}(\widetilde{\phi}_j) = \pi_{[1,i+1]\cup\{j\}}(\mathrm{ass}(\gamma))$.

Compute a stratifying family of polynomials (see [1, Proposition 5.40] for definition), $\mathcal{Q}_j \subset \mathrm{D}[\mathrm{ass}(\mathcal{T})][X_{i+1}, X_j]$ containing $\widetilde{\mathcal{Q}}_j$.

Using Algorithm 13.1 (Computing realizable sign conditions) in [1, pp. 549] determine the set $\Sigma_j \subset \{0, 1, -1\}^{\mathcal{Q}_j}$ of realizable sign conditions of $\mathcal{Q}_j$.

Determine $\Theta_j \subset \Sigma_j$ such that $\pi_{[1,i+1]\cup\{j\}}(\mathrm{ass}(\gamma)) = \bigcup_{\theta \in \Theta_j} \mathcal{R}(\theta)$.

$\Psi_j \leftarrow \bigvee_{\theta \in \Theta_j} \overline{\theta}$.

end for

$\mathcal{Q} \leftarrow \mathbf{F} \cup \bigcup_{i+2 \leq j \leq k} \mathcal{Q}_j$.

$\Psi \leftarrow \overline{\boldsymbol{\sigma}} \wedge \bigwedge_{i+2 \leq j \leq k} \Psi_j$ (see Notation 3 for definition of $\overline{\boldsymbol{\sigma}}$).

Output $\mathcal{Q}, \Psi$.

Complexity: The complexity of the algorithm is bounded by $(k-i)D_1^{O(i)}D_2^{O(1)}$, where $D_1 = \deg(\mathcal{T})$, and $D_2 = \deg(\gamma)$. Moreover, $\mathrm{card}(\mathcal{Q})$ is bounded by $i + (k-i)D_2^{O(1)}$, and the degrees of the polynomials in $\mathcal{Q}$ are bounded by $\max(D_1, D_2^{O(1)})$.

Proof of Correctness of Algorithm 5. The algorithm reduces the problem to obtaining closed formulas describing the closures of the various $\pi_{i+1,j}(\mathrm{ass}(\gamma))$, $i+2 \leq j \leq k$ (cf. Line 2). After obtaining the descriptions of the various $\pi_{i+1,j}(\mathrm{ass}(\gamma))$ using effective quantifier elimination algorithm (Algorithm 14.5 (Quantifier Elimination) in [1, pp. 591] called in Line 3), closed formulas are obtained describing the closure by computing a stratifying family (cf. 4) in each case. The important property of the stratifying families $\mathcal{Q}_j$ is that the closures, $\mathrm{clos}(\pi_{i+1,j}(\mathrm{ass}(\gamma)))$ are unions of realizations of a set of weak sign conditions on $\mathcal{Q}_j$. This is a consequence of the generalized Thom's Lemma (see [1, Proposition 5.39]). Finally, the set of weak sign conditions on $\mathcal{Q}_j$ whose realizations are contained in $\mathrm{clos}(\pi_{i+1,j}(\mathrm{ass}(\gamma)))$ computed using Algorithm 13.9 (Computing realizable sign conditions) in [1, pp. 549] in Line 5. The disjunction of these weak formulas gives a closed formula, Ψ_j describing $\mathrm{clos}(\pi_{i+1,j}(\mathrm{ass}(\gamma)))$ (cf. Line 7). Now Claim 1 together with Lemma 1 (and noting that the conjunction of a finite set of closed formulas is also closed), imply that the conjunction, Ψ, of the formulas $\Psi_j, j = i+2, \ldots, k$ along with the closed formula $\overline{\boldsymbol{\sigma}}$ (Notation 3), describes $\mathrm{clos}(\mathrm{ass}(\gamma))$ (cf. Line 10). This proves the correctness of the algorithm. □

Complexity analysis of Algorithm 5. As before, each arithmetic operation in $\mathrm{D}[\mathrm{ass}(\mathcal{T})]$ costs $D_1^{O(i)}$ arithmetic operations in D (where $D_1 = \deg(\mathcal{T})$). There are $(k-i)$ two dimensional projections (cf. Line 2). The complexity of each of these two-dimensional sub-problems (measured in terms of number of operations in $\mathrm{D}[\mathrm{ass}(\mathcal{T})]$) is bounded by $D_2^{O(1)}$, where $D_2 = \deg(\gamma)$, and this follows from the complexity bounds on the various algorithms used in the different steps (namely, Algorithm 14.5 (Quantifier Elimination) in [1, pp. 591] in Line 3, algorithm for computing stratifying families in 4, and Algorithm 13.9 (Computing realizable sign conditions) in [1, pp. 549] in Line 5). Note that these algorithms are used with the number of variables equal to 2, and hence the complexity of each call (measured in terms of arithmetic operations in $\mathrm{D}[\mathrm{ass}(\mathcal{T})]$) are polynomially bounded in D_1. This completes the complexity analysis of Algorithm 1. $\square$

Before we prove our main result, we must first introduce some definitions and notation required to understand the algorithm in [6].

Definition 3.0.1 (Homological ℓ-equivalences)**.** *We say that a map $f : X \to Y$ between two topological spaces is a* homological ℓ-equivalence *if the induced homomorphisms between the homology groups $f_* : \mathrm{H}_i(X) \to \mathrm{H}_i(Y)$ are isomorphisms for $0 \leq i \leq \ell$.*

The relation of homological ℓ-equivalence as defined above is not an equivalence relation since it is not symmetric. In order to make it symmetric one needs to "formally invert" ℓ-equivalences.

Definition 3.0.2 (Homologically ℓ-equivalent)**.** *We will say that X* is homologically ℓ-equivalent to Y *(denoted $X \sim_\ell Y$), if and only if there exists spaces, $X = X_0, X_1, \ldots, X_n = Y$ and homological ℓ-equivalences $f_1, \ldots, f_n$ as shown below:*

$$\begin{tikzcd}
 & X_1 \arrow[dl, "f_1"'] \arrow[dr, "f_2"] & & X_3 \arrow[dl, "f_3"'] \arrow[dr, "f_4"] & & \cdots & \cdots & & X_{n-1} \arrow[dl, "f_{n-1}"'] \arrow[dr, "f_n"] & \\
X_0 & & X_2 & & \cdots & & & \cdots & & X_n
\end{tikzcd}$$

Note that $\sim_\ell$ is in fact an equivalence relation.

Definition 3.0.3 (Diagrams of topological spaces)**.** *A* diagram of topological spaces *is a functor, $X : J \to \mathbf{Top}$, from a small category J to $\mathbf{Top}$.*

In particular, we need the notion of homological ℓ-equivalence between diagrams of topological spaces. We denote by $\mathbf{Top}$ the category of topological spaces.

Definition 3.0.4 (Homological ℓ-equivalence between diagrams of topological spaces)**.** *Let J be a small category, and $X, Y : J \to \mathbf{Top}$ be two functors.*

We will say that a diagram $X : J \to \mathbf{Top}$ is homologically ℓ-equivalent to the diagram $Y : J \to \mathbf{Top}$ *(denoted as before by $X \sim_\ell Y$), if and only if there exists diagrams $X = X_0, X_1, \ldots, X_n = Y : J \to \mathbf{Top}$ and homological ℓ-equivalences $f_1, \ldots, f_n$ as shown below:*

$$\begin{tikzcd}
 & X_1 \arrow[dl, "f_1"'] \arrow[dr, "f_2"] & & X_3 \arrow[dl, "f_3"'] \arrow[dr, "f_4"] & & \cdots & \cdots & & X_{n-1} \arrow[dl, "f_{n-1}"'] \arrow[dr, "f_n"] & \\
X_0 & & X_2 & & \cdots & & & \cdots & & X_n
\end{tikzcd}$$

Again, $\sim_\ell$ is indeed an equivalence relation.

Notation 6 (Diagram of various unions of a finite number of subspaces)**.** *Let J be a finite set, A a topological space, and $\mathcal{A} = (A_j)_{j \in J}$ a tuple of subspaces of A indexed by J.*

For any subset $J' \subset J$, we denote

$$\mathcal{A}^{J'} \quad = \quad \bigcup_{j' \in J'} A_{j'}.$$

We consider 2^J as a category whose objects are elements of 2^J, and whose only morphisms are given by:

$$\begin{aligned} 2^J(J', J'') &= \emptyset \text{ if } J' \not\subset J'', \\ 2^J(J', J'') &= \{\iota_{J',J''}\} \text{ if } J' \subset J''. \end{aligned}$$

We denote by $\mathbf{Simp}^J(\mathcal{A}) : 2^J \to \mathbf{Top}$ the functor (or the diagram) defined by

$$\mathbf{Simp}^J(\mathcal{A})(J') = \mathcal{A}^{J'}, J' \in 2^J,$$

and $\mathbf{Simp}^J(\mathcal{A})(\iota_{J',J''})$ is the inclusion map $\mathcal{A}^{J'} \hookrightarrow \mathcal{A}^{J''}$.

Now that we have the definition of homological equivalence of diagrams as defined above, we can state the specifications of the simplicial replacement algorithm described in [6].

Algorithm 6 (Simplicial replacement)

Input:

(a) A finite set of polynomials $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$;

(b) An integer $N \geq 0$, and for each $i \in [N]$, a $\mathcal{P}$-closed formula ϕ_i;

(c) $\ell, 0 \leq \ell \leq k$.

Output:

A simplicial complex Δ and for each $I \subset [N]$ a subcomplex $\Delta_I \subset \Delta$ such that there is a diagrammatic homological ℓ-equivalence

$$(I \mapsto \Delta_I)_{I \subset [N]} \overset{h}{\sim}_{\ell} \mathbf{Simp}^{[N]}(\mathcal{R}(\Phi)),$$

where $\Phi(i) = \phi_i, i \in [N]$.

Complexity: The complexity of the algorithm is bounded by $(sd)^{k^{O(\ell)}}$, where $s = \mathrm{card}(\mathcal{P})$ and $d = \max_{P \in \mathcal{P}} \deg(P)$.

With these preliminaries in place, we are now ready to present an algorithm which satisfies the conditions of 3.0.1.

Algorithm 7 (Computing one-dimensional subset)

Input:

1. A triangular Thom encoding $\mathcal{T} = (\mathbf{F}, \boldsymbol{\sigma})$ of size i;
2. a finite set $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$;
3. a $\mathcal{P}$-closed formula Φ such that $\mathcal{R}(\Phi, \mathrm{R}^k)$ is bounded;
4. a finite set $\mathcal{M}$ of real univariate representations over $\mathcal{T}$, whose set of associated points, M, is contained in $\mathcal{R}(\Phi, \mathrm{R}^k)$.

Output:

1. a finite set $\mathcal{U}$ of real univariate representations over $\mathcal{T}$;
2. a finite indexing set I and a finite tuple $(\gamma_j)_{j \in I}$ where each γ_j a curve segment representation over $\mathcal{T}$;
3. mappings $L, R : I \to \mathcal{U}$, defined by $\mathrm{ass}(L(j))$ is the left endpoint of γ_j, and $\mathrm{ass}(R(j))$ is the right endpoint of γ_j.

Procedure:

1: **if** $k - i = 1$ **then**

2: **for** each $u = \{(f, g_0, g_k), \sigma\} \in \mathcal{M}$ **do**

3: Let $R_u \in \mathrm{R}[X_k]$ be the Sylvester resultant (see for example [1, pp. 118]) with respect to the variable T of the polynomials $f, X_k g_0 - g_k$.

4: Use Algorithm 10.96 (Sign Determination) [1, pp. 426] to compute a Thom encoding $\tau_u = (R_u, \sigma_u)$ over $\mathcal{T}$, such that $\mathrm{ass}(\tau_u) = \pi_k(\mathrm{ass}(u))$.

5: $\mathcal{P} \leftarrow \mathcal{P} \cup \{R_u\}$.

6: **end for**

7: Use Algorithm 12.69 (Triangular Comparison of Roots) in [1, pp. 534] repeatedly with inputs $\mathcal{T}$ and pairs of polynomials in $\mathcal{P}$, and order the real roots of the polynomials $P(\mathrm{ass}(\mathcal{T}), X_k), P \in \mathcal{P}$, and hence obtain a partition of R into points and open intervals, and identify those points and open intervals which are contained in $\mathcal{R}(\Phi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})}$.

8: $\mathcal{U} \leftarrow \emptyset$.

9: $I \leftarrow \emptyset$.

10: $j \leftarrow 0$.

11: **for** each Thom encoding (P, σ) over $\mathcal{T}$ obtained in Line 4 whose associated point is in $S = \mathcal{R}(\Phi, \mathrm{R}^k)$ **do**

12: $\mathcal{U} \leftarrow \mathcal{U} \cup \{((P, X_k, 1), \sigma)\}$.

13: **end for**

14: **for** each open interval with end-points described by the Thom encodings $\tau_1 = ((P_1, X_k, 1), \sigma_1), \tau_2 = ((P_2, X_k, 1), \sigma_2) \in \mathcal{U}$ with $\mathrm{ass}(\tau_1) < \mathrm{ass}(\tau_2)$ such that $(\mathrm{ass}(\tau_1), \mathrm{ass}(\tau_2)) \subset \pi_k(\mathcal{R}(\Phi, \mathrm{R}^k))$ **do**

15: $I \leftarrow I \cup \{j\}$.

16: $j \leftarrow j + 1$.

$\gamma_j \leftarrow \gamma$, where γ is the curve segment representation over $\mathcal{T}$ defined by:

$$\begin{aligned}\tau_1(\gamma) &= \tau_1,\\ \tau_2(\gamma) &= \tau_2,\\ u(\gamma) &= ((T,1),(0,1)).\end{aligned}$$

$L(j) \leftarrow \tau_1(\gamma_j)$.

$R(j) \leftarrow \tau_2(\gamma_j)$.

end for

Output $\mathcal{U}, (\gamma_j)_{j\in I}$, and the mappings $L, R : I \to \mathcal{U}$.

end if

Use Algorithm 4 (Curve segments) with input $(\mathcal{T}, \mathcal{P}, \Phi)$ to compute:

(a) A finite tuple $\mathcal{F} = (\tau_1, \ldots, \tau_N)$ of Thom encodings over $\mathcal{T}$, with $\mathrm{ass}(\tau_1) < \cdots < \mathrm{ass}(\tau_N)$;

(b) for each j, $1 \leq j \leq N-1$, a finite tuple $\mathcal{C}_j$ of curve segment representations over $\mathcal{T}$ such that for each $\gamma \in \mathcal{C}_j$, $\tau_1(\gamma) = \tau_j$, $\tau_2(\gamma) = \tau_{j+1}$;

(c) for each $j, 1 \leq j \leq N$ a finite set $\mathcal{U}_j$ of real univariate representations over $\mathcal{T}$, such that for each $u \in \mathcal{U}_j$, the set of points $\{\mathrm{ass}(u) \mid u \in \mathcal{U}_j\}$ is precisely the set of end-points of the curve segments in $\mathcal{C}_{j-1} \cup \mathcal{C}_j$ (with the convention that $\mathcal{C}_0 = \emptyset$);

(d) mappings $L_j, R_{j-1} : \mathcal{C}_j \to \mathcal{U}_j$, such that $\mathrm{ass}(L_j(\gamma))$ is the left end-point of γ, and $\mathrm{ass}(R_j(\gamma))$ is the right end-point of γ.

for $\tau = (f, \sigma) \in \mathcal{F}$ **do**

$\mathcal{M}_\tau \leftarrow \{u \in \mathcal{M} \cup \mathcal{U} \mid \pi_{i+1}(\mathrm{ass}(u)) = \mathrm{ass}(\tau)\}$.

$\mathcal{T}_\tau \leftarrow ((F, f), (\boldsymbol{\sigma}, \sigma)))$.

27: Call Algorithm 7 (Computing one-dimensional subset) recursively with input $(\mathcal{T}_\tau, \mathcal{P}, \Phi, \mathcal{M}_\tau)$ and obtain a set of $\mathcal{U}_\tau$ of real univariate representations over $\mathcal{T}_\tau$, an indexing set I_τ, a tuple $(\gamma_i)_{i \in I_\tau}$ of curve segment representations, and mappings $L_\tau, R_\tau : I_\tau \to \mathcal{U}_\tau$. (Note that for each $i \in I_\tau, \mathrm{ass}(\gamma_i) \subset S_{\mathrm{ass}(\tau)}$.)

28: **end for**

29: $I \leftarrow \bigcup_{\tau \in \mathcal{F}} I_\tau$.

30: $\mathcal{U} \leftarrow \bigcup_{\tau \in \mathcal{F}} \mathcal{U}_\tau$.

31: $L \leftarrow \bigcup_{1 \le j \le N} L_j \cup \bigcup_{\tau \in \mathcal{F}} L_\tau$. (Union of disjoint mappings means the disjoint union of their graphs.)

32: $R \leftarrow \bigcup_{1 \le j \le N} R_j \cup \bigcup_{\tau \in \mathcal{F}} R_\tau$.

33: Output $\mathcal{U}, I, (\gamma_j)_{j \in I}, L, R$.

Complexity: Suppose that $\deg(\mathcal{T}) = d^{O(k)}$. The complexity of the algorithm is bounded by $(sd)^{O(k^2)}$, where $s = \mathrm{card}(\mathcal{P}), d = \max_{P \in \mathcal{P}}$. Moreover, $\mathrm{card}(I) = (sd)^{O((k-i)^2)}$, and the degrees of the elements of $\mathcal{U}$ and $\gamma_j, j \in I$ are bounded by $(d^{O(k)}, d^{O(k)})$.

Proposition 3.0.1. *The output of Algorithm 7 has the following properties. Let* $\Gamma = \bigcup_{j \in I} \mathrm{clos}(\mathrm{ass}(\gamma_j))$, *and* $S = \mathcal{R}(\Phi, \mathrm{R}^k)_{\mathrm{ass}(\mathcal{T})}$.

(a) $M \subset \Gamma$;

(b) $\Gamma \subset S$;

(c) $\dim(\Gamma) \le 1$;

(d) the homomorphism $i_{*,0} : \mathrm{H}_0(\Gamma) \to \mathrm{H}_0(S)$ *induced by the inclusion map* $i : \Gamma \hookrightarrow S$ *is an isomorphism;*

(e) the homomorphism $i_{*,1} : \mathrm{H}_1(\Gamma) \to \mathrm{H}_1(S)$ *induced by the inclusion map* $i : \Gamma \hookrightarrow S$ *is an epimorphism.*

Proof. The property in Part (a) is ensured in Lines 2 and 25. Part (b) follows from the property of the output of Algorithm 4 (Curve segments) (called in Line 23) given in Part (b) of Proposition 2.1.2.

Part (c) holds since Γ is by definition the finite union

$$\bigcup_{j \in I} \mathrm{clos}(\mathrm{ass}(\gamma_j))$$

and $\dim \mathrm{ass}(\gamma_j) = 1$ for each $j \in I$, and taking the closure does not increase the dimension of a semi-algebraic set.

It is a standard exercise (see for example proof of Proposition 15.8 in [1]) to prove that the semi-algebraic set Γ satisfies the properties of being roadmap for $S = \mathcal{R}(\Phi, \mathrm{R}^k)$ (see [1, Chapter 15]), which implies Part (d).

We now prove Part (e).

The proof is by induction on $k - i$.

Base case: $k - i = 1$. In this case the claim is clear since $\mathrm{H}_1(\Gamma) = \mathrm{H}_1(S) = 0$.

Suppose the claim is true for all smaller values of $k - i$. Notice that Algorithm 7 (Computing one-dimensional subset) is called recursively in Line 27. In these calls the triangular Thom encoding $\mathcal{T}_\tau, \tau \in \mathcal{F}$ in the input is of size $i + 1$, while the number of variables is still k. We have also have

$$\Gamma_\tau := \mathrm{clos}\left(\bigcup_{j \in I_\tau} \mathrm{ass}(\gamma_j) \right) = \Gamma_{\mathrm{ass}(\tau)}.$$

Thus, using the induction hypothesis for this recursive call (since $k - (i+1) < k - i$), we obtain that the restriction of the inclusion $\Gamma \to S$ to Γ_τ induces a surjection

$$\mathrm{H}_1(\Gamma_\tau) \to \mathrm{H}_1(S_{\mathrm{ass}(\tau)}). \tag{3.1}$$

For $1 \leq j \leq N$, denote $t_j = \pi_{i+1}(\mathrm{ass}(\tau_j))$.

$$\begin{aligned} a_0 &= t_1, \\ a_i &= \frac{t_i + t_{i+1}}{2}, 1 \leq i < N, \\ a_N &= t_N, \end{aligned}$$

We prove the following claims.

Claim 2. *For each $j, 0 \leq j \leq N$, the inclusion maps induce the following isomorphisms.*

$$\mathrm{H}_*(S_{(-\infty,t_j]}) \to \mathrm{H}_*(S_{(-\infty,a_j]}), \tag{3.2}$$

$$\mathrm{H}_*(S_{t_j}) \to \mathrm{H}_*(S_{[a_{j-1},a_j]}), \tag{3.3}$$

$$\mathrm{H}_*(\Gamma_{(-\infty,t_j]}) \to \mathrm{H}_*(\Gamma_{(-\infty,a_j]}), \tag{3.4}$$

$$\mathrm{H}_*(\Gamma_{t_j}) \to \mathrm{H}_*(\Gamma_{[a_{j-1},a_j]}). \tag{3.5}$$

Proof. Parts (3.2) and (3.3) are consequences of the property of the output of Algorithm 4 (Curve segments) (which is called in Line 23) given in Part (a) of Proposition 2.1.2.

Parts (3.4) and (3.5) follow from the fact that there is a retraction (along the X_{i+1}-coordinate) of $\Gamma_{(-\infty,a_j]}$ to $\Gamma_{(-\infty,t_j]}$ (resp. $\Gamma_{[a_{j-1},a_j]}$ to Γ_{t_j}) making use of the fact that distinct curve segments over the open intervals (t_{j-1}, t_j) do not intersect which is ensured by Part (d) of Proposition 2.1.2. □

Claim 3. *Let $a, b \in \{a_0, \ldots, a_N\}$, with $a \leq b$. The inclusion map $\Gamma \hookrightarrow S$ induces isomorphisms*

$$\mathrm{H}_0(\Gamma_{(-\infty,a]}) \to \mathrm{H}_0(S_{(\infty,a]}),$$

$$\mathrm{H}_0(\Gamma_{[a,b]}) \to \mathrm{H}_0(S_{[a,b]}),$$

Proof. This follows from the fact that $\Gamma_{(-\infty,a]}$ (resp. $\Gamma_{[a,b]}$) satisfy the roadmap property with respect to the set $S_{(\infty,a]}$ (resp. $S_{[a,b]}$). The proof of this fact is standard and omitted. □

Using the claims proved above, we are now going to prove using induction on j, that the inclusion map $\Gamma \to S$ induces an isomorphism,

$$\mathrm{H}_1(\Gamma_{(\infty,a_j]}) \to \mathrm{H}_1(S_{(\infty,a_j]}).$$

The claim is true for $j = 0$ by the global induction hypothesis on i, and hence is also true for $j = 1$ using using (3.2) and (3.4).

We prove it for $j > 1$ by induction. Suppose the claim holds until $j - 1$.

Hence we have isomorphism

$$\mathrm{H}_1(\Gamma_{(\infty,a_{j-1}]}) \to \mathrm{H}_1(S_{(\infty,a_{j-1}]})$$

induced by inclusion.

Observe that for any set $X \subset \mathrm{R}^k$,

$$\begin{aligned} X_{(-\infty,a_j]} &= X_{(-\infty,a_{j-1}]} \cup X_{[a_{j-1},a_j]}, \\ X_{a_{j-1}} &= X_{(-\infty,a_{j-1}]} \cap X_{[a_{j-1},a_j]} X_{a_{j-1}}. \end{aligned}$$

Let

$$\begin{aligned} A_1 &= \Gamma_{(-\infty,a_{j-1}]}, \\ A_2 &= \Gamma_{[a_{j-1},a_j]}, \\ B_1 &= S_{(-\infty,a_{j-1}]}, \\ B_2 &= S_{[a_{j-1},a_j]}. \end{aligned}$$

Also, let A_{12} (resp. B_{12}) denote $A_1 \cap A_2$ (resp. $B_1 \cap B_2$), and A^{12} (resp. B^{12}) denote $A_1 \cup A_2$ (resp. $B_1 \cup B_2$).

The Mayer-Vietoris exact sequence (see for example [1, Theorem 6.35]) yields the following commutative diagrams with exact rows and vertical arrows induced by various restrictions of the inclusion $\Gamma \hookrightarrow S$.

$$\begin{tikzcd} \mathrm{H}_1(A_{12})^a \arrow[r] \arrow[d] & \mathrm{H}_1(A_1) \oplus \mathrm{H}_1(A_2)^{b_1 \oplus b_2} \arrow[r] \arrow[d] & \mathrm{H}_1(A^{12})^c \arrow[r] \arrow[d] & \mathrm{H}_0(A_{12})^d \arrow[r] \arrow[d] & \mathrm{H}_0(A_1) \oplus \mathrm{H}_0(A_2)^{e_1 \oplus e_2} \arrow[d] \\ \mathrm{H}_1(B_{12}) \arrow[r] & \mathrm{H}_1(B_1) \oplus \mathrm{H}_1(B_2) \arrow[r] & \mathrm{H}_1(B^{12}) \arrow[r] & \mathrm{H}_0(B_{12}) \arrow[r] & \mathrm{H}_0(B_1) \oplus \mathrm{H}_0(B_2) \end{tikzcd}$$

By induction hypothesis on j, the map $b_1 : \mathrm{H}_1(A_1) \to \mathrm{H}_1(B_1)$ is surjective. Using (3.1), (3.3), and (3.5) we have that the map $b_2 : \mathrm{H}_1(A_2) \to \mathrm{H}_1(B_2)$ is surjective. Hence the map $b = b_1 \oplus b_2$ is surjective.

Using Proposition 2.1.2 (c) we have that map d is surjective.

Finally, using Claim 3, we have that the maps e_1 and e_2 are both isomorphisms, and hence so is e. In particular, e is injective.

It follows from the above and (one-half of) the "Five-lemma" (see for example [23]), that c is a surjection. $\square$

Now, we can present an algorithm to prove Theorem 3.0.1.

Algorithm 8 (Computing homology basis)

Input:

1. a finite set $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$;
2. a $\mathcal{P}$-closed formula Φ.

Output:

1. a finite set $\mathcal{Q} \subset \mathrm{D}[X_1, \ldots, X_k]$;
2. a finite tuple $(\Psi_j)_{j \in J}$, in which each Ψ_j is a $\mathcal{Q}$-formula, such that the realizations $\Gamma_j = \mathcal{R}(\Psi_j, \mathrm{R}^k)$ have the following properties:
   (a) For each $j \in J$, $\Gamma_j \subset S$ and Γ_j is semi-algebraically homeomorphic to $\mathbf{S}^1$;
   (b) the inclusion map $\Gamma_j \hookrightarrow S$ induces an injective map $\mathbb{F} \cong \mathrm{H}_1(\Gamma_j) \to \mathrm{H}_1(S)$, whose image we denote by $[\Gamma_j]$;
   (c) the tuple $([\Gamma_j])_{j \in J}$ forms a basis of $\mathrm{H}_1(S)$.

Procedure:

1: Use Algorithm 3 (Big enough radius) with input $(\mathcal{P}, \Phi)$ and let $r = \frac{a}{b} > 0, a, b \in \mathrm{D}$ be the output.

2: $\mathcal{P} \leftarrow \mathcal{P} \cup \{b^2(X_1^2 + \cdots + X_k^2) - a^2\}$.

3: $\Phi \leftarrow \Phi \wedge (b^2(X_1^2 + \cdots + X_k^2) - a^2 \leq 0)$.

Call Algorithm 7 (Computing one-dimensional subset) with input $(\mathcal{P}, \Phi)$ to obtain

(a) a finite set $\mathcal{U}$ of real univariate representations over $\mathcal{T}$;

(b) a finite indexing set I and a finite tuple $(\gamma_j)_{j\in I}$ where each γ_j a curve segment representation over $\mathcal{T}$;

(c) mappings $L, R : I \to \mathcal{U}$, defined by $\mathrm{ass}(L(j))$ is the left endpoint of γ_j, and $\mathrm{ass}(R(j))$ is the right endpoint of γ_j.

Using Algorithm 5 (Conversion of curve segment representations to closed formulas) compute for each $j \in I$ a set of polynomials $\mathcal{Q}_j$ and a $\mathcal{Q}_j$-closed formula Θ_j such that $\mathcal{R}(\Theta_j) = \mathrm{clos}(\mathrm{ass}(\gamma_j))$.

$\mathcal{Q} \leftarrow \bigcup_{j\in I} \mathcal{Q}_j$.

$G \leftarrow (E = I, V = \mathcal{U}, \mathrm{head} = L, \mathrm{tail} = R)$.

Using a graph traversal algorithm compute a tuple $(C_1, \ldots, C_N)$ where each $C_h = (i_{h,0}, \ldots, i_{h,q_h-1}) \in I^{q_h}$ and

(a) $\mathrm{tail}(i_{j,h}) = \mathrm{head}(i_{j,h+1 \mod q_j}), h = 0, \ldots, q_j - 1$.

(b) $C_1, \ldots, C_N$ are simple cycles of G.

(c) The cycles $C_1, \ldots, C_N$ form a basis of the cycle space of G (which is isomorphic to $\mathrm{H}_1(|G|)$).

for $1 \le h \le N$ **do**

  $\Psi_h \leftarrow \Theta_{i_{h,0}} \vee \cdots \vee \Theta_{i_{h,q_h-1}}$.

end for

12: Call Algorithm 6 (Simplicial replacement) with input:

1. $\mathcal{Q}$,
2. the tuple of $\mathcal{Q}$-closed formulas $\mathbf{\Phi} = (\phi_0, \ldots, \phi_N) = (\Psi_1, \ldots, \Psi_N, \Phi)$,
3. $\ell = 1$

to obtain a simplicial complex $\Delta_1(\mathbf{\Phi})$ such that

$$(J \mapsto |\Delta_1(\mathbf{\Phi}|_J)|)_{J \subset [N]}$$

is homologically 1-equivalent (cf. Definition 3.0.2) to $\mathbf{Simp}^{[N]}(\mathcal{R}(\mathbf{\Phi}))$ (cf. Notation 6).

13: Using Gauss-Jordan elimination identify a minimal subset $J \subset \{1, \ldots, N\}$ such that the $\mathrm{span}(\{\mathrm{Im}(\mathrm{H}_1(\Delta_1(\mathbf{\Phi}|_{\{h\}})) \to \mathrm{H}_1(\Delta_1(\mathbf{\Phi}))) \mid h \in J\}) = \mathrm{H}_1(\Delta_1(\mathbf{\Phi}))$.

14: Output $(\Psi_h)_{h \in J}$.

Complexity: The complexity of each $\Psi_h, h \in J$ is bounded by $(sd)^{O(k^2)}$, and the complexity of the algorithm is bounded by $(sd)^{k^{O(1)}}$, where $s = \mathrm{card}(\mathcal{P})$ and $d = \max_{P \in \mathcal{P}} \deg(P)$.

Proof of correctness of Algorithm 8. The correctness of Algorithm 8 follows from the correctness of Algorithms 3, 4, 5, and 6. □

Complexity analysis of Algorithm 8. The complexity upper bound is a consequence of the complexity analysis of the Algorithms 3, 4, 5, and 6. □

Proof of Theorem 3.0.1. Theorem 3.0.1 follows from the correctness and complexity analysis of Algorithm 8. □

4. REEB SPACES OF SEMI-ALGEBRAIC MAPS

We begin by defining the Reeb space, a generalization of the Reeb graph, a concept first developed by Georges Reeb as a tool in Morse Theory [8]. In the next two chapters, to ensure that the resulting Reeb spaces are semi-algebraic sets, we restrict ourselves to working with proper maps, that is, maps with the property that inverse images of compact subsets are compact.

While it is possible to compute directly the Reeb space of a semi-algebraic map, the best known algorithm to compute semi-algebraic descriptions of quotients has doubly exponential complexity using cylindrical algebraic decomposition (CAD), which is often used in algorithms for computing topological properties of semi-algebraic sets [1]. For the general case of the Reeb space, our algorithm has similar complexity. In the special case of Reeb graphs, however, we were able to produce an algorithm with singly-exponential complexity to compute the Reeb graph of a function from the roadmap.

Definition 4.0.1. *The* Reeb space *of the map f, henceforth denoted* $\mathrm{Reeb}(f)$*, is the topological space $X/\sim$, equipped with the quotient topology, where $x \sim x'$ if and only if $f(x) = f(x')$, and x, x' belong to the same connected component of $f^{-1}(f(x))$.*

The Reeb space of a function depends highly on on the choice of function $f\colon X \to Y$. To see this, consider the following two functions from the torus into $\mathbb{R}$:

(a) $f\colon T^2 \to \mathbb{R}$

(b) $g\colon T^2 \to \mathbb{R}$

Figure 4.1. Two Reeb graphs of T^2

Letting $\beta(X)$ denote the sum of the Betti numbers of a space X, note that $\beta(\mathrm{Reeb}(f)) = 2$ and $\beta(\mathrm{Reeb}(g)) = 1$, both of which are less than $\beta(T^2) = 3$. The authors of [29] produced the first bound on the Betti numbers of a Reeb graph, showing that, for a manifold M and a Morse function $f\colon M \to \mathbb{R}$, $\beta_1(R(f)) \leq \beta_1(M)$. In fact, the complexity of the Reeb graph of an arbitrary map $f\colon X \to \mathrm{R}$ is bounded by the topological complexity of X, namely, that $\beta(\mathrm{Reeb}(f)) \leq \beta(X)$, as noted in [30, page 141]. However, this does not generalize to Reeb spaces. The following example shows that the topological complexity of a Reeb space can grow arbitrarily large compared to that of the original space.

Example 1. *Consider the closed n-dimensional disk $\mathbf{D}^n$ with $n \geq 1$, and let $\sim$ be the equivalence relation identifying all points on the boundary of $\mathbf{D}^n$, shown in the figure below.*

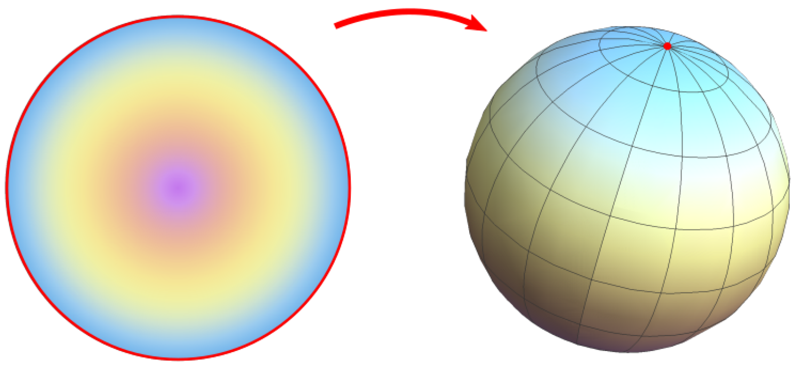


Figure 4.2. Reeb space of $\mathbf{D}^2$

Then $\mathbf{D}^n/\sim\, \cong \mathbf{S}^n$, the n-sphere. Letting f_n denote the quotient map $f_n : \mathbf{D}^n \to \mathbf{S}^n$, the fibers of f_n consist of either one point or the boundary $\mathbf{S}^{n-1}$ of $\mathbf{D}^n$. Therefore $\mathrm{Reeb}(f_n) \cong \mathbf{S}^n$ *for all $n > 1$. Because $\beta_0(\mathbf{D}^n) = 1$ and $\beta_i(\mathbf{D}^n) = 0$ for all $i > 0$ and furthermore, $\beta_0(\mathbf{S}^n) = 1$, $\beta_n(\mathbf{S}^n) = 1$, and $\beta_i(\mathbf{S}^n) = 0, i \neq 0, n$, we have that for $n > 1$,*

$$\begin{aligned} \beta(\mathbf{D}^n) &= 1, \\ \beta(\mathrm{Reeb}(f_n)) &= 2. \end{aligned}$$

Thus the inequality $\beta(\mathrm{Reeb}(f)) \leq \beta(X)$ does not hold for arbitrary maps $f\colon X \to Y$. Indeed, the topological complexity of the Reeb space can be arbitrarily large compared to that of the original space. For $k \geq 0$, let

$$f_{n,k} = \underbrace{f \times \cdots \times f}_{k \text{ times}} : \underbrace{\mathbf{D}^n \times \cdots \times \mathbf{D}^n}_{k \text{ times}} \longrightarrow \underbrace{\mathbf{S}^n \times \cdots \times \mathbf{S}^n}_{k \text{ times}}.$$

Using the same argument as before, for $n > 1$ and $k > 0$,

$$\mathrm{Reeb}(f_{n,k}) \cong \underbrace{\mathbf{S}^n \times \cdots \times \mathbf{S}^n}_{k \text{ times}}.$$

Thus,

$$\begin{aligned} \beta_0(\underbrace{\mathbf{D}^n \times \cdots \times \mathbf{D}^n}_{k \text{ times}}) &= 1, \\ \beta_i((\underbrace{\mathbf{D}^n \times \cdots \times \mathbf{D}^n}_{k \text{ times}})) &= 0,\ i > 0, \end{aligned}$$

and hence

$$\beta(\underbrace{\mathbf{D}^n \times \cdots \times \mathbf{D}^n}_{k \text{ times}}) = 1.$$

Moreover, for $n > 1$,

$$\begin{aligned} \beta_i(\mathrm{Reeb}(f_{n,k})) &= 0 \text{ if } n \not| \, i \text{ or if } i > nk, \\ \beta_i(\mathrm{Reeb}(f_{n,k})) &= \binom{k}{i/n} \text{ otherwise}, \end{aligned}$$

and hence for $n > 1$,

$$\beta(\mathrm{Reeb}(f_{n,k})) = 2^k.$$

Thus, even for definably proper maps $f : X \to Y$, the individual as well as the total Betti numbers of $\mathrm{Reeb}(f)$ can be arbitrarily large compared to those of X.

In this chapter, we determine a bound on the topological complexity of a Reeb space in terms of the complexity of both the complexity of X and f. To do this, we must first discuss definability of Reeb spaces.

4.1 Reeb Spaces as Semi-algebraic Quotients

In this section, we show that the Reeb space of a semi-algebraic map can be realized as a semi-algebraic quotient. This result implies the possible existence of an algorithm to compute the Reeb space of a semi-algebraic map as a semi-algebraic quotient.

Theorem 4.1.1. *Let $X \subset \mathrm{R}^n$ be a closed and bounded semi-algebraic set, and $f : X \to Y$ be a semi-algebraic map. Then, the space* $\mathrm{Reeb}(f) \triangleq X/\sim$ *is a definably proper quotient. In other words, let $X \subset R^n$ be a closed and bounded semi-algebraic set, and $f\colon X \to Y$ be a semi-algebraic map. Then there exists a semi-algebraic set Z, and a proper semi-algebraic map $\psi : X \to Z$ and a homeomorphism $\theta :$* $\mathrm{Reeb}(f) \to Z$ *such that the following diagram commutes:*

$$\begin{tikzcd} X \arrow[r, "q"] \arrow[d] & \mathrm{Reeb}(f) = X/\sim \arrow[dl] \\ Z & \end{tikzcd}$$

(here q is the quotient map). In particular, $\mathrm{Reeb}(f)$ *is homeomorphic to a semi-algebraic set.*

Remark 1. *To see why the assumption that X is closed and bounded is needed, consider the example where $X = \mathrm{R}^2 \setminus \mathbf{0}$ and $f : X \to \mathrm{R}$ is the projection map forgetting the second coordinate. Each fiber $f^{-1}(x)$ has one connected component except where $x = 0$, where $f^{-1}(0)$ has two connected components. The resulting Reeb space of f is homeomorphic to* R *with a doubled point, which is not a semi-algebraic set.*

We now prove Theorem 4.1.1.

Proof of Theorem 4.1.1. We first claim that the relation, "$x \sim x'$ if and only if $f(x) = f(x')$, and x, x' belong to the same connected component of $f^{-1}(f(x))$" is a definably proper equivalence relation. Using Hardt's triviality theorem for o-minimal structures [31], [32], we have that there exists a finite definable partition of Y into locally closed definable sets $(Y_\alpha)_{\alpha\in I}$, $y_\alpha \in Y_\alpha$, and definable homeomorphisms $\phi_\alpha : Y_\alpha \times f^{-1}(y_\alpha) \to f^{-1}(Y_\alpha)$ such that the following diagram commutes for each $\alpha \in I$:

$$\begin{tikzcd} Y_\alpha \times f^{-1}(y_\alpha) \arrow[rr, "\phi_\alpha"] \arrow[dr, "\pi_1"'] & & f^{-1}(Y_\alpha) \arrow[dl, "f|_{f^{-1}(Y_\alpha)}"] \\ & Y_\alpha & \end{tikzcd}$$

(here π_1 is the projection to the first factor in the direct product). For each $\alpha \in I$, let $(C_{\alpha,\beta})_{\beta\in J_\alpha}$ be the connected components of $f^{-1}(y_\alpha)$, and for each $\alpha \in I, \beta \in J_\alpha$, let $D_{\alpha,\beta} = \phi_\alpha(Y_\alpha \times C_{\alpha,\beta})$.

Let

$$E = \bigcup_{\alpha\in I,\beta\in J_\alpha} (\phi_\alpha \times \phi_\alpha)((Y_\alpha \times C_{\alpha,\beta}) \times_{\pi_1} (Y_\alpha \times C_{\alpha,\beta})),$$

where $(Y_\alpha \times C_{\alpha,\beta}) \times_{\pi_1} (Y_\alpha \times C_{\alpha,\beta})$ is the definable subset of $(Y_\alpha \times f^{-1}(y_\alpha)) \times (Y_\alpha \times f^{-1}(y_\alpha))$ defined by

$$((y, x), (y', x')) \in (Y_\alpha \times C_{\alpha,\beta}) \times_{\pi_1} (Y_\alpha \times C_{\alpha,\beta}) \Leftrightarrow y = y', x, x' \in C_{\alpha,\beta}.$$

It is clear that E is a definable subset of $X \times X$, and that $x \sim x'$ if and only if $(x, x') \in E$.

Since X is assumed to be closed and bounded, if we can show that E is closed in $X \times X$, it would follow that E is a definably proper equivalence relation, and we can apply Proposition 4.1.1.

The rest of the proof is devoted to showing that E is a closed definable subset of $X \times X$. For each $\alpha \in I, \beta \in J_\alpha$, let

$$E_{\alpha,\beta} = (\phi_\alpha \times \phi_\alpha)((Y_\alpha \times C_{\alpha,\beta}) \times_{\pi_1} (Y_\alpha \times C_{\alpha,\beta})).$$

Since $E = \bigcup_{\alpha\in I,\beta\in J_\alpha} E_{\alpha,\beta}$, in order to prove that E is closed it suffices to prove that for each $\alpha \in I, \beta \in J_\alpha$,

$$\overline{E_{\alpha,\beta}} \subset E,$$

where $\overline{E_{\alpha,\beta}}$ is the closure of $E_{\alpha,\beta}$ in $X \times X$.

It follows from the curve selection lemma for o-minimal structures [32] that for every $z \in \overline{E_{\alpha,\beta}}$ there exists a definable curve $\gamma : [0,1] \to E_{\alpha,\beta}$ with $\gamma(0) = z$, $\gamma((0,1]) \subset E_{\alpha,\beta}$. Thus, in order to prove that $\overline{E_{\alpha,\beta}} \subset E$, it suffices to show that for each definable curve $\gamma : (0,1] \to E_{\alpha,\beta}$, $z_0 = \lim_{t\to 0} \gamma(t) \in E$.

Let $\gamma : (0,1] \to E_{\alpha,\beta}$ be a definable curve, and suppose that $\lim_{t\to 0} \gamma(0) \notin E_{\alpha,\beta}$. Otherwise, $\lim_{t\to 0} \gamma(0) \in E_{\alpha,\beta} \subset E$, and we are done.

For $t \in (0,1]$, let $y_t = f(\gamma(t))$ and let $(x_t, x'_t) \in (\phi_\alpha \times \phi_\alpha)((Y_\alpha \times C_{\alpha,\beta}) \times_{\pi_1} (Y_\alpha \times C_{\alpha,\beta}))$ be such that $\gamma(t) = (x_t, x'_t)$. Note that $f(x_t) = f(x'_t) = y_t$. Finally, let $z_0 = (x_0, x'_0) = \lim_{t\to 0} \gamma(t)$.

Since, $z_0 \notin E_{\alpha,\beta}$ by assumption and $\gamma((0,1]) \subset E_{\alpha,\beta}$, there exists $t_0 > 0$ such that $\lambda = f \circ \gamma|_{(0,t_0]} : (0,t_0] \to Y_\alpha$ is an injective definable map and $\lim_{t\to 0} \lambda(t) = y_0 = f(x_0) = f(x'_0) \in Y_{\alpha'}$ for some $\alpha' \in I$. We need to show that x_0 and x'_0 belong to the same connected component of $f^{-1}(y_0)$, which would imply that $(x_0, x'_0) \in E$.

Let $D_{\alpha,\beta,\gamma} = f^{-1}(\lambda((0,t_0])) \cap D_{\alpha,\beta}$ and let $g : D_{\alpha,\beta,\gamma} \to (0,t_0]$ be defined by $g(x) = \lambda^{-1}(f(x))$ (which is well defined by the injectivity of λ). Note that for each $t \in (0,t_0]$, $g^{-1}(t)$ is definably homeomorphic to $C_{\alpha,\beta}$, and hence is connected. It also follows from Hardt's triviality theorem that there exists $t'_0 \in (0,t_0]$ and a definable homeomorphism $\theta : g^{-1}(t'_0) \times (0,t'_0] \to g^{-1}((0,t'_0])$ such that the following diagram commutes:

$$\begin{tikzcd}
g^{-1}(t'_0) \times (0,t'_0] \arrow[rr, "\theta"] \arrow[dr, "\pi_2"'] & & g^{-1}((0,t'_0]) \arrow[dl, "g"] \\
& (0,t'_0] &
\end{tikzcd}$$

Extend θ continuously to a definable map $\overline{\theta} : g^{-1}(t'_0) \times [0,t_0] \to \overline{g^{-1}((0,t'_0])}$ by setting $\overline{\theta}(x,0) = \lim_{t\to 0} \theta(x,t)$. Finally, let $\theta' : g^{-1}(t'_0) \to f^{-1}(y_0)$ be the definable map obtained by setting $\theta'(x) = \overline{\theta}(x,0)$.

Note that since $g^{-1}(t'_0)$ is connected, $\theta'(g^{-1}(t'_0))$ is connected as well, since it is the image of a connected set under a continuous map. Also note that for each $t \in (0,t'_0]$, we have that $x_t, x'_t \in D_{\alpha,\beta,\gamma}$ and $f(x,t) = f(x'_t) = \lambda(t)$, hence $x_t, x'_t \in g^{-1}(t)$, and thus $x_0, x'_0 \in \theta'(g^{-1}(t'_0))$.

Moreover, $f(x_0) = f(x'_0) = y_0$. Therefore, since $\theta'(g^{-1}(t'_0))$ is connected, x_0 and x'_0 belong to the same connected component of $f^{-1}(y_0)$.

This shows that $(x_0, x'_0) \in E$, which in turn implies that E is closed in $X \times X$.

The fact that Reeb(f) exists as a definably proper quotient now follows from the following proposition which appears in [31]:

Proposition 4.1.1. *[31, page 166] Let X be a definable set and $E \subset X \times X$ a definably proper equivalence relation on X. Then X/E exists as a definably proper quotient of X.*

□

4.2 A Bound on the Topological Complexity of Reeb Spaces

We now consider the problem of bounding effectively from above the Betti numbers of the Reeb space of a continuous semi-algebraic map. We have seen from Example 1 that, given a continuous semi-algebraic map $f : X \to Y$, $\beta(\text{Reeb}(f))$ can be arbitrarily large compared to $\beta(X)$, unlike in the case of Reeb graphs (i.e. when $\dim(Y) \leq 1$). Because 4.1.1 shows that the Reeb space of a proper semi-algebraic map is indeed a semi-algebraic set, we can make use of results from semi-algebraic geometry to compute a bound on the topological complexity of Reeb(f). In this section, we prove an upper bound on $\beta(\text{Reeb}(f))$ in terms of the "semi-algebraic" complexity of the map f. We present the main result below:

Theorem 4.2.1. *Let $S \subset \mathrm{R}^n$ be a bounded $\mathcal{P}$-closed semi-algebraic set, and $f = (f_1, \ldots, f_m) : S \to \mathrm{R}^m$ be a polynomial map. Suppose that $s = \mathrm{card}(\mathcal{P})$ and the maximum of the degrees of the polynomials in $\mathcal{P}$ and $f_1, \ldots, f_m$ is bounded by d. Then,*

$$\beta(\text{Reeb}(f)) \leq (sd)^{(n+m)^{O(1)}}.$$

4.2.1 Outline of the proof of Theorem 4.2.1

We first replace the map $f : S \to \mathrm{R}^m$, by a new map $\tilde{f} : \tilde{S} \to \mathrm{R}^m$, where $\tilde{S} \subset \mathrm{R}^n \times \mathrm{R}^m$ and $\tilde{f}$ is the restriction to $\tilde{S}$ of the projection map to R^m, such that the following diagram commutes:

$$\begin{tikzcd} \tilde{S} \arrow[rr, hook] \arrow[dr, "\tilde{f}"] & & R^n \times R^m \arrow[dl, "\pi_m"] \\ & R^m & \end{tikzcd}$$

From the definitions it is evident that $\mathrm{Reeb}(f)$ and $\mathrm{Reeb}(\tilde{f})$ are homeomorphic. We next prove that there exists a semi-algebraic partition of R^m of controlled complexity (more precisely given by the connected components of the realizable sign conditions of a family of polynomials of singly exponentially bounded degrees and cardinality) into connected semi-algebraic sets C, such that the connected components of the fibers $\tilde{f}^{-1}(z)$ are in 1-1 correspondence with each other as z varies over C. Moreover, each of these connected components C is described by a quantifier-free first order formula and the complexity of these formulas (i.e. the number of polynomials appearing in the formula and their respective degrees) is bounded singly exponentially (given by Proposition 4.2.2).

Next, we use the fact that the canonical surjection $\phi : \tilde{S} \to \mathrm{Reeb}(\tilde{f})$ is a proper semi-algebraic map. We then use an inequality proved in [33] (see Proposition 4.2.1 below) to obtain an upper bound on the Betti numbers of the image of a proper semi-algebraic map $F : X \to Y$ in terms of the sum of the Betti numbers of various fiber products $X \times_F \cdots \times_F X$ of the same map. Recall that for $p \geq 0$, the $(p+1)$-fold fiber product is given by

$$\underbrace{X \times_F \cdots \times_F X}_{(p+1)\text{-times}} \triangleq \{(x^{(0)}, \ldots, x^{(p)}) \in X^{p+1} \mid F(x^{(0)}) = \cdots = F(x^{(p)})\}.$$

The following proposition proved in [33] allows one to bound the Betti numbers of the image of a closed and bounded definable set X under a definable map F in terms of the Betti numbers of the iterated fibered product of X over F. More precisely:

Proposition 4.2.1. *[33] Let $F : X \to Y$ be a definable continuous map, and X a closed and bounded definable set. Then, for for all $p \geq 0$,*

$$\beta_p(F(X)) \leq \sum_{\substack{i,j \geq 0 \\ i+j=p}} \beta_i(\underbrace{X \times_F \cdots \times_F X}_{(j+1)}).$$

Proposition 4.2.2 provides us with a well controlled description (i.e. by quantifier-free first order formulas involving singly exponentially any polynomials of singly exponentially bounded degrees) of the fibered products $\tilde{S} \times_{\tilde{f}} \cdots \times_{\tilde{f}} \tilde{S}$. Finally, using these descriptions and results on bounding the Betti numbers of general semi-algebraic sets in terms of the number and degrees of polynomials defining them (cf. Proposition 4.2.4 below) we obtain the claimed bound on $\mathrm{Reeb}(f)$.

4.2.2 Technical necessities

Before we proceed to the proof, we need a few preliminary results. We will use the following notation for the rest of this chapter.

Notation 7. *We will denote by $\pi_Y : \mathrm{R}^{k+\ell} \to \mathrm{R}^{\ell}$ the projection to the last ℓ (denoted by $Y = (Y_1, \ldots, Y_\ell)$) coordinates. For any semi-algebraic subset $S \subset \mathrm{R}^{k+\ell}$ and $T \subset \mathrm{R}^{\ell}$, we will denote by $S_T = S \cap \pi_Y^{-1}(T)$. If $T = \{\mathbf{y}\}$, we will write $S_{\mathbf{y}}$ in stead of $S_{\{\mathbf{y}\}}$.*

The following proposition, whose proof can be found in [21], will play a crucial role in the proof of Theorem 4.2.1.

Proposition 4.2.2. *Let R be a real closed field, and let $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k, Y_1, \ldots, Y_\ell]$ be a finite set of polynomials of degrees bounded by d, with $\mathrm{card}(\mathcal{P}) = s$. Let $S \subset \mathrm{R}^k \times \mathrm{R}^\ell$ be a $\mathcal{P}$-semi-algebraic set. Then there exists a finite set of polynomials $\mathcal{Q} \subset \mathrm{R}[Y_1, \ldots, Y_\ell]$ such that $\mathrm{card}(\mathcal{Q})$ and the degrees of polynomials in $\mathcal{Q}$ are bounded by $(sd)^{(k+\ell)^{O(1)}}$, and $\mathcal{Q}$ has the following additional property.*

For each $\sigma \in \mathbf{sign}(\mathcal{Q}) \subset \{0, 1, -1\}^{\mathcal{Q}}$ and $C \in \mathrm{Cc}(\mathcal{R}(\sigma, \mathrm{R}^\ell))$, there exists

(i) an index set $I_{\sigma,C}$,

(ii) a finite family of polynomials $\mathcal{P}_{\sigma,C} \subset \mathrm{R}[X_1, \ldots, X_k, Y_1, \ldots, Y_\ell]$, and

(iii) $\mathcal{P}_{\sigma,C}$-formulas, $(\Theta_\alpha(\overline{X}, \overline{Y}))_{\alpha \in I_{\sigma,C}}$,

such that

1. $\Theta_\alpha(\mathbf{x}, \mathbf{y}) \Rightarrow \mathbf{y} \in C$;

2. for each $\mathbf{y} \in C$ *and each* $D \in \mathrm{Cc}(S_C)$*, there exists a unique* $\alpha \in I_{\sigma,C}$ *such that* $\mathcal{R}(\Theta_\alpha(\cdot, \mathbf{y})) = D_{\mathbf{y}}$ *and* $D_{\mathbf{y}} \in \mathrm{Cc}(S_{\mathbf{y}})$*.*

In order to prove Theorem 4.2.1, we will need singly exponential upper bounds on the Betti numbers of semi-algebraic sets in terms of the number and degrees of the polynomials appearing in any quantifier-free formula defining the set. The first results on bounding the Betti numbers of real varieties were proved by Oleĭnik and Petrovskiĭ [34], Thom [35], and Milnor [36]. Using a Morse-theoretic argument and Bezout's theorem they proved the following proposition which appears in [37] and makes more precise an earlier result which appeared in [38]:

Proposition 4.2.3. *[37] If* $S \subset \mathrm{R}^k$ *is a* $\mathcal{P}$*-closed semi-algebraic set, then*

$$\beta(S) \leq \sum_{i=0}^{k} \sum_{j=0}^{k-i} \binom{s+1}{j} 6^j d(2d-1)^{k-1}, \tag{4.1}$$

where $s = \mathrm{card}(\mathcal{P}) > 0$ *and* $d = \max_{P \in \mathcal{P}} \deg(P)$*.*

Using a technique to replace an arbitrary semi-algebraic set by a locally closed one with a very controlled increase in the number of polynomials used to describe the given set, Gabrielov and Vorobjov [39] extended Proposition 4.2.3 to arbitrary $\mathcal{P}$-semi-algebraic sets with only a small increase in the bound. Their result, in conjunction with Proposition 4.2.3, gives the following proposition.

Proposition 4.2.4. *[1], [40] If* $S \subset \mathrm{R}^k$ *is a* $\mathcal{P}$*-semi-algebraic set, then*

$$\beta(S) \leq \sum_{i=0}^{k} \sum_{j=0}^{k-i} \binom{2ks+1}{j} 6^j d(2d-1)^{k-1}, \tag{4.2}$$

where $s = \mathrm{card}(\mathcal{P})$ *and* $d = \max_{P \in \mathcal{P}} \deg(P)$*.*

We will also use the following bound on the number of connected components of the realizations of all realizable sign conditions of a family of polynomials proved in [37].

Proposition 4.2.5. *Let $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]_{\leq d}$ and let $s = \mathrm{card}(\mathcal{P})$. Then*

$$\mathrm{card}\left(\bigcup_{\sigma \in \mathbf{sign}(\mathcal{P})} \mathrm{Cc}(\mathcal{R}(\sigma, \mathrm{R}^k))\right) \leq \sum_{1 \leq j \leq k} \binom{s}{j} 4^j d(2d-1)^{k-1}.$$

With these preliminary results in place, we proceed to the proof of Theorem 4.2.1:

4.2.3 Proof of Theorem 4.2.1

Proof of Theorem 4.2.1. Let Φ be the $\mathcal{P}$-closed formula defining S. Introducing new variables $Z_1, \ldots, Z_m$, let $\tilde{S} \subset \mathrm{R}^n \times \mathrm{R}^m$ be the $\tilde{\mathcal{P}}$-formula

$$\Phi \wedge \bigwedge_{1 \leq i \leq m} (Z_i - f_i = 0).$$

Let $\tilde{f} : \tilde{S} \to \mathrm{R}^m$ denote the restriction to $\tilde{S}$ of the projection map $\pi_Z : \mathrm{R}^m \times \mathrm{R}^n \to \mathrm{R}^m$ to the Z-coordinates. Then clearly S is semi-algebraically homeomorphic to $\tilde{S}$, $f(S) = \tilde{f}(\tilde{S})$, and $\mathrm{Reeb}(f)$ is semi-algebraically homeomorphic to $\mathrm{Reeb}(\tilde{f})$. We have the following commutative square where the horizontal arrows are homeomorphisms and the vertical arrows are the quotient maps.

$$\begin{tikzcd} S \arrow[r, "\cong"] \arrow[d, "\phi"] & \tilde{S} \arrow[d, "\tilde{\phi}"] \\ \mathrm{Reeb}(f) \arrow[r, "\cong"] & \mathrm{Reeb}(\tilde{f}) \end{tikzcd}$$

Now it follows from Proposition 4.2.2 that there exists a finite set of polynomials $\mathcal{Q} \subset \mathrm{R}[Z_1, \ldots, Z_m]$, with

$$\mathrm{card}(\mathcal{Q}), \max_{Q \in \mathcal{Q}} \deg(Q) \quad \leq \quad (sd)^{(n+m)^{O(1)}} \tag{4.3}$$

having the following property: for each $\sigma \in \mathbf{sign}(\mathcal{Q})$ and each $C \in \mathrm{Cc}(\mathcal{R}(\sigma, \mathrm{R}^m))$, there exists an index set $I_{\sigma,C}$, a finite family of polynomials

$$\mathcal{P}_{\sigma,C} \subset \mathrm{R}[X_1, \ldots, X_n, Z_1, \ldots, Z_m],$$

and $\mathcal{P}_{\sigma,C}$ formulas $(\Theta_\alpha(\overline{X},\overline{Z}))_{\alpha\in I_{\sigma,C}}$ such that $\Theta_\alpha(x,z)\Rightarrow z\in C$, and for each $z\in C$, and each connected component D of $\pi_Z^{-1}(C)\cap\tilde{S}$, there exists a unique $\alpha\in I_{\sigma,C}$ (which does not depend on z) with $\mathcal{R}(\Theta_\alpha(\cdot,z))=\pi_Z^{-1}(z)\cap D$.

Moreover, the cardinalities of $I_{\sigma,C}$ and $\mathcal{P}_{\sigma,C}$ and the degrees of the polynomials in $\mathcal{P}_{\sigma,C}$ are all bounded by $(sd)^{(n+m)^{O(1)}}$.

Let ϕ (resp. $\tilde{\phi}$) be the canonical surjection $\phi : S \to \mathrm{Reeb}(f)\cong S/\sim$ (resp. $\tilde{\phi}:\tilde{S}\to\mathrm{Reeb}(\tilde{f})\cong\tilde{S}/\sim$). From Theorem 4.1.1 it follows that we can assume that ϕ is a proper semi-algebraic map. For each $i\geq 0$, we have the inequality (cf. Proposition 4.2.1)

$$\beta_i(\mathrm{Reeb}(f)) \leq \sum_{p+q=i}\beta_q(\underbrace{S\times_\phi\cdots\times_\phi S}_{(p+1)\text{ times}}). \tag{4.4}$$

Now observe that $\underbrace{\tilde{S}\times_{\tilde{\phi}}\cdots\times_{\tilde{\phi}}\tilde{S}}_{(p+1)\text{ times}}$ (and hence $\underbrace{S\times_\phi\cdots\times_\phi S}_{(p+1)\text{ times}}$) is semi-algebraically homeomorphic to the semi-algebraic set defined by the formula

$$\Theta(\overline{X}^{(0)},\ldots,\overline{X}^{(p)},\overline{Z}) = \bigvee_{\substack{\sigma\in\mathbf{sign}(\mathcal{Q})\\ C\in\mathrm{Cc}(\mathcal{R}(\sigma,\mathrm{R}^m))\\ \alpha\in I_{\sigma,C}}}\ \bigwedge_{0\leq j\leq p}\Theta_\alpha(\overline{X}^{(j)},\overline{Z}). \tag{4.5}$$

To see this observe that

$$((x^{(0)},z^{(0)}),\ldots,(x^{(p)},z^{(p)}))\in\underbrace{\tilde{S}\times_{\tilde{\phi}}\cdots\times_{\tilde{\phi}}\tilde{S}}_{(p+1)\text{ times}}$$

if and only if

$$z^{(0)}=\cdots=z^{(p)}=z,$$

for some z, and $x^{(0)},\ldots,x^{(p)}$ belong to the same connected component of $\tilde{f}^{-1}(z)$.

It is easy to verify this last equivalence using the properties of the decomposition given by Proposition 4.2.2.

We now claim that each of the formulas

$$\Theta(\overline{X}^0, \ldots, \overline{X}^{(p)}, \overline{Z}),\ 0 \leq p \leq m,$$

is a $\tilde{\mathcal{P}}_p$-formula for some finite set $\tilde{\mathcal{P}}_p \subset \mathrm{R}[\overline{X}^0, \ldots, \overline{X}^{(p)}, \overline{Z}]$ with $\mathrm{card}(\tilde{\mathcal{P}}_p)$ and the degrees of the polynomials in $\tilde{\mathcal{P}}_p$ being bounded singly exponentially.

In order to prove the claim first observe that the cardinality of the set

$$\bigcup_{\sigma \in \mathbf{sign}(\mathcal{Q})} \mathrm{Cc}(\mathcal{R}(\sigma, \mathrm{R}^m))$$

is bounded singly exponentially, once the number of polynomials in $\mathcal{Q}$, and their degrees are bounded singly exponentially (using Proposition 4.2.5). The fact that the number of polynomials in $\mathcal{Q}$ and their degrees are bounded singly exponentially follows from (4.3). Moreover, for similar reasons the cardinalities of the index sets $I_{\sigma,C}$ are also bounded singly exponentially. The claim now follows from Eqn. (4.5).

Finally, to prove the theorem we first apply inequality (4.4) and then apply Proposition 4.2.4 to bound the right hand side of the inequality (4.4). $\square$

5. AN EFFICIENT ALGORITHM FOR THE COMPUTATION OF REEB SPACES FROM ROADMAPS

The field of algorithmic semi-algebraic geometry is rich and well-studied. Born out of a desire to develop an algorithm to count the number of real roots of a real-valued polynomial, algorithmic semi-algebraic geometry been the area of exploration for over 350 years. Algorithmic problems in semi-algebraic geometry generally take as input a finite family of polynomials and formulas defining a semi-algebraic set S, and either decide whether or not certain topological properties hold for S, or compute topological invariants of S. Many algorithms in semi-algebraic geometry rely on cylindrical algebraic decomposition (CAD), a technique with doubly-exponential algorithmic complexity in terms of the number of polynomials defining S and their degrees.

In fact, one could apply general the quotienting algorithm, which would rely on CAD which has doubly exponential complexity, to yield a description of the Reeb space with doubly exponential complexity. The point of much research in other problems has been to obtain more refined algorithms with singly exponential complexity, for example, computing descriptions of connected components or Euler characteristics of semi-algebraic sets. In a similar direction, in this paper, we are designing an algorithm with singly exponential complexity for computing Reeb graphs.

Because there is a meta theorem in algorithmic semi-algebraic geometry relating upper bounds on topological complexity of semi-algebraic sets with worst-case complexity of algorithms to compute their topological invariants, the singly-exponential upper bound in Theorem 4.2.1 raises the possibility of constructing an algorithm with singly-exponential complexity to compute a semi-algebraic description of the Reeb space the Reeb space of a semi-algebraic map. We begin to prove this result by presenting an algorithm with singly exponential complexity to compute the Reeb graph of a semi-algebraic map.

Theorem 5.0.1. *There is an algorithm that takes as input a family $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ of polynomials and formulas describing a semi-algebraic set and map f and computes as output a semi-algebraic description of the Reeb graph with complexity $s^{k+1}d^{\mathcal{O}(k^2)}$ where s is a bound on the number of polynomials in $\mathcal{P}$ and d is a bound on their degree.*

5.1 Example

We demonstrate how our algorithm works with the following examples.

Example 2. *Consider the curves show in Figure 5.1. Starting with Figure 5.1, we run algorithm 9 and obtain Figure 5.2 as an intermediate step.*

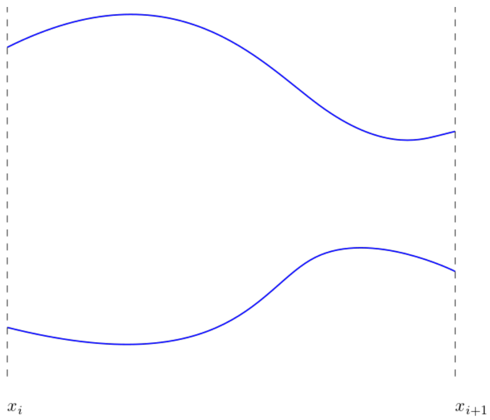


Figure 5.1.

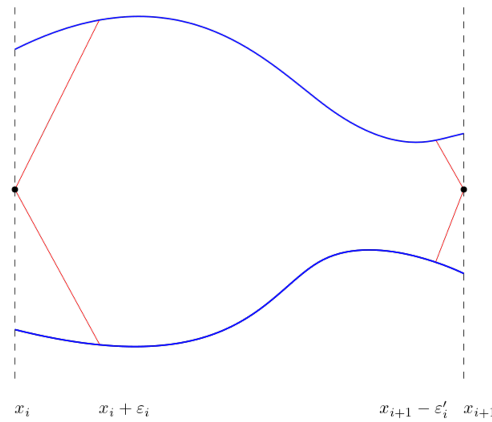


Figure 5.2.

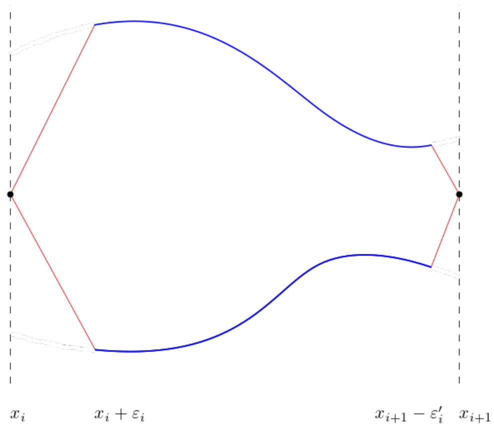
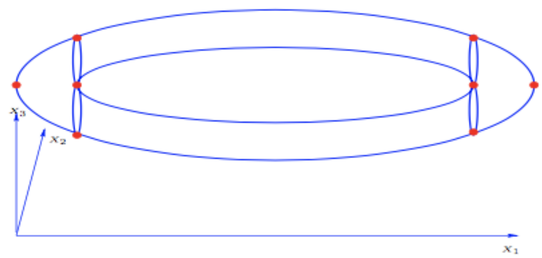


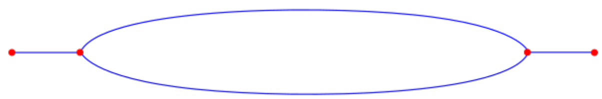
Figure 5.3.

The output of the algorithm is Figure 5.3.

Example 3. *Consider a roadmap of the torus given in [1] corresponding to the projection map π onto the first coordinate.*



The Reeb graph $\mathrm{Reeb}(\pi)$ *is the set*



By choosing one curve per connected component of the fiber of each point on X_1*, as shown in Figure 5.4 and joining them at their endpoints as in Figure 5.5, we obtain a set homeomorphic to* $\mathrm{Reeb}(\pi)$*. In the case of the roadmap of the torus, this subset in Figure 5.4 is not unique; any subset satisfying this criterion will produce the Reeb graph after joining the curves in this manner.*

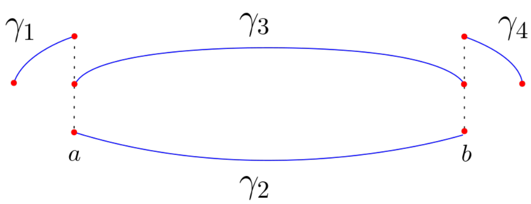


Figure 5.4.

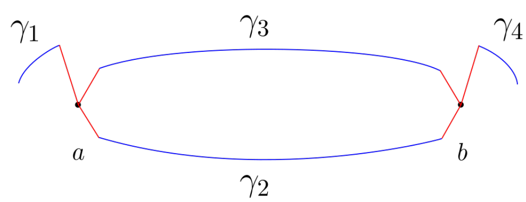


Figure 5.5.

5.2 Algorithm to Compute the Reeb Graph

We begin by defining Puiseux series, a commonly used tool in algorithmic semi-algebraic geometry.

Definition 5.2.1. *A* Puiseux series in ε with coefficients in R *is a series of the form* $\overline{r} = \sum_{i \geq k} r_i \varepsilon^{i/q}$ *with* $k, i \in \mathbb{Z}$*,* $q \in \mathbb{N}$*, and* $r_i \in \mathrm{R}$*. We denote the set of algebraic Puiseux series in* ε *with coefficients in* R *as* $\mathrm{R}\langle\varepsilon\rangle$*. This is a real closed field whose unique order is the one whose restriction to* $\mathrm{R}(\varepsilon)$ *is* $>_{0^+}$*.*

We can now consider the field $\mathrm{R}(\varepsilon) = \left\{ \frac{f(\varepsilon)}{g(\varepsilon)} \mid f, g \in \mathrm{R}[\varepsilon] \right\}$ with order $>_{0^+}$, defined on polynomials in $\mathrm{R}(\varepsilon)$, as follows: $f > 0$ if and only if it is positive to the right of the origin, and $f < 0$ if and only if it is negative to the right of the origin. Equivalently, if $f(\varepsilon) = a_p \varepsilon^p + \ldots + a_n \varepsilon^n$, $a_i \neq 0$, then $f(\varepsilon) > 0 \iff a_p > 0$. In particular, let $f = a - \varepsilon$ for any $a > 0$. Then $f > 0$, and hence $a > \varepsilon$, so ε is indeed smaller than every element in R.

Definition 5.2.2. *We define a variable $\varepsilon > 0$ to be* infinitesimal *if $\varepsilon < r$ for any $r \in \mathrm{R}$.*

Because the field $\mathrm{R}(\varepsilon)$ is not a real closed field (there is no unique order on it), we work with the field $\mathrm{R}\langle\varepsilon\rangle$, Puiseux series in ε.

Our algorithm makes use of a generalized version [31] of Hardt's triviality theorem [41]:

Theorem 5.2.1. *Let $S \subset \mathrm{R}^k$ and $T \subset \mathrm{R}^m$ be semi-algebraic sets. Given a continuous semi-algebraic function $f \colon S \to T$, there exists a finite partition of T into semi-algebraic sets $T = \bigcup_{i=1}^{r} T_i$, so that for each i and any $x_i \in T_i$, $T_i \times f^{-1}(x_i)$ is semi-algebraically homeomorphic to $f^{-1}(T_i)$.*

Definition 5.2.3. *Consider a finite partition of $[0, 1]$ into intervals of the form (x_i, x_{i+1}), $0 = x_0 < x_1 < \ldots x_{n-1} < x_n = 1$. Suppose we are given a a finite set Γ of semi-algebraic continuous maps $\{\gamma_k\}$ where each $\gamma_k \colon (x_i, x_{i+1}) \to \mathrm{R}^k$ is parametrized by X_1. Suppose also that for each curve γ_k, we are given two points, z_{k_i} and $z_{k_{i+1}}$, whose first coordinates are equal to the first coordinates of the left and right endpoints of the image of γ_k. Let U be the union of the images of the curves in Γ. Define an equivalence relation $\sim$ on points in U by letting $\gamma_k(x_i) \sim \gamma_j(x_i)$ if and only if $z_{k_i} = z_{j_i}$. Letting Z denote the set $\{z_{k_i}\}$, the set $U/\sim$, denoted $\mathcal{G}_{U_Z}$, is called the gluing of U with respect to Z, and will play an important role in the proof of Theorem 5.0.1.*

By Hardt's Triviality Theorem, there exists an $r_0 > 0$ such that the images of any two curves $\gamma_1(t)$ and $\gamma_2(t)$ do not intersect on $(0, r_0)$. By definition, $\varepsilon < r_0$. Thus no two curves intersect on the intervals $(x_i - \varepsilon, x_i)$ and $(x_i, x_i + \varepsilon)$ for all x_i.

Algorithm 9 (Gluing)

Input:

1. A finite partition of $[0,1]$ into intervals of the form (x_i, x_{i+1}), $0 = x_0 < x_1 < \ldots x_{n-1} < x_n = 1$;
2. a set U and a set Z as described in Definition 5.2.3.

Output: A semi-algebraic set $\tilde{U}$ of dimension 1 over $\mathrm{R}\langle\varepsilon\rangle$ and a map $p\colon \tilde{U} \to \mathrm{R}\langle\varepsilon\rangle$ such that the following diagram commutes:

$$\begin{tikzcd} \mathcal{G}_{S_Z} \arrow[r, "\sim"] \arrow[d] & \tilde{U} \arrow[dl] \\ \mathrm{R}\langle\varepsilon\rangle & \end{tikzcd}$$

Complexity: $\mathrm{card}(\Gamma) d^{O(k)}$, where d is a bound on the degree of the curve segments in Γ.

Procedure:

for each each $\gamma_j \in \Gamma$ with endpoints x_i, x_{i+1} **do**,
    construct a line L_{j_1} between z_i and $\gamma_j(x_i + \varepsilon)$.
    construct a line L_{j_2} between z_{i+1} and $\gamma_j(x_{i+1} - \varepsilon)$.
end for
if n line segments $L_1, L_2, \ldots, L_n$ intersect at a point p **then**
    construct balls $B^2(p, \varepsilon), B^2(p, 2\varepsilon), \ldots, B^2(p, (n-1)\varepsilon)$.
    for each k in $1, \ldots, n$ **do**
        $L_k \leftarrow L_k \setminus B^2(p, k\varepsilon) \cup \{(x, y, z) \mid (x, y) \in L_k \cap B^2(p, k\varepsilon) \wedge z = \sqrt{(k\varepsilon)^2 - x^2 - y^2}\}$.
    end for
end if
$\gamma'_j \leftarrow L_{j_1} \cup \{(x, y) \mid (x, y) \in \gamma_j \wedge x \in (x_i + \varepsilon, x_{i+1} - \varepsilon)\} \cup L_{j_2}$.
$\tilde{U} \leftarrow \bigcup \gamma'_j$.

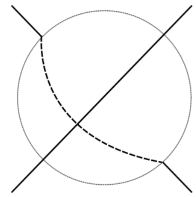


Figure 5.6. Re-routing lines around a ball, as described in Step 8 in Algorithm 9

Proof of correctness. First, we show that there exists a homeomorphism $f\colon \mathcal{G}_{S_Z} \to \tilde{U}$. Consider the diagram

$$\begin{tikzcd}
U \arrow[d, "q"] \arrow[dr, "g"] & \\
\mathcal{G}_{S_Z} \arrow[r, "f"] \arrow[d, "s"] & \tilde{U} \arrow[dl, "p"] \\
\mathrm{R}\langle\varepsilon\rangle &
\end{tikzcd}$$

Define the map $g\colon U \to \tilde{U}$ as follows: consider a point $(x, y) \in \gamma_i$ where $x \in \mathrm{R}$ and $y \in \mathrm{R}^{k-1}$. Define $g(x, y) = (x, y')$, where y' is the unique point such that $(x, y') \in \tilde{U}$; uniqueness is guaranteed by condition 3 of the curves being glueable.

To define the map $f\colon \mathcal{G}_{U_Z} \to \tilde{U}$, let $(x, y) \in \mathcal{G}_{U_Z}$ where $x \in \mathrm{R}$ and $y \in \mathrm{R}^{k-1}$. Let $q\colon U \to \mathcal{G}_{U_Z}$ be the quotient map. Pick an arbitrary point $(x', y') \in U$ that is contained in the preimage of (x, y) under q. Then $f(x, y) = g(x', y')$. This is well defined because for any two points (x', y'), (x'', y'') in the preimage of (x, y) under q, $x' = x''$ so $g(x', y') = g(x'', y'')$. This construction also guarantees that f is surjective. Similarly, f^{-1} is well-defined: for any point $(x, y) \in \tilde{U}$, the map q sends all points in the preimage of (x, y) under g to the same point $(x', y') \in \mathcal{G}_{U_Z}$. Let $U \subseteq \tilde{U}$ be a basic open set. Then $U = B^k \cap g(\gamma_i)$ for some $\gamma_i \in \Gamma$. Then $f^{-1}(U)$ is of the form $B^k \cap q(\gamma_i)$ because f^{-1} acts as the identity on the first coordinate. Therefore $f^{-1}(U)$ is open, and hence f^{-1} is continuous. To show that f is a homeomorphism, it remains to show that f is bijective. Note that g is surjective because each point $(x, y') \in \tilde{U}$ corresponds with exactly one point $(x, y) \in S$ by construction.

Define $s\colon \mathcal{G}_{U_Z} \to \mathrm{R}\langle\varepsilon\rangle$ and $p\colon \tilde{U} \to \mathrm{R}\langle\varepsilon\rangle$ to be projection maps onto the first coordinate. Because the map f acts as the identity on the first coordinate, the lower triangle commutes.

By construction, it is straightforward to see that the resulting set is dimension 1. Because the input curves are semi-algebraic sets, the truncated curves are as well. Moreover, since lines are semi-algebraic sets, the finite union of the lines with the semi-algebraic curves is a semi-algebraic set.

□

In Example 3, it was shown that by choosing a particular subset of curves of a roadmap of a function and applying Algorithm 9, we can obtain the Reeb graph of that function. The following algorithm includes a procedure to select curves and produce an equivalence relation on them so that after applying Algorithm 9, our Algorithm will output a semi-algebraic description of the Reeb graph of the given function.

Our algorithm works as follows: let S be a semi-algebraic set and let $\pi\colon S \to \mathrm{R}$ be the projection map. From Algorithm 15.12 in [1], we obtain a semi-algebraic description Γ of the curves in the roadmap of π. We then obtain a subset U of Γ, which we then equip with an equivalence relation using Algorithm 9 to obtain a set Γ homeomorphic to the Reeb graph $\mathrm{Reeb}(\pi)$ and a map p which makes the following diagram commute, where s is the projection map onto the first coordinate:

$$\begin{tikzcd}
\mathrm{Reeb}(\pi) \arrow[r, "f"] \arrow[d, "s"] & \Gamma'' \arrow[dl, "t"] \\
\mathrm{R}\langle\varepsilon\rangle &
\end{tikzcd}$$

The following algorithm proves Theorem 5.0.1.

Algorithm 10 (Construction of Reeb graphs)

Input:

A family of polynomials $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ and a $\mathcal{P}$-semi-algebraic set S.

Output:

A semi-algebraic set Γ of dimension 1 over $\mathrm{R}\langle\varepsilon\rangle$ homeomorphic to $\mathrm{Reeb}(\pi)$ and a map $p\colon \Gamma \to \mathrm{R}\langle\varepsilon\rangle$ such that the diagram,

$$\begin{tikzcd}
\mathrm{Reeb}(\pi) \arrow[r, "r"] \arrow[d] & \Gamma \arrow[dl, "p"] \\
\mathrm{R}\langle\varepsilon\rangle &
\end{tikzcd}$$

where r is a homeomorphism and s denotes projection onto the first coordinate, commutes.

Complexity: The complexity of the algorithm is dominated by the complexity of the roadmap algorithm in [1], which has complexity $s^{k+1}d^{\mathcal{O}(k^2)}$ where s is a bound on the number of polynomials in $\mathcal{P}$ and d is a bound on their degree.

Procedure:

1: Call Algorithm 16.26 (General Roadmap) with input S to obtain a set of curve segments $\{\mathcal{C}_i\}$, critical points $\{u_j\}_{j=1}^{\ell}$, and a set Z as described in Definition 5.2.3.

2: For each $j = 1, \ldots, \ell - 1$, select one $x_j \in (u_j, u_{j+1})$, where each u_i is the point associated to $\mathcal{D}_j$.

3: For each $\pi^{-1}(x_j) \cap U \subset S$, and select a single point $p \in \pi^{-1}(x_j) \cap U$. Denote by P the collection of all such points p, and let Γ' be the subset of U of curves containing some point $p \in P$.

4: Perform Algorithm 9 with inputs Γ' and the points output in step 3 to obtain a gluing Γ of the curves represented by the $\mathcal{C}_j$ and a map $p\colon \Gamma \to \mathrm{R}\langle\varepsilon\rangle$.

Proof of correctness. We begin by showing that $\mathrm{Reeb}(\pi) \cong \mathcal{G}_{\Gamma'_Z}$. Consider the diagram

$$\begin{tikzcd}
S \arrow[d, "r"] & \Gamma' \arrow[l, "\iota"'] \arrow[d, "q"] \\
\mathrm{Reeb}(\pi) \arrow[r, "f"] & \mathcal{G}_{\Gamma'_Z}
\end{tikzcd}$$

where $r\colon S \to \mathrm{Reeb}(\pi)$ is the quotient map. We define the map $f\colon \mathrm{Reeb}(\pi) \to \mathcal{G}_{\Gamma'_Z}$ as follows: let $(x,y) \in \mathrm{Reeb}(\pi)$, where $x \in \mathrm{R}$ and $y \in \mathrm{R}^{k-1}$. Pick an element $(x,y') \in r^{-1}(x,y)$. By construction, this element is contained in $\pi^{-1}(x)$. Let p the point in step 4 such that $p \in \pi^{-1}(x) \cap U$. Then $p \in \Gamma'$ by definition, so let $f(x,y) = q(p)$. This is well defined since any element in $p^{-1}(x,y)$ is in the same connected component of $\pi^{-1}(x) \in S$, and hence will get mapped to the same $p \in \Gamma'$.

To show that f is continuous, let $U \subseteq \mathcal{G}_{S_Z}$ be a basic open set. Then $U = B^k \cap q(\gamma_i)$ for some $\gamma_i \in \Gamma'$. Because f^{-1} acts as the identity on the first coordinate, $f^{-1}(U)$ is of the form $B^k \cap \mathrm{Reeb}(\pi)$. Therefore $f^{-1}(U)$ is open, and hence f^{-1} is continuous.

To show that f is injective, suppose that $(x,y),(x',y') \in \mathrm{Reeb}(\pi)$ with $f(x,y) = f(x',y')$. Because f acts as the identity on the first coordinate, $x = x'$. Let p, p' be the points corresponding to (x,y) and (x',y') in the construction of f. First, suppose that $x = x'$ are not critical points. If $p \neq p'$, then p and p' are not in the same connected component of $\pi^{-1}(x)$, and hence not equal in $\mathrm{Reeb}(\pi)$, a contradiction. Now, suppose that $x = x'$ are not critical points. Then $p = p'$, since q is injective outside of critical points. Thus $(x,y) = (x',y')$.

Lastly, to show that f is surjective, let $(x,y) \in \mathcal{G}_{\Gamma'_Z}$. By surjectivity of q, there is some point $p \in \Gamma'$ such that $q(p) = (x,y)$, which can be expressed as some $(x,y') \in S$. Then $p(x,y') = (x,y'') \in \mathrm{Reeb}(\pi)$. Thus, by construction of f, $f(x,y'') = (x,y)$, as desired.

By Algorithm 9, the diagram

$$\begin{tikzcd} \mathcal{G}_{\Gamma'_Z} \arrow[r, "f"] \arrow[d, "s'"] & \Gamma \arrow[dl, "p"] \\ \mathrm{R}\langle\varepsilon\rangle & \end{tikzcd}$$

commutes. Because $\mathrm{Reeb}(\pi) \cong \mathcal{G}_{\Gamma'_Z}$, the diagram

$$\begin{tikzcd} \mathrm{Reeb}(\pi) \arrow[r, "g\circ f"] \arrow[d, "s"] & \Gamma \arrow[dl, "p"] \\ \mathrm{R}\langle\varepsilon\rangle & \end{tikzcd}$$

commutes and $g \circ f$ is a homeomorphism, as desired.

$\square$

Complexity analysis. The complexity of the algorithm is dominated by the complexity of the roadmap algorithms, which has complexity $s^{k+1}d^{\mathcal{O}(k^2)}$ [1], where s is a bound on the number of polynomials in $\mathcal{P}$ and d is a bound on their degree. $\square$

Algorithm 10 can be generalized to compute Reeb spaces using similar ideas, but the complexity is doubly exponential in n.

REFERENCES

[1] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in Real Algebraic Geometry*. Springer-Verlag, 2011.

[2] D. Y. Grigor'ev and N. Vorobjov, "Solving systems of polynomial inequalities in subexponential time," *Journal of Symbolic Computation*, vol. 5, no. 1, pp. 37–64, 1988, ISSN: 0747-7171. DOI: https://doi.org/10.1016/S0747-7171(88)80005-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747717188800051.

[3] J. F. Canny, "Some algebraic and geometric computations in pspace," EECS Department, University of California, Berkeley, Tech. Rep. UCB/CSD-88-439, Aug. 1988. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1988/6041.html.

[4] J. Heintz, M.-F. Roy, and P. Solerno, "Sur la complexité du principe de tarski-seidenberg," fre, *Publications mathématiques et informatique de Rennes*, vol. 309, no. 4, pp. 103–120, 1989. [Online]. Available: http://eudml.org/doc/274810.

[5] J. Renegar, "On the computational complexity and geometry of the first-order theory of the reals. part iii: Quantifier elimination," *Journal of Symbolic Computation*, vol. 13, no. 3, pp. 329–352, 1992, ISSN: 0747-7171. DOI: https://doi.org/10.1016/S0747-7171(10)80005-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747717110800057.

[6] S. Basu and N. Karisani, *Efficient simplicial replacement of semi-algebraic sets and applications*, 2020. arXiv: 2009.13365 `[math.AT]`.

[7] S. Basu and S. Percival, "Efficient computation of a semi-algebraic basis of the first homology group of a semi-algebraic set," submitted to the ACM-SIAM Symposium on Discrete Algorithms.

[8] G. Reeb, "Sur les points singuliers d'une forme de pfaff complètement intégrable ou d'une fonction numérique," *Comptes Rendus de l'Académie des Sciences*, vol. 222, pp. 847–849, 1946.

[9] A. Patel, "Reeb spaces and the robustness of preimages," 2010.

[10] O. Saeki, *Reeb spaces of smooth functions on manifolds*, 2020. arXiv: 2006.01689 `[math.GT]`.

[11] N. Kitazawa, *Global topologies of reeb spaces of stable fold maps with non-trivial top homology groups*, 2021. arXiv: 2105.12934 `[math.GN]`.

[12] O. Burlet and G. de Rham, “Sur certaines applications génériques d’une variété sur certaines applications génériques d’une variété close à 2 dimensions dans le plan,” *L’Ensiegnement Mathématique*, vol. 20, pp. 275–292, 1974.

[13] H. Edelsbrunner, J. Harer, and A. K. Patel, “Reeb spaces of piecewise linear mappings,” in *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, ser. SCG ’08, College Park, MD, USA: ACM, 2008, pp. 242–250, ISBN: 978-1-60558-071-5. DOI: 10.1145/1377676.1377720. [Online]. Available: http://doi.acm.org/10.1145/1377676.1377720.

[14] Y. Shinagawa, T. L. Kunii, and Y. L. Kergosien, “Surface coding based on morse theory,” *IEEE Computer Graphics and Applications*, vol. 11, no. 5, pp. 66–78, 1991.

[15] G. Singh, F. Mémoli, and G. Carlsson, “Topological methods for the analysis of high dimensional data sets and 3d object recognition,” in *PBG@Eurographics*, 2007.

[16] E. Munch and B. Wang, “Convergence between Categorical Representations of Reeb Space and Mapper,” *ArXiv e-prints*, Dec. 2015. arXiv: 1512.04108 [cs.CG].

[17] M. Nicolau, A. J. Levine, and G. Carlsson, “Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 17, pp. 7265–7270, 2011, ISSN: 0027-8424. DOI: 10.1073/pnas.1102826108. eprint: https://www.pnas.org/content/108/17/7265.full.pdf. [Online]. Available: https://www.pnas.org/content/108/17/7265.

[18] M. Carrière and B. Michel, *Approximation of reeb spaces with mappers and applications to stochastic filters*, 2019. arXiv: 1912.10742 [math.AT].

[19] H. Doraiswamy and V. Natarajan, *Computing reeb graphs as a union of contour trees*.

[20] J. Tierny and H. Carr, “Jacobi fiber surfaces for bivariate reeb space computation,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 960–969, 2017. DOI: 10.1109/TVCG.2016.2599017.

[21] S. Basu, N. Cox, and S. Percival, “On the Reeb spaces of definable maps,” *arXiv e-prints*, Apr. 2018. arXiv: 1804.00605 [math.AT].

[22] J. Canny, *The Complexity of Robot Motion Planning*. MIT Press, 1987.

[23] N. Bourbaki, *Éléments de mathématique. Algèbre. Chapitre 10. Algèbre homologique*. Springer-Verlag, Berlin, 2007, pp. viii+216, Reprint of the 1980 original [Masson, Paris; MR0610795].

[24] J. Canny, “Computing road maps in general semi-algebraic sets,” *The Computer Journal*, vol. 36, pp. 504–514, 1993.

[25] L. Gournay and J. J. Risler, “Construction of roadmaps of semi-algebraic sets,” *Appl. Algebra Eng. Commun. Comput.*, vol. 4, no. 4, pp. 239–252, 1993.

[26] D. Grigoriev and N. Vorobjov, “Counting connected components of a semi-algebraic set in subexponential time,” *Comput. Complexity*, vol. 2, no. 2, pp. 133–186, 1992.

[27] J. Heintz, M.-F. Roy, and P. Solernò, “Single exponential path finding in semi-algebraic sets ii: The general case,” in *Algebraic geometry and its applications*, C. L. Bajaj, Ed., Shreeram S. Abhyankar’s 60th birthday conference, 1990, Springer-Verlag, 1994, pp. 449–465.

[28] S. Basu, R. Pollack, and M.-F. Roy, “Computing roadmaps of semi-algebraic sets on a variety,” *J. Amer. Math. Soc.*, vol. 13, no. 1, pp. 55–82, 2000, ISSN: 0894-0347.

[29] K. Cole-McLaughlin, H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, “Loops in reeb graphs of 2-manifolds,” *Discrete & Computational Geometry*, vol. 32, no. 2, pp. 231–244, Jul. 2004.

[30] H. Edelsbrunner and J. L. Harer, *Computational topology*. American Mathematical Society, Providence, RI, 2010, pp. xii+241, An introduction, ISBN: 978-0-8218-4925-5.

[31] L. van den Dries, *Tame topology and o-minimal structures*, ser. London Mathematical Society Lecture Note Series. Cambridge: Cambridge University Press, 1998, vol. 248, pp. x+180, ISBN: 0-521-59838-9.

[32] M. Coste, *An introduction to o-minimal geometry*. Pisa: Istituti Editoriali e Poligrafici Internazionali, 2000, Dip. Mat. Univ. Pisa, Dottorato di Ricerca in Matematica.

[33] A. Gabrielov, N. Vorobjov, and T. Zell, “Betti numbers of semialgebraic and sub-Pfaffian sets,” *J. London Math. Soc. (2)*, vol. 69, no. 1, pp. 27–43, 2004, ISSN: 0024-6107.

[34] I. G. Petrovskiĭ and O. A. Oleĭnik, “On the topology of real algebraic surfaces,” *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, vol. 13, pp. 389–402, 1949, ISSN: 0373-2436.

[35] R. Thom, “Sur l’homologie des variétés algébriques réelles,” in *Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse)*, Princeton, N.J.: Princeton Univ. Press, 1965, pp. 255–265.

[36] J. Milnor, “On the Betti numbers of real varieties,” *Proc. Amer. Math. Soc.*, vol. 15, pp. 275–280, 1964, ISSN: 0002-9939.

[37] S. Basu, R. Pollack, and M.-F. Roy, "On the Betti numbers of sign conditions," *Proc. Amer. Math. Soc.*, vol. 133, no. 4, 965–974 (electronic), 2005, ISSN: 0002-9939.

[38] S. Basu, "On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets," *Discrete Comput. Geom.*, vol. 22, no. 1, pp. 1–18, 1999, ISSN: 0179-5376.

[39] A. Gabrielov and N. Vorobjov, "Betti numbers of semialgebraic sets defined by quantifier-free formulae," *Discrete Comput. Geom.*, vol. 33, no. 3, pp. 395–401, 2005, ISSN: 0179-5376.

[40] A. Gabrielov and N. Vorobjov, "Approximation of definable sets by compact families, and upper bounds on homotopy and homology," *J. Lond. Math. Soc. (2)*, vol. 80, no. 1, pp. 35–54, 2009, ISSN: 0024-6107. DOI: 10.1112/jlms/jdp006. [Online]. Available: http://dx.doi.org/10.1112/jlms/jdp006.

[41] R. M. Hardt, "Semi-algebraic local-triviality in semi-algebraic mappings," *American Journal of Mathematics*, vol. 102, no. 2, pp. 291–302, 1980, ISSN: 00029327, 10806377. [Online]. Available: http://www.jstor.org/stable/2374240.