ALGORITHMIC SEMI-ALGEBRAIC GEOMETRY AND TOPOLOGY : LECTURE 1

SAUGATA BASU

Abstract. In this lecture we introduce semi-algebraic sets, Tarski-Seidenberg principle, give basic definitions of homology and co-homology groups of semi-algebraic sets, and state certain quantitative results which give tight bounds on the ranks of these groups. We also state several open problems and exercises.

Contents

1. Semi-algebraic Geometry: Background

1.1. **Notation.** We first fix some notation. Let R be a ***real closed field*** (for example, the field $\mathbb{R}$ of real numbers or $\mathbb{R}_{\mathrm{alg}}$ of real algebraic numbers). A ***semi-algebraic subset of*** R^k is a set defined by a finite system of polynomial equalities and inequalities, or more generally by a Boolean formula whose atoms are polynomial equalities and inequalities. Given a finite set $\mathcal{P}$ of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$, a subset S of R^k is ***$\mathcal{P}$-semi-algebraic*** if S is the realization of a Boolean formula with atoms $P = 0$, $P > 0$ or $P < 0$ with $P \in \mathcal{P}$. It is clear that for every semi-algebraic subset S of R^k there exists a finite set $\mathcal{P}$ of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$ such that S is $\mathcal{P}$-semi-algebraic. We call a semi-algebraic set a ***$\mathcal{P}$-closed semi-algebraic set*** if

2000 MATHEMATICS SUBJECT CLASSIFICATION PRIMARY 14P10, 14P25; SECONDARY 68W30

it is defined by a Boolean formula with no negations with atoms $P = 0$, $P \geq 0$, or $P \leq 0$ with $P \in \mathcal{P}$.

For an element $a \in \mathrm{R}$ we let

$$\mathrm{sign}(a) = \begin{cases} 0 & \text{if } a = 0, \\ 1 & \text{if } a > 0, \\ -1 & \text{if } a < 0. \end{cases}$$

A ***sign condition*** on $\mathcal{P}$ is an element of $\{0, 1, -1\}^{\mathcal{P}}$. For any semi-algebraic set $Z \subset \mathrm{R}^k$ the ***realization of the sign condition*** σ ***over*** Z, $\mathcal{R}(\sigma, Z)$, is the semi-algebraic set

$$\{x \in Z \mid \bigwedge_{P \in \mathcal{P}} \mathrm{sign}(P(x)) = \sigma(P)\},$$

and in case $Z = \mathrm{R}^k$ we will denote $\mathcal{R}(\sigma, Z)$ by just $\mathcal{R}(\sigma)$.

If $\mathcal{P}$ is a finite subset of $\mathrm{R}[X_1, \ldots, X_k]$, we write the set of zeros of $\mathcal{P}$ in R^k as

$$\mathrm{Z}(\mathcal{P}, \mathrm{R}^k) = \{x \in \mathrm{R}^k \mid \bigwedge_{P \in \mathcal{P}} P(x) = 0\}.$$

We will denote by $B_k(0, r)$ the open ball with center 0 and radius r in R^k. We will also denote by $\mathbf{S}^k$ the unit sphere in R^{k+1} centered at the origin. Notice that these sets are semi-algebraic.

For any semi-algebraic set X, we denote by $\overline{X}$ the closure of X, which is also a semi-algebraic set by the Tarksi-Seidenberg principle [14, 12] (see [5] for a modern treatment). The ***Tarksi-Seidenberg principle*** states that the class of semi-algebraic sets is closed under linear projections or equivalently that ***the first-order theory of the reals admits quantifier elimination***.

Exercise 1.1. Prove using the Tarski-Seidenberg principle that the closure of a semi-algebraic set is again a semi-algebraic set.

Exercise 1.2. Consider the semi-algebraic subset of R defined by the single (strict) polynomial inequality $X^3 - X > 0$. Notice that the closure of this set is *not* defined by the corresponding weak inequality $X^3 - X \geq 0$.

1.2. **Topological Preliminaries.** The purpose of this section is to provide a self-contained introduction to the basic mathematical machinery needed later. Some of the topics would be familiar to most readers while a few others perhaps less so. The sophisticated reader can choose to skip this whole section and proceed directly to the descriptions of the various algorithms in the later sections.

1.2.1. *Homology and Cohomology Groups.* Before we get to the precise definitions of these groups it is good to have some intuition about them. Closed and bounded semi-algebraic sets are ***finitely triangulable***. This means that each closed and bounded semi-algebraic set $S \subset \mathrm{R}^k$ is homeomorphic (in fact, by a semi-algebraic map) to the polyhedron $|K|$ associated to a finite simplicial complex K. In fact K can be chosen such that $|K| \subset \mathrm{R}^k$, and there is an effective algorithm for computing K given S. The ***simplicial cohomology (resp. homology groups)*** of S are defined in terms of K and are well-defined (i.e they are independent of the chosen triangulation which is of course very far from being unique).

Roughly speaking the simplicial homology groups of a finite simplicial complex K with coefficients in a field $\mathbb{F}$ (which we assume to be $\mathbb{Q}$ in this survey) are finite dimensional $\mathbb{F}$-vector spaces and measure the *connectivity* of $|K|$ in various dimensions. For example, the zero-th simplicial homology group, $\mathrm{H}_0(K)$, has a generator corresponding to each connected component of K and its dimension gives the number of connected components of $|K|$. Similarly the first simplicial homology group, $\mathrm{H}_1(K)$, is generated by the "one-dimensional holes" of $|K|$, and its dimension is the number of "independent" one-dimensional holes of $|K|$. If K is one-dimensional (that is a finite graph) the dimension of $\mathrm{H}_1(K)$ is the number of independent cycles in K. Analogously, the i-th the simplicial homology group, $\mathrm{H}_i(K)$, is generated by the "i-dimensional holes" of $|K|$, and its dimension is the number of independent i-dimensional holes of $|K|$. Intuitively an i-dimensional hole is an i-dimensional closed surface in K (technically called a ***cycle***) which does not *bound* any $(i+1)$-dimensional subset of $|K|$.

The simplicial cohomology groups of K are dual (and isomorphic) to the simplicial homology groups of K as groups. However, in addition to the group structure they also carry a multiplicative structure (the so called cup-product) which makes them a finer topological invariant than the homology groups. We are not going to use this multiplicative structure. Cohomology groups also have nice but less geometric interpretations. Roughly speaking the cohomology groups of K represent spaces of globally defined objects satisfying certain local conditions. For example, the zero-th cohomology group, $\mathrm{H}^0(K)$, can be interpreted as the vector space of global functions on $|K|$ which are locally constant. It is easy to see from this interpretation that the dimension of $\mathrm{H}^0(K)$ is the number of connected components of K. Similar geometric interpretations can be given for the higher cohomology groups, in terms of vector spaces of (globally defined) differential forms satisfying certain local condition (of being closed). In literature this cohomology theory is referred to as ***de Rham cohomology theory*** and it is usually defined for smooth manifolds, but it can also be defined for simplicial complexes (see for example [10, Section 1.3.1]).

It turns out that the cohomological point of view gives better intuition in designing algorithms described later in the paper. This is our primary reason behind preferring cohomology over homology. Another reason for preferring the cohomology groups over the homology groups is that their interpretations continue to make sense in applications outside of semi-algebraic geometry where the notions of holes is meaningless (for instance, think of algebraic varieties defined over fields of positive characteristics) but the notion of global functions (or for instance differential forms) continue to make sense.

1.2.2. *Definition of the Cohomology Groups of a Simplicial Complex.* We now give precise definitions of the cohomology groups of simplicial complexes.

In order to do so we first need to introduce some amount of algebraic machinery – namely the concept of complexes of vector spaces and homomorphisms between them.

1.2.3. *Complex of Vector Spaces.* A ***complex*** of vector spaces is just a sequence of vector spaces and linear transformations satisfying the property that the composition of two successive linear transformations is 0.

More precisely

Definition 1.3 (Complex of Vector Spaces)**.** A sequence $\{\mathrm{C}^p\}$, $p \in \mathbb{Z}$, of $\mathbb{Q}$-vector spaces together with a sequence $\{\delta^p\}$ of homomorphisms $\delta^p : \mathrm{C}^p \to \mathrm{C}^{p+1}$ (called differentials) for which

$$\delta^{p+1} \circ \delta^p = 0 \tag{1.1}$$

for all p is called a complex.

The most important example for us of a complex of vector spaces is the ***co-chain complex*** of a simplicial complex K denoted by $\mathrm{C}^\bullet(K)$. It is defined as follows.

Definition 1.4 (Simplicial cochain complex)**.** For each $p \geq 0$, $\mathrm{C}^p(K)$ is a linear functional on the $\mathbb{Q}$-vector-space generated by the p-simplices of K. Given $\phi \in \mathrm{C}^p(K)$, $\delta^p(\phi)$ is specified by its values on the $(p+1)$-dimensional simplices of K. Given a $(p+1)$-dimensional simplex $\sigma = [a_0, \ldots, a_{p+1}]$ of K

$$(\delta^p \phi)([a_0, \ldots, a_{p+1}]) = \sum_{i=0}^{p+1} (-1)^i \phi([a_0, \ldots, \hat{a}_i, \ldots, a_{p+1}]), \tag{1.2}$$

where $\hat{}$ denotes omission.

Notice that each $[a_0, \ldots, \hat{a}_i, \ldots, a_{p+1}]$ is a p-dimensional simplex of K and since $\phi \in \mathrm{C}^p(K)$, $\phi([a_0, \ldots, \hat{a}_i, \ldots, a_{p+1}]) \in \mathbb{Q}$ is well-defined. It is an exercise now to check that the homomorphisms $\delta^p : \mathrm{C}^p(K) \to \mathrm{C}^{p+1}(K)$ indeed satisfy Eqn. 1.1 in the definition of a complex.

Now let K be a simplicial complex and $L \subset K$ a sub-complex of K – we will denote such a pair simply by (K, L). Then for each $p \geq 0$ we have a surjective homomorphism $\mathrm{C}^p(K) \to \mathrm{C}^p(L)$ and we denote by $\mathrm{C}^p(K, L)$ the kernel of this homomorphism. It is now an easy exercise to verify that the differentials δ^p in the complex $\mathrm{C}^p(K)$ descend to $\mathrm{C}^p(K, L)$ and we define

Definition 1.5 (Simplicial cochain complex of a pair)**.** The simplicial cochain complex of the pair (K, L) to be the complex $\mathrm{C}^\bullet(K, L)$ whose terms, $\mathrm{C}^p(K, L)$, and differentials, δ^p, are defined as above.

Often, particularly in the context of algorithmic applications it is more economical to use cellular complexes instead of simplicial complexes. We recall here the definition of a finite regular cell complex referring the reader to standard sources in algebraic topology for more in-depth study of cellular theory (see [16, pp. 81]).

Definition 1.6 (Regular cell complex)**.** An ℓ-dimensional ***cell*** in R^k is a subset of R^k homeomorphic to $\overline{B_\ell(0,1)}$. A regular cell complex Σ in R^k is a finite collection of cells satisfying the following properties:

(1) If $c_1, c_2 \in \Sigma$, then either $c_1 \cap c_2 = \emptyset$ or $c_1 \subset \partial c_2$ or $c_2 \subset \partial c_1$.
(2) The boundary of each cell of Σ is a union of cells of Σ.

We denote by $|\Sigma|$ the set $\displaystyle\bigcup_{c\in\Sigma} c$.

Remark 1.7. Notice that every simplicial complex K may be considered as a regular cell complex whose cells are the closures of the simplices of K.

As in the case of simplicial complexes it is possible to associate a complex, $\mathrm{C}^\bullet(\Sigma)$ (the co-chain complex of K), to each regular cell complex K which is defined in an analogous manner. In order to avoid technicalities we omit the precise definition of this complex referring the interested reader to [16, pp. 82] instead. We remark that the dimension of $\mathrm{C}^p(\Sigma)$ is equal to the number of p-dimensional cells in Σ and the matrix entries for the differentials in the complex with respect to the standard basis comes from $\{0, 1, -1\}$ just as in the case of simplicial co-chain complexes.

The advantage of using cell complexes instead of simplicial complexes can be seen in the following example.

Example 1.8. Consider the unit sphere $\mathbf{S}^k \subset \mathrm{R}^{k+1}$. For $0 \le j \le k+1$ and $\varepsilon \in \{+,-\}$ let

$$c_j^\varepsilon = \{x \in \mathbf{S}^k \mid X_0 = \cdots = X_{j-1} = 0, \varepsilon X_j \ge 0\}. \tag{1.3}$$

Then it is easy to check that each c_j^ε is a $k-j$ dimensional cell and the collection, $\Sigma_k = \{c_j^\varepsilon \mid 0 \le j \le k, \varepsilon \in \{-,+\}\}$ is a regular cell complex with $|\Sigma_k| = \mathbf{S}^k$ (see Figure 1 for the case $k=2$).

Notice that $\#\Sigma_k = 2k$. On the other hand if we consider the sphere as homeomorphic to the boundary of a standard $(k+1)$-dimensional simplex, then the corresponding simplicial complex will contain $(2^{k+2}-2)$ simplices (which is exponentially large in k).

We now associate to each complex, $\mathrm{C}^\bullet$, a sequence of vector spaces, $\mathrm{H}^p(\mathrm{C}^\bullet)$, called the cohomology groups of $\mathrm{C}^\bullet$. Note that it follows from Eqn. 1.1 that for a complex $\mathrm{C}^\bullet$ with differentials $\delta^p : \mathrm{C}^p \to \mathrm{C}^{p+1}$ the subspace $B^p(\mathrm{C}^\bullet) = \mathrm{Im}(\delta^{p-1}) \subset \mathrm{C}^p$ is contained in the subspace $Z^p(\mathrm{C}^\bullet) = \mathrm{Ker}(\delta^p) \subset \mathrm{C}^p$. The subspaces $B^p(\mathrm{C}^\bullet)$ (resp. $Z^p(\mathrm{C}^\bullet)$) are usually referred to as the ***co-boundaries*** (resp. ***co-cycles***) of the complex $\mathrm{C}^\bullet$. Moreover,

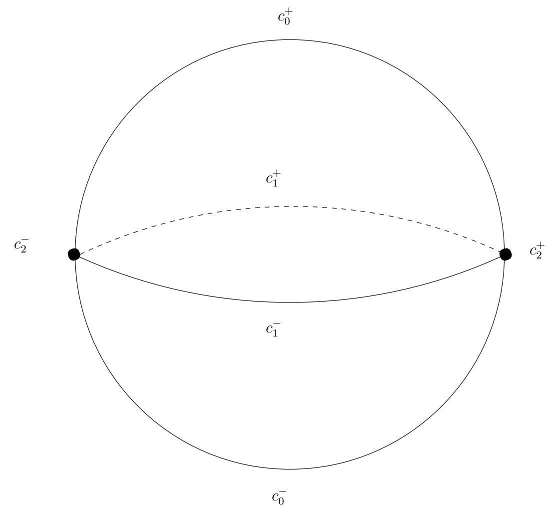

FIGURE 1. Cell decomposition of $\mathbf{S}^2$

Definition 1.9 (Cohomology groups of a complex)**.** The ***cohomology groups,*** $\mathrm{H}^p(\mathrm{C}^\bullet)$, are defined by

$$\mathrm{H}^p(\mathrm{C}^\bullet) = Z^p(\mathrm{C}^\bullet)/B^p(\mathrm{C}^\bullet). \tag{1.4}$$

We will denote by $\mathrm{H}^*(\mathrm{C}^\bullet)$ the graded vector space $\bigoplus_p \mathrm{H}^p(\mathrm{C}^\bullet)$.

Note that the cohomology groups, $\mathrm{H}^p(\mathrm{C}^\bullet)$, are all $\mathbb{Q}$-vector spaces (finite dimensional if the vector spaces C^p's are themselves finite dimensional).

Definition 1.10 (Exact sequence)**.** A complex $\mathrm{C}^\bullet$ is called ***acyclic*** and the corresponding sequence of vector space homomorphisms is called an ***exact sequence*** if $\mathrm{H}^*(\mathrm{C}^\bullet) = 0$.

Exercise 1.11. Look up and prove for yourself the "snake lemma" and the "five lemma".

Applying Definition 1.9 to the particular case of the co-chain complex of a simplicial complex K (cf. Definition 1.4) we obtain

1.2.4. *Cohomology of a Simplicial Complex.*

Definition 1.12 (Cohomology of a simplicial complex)**.** The cohomology groups of a simplicial complex K are by definition the cohomology groups, $\mathrm{H}^p(\mathrm{C}^\bullet(K))$, of its co-chain complex.

Similarly, given a pair of simplicial complexes (K, L), we define

Definition 1.13 (Cohomology of a pair)**.** The cohomology groups of the pair (K, L) are by definition the cohomology groups, $\mathrm{H}^p(\mathrm{C}^\bullet(K, L))$, of its co-chain complex.

Example 1.14. Let Δ_n be the simplicial complex corresponding to an n-simplex. In other words the simplices of Δ_n consist of $[i_0, \ldots, i_\ell], 0 \leq i_0 < \cdots < i_\ell \leq n$. The polyhedron $|\Delta_n|$ is just the n-dimensional simplex. Then using Definition 1.12 one can verify that

$$\mathrm{H}^i(\Delta_n) = \mathbb{Q},\ i = 0,$$
$$\mathrm{H}^i(\Delta_n) = 0,\ i > 0.$$

Example 1.15. Let $\partial\Delta_n$ be the simplicial complex corresponding to the boundary of the n-simplex. In other words the simplices of $\partial\Delta_n$ consist of $[i_0, \ldots, i_\ell], 0 \leq i_0 < \cdots < i_\ell \leq n, \ell < n$. Then again by a direct application of Definition 1.12 one can verify that

$$\mathrm{H}^i(\partial\Delta_n) = \mathbb{Q},\ i = 0, n-1$$
$$\mathrm{H}^i(\partial\Delta_n) = 0,\ \text{ else.}$$

The above examples serve to confirm our geometric intuition behind the homology groups of the spaces $|\Delta_n|$ and $|\partial\Delta_n|$ explained in Section 1.2.1 above – namely that they are both connected and $|\Delta_n|$ has no holes in dimension > 0, and $|\partial\Delta_n|$ has a single $(n-1)$-dimensional hole.

Exercise 1.16. Verify that

$$\mathrm{H}^i(\Delta_n, \partial\Delta_n) = \mathbb{Q},\ i = n$$
$$\mathrm{H}^i(\Delta_n, \partial\Delta_n) = 0,\ \text{ else.}$$

1.2.5. *Homomorphisms of Complexes.* We will also need the notion of homomorphisms of complexes which generalizes the notion of ordinary vector space homomorphisms.

Definition 1.17 (Homomorphisms of complexes)**.** Given two complexes, $\mathrm{C}^\bullet = (\mathrm{C}^p, \delta^p)$ and $\mathrm{D}^\bullet = (\mathrm{D}^p, \delta^p)$, ***a homomorphism of complexes***, $\phi^\bullet : \mathrm{C}^\bullet \to \mathrm{D}^\bullet$, is a sequence of homomorphisms $\phi^p : \mathrm{C}^p \to \mathrm{D}^p$ for which $\delta^p \circ \phi^p = \phi^{p+1} \circ \delta^p$ for all p.

In other words the following diagram is commutative.

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \mathrm{C}^p & \xrightarrow{\delta^p} & \mathrm{C}^{p+1} & \longrightarrow & \cdots \\ & & \Big\downarrow{\scriptstyle \phi^p} & & \Big\downarrow{\scriptstyle \phi^{p+1}} & & \\ \cdots & \longrightarrow & \mathrm{D}^p & \xrightarrow{\delta^p} & \mathrm{D}^{p+1} & \longrightarrow & \cdots \end{array} \tag{1.5}$$

A homomorphism of complexes $\phi^\bullet : \mathrm{C}^\bullet \to \mathrm{D}^\bullet$ induces homomorphisms $\phi^i : \mathrm{H}^i(\mathrm{C}^\bullet) \to \mathrm{H}^i(\mathrm{D}^\bullet)$ and we will denote the corresponding homomorphism between the graded vector spaces $\mathrm{H}^*(\mathrm{C}^\bullet), \mathrm{H}^*(\mathrm{D}^\bullet)$ by ϕ^*.

Definition 1.18 (Quasi-isomorphism)**.** The homomorphism $\phi^\bullet$ is called a ***quasi-isomorphism*** if the homomorphism ϕ^* is an isomorphism.

Having introduced the algebraic machinery of complexes of vector spaces, we now define the cohomology groups of semi-algebraic sets in terms of their triangulations and their associated simplicial complexes.

1.2.6. *Cohomology Groups of Semi-algebraic Sets.* A closed and bounded semi-algebraic set $S \subset \mathrm{R}^k$ is semi-algebraically triangulable.

Definition 1.19 (Cohomology groups of closed and bounded semi-algebraic sets)**.** Given a triangulation, $h : |K| \to S$, where K is a simplicial complex, we define the i-th simplicial cohomology group of S, by $\mathrm{H}^i(S) = \mathrm{H}^i(\mathrm{C}^\bullet(K))$, where $\mathrm{C}^\bullet(K)$ is the co-chain complex of K. The groups $\mathrm{H}^i(S)$ are ***invariant under semi-algebraic homeomorphisms*** (and they coincide with the corresponding singular cohomology groups when $\mathrm{R} = \mathbb{R}$). We denote by $b_i(S)$ the ***i-th Betti number of S*** (i.e. the dimension of $\mathrm{H}^i(S)$ as a vector space).

Exercise 1.20. Prove from the above definition that if S is a closed and bounded semi-algebraic set then $b_0(S)$ is the number of semi-algebraically connected components of S.

Remark 1.21*.* For a closed but not necessarily bounded semi-algebraic set $S \subset \mathrm{R}^k$ we will denote by $\mathrm{H}^i(S)$ the i-th simplicial cohomology group of $S \cap \overline{B_k(0,r)}$ for sufficiently large $r > 0$. The sets $S \cap \overline{B_k(0,r)}$ are semi-algebraically homeomorphic for all sufficiently large $r > 0$ and hence this definition makes sense. (The last property is usually referred to as ***the local conic structure at infinity*** of semi-algebraic sets [5, Theorem 5.48]). The definition of cohomology groups of arbitrary semi-algebraic sets in R^k requires some care and several possibilities exist and we refer the reader to [5, Section 6.3] where one such definition is given which agrees with singular cohomology in case $\mathrm{R} = \mathbb{R}$.

1.2.7. *The Euler-Poincaré Characteristic: Definition and Basic Properties.* An useful topological invariant of semi-algebraic sets which is often easier to compute than their Betti numbers is the ***Euler-Poincaré characteristic***.

Definition 1.22 (Euler-Poincaré characteristic of a closed and bounded semi-algebraic set)**.** Let $S \subset \mathrm{R}^k$, be a closed and bounded semi-algebraic set. Then the Euler-Poincaré characteristic of S is defined by

$$\chi(S) = \sum_{i \geq 0} (-1)^i b_i(S). \tag{1.6}$$

From the point of view of designing algorithms, it is useful to define Euler-Poincaré characteristic also for ***locally closed*** semi-algebraic sets. A semi-algebraic set is locally closed if it is the intersection of a closed semi-algebraic set with an open one. A standard example of a locally closed semi-algebraic set is the realization, $\mathcal{R}(\sigma)$, of a sign-condition σ on a family of polynomials.

We now define Euler-Poincaré characteristic for locally closed semi-algebraic sets in terms of the ***Borel-Moore cohomology*** groups of such sets (defined below). This definition agrees with the definition of Euler-Poincaré characteristic stated above for closed and bounded semi-algebraic sets. They may be distinct for semi-algebraic sets which are closed but not bounded.

Definition 1.23. The simplicial cohomology groups of a pair of closed and bounded semi-algebraic sets $T \subset S \subset \mathrm{R}^k$ are defined as follows. Such a pair of closed and bounded semi-algebraic sets can be triangulated using a pair of simplicial complexes (K, A) where A is a sub-complex of K. The p-th simplicial cohomology group of the pair (S, T), $\mathrm{H}^p(S, T)$, is by definition to be $\mathrm{H}^p(K, A)$. The dimension of $\mathrm{H}^p(S, T)$ as a $\mathbb{Q}$-vector space is called the p-th Betti number of the pair (S, T) and denoted $b_p(S, T)$. The Euler-Poincaré characteristic of the pair (S, T) is

$$\chi(S, T) = \sum_i (-1)^i b_i(S, T).$$

Definition 1.24 (Borel-Moore cohomology group)**.** The ***p-th Borel-Moore cohomology*** group of $S \subset \mathrm{R}^k$, denoted $\mathrm{H}^p_{BM}(S)$, is defined in terms of the cohomology groups of a pair of closed and bounded semi-algebraic sets as follows. For any $r > 0$ let $S_r = S \cap B_k(0, r)$. Note that for a locally closed semi-algebraic set S both $\overline{S_r}$ and $\overline{S_r} \setminus S_r$ are closed and bounded, and hence $\mathrm{H}^p(\overline{S_r}, \overline{S_r} \setminus S_r)$ is well defined. Moreover, it is a consequence of the local conic structure at infinity of semi-algebraic sets (see Remark 1.21 above) that the cohomology group $H^p(\overline{S_r}, \overline{S_r} \setminus S_r)$ is invariant for all sufficiently large $r > 0$. We define $\mathrm{H}^p_{BM}(S) = \mathrm{H}^p(\overline{S_r}, \overline{S_r} \setminus S_r)$ for $r > 0$ sufficiently large and it follows from the above remark that it is well defined.

Exercise 1.25. Compute the Borel-Moore cohomology groups of the following locally closed semi-algebraic sets.

(1) $B_k(0, 1)$.
(2) $B_k(0, 1) \setminus \{0\}$.
(3) $\overline{B_k(0, 1)}$.
(4) $\overline{B_k(0, 1)} \setminus \{0\}$.

The Borel-Moore cohomology groups are invariant under semi-algebraic homeomorphisms (see [7]. It also follows clearly from the definition that for a closed and bounded semi-algebraic set the Borel-Moore cohomology groups coincide with the simplicial cohomology groups.

Definition 1.26 (Borel-Moore Euler-Poincaré characteristic)**.** For a locally closed semi-algebraic set S we define the Borel-Moore Euler-Poincaré characteristic by

$$\chi^{BM}(S) = \sum_{i=0}^{k} (-1)^i \, b_i^{BM}(S) \tag{1.7}$$

where $b_i^{BM}(S)$ denotes the dimension of $\mathrm{H}^i_{BM}(S)$.

Exercise 1.27. Compute $\chi^{BM}([0,1)), \chi^{BM}((0,1)), \chi^{BM}(B_k(0,1))$.

If S is closed and bounded then $\chi^{BM}(S) = \chi(S)$.

The Borel-Moore Euler-Poincaré characteristic has the following additivity property (reminiscent of the similar property of volumes) which makes them particularly useful in algorithmic applications.

Proposition 1.28. *Let X, X_1 and X_2 be locally closed semi-algebraic sets such that*

$$X_1 \cup X_2 = X, X_1 \cap X_2 = \emptyset.$$

Then

$$\chi^{BM}(X) = \chi^{BM}(X_1) + \chi^{BM}(X_2). \tag{1.8}$$

Since for closed and bounded semi-algebraic sets, the Borel-Moore Euler-Poincaré characteristic agrees with the ordinary Euler-Poincaré characteristic, it is easy to derive the following additivity property of the Euler-Poincaré characteristic of closed and bounded sets.

Proposition 1.29. *Let X_1 and X_2 be closed and bounded semi-algebraic sets. Then*

$$\chi(X_1 \cup X_2) = \chi(X_1) + \chi(X_2) - \chi(X_1 \cap X_2). \tag{1.9}$$

Note that Proposition 1.29 is an immediate consequence of Proposition 1.28 once we notice that the sets $Y_1 = X_1 \setminus (X_1 \cap X_2)$ and $Y_2 = X_2 \setminus (X_1 \cap X_2)$ are locally closed, the set $X_1 \cup X_2$ is the disjoint union of the locally closed sets Y_1, Y_2 and $X_1 \cap X_2$, and

$$\chi^{BM}(Y_i) = \chi^{BM}(X_i) - \chi^{BM}(X_1 \cap X_2) = \chi(X_i) - \chi(X_1 \cap X_2), \text{ for } i = 1, 2.$$

More generally by applying Proposition 1.29 inductively we get the following inclusion-exclusion property of the (ordinary) Euler-Poincaré characteristic.

For any $n \in \mathbb{Z}_{\geq 0}$ we denote by $[n]$ the set $\{1, \ldots, n\}$.

Proposition 1.30. *Let $X_1, \ldots, X_n$ be closed and bounded semi-algebraic sets. Then denoting by X_I the semi-algebraic set $\bigcap_{i \in I} X_i$ for $I \subset [n]$, we have*

$$\chi(\bigcup_{i \in [n]} X_i) = \sum_{I \subset [n]} (-1)^{(\#I+1)} \chi(X_I). \tag{1.10}$$

1.2.8. *Homotopy Invariance.* The cohomology groups of semi-algebraic sets as defined above (Definition 1.19) are obviously invariant under semi-algebraic homeomorphisms. But, in fact, they are invariant under a weaker equivalence relation – namely, semi-algebraic ***homotopy equivalence*** (defined below). This property is crucial in the design of efficient algorithms for computing Betti numbers of semi-algebraic sets since it allows us to replace a given set by one that is better behaved from the algorithmic point of view but having the same homotopy type as the original set. This technique is ubiquitous in algorithmic semi-algebraic geometry and we will see some version of it in almost every algorithm described in the following lectures.

Remark 1.31. The reason behind insisting on the prefix *"semi-algebraic"* with regard to homeomorphisms and homotopy equivalences here and in the rest of the paper, is that for general real closed fields, the ordinary Euclidean topology could be rather strange. For example, the real closed field, $\mathbb{R}_{\mathrm{alg}}$, of real algebraic numbers is totally disconnected as a topological space under the Euclidean topology. On the other hand, if the ground field $\mathrm{R} = \mathbb{R}$, then we can safely drop the prefix "semi-algebraic" in the statements made above. However, even if we start with $\mathrm{R} = \mathbb{R}$, in many applications described below we enlarge the field by taking non-archimedean extensions of R made above would again apply to these field extensions.

Definition 1.32 (Semi-algebraic homotopy)**.** Let X, Y be two closed and bounded semi-algebraic sets. Two semi-algebraic continuous functions $f, g : X \to Y$ are ***semi-algebraically homotopic***, $f \sim_{sa} g$, if there is a continuous semi-algebraic function $F : X \times [0,1] \to Y$ such that $F(x,0) = f(x)$ and $F(x,1) = g(x)$ for all $x \in X$.

Clearly, semi-algebraic homotopy is an equivalence relation among semi-algebraic continuous maps from X to Y.

Definition 1.33 (Semi-algebraic homotopy equivalence)**.** The sets X, Y are semi-algebraically homotopy equivalent if there exist semi-algebraic continuous functions $f : X \to Y$, $g : Y \to X$ such that $g \circ f \sim_{sa} \mathrm{Id}_X$, $f \circ g \sim_{sa} \mathrm{Id}_Y$.

We have

Proposition 1.34 (Homotopy Invariance of the Cohomology Groups)**.** *Let X, Y be two closed and bounded semi-algebraic sets of R^k that are semi-algebraically homotopy equivalent. Then,* $\mathrm{H}^*(X) \cong \mathrm{H}^*(Y)$.

For any semi-algebraic set S, we will denote by $b_i(S)$ its i-th Betti number, which is the dimension of the i-th cohomology group, $\mathrm{H}^i(S, \mathbb{Q})$, taken with rational coefficients, which in our setting is also isomorphic to the i-th homology group, $\mathrm{H}_i(S, \mathbb{Q})$ (see Section 1.2.6 below for precise definitions of these groups). In particular, $b_0(S)$ is the number of semi-algebraically connected components of S. We will sometimes refer to the sum $b(S) = \sum_{i \geq 0} b_i(S)$ as the ***topological complexity*** of a semi-algebraic set S.

2. Bounds

After having defined the Betti numbers of semi-algebraic sets we now proceed to discuss the best known bounds on these numbers.

2.1. Bounds on the Betti Numbers of Algebraic Sets.

We first consider the simplest case of a non-singular compact hyper-surface in R^k.

Theorem 2.1. *Let $P \subset \mathrm{R}[X_1, \ldots, X_k]$ such that $\deg(P) \leq d$ and $\mathrm{Z}(P, \mathrm{R}^k)$ is a non-singular, compact hypersurface. Then,*

$$b(S) \leq d(d-1)^{k-1}.$$

More generally, we have that

Theorem 2.2. [11, 15, 9] *Let $P \subset \mathrm{R}[X_1, \ldots, X_k]$ such that $\deg(P) \leq d$. Then,*

$$b(S) \leq d(2d-1)^{k-1}.$$

2.2. **Bounds on the Betti numbers of Semi-algebraic Sets.** We now consider the case of semi-algebraic sets.

For the special case of $\mathcal{P}$-closed semi-algebraic sets we have the following bound. (and this bound is used in an essential way in the proof of Theorem 2.4). Using the same notation as in Theorem 2.4 above we have

Theorem 2.3. [2, 4] *For every $\mathcal{P}$-closed semi-algebraic set $S \subset \mathrm{R}^k$, with $s = \#\mathcal{P}$, and $d = \max_{P \in \mathcal{P}} \deg(P)$,*

$$b(S) \leq \sum_{i=0}^{k} \sum_{j=0}^{k-i} \binom{s}{j} 6^j d(2d-1)^{k-1}.$$

Theorem 2.4. [8] *For a general $\mathcal{P}$-semi-algebraic set $S \subset \mathrm{R}^k$, with $s = \#\mathcal{P}$, and $d = \max_{P \in \mathcal{P}} \deg(P)$,*

$$b(S) \leq \sum_{i=0}^{k} \sum_{j=0}^{k-i} \binom{2s^2+1}{j} 6^j d(2d-1)^{k-1}.$$

Remark 2.5. The bound gets worse (compared to Theorem 2.3) because of an initial reduction step which entails replacing an arbitrary semi-algebraic $\mathcal{P}$-semi-algebraic set by a $\mathcal{P}'$-*closed* semi-algebraic set having the same homotopy type as the original set, where the polynomials in $\mathcal{P}'$ satisfy the same degree bound as the ones in $\mathcal{P}$, but $\#\mathcal{P}'$ is larger than $\#\mathcal{P}$.

Exercise 2.6. Prove that every semi-algebraic set is (semi-algebraically) homotopy equivalent to a closed and bounded one.

Exercise 2.7. For $k = 1, 2$ design a scheme for replacing an arbitrary $\mathcal{P}$-semi-algebraic by a $\mathcal{P}'$-closed semi-algebraic set having the same homotopy type, and such that the polynomials in $\mathcal{P}'$ are of the form $P \pm \varepsilon_j, P \in \mathcal{P}$ and the ε_j's are infinitesimals. How about a general k ?

Exercise 2.8 (Open problem)**.** Notice that the bound on the Betti numbers for an arbibtrary $\mathcal{P}$-semi-algebraic set in Theorem 2.4 is significantly worse than the bound for $\mathcal{P}$-closed semi-algebraic set in Theorem 2.3. Does this have to be so ? In other words construct an arbitrary $\mathcal{P}$-semi-algebraic set for which the sum of the Betti numbers is bigger than the bound in Theorem 2.3 (for $\mathcal{P}$-closed semi-algebraic sets) (or improve Theorem 2.4).

2.3. **Betti numbers of Sign Conditions.** A family of polynomials $\mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$ induces a partition of R^k into connected regions where the polynomials maintain their signs. Bounding the the number of elements of such a partition (and also their topological complexity) is often very important in practice. The following theorem gives a bound on the different Betti numbers of the elements of this partition.

For $\sigma \in \{0, 1, -1\}^{\mathcal{P}}$, let $b_i(\sigma)$ denote the i-th Betti number of $\mathcal{R}(\sigma)$ and let

$$b_i(\mathcal{P}) = \sum_{\sigma} b_i(\sigma).$$

Note that, $b_0(\mathcal{P})$ is the total number of semi-algebraically connected components of the realizations of all realizable sign conditions of $\mathcal{P}$.

With same notation as before we have

Theorem 2.9. [4] *For* $0 \leq i \leq k$,

$$b_i(\mathcal{P}) \leq \sum_{j=0}^{k-i} \binom{s}{j} 4^j d(2d-1)^{k-1}.$$

Remark 2.10. These bounds are asymptotically tight, as can be already seen from the example where each $P \in \mathcal{P}$ is a product of d generic polynomials of degree one. The number of connected components of the $\mathcal{P}$-semi-algebraic set defined as the subset of R^k where all polynomials in $\mathcal{P}$ are non-zero is clearly bounded from below by $(\Omega(sd))^k$.

Exercise 2.11 (Open ended)**.** While the bound in Theorem 2.9 is asymptotically tight there is still a substantial gap between the actual (non-asymptotic) upper and lower bounds. Can you close this gap at least for small values of k (starting from $k = 2$) ?

2.4. **On the Number of Connected Components of Sign Conditions.** Often the number of connected components (i.e. $b_0(\mathcal{P})$) is the quantity of interest and it is an interesting problem to derive precise bounds just in this case. To obtain absolutely tight bounds is still an open problem. However, if we only consider asymptotics in $s = \#\mathcal{P}$ and consider d, k to be fixed we have tight bounds.

With the notation of the previous paragraphs we have

Theorem 2.12. [6]

$$\begin{aligned} b_0(\mathcal{P}) &\leq \sum_{0 \leq \ell \leq k} \left(\binom{s}{k-\ell} \binom{s}{\ell} d^k + \sum_{1 \leq i \leq \ell} \binom{s}{k-\ell} \binom{s}{\ell - i} d^{O(k^2)} \right) \\ &= \sum_{\ell=0}^{k} \binom{s}{k-\ell} \binom{s}{\ell} d^k + O(s^{k-1}) \\ &= \left(\sum_{\ell=0}^{k} \frac{d^k}{\ell!(k-\ell)!} \right) s^k + O(s^{k-1}) \end{aligned}$$

$$= \frac{(2d)^k}{k!} s^k + O(s^{k-1}).$$

Remark 2.13. For $0 < d, k \ll s$, this gives

$$b_0(\mathcal{P}) \leq \left(\frac{2esd}{k}\right)^k.$$

This matches asymptotically the lower bound obtained by taking s polynomials each of which is a product of d linear polynomials in general position.

2.5. **Certain Restricted Classes of Semi-algebraic Sets.** Since general semi-algebraic sets can have exponential topological complexity (cf. Remark 2.10), it is natural to consider certain restricted classes of semi-algebraic sets. One natural class consists of semi-algebraic sets defined by a conjunction of quadratic inequalities.

2.5.1. *Quantitative Bounds for Sets Defined by Quadratic Inequalities.* Since sets defined by linear inequalities have no interesting topology, sets defined by quadratic inequalities can be considered to be the simplest class of semi-algebraic sets which can have non-trivial topology. Such sets are in fact quite general, since every semi-algebraic set can be defined by a (quantified) formula involving only quadratic polynomials (at the cost of increasing the number of variables and the size of the formula). Moreover, as in the case of general semi-algebraic sets, the Betti numbers of such sets can be exponentially large. For example, the set $S \subset \mathrm{R}^k$ defined by

$$X_1(1 - X_1) \leq 0, \ldots, X_k(1 - X_k) \leq 0,$$

has $b_0(S) = 2^k$.

Hence, it is somewhat surprising that for any constant $\ell \geq 0$, the Betti numbers $b_{k-1}(S), \ldots, b_{k-\ell}(S)$, of a basic closed semi-algebraic set $S \subset \mathrm{R}^k$ defined by quadratic inequalities, are polynomially bounded. The following theorem which appears in [3] is derived using a bound proved by Barvinok [1] on the Betti numbers of sets defined by few quadratic equations.

Theorem 2.14. [3] *Let* R *a real closed field and* $S \subset \mathrm{R}^k$ *be defined by*

$$P_1 \leq 0, \ldots, P_s \leq 0, \deg(P_i) \leq 2, 1 \leq i \leq s.$$

Then, for any $\ell \geq 0$,

$$b_{k-\ell}(S) \leq \binom{s}{\ell} k^{O(\ell)}.$$

Notice that for fixed ℓ this gives a polynomial bound on the highest ℓ Betti numbers of S (which could possibly be non-zero). Observe also that similar bounds do not hold for sets defined by polynomials of degree greater than two. For instance, the set $V \subset \mathrm{R}^k$ defined by the single quartic inequality,

$$\sum_{i=1}^{k} X_i^2 (X_i - 1)^2 - \varepsilon \geq 0,$$

will have $b_{k-1}(V) = 2^k$, for all small enough $\varepsilon > 0$.

To see this observe that for all sufficiently small $\varepsilon > 0$, $\mathrm{R}^k \setminus V$ is defined by

$$\sum_{i=1}^{k} X_i^2 (X_i - 1)^2 < \varepsilon$$

and has 2^k connected components since it retracts onto the set $\{0, 1\}^k$. It now follows that

$$b_{k-1}(V) = b_0(\mathrm{R}^k \setminus V) = 2^k,$$

where the first equality is a consequence of the well-known Alexander duality theorem (see [13, pp. 296]).

References

1. A. I. Barvinok, *On the Betti numbers of semialgebraic sets defined by few quadratic inequalities*, Math. Z. **225** (1997), no. 2, 231–244. MR 98f:14044
2. S. Basu, *On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets*, Discrete Comput. Geom. **22** (1999), no. 1, 1–18. MR 1692627 (2000d:14061)
3. S. Basu, *Different bounds on the different Betti numbers of semi-algebraic sets*, Discrete Comput. Geom. **30** (2003), no. 1, 65–85, ACM Symposium on Computational Geometry (Medford, MA, 2001). MR 1991587 (2004f:14082)
4. S. Basu, R. Pollack, and M.-F. Roy, *On the Betti numbers of sign conditions*, Proc. Amer. Math. Soc. **133** (2005), no. 4, 965–974 (electronic). MR 2117195 (2006a:14096)
5. ______, *Algorithms in real algebraic geometry*, Algorithms and Computation in Mathematics, vol. 10, Springer-Verlag, Berlin, 2009, online version. MR 1998147 (2004g:14064)
6. ______, *An asymptotically tight bound on the number of connected components of realizable sign conditions*, Combinatorica (to appear), preprint at [arXiv:math.CO/0603256].
7. J. Bochnak, M. Coste, and M.-F. Roy, *Géométrie algébrique réelle (second edition in english: Real algebraic geometry)*, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], vol. 12 (36), Springer-Verlag, Berlin, 1987 (1998). MR 949442 (90b:14030)
8. A. Gabrielov and N. Vorobjov, *Betti numbers of semialgebraic sets defined by quantifier-free formulae*, Discrete Comput. Geom. **33** (2005), no. 3, 395–401. MR 2121987 (2005i:14075)
9. J. Milnor, *On the Betti numbers of real varieties*, Proc. Amer. Math. Soc. **15** (1964), 275–280. MR 0161339 (28 #4547)
10. S. Morita, *Geometry of characteristic classes*, Translations of Mathematical Monographs, vol. 199, American Mathematical Society, 1999.
11. I. G. Petrovskiĭ and O. A. Oleĭnik, *On the topology of real algebraic surfaces*, Izvestiya Akad. Nauk SSSR. Ser. Mat. **13** (1949), 389–402. MR 0034600 (11,613h)
12. A. Seidenberg, *A new decision method for elementary algebra*, Annals of Mathematics **60** (1954), 365–374.
13. Edwin H. Spanier, *Algebraic topology*, McGraw-Hill Book Co., New York, 1966. MR 0210112 (35 #1007)
14. A. Tarski, *A decision method for elementary algebra and geometry*, University of California Press, Berkeley and Los Angeles, Calif., 1951, 2nd ed. MR 13,423a

15. R. Thom, *Sur l'homologie des variétés algébriques réelles*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 255–265. MR 0200942 (34 #828)
16. G. Whitehead, *Elements of homotopy theory*, Graduate Texts in Mathematics, vol. 61, Springer-Verlag, New York, 1978. MR 516508 (80b:55001)

3. List of Exercises

(1) Prove using the Tarski-Seidenberg principle that the closure of a semi-algebraic set is again a semi-algebraic set.
(2) Consider the semi-algebraic subset of R defined by the single (strict) polynomial inequality $X^3 - X > 0$. Notice that the closure of this set is *not* defined by the corresponding weak inequality $X^3 - X \geq 0$.
(3) Verify that

$$\mathrm{H}^i(\Delta_n, \partial\Delta_n) = \mathbb{Q},\ i = n$$
$$\mathrm{H}^i(\Delta_n, \partial\Delta_n) = 0, \quad \text{else}.$$

(4) Prove from the above definition that if S is a closed and bounded semi-algebraic set then $b_0(S)$ is the number of semi-algebraically connected components of S.
(5) Compute the Borel-Moore cohomology groups of the following locally closed semi-algebraic sets.
 (a) $B_k(0,1)$.
 (b) $B_k(0,1) \setminus \{0\}$.
 (c) $\overline{B_k(0,1)}$.
 (d) $\overline{B_k(0,1)} \setminus \{0\}$.
(6) Compute $\chi^{BM}([0,1)), \chi^{BM}((0,1)), \chi^{BM}(B_k(0,1))$.
(7) Prove that every semi-algebraic set is (semi-algebraically) homotopy equivalent to a closed and bounded one.
(8) For $k = 1, 2$ design a scheme for replacing an arbitrary $\mathcal{P}$-semi-algebraic by a $\mathcal{P}'$-closed semi-algebraic set having the same homotopy type, and such that the polynomials in $\mathcal{P}'$ are of the form $P \pm \varepsilon_j, P \in \mathcal{P}$ and the ε_j's are infinitesimals. How about a general k ?
(9) (Open problem) Notice that the bound on the Betti numbers for an aribbtrary $\mathcal{P}$-semi-algebraic set in Theorem 2.4 is significantly worse than the bound for $\mathcal{P}$-closed semi-algebraic set in Theorem 2.3. Does this have to be so ? In other words construct an arbitrary $\mathcal{P}$-semi-algebraic set for which the sum of the Betti numbers is bigger than the bound in Theorem 2.3 (for $\mathcal{P}$-closed semi-algebraic sets) (or improve Theorem 2.4).
(10) (Open ended) While the bound in Theorem 2.9 is asymptotically tight there is still a substantial gap between the actual (non-asymptotic) upper and lower bounds. Can you close this gap at least for small values of k (starting from $k = 2$) ?

Department of Mathematics Purdue University, West Lafayette, IN 47907, U.S.A.

E-mail address: sbasu@math.purdue.edu