Computing the Homology of Semialgebraic Sets

Peter Bürgisser, Felipe Cucker, Josué Tonelli-Cueto

Berlin Mathematical School · City University of Hong Kong · Technische Universität Berlin

Semialgebraic sets

Semialgebraic sets are the class of geometric objects that can be described by real polynomials and inequalities. A way of describing a semialgebraic set is to use *formulas*. Formulas are expression obtained by combining *atoms* of the form

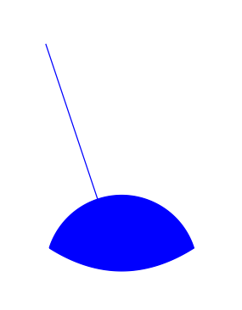
- $(p(x) < 0)$, • $(p(x) = 0)$, • $(p(x) > 0)$ and
- $(p(x) \leq 0)$, • $(p(x) \geq 0)$, • $(p(x) \neq 0)$,

which represent the most basic semialgebraic sets; using

- *negations* ($\neg$), which represent complements;
- *conjunctions* ($\wedge$), which represent intersections; and
- *disjunctions* ($\vee$), which represent unions.

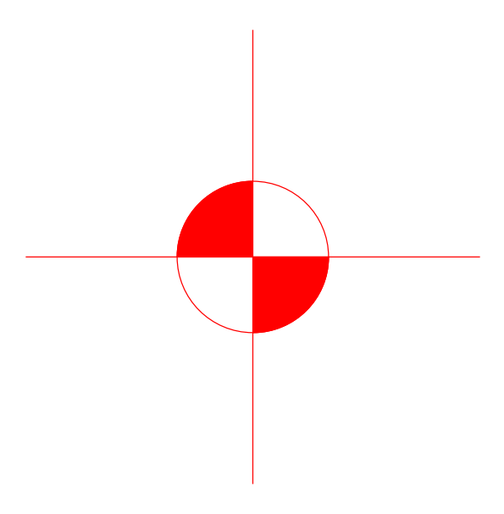
Formulas should be seen as "recipes" telling us how to construct the described set from the most basic ones.

Example 1



$$((x^2 + y^2 \leq 1) \vee (x + 3y = 0)) \wedge (\neg(3y - x^2 < 0))$$

Example 2



$$((xy \leq 0) \wedge ((x^2 + y^2 \leq 1) \vee (xy = 0))) \vee (x^2 + y^2 = 1)$$

Why do we care?

1. Semialgebraic sets are a large class of geometric objects preserved under many of the usual operations that one can do with sets (intersection, union, complementation, projection,...).
2. Semialgebraic sets can be used to describe:
   - Configuration space of a robotic arm.
   - Realization space of a polytope.
   - Configuration space of a molecule.
   - Regions of behavior of a real algebraic object.

How complex is the homology?

(Gabrielov, Vorobjov; 2005) showed that the sum of the Betti numbers (i.e., rank of the free part of the homology groups) of a semialgebraic set S is

$$O(q^2 D)^n$$

where q is the number of distinct polynomials in the description, D the maximum degree and n the number of variables.

Due to this, we expect that algorithms computing generators of homology should have a complexity

$$\mathrm{poly}(q, D)^{\mathrm{poly}(n)},$$

i.e., exponential in n, but all symbolic algorithms until today turns out to have complexity

$$\mathrm{poly}(q, D)^{\exp(n)},$$

i.e. doubly exponential in n.

Numerical algorithm

Our algorithm is numerical. Therefore, its performance (i.e. how much precision it needs and how many arithmetic operations it uses) is dominated by a condition number.

What is a condition number? A condition number is a quantity depending on the input that indicates how much the output varies when the input varies. However, for discrete outputs, such as homology, the condition number becomes easier. It measures how much we can modify the input without altering the homology.

Ill-posed inputs Those inputs for which the condition is infinite are called *ill-posed*. For them, an arbitrarily small perturbation of the coefficients of the polynomials changes what we want to compute (the homology groups). Using a numerical algorithm means accepting that for these inputs, there will be no numerical algorithm computing the answer.

Example. An example of an ill-posed input on our case is the parabola. This is so, because arbitrarily small perturbation of the equation of a parabola can turn it into either an hyperbola or an ellipse. All of these conics have different homology groups.

Probabilistic complexity

The condition number allows us to understand the performance of the algorithm for each input, but not uniformly as we can do with symbolic algorithms. However, can we get such an understanding that is not input-dependent? In other words, can we get rid of the condition?

To do so, we endow the inputs with a probability distribution and we compute probabilistic estimates of the condition.

Weak complexity There are several ways of computing probabilistic estimates on the condition. One of the most common ones is to compute bounds on the expectation and other moments of the conditions. However, recently, a new approach was developed by (Amelunxen, Lotz; 2015): *weak complexity.*

In this approach, one constructs a "good" bound on the performance which hold for all inputs excepts for those (called *black swans* as they are very improbable) in an exceptional set with exponentially small probability. This approach can help to explain the performance of many algorithms in practice, as near ill-posed inputs are uncommon.

Example. The power iterated method for computing an eigenvalue of an Hermitian method can be shown to have weak polynomial time complexity.

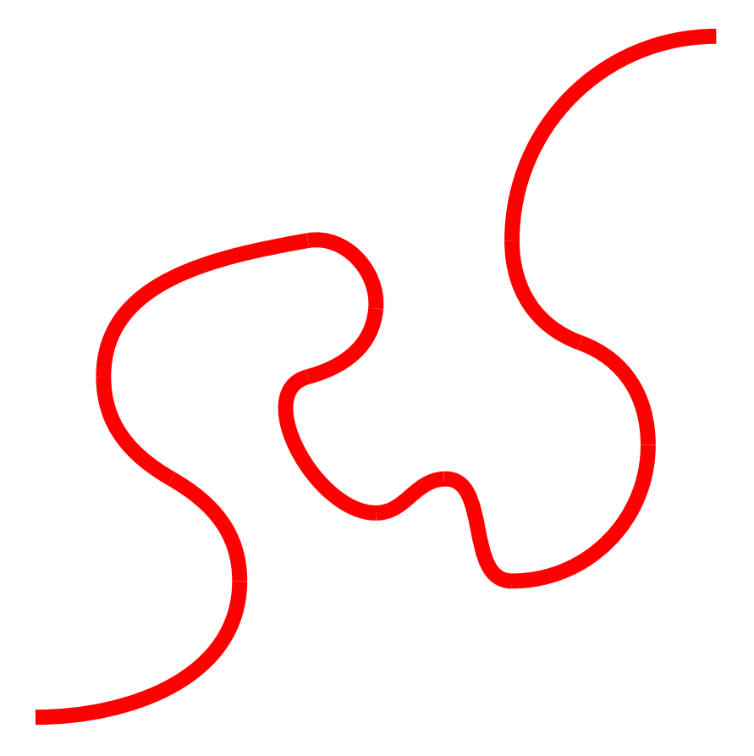
State of the art

Theorem. *There is a numerical algorithm, numerically stable, with input polynomial q-tuples f and lax formulas Φ (i.e. without "$<$", "$>$", "$\neq$" and "$\neg$") using the polynomial in f that computes the homology groups of the semialgebraic set described by Φ in weak exponential time for f uniformly distributed on the sphere with respect the Bombieri-Weyl norm.*

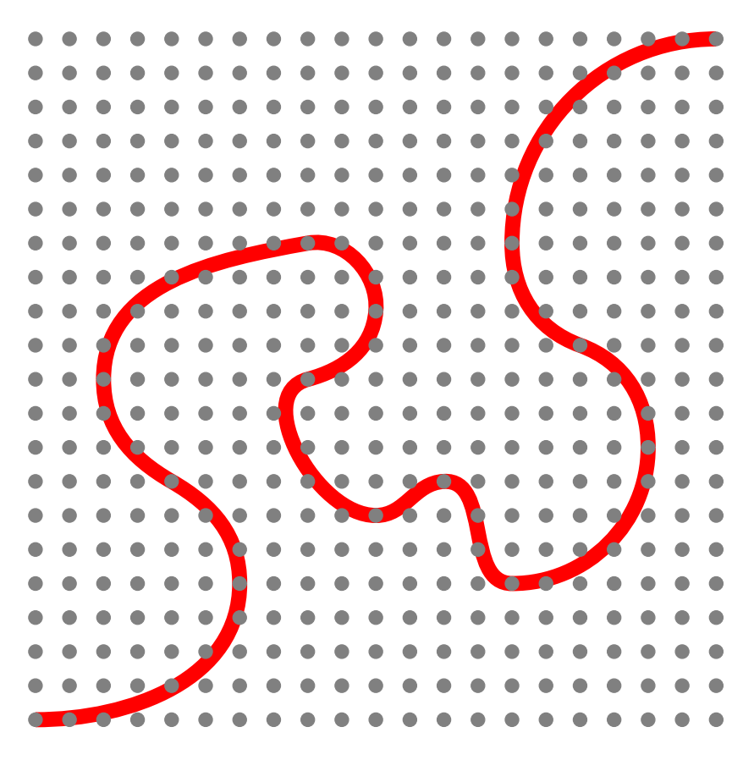
Even more, the algorithm can be parallelized so that it runs in weak polynomial time for f uniformly distributed on the sphere.

Core of the algorithm

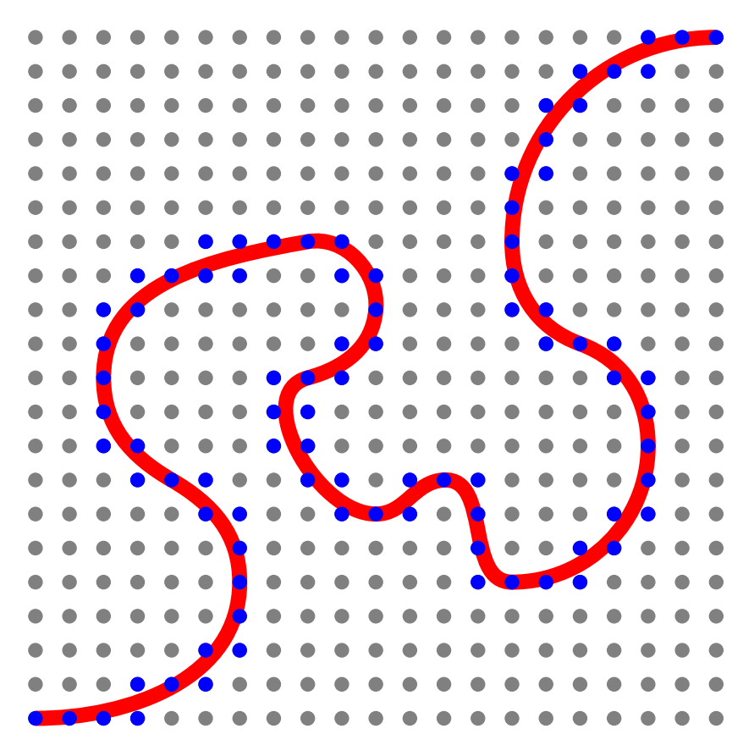
1. Take the geometric object.



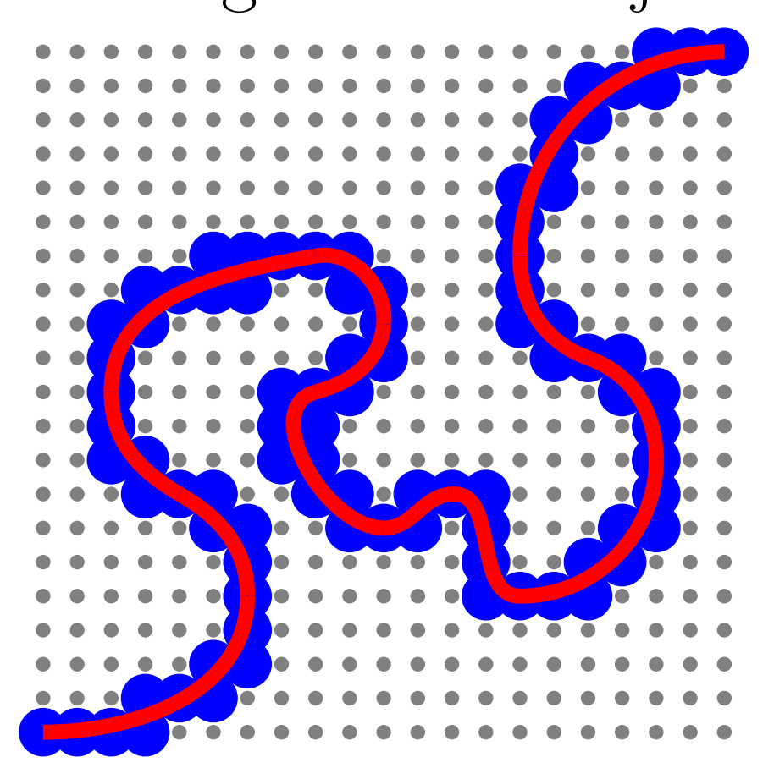
2. Construct a sufficiently fine grid. The size of the grid's mesh is controlled by the condition number.



3. Sample from the grid a cloud of points near enough the geometric object.



4. A thickening of the sampled cloud of points will capture the shape of the geometric object.



5. Standard techniques of algebraic topology allow the computation of the homology groups.

Open challenges

1. Can we extend the algorithm to work on general formulas and not only on lax formulas?
2. Can we extend the probabilistic analysis of the condition number for more general probability distributions of f?
3. Implement the algorithm and evaluate its performance in practice.

References for further reading

F. Cucker, T. Krick, M. Shub. Computing the Homology of Real Projective Sets. *Found Comput Math*, 2017.

P. Bürgisser, F. Cucker, P. Lairez. Computing the homology of basic semialgebraic sets in weak exponential time. To appear in *Journal of the ACM*, 2018.

P. Bürgisser, F. Cucker, J. Tonelli-Cueto. Computing the Homology of Semialgebraic Sets. I: The Lax Case. arXiv:1807.06435.

The research has been funded by Felipe Cucker's Einstein Visiting Fellowship "Complexity and accuracy of numerical algorithms in algebra and geometry" given by the

EINSTEIN Foundation.de

Einstein Stiftung Berlin
Einstein Foundation Berlin