Panoramas & Synthèses
24, 2007, p. 1–32

REAL ALGEBRAIC SETS

by

Michel Coste

***Abstract.* —** The aim of these notes is to present the material needed for the study of the topology of singular real algebraic sets via algebraically constructible functions. The first chapter reviews basic results of semialgebraic geometry, notably the triangulation theorem and triviality results which are crucial for the notion of link, which plays an important role in these notes. The second chapter presents some results on real algebraic sets, including Sullivan's theorem stating that the Euler characteristic of a link is even, and the existence of a fundamental class. The third chapter is devoted to constructible and algebraically constructible functions; the main tool which makes these functions useful is integration against Euler characteristic. We give an idea of how algebraically constructible functions give rise to combinatorial topological invariants which can be used to characterize real algebraic sets in low dimensions.

***Résumé* (Ensembles algébriques réels). —** Ces notes ont pour ambition d'expliquer les outils nécessaires à l'étude de la topologie des ensembles algébriques réels singuliers au moyen des fonctions algébriquement constructibles. La première section passe en revue les faits de base de la géométrie semialgébrique, notamment le théorème de triangulation et les résultats de trivialisation, cruciaux pour la notion d'entrelacs qui joue un rôle important dans ces notes. La deuxième section présente quelques résultats sur les ensembles algébriques réels, dont le théorème de Sullivan qui dit que la caractéristique d'Euler de l'entrelacs est paire, et l'existence d'une classe fondamentale. La troisième section est consacrée aux fonctions constructibles et algébriquement constructibles ; l'outil principal qui rend ces fonctions utiles est l'intégration par rapport à la caractéristique d'Euler. On donne une idée de la façon dont les fonctions algébriquement constructibles donnent des invariants topologiques combinatoires qui permettent de caractériser les ensembles algébriques réels de petite dimension.

1. Semialgebraic sets

In this section we present some basic topological facts concerning semialgebraic sets, which are subsets of $\mathbb{R}^n$ defined by combinations of polynomial equations and

2000 *Mathematics Subject Classification.* — 14P25, 14B05, 14P10.
***Key words and phrases.* —** Real algebraic sets, semi-algebraic sets, Euler integral, algebraically constructible functions, topological invariants.

inequalities. One of the main properties is the fact that a compact semialgebraic set can be triangulated. We introduce the notion of link, which is an important invariant in the local study of singular semialgebraic sets. We also define a variant of Euler characteristic on the category of semialgebraic sets, which satisfies nice additivity properties (this will be useful for integration in section 3).

We do not give here the proofs of the main results. We refer the reader to [**2**], [**3**] and [**5**].

Most of the results presented here hold also for definable sets in o-minimal structures (this covers, for instance, sets defined with the exponential function and with any real analytic function defined on a compact set). We refer the reader to [**6**] and [**4**].

1.1. Semialgebraic sets, Tarski-Seidenberg. — A *semialgebraic subset* of $\mathbb{R}^n$ is the subset of $(x_1, \dots, x_n)$ in $\mathbb{R}^n$ satisfying a boolean combination of polynomial equations and inequalities with real coefficients. In other words, the semialgebraic subsets of $\mathbb{R}^n$ form the smallest class $\mathcal{S}A_n$ of subsets of $\mathbb{R}^n$ such that:

1) If $P \in \mathbb{R}[X_1, \dots, X_n]$, then

$$\big\{x \in \mathbb{R}^n \;;\; P(x) = 0\big\} \in \mathcal{S}A_n \quad \text{and} \quad \big\{x \in \mathbb{R}^n \;;\; P(x) > 0\big\} \in \mathcal{S}A_n.$$

2) If $A \in \mathcal{S}A_n$ and $B \in \mathcal{S}A_n$, then

$$A \cup B \in \mathcal{S}A_n, \quad A \cap B \in \mathcal{S}A_n \quad \text{and} \quad \mathbb{R}^n \setminus A \in \mathcal{S}A_n.$$

The fact that a subset of $\mathbb{R}^n$ is semialgebraic does not depend on the choice of affine coordinates. Some stability properties of the class of semialgebraic sets follow immediately from the definition.

1) All algebraic subsets of $\mathbb{R}^n$ are in $\mathcal{S}A_n$. Recall that an algebraic subset is a subset defined by a finite number of polynomial equations

$$P_1(x_1, \dots, x_n) = \cdots = P_k(x_1, \dots, x_n) = 0.$$

2) $\mathcal{S}A_n$ is stable under the boolean operations, i.e. finite unions and intersections and taking complement. In other words, $\mathcal{S}A_n$ is a Boolean subalgebra of the powerset $\mathcal{P}(\mathbb{R}^n)$.
3) The cartesian product of semialgebraic sets is semialgebraic. If $A \in \mathcal{S}A_n$ and $B \in \mathcal{S}A_p$, then $A \times B \in \mathcal{S}A_{n+p}$.

Sets are not sufficient, we need also maps. Let $A \subset \mathbb{R}^n$ be a semialgebraic set.

A map $f : A \to \mathbb{R}^p$ is said to be *semialgebraic* if its graph $\Gamma(f) \subset \mathbb{R}^n \times \mathbb{R}^p = \mathbb{R}^{n+p}$ is semialgebraic.

For instance, the polynomial maps and the regular maps (i.e. those maps whose coordinates are rational functions such that the denominator does not vanish) are semialgebraic. The function $x \mapsto \sqrt{1 - x^2}$ for $|x| \leq 1$ is semialgebraic.

The most important stability property of semialgebraic sets is known as "Tarski-Seidenberg theorem". This central result in semialgebraic geometry is not obvious from the definition.

***Theorem 1.1* (Tarski-Seidenberg).** — *Let A be a semialgebraic subset of $\mathbb{R}^{n+1}$ and*

$$\pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n,$$

the projection on the first n coordinates. Then $\pi(A)$ is a semialgebraic subset of $\mathbb{R}^n$.

It follows from the Tarski-Seidenberg theorem that images and inverse images of semialgebraic sets by semialgebraic maps are semialgebraic. Also, the composition of semialgebraic maps is semialgebraic. Other consequences are the following.

Let $A \subset \mathbb{R}^n$ be a semialgebraic set; then its closure $\mathrm{clos}(A)$ is semialgebraic and the function "distance to A" on $\mathbb{R}^n$ is semialgebraic.

- ▷ A *Nash manifold* $M \subset \mathbb{R}^n$ is an analytic submanifold which is a semialgebraic subset.
- ▷ A *Nash map* $M \to \mathbb{R}^p$ is a map which is analytic and semialgebraic.

1.2. Cell decomposition and stratification. — The semialgebraic subsets of the line are very simple to describe: they are the finite unions of points and open intervals. We cannot hope for such a simple description of semialgebraic subsets of $\mathbb{R}^n$, $n > 1$. However, we have that every semialgebraic set has a finite partition into semialgebraic subsets diffeomorphic to open boxes (i.e. cartesian product of open intervals). We give a name to these pieces:

***Definition 1.2*.** — *A* (Nash) cell *in $\mathbb{R}^n$ is a (Nash) submanifold of $\mathbb{R}^n$ which is (Nash) diffeomorphic to an open box $(-1,1)^d$ (d is the* dimension *of the cell).*

Every semialgebraic set can be decomposed into a disjoint union of Nash cells. More precisely:

***Theorem 1.3*.** — *Let $A_1, \dots, A_p$ be semialgebraic subsets of $\mathbb{R}^n$. Then there exists a finite semialgebraic partition of $\mathbb{R}^n$ into Nash cells such that each A_j is a union of some of these cells.*

This cell decomposition is a consequence of the so-called "cylindrical algebraic decomposition" (cad), which is the main tool in the study of semialgebraic sets. Actually, the Tarski-Seidenberg theorem can be proved by using cad.

A *cad* of $\mathbb{R}^n$ is a partition of $\mathbb{R}^n$ into finitely many semialgebraic subsets (the cells of the cad), satisfying certain properties. We define a cad of $\mathbb{R}^n$ by induction on n.

- ▷ A cad of $\mathbb{R}$ is a subdivision by finitely many points $a_1 < \cdots < a_\ell$. The cells are the singletons $\{a_i\}$ and the open intervals delimited by these points.

- ▷ For $n > 1$, a cad of $\mathbb{R}^n$ is given by a cad of $\mathbb{R}^{n-1}$ and, for each cell C of $\mathbb{R}^{n-1}$, Nash functions
$$\zeta_{C,1} < \cdots < \zeta_{C,\ell_C} : C \longrightarrow \mathbb{R}.$$
The cells of the cad of $\mathbb{R}^n$ are the graphs of the $\zeta_{C,j}$ and the bands in the cylinders $C \times \mathbb{R}$ delimited by these graphs.

Observe that every cell of a cad is indeed Nash diffeomorphic to an open box. This is easily proved by induction on n.

The main result about cad is that, given any finite family $A_1, \ldots, A_p$ of semialgebraic subsets of $\mathbb{R}^n$, one can construct a cad of $\mathbb{R}^n$ such that every A_j is a union of cells of this cad. This gives theorem 1.3.

Moreover, the cell decomposition of theorem 1.3 can be assumed to be a *stratification*: this means that for each cell C, the closure $\mathrm{clos}(C)$ is the union of C and of cells of smaller dimension. This property of incidence between the cells may not be satisfied by a cad (where the cells have to be arranged in cylinders whose directions are given by the coordinate axes), but it can be obtained after a generic linear change of coordinates in $\mathbb{R}^n$. In addition, one can ask the stratification to satisfy a local triviality condition.

Definition 1.4. — *Let $\mathcal{S}$ a finite stratification of $\mathbb{R}^n$ into Nash cells; then we say that $\mathcal{S}$ is* locally (semialgebraically) trivial *if for every cell C of $\mathcal{S}$, there exist a neighborhood U of C and a* (*semialgebraic*) *homeomorphism*
$$h : U \longrightarrow C \times F,$$
where F is the intersection of U with the normal space to C at a generic point of C, and, for every cell D of $\mathcal{S}$, $h(D \cap U) = C \times (D \cap F)$.

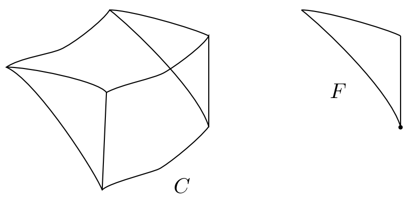

FIGURE 1. Local triviality: a neighborhood of C is homeomorphic to $C \times F$.

Theorem 1.5. — *Let $A_1, \ldots, A_p$ be semialgebraic subsets of $\mathbb{R}^n$. Then there exists a finite semialgebraic stratification $\mathcal{S}$ of $\mathbb{R}^n$ into Nash cells such that*

- ▷ *$\mathcal{S}$ is locally trivial,*
- ▷ *every A_j is a union of cells of $\mathcal{S}$.*

1.3. Connected components, dimension. — Every Nash cell is obviously arc-wise connected. Hence, from the decomposition of a semialgebraic set into finitely many Nash cells, we obtain:

Proposition 1.6. — *A semialgebraic set has finitely many connected components, which are semialgebraic.*

The cell decomposition also leads to the definition of the dimension of a semi-algebraic set as the maximum of the dimensions of its cells. This works well.

Proposition 1.7. — *Let $A \subset \mathbb{R}^n$ be a semialgebraic set, and let $A = \bigcup_{i=1}^{p} C_i$ be a decomposition of A into a disjoint union of Nash cells C_i. The number*

$$\max \big\{ \dim(C_i),\ i = 1, \ldots, p \big\}$$

does not depend on the decomposition. The dimension *of A is defined to be this number.*

The dimension is even invariant by any semialgebraic bijection (not necessarily continuous):

Proposition 1.8. — *Let A be a semialgebraic subset of $\mathbb{R}^n$, and $f : S \to \mathbb{R}^k$ a semi-algebraic map (not necessarily continuous). Then*

$$\dim f(A) \leq \dim A.$$

If f is one-to-one, then

$$\dim f(A) = \dim A.$$

Using a stratification, we obtain immediately the following result.

Proposition 1.9. — *Let A be a semialgebraic subset of $\mathbb{R}^n$. Then*

$$\dim \big(\operatorname{clos}(A) \setminus A \big) < \dim(A).$$

1.4. Triangulation. — First we fix the notation. A k-simplex σ in $\mathbb{R}^n$ (where $0 \leq k \leq n$) is the convex hull of $k+1$ points $a_0, \ldots, a_k$ which are not contained in a $(k-1)$-affine subspace; the points a_i are the vertices of σ. A (proper) face of σ is a simplex whose vertices form a (proper) subset of the set of vertices of σ. The open simplex $\mathring{\sigma}$ associated to a simplex σ is σ minus the union of its proper faces. A map $f : \sigma \to \mathbb{R}^k$ is called linear if $f(\sum_{i=0}^{k} \lambda_i a_i) = \sum_{i=0}^{k} \lambda_i f(a_i)$ for every $(k+1)$-tuple of nonnegative real numbers λ_i such that $\sum_{i=0}^{k} \lambda_i = 1$.

A finite simplicial complex in $\mathbb{R}^n$ is a finite set K of simplices such that

- ▹ every face of a simplex of K is in K,
- ▹ the intersection of two simplices in K is either empty or a common face of these two simplices.