ON THE NUMBER OF CELLS DEFINED BY A FAMILY OF POLYNOMIALS ON A VARIETY

SAUGATA BASU, RICHARD POLLACK AND MARIE-FRANÇOISE ROY

Abstract. Let R be a real closed field and $\mathcal{V}$ a variety of real dimension k' which is the zero set of a polynomial $Q \in R[X_1, \ldots, X_k]$ of degree at most d. Given a family of s polynomials $\mathcal{P} = \{P_1, \ldots, P_s\} \subset R[X_1, \ldots, X_k]$ where each polynomial in $\mathcal{P}$ has degree at most d, we prove that the number of cells defined by $\mathcal{P}$ over $\mathcal{V}$ is $\binom{s}{k'}(O(d))^k$. Note that the combinatorial part of the bound depends on the dimension of the variety rather than on the dimension of the ambient space.

§1. *Introduction.*

§1.1. *Notation.* A *sign condition* for a set of s polynomials $\mathcal{P} = \{P_1, \ldots, P_s\}$ is a vector $\sigma \in \{-1, 0, +1\}^s$ and the sign condition σ is called *strict* if $\sigma \in \{-1, +1\}^s$. We call the sign condition σ *non-empty over a variety* $\mathcal{V}$ with respect to $\mathcal{P}$ if there is a point $x \in \mathcal{V}$ which *realizes* the sign condition, *i.e.*, $(\text{sign}(P_1(x)), \ldots, \text{sign}(P_s(x))) = \sigma$.

The set, $\sigma_{\mathcal{P},\mathcal{V}} = \{x | x \in \mathcal{V}, (\text{sign}(P_1(x)), \ldots, \text{sign}(P_s(x))) = \sigma\}$ is the *realization* space of σ over $\mathcal{V}$ with respect to $\mathcal{P}$ and its non-empty semi-algebraically connected components are the *cells* of the sign condition σ for $\mathcal{P}$ over $\mathcal{V}$. The number of these cells is denoted by $|\sigma_{\mathcal{P},\mathcal{V}}|$ and thus

$$C(\mathcal{P}, \mathcal{V}) = \sum_{\sigma_{\mathcal{P},\mathcal{V}} \neq \emptyset} |\sigma_{\mathcal{P},\mathcal{V}}|$$

is the number of cells defined by $\mathcal{P}$ over $\mathcal{V}$.

We write $f(d, k, k', s)$ for the maximum of $C(\mathcal{P}, \mathcal{V})$ over all varieties, $\mathcal{V} \subset R^k$ of dimension k', defined by polynomial equations of degree at most d and over all $\mathcal{P}$ consisting of s polynomials in k variables, each of degree at most d.

Remark 1. It is no restriction to consider only varieties defined by a single polynomial. If the variety is the zero set of a finite family of polynomial $\mathcal{Q}$ we can just as well consider the zero set of the single polynomial $Q = \sum_{q \in \mathcal{Q}} q^2$.

§1.2. *Background.* Previous work considered only the case $k = k'$. In particular, the problem of determining the complexity of an arrangement of s hyperplanes in R^k, which is the same as determining $f(1, k, k, s)$, is well known to be $\Theta(\binom{s}{k})$ (see [8] for example). This bound has played an important role in discrete and computational geometry for many years.

For $f(d, k, k, s)$, the best bound had been $(sd)^{O(k)}$, which was based on a result of Heintz [10]. Since the set of cells of sd hyperplanes is the same as the set of cells of s polynomials, each the product of d of the given linear polynomials, a lower bound of $\Omega(\binom{sd}{k})$ follows. This lower bound was recently shown to be an upper bound as well [14].

For the case $f(1, k, k', s)$, the variety is a k'-flat and we can linearly eliminate $k-k'$ variables. This reduces the problem to that of bounding $f(1, k', k', s) = \Theta(\binom{s}{k'})$.

Our result is

THEOREM 1. $f(d, k, k', s) = \binom{s}{k'}(O(d))^k$.

The main contribution of this paper is that the bound $\binom{s}{k'}$ on the combinatorial part of $f(d, k, k', s)$ depends only on k' and not at all on k. We have seen that this bound is sharp for the case $d=1$. The bound of $(O(d))^k$ on the algebraic part of $f(d, k, k', s)$ is also sharp in the case $k'=0$ and matches the known upper bounds for arbitrary k' that follow from the well known results of Milnor–Oleĭnik–Petrovsky–Thom [11, 12, 13, 16].

The ideas that make possible the separation of this bound into a combinatorial part and an algebraic part have also played a key role in recent improvements for related algorithmic problems [1, 2, 3, 5, 6, 7].

Our bound has proved useful in a recent result in geometric transversal theory [9]. There, the relevant variety $\mathscr{V}$ is the Grassmannian $G_{k,d}$ of k subspaces of R^d.

§1.3. *Outline of the argument.* In our argument, we perturb the polynomials using various *infinitesimals*. We then use basic properties of the field of Puiseux series in these infinitesimals. We write $R\langle\varepsilon\rangle$ for the real closed field of Puiseux series in ε with coefficients in R [4]. This field is *uniquely* orderable in the following way: the sign of an element in this field agrees with the sign of the coefficient of its lowest degree term in ε. This order makes ε positive and smaller than any positive element of R. We also iterate this notation in the usual way so that $R\langle\varepsilon_1, \varepsilon_2\rangle = R\langle\varepsilon_1\rangle\langle\varepsilon_2\rangle$ and, thus, $1 \gg \varepsilon_1 \gg \varepsilon_2$ *i.e.*, ε_1 is smaller than any positive element of R and ε_2 is positive and smaller than any positive element in $R\langle\varepsilon_1\rangle$. The valuation ring, V, consists of those Puiseux series that are bounded over R *i.e.*, the Puiseux series with no negative powers of ε. The map $\text{eval}_\varepsilon : V \to R$ maps an element of V to its constant term.

If R' is a real closed field extending R, and S is a semi-algebraic set defined over R, then we denote by $S_{R'}$ the solution set in R'^k of the same polynomial equalities and inequalities that define S. Both S and $S_{R'}$, the *extension* of S to R', have the same number of semi-algebraically connected components [4].

Throughout the paper, a *cell* of a semi-algebraic S set will be a non-empty semi-algebraic connected component of S (see [4]).

The idea of the proof of our theorem is to first observe (in Proposition 1) that the extension of every cell of a sign condition for $\mathscr{P}$ over $\mathscr{V}$ to $R\langle\varepsilon\rangle$ contains a cell of an algebraic set defined by a set of equalities chosen from the extended family of polynomials $\mathscr{P}' = \bigcup_{P\in\mathscr{P}} \{P-\varepsilon, P, P+\varepsilon\}$. Thus, the cells defined by $\mathscr{P}$ on $\mathscr{V}$ are all accounted for by counting the number of cells in each algebraic set determined by Q and some subset of $\mathscr{P}'$. Recall that, by the Milnor–Oleǐnik–Petrovsky–Thom bounds **[11, 12, 13, 16]**, any of these algebraic sets has at most $O(d)^k$ cells. We make the observation that if the family $\mathscr{P}'$ is in *general position with respect to* $\mathscr{V}$, *i.e.*, no more than k' polynomials of $\mathscr{P}'$ have a common zero on $\mathscr{V}$, then the number of cells defined by $\mathscr{P}$ on $\mathscr{V}$ is at most $\binom{3s}{k'}O(d)^k$ and our claimed bound would follow.

With this in mind, we perturb the set of polynomials $\mathscr{P}$ with infinitesimals $1/\Omega \gg \delta_1 \gg \ldots \gg \delta_s \gg \delta$ to obtain the family of polynomials $\mathscr{P}^* = \bigcup_{1\leqslant i\leqslant s} \{P_i-\delta_i, P_i+\delta_i, P_i-\delta\delta_i, P_i+\delta\delta_i\}$ and show, in Corollary 1 that $\mathscr{P}^*$ is in general position with respect to $\mathscr{V}$ so that we obtain the claimed bound for the family $\mathscr{P}^*$. We then show (Proposition 4) that the extension of every cell defined by $\mathscr{P}$ over $\mathscr{V}$ to $R\langle\delta_1 \ldots \delta_s\rangle$ contains the image under the eval$_\delta$ map of a cell of this perturbed family. Since we also know (Proposition 3) that the eval map takes bounded semi-algebraically connected sets to semi-algebraically connected sets, it follows that the number of cells of this perturbed family $\mathscr{P}^*$ bounds the number of cells of the original family $\mathscr{P}$.

We wish to acknowledge many useful conversations with Michel Coste. The first named author thanks the US NSF for a grant; the second thanks the US NSF and NSA for grants; and the third thanks the European Esprit programme for a grant.

§2. *Propositions and proofs.*

PROPOSITION 1. *Let C be a cell of a semi-algebraic set of the form $P_1 = \ldots = P_l = 0$, $P_{l+1} > 0, \ldots, P_s > 0$, then we can find an algebraic set V in $R\langle\varepsilon\rangle^k$ defined by equations $P_1 = \ldots = P_l = P_{i_1} - \varepsilon = \ldots = P_{i_m} - \varepsilon = 0$, such that a cell of V, say C', is contained in $C_{R\langle\varepsilon\rangle}$.*

Proof. If C is closed, it is a cell of the algebraic set defined by $P_1 = \ldots = P_l = 0$. If not, consider Γ, the set of all semi-algebraic paths γ in R^k going from some point $x(\gamma)$ in C to a $y(\gamma)$ in $\bar{C}\setminus C$ such that $\gamma\setminus\{y(\gamma)\}$ is entirely contained in C. For each $\gamma\in\Gamma$, there is an $i>l$ such that P_i vanishes at $y(\gamma)$. Then on $\gamma_{R\langle\varepsilon\rangle}$ there is a point $z(\gamma, \varepsilon)$ and an $i>l$ such that $P_i-\varepsilon$ vanishes at $z(\gamma,\varepsilon)$ and that on the portion of the path between x and $z(\gamma,\varepsilon)$ no such $P_i-\varepsilon$ with $i>l$ vanishes. Let $I_\gamma = \{i \mid i>l, P_i(z(\gamma,\varepsilon))-\varepsilon=0\}$. Now choose a path $\gamma\in\Gamma$ so that the set $I_\gamma = \{i_1, \ldots, i_m\}$ is maximal under set inclusion and let V be defined by $P_1 = \ldots = P_l = P_{i_1} - \varepsilon = \ldots = P_{i_m} - \varepsilon = 0$.

It is clear that at $z(\gamma,\varepsilon)$, defined above, we have $P_{l+1}>0, \ldots, P_s>0$ and $P_j-\varepsilon>0$ for every $j\notin I_\gamma$ which is $>l$. Let C' be the cell of V containing $z(\gamma,\varepsilon)$. We shall prove that no polynomial $P_{l+1}, \ldots, P_s$ vanishes on this cell, and thus that C' is contained in $C_{R\langle\varepsilon\rangle}$. Suppose not, then some new P_i $(i>l, i\notin I_\gamma)$ vanishes on C', say at y_ε. We can suppose without loss of generality that the coordinates of y_ε are algebraic over $R[\varepsilon]$. Take a semi-algebraic path γ_ε defined over $R[\varepsilon]$ connecting $z(\gamma,\varepsilon)$ to y_ε with $\gamma_\varepsilon\subset C'$. Denote by $z(\gamma_\varepsilon,\varepsilon)$ the first point of γ_ε with $P_1 = \ldots = P_l = P_{i_1} - \varepsilon = \ldots = P_{i_m} - \varepsilon = P_j - \varepsilon = 0$ for some new j not in I_γ.

For t in R small enough, the set γ_t (obtained by replacing ε by t in γ_ε) defines a semi-algebraic path from $z(\gamma, t)$ to $z(\gamma_\varepsilon, t)$ contained in C. Replacing ε by t in the Puiseux series which give the coordinates of $z(\gamma_\varepsilon, \varepsilon)$ defines a path γ' containing $z(\gamma_\varepsilon, \varepsilon)$ from $z(\gamma_\varepsilon, t)$ to $y = \text{eval}\,(z(\gamma_\varepsilon, \varepsilon))$ (which is a point of $\bar{C} \backslash C$). Let us consider the new path γ^* consisting of the beginning of γ (up to the point z_t for which $P_{i_1} = \ldots, P_{i_m} = t$), followed by γ_t and then followed by γ'. Now the first point in γ^* such that there exists a new j with $P_j - \varepsilon = 0$ is $z(\gamma_\varepsilon, \varepsilon)$ and thus $\gamma^* \in \Gamma$ with I_{γ^*} strictly larger than I_γ. This is impossible by the maximality of I_γ.

Remark 2. Somewhat more is true. It is easy to see that $\text{eval}_\varepsilon\,(C') \neq \emptyset$. That is to say that C' contains points bounded over R. In consequence, if we know that $\mathscr{P}$ is in general position with respect to $\mathscr{V}$ we need only consider the zero sets of at most k' polynomials chosen from $\mathscr{P}'$. If more than k' polynomials in $\mathscr{P}'$ had a common zero *bounded over* R, then its eval would be a point on $\mathscr{V}$ satisfying more than k' polynomials in $\mathscr{P}$ which is impossible. This does not mean that if $\mathscr{P}$ is in general position with respect to $\mathscr{V}$ then $\mathscr{P}'$ is in general position with respect to $\mathscr{V}$. It only means that these additional zeros are not bounded over R.

PROPOSITION 2. *Given a family* $\{P_1, \ldots, P_s\}$ *of polynomials in* $R[X_1, \ldots, X_k]$ *and a variety* $\mathscr{V}$ *of real dimension* k', *let* R' *be a real closed field containing* R, *and let* $\delta_1, \ldots, \delta_s$, *be elements of* R' *that are algebraically independent over* R. *Then the perturbed family* $\mathscr{P}^* = \bigcup_{1 \leq i \leq s} \{P_i - \delta_i\}$, *is in general position with respect to the variety* $\mathscr{V}_{R'}$.

Proof. The result follows from the following simple observations.

If $\mathscr{V}$ has real dimension k' then $\mathscr{V}$ is the union of a finite number of semi-algebraically connected semi-algebraic sets of real dimension less than or equal to k' whose Zariski closures are irreducible [**4**].

If C is a semi-algebraically connected semi-algebraic set whose Zariski closure is irreducible then any polynomial is either constant on C or its zero set meets C in a semi-algebraic set of real dimension less than the dimension of C. This is immediate from the definition of irreducibility.

As a consequence, we see that the zero set of any of the perturbed polynomials meets the variety $\mathscr{V}$ in a variety of lower real dimension. The proposition is proved by repeating this argument at most k' times.

COROLLARY 1. *Given a family* $\{P_1, \ldots, P_s\}$, *of polynomials in* $R[X_1, \ldots, X_k]$ *and a variety* $\mathscr{V}$ *of real dimension* k', *let* R' *be a real closed field containing* R, *and let* $\delta, \delta_1, \ldots, \delta_s$, *be elements of* R' *algebraically independent*

https://doi.org/10.1112/S0025579300011621 Published online by Cambridge University Press

over R. Then the perturbed family

$$\mathscr{P}^* = \bigcup_{1 \leqslant i \leqslant s} \{P_i - \delta_i, P_i + \delta_i, P_i - \delta\delta_i, P_i + \delta\delta_i\} \cup \left\{ \sum_{1 \leqslant i \leqslant k} X_i^2 - \Omega^2 \right\}$$

is in general position with respect to the variety $\mathscr{V}_{R'}$.

PROPOSITION 3. *If $S' \subset R\langle\varepsilon\rangle^k$ is a semi-algebraic set defined over $R[\varepsilon]$ and $S = \text{eval}_\varepsilon(S')$, then S is a semi-algebraic set. Moreover, if S′ is bounded over R and is semi-algebraically connected then S is semi-algebraically connected.*

Proof. Suppose that $S' \subset (R\langle\varepsilon\rangle)^k$ is described by a quantifier-free formula $\Phi(\varepsilon)(X_1, \ldots, X_k)$. Introduce a new variable X_{k+1} and denote by $\Phi(X_1, \ldots, X_k, X_{k+1})$ the result of substituting X_{k+1} for ε in $\Phi(\varepsilon)(X_1, \ldots, X_k)$. Embed R^k in R^{k+1} by sending $(X_1, \ldots, X_k)$ to $(X_1, \ldots, X_k, 0)$. Thus, S is a subset of $Z(X_{k+1})$. We prove that $S = \bar{T} \cap Z(X_{k+1})$ where

$$T = \{(x_1, \ldots, x_k, x_{k+1}) \in R^{k+1} | \Phi((x_1, \ldots, x_k, x_{k+1}) \text{ and } x_{k+1} > 0\}$$

and $\bar{T}$ is the closure of T in the euclidean topology.

If $x \in S$ there is a $z \in S'$ such that $\text{eval}_\varepsilon(z) = x$. Let $B_x(r)$ denote the open ball of radius r centred at x. Since (z, ε) belongs to the extension of $B_x(r) \cap T$ to $R\langle\varepsilon\rangle$ it follows that $B_x(r) \cap T$ is non-empty, and hence that $x \in \bar{T}$.

Conversely, let x be in $\bar{T} \cap Z(X_{k+1})$. The semi-algebraic curve selection lemma [4] asserts the existence of a semi-algebraic function f from [0, 1] to $\bar{T}$ with $f(0) = x$ and $f((0, 1]) \subset T$. This semi-algebraic function defines a point $z = f(\varepsilon)$ whose coordinates lie in $R\langle\varepsilon\rangle$ and belongs to S' and moreover $\text{eval}_\varepsilon(z) = x$.

If S' is bounded by M in R and semi-algebraically connected then there exists a positive t in R such that $T \cap (B_0(M) \times [0, t])$ is semi-algebraically connected. It follows easily that $S = \bar{T} \cap Z(X_{k+1})$ is semi-algebraically connected.

PROPOSITION 4. *Let C be a non-empty cell in $\mathscr{V} = Z(Q)$, of the semi-algebraic set defined by $P_1 = \ldots = P_l = 0$, $P_{l+1} > 0, \ldots, P_s > 0$, and let C′ be the extension of C to $R\langle\delta_1, \ldots, \delta_s\rangle$. Then C′ contains some $\text{eval}_\delta(C'')$, where C″ is a cell of the semi-algebraic set defined by the sign conditions*

$$(*)\begin{cases} Q = 0, \quad -\delta\delta_1 < P_1 < \delta\delta_1, \ldots, -\delta\delta_l < P_l < \delta\delta_l, \\ \qquad P_{l+1} > \delta_{l+1}, \ldots, P_s > \delta_s, \\ \qquad\qquad X_1^2 + \ldots X_k^2 < 1, \end{cases}$$

over $R\langle 1/\Omega, \delta_1, \ldots, \delta_s, \delta\rangle$.

Proof. If $x \in C$, then x satisfies $(*)$. Let C'' be the cell of the semi-algebraic set in $(R\langle 1/\Omega, \delta_1, \ldots, \delta_s, \delta\rangle)^k$ defined by these equalities and inequalities, which contains x.

It is clear that $\mathrm{eval}_\delta(C'')$ is contained in the semi-algebraic set defined by the sign condition $Q=P_1=\ldots=P_l=0$, $P_{l+1}>0,\ldots,P_s>0$, in $(R\langle\delta_1,\ldots,\delta_s\rangle)^k$ and that it also contains $x\in C'$. Since, by Proposition 3, $\mathrm{eval}_\delta(C'')$ is also semi-algebraically connected the statement of the lemma follows.

§2.1. *Proof of the theorem.* The family of polynomials,

$$\mathscr{P}^*=\bigcup_{1\leqslant i\leqslant s}\{P_i-\delta_i, P_i+\delta_i, P_i-\delta\delta_i, P_i+\delta\delta_i\}\cup\left\{\sum_{1\leqslant i\leqslant k}X_i^2-\Omega^2\right\}$$

is in general position with respect to $\mathscr{V}$ by Corollary 1. Hence, by Proposition 1, the extension of every cell of a *strict* sign condition for $\mathscr{P}^*$ over $\mathscr{V}$ to $R\langle\delta_1,\ldots,\delta_s,\delta,\varepsilon\rangle$ contains a cell of an algebraic variety defined by $\{Q\}\cup\bar{\mathscr{P}}^*$ where $\bar{\mathscr{P}}^*$ is a subset of $\bigcup_{P\in\mathscr{P}^*}\{P-\varepsilon, P, P+\varepsilon\}$. As noted in Remark 2, we can assume that the cardinality of $\bar{\mathscr{P}}^*$ is at most k'. There are $\sum_{1\leqslant l\leqslant k'}\binom{12s}{l}=\binom{O(s)}{k'}$ of these varieties and each has at most $O(d)^k$ cells by the well-known bounds of Milnor–Oleĭnik–Petrovsky–Thom [11, 12, 13, 16]. Hence the number of cells of strict sign conditions for $\mathscr{P}^*$ over $\mathscr{V}$ is $\binom{s}{k'}O(d)^k$. Finally, by Proposition 4, the extension of each cell of a sign condition for $\mathscr{P}$ over $\mathscr{V}$ to $R\langle\delta_1,\ldots,\delta_s\rangle$ contains the eval_δ of one of these $\binom{s}{k'}O(d)^k$ cells of strict sign conditions for $\mathscr{P}^*$ over $\mathscr{V}$. Since these are semi-algebraically connected by Proposition 3 it follows that there are no more than $\binom{s}{k'}O(d)^k$ cells defined by $\mathscr{P}$ over $\mathscr{V}$.

References

1. S. Basu, R. Pollack and M.-F. Roy. A new algorithm to find a point in every cell defined by a family of polynomials. In *Quantifier Elimination and Cylindrical Algebraic Decomposition*, edited by B. Caviness and J. Johnson (Springer-Verlag) to appear.
2. S. Basu, R. Pollack and M.-F. Roy. Computing points meeting every cell on a variety. In *The Algorithmic Foundations of Robotics*, edited by K. Goldberg, D. Halperin, J. C. Latombe and R. Wilson (A. K. Peters, Boston, MA, 1955), 537–555.
3. S. Basu, R. Pollack and M.-F. Roy. On the combinatorial and algebraic complexity of Quantifier Elimination. In *Proc. 35th Annual IEEE Sympos. on the Foundations of Computer Science*, 632–641 (1994).
4. J. Bochnak, M. Coste and M.-F. Roy. *Géométrie algébrique réelle* (Springer-Verlag, 1987).
5. J. Canny. Some Practical Tools for Algebraic Geometry. In *Technical report in Spring school on robot motion planning* (PROMOTION ESPRIT, 1993).
6. J. Canny. Computing road maps in general semi-algebraic sets. *The Computer Journal*, 36 (1993), 504–514.
7. J. Canny. Improved algorithms for sign determination and existential quantifier elimination. *The Computer Journal*, 36 (1993), 409–418.
8. H. Edelsbrunner. *Algorithms in Combinatorial Geometry* (Springer-Verlag, Berlin, 1987).
9. J. E. Goodman, R. Pollack and R. Wenger. Bounding the number of geometric permutations induced by k-transversals. In *Proc. 10th Ann. ACM Sympos. Comput. Geom.* (1994), 192–197.
10. J. Heintz, M.-F. Roy and P. Solernó. On the complexity of semi-algebraic sets. In *Proc. IFIP San Francisco* (North-Holland, 1989), 293–298.
11. J. Milnor. On the Betti numbers of real varieties. *Proc. Amer. Math. Soc.*, 15 (1964), 275–280.
12. O. A. Oleĭnik. Estimates of the Betti numbers of real algebraic hypersurfaces (Russian). *Mat. Sb. (N.S.)*, 28 (70) (1951), 635–640.

13. I. G. Petrovsky and O. A. Oleĭnik. On the topology of real algebraic surfaces. *Izvestiaya Akademii Nauk SSSR. Serija Matematičeskaya*, 13 (1949), 389–402.
14. R. Pollack and M.-F. Roy. On the number of cells defined by a set of polynomials. *C. R. Acad. Sci. Paris*, 316 (1993), 573–577.
15. J. Renegar. On the computational complexity and geometry of the first order theory of the reals. *J. of Symbolic Comput.*, 13 (1992), 255–352.
16. R. Thom. Sur l'homologie des variétés algébriques réelles. In *Differential and Combinatorial Topology* (Princeton University Press, Princeton, 1965), 255–265.

Dr. S. Basu,
Courant Institute
of Mathematical Science,
New York University,
New York, NY 10012,
U.S.A.

Professor R. Pollack,
Courant Institute
of Mathematical Science,
New York University,
New York, NY 10012,
U.S.A.

Professor M.-F. Roy,
IRMAR (URA CNRS 305),
Université de Rennes,
Campus de Beaulieu,
35042 Rennes cedex,
France.

14P10: *ALGEBRAIC GEOMETRY; Real algebraic and real analytic geometry; Semialgebraic sets and related spaces.*

Received on the 3rd of January, 1995.