Improved Lower Bound on Testing Membership to a Polyhedron by Algebraic Decision Trees

Dima Grigoriev * Marek Karpinski† Nicolai Vorobjov‡

Abstract

We introduce a new method of proving lower bounds on the depth of algebraic decision trees of degree d and apply it to prove a lower bound $\Omega(\log N)$ for testing membership to an n-dimensional convex polyhedron having N faces of all dimensions, provided that $N > (nd)^{\Omega(n)}$. This weakens considerably the restriction on N previously imposed by the authors in [GKV 94] and opens a possibility to apply the bound to some naturally appearing polyhedra.

Introduction

We study the problem of deciding membership to a convex polyhedron. The problem of testing membership to a semialgebraic set Σ was considered by many authors (see, e.g., [B 83], [B 92], [BKL 92], [BL 92], [BLY 92], [MH 85], [GKV 94], [Y 92], [GK 93], [GK 94], [Y 93], [YR 80] and the references there). We consider a problem of testing membership to a convex polyhedron P in n-dimensional space $\mathbf{R}^n$. Let P have N faces of all the dimensions. In [MH 85] it was shown, in particular, that for this problem $O(\log N)n^{O(1)}$ upper bound is valid for the depth of linear decision trees, in [YR 80] a lower bound $\Omega(\log N)$ was obtained. A similar question was open for algebraic decision trees. In [GKV 94] we have proved a lower bound $\Omega(\log N)$ for the depth of algebraic decision trees testing membership to P, provided that $N > (dn)^{\Omega(n^2)}$. In the present paper we weaken the latter assumption to $N \geq (dn)^{\Omega(n)}$. In this new form the bound looks plausible to be applicable to polyhedra given by $2^{O(n)}$ linear constraints (like in "knapsack" problem), thus having $2^{O(n^2)}$ faces. In the present note we apply the obtained lower bound to a concrete class of polyhedra given by $\Omega(n^2)$ linear constraints and with $n^{\Omega(n)}$ faces.

In [GV 94] the lower bound $\Omega(\sqrt{\log N})$ was proved for the Pfaffian computation tree model. This model uses at gates *Pfaffian functions*, the latter include all major elementary transcendental and algebraic functions.

Several topological methods were introduced for obtaining lower bounds for the complexity of testing membership to Σ by linear decision trees, algebraic decision trees, algebraic computation trees (see the definitions, e.g., [B 83]).

In [B 83] a lower bound $\Omega(\log C)$ was proved for the most powerful among the considered in this area computational models, namely algebraic computation trees, where C is the number of connected components of Σ or of the complement of Σ. Later, in [BLY 92], a lower bound $\Omega(\log \chi)$ for linear decision trees was proved, where χ is Euler characteristic of Σ, in [Y 92] this lower bound was extended to algebraic computation trees. A stronger lower bound $\Omega(\log B)$ was proved later in [BL 92], [B 92] for linear decision trees, where B is the sum of Betti numbers of Σ (obviously, $C, \chi \leq B$). In the recent paper [Y 94] the latter lower bound was extended to the algebraic decision trees.

All the mentioned topological tools fail when Σ is a convex polyhedron, because $B = 1$ in this situation. The same is true for the method developed in [BLY 92] for linear decision trees, based on the minimal number of convex polyhedra onto which Σ can be partitioned.

To handle the case of a convex polyhedron, we introduce in Sections 1, 3 another approach which differs drastically from [GKV 94]. Let W be a semialgebraic set accepted by a branch of an algebraic decision tree. In Section 3 we make an "infinitesimal perturbation" of W which transforms this set into a smooth hypersurface. Then we describe the semialgebraic subset of all the points of the hypersurface in which all its principal curvatures are "infinitely large" (the set $\mathcal{K}_0$ in Sec-

*Dept. of Compter Science and Mathematics, Penn State University, University Park, PA 16802. Supported in part by Volkwagen-Stiftung. e-mail: dima@cse.psu.edu

†Dept. of Computer Science, University of Bonn, 53117 Bonn, and International Computer Science Institute, Berkeley, CA. Supported in part by DFG Grant KA673/4-1, by the ESPRIT BR Grant 7097 and by the Volkwagen-Stiftung. email: marek@cs.bonn.edu

‡Dept. of Compter Science and Mathematics, Penn State University, University Park, PA 16802. email: vorobjov@cse.psu.edu

tion 3). We also construct a more general set $\mathcal{K}_i$ (for each $0 \le i \le n-1$) of the points with infinitely large curvatures in the shifts of a fixed $(n-i)$-dimensional plane. Section 1 provides a short system of inequalities for determining $\mathcal{K}_i$. It is done by developing an explicit symbolic calculis for principal curvatures.

In Section 2 we introduce some necessary notions concerning infinitesimals and in Section 3 apply them to define the "standard part" $K_i = \mathrm{st}(\mathcal{K}_i) \subset \mathbf{R}^n$. We show (Corollary to Lemma 3) that to obtain the required bound for the number of i-faces P_i of P such that $\dim(P_i \cap W) = i$ it is sufficient to estimate the number of faces P_i with $\dim(P_i \cap K_i) = i$. In Section 4 we reduce the latter bound to an estimate of the number of local maxima of a generic linear function L on $\mathcal{K}_i$ with the help of a Whitney stratification of K_i. To estimate these local maxima we introduce in Section 5 another infinitesimal perturbation of $\mathcal{K}_i$ and obtain a new smooth hypersurface. At this point a difficulty arises due to the fact that $\mathcal{K}_i$ (and therefore, the related smooth hypersurface) are defined by systems of inequalities involving algebraic functions, rather than polynomials, because in the expressions for curvatures (in Section 1) square roots of polynomials appear. We represent the set of local maxima of L on the smooth hypersurface by a formula of the first-order theory of real closed fields with merely existential quantifiers and quantifier-free part Φ. We estimate in Section 5 (invoking [Mi 64] in a usual way) the number of the connected components of the semialgebraic set defined by Φ.

In Section 6 we describe a particular class of polyhedra (dual to cyclic polyhedra [MS 71]) having large numbers of facets, for which Theorem 1 provides a nontrivial lower bound.

Now let us formulate precisely the main result. We consider algebraic decision trees of a fixed degree d (see, e.g., [B 83], [Y 93]). Suppose that such a tree T, of the depth k, tests a membership to a convex polyhedron $P \subset \mathbf{R}^n$. Denote by N the number of faces of P of all dimensions from zero to $n-1$. In this paper we agree that a face is "open", i.e., does not contain faces of smaller dimensions.

Theorem 1.
$$k \ge \Omega(\log N),$$
provided that $N \ge (dn)^{cn}$ for a suitable $c > 0$.

Let us fix a branch of T which returns "yes". Denote by $f_i \in \mathbf{R}[X_1, \dots, X_n]$, $1 \le i \le k$ the polynomials of degrees $\deg(f_i) \le d$, attached to the vertices of T along the fixed branch. Without loss of generality, we can assume that the corresponding signs of polynomials along the branch are

$$f_1 = \dots = f_{k_1} = 0,\ f_{k_1+1} > 0, \dots, f_k > 0.$$

Then the (accepted) semialgebraic set

$$W = \{f_1 = \dots = f_{k_1} = 0,\ f_{k_1+1} > 0, \dots, f_k > 0\}$$

lies in P.

Our main technical tool is the following theorem.

Theorem 2. *The number of faces P' of P such that $\dim(P') = \dim(P' \cap W)$ is bounded from above by $(knd)^{O(n)}$.*

Let us deduce Theorem 1 from Theorem 2.

For each face P' of P there exists at least one branch of the tree T with the output "yes" and having an accepted set $W_1 \subset \mathbf{R}^n$ such that

$$\dim(W_1 \cap P') = \dim(P').$$

Since there are at most 3^k different branches of T, the inequality

$$N < 3^k (knd)^{O(n)}$$

follows from Theorem 2. This inequality and the assumption $N > (dn)^{cn}$ (for a suitable c) imply $k \ge \Omega(\log N)$, which proves Theorem 1.

Note that in the case $k_1 = 0$ for an open set W and each face P' of P we have $P' \cap W = \emptyset$. Thus in what follows we can suppose that $k_1 \ge 1$.

1 Computer algebra for curvatures

Let a polynomial $F \in \mathbf{R}[X_1, \dots, X_n]$ with $\deg(F) < d$. Assume that at a point $x \in \{F = 0\} \subset \mathbf{R}^n$ the gradient $\mathrm{grad}_x(F) = \left(\frac{\partial F}{\partial X_1}, \dots, \frac{\partial F}{\partial X_n}\right)(x) \neq 0$. Then, according to the implicit function theorem, the real algebraic variety $\{F = 0\} \subset \mathbf{R}^n$ is a smooth hypersurface in a neighborhood of x.

Fix a point $x \in \{F = 0\}$. Consider a linear transformation $X \longrightarrow A_x X + x$, where A_x is an arbitrary orthogonal matrix such that

$$u_1 = A_x e_1 + x = \frac{\mathrm{grad}_x(F)}{||\mathrm{grad}_x(F)||}$$

is the normalized gradient and $e_1, \dots, e_n$ is the coordinate basis at the origin. Then the linear hull of vectors $u_j = Ae_j + x$, $2 \le j \le n$ is the tangent space T_x to $\{F = 0\}$ at x. Denote by $U_1, \dots, U_n$ the coordinate variables in the basis $u_1, \dots, u_n$. By the implicit function theorem, there exists a smooth function

$H_x(U_2,\dots,U_n)$ defined in a neighborhood of x on T_x such that $\{F=0\}=\{U_1=H_x(U_2,\dots,U_n)\}$ in this neighborhood.

Let $\mathrm{grad}_x(F)=(\tilde{\alpha}_1,\dots,\tilde{\alpha}_n)$ with $\tilde{\alpha}_{i_0}\neq 0$. Take any permutation π_{i_0} of $\{1,\dots,n\}$ such that $\pi_{i_0}(1)=i_0$. Denote $(\alpha_1,\dots,\alpha_n)=(\tilde{\alpha}_{\pi_{i_0}(1)},\dots,\tilde{\alpha}_{\pi_{i_0}(n)})$ (thus $\alpha_1\neq 0$) and $\beta_i=\sqrt{\alpha_1^2+\cdots+\alpha_i^2}$, $1\le i\le n$. Obviously $\beta_i>0$ and $\beta_n=\|\mathrm{grad}_x(F)\|$.

As A_x one can take the following product of $(n-1)$ orthogonal matrices:

$$\prod_{0\le k\le n-2}\begin{pmatrix} \frac{\beta_{n-k-1}}{\beta_{n-k}} & 0 & \cdots & 0 & \frac{\alpha_{n-k}}{\beta_{n-k}} & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0\\ -\frac{\alpha_{n-k}}{\beta_{n-k}} & 0 & \cdots & 0 & \frac{\beta_{n-k-1}}{\beta_{n-k}} & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

(in kth matrix of this product the element $\frac{\beta_{n-k-1}}{\beta_{n-k}}$ occurs at the positions $(1,1)$ and $(n-k,n-k)$).

Denote $F_x(U_1,\dots,U_n)=F(A_x^T(U_1,\dots,U_n)+x)$. Differentiating this function twice and taking into the account that $F_x(H_x(U_2,\dots,U_n),U_2,\dots,U_n)=0$ in a neighborhood of x in T_x we get

$$\frac{\partial^2 F_x}{\partial U_1\partial U_j}\frac{\partial H_x}{\partial U_i}+\frac{\partial F_x}{\partial U_1}\frac{\partial^2 H_x}{\partial U_i\partial U_j}+\frac{\partial^2 F_x}{\partial U_i\partial U_j}=0 \qquad (1)$$

for $2\le i,\ j\le n$.

Since

$$\frac{\partial H_x}{\partial U_i}\Big|_{(U_2,\dots,U_n)=0}=0 \quad \text{and}$$

$$\frac{\partial F_x}{\partial U_1}\Big|_{(U_1,\dots,U_n)=0}=\|\mathrm{grad}_x(F)\|\neq 0,$$

evaluating the equality (1) at x (i.e., substituting $(U_1,\dots,U_n)=0$) we obtain (cf. [Mi 64]):

$$\Big(\frac{\partial^2 H_x}{\partial U_i\partial U_j}\Big)\Big|_{(U_2,\dots,U_n)=0}=$$

$$(\|\mathrm{grad}_x(F)\|)^{-1}\Big(\frac{\partial^2 F_x}{\partial U_i\partial U_j}\Big)\Big|_{(U_1,\dots,U_n)=0}. \qquad (2)$$

Introduce the symmetric $(n-1)\times(n-1)$-matrix

$$\mathcal{H}_x=\Big(\frac{\partial^2 H_x}{\partial U_i\partial U_j}\Big)\Big|_{(U_2,\dots,U_n)=0}.$$

Its eigenvalues $\lambda_2,\dots,\lambda_n$ belong to $\mathbf{R}$ and are called the *principal curvatures* of the hypersurface $\{F=0\}$ at x [Th 77].

Now we describe symbolically the set of all points x with all principal curvatures greater than some parameter κ.

Denote by $\chi(Z)$ the characteristic polynomial of the matrix $\mathcal{H}_x$. The roots of χ are exactly $\lambda_2,\dots,\lambda_n$. Due to Sturm theorem, every $\lambda_2,\dots,\lambda_n$ is greater than κ if and only if $\chi_l(\kappa)\chi_{l+1}(\kappa)<0$, $0\le l\le n-2$, where $\chi_0=\chi$, $\chi_1=\chi_0'$ and $\chi_2,\dots,\chi_{n-1}$ is the polynomial remainder sequence of χ_0, χ_1 [Lo 82]. Obviously $\deg_Z(\chi_l)=n-l-1$.

Observe that every element of the matrix A_x can be represented as a fraction γ_1/γ_2 where $\gamma_2=\beta_1^{\nu_1}\cdots\beta_n^{\nu_n}$, $\nu_1\ge 0,\dots,\nu_n\ge 0$ are integers and

$$\gamma_1=\Gamma(\beta_1,\dots,\beta_{n-1},X_1,\dots,X_n)$$

is a polynomial in

$$\beta_1(X_1,\dots,X_n),\dots,\beta_{n-1}(X_1,\dots,X_n),X_1,\dots,X_n$$

with $\Gamma\in\mathbf{R}[Z_1,\dots,Z_{n-1},X_1,\dots,X_n]$. Moreover, $\nu_1+\cdots+\nu_n\le 2(n-1)$ and $\deg(\Gamma)\le d(n-1)$. Hence all elements of A_x are algebraic functions in $X_1,\dots,X_n$ of quadratic-irrational type. By the degree of such quadratic-irrational function we mean

$$\max\{\deg(\Gamma),\ \nu_1+\cdots+\nu_n\}.$$

Since an inequality for fraction one could rewrite as a system of inequalities for its numerator and denominator, in what follows we deal with more special algebraic functions in $X_1,\dots,X_n$, namely of the type γ_1.

Formula (2) and Habicht's theorem [Lo 82] imply that $\deg(\chi_l)\le(nd)^{O(1)}$.

We summarize a description of the set of all points with large principal curvatures in the following lemma.

Lemma 1. *Fix $1\le i_0\le n$. The set of all points $x\in\{F=0\}$ such that $\mathrm{grad}_x(F)=(\hat{\alpha}_1,\dots,\hat{\alpha}_n)$ has $\hat{\alpha}_{i_0}\neq 0$ and all principal curvatures of the hypersurface $\{F=0\}$ at x are greater than κ can be represented as $\{F=0,\ g_1>0,\dots,g_n>0\}$. Here $g_1=\hat{\alpha}_{i_0}^2,g_2\dots,g_n$ are polynomials in κ of degrees at most $2n$ with coefficients being quadratic-irrational algebraic functions (see above) of degrees less than $(nd)^{O(1)}$.*

Remark. Observe that a set given by a system of inequalities involving real algebraic functions is semialgebraic. Hence the set introduced in Lemma 1 is semialgebraic.

2 Calculis with infinitesimals

The definitions below concerning infinitesimals follow [GV 88].

Let $\mathbf{F}$ be an arbitrary real closed field (see, e.g., [L 65]) and an element ε be infinitesimal relative to elements of $\mathbf{F}$. The latter means that for any positive element $a \in \mathbf{F}$ inequalities $0 < \varepsilon < a$ are valid in the ordered field $\mathbf{F}(\varepsilon)$. Obviously, the element ε is transcendental over $\mathbf{F}$. For an ordered field $\mathbf{F}'$ we denote by $\tilde{\mathbf{F}}'$ its (unique up to isomorphism) real closure, preserving the order on $\mathbf{F}'$ [L 65].

Let us remind some other well-known statements concerning real closed fields. A Puiseux (formal power-fractional) series over $\mathbf{F}$ is series of the kind

$$b = \sum_{i \geq 0} a_i \varepsilon^{\nu_i/\mu},$$

where $0 \neq a_i \in \mathbf{F}$ for all $i \geq 0$, integers $\nu_0 < \nu_1 < \ldots$ increase and the natural number $\mu \geq 1$. The field $\mathbf{F}((\varepsilon^{1/\infty}))$ consisting of all Puiseux series (appended by zero) is real closed, hence $\mathbf{F}((\varepsilon^{1/\infty})) \supset \widetilde{\mathbf{F}(\varepsilon)} \supset \mathbf{F}(\varepsilon)$. Besides the field $\mathbf{F}[\sqrt{-1}]((\varepsilon^{1/\infty}))$ is algebraically closed.

If $\nu_0 < 0$, then the element $b \in \mathbf{F}((\varepsilon^{1/\infty}))$ is infinitely large. If $\nu_0 > 0$, then b is infinitesimal relative to elements of the field $\mathbf{F}$. A vector $(b_1, \ldots, b_n) \in \left(\mathbf{F}((\varepsilon^{1/\infty}))\right)^n$ is called $\mathbf{F}$-finite if each coordinate b_i, $1 \leq i \leq n$ is not infinitely large relative to elements of $\mathbf{F}$.

For any $\mathbf{F}$-finite element $b \in \mathbf{F}((\varepsilon^{1/\infty}))$ its *standard part* $\mathrm{st}(b)$ is definable, namely $\mathrm{st}(b) = a_0$ in the case $\nu_0 = 0$ and $\mathrm{st}(b) = 0$ if $\nu_0 > 0$. For any $\mathbf{F}$-finite vector $(b_1, \ldots, b_n) \in \left(\mathbf{F}((\varepsilon^{1/\infty}))\right)^n$ its standard part is defined by the equality

$$\mathrm{st}(b_1, \ldots, b_n) = (\mathrm{st}(b_1), \ldots, \mathrm{st}(b_n)).$$

For a set $\mathcal{W} \subset \left(\mathbf{F}((\varepsilon^{1/\infty}))\right)^n$ consisting of only $\mathbf{F}$-finite vectors we define

$$\mathrm{st}(\mathcal{W}) = \{\mathrm{st}(w) : \ w \in \mathcal{W} \text{ and } w \text{ is } \mathbf{F}\text{−finite}\}.$$

The following "transfer principle" is true [T 51]. If $\mathbf{F}'$, $\mathbf{F}''$ are real closed fields with $\mathbf{F}' \subset \mathbf{F}''$ and Φ is a closed (without free variables) formula of the first order theory of the field $\mathbf{F}'$, then Φ is true over $\mathbf{F}'$ if and only if $\mathcal{P}$ is true over $\mathbf{F}''$.

In the sequel we consider infinitesimals $\varepsilon_1, \varepsilon_2, \ldots$ such that ε_{i+1} is infinitesimal relative to the real closure $\mathbf{R}_i$ of the field $\mathbf{R}(\varepsilon_1, \ldots, \varepsilon_i)$ for each $i \geq 0$. We assume that $\mathbf{R}_0 = \mathbf{R}$.

For an $\mathbf{R}_i$-finite element $b \in \mathbf{R}_{i+1}$ its standard part (relative to $\mathbf{R}_i$) denote by $\mathrm{st}_i(b) \in \mathbf{R}_i$. For any $b \in \mathbf{R}_j$, $j > i$ we define $\mathrm{st}_i(b) = \mathrm{st}_i(\mathrm{st}_{i+1}(\ldots \mathrm{st}_{j-1}(b) \ldots))$. For a semialgebraic set $V \subset \mathbf{F}_1^n$ defined by a certain formula Φ of the first order theory of the field $\mathbf{F}_1$ and for a real closed $\mathbf{F}_2 \supset \mathbf{F}_1$ we define the completion $V^{(\mathbf{F}_2)} \subset \mathbf{F}_2^n$ of V as the semialgebraic set given in $\mathbf{F}_2^n$ by the same formula Φ (we say that $V^{(\mathbf{F}_2)}$ is *defined over* $\mathbf{F}_1$). In a similar way one can define completions of polynomials and algebraic functions.

Note that one can apply the transfer principle also to a formula containing quadratic irrational functions since any such formula can be replaced by an equivalent formula of first-order theory. This can be done with replacing each occurrence of a square root $\sqrt{\varphi}$ by new variable Z, adding the quantifier prefix $\exists Z$ and inequalities $Z > 0$, $Z^2 = \varphi$.

Lemma 2 (cf. Lemma 4a) in [GV 88]). *Let F be a smooth algebraic function defined on an open semialgebraic set $U \subset \mathbf{R}_i^n$ and determined by a polynomial with coefficients from $\mathbf{R}_i$. Then ε_{i+1} is not a critical value of F (i.e., $\mathrm{grad}_y(F)$ does not vanish at any point $y \in \{F = \varepsilon_{i+1}\} \cap U^{(\mathbf{R}_{i+1})}$).*

To prove Lemma 2 note that Sard's theorem [Hi 76] and the transfer principle imply the finiteness of the set of all critical values of F in $U^{(\mathbf{R}_{i+1})}$, moreover this set lies in $\mathbf{R}_i$.

3 Curved points

In what follows we assume w.l.o.g. that polyhedron P is compact, a reduction of a general case to this one is described in Section 2 of [GKV 94].

For an m-plane $Q \subset \mathbf{R}_j^n$ and a point $x \in \mathbf{R}_j^n$ denote by $Q(x)$ the m-plane collinear to Q and containing x.

Two planes Q_1, Q_2 of arbitrary dimensions are called *transversal* if

$$\dim\bigl(Q_1(0) \cap Q_2(0)\bigr) =$$

$$\max\{0, \ \dim\bigl(Q_1(0)\bigr) + \dim\bigl(Q_2(0)\bigr) - n\}.$$

For every $0 \leq i < n$ choose an $(n-i)$-plane Π_{n-i} (defined over R) transversal to any facet of the polyhedron P.

Denote $f = f_1^2 + \cdots + f_{k_1}^2$.

Fix $0 \leq i < n$ and denote by $f^{(x)}$ the restriction of f on $\Pi_{n-i}(x)$ (for $x \in \mathbf{R}_j^n$).

Definition. *A point $y \in \{f = \varepsilon_3\}$ is called i-curved if $\mathrm{grad}_y(f^{(y)} - \varepsilon_3) \neq 0$, all principal curvatures of the* *variety* $\{f^{(y)} = \varepsilon_3\} \subset \Pi_{n-i}(y)$ *at* y *are greater than* ε_2^{-1} *and* $f_{k_1+1}(y) > \varepsilon_2, \ldots, f_k(y) > \varepsilon_2$.

Remark. We fix an orthogonal basis in $\Pi_{n-i}(0)$ with coordinates belonging to $\mathbf{R}$. Then in Definition we consider curvatures in $\Pi_{n-i}(y)$ with respect to the basis obtained from the fixed one by the shift $Y \longrightarrow Y + y$.

One can consider this definition as a kind of "localization" of the key concept of an *angle point* from [GV 94].

Denote the set of all i-curved points by $\mathcal{K}_i \subset \mathbf{R}_3^n$. Observe that $\mathcal{K}_i$ is semialgebraic due to the remark at the end of Section 1. Denote $K_i = \mathrm{st}_0(\mathcal{K}_i) \subset \mathbf{R}^n$, this set is also semialgebraic by Lemma 5.1 from [RV 94].

Lemma 3. *Let for an* i*-facet* P_i *of* P *the dimension* $\dim(W \cap P_i) = i$*. Then* $W \cap P_i \subset K_i$.

Corollary. *If* $\dim(W \cap P_i) = i$ *then* $\dim(K_i \cap P_i) = i$.

This Corollary implies that in order to prove Theorem 2 it is sufficient to bound the number of i-facets P_i for which $\dim(K_i \cap P_i) = i$.

Lemma 4. *For any smooth point* $z \in K_i$ *with the dimension* $\dim_z(K_i) \geq i+1$ *the tangent plane* T_z *to* K_i *at* z *is not transversal to* Π_{n-i}.

Remark. In the particular case $i = 0$ Lemma 4 states that K_0 consists of a finite number of points.

4 Faces of P and Whitney stratification of K_i

Denote by $B_x(r)$ the open ball in $\mathbf{R}_i^n$ centered at x and of the radius r.

For a subset $E \subset \mathbf{R}_i^n$ denote by $cl(E)$ its closure in the topology with the base of all open balls. Denote by ∂E the boundary

$$\{y \in \mathbf{R}_i^n : \text{for any } 0 < r \in \mathbf{R}_i\ \emptyset \neq B_y(r) \cap E \neq B_y(r)\}.$$

Recall that K_i, as any semialgebraic set, admits a Whitney stratification (see, e.g., [GM 88]). Namely, K_i can be represented as a disjoint union $K_i = \bigcup_j S_j$ of a finite number of semialgebraic sets, called strata, which are smooth manifolds and such that:

(1) (frontier condition) $S_{j_1} \cap cl(S_{j_2}) \neq \emptyset$ if and only if $S_{j_1} \subset cl(S_{j_2})$;

(2) (Whitney condition A) Let $S_{j_1} \subset cl(S_{j_2})$ and a sequence of points $x_l \in S_{j_2}$ tends to a point $y \in S_{j_1}$ when $l \longrightarrow \infty$. Assume that the sequence of tangent planes T_{x_l} to S_{j_2} at points x_l tends to a certain plane T. Then $T_y \subset T$ where T_y is a tangent plane to S_{j_1} at y.

Lemma 5. *Let for an* i*-face* P_i *of* P *the dimension* $\dim(K_i \cap P_i) = i$*. Assume that* S'_j *is a connected component of a stratum* S_j *of* K_i *such that* $\dim\big(cl(S'_j) \cap K_i \cap P_i\big) = i$*. Then* $S'_j \subset P_i$.

Denote $g = f_{k_1+1} \cdots f_k$. Choose $0 < \mu \in \mathbf{R}$ satisfying the following properties:

(a) μ is less than the absolute values of all critical values of the restrictions of g on i-faces P_i (note that Sard's theorem implies the finiteness of the number of all critical values, moreover they all belong to $\mathbf{R}$);

(b) for any P_i such that $\dim(K_i \cap P_i) = i$ the dimension

$$\dim\big(\{g = \mu\} \cap cl(S_j) \cap K_i \cap P_i\big) \leq i - 2$$

for every connected component S'_j of a stratum S_j such that S'_j is not contained in P_i (observe that due to Lemma 5 there exists at most finite number of μ violating this condition).

For any i-face P_i denote by $\overline{P}_i$ the i-plane containing P_i. Denote $K'_i = K_i \cap \{g = \mu\}$.

From the properties (a), (b) using Lemma 3 we deduce the following lemma.

Lemma 6. *Let for an* i*-face* P_i *of* P *the dimension* $\dim(W \cap P_i) = i$*. The following equality of the varieties holds:*

$$K'_i \cap P_i = \{g = \mu\} \cap \{f_{k_1+1} > 0, \ldots, f_k > 0\} \cap P_i,$$

and, moreover, this variety is a nonempty smooth compact hypersurface in $\overline{P}_i$*. Besides,*

$$\dim\big((cl(K'_i \setminus P_i)) \cap (K'_i \cap P_i)\big) \leq i - 2.$$

The next important step is the proof of the following lemma.

Lemma 7. *The number of* i*-faces* P_i *such that* $K'_i \cap P_i$ *is a nonempty smooth hypersurface in* $\overline{P}_i$ *and*

$$\dim\big((cl(K'_i \setminus P_i)) \cap (K'_i \cap P_i)\big) \leq i - 2,$$

does not exceed $(nkd)^{O(n)}$.

Theorem 2 immediately follows from Lemmas 6 and 7. A sketch of a proof of Lemma 7 is given in the next section.

Lemma 8.

$$K_i' = \mathrm{st}_0(\mathcal{K}_i \cap \{|g-\mu| \leq \varepsilon_1\}).$$

5 Extremal points of a linear function on K_i'

Take a generic linear function $L = \gamma_1 X_1 + \cdots + \gamma_n X_n$ with coefficients $\gamma_1, \ldots, \gamma_n \in \mathbf{R}$. Fix P_i satisfying the conditions of Lemma 8 and denote by $L^{(P_i)}$ the restriction of L on $\overline{P}_i$. Then $L^{(P_i)}$ attains its maximal value, say, $\theta_0^{(P_i)}$ on the compact set $K_i' \cap P_i$ at a certain point v. Denote by V a connected component of $K_i' \cap P_i$ which contains v. There exists $0 < r \in \mathbf{R}$ such that $B_v(r) \cap K_i' = B_v(r) \cap V$ due to the property (b) (see Section 4). Moreover, there exists $0 < \zeta^{(P_i)} \in \mathbf{R}$ such that the values of L on the set $K_i' \cap \partial B_v(r/2)$ are less than $\theta_0 - \zeta^{(P_i)}$. This implies, using Lemma 8, the following lemma.

Lemma 9. *The linear form L attains its maximal value $\theta^{(P_i)}$ on the set*

$$cl\big(\mathcal{K}_i \cap \{|g-\mu| \leq \varepsilon_1\}\big) \cap B_v(r/2)$$

(at a point, say, w) and the values of L on the set

$$cl\big(\mathcal{K}_i \cap \{|g-\mu| \leq \varepsilon_1\}\big) \cap \partial B_v(r/2)$$

are less than $\mathrm{st}_0(\theta^{(P_i)} - \zeta^{(P_i)})$. *Moreover,* $\mathrm{st}_0(\theta^{(P_i)}) = \theta_0^{(P_i)}$ *and* $\mathrm{st}_0(w) = v \in P_i$.

For a point y let

$$\mathrm{grad}_y(f^{(y)} - \varepsilon_3) = (u_1, \ldots, u_{n-i})$$

(cf. Definition). The set $\mathcal{K}_i \cap \{|g-\mu| \leq \varepsilon_1\}$ of the points $y = (y_1, \ldots, y_n)$ can be represented as a union of $n-i$ semialgebraic sets of the form

$$U^{(i_0)} = \{f - \varepsilon_3 = 0, u_{i_0}^2 > 0,$$

$$p_1 > 0, \ldots, p_s > 0\} \subset \mathbf{R}_3^n, \quad 1 \leq i_0 \leq n-i$$

for some algebraic functions $p_1, \ldots, p_s$ of the quadratic-irrational type introduced in Section 1, i.e., polynomials (with coefficients from $\mathbf{R}_2$) in $y_1, \ldots, y_n$ and in

$$\sqrt{u_{i_0}^2},\ \sqrt{u_{i_0}^2 + u_{\pi_{i_0}(2)}^2},$$
$$\ldots, \sqrt{u_{i_0}^2 + u_{\pi_{i_0}(2)}^2 + \ldots + u_{\pi_{i_0}(n-i)}^2} \qquad (3)$$

(see Lemma 1). Here π_{i_0} is a permutation of $\{1, 2, \ldots, n-i\}$ such that $\pi_{i_0}(1) = i_0$ (cf. Section 1).

Denote

$$q = \big(\varepsilon_5^2 - (f - \varepsilon_3)^2\big)(u_{i_0}^2 - \varepsilon_4)(p_1 - \varepsilon_4) \cdots (p_s - \varepsilon_4).$$

Introduce the semialgebraic set

$$\mathcal{U}_0^{(i_0)} = \{\varepsilon_5^2 > (f - \varepsilon_3)^2,\ u_{i_0}^2 > \varepsilon_4,$$
$$p_1 > \varepsilon_4, \ldots, p_s > \varepsilon_4\} \subset \mathbf{R}_5^n$$

and

$$\mathcal{U}^{(i_0)} = \{q = \varepsilon_6\} \cap (\mathcal{U}_0^{(i_0)})^{(\mathbf{R}_6)} \subset \mathbf{R}_6.$$

The next lemma follows from Lemmas 1, 4 in [GV 92].

Lemma 10.

$$\mathrm{st}_3(\mathcal{U}^{(i_0)}) = cl(U^{(i_0)}).$$

Lemma 11. *For a certain $1 \leq i_0 \leq n-i$ the linear form L attains its maximal value $\theta_1^{(P_i)}$ on the set $\mathcal{U}^{(i_0)} \cap B_v(r/2)$ at a certain point w_1, and the values of L on the set $\mathcal{U}^{(i_0)} \cap \partial B_v(r/2)$ are less than* $\mathrm{st}_0(\theta_1^{(P_i)}) - \zeta^{(P_i)}$. *Moreover,* $\mathrm{st}_3(\theta_1^{(P_i)}) = \theta^{(P_i)}$ *and* $\mathrm{st}_0(w_1) = v \in P_i$.

Lemma 11 follows from Lemmas 9, 10. For a proof, take $1 \leq i_0 \leq n-i$ such that the corresponding point w (see Lemma 9) lies in $cl(U^{(i_0)})$.

Corollary *The number of i-faces P_i satisfying the conditions of Lemma 7 does not exceed the number of local maxima of L on the set*

$$\bigcup_{1 \leq i_0 \leq n-i} cl(\mathcal{U}^{(i_0)}).$$

Observe that in the open semialgebraic set $\{u_{i_0}^2 > 0\}$ all the square roots (3) are positive. Therefore all algebraic functions $p_1, \ldots, p_s$ occurring in $\mathcal{U}_0^{(i_0)}$ are smooth, hence q is smooth as well. Because of Lemma 2 ε_6 is not a critical value of q in the set

$\{u_{i_0}^2 > 0\}$. Then the implicit function theorem implies the following lemma.

Lemma 12. *$\mathcal{U}^{(i_0)}$ is a smooth hypersurface, namely for each point $x \in \mathcal{U}^{(i_0)}$ there is a neighborhood of x in which $\mathcal{U}^{(i_0)}$ is defined by the equation $q = \varepsilon_6$ and the gradient $\mathrm{grad}_x(q - \varepsilon_6)$ does not vanish.*

Finally, let us prove the following lemma.

Lemma 13. *The number ν of local maxima of L on $\mathcal{U}^{(i_0)}$ does not exceed $(nkd)^{O(n)}$.*

Together with Corollary to Lemma 11 this implies Lemma 7 (and hence Theorem 2).

Because of Lemma 12, ν does not exceed the number of connected components of the semialgebraic set

$$M = \{0 = q - \varepsilon_6 = \gamma_i \frac{\partial q}{\partial X_j} - \gamma_j \frac{\partial q}{\partial X_i},$$

$$1 \le i < j \le n\} \subset \mathbf{R}_6^n.$$

Replace each occurrence of the square root

$$\sqrt{u_{i_0}^2 + u_{\pi_{i_0}(2)}^2 + \cdots + u_{\pi_{i_0}(m)}^2},$$

$1 \le m \le n - i$ in q by a new variable Z_m. Denote the resulting polynomial by $Q \in \mathbf{R}_5[X_1, \ldots, X_n, Z_1, \ldots, Z_m]$ (cf. Section 1).

Introduce the semialgebraic set

$$\mathcal{M} = \{0 = Q - \varepsilon_6 = \gamma_i \frac{\partial Q}{\partial X_j} - \gamma_j \frac{\partial Q}{\partial X_i},\ 1 \le i < j \le n,$$

$$Z_m > 0,\ Z_m^2 = u_{i_0}^2 + u_{\pi_{i_0}(2)}^2 + \cdots + u_{\pi_{i_0}(m)}^2,$$

$$1 \le m \le n - i\} \subset \mathbf{R}_6^{2n-i}.$$

Consider the linear projection

$$\rho : \mathbf{R}_6^{2n-i} \longrightarrow \mathbf{R}_6^n,$$

$$\rho(X_1, \ldots, X_n, Z_1, \ldots, Z_m) = (X_1, \ldots, X_n).$$

Then $\rho(\mathcal{M}) = M$. Hence the number of connected components of M is less than or equal to the number of connected components of $\mathcal{M}$.

Observe that the degrees of rational functions occurring in $\mathcal{M}$ can be bounded from above by $(knd)^{O(1)}$ due to Lemma 1. Therefore, the number of connected components of $\mathcal{M}$ does not exceed $(knd)^{O(n)}$ by [Mi 64].

This completes the proof of Lemma 13 and thereby Theorems 2 and 1.

6 Lower bounds for concrete polyhedra

In this section we give an application of the lower bound from Theorem 1 to a concrete class of polyhedra. We follow the construction of cyclic polyhedra (see [MS 71]), used in the analysis of the simplex method.

Take any $m > \Omega(n^2)$ points in $\mathbf{R}^n$ of the form $(t_j, t_j^2, \ldots, t_j^n)$ for pairwise distinct t_j, $1 \le j \le m$. Consider the convex hull of these points and denote by $P_{n,m} \subset \mathbf{R}^n$ its dual polyhedron [MS 71]. Then $P_{n,m}$ has m faces of the highest dimension $n-1$ and the number of faces of all dimensions

$$N > \binom{m - \lfloor n/2 \rfloor}{\lfloor n/2 \rfloor} > m^{\Omega(n)}$$

(see [MS 71]).

Therefore, Theorem 1 implies that the complexity of testing membership to $P_{n,m}$ is bounded by $\Omega(\log N) > \Omega(n \log m)$.

We would like to mention that Section 4 of [GKV 94] provides a weaker bound $\Omega(\log m)$ even for algebraic *computation* trees.

Acknowledgments

We thank Dima Burago, Friedhelm Meyer auf der Heide, Kolia Ivanov, Andy Yao for a number of useful discussions and Anders Björner for a help with the example from Section 6.

References

[B 83] M. Ben-Or, *"Lower bounds for algebraic computation trees, "* Proceedings of ACM Symposium on Theory of Computing, 1983, 80-86.

[B 92] A. Björner, *"Subspace arrangements,"* Proceedings of 1st European Congress of Mathematicians, Paris, 1992.

[BKL 92] P. Buergisser, M. Karpinski, and T. Lickteig *"On randomized algebraic test complexity,"* Technical Report TR-92-070, International Computer Science Institute, Berkeley, 1992; in: Journal of Complexity (1993), pp. 231-251.

[BL 92] A. Björner, and L. Lovasz *"Linear decision trees, subspace arrangements and Möbius functions,"* Preprint, 1992.

[BLY 92] A. Björner, L. Lovasz, and A. Yao *"Linear decision trees: Volume estimates and topological bounds,"* Proceedings of ACM Symposium on Theory of Computing, 1992, 170-177.

[G 88] D. Grigoriev *"Complexity of deciding Tarski algebra,"* Journal Symbolic Comp., V. 5, 1988, 65-108.

[GK 93] D. Grigoriev and M. Karpinski *"Lower Bounds on Testing Membership to a Polygon for Algebraic and Randomized Computation Trees"* Technical Report TR-93-042, International Computer Science Institute, Berkeley, 1993.

[GK 94] D. Grigoriev and M. Karpinski *"Lower Bound for Randomized Linear Decision Trees Recognizing a Union of Hyperplanes in a Generic Position"* Research Report No. 85144-CS, University of Bonn, 1994.

[GKS 93] D. Grigoriev, M. Karpinski, and M. Singer *"Computational complexity of sparse real algebraic function interpolation,"* Proceedings MEGA'92, in: Progress in Mathematics, Birkhäuser, V. 109, 1993, 91-104.

[GKV 94] D. Grigoriev, M. Karpinski, and N. Vorobjov *"Lower bounds on testing membership to a polyhedron by algebraic decision trees,"* Proceedings of ACM Symposium on Theory of Computing, 1994, 635-644.

[GM 88] M. Goresky, and R. MacPherson *"Stratified Morse Theory,"* Springer-Verlag, Berlin, 1988.

[GV 88] D. Grigoriev, and N. Vorobjov *"Solving systems of polynomial inequalities in subexponential time,"* Journal Symbolic Comp., V. 5, 1988, 37-64.

[GV 92] D. Grigoriev, and N. Vorobjov *"Counting connected components of a semialgebraic set in subexponential time,"* Computational Complexity, V. 2, 1992, 133-186.

[GV 94] D. Grigoriev, and N. Vorobjov *"Complexity lower bounds for computation trees with elementary transcendental function gates,"* Proceedings of IEEE Symposium on Foundations of Computer Science, 1994, 548-552.

[Hi 76] M. Hirsch *"Differential Topology,"* Springer-Verlag, 1976.

[L 65] S. Lang *"Algebra,"* Addison-Wesley, New York, 1965.

[Lo 82] R. Loos *"Generalized polynomial remainder sequences,"* in: Symbolic and Algebraic Computation, B. Buchberger et al., eds., Springer-Verlag, New York, 1982.

[M 64] J. Milnor *"On the Betti numbers of real varieties,"* Proc. Amer. Math. Soc., V.15, 1964, 275-280.

[MH 85] F. Meyer auf der Heide *"Fast algorithms for n-dimensional restrictions of hard problems,"* Proceedings of ACM Symposium on Theory of Computing, 1985, 413-420.

[MS 71] P. McMullen, and G. Shephard *"Convex Polythopes and the Upper Bound Conjecture,"* Cambridge University Press, Cambridge, 1971.

[RV 94] M. -F. Roy, and N. Vorobjov *"Finding irreducible components of some real transcendental varieties,"* Computational Complexity, V. 4, 1994, 107-132.

[T 51] A. Tarski *"A Decision Method for Elementary Algebra and Geometry,"* University of California Press, 1951.

[Th 77] J. A. Thorpe *"Elementary Topics in Differential Geometry,"* Springer-Verlag, Berlin, 1977.

[Y 92] A. Yao *"Algebraic decision trees and Euler characteristics,"* Proceedings of IEEE Symposium on Foundations of Computer Science, 1992, 268-277.

[Y 94] A. Yao *"Decision tree complexity and Betti numbers,"* Proceedings of ACM Symposium on Theory of Computing, 1994, 615-624.

[YR 80] A. Yao, and R. Rivest *"On the polyhedral decision problem,"* SIAM J. Comput., V. 9, 1980, 343-347.