Some Algebraic and Geometric Computations in PSPACE

John Canny

543 Evans Hall,
Computer Science Division,
University of California, Berkeley

Abstract

We give a PSPACE algorithm for determining the signs of multivariate polynomials at the common zeros of a system of polynomial equations. One of the consequences of this result is that the "Generalized Movers' Problem" in robotics drops from EXPTIME into PSPACE, and is therefore PSPACE-complete by a previous hardness result [Rei]. We also show that the existential theory of the real numbers can be decided in PSPACE. Other geometric problems that also drop into PSPACE include the 3-d Euclidean Shortest Path Problem, and the "2-d Asteroid Avoidance Problem" described in [RS]. Our method combines the theorem of the primitive element from classical algebra with a symbolic polynomial evaluation lemma from [BKR]. A decision problem involving several algebraic numbers is reduced to a problem involving a single algebraic number or primitive element, which rationally generates all the given algebraic numbers.

1 Introduction

Ever since Tarski's paper [Tar] on the decidability of the theory of real closed fields, there have been steady improvements in the time and space bounds of algebraic algorithms. A double-exponential time algorithm for the theory was given by Collins [Col]. Collins' method was later improved by Ben-Or, Kozen and Reif [BKR] to single-exponential space. If only existential quantification is allowed, the theory of the reals has an EXPTIME solution, described in [GV]. The [BKR] method was interesting in that it was purely symbolic, i.e. it did not require numerical root isolation, as did Collins' method. Instead [BKR] made use of a polylog space algorithm for finding the signs of a set of polynomials at the roots of a given polynomial.

In this paper we generalize the [BKR] lemma, which applied to the roots of a single univariate polynomial, to the common zeros of several polynomials in several variables. Our generalized algorithm requires only polynomial space (the problem it solves is at least NP-hard). Using this result, we obtain a PSPACE algorithm for the existential theory of the reals. We also apply the lemma to obtain PSPACE upper bounds for several problems in robotics.

The "Generalized Movers' Problem" which is the problem of finding a collision-free path for a robot with n degrees of freedom moving among polyhedral obstacles, was shown to be PSPACE-hard by Reif [Rei]. The problem was treated by [Loz] for the restricted cases of motion without rotation in two and three dimensions. Later [SS] gave a double-exponential time algorithm for the general problem based on Collins' cellular algebraic decomposition [Col]. The space bound was improved to single-exponential by Kozen and Yap [KY].

Eventually, a single exponential time algorithm for the movers' problem was described in [C87b]. In this paper, we bring that algorithm into PSPACE, showing that the movers' problem is PSPACE-complete.

The *3-d euclidean shortest path problem* is the problem of finding the shortest path between two given points which avoids some polyhedral obstacles. The problem has been dealt with by Sharir and Schorr [SSc] who gave an $2^{2^{O(n)}}$ algorithm by reducing the problem to an algebraic decision problem in the theory of real closed fields. This was improved by Reif and Storer [RSt] who gave $2^{n^{O(1)}}$-time ($n^{O(\log n)}$)-space algorithm using the same theory but with a more efficient reduction.

We define the *2-d asteroid avoidance problem* as the problem of determining a collision-free path for a point in the plane with bounded velocity magnitude, with convex polygonal obstacles moving with fixed linear velocity (no rotation). The obstacles are assumed not to collide. In [RS] the problem was shown to be solvable in time $2^{n^{O(1)}}$.

In [CR] both the 3-d euclidean shortest path and the 2-d asteroid avoidance problems were shown to be NP-hard. In [C87a] improvements in the exponents of upper bounds for both problems were given. These improvements were based on the roadmap algorithm and on the multivariate resultant as an equation solving tool. Here we improve the upper bounds for both problems to PSPACE.

© 1988 ACM-0-89791-264-0/88/0005/0460 $1.50

In section 2 we give three lemmas on primitive elements. The first is a polylog-space algorithm for computing a primitive element polynomial for a pair of univariate polynomials. The second is a PSPACE algorithm for computing a primitive element polynomial for a system of multivariate polynomials. Using the latter result and the main lemma from [BKR], we give our main lemma which is a PSPACE algorithm for computing the signs of a set of multivariate polynomials at the common zeros of a system of polynomial equations.

In section 3 we use the main lemma to give a PSPACE decision algorithm for the existential theory of the reals. PSPACE upper bounds for the 3-d euclidean shortest path problem and the asteroid-avoidance problem follow as straightforward corollaries. We also describe modifications of the roadmap algorithm of [C87b] which runs in PSPACE.

2 Computing with Primitive Elements

Let $p_1(x)$ and $p_2(x)$ be polynomials of degree d_1 and d_2 respectively, each having only simple roots. We denote the roots of $p_1(x)$ as

$$\alpha_1, \ldots, \alpha_{d_1} \tag{1}$$

and the roots of $p_2(x)$ as

$$\beta_1, \ldots, \beta_{d_2} \tag{2}$$

then we have the following constructive version of the theorem of the primitive element [Wae]:

Lemma 2.1 *Let $p_1(x)$ and $p_2(x)$ be square-free polynomials of degree d_1 and d_2 respectively. Then there is a polynomial $q(x)$ of degree d_1d_2 and rational functions $r_1(x)$ and $r_2(x)$, such that every pair (α_i, β_j) of roots of p_1, p_2 equals $(r_1(\theta_{i,j}), r_2(\theta_{i,j}))$ for some root $\theta_{i,j}$ of $q(x)$. The computation requires polylog space.*

In other words there is a one-to-one correspondence between the d_1d_2 pairs (α_i, β_j) and the d_1d_2 roots $\theta_{i,j}$ of q. Furthermore each pair is rational in the corresponding root $\theta_{i,j}$.

Proof The proof closely follows [Wae]. We choose a constant c and set

$$\theta_{i,j} = \alpha_i + c\beta_j \tag{3}$$

Clearly if we want to recover α_i and β_j from $\theta_{i,j}$, we must ensure that $\theta_{i,j}$ takes on distinct values for distinct pairs (α_i, β_j). So we must have

$$\alpha_i + c\beta_j \neq \alpha_k + c\beta_l \tag{4}$$

for $i \neq k$ and $j \neq l$. Assume for the moment that c satisfies (4), we compute the greatest common divisor of $p_1(\theta - cx)$ and $p_2(x)$ as polynomials in x. If $\theta = \alpha_i + c\beta_j$, then $x = \beta_j$ is a common root, and by the inequalities (4) there are no other common roots, so the GCD is linear.

From the theory of subresultants [BT] we have the following facts:

- The resultant of two polynomials vanishes if and only they have a common root.
- If the GCD of two polynomials is linear, then it is *similar to* (in this case a constant multiple of) the first subresultant of the polynomials.

We denote the resultant (with respect to x) of polynomials $a(x)$ and $b(x)$ as $\mathrm{RES}_x(a(x), b(x))$. Then from the first property above, the polynomial $q(\theta) = \mathrm{RES}_x(p_1(\theta - cx), p_2(x))$ vanishes only at $\theta = \theta_{i,j}$, so $q(x)$ is the primitive element polynomial we were looking for. The rational functions are found from the first subresultant which is of the form

$$d(\theta)x + n(\theta) \tag{5}$$

and which vanishes at the common root $x = \beta_j$. The rational function r_2 is just

$$r_2(\theta) = -\frac{n(\theta)}{d(\theta)} \tag{6}$$

and since $\alpha_i = (\theta_{i,j} - c\beta_j)$, r_1 is given by

$$r_1(\theta) = \theta + c\frac{n(\theta)}{d(\theta)} \tag{7}$$

Both the resultant and the first subresultant can be computed in polylog space from the Sylvester matrix using the determinant algorithm of [Csa]. Inspection of the Sylvester matrix for $p_1(\theta - cx)$ and $p_2(x)$ shows that the degree of $q(x)$ is d_1d_2. The degrees of r_1 and r_2 are less than d_1d_2, where we define the degree of a rational function as the maximum of the degrees of its numerator and denominator.

It remains to show that we can compute a suitable constant c. There are only finitely many bad values of c, each one failing to satisfy one of the inequalities (4). Given p_1 and p_2 we choose c as follows:

First, we construct for both p_1 and p_2 "difference" polynomials $\Delta_1(y)$ and $\Delta_2(y)$. Each Δ_i is the resultant with respect to x of the polynomials $p_i(x)$ and $p_i(x + y)$. This resultant vanishes if and only if the two polynomials have a common root, that is if there is some α such that both $x = \alpha$ and $x = \alpha + y$ are roots. Every root of $\Delta_1(y)$ is therefore of the form $(\alpha_i - \alpha_j)$ for some pair of roots of p_1, and similarly for $\Delta_2(y)$.

We denote the non-zero roots of $\Delta_1(y)$ as $\delta^{\alpha}_{i,k}$ and the non-zero roots of $\Delta_2(y)$ as $\delta^{\beta}_{j,l}$, then (4) can be expressed as

$$\delta^{\alpha}_{i,k} \neq c\delta^{\beta}_{j,l} \tag{8}$$

for all $i \neq k$, $j \neq l$. Since we are only interested in non-zero roots of $\Delta_i(y)$, we divide each $\Delta_i(y)$ by its lowest power of y, giving polynomials $\Delta'_i(y)$ having only non-zero roots.

Now we compute the resultant with respect to y of the polynomials $\Delta'_1(zy)$ and $\Delta'_2(y)$. This gives us a resultant polynomial $R(z)$ which is zero for each value $z = c$ which violates some inequality in (8). The resultant $R(z)$ has polynomial degree and coefficient length in the degree and coefficient lengths of p_1 and p_2, and knowing only these bounds we can readily obtain upper bounds on the size of any root of $R(z)$, [Mig]. These bounds are again polynomial, and choosing a value of c slightly larger than the bound guarantees that it is not a root of $R(z)$ and that the inequalities (4) are satisfied. The computations for c can be performed in polylog space. □

The following lemma shows that the solution rays of a system of polynomials can be represented as the values of a series of rational functions evaluated at the roots of a univariate polynomial. A similar result is proved in [Ren], although the rational functions defined there are different. The computations in [Ren] were not shown to be in PSPACE, but it is easily shown that they are by Csanky's result [Csa]. The result proved here is more general, and does not require that there be only finitely many solution rays at infinity.

Lemma 2.2 *Let $p_1(x_0,\ldots,x_n),\ldots,p_n(x_0,\ldots,x_n)$ be homogeneous polynomials with $D \le \prod d_i$ isolated solution rays $(\alpha_{0,j},\ldots,\alpha_{n,j})$, $j=1,\ldots,D$. Let $N \le D$ be the number of solution rays not "at infinity" i.e. those which have, say $\alpha_{0,j} \neq 0$. Then there is a polynomial $q(x)$ of degree N, and rational functions $r_1(x),\ldots,r_n(x)$, such that every solution ray not at infinity is of the form $(1, r_1(\theta),\ldots,r_n(\theta))$ for some root θ of $q(x)$. The polynomials $q(x)$ and $r_k(x)$ can be computed in polynomial space.*

Proof We add the linear polynomial $u_0x_0 + \cdots + u_nx_n$ to the system of polynomials, giving a system of $n+1$ polynomials in $n+1$ variables. The resultant polynomial $R(u_0,\ldots,u_n)$ is called the u-resultant [Wae]. It factors into linear factors, the coefficients of which are the solution rays. That is, the linear factors of $R(u_0,\ldots,u_n)$ are of the form

$$(\alpha_{0,j}u_0 + \alpha_{1,j}u_1 + \cdots + \alpha_{n,j}u_n) \tag{9}$$

Note that this method does not work if there are infinitely many solution rays at infinity. However, as shown in [C88b], this case can be dealt with by using the lowest degree term of the generalized characteristic polynomial instead of the resultant, which takes essentially the same time to compute, and is no larger than the resultant.

We are looking for a single polynomial whose roots generate the coordinates of the solution rays. We construct it by defining a polynomial whose roots are linear combinations of the coordinates. We choose n constants $c_1,\ldots,c_n$ and set

$$q(x) = R(-x, c_1,\ldots,c_n) \tag{10}$$

In this case and later cases, R denotes the resultant if it is non-vanishing, or else the lowest degree coefficient of the generalized characteristic polynomial [C88a]. With this substitution, the factors of R corresponding to roots at infinity become constants, while the linear factors are of the form

$$(\alpha_{0,j}x - \sum_{i=1,\ldots,n} c_i\alpha_{i,j}) \tag{11}$$

So $q(x)$ has degree N. We assume without loss of generality that $\alpha_{0,j} = 1$ for rays not at infinity. Then the roots of $q(x)$ are θ_j where

$$\theta_j = \sum_{i=1,\ldots,n} c_i\alpha_{i,j} \tag{12}$$

To construct each rational function r_k, we make two other substitutions in the u-resultant. We set

$$\begin{aligned} q_k^+(x) &= R(-x, c_1,\ldots,(c_k+1),\ldots,c_n) \\ q_k^-(x) &= R(-x, c_1,\ldots,(c_k-1),\ldots,c_n) \end{aligned} \tag{13}$$

This may lead to polynomials with multiple roots, so we extract the square-free parts of q_k^+ and q_k^- and denote them $\hat{q}_k^+$ and $\hat{q}_k^-$ respectively. If $\theta = \theta_j$, then $\hat{q}_k^-(x)$ and $\hat{q}_k^+(2\theta - x)$ have a common root if and only if

$$\sum_{i=1,\ldots,n} c_i\alpha_{i,m} - \alpha_{k,m} = 2\sum_{i=1,\ldots,n} c_i\alpha_{i,j} - \sum_{i=1,\ldots,n} c_i\alpha_{i,l} - \alpha_{k,l} \tag{14}$$

There will certainly be a common root if $l = m = j$, but there can also be spurious roots if the c_i satisfy (14) for some $\alpha_{i,m} \neq \alpha_{i,j}$. Assuming for the moment that they do not, then the common root must be $x = \theta_j - \alpha_j$, and it will be a root of the first subresultant. We now compute the first subresultant of $\hat{q}_k^-(x)$ and $\hat{q}_k^+(2\theta - x)$, and let it be $d_k(\theta)x + n_k(\theta)$. The rational function $r_k(\theta)$ is given by

$$r_k(\theta) = \frac{n_k(\theta)}{d_k(\theta)} + \theta \tag{15}$$

In [Ren] a rather different method is given for finding the rational functions. That method gives rational functions of much lower degree, although in the present case, it is possible to keep the degree under control by repeatedly reducing polynomials mod $q(x)$ during GCD computations.

The computation of $q(x)$, $q_k^-(x)$ and $q_k^+(x)$ from the input polynomials requires polynomial space using the multivariate resultant algorithm described in [C87b]. These polynomials have degree D and coefficient length $O(bD)$. The subresultant and GCD computations for $d_k(x)$ and $n_k(x)$ can be performed in space polynomial in $\log(bD)$ using [Csa]. Thus overall the space required is polynomial in the input size.

We have postponed until now the question of how to find a suitable tuple $C = (c_1,\ldots,c_n)$ of constants. The bad choices of C are of two types, and both can be detected during execution. The first type are those that cause $q(x)$ or $q_k^+(x)$ or $q_k^-(x)$ to vanish identically. $q(x)$ will vanish identically if inner product of the tuple of c_i's and a solution ray at infinity is zero, i.e. if

$$\sum \alpha_{i,j}c_i = 0 \quad \text{and} \quad \alpha_{0,j} = 0 \tag{16}$$

for some j. Similar conditions apply to $q_k^+(x)$ where c_k is replaced by $c_k + 1$ and to $q_k^-(x)$ with c_k replaced by $c_k - 1$.

Other bad choices of c_i are those that satisfy (14) for distinct j, l, m. All these equations are linear in the c_i. For any bad choice of the c_i, there will be a spurious root at $\theta = \theta_j$. So $\hat{q}_k^-(x)$ and $\hat{q}_k^+(2\theta - x)$ will have a common factor of degree two or greater, and their first subresultant will vanish identically. Thus $q(\theta)$, $d_k(\theta)$ and $n_k(\theta)$ would all vanish at $\theta = \theta_j$. i.e. Their GCD would be non-trivial. A necessary and sufficient test for the absence of spurious roots is to compute the GCD of these polynomials and determine if it is a constant.

Any tuple of c_i's which does not satisfy equation (14) for distinct j, l, m or equation (16) will work. Since all the bad choices satisfy certain equations, the set of bad values has measure zero. If a random integer tuple C is chosen it should work with probability one. Since we gave tests for both types of failure, we can construct a Las Vegas type probabilistic algorithm by repeatedly trying random C's until one succeeds.

A less efficient but purely deterministic approach is to introduce an indeterminate β and set $c_i = \beta^i$ for $i = 1, \ldots, n$. Then all the equations for bad choices of c_i's become non-vanishing polynomials in β. Thus there are some values of β (in fact all sufficiently small values of β) which do not satisfy any of the equations. This guarantees that when GCD's are computed, spurious roots will not occur.

The result of the algorithm for this choice of c_i's will be a polynomial $q(x)$ and rational functions $r_k(x)$ whose coefficients are polynomials in β. Manipulation of polynomial coefficients rather than integers causes the running time of the algorithm to increase by a polynomial factor. The space bound is still polynomial by the earlier argument for integer c_i's. □

The following lemma is due to Ben-Or, Kozen and Reif [BKR]:

Lemma 2.3 *Given a polynomial $p(x)$ and polynomials $q_i(x)$, $i = 1, \ldots, m$ the sign sequences $(\mathrm{sign}(q_1(\alpha_j)), \ldots, \mathrm{sign}(q_m(\alpha_j)))$ at real roots α_j of $p(x)$ can be computed in polylog space.*

We can now state the main lemma on the computation of sign sequences of multivariate polynomial systems.

Lemma 2.4 *Let $p_1(x_1, \ldots, x_n) = 0, \ldots, p_n(x_1, \ldots, x_n) = 0$ be a system of integer coefficient polynomial equations having a finite number of solution points. Denote the N real solution points not at infinity as $a_j \in \Re^n$, $j = 1, \ldots, N$. Let $q_1(x_1, \ldots, x_n), \ldots, q_m(x_1, \ldots, x_n)$ be a set of polynomials. Then the set of sign sequences $(\mathrm{sign}(q_1(a_j)), \ldots, \mathrm{sign}(q_m(a_j)))$, $j = 1, \ldots, N$ can be computed in polynomial space.*

Proof We first introduce a homogenizing variable x_0, and for each $p_i(x_1, \ldots, x_n)$, we multiply every term by a power of x_0 so that the degree of the term equals the degree of p_i. This gives a set of polynomials p'_k which is homogeneous, and satisfies the conditions for lemma 2.2. We use that lemma to compute a polynomial $q(x)$ whose roots generate all the solution points not at infinity via rational functions $r_k(x)$. We use the deterministic version of that lemma, where the c_i's are powers of an indeterminate β.

Now we show that a root θ_j of $q(x)$ is real if and only if all the coordinates of the corresponding solution point $(\alpha_{1,j}, \ldots, \alpha_{n,j})$ are real. If the coordinates $\alpha_{i,j}$ are all real, so is θ_j since it is just the linear combination $\sum_{i=1,\ldots,n} c_i \alpha_{i,j}$ with real weights c_i. The converse is true if the rational functions r_k have all real coefficients. They will because they are computed from the coefficients of the p_i by rational operations only (resultant and subresultant computations).

We next substitute $x_k = r_k(\theta)$ in each polynomial q_l. This produces a rational function $R_l(\theta)$ whose denominator is a product of powers of the denominators of the $r_k(\theta)$. If the denominator of $R_l(\theta)$ is an odd power of the denominator $d_k(\theta)$ of $r_k(\theta)$, we multiply numerator and denominator of $R_l(\theta)$ by $d_k(\theta)$. This does not change the sign of $R_l(\theta)$, but ensures that its denominator is positive. Let the numerator of $R_l(\theta)$ after the above multiplications be $Q_l(\theta)$. Then $Q_l(\theta)$ is a polynomial with the same sign as the rational function $R_l(\theta)$, (so long as the denominator of $R_l(\theta)$ is non-zero, which it will be at any root $\theta = \theta_j$ of $q(\theta)$)

We now have a system of polynomials $Q_l(\theta)$ such that the sign sequence $(\mathrm{sign}(Q_1(\theta_j)), \ldots, \mathrm{sign}(Q_m(\theta_j)))$ at some real root θ_j of $q(\theta)$ is the same as the sign sequence $(\mathrm{sign}(q_1(a_j)), \ldots, \mathrm{sign}(q_m(a_j)))$ at the real solution point $a_j \in \Re^n$ corresponding to θ_j.

So our original problem reduces to the problem of enumerating the sign sequences of the $Q_l(\theta)$ at the real roots of the *single* polynomial $q(\theta)$. For this we can make use of the last lemma from [BKR]. The only caveat is that [BKR] requires determination of the signs of certain polynomials in the coefficients of q and the Q_l's. Because we introduced the indeterminate β, these latter polynomials will be polynomials in β. To determine their signs, we take the sign of the coefficient of the lowest degree term in β. This term correctly gives the sign of the polynomial for arbitrarily small β. Recall that we interpreted β as an arbitrarily small quantity so that the c_i's not satisfy (14) or (16).

Now $q(\theta)$ and the $Q_l(\theta)$ will have degree $O(D^2)$ and coefficient size $O(bD^4)$, and the [BKR] algorithm requires space polylogarithmic in the degree and coefficient size of the polynomials passed to it. In our case these polynomials have exponential size in the input, but polylog of their size is polynomial in the input size. Thus the computation of the sign sequences can be accomplished in polynomial space. □

3 Applications

The lack of an efficient method for computation with algebraic numbers defined by *systems* of polynomials has been a barrier to efficient parallel solutions to a number of algebraic and geometric problems. The main lemma of the last section removes this obstacle and leads to PSPACE algorithms for these problems.

3.1 The Existential Theory of the Reals

An existentially-quantified formula in the first order theory of the reals is a formula of the form:

$$\exists x_1 \exists x_2 \ldots \exists x_n \; P(x_1, x_2, \ldots, x_n)$$

where each $x_i \in \Re$, and where $P(x_1, x_2, \ldots, x_n)$ is a predicate which is a monotone boolean function of atomic predicates of the form

$$f_i(x_1, x_2, \ldots, x_n) \geq 0 \quad \text{or} \quad g_i(x_1, x_2, \ldots, x_n) > 0$$

There is no loss of generality in this form for P over forms with arbitrary boolean functions and arbitrary inequalities. Deciding the existential theory is equivalent to deciding whether sets of the form

$$S_P = \{(x_1, x_2, \ldots, x_n) \in \Re^n \mid P(x_1, x_2, \ldots, x_n)\} \tag{17}$$

are non-empty. Such a set is called a *semi-algebraic set.* The basic method is to test the predicate at a collection of sample points in $\Re^n$ one of which is guaranteed to be in the set if it is non-empty. The sample points are extremal points in some direction of the closures of *sign-invariant sets.* A sign-invariant set is the set of points where all the polynomials defining S_P have particular signs. There are two obstacles to applying this method directly: (i) The set S_P may not be compact, so that extremal points may not exist, and (ii) intersections between surfaces given by the defining polynomials may be singular, in which case it may be difficult to compute the extremal points.

Step 1: Compactness

To get compactness, we first need to bound the quantifiers, which we do by substituting

$$x_i = \frac{y_i}{1 - y_i^2} \qquad \text{for } i = 1, \ldots, n \tag{18}$$

in all the polynomials in the predicate P. We clear denominators, and adjoin the inequalities $-1 < y_i < 1$ for $i = 1, 2, \ldots, n$ with conjunctions, and call the new predicate $Q(y_1, y_2, \ldots, y_n)$. This predicate defines a semi-algebraic set in the cube $(-1, 1)^n$. Adding existential quantification over the y_i's we get a new formula which has size polynomial in the original, and is true if and only if the original was true.

Next we replace all strict inequalities $g_i(y_1, y_2, \ldots, y_n) > 0$ with inequalities $g_i(x_1, x_2, \ldots, x_n) \geq \epsilon$. Write this new predicate as $Q^\epsilon(y_1, y_2, \ldots, y_n)$. We write S_Q for the set of points where Q is true and similarly for S_{Q^ϵ}, and we claim that

Lemma 3.1 *S_P is non-empty if and only if there exists an $\epsilon > 0$ such that S_{Q^ϵ} is non-empty. Furthermore, if S_{Q^ϵ} is non-empty, then so is $S_{Q^{\epsilon'}}$ for any positive $\epsilon' < \epsilon$.*

Proof

We have already seen that S_P is non-empty if and only if S_Q is. If S_Q is non-empty, there must exist some positive ϵ such that S_{Q^ϵ} is non-empty. To see this, choose any point $y \in S_Q$, and let ϵ be the minimum over i of the $g_i(y)$'s which are *positive*. Then for each i, $g_i(y) \geq \epsilon$ if and only if $g_i(y) > 0$. Conversely, if S_{Q^ϵ} contains some point y for any $\epsilon > 0$, then that point also lies in S_Q.

Since the predicate is a monotone boolean function, and since ϵ always appears on the right of a $\geq$, if the predicate is true of a point p for some value of ϵ, it will be true of that point for all ϵ less than that value. □

Taking ϵ to be some fixed positive value, we notice that S_{Q^ϵ} is constructed by finite union and intersection of sets of the form $f_i(y) \geq 0$ or $g_i(y) \geq \epsilon$ which are closed, so S_{Q^ϵ} is closed. Since it also lies in the bounded cube $(-1, 1)^n$, it is compact.

Step 2: Reduction to the non-singular case

We assume that the compact set S_{Q^ϵ} is defined by a predicate which is a monotone boolean function of inequalities $f_i(y) \geq 0$, all $>$'s having been converted to $\geq$'s. Suppose there are m polynomials $f_i : \Re^n \to \Re$. Then they define a map $f : \Re^n \to \Re^m$ We first partition $\Re^m$ into regions of fixed sign for each coordinate, and then look at the preimages of these regions in $\Re^n$. The latter are the sign-invariant sets.

Let $\Re^-$ denote the set $(-\infty, 0)$ and $\Re^+$ denote $(0, \infty)$. The real line $\Re$ can be partitioned into $\underline{\Re} = \{\Re^-, \{0\}, \Re^+\}$ which is a Whitney regular stratification [GWD], [C87b]. A Whitney regular stratification is a partition of a set into manifolds (strata) which satisfies certain tangency conditions where two strata meet. Similarly we partition $\Re^m$ as $\underline{\Re}^m$. This is a regular stratification of $\Re^m$. Its elements are of the form $\sigma = \sigma_1 \times \sigma_2 \cdots \sigma_m$ where each σ_i is either $\Re^+$, $\Re^-$ or $\{0\}$. σ may be thought of as a sign sequence for the f_i's. The preimage $f^{-1}(\sigma)$ of a sign sequence $\sigma \in \underline{\Re}^m$ is called a *sign-invariant* set. We write $f^{-1}(\underline{\Re}^m)$ for the collection of all sign-invariant sets. If the map f is transversal to $\underline{\Re}^m$ (meaning it is transversal to all the strata in $\underline{\Re}^m$, [GWD]), then the preimage $f^{-1}(\underline{\Re}^m)$ is a regular stratification of $\Re^n$.

The map f that defines S_{Q^ϵ} will not in general be transversal to $\underline{\Re}^m$. The sign-invariant sets may not be manifolds, and in these cases it is difficult to compute sample points. However, if we introduce m extra variables $a_1, \ldots, a_m$ and set

$$f_i^a(y_1, \ldots, y_n, a_1, \ldots, a_m) = f_i(y_1, \ldots, y_n) + a_i \tag{19}$$

then the map $f^a : \Re^{(n+m)} \to \Re^m$ is transversal to all the strata. Thus the preimage $f^{-1}(\underline{\Re}^m)$ is a Whitney regular stratification of $\Re^{(n+m)}$.

If we define a real positive quantity δ as

$$a_1^2 + a_2^2 + \cdots + a_m^2 = \delta^2 \tag{20}$$

then the manifold of non-zero vectors $(a_1, \ldots, a_m)$ is diffeomorphic to the product of the $m - 1$ sphere Σ^{m-1}, giving the directions of the vectors, and $\Re^+$ giving their δ values (euclidean norms). By abuse of notation, we write f^a again for the function f^a (restricted to the set of non-zero a vectors) composed with this diffeomorphism, and write $f^a : \Re^n \times \Sigma^{m-1} \times \Re^+ \to \Re^m$. f^a is still a submersion, and therefore transversal to all the strata of $\underline{\Re}^m$.

We can also think of f^a as a family of maps f_δ^a, parameterized by δ, so that $f_\delta^a : \Re^n \times \Sigma^{m-1} \to \Re^m$. Now we can make use of the generic map lemma [GG] lemma 4.6, which states that if a parametrized family of maps is transversal to a manifold, then the individual maps are transversal to that manifold *for a dense subset of the parameters.* In our case this means that for a dense subset of $\Re^+$, choosing a fixed δ still gives us a map transversal to some stratum of $\underline{\Re}^m$. The set of δ values which give us a map transversal to *all* of the strata in $\underline{\Re}^m$ is the intersection of finitely many dense subsets, and is therefore dense. Finally, the set of "bad" values of δ is semi-algebraic, and since its complement is dense, it must be finite. Thus f_δ^a is transversal to $\underline{\Re}^m$ for *all* sufficiently small $\delta > 0$.

Now consider the predicate $T^\epsilon_\delta(y_1, \ldots, y_n, a_1, \ldots, a_m)$ formed from the predicate $Q^\epsilon(y_1, \ldots, y_n)$ by replacing each polynomial f_i with $f^a_i = f_i + a_i$ and adding the constraint (20). We let $S_{T^\epsilon_\delta} \subset \Re^n \times \Re^m$ denote the set of points for which T^ϵ_δ is true. Now $S_{T^\epsilon_\delta}$ is a closed set, since all the inequalities defining it are $\geq$'s or $=$. Furthermore, the y values are all in the range $(-1-\delta, 1+\delta)$, while all a values are in the range $(-\delta, \delta)$. Thus $S_{T^\epsilon_\delta}$ is bounded and therefore compact.

Lemma 3.2 *S_{Q^ϵ} is non-empty if and only if there exists a $d(\epsilon) > 0$ such that for all positive $\delta < d(\epsilon)$, $S^{T^\epsilon_\delta}$ is non-empty.*

Proof

We replace the equality (20) in the predicate T^ϵ_δ with $\sum a_i^2 < d^2$, giving a new predicate $T^\epsilon_{\leq d}$. Now $S_{T^\epsilon_{\leq d}}$ is a compact set, and contains all the sets $S_{T^\epsilon_\delta}$ for all positive $\delta \leq d$.

To show $S_{T^\epsilon_\delta}$ non-empty for small enough δ implies S_{Q^ϵ} non-empty, we pick a sequence $(\delta_i) \rightarrow 0$, with all $0 < \delta_i < d$. For each δ_i there is a point p_i in $S_{T_{\delta_i}}$. The sequence p_i, since it lies in the compact set $S_{T^\epsilon_{\leq d}}$, has a convergent subsequence with some limit point p also in $S_{T^\epsilon_{\leq d}}$. In fact p lies in $S_{T^\epsilon_{\leq 0}}$ which equals $S_{Q^\epsilon} \times \{0\}^m$, which shows that S_{Q^ϵ} is non-empty.

Sketch of proof of the converse: we start from any point in $S_{Q^\epsilon} \times \{0\}^m$ and move in some direction that changes the values of some of the f_i, while adjusting the values of the a_i so that the f^a_i's are unchanged. This gives points in $S_{T_{\leq\delta}}$ for all sufficiently small δ.

In detail we define a path $q : [0,1] \rightarrow \Re^n \times \Re^m$ in terms of a path $p : [0,1] \rightarrow \Re^n$ as follows: let $p(0)$ be any point in S_{Q^ϵ}. Choose any f_j which is not identically zero, and let $\eta \in \Re^n$ be any vector not in the tangent cone of the zeros set of f_j at p. We define $p(s) = p(0) + s\eta$, and $q(s) = p(s) \times (f(p(0)) - f(p(s)))$. Then for any s, $f^a(q(s)) = f(p(0))$ so clearly all the points $q(s)$ lie in $S_{T^\epsilon_{\leq d}}$ for large enough d.

We choose $d = \max(\|f(p(s)) - f(p(0))\|)$ for $s \in [0,1]$, where $\|\cdot\|$ denotes euclidean norm. We know $d > 0$ since the tangent (η) to $p(s)$ at $p(0)$ is not in the tangent cone of the zeros set of f_j, so $f_j(p(s))$ must differ from $f_j(p(0))$ for some small s. d gives the euclidean distance of the furthest point (in a coordinates) on the path $q(s)$ from $q(0)$. Let $q_{\max}$ be any point on the path $q(s)$ where this maximum is attained. Then some segment of the path $q(s)$ joins $q(0)$ to the point $q_{\max}$. The euclidean magnitude of the a-coordinates of points (which equals their δ-values from (20)) on this segment varies continuously from 0 to d, and so the segment contains points in $S_{T^\epsilon_\delta}$ for all $0 \leq \delta \leq d$. □

Theorem 3.3 *The existential theory of the reals is decidable in PSPACE.*

Proof

Given a predicate $P(x_1, \ldots, x_n)$, using the last two lemmas we can define a new predicate $T^\epsilon_\delta(y_1, \ldots, y_n, a_1, \ldots, a_m)$ such that S_P is non-empty if and only if there exists some $\epsilon > 0$ and a $d(\epsilon) > 0$ such that for all $0 < \delta < d$, the set $S_{T^\epsilon_\delta}$ is non-empty. Under reasonable measures of formula size, T^ϵ_δ has size polynomial in the size of P.

Now T^ϵ_δ defines a compact set $S_{T^\epsilon_\delta} \subset \Re^n \times \Re^m$, which has a regular stratification into sign-invariant sets. Once we have a compact, non-singular set, it is easy to compute sample points. We pick a non-zero vector $v \in \Re^n \times \Re^m$ and define a projection map $\pi_v : \Re^n \times \Re^m \rightarrow \Re$ as $\pi_v(x) = x \cdot v$. Now, since $S_{T^\epsilon_\delta}$ is compact, π_v attains a maximum value when restricted to it. Let p be the point where this maximum is attained. p lies in some smooth sign-invariant stratum of $S_{T^\epsilon_\delta}$.

We would like to find a finite set of points which is guaranteed to contain the point p. The tangent space of the stratum which contains p is determined by the polynomials that are zero at p, those that are non-zero are irrelevant. Let σ be the set of common zeros of the f^a_i that are zero at p. If p is an extremal point in direction v, then it will be a critical point of the map $\pi_v|_\sigma$, see [GG] or [C87b]. By enumerating the critical points of $\pi_v|_\sigma$ for every σ i.e. for every subset of the polynomials, we are guaranteed to get one sample point that lies in the set $S_{T^\epsilon_\delta}$ if it is non-empty.

To find the critical points, we can make use of the tube construction of [C87b] and [C88b]. Specifically, to find the critical points of the manifold σ defined by polynomials $f_{i_1}, \ldots, f_{i_k}$, we define the polynomial

$$g^\alpha = f_{i_1}^2 + \cdots + f_{i_k}^2 - \alpha \tag{21}$$

and observe that for all sufficiently small α, $g^\alpha = 0$ defines a smooth hypersurface. The vector $v \in \Re^n \times \Re^m$ gives the direction in which extremals are computed, and it can be shown that the set of directions for which all the critical points are isolated is dense [C87b]. We can therefore assume without loss of generality that $v_1 = 1$, and then critical points of π_v restricted to the hypersurface $g^\alpha = 0$ are defined by the following conditions:

$$\begin{aligned} g^\alpha &= 0 \\ \frac{\partial}{\partial x_j} g^\alpha - v_j \frac{\partial}{\partial x_1} g^\alpha &= 0 \quad \text{for } j = 2, \ldots, n+m \end{aligned} \tag{22}$$

as demonstrated in [C88b], as $\alpha \rightarrow 0$ there is a sequence of critical points of $\pi_v|_{(g^\alpha = 0)}$ converging to each non-degenerate critical point of $\pi_v|_\sigma$.

Using the main lemma in this paper, we can find the signs of all the other polynomials at each critical point of $\pi_v|_{(g^\alpha=0)}$, if ϵ, δ and α are fixed. We are interested though, in the signs of the polynomials at "vanishingly small" positive values of these variables. The main lemma reduces computation of the signs of the other f^a_i to determination of the signs of certain polynomials in the *coefficients* of g^α and the f^a_i. These will be polynomials in α, δ and ϵ. Now δ must be less than $d(\epsilon)$, and α must be sufficiently small that it defines a smooth tube around σ. Since σ may depend on both ϵ and δ, we need to choose α smaller than some function of ϵ and δ.

We get correct results by assuming that $0 < \alpha \ll \delta \ll \epsilon \ll 1$ where $\ll$ indicates that the quantity on the left of the inequality is much smaller than that on the right, and that they are not related by any polynomial, i.e. they are independent transcendentals. In concrete terms, the sign of a polynomial in α, δ and ϵ for this specialization equals the sign of the term of lowest degree under the lexicographic order of variables $\alpha \succ \delta \succ \epsilon$. This means we collect together all terms of lowest degree in α, and from these select the ones with lowest degree in δ, and finally pick the term of lowest degree in ϵ. It should be clear that if $0 < \alpha \ll \delta \ll \epsilon \ll 1$, all the other terms are vanishingly small compared to this term.

Once a set of signs is enumerated at a sample point, we can readily determine if the predicate T_δ^ϵ is true at that point (for $0 < \delta \ll \epsilon \ll 1$). By enumerating critical points of all σ, we are guaranteed to find one for which T_δ^ϵ is true if $S_{T_\delta^\epsilon}$ is non-empty. If T_δ^ϵ is not true at any sample point, then $S_{T_\delta^\epsilon}$ must be empty.

From the last two lemmas, we know that $S_{T_\delta^\epsilon}$ is non-empty for $0 < \delta \ll \epsilon \ll 1$ if and only if S_P is non-empty. Thus we have a decision method for the existential theory of the reals. Conversion from P to the predicate T_δ^ϵ can be done in polynomial time. Enumeration of subsets of the polynomials defining T_δ^ϵ for sample point computation requires polynomial space. The determination of sign sequences for sample points for one of these subsets makes use of our main lemma. Although we are applying the main lemma to polynomials whose coefficients contain α, δ and ϵ rather than integers, this only increases the running time of the algorithm by a polynomial factor, and it still requires polynomial space. Since each of the three steps requires at most polynomial space, the existential theory of the reals is decidable in polynomial space. □

3.2 Robot Motion Planning Problems

We obtain the following result as a corollary to the theorem just proved:

Corollary 3.4 *The 3-d euclidean shortest path problem and the 2-d asteroid avoidance problem are solvable in PSPACE.*

Proof Both these problems can be reduced to decision problems in the existential theory of the reals. For the euclidean shortest path problem, we define a predicate $PL(l)$ which is true if and only if there is a path of length $\leq l$ in a given polyhedral environment. We can obtain bits of the shortest path by binary search with repeated calls to $PL(l)$.

Given a polyhedral environment, it is not difficult to define a predicate $F(p_1, p_2, l_1)$, where $p_1, p_2 \in \Re^3$ and $l_1 \in \Re$, which is true if and only if the path segment from p_1 to p_2 is clear of obstacles and has length l_1. The shortest path in an environment with n obstacles edges consists of straight line segments with at most n bending or "via" points [SSc]. So the predicate $PL(l)$ can be defined as

$$\exists l_0, p_1, l_1, \ldots, p_n, l_n \bigcap_{i=0,\ldots,n} F(p_i, p_{i+1}, l_i) \cap \left(l \geq \sum_{i=0}^{n} l_i \right) \quad (23)$$

where p_0 and p_{n+1} are the start and end-point of the path respectively. The formula $PL(l)$ has size polynomial in the environment description, and for fixed l, it can be decided in PSPACE. Thus we can find polynomially many bits of the shortest path length in PSPACE.

The 2-d asteroid avoidance problem is an existence problem, and can be solved by defining a similar predicate. See e.g. [RS] or [C87a]. The only difference is that each path segment is in 2-d space plus time instead of 3-d, and that path segments must satisfy a velocity constraint, which amounts to a constraint in the slope of the segment in space-time. Once again the predicate has polynomial size in the size of the environment description, and we conclude that 2-d asteroid avoidance is decidable in PSPACE. □

Theorem 3.5 *The roadmap algorithm described in [C87b] can be programmed to run in PSPACE. It follows that the generalized movers' problem is PSPACE-complete.*

Proof The input to the roadmap algorithm is a semi-algebraic set specified by n polynomials $p_i(x_1, \ldots, x_r)$ of degree d in r variables. The roadmap algorithm as described in [C87b] uses finite length binary approximations to real algebraic numbers, which are solutions of some set of polynomial equations $q_j = 0$ derived from the p_i. All the algebraic computations in the algorithm reduce to substitution of these approximations into other polynomials and testing their signs. For our PSPACE result, we compute the signs of the latter polynomials directly using lemma 2.4.

Specifically, we replace the calls to algorithm 1 of [C87b] and subsequent evaluation of polynomials at the solution points, with a call to lemma 2.4. From the description of algorithm 3,4 and 5, we find that the calls to algorithm 1 involve at most r polynomials. These polynomials have degree at most $(d^{O(r^2)})$, the highest degree polynomials being those that define the slices $x_r = \alpha_r$ on which the algorithm is being called recursively. The algorithm described in lemma 2.4 runs in space polylogarithmic in the product of the degrees of the polynomial equations, that is polylogarithmic in $(d^{O(r^3)})$ in the present context. This amount of space is still polynomial in the input size.

The other steps in the roadmap algorithm, such as enumeration of silhouette curves, and ordering of points along those curves, and search of the (exponential size) adjacency graph of the roadmap can also be done in polynomial space. Thus the generalized movers' problem is solvable in PSPACE. PSPACE-completeness follows from the PSPACE-hardness result of [Rei]. □

4 Conclusions

We gave a method for reducing computations involving several algebraic numbers to computation with only a single algebraic number or primitive element. This gave us a symbolic method for evaluating the signs of a collection of polynomials at the common zeros of some polynomial equations. We saw that this lead to a space-efficient, and therefore parallelizable algorithm.

Using this lemma, we were able to give a PSPACE decision algorithm for the existential theory of the real numbers. Our result required two preparation steps to reduce the non-emptiness test for an arbitrary semi-algebraic set to a test of non-emptiness for a compact non-singular semi-algebraic set. We were then able to test for non-emptiness by enumerating certain extremal points.

This result gave us PSPACE decision algorithms for the 3-d euclidean shortest path problem, and for the 2-d asteroid-avoidance problem. We also showed that the roadmap algorithm of [C87b] can be modified to run in PSPACE, and so the generalized movers' problem can be decided in PSPACE.

References

[BKR] Ben-Or M., Kozen D., and Reif J., "The Complexity of Elementary Algebra and Geometry", J. Comp. and Sys. Sciences, Vol. 32, (1986), pp. 251-264.

[BT] Brown W. S. and Traub J. F., "On Euclid's Algorithm and the Theory of Subresultants", JACM, Vol. 18, No. 4., (1971), pp 505-514.

[C87a] Canny J. F., "Exact Solution of Some Robot Motion Planning Problems", Proc. 4th International Symposium on Robotics Research, Santa Cruz, California, (1987).

[C87b] Canny J. F., "A New Algebraic Method for Motion Planning and Real Geometry", Proc. 28th IEEE Symp. FOCS, Los Angeles, (1987), pp 39-48.

[C88a] Canny J. F., "Generalized Characteristic Polynomials", submitted to ISSAC-88/AAECC-6, Rome, (July 1988).

[C88b] Canny J. F., "The Complexity of Robot Motion Planning", ACM Doctoral Dissertation Series, MIT Press, Cambridge, Mass. (1988).

[Col] Collins G.E. "Quantifier Elimination for Real Closed Fields by Cylindrical Algebraic Decomposition" Lecture Notes in Computer Science, No. 33, Springer-Verlag, New York, (1975).

[Csa] Csanky L., "Fast Parallel Matrix Inversion Algorithms" SIAM J. Comp., Vol. 5, No. 4, (Dec. 1976) pp 618-623.

[GWD] Gibson C. G., Wirthmüller K., Du Plessis A. A., Looijenga E. J. N., "Topological Stability of Smooth Mappings", Lecture Notes in Mathematics, No. 552, Springer-Verlag, New York, (1976).

[GG] Guilleman V., and Golubitsky M., "Stable Mappings and Their Singularities", GTM-14, Springer Verlag, New York, (1973).

[GV] Grigoryev D. Y. and Vorobjov N. N., "Solving Systems of Polynomial Inequalities in Subexponential Time" Jour. Symbolic Computation, special issue on decision algorithms for the theory of real closed fields, to appear (1988).

[KY] Kozen D., and Yap C. "Algebraic Cell Decomposition in NC", Proc 25th IEEE Symp. FOCS, (1985), pp. 515-521.

[Loz] Lozano-Pérez T., "Spatial Planning: A Configuration Space Approach," IEEE Trans. Computers, C-32, No 2 (Feb 1983), pp 108-120.

[Mig] Mignotte M., "Some Useful Bounds", in "Computer Algebra, Symbolic and Algebraic Computation", Buchberger et al. ed., Springer-Verlag, New York, (1982).

[Rei] Reif J., "Complexity of the Mover's Problem and Generalizations," Proc. 20th IEEE Symp. FOCS, (1979).

[RS] Reif J., and Sharir M., "Motion Planning in the Presence of Moving Obstacles", Proc. 25th IEEE Symp. FOCS, (1985), pp. 144-154.

[RSt] Reif J., and Storer J., "Shortest Paths in Euclidean Space with Polyhedral Obstacles", Tech. Rep. CS-85-121, Comp. Sci. Dept., Brandeis University, (April 1985).

[Ren] Renegar J., "On the Worst-Case Arithmetic Complexity of Approximating Zeros of Systems of Polynomials", Technical Report, School of Operations Research and Industrial Engineering, Cornell University, (May 1987).

[SS] Schwartz J. and Sharir M., "On the 'Piano Movers' Problem, II. General Techniques for Computing Topological Properties of Real Algebraic Mani folds," Computer Science Department, New York University report 41, (1982).

[SSc] Sharir M., and Schorr A., "On Shortest Paths in Polyhedral Spaces", Proc. 16th ACM STOC, (1984), pp. 144-153.

[Wae] van der Waerden B. L., "Modern Algebra", (third edition) F. Ungar Publishing Co., New York (1950).