Proving Polynomial-Time for Sphere-Constrained Quadratic Programming*

Stephen A. Vavasis
Richard Zippel

TR 90-1182
December 1990

Department of Computer Science
Cornell University
Ithaca, NY 14853-7501

*This report represents material drawn from a forthcoming book by the first author entitled *Nonlinear Optimization: Complexity Issues.* This work supported by the Applied Mathematical Sciences Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under Grant DE-FG02-86ER25013.A000 and in part by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

Proving polynomial-time for sphere-constrained quadratic programming*

Stephen A. Vavasis
Richard Zippel

Department of Computer Science
Cornell University

December 18, 1990

Abstract

Recently Ye and Karmarkar have proposed similar algorithms for minimizing a nonconvex quadratic function on a sphere. These algorithms are based on trust-region work going back to Levenberg and Marquardt. Although both authors state that their algorithm is polynomial time, neither makes estimates necessary to prove that conclusion in a formal sense. In this report we derive estimates for the convergence of the algorithm. Our estimates are based on bounds for separation of roots of polynomials. These bounds prove that the underlying decision problem is polynomial time in the Turing machine sense.

*This report represents material drawn from an forthcoming book by the first author entitled *Nonlinear Optimization: Complexity Issues.* This work supported by the Applied Mathematical Sciences Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under grant DE-FG02-86ER25013.A000 and in part by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

1 Quadratic programming on a sphere

The problem we consider is

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{x}^T\mathbf{x} \leq 1. \end{array} \tag{1}$$

Here H is an $n \times n$ symmetric matrix and $\mathbf{c}$ is an n-vector. This is the problem of minimizing a quadratic objective function on an n-dimensional sphere.

Problem (1) is of interest because it arises as a subproblem in more general optimization algorithms. Specifically, there is a class of nonlinear optimization algorithms called *model trust region methods.* In these algorithms, a quadratic function is used as a "model" of the nonlinear objective function around the current iterate $\mathbf{x}_c$. The next step is to minimize the model function. In general, however, the model is expected to be accurate (the model is "trusted") only in a neighborhood of $\mathbf{x}_c$. Accordingly, the model quadratic function is minimized over a neighborhood of $\mathbf{x}_c$. If the 2-norm is used to define the neighborhood, a problem exactly of the form (1) is produced. See Figure 1 for a 1-dimensional trust region step.

The model-trust region method is due to Levenberg [1944] and Marquardt [1963]. These authors considered only the case where H is positive definite. Moré [1978] proposed a robust algorithm with a convergence proof for the positive definite case. Gay [1979] and Sorenson [1982] proposed algorithms for the general case.

The solution to (1) will in general be irrational. For example, if H is negative definite and $\mathbf{c} = \mathbf{0}$ then the solution to (1) will be the eigenvector $\mathbf{x}$ of H with the smallest eigenvalue. The objective function optimum in this case is half the smallest eigenvalue. The smallest eigenvalue of a negative definite matrix will usually be irrational.

Accordingly, we cannot expect to exactly compute the minimum value of the objective function nor the optimum $\mathbf{x}$. The first goal of this report is therefore to show that $\mathbf{x}$ can be approximated efficiently. In particular, we show that p digits of $\mathbf{x}$ can be computed in time polynomial in p. This is a result due to Ye [1988] and Karmarkar [1989].

The second goal is to prove that the decision problem associated with (1) lies in $\mathbf{P}$, the class of polynomial time decision problems in the standard Turing machine definition (see Garey and Johnson [1979]). The decision

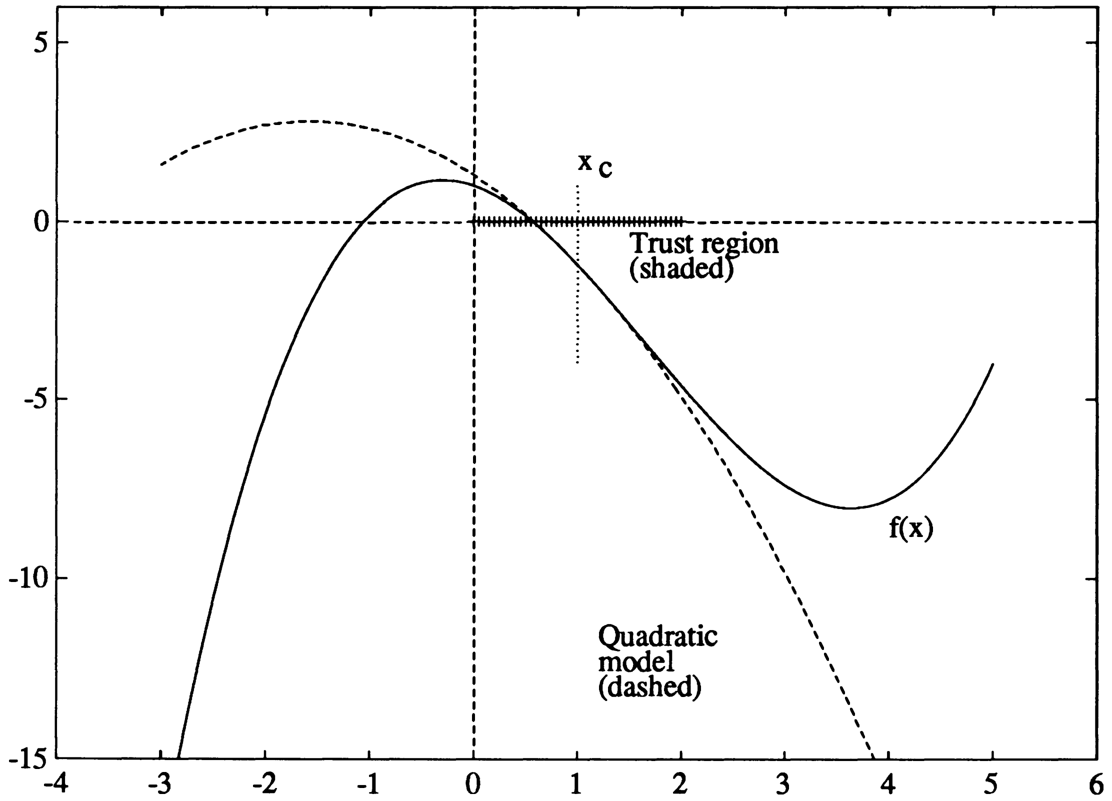

Figure 1: Model trust region when $n = 1$.

problem, which we call QPSPHERE, is as follows. Given an $n \times n$ symmetric matrix H with integer entries, an n-vector $\mathbf{c}$ with integer entries, and a rational number ζ, does there exist an $\mathbf{x} \in \mathbb{R}^n$ such that $\frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{c}^T\mathbf{x} \leq \zeta$ and $\mathbf{x}^T\mathbf{x} \leq 1$? To show that this decision problem lies in **P**, we call upon results concerning separation of roots of polynomials.

2 An algorithm for minimization

We now propose an algorithm for (1) due to Ye [1988]. A similar algorithm is due to Karmarkar [1989]. This algorithm is in practice less efficient than the algorithms mentioned in the last section, but it has the advantage of leading to a simple polynomial time analysis. The algorithm starts from the first and second-order necessary conditions for local minimality of (1).

Before writing the conditions, we first dispense with the case that $\mathbf{x}^*$ is interior to the feasible region. In this case, $\mathbf{x}^*$ is a local minimum for an unconstrained quadratic function. This is possible only if $\mathbf{x}^*$ is the unconstrained global minimum of the objective function. Such a minimum exists only if H is positive semidefinite. If H is positive definite, then the global minimum is the root to $H\mathbf{x}^* = -\mathbf{c}$. Accordingly, this case is handled by (1) testing whether H is positive definite, and if so, (2) solving $H\mathbf{x}^* = -\mathbf{c}$, then (3) testing whether $\mathbf{x}^{*T}\mathbf{x}^* \leq 1$. If all these conditions hold, we have found the minimum. Solving a system of linear equations with integer coefficients can be done in polynomial time by a Turing machine; this result goes back to Edmonds [1967]. The relevant theorems are in Section 3.

If H is positive semidefinite but not positive definite and has a minimum $\mathbf{x}^*$ interior to the unit sphere, then there must be a nonzero direction $\mathbf{d}$ from $\mathbf{x}^*$ such that $\mathbf{x}^* + t\mathbf{d}$ has the same value of the objective function as $\mathbf{x}^*$. This direction $\mathbf{d}$ is a eigenvector of H corresponding to the eigenvalue 0. Therefore, in this case, there is also a global minimum $\mathbf{x}^*$ satisfying $\mathbf{x}^{*T}\mathbf{x}^* = 1$.

Thus, we can focus attention on the case that $\mathbf{x}^{*T}\mathbf{x}^* = 1$. In this situation the first order conditions are as follows. There exists μ^* such that

$$\begin{aligned} H\mathbf{x}^* + \mathbf{c} &= -\mu^*\mathbf{x}^*, \\ \mathbf{x}^{*T}\mathbf{x}^* &= 1, \\ \mu^* &\geq 0. \end{aligned} \tag{2}$$

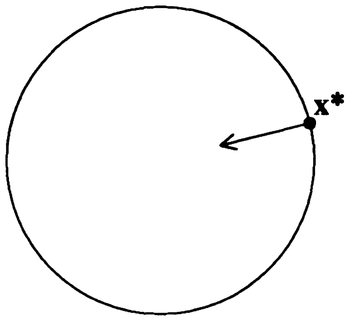

Figure 2: The KKT conditions in the $n = 2$ case.

It is a well-known fact that these conditions must hold at a local optimum (and therefore global optimum) of (1). Conditions like this are called the *Karush-Kuhn-Tucker* conditions and can be applied to general constrained nonlinear optimization (although they are not always necessary for local minimality; see Fletcher [1987]).

We do not prove the necessity of these conditions here, but we indicate briefly where they come from. The quantity $H\mathbf{x}^* + \mathbf{c}$ is seen to be the gradient of the objective function of (1) evaluated at $\mathbf{x}^*$. The first condition and third condition imply that the direction of the gradient must be exactly aligned with the inward radial direction (direct towards the center of the sphere). If this failed to hold, then there would be a small perturbation of $\mathbf{x}^*$ with a lower value of the objective function. This can be formally proved by considering directional derivatives of the objective function at $\mathbf{x}^*$. See Figure 2 for the picture in two dimensions.

We now derive second-order necessary conditions for this problem. Let $\mathbf{v}_0$ be a vector such that $\mathbf{x}^{*T}\mathbf{v}_0 < 0$. Let

$$\mathbf{v} = \mathbf{v}_0 \cdot \frac{-2\mathbf{x}^{*T}\mathbf{v}_0}{\mathbf{v}_0^T\mathbf{v}_0}.$$

Then $\mathbf{v}$ is a positive scalar multiple of $\mathbf{v}_0$ satisfying $-2\mathbf{x}^{*T}\mathbf{v} = \mathbf{v}^T\mathbf{v}$. Note that

$$\begin{aligned}(\mathbf{x}^* + \mathbf{v})^T(\mathbf{x}^* + \mathbf{v}) &= \mathbf{x}^{*T}\mathbf{x}^* + 2\mathbf{x}^{*T}\mathbf{v} + \mathbf{v}^T\mathbf{v} \\ &= 1 + 2\mathbf{x}^{*T}\mathbf{v} - 2\mathbf{x}^{*T}\mathbf{v} \\ &= 1.\end{aligned}$$

Therefore, $\mathbf{x}^* + \mathbf{v}$ is feasible for (1). Since $\mathbf{x}^*$ is supposed to be a global minimum, the value of the objective function at $\mathbf{x}^* + \mathbf{v}$ must be no less than the value at $\mathbf{x}^*$. Evaluating the objective function and $\mathbf{x}^* + \mathbf{v}$ and simplifying, we are left with the inequality

$$\mathbf{v}^T H\mathbf{x}^* + \frac{1}{2}\mathbf{v}^T H\mathbf{v} + \mathbf{c}^T\mathbf{v} \geq 0.$$

We can write the first and last terms as $(H\mathbf{x}^* + \mathbf{c})^T\mathbf{v}$ and plug in the first equation of (2) to get

$$-\mu^*\mathbf{x}^{*T}\mathbf{v} + \frac{1}{2}\mathbf{v}^T H\mathbf{v} \geq 0.$$

Since $-2\mathbf{x}^{*T}\mathbf{v} = \mathbf{v}^T\mathbf{v}$, we can substitute to get

$$\frac{1}{2}\mu^*\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{v}^T H\mathbf{v} \geq 0$$

which may be rewritten as

$$\mathbf{v}^T(H + \mu^* I)\mathbf{v} \geq 0. \tag{3}$$

This holds for $\mathbf{v}$, and hence also for $\mathbf{v}_0$. Therefore, we can conclude that (3) holds for all $\mathbf{v}$ such that $\mathbf{x}^{*T}\mathbf{v} < 0$. But then by changing the sign it holds for all $\mathbf{v}$ such that $\mathbf{x}^{*T}\mathbf{v} > 0$ as well. Finally, since the function $g(\mathbf{v}) = \mathbf{v}^T(H + \mu^* I)\mathbf{v}$ is continuous, it holds for all $\mathbf{v}$ such that $\mathbf{x}^{*T}\mathbf{v} = 0$ as well. Thus, (3) holds for all vectors $\mathbf{v}$. This is a condition that must be satisfied at a global minimum; this condition is equivalent to

$$H + \mu^* I \text{ is positive semidefinite.} \tag{4}$$

This condition implies, among other things, that the level curve of the objective function passing through $\mathbf{x}^*$ bends "outward" from the unit sphere.

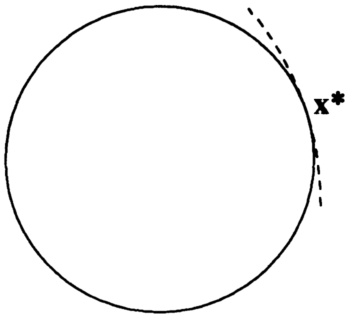

Figure 3: A level curve at the optimum point in two dimensions.

See Figure 3 for an example in two dimensions. Note that this level curve is tangent to the unit sphere because of the KKT conditions.

It is a unique feature of the sphere-constrained quadratic problem that (2) and (4) are in fact sufficient for global minimality. In particular, it is not possible for (2) and (4) to be satisfied with differing values of μ or of the objective function. This is the content of the following two lemmas.

Let H be factored as $Q^T DQ$ where Q is an $n \times n$ orthogonal matrix and D is diagonal. Note that such a factorization always exists for a symmetric real matrix; see, for example, Golub and Van Loan [1989]. We will frequently refer to Q and D in the analysis that follows; they are not needed, however, for the algorithm itself. The diagonal entries of D are denoted as d_{ii}. Let μ_0 be the negative of the smallest d_{ii}, i.e., the negative of the smallest eigenvalue of H. Then clearly $H + \mu_0 I$ is positive semidefinite, and $H + \mu I$ is positive definite if and only if $\mu > \mu_0$.

Lemma 1 *Let* $\mathbf{w}$ *be a nonzero vector, and suppose* μ_1 *and* μ_2 *are real numbers such that* $\mu_0 \leq \mu_1 < \mu_2$. *Suppose* $\mathbf{x}_1$ *is a solution to* $(H + \mu_1 I)\mathbf{x}_1 = \mathbf{w}$ *and* $\mathbf{x}_2$ *is a solution to* $(H + \mu_2 I)\mathbf{x}_2 = \mathbf{w}$. *Then*

$$\|\mathbf{x}_1\|_2 > \|\mathbf{x}_2\|_2.$$

PROOF. Since $\mathbf{w} \neq \mathbf{0}$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are both nonzero. Since $\mu_2 > \mu_0$, the matrix $H + \mu_2 I$ is positive definite and hence invertible. We compute

$$\begin{aligned} Q\mathbf{x}_2 &= Q(H+\mu_2 I)^{-1}\mathbf{w} \\ &= Q(H+\mu_2 I)^{-1}(H+\mu_1 I)\mathbf{x}_1 \\ &= (D+\mu_2 I)^{-1}(D+\mu_1 I)Q\mathbf{x}_1 \end{aligned}$$

Since Q is orthogonal, $Q\mathbf{x}_2$ has the same norm as $\mathbf{x}_2$ and a similar statement holds for $\mathbf{x}_1$. Let $\mathbf{y}_1 = Q\mathbf{x}_1$ and $\mathbf{y}_2 = Q\mathbf{x}_2$ (both nonzero), and let y_{1i} and y_{2i} denote their entries. Then $\mathbf{y}_2$ is equal to the above diagonal matrix product applied to $\mathbf{y}_1$, hence

$$y_{2i} = \frac{d_{ii}+\mu_1}{d_{ii}+\mu_2} y_{1i}.$$

For each i the above fraction is nonnegative and less than 1 since $d_{ii}+\mu_1 \geq d_{ii}+\mu_0 \geq 0$ and $\mu_2 > \mu_1$. Therefore, the 2-norm of $\mathbf{y}_2$ is strictly less than the 2-norm of $\mathbf{y}_1$. ∎

Lemma 2 *Suppose* $(\mathbf{x}_1, \mu_1)$ *and* $(\mathbf{x}_2, \mu_2)$ *are each pairs of solutions to* (2) *and* (4). *Then* $\mu_1 = \mu_2$ *and*

$$\frac{1}{2}\mathbf{x}_1^T H \mathbf{x}_1 + \mathbf{c}^T\mathbf{x}_1 = \frac{1}{2}\mathbf{x}_2^T H \mathbf{x}_2 + \mathbf{c}^T\mathbf{x}_2.$$

PROOF. We prove the first statement by contradiction. Suppose this statement is false; then, without loss of generality $\mu_1 < \mu_2$. By (4), $\mu_1 \geq \mu_0$. We notice that $\mathbf{x}_i$ for $i = 1, 2$ is a solution of $(H + \mu_i I)\mathbf{x}_i = -\mathbf{c}$. Moreover, $-\mathbf{c} \neq \mathbf{0}$ since the equation involving $\mathbf{x}_2$ is nonsingular, and we know that $\mathbf{x}_2 \neq \mathbf{0}$. Therefore, the previous lemma tells us that $\mathbf{x}_1$ and $\mathbf{x}_2$ have different 2-norms. This contradicts (2), where we required that they both had 2-norms equal to 1.

Thus, we have shown that $\mu_1 = \mu_2$. Call this number μ^*. Now we prove that the objective functions are equal with a long calculation. If we let

$f(\mathbf{x})$ denote the objective function, then

$$\begin{aligned}
f(\mathbf{x}_1) - f(\mathbf{x}_2) &= \tfrac{1}{2}\mathbf{x}_1^T H\mathbf{x}_1 + \mathbf{c}^T\mathbf{x}_1 - \tfrac{1}{2}\mathbf{x}_2^T H\mathbf{x}_2 - \mathbf{c}^T\mathbf{x}_2 \\
&= \tfrac{1}{2}\mathbf{x}_1^T H(\mathbf{x}_1 - \mathbf{x}_2) + \tfrac{1}{2}\mathbf{x}_2^T H(\mathbf{x}_1 - \mathbf{x}_2) + \mathbf{c}^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= \tfrac{1}{2}(H\mathbf{x}_1 + \mathbf{c})^T(\mathbf{x}_1 - \mathbf{x}_2) + \tfrac{1}{2}(H\mathbf{x}_2 + \mathbf{c})^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= -\mu^*\mathbf{x}_1^T(\mathbf{x}_1 - \mathbf{x}_2) - \mu^*\mathbf{x}_2^T(\mathbf{x}_1 - \mathbf{x}_2) \qquad (5) \\
&= -\mu^*(\mathbf{x}_1 + \mathbf{x}_2)^T(\mathbf{x}_1 - \mathbf{x}_2) \\
&= -\mu^*\left(\mathbf{x}_1^T\mathbf{x}_1 - \mathbf{x}_1^T\mathbf{x}_2 + \mathbf{x}_2^T\mathbf{x}_1 - \mathbf{x}_2^T\mathbf{x}_2\right) \\
&= 0.
\end{aligned}$$

∎

Therefore, the problem of solving (1) is reduced to the problem of finding μ^* and $\mathbf{x}^*$ to satisfy (2) and (4). We now prove that this pair always exists. (Actually, existence is already proved since (1) has a solution by compactness, and the two conditions are satisfied at this solution. The goal for now, however, is to come up with a more constructive proof of existence.) Define

$$p(\mu) = \|(H + \mu I)^{-1}\mathbf{c}\|_2$$

for $\mu > \mu_0$. Notice that p is a nonincreasing function (by Lemma 1) and is continuous. We want to show that it crosses through the value 1 for a $\mu \geq \max(0, \mu_0)$; this would establish the existence of a solution to (2) and (4). First, we show that it definitely achieves values smaller than 1. Define

$$m = \|H\|_\infty + \|\mathbf{c}\| + 1;$$

then one checks using standard matrix norm estimates that $\|(H+mI)^{-1}\mathbf{c}\|_2$ is less than 1.

Now, we have to show that $p(\mu)$ is bigger than 1 for μ small enough. There are several cases to consider. First suppose that $\mu_0 \geq 0$. Then we investigate what happens to $p(\mu)$ as $\mu \downarrow \mu_0$. If we let $\mathbf{c}' = Q\mathbf{c}$, then we see that $p(\mu) = \|(D + \mu I)^{-1}\mathbf{c}'\|_2$, and μ is tending to the negative of the smallest diagonal entry of D. Therefore, we conclude that $p(\mu)$ tends to $+\infty$ in this situation—and, in particular, is greater than 1—unless $\mathbf{c}'$ has zeros in all the positions in which $D + \mu_0 I$ has zeros. The condition that $c_i' = 0$ for all i such that $d_{ii} = 0$ is equivalent to the condition that $\mathbf{c}$ lies in the image of $H + \mu_0 I$. Accordingly, we have showed that $p(\mu)$ increases past 1 as long as $\mathbf{c}$ does not lie in the image of $H + \mu_0 I$.

Suppose we are in the case that $\mathbf{c}$ lies in the image of $H + \mu_0 I$. If $p(\mu)$ surpasses 1 for $\mu > \mu_0$ then we are done (we have proved it must cross through 1). Otherwise, suppose $p(\mu) \leq 1$ for all $\mu > \mu_0$. Then we can take $\mu^* = \mu_0$, and we take $\mathbf{x}_0$ to be the limit point of the sequence of vectors $-(H + \mu I)^{-1}\mathbf{c}$ as $\mu \downarrow \mu_0$. This limit point exists and lies in the unit ball because $p(\mu) \leq 1$. Moreover, it satisfies $(H + \mu_0 I)\mathbf{x}_0 = \mathbf{c}$ because the mapping $(\mathbf{x}, \mu) \mapsto (H + \mu I)\mathbf{x}$ is continuous everywhere.

As noted above, we can now get a point $\mathbf{x}^*$ satisfying $\mathbf{x}^{*T}\mathbf{x}^* = 1$ (and therefore the conditions for global minimality) by adding to $\mathbf{x}_0$ a vector $\mathbf{w}$ in the null space of $H + \mu_0 I$.

The last case is when $\mu_0 < 0$, so we consider $p(0)$. Suppose $p(0) < 1$; this means that $H^{-1}\mathbf{c}$ is contained in interior the unit ball. But then $-H^{-1}\mathbf{c}$ is the global minimum of the objective function when it is unconstrained, and hence we are in the case settled earlier of a unique interior global minimum.

Thus, we have shown in all cases that $\mathbf{x}^*$ and μ^* exist to satisfy (2) and (4). Moreover, the existence argument suggests a bisection algorithm to find μ^*. We start with μ^* known to lie in the interval $[0, m]$. We choose the midpoint of the interval, say μ_m, and perform the following tests. First, we check whether $\mu_m \geq \mu_0$. This test is equivalent to testing whether symmetric matrix $H + \mu_m I$ is positive semidefinite. This test is done via LDL^T factorization (see Golub and Van Loan [1989] as well as the discussion in the next section). If the matrix is positive semidefinite, then we know $\mu_m \geq \mu_0$.

If it turns out that $\mu_m < \mu_0$, then we know that μ_m is too small to be μ^* (because (4) would fail to be satisfied).

If it turns out that $\mu_m = \mu_0$, this case can be handled with a complicated analysis. The simplest approach, however, is to simply reject this trial and pick μ_m to be a slight perturbation of μ_0.

Finally, if it turns out that $\mu_m > \mu_0$, then we solve the linear system

$$(H + \mu_m I)\mathbf{x} = -\mathbf{c}$$

to yield $\mathbf{x}$. Then we compare $\mathbf{x}^T\mathbf{x}$ to 1. If $\mathbf{x}^T\mathbf{x} > 1$ then by Lemma 1 $\mu^* > \mu_m$. If $\mathbf{x}^T\mathbf{x} \leq 1$ then $\mu^* < \mu_m$.

Thus, for any μ_m we can determine which side of μ^* it lies on; this is the main ingredient for a binary search to compute μ^*. The bracket containing μ^* is of width $2^{-k}m$ on the kth iteration.

Let the right endpoints of the intervals generated by the binary search be denoted as $\mu^{(k)}$. The preceding analysis shows that we can get a bound on the convergence rate of $\mu^{(k)}$ to μ^*. We would now like to establish a similar result for the convergence of $\mathbf{x}^{(k)}$ to $\mathbf{x}^*$.

In this regard, define

$$\mathbf{x}^{(k)} = -(H + \mu^{(k)} I)^{-1}\mathbf{c}.$$

How fast does $\mathbf{x}^{(k)}$ converge to the optimum? We can get a bound on the objective function using a calculation similar to (5).

We first treat the case that $\mu^* > \mu_0$, so that $p(\mu^*) = 1$. We have the following calculation:

$$\begin{aligned}
f(\mathbf{x}^{(k)}) - f(\mathbf{x}^*) &= \tfrac{1}{2}\mathbf{x}^{(k)T} H\mathbf{x}^{(k)} + \mathbf{c}^T\mathbf{x}^{(k)} - \tfrac{1}{2}\mathbf{x}^{*T} H\mathbf{x}^* - \mathbf{c}^T\mathbf{x}^* \\
&= \tfrac{1}{2}\mathbf{x}^{(k)T} H(\mathbf{x}^{(k)} - \mathbf{x}^*) + \tfrac{1}{2}\mathbf{x}^{*T} H(\mathbf{x}^{(k)} - \mathbf{x}^*) + \mathbf{c}^T(\mathbf{x}^{(k)} - \mathbf{x}^*) \\
&= \tfrac{1}{2}(H\mathbf{x}^{(k)} + \mathbf{c})^T(\mathbf{x}^{(k)} - \mathbf{x}^*) + \tfrac{1}{2}(H\mathbf{x}^* + \mathbf{c})^T(\mathbf{x}^{(k)} - \mathbf{x}^*) \\
&= -\tfrac{1}{2}\mu^{(k)}\mathbf{x}^{(k)T}(\mathbf{x}^{(k)} - \mathbf{x}^*) - \tfrac{1}{2}\mu^*\mathbf{x}^{*T}(\mathbf{x}^{(k)} - \mathbf{x}^*) \\
&= \tfrac{1}{2}(\mu^* - \mu^{(k)})\mathbf{x}^{(k)T}(\mathbf{x}^{(k)} - \mathbf{x}^*) - \tfrac{1}{2}\mu^*(\mathbf{x}^{(k)} + \mathbf{x}^*)^T(\mathbf{x}^{(k)} - \mathbf{x}^*) \\
&= \tfrac{1}{2}(\mu^* - \mu^{(k)})\mathbf{x}^{(k)T}(\mathbf{x}^{(k)} - \mathbf{x}^*) - \tfrac{1}{2}\mu^*(\mathbf{x}^{(k)T}\mathbf{x}^{(k)} - 1).
\end{aligned} \tag{6}$$

Now we can bound these two terms separately. We start with the first term; the vectors $\mathbf{x}^{(k)}$ and $\mathbf{x}^*$ have norm at most 1, and the quantity $(\mu^* - \mu_k)$ is at most $2^{-k}m$, hence this term is at most $2^{-k}m$.

The second term is equal to $-\frac{1}{2}\mu^*(p(\mu^{(k)})^2 - 1)$. The number $-\mu^*$ is at most m in magnitude. We know that $p(\mu^{(k)})^2 \in [0,1]$, and moreover, $p(\mu^{(k)})^2$ tends to 1; the only question is how fast. Letting $\mathbf{c}'$ be as above, we notice that

$$Q\mathbf{x}^{(k)} = -(D + \mu^{(k)} I)^{-1}\mathbf{c}'$$

so

$$\begin{aligned}
p(\mu^{(k)})^2 &= \sum_{i=1}^{n}\left(\frac{c'_i}{d_{ii} + \mu^{(k)}}\right)^2 \\
&= \sum_{i=1}^{n}\left(\frac{c'_i}{d_{ii} + \mu^*}\right)^2\left(\frac{d_{ii} + \mu^*}{d_{ii} + \mu^{(k)}}\right)^2.
\end{aligned} \tag{7}$$

The summation on the right hand side would be 1 if the second parenthesized factor were absent. Let that factor be t_i^2, i.e.,

$$\begin{aligned} t_i &= \frac{d_{ii} + \mu^*}{d_{ii} + \mu^{(k)}} \\ &= 1 - \frac{\mu^{(k)} - \mu^*}{d_{ii} + \mu^{(k)}} \\ &\geq 1 - \frac{\mu^{(k)} - \mu^*}{\mu^{(k)} - \mu_0} \\ &\geq 1 - \frac{\mu^{(k)} - \mu^*}{\mu^* - \mu_0}. \end{aligned} \tag{8}$$

We want to prove that t_i tends to 1. Clearly $t_i \leq 1$, and the preceding inequality gives a lower bound on t_i. To make use of this inequality, we need an an upper bound for $\mu^{(k)} - \mu^*$ (which has already been established), and we also need a lower bound for $\mu^* - \mu_0$.

In order to get a lower bound on the difference $\mu^* - \mu_0$, two irrational numbers, we need to set up machinery. We now pause in the development of the material for background on systems of linear equations and on roots of polynomials. We remark that this is roughly the point in the development of the algorithms at which Ye and Karmarkar end their analysis.

3 Linear and polynomial algebra

We first start with a lemma and a theorem to show that systems of linear equations can be solved in polynomial time. These are variants of results due to Edmonds [1967].

Lemma 3 *Let A be an $n \times n$ matrix with integer entries. Suppose that L_0 bits are required to write the entries of A. Then*

$$|\det(A)| \leq n! 2^{L_0}.$$

PROOF. The determinant of A is the sum of $n!$ terms. Each term is the product of n matrix elements. The product of all the matrix elements is bounded by 2^{L_0} in magnitude. Therefore, each of the $n!$ terms that make up the determinant is bounded by 2^{L_0} in magnitude. ∎

Theorem 1 *Let A be a nonsingular integer $n \times n$ matrix and let $\mathbf{b}$ be an integer n-vector. Suppose $(A, \mathbf{b})$ requires L_0 bits to write. Then a Turing machine can compute a solution to $A\mathbf{x} = \mathbf{b}$ in time bounded by a polynomial in n, L_0.*

PROOF. The Turing machine can execute a standard algorithm for solving linear systems, say Gaussian elimination with partial pivoting (see Golub and Van Loan [1989]), using exact rational arithmetic (recall that Gaussian elimination requires only rational operations). The number of arithmetic operations required by Gaussian elimination is $O(n^3)$. The Turing machine can carry out each arithmetic operation ($+$, $*$, $-$, $/$) in time that is polynomial in the length of the operands.

Accordingly, the only remaining hurdle is to show that every rational number that occurs during Gaussian elimination has length that is bounded by a polynomial in n, L_0. This follows because every rational number that occurs during Gaussian elimination is in fact a ratio of two subdeterminants, where a *subdeterminant* refers to the determinant of a square submatrix of $(A, \mathbf{b})$.

In particular, Gaussian elimination with partial pivoting computes a factorization $PA = LU$, where P is a permutation matrix, L is a unit lower triangular matrix, and U is an upper triangular matrix. Let us consider u_{kk}. If we let A_1 denote the upper $k \times k$ submatrix of PA, L_1 the upper left $k \times k$ submatrix of L, and U_1 the upper $k \times k$ submatrix of U, then $A_1 = L_1 U_1$, so $\det(A_1) = \det(L_1)\det(U_1)$. Since L_1 is unit lower triangular, $\det(L_1) = 1$. Since U_1 upper triangular, $\det(A_1) = \det(U_1) = u_{11} \cdots u_{kk}$. An analogous formula can be derived for the determinant of the upper left $(k-1) \times (k-1)$ submatrix of PA, say A_0. Thus, we conclude that $u_{kk} = \det(A_1)/\det(A_0)$. By considering additional permutation matrices, a similar argument holds for the rest of U. By factoring a diagonal matrix out of U, we can argue symmetrically about L.

Thus, we conclude that every number encountered during Gaussian elimination is a subdeterminant and therefore requires at most $L_0 + n \lg n$ bits to write according to the preceding lemma. This length is bounded by a polynomial in n and L_0. Note that the rational numbers encountered by the Turing machine must be continually reduced to lowest terms for this bound to apply; reducing a fraction to lowest terms also requires polynomial time by Turing machine. ∎

The preceding proof can be used to derive bounds on the size of the rational numbers occurring in the solution $\mathbf{x}$ to the system $A\mathbf{x} = \mathbf{b}$. Better bounds may be obtained by an argument based on Cramer's rule. We introduce the following notation. If A is an $n \times n$ matrix and $\mathbf{b}$ a vector, let $\langle A, \mathbf{b} \rangle_i$ denote the $n \times n$ matrix resulting when $\mathbf{b}$ is substituted for the ith column of A. Then *Cramer's rule* says that if A is nonsingular and $A\mathbf{x} = \mathbf{b}$, then

$$x_i = \det\langle A, \mathbf{b} \rangle_i / \det(A).$$

Cramer's rule gives the following bound. The proof, which follows directly from Lemma 3, is omitted.

Lemma 4 *Let A be an $n \times n$ nonsingular integer matrix and $\mathbf{b}$ an integer n-vector. Suppose that L_0 bits are required to write down the entries of $(A, \mathbf{b})$. Then the solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$ is a vector with rational entries. Moreover, these rational entries have a common denominator d less than $2^{L_0 + n \lg n}$ in magnitude, and the numerators with respect to d are each less than $2^{L_0 + n \lg n}$ in magnitude.*

Next we turn to polynomials. For the remainder of this section, the emphasis is on deriving bounds rather than on proposing algorithms. Most of the existence statements, however, can be turned into constructive algorithms. All of the material of this section is explained in greater depth in Zippel [1990].

The set of all polynomials with indeterminate x and coefficients chosen from a ring R is denoted $R[x]$. In many of our examples R will be one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. For results that apply to polynomials over a general field, we write $k[x]$ for polynomials over field k. We write $k(x)$ for rational functions over field k. A polynomial $f(x)$ is said to be *irreducible* over $R[x]$ if it cannot be written as a product of polynomials of lower degree. The following facts are proved in Jacobson [1974]. First, any polynomial $f(x) \in R[x]$ can be factored as a product of irreducible polynomials for $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Moreover, this factorization is unique up to constant multiples (where the constant lies in R).

As an application of this theory, suppose $f(x)$ and $g(x)$ are nonzero polynomials over $\mathbb{Q}$ or $\mathbb{R}$ of degree d, e respectively, and suppose there exist polynomials $r(x), s(x)$ such that $r(x)f(x) = s(x)g(x)$, and suppose the degree of $r(x)$ is k such that $k < e$. Then we can conclude that $f(x)$

and $g(x)$ have a common factor. To prove this, let $c_1(x), \ldots, c_i(x)$ be a factorization of $r(x)$ into irreducible polynomials, and let $c_{i+1}(x), \ldots, c_n(x)$ be a factorization of $f(x)$ into irreducibles. Then the sum of degrees of $c_1(x), \ldots, c_i(x)$ must be k and the sum of the degrees of $c_{i+1}(x), \ldots, c_n(x)$ must be d. This is also the (unique) factorization of $s(x)g(x)$, hence $g(x) = \alpha c_{j_1}(x) \cdots c_{j_p}(x)$, where α lies in $\mathbb{Q}$ or $\mathbb{R}$ (depending which field we started with) and $j_1, \ldots, j_p$ is some subsequence of $1, \ldots, n$. Now, suppose that $\{j_1, \ldots, j_p\} \subset \{1, \ldots, i\}$; this would mean that the degree of g is at most the degree of $c_1(x) \cdots c_i(x) = r(x)$, i.e., the degree is at most k. This contradicts our assumption that $k < e$. Thus, for some index k, $j_k \geq i$. We conclude that $c_{j_k}(x)$ is a common factor to f and g.

Another well known fact about polynomials is that any polynomial $p(x) \in \mathbb{C}[x]$ (and, in particular, any polynomial in $\mathbb{Z}[x]$) can be factored uniquely as $a(x - \alpha_1) \cdots (x - \alpha_d)$ where d is the degree of the polynomial and $\alpha_1, \ldots, \alpha_d$ are the roots of p lying in $\mathbb{C}$.

Next, we prove a lemma taken from Mignotte [1982]. The theorem is due to Cauchy [1829]. Mignotte goes on to prove stronger bounds than Theorem 3 for root separation, but more advanced techniques are needed.

Lemma 5 *Let $p(x) \in \mathbb{R}[x]$ be the nonzero polynomial*

$$f(x) = c_d x^d + \cdots + c_1 x + c_0.$$

Let $\alpha \in \mathbb{C}$ be a root of f. Then

$$|\alpha| \geq \frac{|c_0|}{|c_0| + \max(|c_1|, \ldots, |c_d|)}$$

PROOF. If $|\alpha| \geq 1$ then the lemma is trivially true. For the other case, let $m = \max(|c_1|, \ldots, |c_d|)$. By hypothesis α satisfies $c_0 = -c_d\alpha^d - \cdots - c_1\alpha$, so

$$\begin{aligned} |c_0| &\leq m(|\alpha|^d + \cdots + |\alpha|) \\ &\leq m \cdot |\alpha| \cdot (1 + |\alpha| + \cdots + |\alpha|^{d-1}) \\ &\leq \frac{m \cdot |\alpha|}{1 - |\alpha|}. \end{aligned}$$

Noting that $1 - |\alpha| > 0$ by assumption, we can simplify this to get $|\alpha| \geq |c_0|/(|c_0| + m)$, proving the lemma. ∎

The next definition needed for the development of our subject is the resultant.

DEFINITION. *Let k be a field, and let $p(x), q(x)$ be polynomials of degree d, e in $k[x]$. Suppose that*

$$p(x) = a_d x^d + \cdots + a_0$$

and

$$q(x) = b_e x^e + \cdots + b_0.$$

Then the **resultant** *of $p(x), q(x)$, denoted by $\mathrm{res}_x(p, q)$, is the determinant of the following matrix*

$$\left[\begin{array}{ccccc|ccccc}
a_d & 0 & \cdots & \cdots & 0 & b_e & 0 & \cdots & \cdots & 0 \\
a_{d-1} & a_d & \ddots & & \vdots & b_{e-1} & b_e & \ddots & & \vdots \\
\vdots & a_{d-1} & \ddots & 0 & \vdots & \vdots & b_{e-1} & \ddots & 0 & \vdots \\
a_0 & \vdots & \ddots & a_d & 0 & b_0 & \vdots & \ddots & b_e & 0 \\
0 & a_0 & & a_{d-1} & a_d & 0 & b_0 & & b_{e-1} & b_e \\
\vdots & 0 & \ddots & \vdots & a_{d-1} & \vdots & 0 & \ddots & \vdots & b_{e-1} \\
\vdots & & \ddots & a_0 & \vdots & \vdots & & \ddots & b_0 & \vdots \\
0 & \cdots & \cdots & 0 & a_0 & 0 & \cdots & \cdots & 0 & b_0
\end{array}\right].$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{e\ \text{columns}} \qquad \underbrace{\qquad\qquad\qquad\qquad}_{d\ \text{columns}}$$

This matrix is known as the **Sylvester matrix** *for p and q.*

We now prove a lemma giving the main property of resultants. This lemma is taken from Jacobson. Although we use resultants only to derive bounds, resultants have been used to construct algorithms. See, for example, Netto [1908] for some old uses, or Ierardi and Kozen [1990] for a very recent use.

Lemma 6 *If $p(x)$ and $q(x)$ are two polynomials over k, then $\mathrm{res}_x(p, q) = 0$ iff p and q have a common factor.*

PROOF. Let A be the Sylvester matrix in the previous definition. We claim this matrix is singular iff p and q have a common factor. Let the common factor be $c(x)$. Write $p(x) = p_1(x)c(x)$ and $q(x) = q_1(x)c(x)$. Suppose the degree of $c(x)$ is k (where $1 \leq k \leq \min(d, e)$). Write the coefficients of $q_1(x)$ as an $e-k+1$ vector and concatenate $k-1$ zeros. Call this e-vector $\mathbf{x}$. Now write the coefficients of $-p_1(x)$ as a $d-k+1$ vector and concatenate $k-1$ zeros. Call this d-vector $\mathbf{y}$. Now, finally, let $\mathbf{w}$ be the $d+e$ vector resulting from the concatenation of $\mathbf{x}$ and $\mathbf{y}$. We observe that $A\mathbf{w} = \mathbf{0}$. In particular, the ith entry of $A\mathbf{w}$ is the ith coefficient of $x^{k-1}(q_1(x)p(x) - p_1(x)q(x))$ which is the zero polynomial. Thus, the matrix has a nontrivial null space so it is singular.

Conversely, suppose that matrix A is singular. Let $\mathbf{w}$ be a nontrivial null vector. Then if we split $\mathbf{w}$ into two subvectors, one with e entries and the other with d, these correspond to two polynomials $r(x)$ and $s(x)$ such that $r(x)p(x) + s(x)q(x) = 0$ (identically). Moreover, the degree of r is at most $e-1$ and the degree of s is at most $d-1$, hence $p(x)$ and $q(x)$ have a least common multiple (namely $r(x)p(x) = -s(x)q(x)$) of degree at most $d+e-1$. ∎

With this lemma we can establish a definition for resultant in terms of roots.

Lemma 7 *Suppose that $p(x), q(x)$ are polynomials of degree d, e, respectively over a field k such that all the roots of p, q lie in k. Denote the leading coefficients of p, q as a_d, b_e respectively. Denote the roots of p as $\alpha_1, \ldots \alpha_d$ and of q as $\beta_1, \ldots, \beta_e$. Then*

$$\mathrm{res}_x(p,q) = b_e^d p(\beta_1)\cdots p(\beta_e) = (-1)^{de} a_d^e q(\alpha_1)\cdots q(\alpha_d).$$

PROOF. Notice that both formulas in the statement of the lemma are equal to the following expression:

$$a_d^e b_e^d \prod_{i=1}^{d} \prod_{j=1}^{e} (\beta_j - \alpha_i).$$

Thus, we need to prove only that $\mathrm{res}_x(p(x), q(x))$ is equal to this expression. Let y be another indeterminant, and consider the polynomial $q(x+y) \in k[x,y]$. We can think of $q(x+y)$ as lying in $k(y)[x]$, and similarly for $p(x)$. The coefficient of x^i in $q(x+y)$ in $k(y)[x]$ is a degree $e-i$ polynomial in y. (This fact and others mentioned in this proof follow from writing $q(x+y)$ using the binomial formula.) Then we can consider $\mathrm{res}_x(p(x), q(x+$

$y)$), which will be an element of the base field $k(y)$. However, since the determinant operation requires only ring operations, $\mathrm{res}_x(p(x), q(x + y))$ actually lies in $k[y]$. Let us call this polynomial $h(y)$. Then a simple degree-counting argument shows that $h(y)$ has degree de. Now, observe that $\mathrm{res}_x(p(x), q(x)) = h(0)$ because substituting 0 for y is an operation preserving the definition of resultant (also, see Lemma 8 below).

Thus, $\mathrm{res}_x(p(x), q(x))$ is the constant coefficient of $h(y)$, i.e., the product of the de roots of $h(y)$ multiplied by the leading coefficient of $h(y)$. It is easy to verify that the leading term of $h(y)$ is $a_d^e b_e^d$. Finally, for all i, j, $\beta_j - \alpha_i$ is a root to $h(y)$ since $p(x)$ and $q(x + \beta_j - \alpha_i)$ have α_i as a common root (and hence the resultant implicit in $h(y)$ is zero as in the previous lemma). ∎

For the next theorem, assume that e is at least $3 \log d$. We introduce new notation: if $p(x)$ is a polynomial, then $\|p\|$ denotes the maximum absolute value among coefficients of p. This is only defined in the case that the field k has a notion of absolute value satisfying the usual axioms. For example, we could have $k = \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$. These results could be generalized to the case that k is a ring with a valuation (see Lang [1971]).

Theorem 2 *Let $p(x)$ and $q(x)$ be polynomials of degree d, e respectively over a field k. Let $h(y)$ be defined as in the previous proof, i.e.,*

$$h(y) = \mathrm{res}_x(p(x), q(x + y)).$$

Then $h(y)$ is a polynomial function of y of degree de, and

$$\|h\| \leq 2^{2de} \|p\|^e \cdot \|q\|^d.$$

PROOF. We explained in the previous proof that $h(y)$ is a polynomial of degree de.

We next compute the largest coefficient of $h(y)$. Since $h(y)$ is a determinant of a $(d + e) \times (d + e)$ matrix, it is a sum of $(d + e)!$ terms; focus on one of these terms. Such a term will be a product of e coefficients of p multiplied by d polynomials that are coefficients of x in $q(x + y)$. Write $q(x + y) = c_e(y)x^e + \cdots + c_0(y)$. An upper bound on the product of e coefficients of p is $\|p\|^e$. Let us consider the coefficient of y^k in this product of d polynomials. The number of terms contributing to this coefficient is the number of ways to partition the integer k into d nonnegative integers; we can put an upper bound of d^d on the number of terms. Each term is the product of at most k coefficients chosen from c_i. Notice that $||c_i|| \leq ||q|| \cdot 2^e$, as we can see from the binomial formula. Thus, the term is at most $||q||^d 2^{de}$; the number of terms is d^d contributing to y^k, and the number of terms contributing to the determinant is $(d+e)!$. This gives an upper bound of $(d+e)!d^d 2^{de}||q||^e \cdot ||p||^d$. We can estimate the first three factors as 2^{2de} provided that e is larger than $3 \log d$. This gives the bound stated in the theorem. ∎

Now we state the main result on the separation of polynomial roots.

Theorem 3 *Let α be a root of a nonzero polynomial $p(x) \in \mathbb{Z}[x]$, and let β be a root of a nonzero polynomial $q(x) \in \mathbb{Z}[x]$. Suppose $\alpha \neq \beta$, and suppose $p(x), q(x)$ have degree d, e respectively. Then*

$$|\alpha - \beta| \geq 2^{-2de-1}||p||^{-e} \cdot ||q||^{-d}.$$

PROOF. Apply the previous theorem with $k = \mathbb{Q}$ to get a polynomial $h(y)$ of degree de that has $\beta - \alpha$ as root and has coefficients at most $2^{2de}||p||^e \cdot ||q||^d$. (The fact that $h(y)$ has $\beta - \alpha$ as a root is comes from the proof of Lemma 7.) Note that $h(y) \in \mathbb{Z}[y]$ since, as mentioned earlier, taking determinants does not involve division. Divide this polynomial by factors of y until the the constant term becomes nonzero; call this new polynomial h_1. Polynomial h_1 still has $\beta - \alpha$ as a root and satisfies $||h_1|| = ||h||$.

Now we apply 5 to conclude that all roots of polynomial $h_1(y)$ have magnitude at least $a/(a + ||h_1||)$, where a is the magnitude of the constant coefficient of h_1. We know that $a \geq 1$ since h_1 has integer coefficients, and $a, ||h_1|| \leq 2^{2de}||p||^e \cdot ||q||^d$. This gives a lower bound on $|\beta - \alpha|$, proving the theorem. ∎

For the upcoming theorems we need a lemma. We have already used the simple case of this lemma implicitly in earlier proofs.

Lemma 8 *Let $p(x)$ be a polynomial of degree d in $k[x]$, and let $q(x, y)$ be a polynomial of degree e with respect to x in $k[x, y]$. Let $r(y)$ denote $\mathrm{res}_x(p(x), q(x, y))$; hence $r(y) \in k[y]$. Suppose $a \in k$. Then $r(a) = 0$ if and only if $q(x, a)$ and $p(x)$ have a common factor.*

PROOF. This lemma is trivial if $q(x, a)$ and $q(x, y)$ have the same degree when regarded as polynomials of x. It follows in this case because substitution of a for y is a homomorphism from $k[y]$ to k and hence $r(a)$ is equal to $\text{res}_x(p(x), q(x, a))$ in this case. As proved earlier, this is zero if and only if $p(x)$ and $q(x, a)$ have a common root.

The more complicated case is when $q(x, a)$ has a lower degree than $q(x, y)$ (this happens when a is a root of the coefficient of leading term of x in $q(x, y)$). In this case the equation $r(a) = \text{res}_x(p(x), q(x, a))$ fails to hold. To prove the theorem, we need to redo the proof of Lemma 6 more carefully. Let e' be the actual degree of $q(x, a)$, so that $e' \leq e$.

In this case, we know that $r(a) = 0$ iff the Sylvester matrix of $p(x)$ and $q(x, a)$ is singular, where $q(x, a)$ is treated as having degree e (i.e., leading coefficients are zeros) for the purpose of writing the Sylvester matrix. Let A denote this $(d + e) \times (d + e)$ Sylvester matrix.

First, suppose that $p(x)$ and $q(x, a)$ have a common factor. Let the common factor be $c(x)$. Write $p(x) = p_1(x)c(x)$ and $q(x) = q_1(x)c(x)$. Suppose the degree of $c(x)$ is k (where $1 \leq k \leq \min(d, e') \leq \min(d, e)$). Write the coefficients of $q_1(x)$ as an $e - k + 1$ vector and concatenate $k - 1$ zeros. Call this e-vector $\mathbf{x}$. Now write the coefficients of $-p_1(x)$ as a $d-k+1$ vector and concatenate $k-1$ zeros. Call this d-vector $\mathbf{y}$. Now, finally, let $\mathbf{w}$ be the $d+e$ vector resulting from the concatenation of $\mathbf{x}$ and $\mathbf{y}$. We observe that $A\mathbf{w} = \mathbf{0}$. In particular, the ith entry of $A\mathbf{w}$ is the ith coefficient of $x^{k-1}(q_1(x)p(x) - p_1(x)q(x))$ which is the zero polynomial. Thus, the matrix has a nontrivial null space so it is singular.

Conversely, suppose that matrix A is singular. Let $\mathbf{w}$ be a nontrivial null vector. Then if we split $\mathbf{w}$ into two subvectors, one with e entries and the other with d, these vectors make up the coefficients of two polynomials $r(x)$ and $s(x)$ such that $r(x)p(x) + s(x)q(x, a) = 0$ (identically). Moreover, the degree of r is at most $e - 1$ and the degree of s is at most $d - 1$. This means that the degree of $s(x)q(x, a)$ is at most $d-1+e'$, hence $p(x)$ and $q(x)$ have a least common multiple (namely $r(x)p(x) = -s(x)q(x)$) of degree at most $d + e' - 1$. ∎

It will become necessary below to operate on roots of polynomials. We would like to prove that rational operations on roots of polynomials yield other roots of polynomials.

Theorem 4 *Let α be a root to a nonzero polynomial $p(x)$. Suppose $\beta = q(\alpha)/r(\alpha)$, where $q(x), r(x)$ are nonzero polynomials and α is not a root of r. Suppose $p(x), q(x), r(x) \in \mathbb{Z}[x]$. Suppose $q(x)$ and $r(x)$ have degree at most d and $p(x)$ has degree at most e. Then β is the root to a polynomial w in $\mathbb{Z}[x]$ of degree at most e, and*

$$\|w\| \leq (d+e)!(3\|p\|)^e \max(\|q\|, \|r\|)^d.$$

PROOF. Let $\beta(z)$ be the ratio $(q(\alpha)+z)/r(\alpha)$. Then the quantity we are after is $\beta(0)$. Note that $\beta(z)$ depends continuously on z.

Next, define $v(y,z) = \mathrm{res}_x(p(x), yr(x) - q(x) - z)$. Notice that $v(y,z)$ is a polynomial in both y and z as discussed earlier. Notice also that $v(\beta(z), z) = 0$; this is because polynomials $\beta(z)r(x) - q(x) - z$ and $p(x)$ have α as a common root—this follows from the previous lemma.

We can show that $v(y,z)$ is a nonzero polynomial by exhibiting z_1 so that $v(0, z_1) \neq 0$. Pick z_1 so that $q(x) + z_1$ has no roots in common with $p(x)$. Clearly this is possible since choosing z_1 large enough can force the roots of $q(x) + z_1$ arbitrarily far away from the origin. The absence of common roots means $v(0, z_1) \neq 0$.

Notice also that $v(y,z)$ has integer coefficients because the p, q, r have integer coefficients.

Next, let z^k be the highest power of the indeterminate z that divides each term of $v(y,z)$. Then we can write $v(y,z) = z^k w(y,z)$ where $w(y,z)$ is another polynomial of two variables. Notice that for $z \neq 0$, $w(\beta(z), z) = 0$ because the same result holds for polynomial v. Since $\beta(z)$ depends continuously on z, we conclude $w(\beta(0), 0) = 0$. Thus, $w(y, 0)$ is a polynomial in y that has $\beta(0)$ as a root.

Now we investigate the polynomial $w(y, 0)$. First notice that this polynomial is nonzero by the following argument. Polynomial $v(y,z)$ was nonzero; by factoring out the highest power of z, we ensured that there is at least one term of in $w(y,z)$ that does not depend on z—this term is either constant or depends only on y. In either case, the presence of this term implies that $w(y, 0)$ is a nonzero polynomial.

Next, we look at the degree of $w(y, 0)$. Notice that the highest power of y in $v(y,z)$ is e because entries in the first e columns of the matrix with which we compute the resultant depend linearly on y. This means that the highest power of y in $w(y,0)$ is also at most e.

Next we must analyze the size of the maximum coefficient in $w(y,0)$. The largest coefficient in $w(y,0)$ is no larger than the largest coefficient in $v(y,z)$; this is because the removal of the powers of z does not change coefficients, and the setting of z to zero also cannot cause coefficient growth.

Accordingly, let us consider a term of $v(y,z)$. The determinant used to compute this resultant is the sum of $(d+e)!$ terms; focus on one term. This term is the product of e matrix entries that are polynomials in y and z (in particular, a polynomial of the form $c_1y - c_2z - c_3$ where c_1 is a coefficient from q, c_2 is 0 or 1, and c_3 is a coefficient from r). Multiplying out these polynomials yields at most 3^e distinct terms involving y and z. The coefficient preceding this single term is at most $\|p\|^e \max(\|q\|,\|r\|)^d$; the second factor comes from coefficients of q and r in the resultant matrix, and the first factor comes from coefficients of p in the matrix.

Subsets of these terms combine to give the terms of y^iz^j in $v(y,z)$. We can, however, get an upper bound on any one coefficient by multiplying all the bounds together in the previous paragraph; we conclude that no coefficient can be larger than the bound stated in the theorem.

We remark that this proof could be greatly simplified if we assumed that q, r have no common roots with p. In this case, we could started with $z = 0$ in the proof. ∎

4 Convergence rate of the objective function

We return to the development of the algorithm for the sphere constrained problem. We left off requiring a lower bound for $\mu^* - \mu_0$. (Recall that for now we are considering the case that $\mu^* > \mu_0$.) Assume that H, $\mathbf{c}$ have integer entries. Assume that H, $\mathbf{c}$, and ζ require a total of L bits to write. (Our results up to now have assumed merely that H and $\mathbf{c}$ have real entries.)

Notice that μ_0 is a solution to $\det(H + \mu I) = 0$, which is a degree-n polynomial equation in μ such that $\|\det(H+\mu I)\| \le n!2^{n+L}$. (Here, $\|\cdot\|$ refers to the polynomial norm introduced in the previous section.) Call this bound s.

Similarly, μ^* is a solution to a polynomial equation. Specifically, we apply Cramer's rule to the equation $(H + \mu^* I)\mathbf{x}^* = -\mathbf{c}$, so that

$$x_i^* = -\frac{\det\langle H + \mu^* I, \mathbf{c}\rangle_i}{\det(H + \mu^* I)} \tag{9}$$

(recall that $H + \mu^* I$ is nonsingular since we are treating the case that $\mu^* > \mu_0$). Then μ^* is a root to

$$\sum_{i=1}^{n} \left(\frac{\det\langle H + \mu I, \mathbf{c}\rangle_i}{\det(H + \mu I)} \right)^2 = 1.$$

Multiplying both sides by $\det(H + \mu I)^2$ gives a polynomial of degree $2n$ in μ. This norm of this polynomial is at most $n(n!)^2 2^{2n+2L}$; call this bound t.

Now we apply Theorem 3 to conclude $\mu^* - \mu_0$ is at least $2^{-4n^2-1}s^{-2n}t^{-n}$, which is at least 2^{-q} where $q = O(n^2 \log n + nL)$. Since $\mu^{(k)} - \mu^*$ at most $2^{-k}m$, a lower bound on t_i in (8) is $1 - 2^{q-k}m$, so a lower bound on t_i^2 would be $1 - 2^{q-k+1}m$. This is therefore a lower bound on (7). Thus, the second term of (6) differs from zero by at most $2^{q-k+1}m^2$.

Accordingly the value of the objective function converges linearly, and the kth iteration is at most $2^{q'-k}$ above optimal, where $q' = O(n^2 \log n + nL)$.

So far we have been looking at the case that $\mu^* > \mu_0$. We next consider the more complicated case that $\mu^* = \mu_0$. This case can be detected after $O(n^2 \log n + nL)$ iterations, because $\mu^{(k)}$ will come too close to μ_0 in this situation (i.e., closer than the distance permitted by Theorem 3).

Recall that if $\mu^* = \mu_0$ then $\mathbf{c}$ lies in the image of the singular matrix $H + \mu_0 I$. We consider how the sequence $\mathbf{x}^{(k)}$ defined above behaves in this situation. By examining the transformed vector $Q\mathbf{x}^{(k)}$ we can conclude that $\mathbf{x}^{(k)}$ must converge to a solution of $(H + \mu_0 I)\mathbf{x} = -\mathbf{c}$ that satisfies $\mathbf{x}^T\mathbf{x} \le 1$.

To solve (1) we must find a vector in the null space of $H + \mu_0 I$; this must be accomplished without knowing μ_0 exactly. We now show how to compute a vector $\mathbf{y}$ that is an approximation a null-space vector.

Let $\mathbf{b}$ be a unit vector *not* in the image of $H + \mu_0 I$. To come up with such a vector $\mathbf{b}$, let $\mathbf{q}$ be denote the ith column of Q^T, where i is an index such that $d_{ii} = -\mu_0$. Note that this $\mathbf{q}$ is an eigenvector of H. Suppose $\mathbf{q}^T\mathbf{b} \ne 0$. Then $(H + \mu_0 I)\mathbf{x} = \mathbf{b}$ has no solution (multiply both sides by $\mathbf{q}^T$; the left side is zero). Thus, $\mathbf{b}$ is not in the image of $H + \mu_0 I$.

Accordingly, we need a unit vector $\mathbf{b}$ not orthogonal to $\mathbf{q}$. We have not computed $\mathbf{q}$, but we know that among $\mathbf{e}_1, \ldots, \mathbf{e}_n$ (the columns of the identity matrix), at least one of these vectors will not be orthogonal to $\mathbf{q}$. Moreover, at least one of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ will have an inner product of absolute value at least $1/\sqrt{n}$ with $\mathbf{q}$. Therefore, in the material that follows, we can assume that each $\mathbf{e}_i$ is substituted for $\mathbf{b}$; at least one will work. Indeed, a randomly chosen unit-length vector $\mathbf{b}$ will probably do. The expected value of $|\mathbf{q}^T\mathbf{b}|$ is about $2/(\pi\sqrt{n})$ for large n for $\mathbf{b}$ chosen uniformly at random (see Kuczyński and Woźniakowski [1990] for results in this direction).

Thus, assume we have a unit-length vector $\mathbf{b}$ such that $|\mathbf{q}^T\mathbf{b}| \geq 1/\sqrt{n}$.

Next, we compute the solution to the system $(H + \mu_1 I)\mathbf{y} = \mathbf{b}$, where μ_1 is greater than μ_0 but very close to it (see below). If we let $\mathbf{z} = Q\mathbf{y}$ and $\mathbf{b}' = Q\mathbf{b}$ then we see that $(D + \mu_1 I)\mathbf{z} = \mathbf{b}'$, i.e.,

$$z_j = \frac{b'_j}{d_{jj} + \mu_1}.$$

If we let i be the index of the distinguished column of Q^T selected above, then we see that b'_i is exactly equal to $\mathbf{q}^T\mathbf{b}$. Thus, $z_i = b'_i/(\mu_1 - \mu_0) \geq 1/[\sqrt{n}(\mu_1 - \mu_0)]$. Thus, $\mathbf{z}$ has 2-norm at least $1/[\sqrt{n}(\mu_1 - \mu_0)]$. This means that $\mathbf{y}$ has the same lower bound on its 2-norm. Thus, $\mathbf{y}$ will be a large vector since $\mu_1 - \mu_0$ is small.

Next, assume $\mathbf{x}$ is the solution to $(H + \mu_1 I)\mathbf{x} = -\mathbf{c}$. As noted above, $\mathbf{x}$ lies within the unit sphere. Then compute a vector $\mathbf{w} = \mathbf{x} + \alpha\mathbf{y}$ such that $\mathbf{w}^T\mathbf{w} = 1 - \epsilon$, where $\epsilon \geq 0$ is small. Note that this is possible because $\mathbf{x}^T\mathbf{x} \leq 1$ and $\mathbf{y} \neq \mathbf{0}$. The scalar α is computed as a rational approximate solution to the one-variable quadratic equation $(\mathbf{x}+\alpha\mathbf{y})^T(\mathbf{x}+\alpha\mathbf{y}) = 1$. (We would ideally want the exact solution, but the exact solution is probably irrational.) Note that by the triangle inequality, $\|\alpha\mathbf{y}\| \leq 2 - \epsilon$, hence $|\alpha| \leq 2/\|\mathbf{y}\|$. This implies that

$$|\alpha| \leq 2(\mu_1 - \mu_0)\sqrt{n}.$$

Note that $H\mathbf{w} + \mu_1\mathbf{w} = -\mathbf{c} + \alpha\mathbf{b}$ by choice of $\mathbf{w}$.

The claim is that $\mathbf{w}$ is an approximation to $\mathbf{x}^*$ because it is the sum of a "near" solution to $(H + \mu_0 I)\mathbf{x} = -\mathbf{c}$ and a vector $\alpha\mathbf{y}$ "nearly" in the null space of $H + \mu_0 I$. Now we do a calculation to determine how far $\mathbf{w}$ is from the optimal $\mathbf{x}^*$ in terms of the objective function. The second line is derived from the first using the same intermediate steps as (6).

$$
\begin{aligned}
f(\mathbf{w}) - f(\mathbf{x}^*) &= \frac{1}{2}\mathbf{w}^T H\mathbf{w} + \mathbf{c}^T\mathbf{w} - \frac{1}{2}\mathbf{x}^{*T} H\mathbf{x}^* - \mathbf{c}^T\mathbf{x}^* \\
&= \frac{1}{2}(H\mathbf{w} + \mathbf{c})^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}(H\mathbf{x}^* + \mathbf{c})^T(\mathbf{w} - \mathbf{x}^*) \\
&= \frac{1}{2}(H\mathbf{w} + \mathbf{c} - \alpha\mathbf{b})^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*) - \frac{1}{2}\mu_0\mathbf{x}^{*T}(\mathbf{w} - \mathbf{x}^*) \\
&= -\frac{1}{2}\mu_1\mathbf{w}^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*) - \frac{1}{2}\mu_0\mathbf{x}^{*T}(\mathbf{w} - \mathbf{x}^*) \\
&= -\frac{1}{2}(\mu_1\mathbf{w} + \mu_0\mathbf{x}^*)^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*) \\
&= -\frac{1}{2}\mu_0(\mathbf{w} + \mathbf{x}^*)^T(\mathbf{w} - \mathbf{x}^*) - \frac{1}{2}(\mu_1 - \mu_0)\mathbf{w}^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*) \\
&= \frac{1}{2}\mu_0\epsilon - \frac{1}{2}(\mu_1 - \mu_0)\mathbf{w}^T(\mathbf{w} - \mathbf{x}^*) + \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*) \\
&= T_1 + T_2 + T_3
\end{aligned}
$$

where

$$
\begin{aligned}
T_1 &= \frac{1}{2}\mu_0\epsilon, \\
T_2 &= -\frac{1}{2}(\mu_1 - \mu_0)\mathbf{w}^T(\mathbf{w} - \mathbf{x}^*), \\
T_3 &= \frac{1}{2}\alpha\mathbf{b}^T(\mathbf{w} - \mathbf{x}^*).
\end{aligned}
$$

Now, each of these terms can be made as small as we want; μ_1 is within $2^{-k}m$ of μ_0, ϵ was chosen above and can be arbitrarily small, and α is also bounded by a multiple of $\mu_1 - \mu_0$.

5 Polynomial time

In the previous section we concluded in k iterations we can compute a feasible vector $\mathbf{x}$ for (1) such that $f(\mathbf{x}) - f(\mathbf{x}^*) \leq 2^{q-k}$, where q is polynomial in the size of the input. We want to show how to use this to solve the decision problem QPSPHERE. In particular, let $\theta = f(\mathbf{x}^*)$. Then we want show that either $\theta = \zeta$ or else $|\theta - \zeta|$ has a lower bound that can be determined after a polynomial number of operations.

First, notice that ζ is the root to a polynomial with integer coefficients. In particular, since ζ is rational, it is the root to a polynomial of the form $ax - b = 0$ where a, b are integers with fewer than L bits.

Next we consider θ. As usual, we take two cases. First, we consider the simple case that $\mu^* > \mu_0$. In this case, we have already given a formula for x_i^* in terms of μ^*, namely (9). This means that

$$\begin{aligned}\theta &= \frac{1}{2}\mathbf{x}^{*T}H\mathbf{x}^* + \mathbf{c}^T\mathbf{x}^* \\ &= \frac{1}{2}\sum_{i,j=1}^{n} x_i{}^* x_j{}^* h_{ij} + \sum_{i=1}^{n} c_i x_i^* \\ &= \frac{q(\mu^*)}{2r(\mu^*)^2}\end{aligned}$$

where $r(x) = \det(H + xI)$ and $q(x)$ is a sum of polynomials of the form $h_{ij}\det\langle H+xI, \mathbf{c}\rangle_i \cdot \det\langle H+xI, \mathbf{c}, \rangle_j$ or of the the form $2c_i \det\langle H+xI, \mathbf{c}\rangle_i \cdot r(x)$.

Thus, θ is exactly the quotient of two polynomial functions of μ^*. This means that θ itself is the root of a polynomial with bounds given by Theorem 4. In particular, μ^* is the root to a polynomial of degree $2n$ and maximum coefficient $n(n!)^2 2^{2n+2L}$. We can express θ as a rational function $q(\mu^*)/(2r(\mu^*)^2)$, where the degree of $q(x)$ is $2n$ and the maximum coefficient is $n^3(n!)^2 2^{2n+2L}$, and the same bounds apply to $2r(x)^2$.

The theorem gives a bound of $2n$ on the degree of $s(x)$, the polynomial that has θ as a root. It also gives a bound:

$$\|s\| \leq (4n)!(3n^3(n!)^2 2^{2n+2L})^{2n}(n(n!)^2 2^{2n+2L})^{2n}.$$

This number has the form $2^{O(n^2 \lg n + nL)}$.

Now finally, we can get a gap on the distance between θ and ζ, assuming they are distinct. If they are distinct, then Theorem 3 says that their distance is at least $2^{-O(n^2 \lg n + nL)}$.

Thus, we need to evaluate θ to at most $O(n^2 \lg n + nL)$ digits before we can tell whether $\theta < \zeta$, $\theta = \zeta$, or $\theta > \zeta$.

We now turn to the case that $\mu^* = \mu_0$. The goal once again is to show that θ, defined to be $f(\mathbf{x}^*)$, is the root to a polynomial. Recall that in this case $\mathbf{c}$ lies in the image of $H + \mu_0 I$. Let $\mathbf{b}$ be a column of the identity matrix such that $\mathbf{b}$ does not lie in the image of $H + \mu_0 I$. Let $\mathbf{x}^*(\epsilon)$ be the optimal solution to the problem of minimizing $\frac{1}{2}\mathbf{x}^T H\mathbf{x} + (\mathbf{c} + \epsilon\mathbf{b})^T\mathbf{x}$ subject to $\mathbf{x}^T\mathbf{x} = 1$, where ϵ is a positive parameter. Let $\theta(\epsilon)$ be the optimal value of the objective function parametrized by ϵ. Finally, let $\mu(\epsilon)$ be the value of the KKT multiplier for the parametrized problem. As we proved last section, $\theta(\epsilon)$ tends to $\theta(0)$, i.e., to θ, as ϵ tends to zero.

Note that $\mu(\epsilon)$ for $\epsilon > 0$ is a solution to the equation:

$$\sum_{i=1}^{n} \left(\frac{\det\langle H + \mu(\epsilon)I, \mathbf{c} + \epsilon\mathbf{b}\rangle_i}{\det(H + \mu(\epsilon)I)} \right)^2 = 1.$$

We can write this is as a polynomial relation by clearing the denominator: $p(\mu(\epsilon), \epsilon) = 0$ where $p(x, y)$ is a polynomial in two variables with integer coefficients.

Next, note that for $\epsilon > 0$ we can write $\theta(\epsilon)$ as a rational function $q(\mu(\epsilon), \epsilon)/(2r(\mu(\epsilon), \epsilon)^2)$, where q and r are polynomials of two variables with integer coefficients. Let $\theta(\epsilon, z)$ be the rational function

$$\frac{q(\mu(\epsilon), \epsilon) + z}{2r(\mu(\epsilon), \epsilon)^2}.$$

Now we consider

$$v(y, z, \epsilon) = \mathrm{res}_\mu(p(\mu, \epsilon), 2r(\mu, \epsilon)^2 y - q(\mu, \epsilon) - z).$$

Note that for any choice of z, ϵ, we have $v(\theta(\epsilon, z), z, \epsilon) = 0$. This is because $2r(\mu, \epsilon)^2\theta(\epsilon, z) - q(\mu, \epsilon) - z$ and $p(\mu, \epsilon)$ have a common root for μ (namely, $\mu = \mu(\epsilon)$).

We claim that $v(y, z, \epsilon)$ is a nonzero polynomial. To see this, note that for $y = 0$, ϵ fixed and z large enough, $q(\mu, \epsilon) + z$ and $p(\mu, \epsilon)$ will not have common roots. This means $v(0, z, \epsilon) \neq 0$.

Next, factor out the highest power of z from all terms of $v(y, z, \epsilon)$ so that we can write $v(y, z, \epsilon) = z^k w(y, z, \epsilon)$. Notice that the polynomial $w(y, 0, \epsilon)$ is a nonzero polynomial of y and ϵ (see the proof of Theorem 4). Notice also that for $z \neq 0$, since $v(\theta(\epsilon, z), z, \epsilon) = 0$, then $w(\theta(\epsilon, z), z, \epsilon) = 0$. Then by continuity in z, this means that $w(\theta(\epsilon, 0), 0, \epsilon) = 0$.

Thus, we have a nonzero polynomial $w(y, 0, \epsilon)$ such that $w(\theta(\epsilon), 0, \epsilon) = 0$. Let $h(y, 0, \epsilon)$ be $w(y, 0, \epsilon)$ with the highest power of ϵ factored out (i.e.,

$w(y,0,\epsilon) = \epsilon^k h(y,0,\epsilon))$. For $\epsilon \neq 0$ if $w(y,0,\epsilon) = 0$ then $h(y,0,\epsilon) = 0$ as well. In particular, for $\epsilon \neq 0$ $h(\theta(\epsilon),0,\epsilon) = 0$. But since $\theta(\epsilon)$ tends to θ as $\epsilon \to 0$, then $h(\theta,0,0) = 0$ by continuity. We claim by arguments used before that $h(y,0,0)$ is a nonzero polynomial of y.

Thus, finally, we have produced a polynomial that has θ as a root. We omit the estimates to show that in fact the coefficients and degree of this polynomial $h(y,0,0)$ are bounded well enough so that the procedure is polynomial time.

Thus, we have proved that after a polynomial number of iterations of the algorithm of Section 2, we can compute an estimate θ_1 to θ such that from the estimate it can be determined whether $\theta < \zeta$, $\theta = \zeta$, or $\theta > \zeta$.

To prove that problem QPSPHERE is solvable in polynomial time, one more issue must be addressed: the number of digits needed for the arithmetic in the various steps. We assume that all operations of QPSPHERE are carried out in exact rational arithmetic by the Turing machine, and that rational numbers are reduced to lowest terms after every operation. (This is in contrast to interior point methods for linear programming, in which case it is necessary to work with approximations to the quantities involved.) The main step of the algorithm is computing the LDL^T factorization of a matrix $H + \mu I$. We proved in Section 3 that the number of digits in the rational numbers involved in LDL^T factorization of A is polynomial in the number of digits needed to write A. Thus, we need only argue that $H + \mu I$ has a polynomial number of digits in the length of the input. Since H is part of the input, the claim follows immediately for the digits contributed by this term. As for μI, we observe that μ can be written as rational number with a polynomial number of digits (this follows from the estimates of how many digits of accuracy are needed in μ).

References

AUGUSTIN CAUCHY [1829], *Exercises de Mathematiques, Quatrième Annèe,* Bibliothèque Nationale, Paris.

JACK EDMONDS [1967], Systems of distinct representatives and linear algebra, *J. Res. Nat. Bur. Standards*, 71B:241–245.

R. FLETCHER [1987], *Practical Methods of optimization*, Wiley, New York.

MICHAEL R. GAREY AND DAVID S. JOHNSON [1979], *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco.

DAVID M. GAY [1979], Computing optimal locally constrained steps, Technical Summary Report 2013, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin.

GENE H. GOLUB AND CHARLES F. VAN LOAN [1989], *Matrix Computations, Second Edition*, Johns Hopkins University Press, Baltimore.

DOUG IERARDI AND DEXTER KOZEN [1990], Parallel resultant computation, Technical Report 90-1087, Department of Computer Science, Cornell University, Ithaca, New York.

NATHAN JACOBSON [1974], *Basic Algebra I*, W. H. Freeman, San Francisco.

N. KARMARKAR [1989], An interior-point approach to NP-complete problems, lecture, 3rd SIAM Conference on Optimization, Boston.

J. KUCZYŃSKI AND H. WOŹNIAKOWSKI [1990], Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start, preprint..

SERGE LANG [1971], *Algebra*, Addison Wesley, Reading, Massachusetts.

K. LEVENBERG [1944], A method for the solution of certain nonlinear problems in least squares, *Quarterly Appl. Math.* 2:164–168.

DONALD W. MARQUARDT [1963], An algorithm for least-squares estimation of nonlinear parameters, *J. SIAM* 11:431–441.

M. MIGNOTTE [1982], Some useful bounds, pp. 259–263 in *Computer Algebra: Symbolic and Algebraic Computation* edited by B. Buchberger, G. E. Collins and R. Loos, Springer Verlag, Wien.

JORGE J. MORÉ [1978], The Levenberg-Marquardt algorithm: Implementation and theory, pp. 105–116 in *Numerical Analysis: Proceedings of the Biennial Conference, Dundee, 1977*, edited by G. A. Watson, Springer Verlag Lecture Notes in Mathematics Volume 630, Berlin.

EUGEN NETTO [1908], *Gruppen und Substitutionentheorie*, G. J. Göschen, Leipzig.

D. C. SORENSON [1982], Newton's method with a model trust region modification, *SIAM J. Numer. Anal.* 19:409–426.

YINYU YE [1988], On the interior algorithms for nonconvex quadratic programming, preprint.

RICHARD ZIPPEL [1990], *Algebraic Computation*, manuscript.