Algorithmica (1993) 10: 365–382

Algorithmica

A Long-Step Barrier Method for Convex Quadratic Programming

K. M. Anstreicher,[1] D. den Hertog,[2] C. Roos,[2] and T. Terlaky[2,3]

Abstract. In this paper we propose a long-step logarithmic barrier function method for convex quadratic programming with linear equality constraints. After a reduction of the barrier parameter, a series of long steps along projected Newton directions are taken until the iterate is in the vicinity of the center associated with the current value of the barrier parameter. We prove that the total number of iterations is $O(\sqrt{n}L)$ or $O(nL)$, depending on how the barrier parameter is updated.

Key Words. Convex quadratic programming, Interior-point method, Logarithmic barrier function, Polynomial algorithm.

1. Introduction. Karmarkar's [14] invention of the projective method for linear programming (LP) has given rise to active research in interior-point algorithms. At this moment, the variants can be roughly categorized into four classes: projective, affine scaling, path-following, and potential reduction methods.

Researchers have also extended interior-point methods to other problems, including convex quadratic programming (QP). Kapoor and Vaidya [12], [13] and Ye and Tse [31] extended Karmarkar's projective method for QP. They proved that their method requires $O(nL)$ iterations, where L is the size of the problem. Dikin [5] had already proved convergence for the affine scaling method applied to QP. Ye [29] further analysed this method.

Monteiro and Adler [21] developed a primal–dual small-step path-following method for QP. Ye [28], Ben Daya and Shetty [1], and Goldfarb and Liu [8] proposed small-step logarithmic barrier methods for QP. All these methods require $O(\sqrt{n}L)$ iterations. Renegar and Shub [26] simplified and unified the complexity analysis for these methods. Monteiro *et al.* [22] proposed a small-step path-following method for QP in which only the primal–dual affine scaling direction is used.

Nesterov and Nemirovsky [24] proposed small-step barrier methods for smooth convex programming problems. In Chapter 5 of their monograph they also work out long-step barrier methods for convex quadratic programming problems with linear inequality constraints. Jarre [10], [11] and Mehrotra and Sun [18], [19] analysed a small-step path-following method for quadratically constrained QP

[1] Department of Management Sciences, The University of Iowa, Iowa City, IA 52242, USA.
[2] Faculty of Technical Mathematics and Computer Science, Delft University of Technology, P.O. Box 356, 2600 AJ Delft, The Netherlands.
[3] On leave from Eötvös University, Budapest and partially supported by OTKA 2116.

Received July 20, 1990; revised August 17, 1991. Communicated by Nimrod Megiddo.

and smooth convex problems, based on Huard's method of centers. Den Hertog *et al.* [2] extended this method to a long-step method. In [3] they also proposed a long-step method for smooth convex problems, based on the logarithmic barrier-function approach.

It is well known that QP can be formulated as a Linear Complementarity Problem (LCP). So, methods for solving LCPs are indirect methods for QP. However, the transformation from QP to LCP may make these methods inefficient for QP, due to the increase in the size of the matrices involved. Many interior-point methods for solving LCP have appeared in the literature, e.g., [16], [17], [20], and [30]. Kojima *et al.* [15] gave a unified approach for interior-point methods for LCP.

In this paper we propose a *long-step* logarithmic barrier method for QP with linear equality constraints. Our method is an extension of Den Hertog *et al.*'s [4] long-step barrier method for linear programming. The structure of the algorithm, and its analysis, are similar to that of [4]. Steps along projected Newton directions are taken, until the iterate is sufficiently close to the current center. After that, the barrier parameter is reduced and the process is repeated. The main difference with [4] is that three different proximity criteria, which are all equal in the LP case, are required in the analysis. Two of these criteria are directly based on the Newton step, while the third is the QP extension of a measure developed for LP by Roos and Vial [27]. Through careful use of these criteria we show that in a certain neighborhood of a center the Newton process for the barrier function quadratically converges. Our final result is that the total number of iterations is $O(\sqrt{n}L)$ or $O(nL)$, depending on how the barrier parameter is updated.

The method developed here is also related to the long-step algorithms in [2] and [3], but there are several substantial differences. The first is that [2] and [3], which apply to more general convex programming problems, have no quadratic convergence result. Using the quadratic convergence result proved here, we are able to give an analysis which is both sharper and much simpler than that of [2] and [3]. In addition, to apply the algorithms of [2] or [3] to the QP problem considered here requires an awkward transformation which makes the objective function linear, while adding a nonlinear constraint. Our analysis in this paper deals directly with the original, quadratic objective. Finally, it may be noted that the latter papers deal with inequality constraints, and consequently require additional transformations to handle the equality constraints of our problem. The biggest difference between our paper and [24], which develops a long-step algorithm for inequality constrained QP, is that our iteration bound is much better than the one obtained by Nesterov and Nemirovsky [24]. (They obtained an $O(n^4 L \ln n)$ iteration bound for the long-step method, while we have an $O(nL)$ bound.) Moreover, our analysis is substantially simpler. Our quadratic convergence result is also somewhat sharper than the quadratic convergence result in [24].

The paper is organized as follows. In Section 2 we prove some properties of the central path. In Section 3 proximity criteria are introduced, and some properties of nearly centered points are proved. The quadratic convergence property is shown to be highly expedient for proving other important bounds. In Section 4 we describe our algorithm, and in Section 5 we derive an upper bound for the total number of iterations. Finally, in Section 6 we end up with some remarks.

NOTATION. As far as notations are concerned, e denotes the vector of all ones and I the identity matrix. Given an n-dimensional vector x we denote by X the $n \times n$ diagonal matrix whose diagonal entries are the coordinates x_j of x; x^T is the transpose of the vector x and the same notation holds for matrices. Finally, $\|x\|$ denotes the l_2 norm of x.

2. Properties of the Central Path. We consider the convex quadratic programming problem in the following standard form:

$$\text{(QP)} \qquad \begin{cases} \min q(x) = c^T x + \frac{1}{2} x^T Q x, \\ \quad Ax = b, \qquad x \geq 0. \end{cases}$$

Here Q is a symmetric, positive semidefinite $n \times n$ matrix, A is an $m \times n$ matrix, and b and c are m- and n-dimensional vectors, respectively; the n-dimensional vector x is the variable in which the minimization is done. The dual formulation to (QP) is

$$\text{(QD)} \qquad \begin{cases} \quad \max d(x, y) = b^T y - \frac{1}{2} x^T Q x, \\ A^T y - Qx + s = c, \qquad s \geq 0, \end{cases}$$

where y is an m-dimensional vector.

It is well known that, for all x and y that are feasible for (QP) and (QD), we have

$$d(x, y) \leq z^* \leq q(x),$$

where z^* denotes the optimal objective value for (QP). Optimality holds if and only if the complementary slackness relation $x^T s = 0$ is satisfied.

Without loss of generality we assume that all the coefficients are integer. We denote by L the length of the input data of (QP). We make the standard assumption that the feasible set of (QP) is bounded and has a nonempty relative interior. In order to simplify the analysis we also assume that A has full rank, though this assumption is not essential.

We consider the logarithmic barrier function

$$(1) \qquad f(x, \mu) := \frac{c^T x + \frac{1}{2} x^T Q x}{\mu} - \sum_{j=1}^{n} \ln x_j,$$

where μ is a positive parameter. The first- and second-order derivatives of f are

$$\nabla f(x, \mu) = \frac{c + Qx}{\mu} - X^{-1} e,$$

$$\nabla^2 f(x, \mu) = \frac{1}{\mu} Q + X^{-2}.$$

Consequently, f is strictly convex on the relative interior of the feasible set. It also takes infinite values on the boundary of the feasible set. Thus it achieves a minimum value at a unique point. The necessary and sufficient first-order optimality conditions for this point are

(2) $$A^{\mathrm{T}}y - Qx + s = c, \qquad s \geq 0,$$

(3) $$Ax = b, \qquad x \geq 0,$$

(4) $$Xs = \mu e,$$

where y and s are m- and n-dimensional vectors, respectively.

Let us denote the unique solution of this system by $(x(\mu), y(\mu), s(\mu))$. Then the primal and dual central path is defined as the solution set $x(\mu)$ and $y(\mu)$, respectively, for $\mu > 0$. It is easy to see that the duality gap in $(x(\mu), y(\mu), s(\mu))$ satisfies

(5) $$x(\mu)^{\mathrm{T}}s(\mu) = n\mu.$$

It is well known that $x(\mu)$ and $y(\mu)$ are continuous and differentiable. Hence, if $\mu \to 0$, then $x(\mu)$ and $(x(\mu), y(\mu))$ will converge to optimal primal and dual solutions, respectively.

The following lemma states that the primal objective decreases along the primal path and the dual objective increases along the dual path. The first part is a classical result of Fiacco and McCormick [6]. The second part also follows from this result, since $(x(\mu), y(\mu))$ is the minimizer of the (convex) dual logarithmic barrier function. Our proof is completely different from the classical one.

LEMMA 1. *The objective $q(x(\mu))$ of the primal problem* (QP) *is monotonically decreasing and the objective $d(x(\mu), y(\mu))$ of the dual problem* (QD) *is monotonically increasing if μ decreases.*

PROOF. To prove the first part of the lemma it suffices to show that

(6) $$q(x(\mu))' = c^{\mathrm{T}}x' + x^{\mathrm{T}}Qx' \geq 0,$$

where the prime denotes the derivative with respect to μ. Using that $x(\mu)$ and $y(\mu)$ satisfy (2)–(4) and taking derivatives with respect to μ we obtain

(7) $$A^{\mathrm{T}}y' - Qx' + s' = 0,$$

(8) $$Ax' = 0,$$

(9) $$Xs' + Sx' = e.$$

Now, using (2) and (8), we have

$$c^{\mathrm{T}}x' = (s - Qx + A^{\mathrm{T}}y)^{\mathrm{T}}x' = s^{\mathrm{T}}x' - x^{\mathrm{T}}Qx'.$$

Using (9), (4), (7), and (8), respectively, this results into

$$\begin{aligned} c^{\mathrm{T}}x' + x^{\mathrm{T}}Qx' = s^{\mathrm{T}}x' = e^{\mathrm{T}}Sx' &= (Sx' + Xs')^{\mathrm{T}}Sx' \\ &= (x')^{\mathrm{T}}S^2x' + \mu(x')^{\mathrm{T}}s' \\ &= (x')^{\mathrm{T}}S^2x' + \mu(x')^{\mathrm{T}}Qx' \geq 0. \end{aligned}$$

Thus, the first part of the lemma follows because of (6).

To prove the second part of the lemma it suffices to show that

$$d(x(\mu), y(\mu))' = b^{\mathrm{T}}y' - (x')^{\mathrm{T}}Qx \leq 0. \tag{10}$$

Multiplying (9) by AS^{-1} we obtain

$$AS^{-1}Xs' + Ax' = AS^{-1}e,$$

which reduces to $AX^2s' = b$. Now, taking the inner product with y' results into

$$\begin{aligned} b^{\mathrm{T}}y' &= (A^{\mathrm{T}}y')^{\mathrm{T}}X^2s' \\ &= (x')^{\mathrm{T}}QX^2s' - (s')^{\mathrm{T}}X^2s' \\ &= (x')^{\mathrm{T}}QX(e - Sx') - (s')^{\mathrm{T}}X^2s'. \end{aligned}$$

Consequently, we have

$$b^{\mathrm{T}}y' - (x')^{\mathrm{T}}Qx = -\mu(x')^{\mathrm{T}}Qx' - (s')^{\mathrm{T}}X^2s' \leq 0.$$

Together with (10), this proves the second part of the lemma. □

3. Properties Near the Central Path. To measure the distance to the central path of noncentered points, we introduce three measures. The first is analogous to Roos and Vial's [27] appealing measure for linear programming:

$$\delta(x, \mu) := \min_y \left\| \frac{X}{\mu}(c + Qx - A^{\mathrm{T}}y) - e \right\|. \tag{11}$$

Loosely speaking, $\delta(x, \mu)$ measures the deviation from optimality condition (4). The unique solution of the minimization problem in the definition of $\delta(x, \mu)$ is denoted by $y(x, \mu)$ and the corresponding slack variable by $s(x, \mu)$ (i.e.,

$s(x, \mu) = c + Qx - A^T y(x, \mu)$). It can be easily verified that

$$x = x(\mu) \quad \Leftrightarrow \quad \delta(x, \mu) = 0 \quad \Rightarrow \quad y(x, \mu) = y(\mu).$$

In the sequel of this paper we will deal with the scaled version of (QP):

$$(\overline{\mathrm{QP}}) \qquad \begin{cases} \min \bar{q}(\bar{x}) = \bar{c}^T \bar{x} + \frac{1}{2}\bar{x}^T \bar{Q} \bar{x}, \\ \quad \bar{A}\bar{x} = b, \qquad \bar{x} \geq 0, \end{cases}$$

where $\bar{A} = AX$, $\bar{c} = Xc$, $\bar{Q} = XQX$, and x is the current iterate. Hence, a step starting at e in $(\overline{\mathrm{QP}})$ corresponds to a step starting at x in (QP). The scaled versions of $\nabla f(x, \mu)$ and $\nabla^2 f(x, \mu)$ are denoted by $g = g(x, \mu)$ and $H = H(x, \mu)$, respectively. So

$$g = \frac{\bar{c} + \bar{Q}e}{\mu} - e$$

and

$$H = \frac{\bar{Q}}{\mu} + I.$$

In the algorithm, described in the following section, we do steps along projected Newton directions to minimize f for fixed μ, i.e., the directions correspond to exact minimization of the quadratic approximation to f on the affine space $Ax = b$. This means that the Newton direction is determined by

$$(12) \qquad g + Hp = \bar{A}^T \bar{y}, \qquad \bar{A}p = 0.$$

Gill *et al.* [7] give two alternative (equivalent) forms for this direction:

- The range-space form

$$(13) \qquad p(x, \mu) = -H^{-1}(I - \bar{A}^T(\bar{A}H^{-1}\bar{A}^T)^{-1}\bar{A}H^{-1})g.$$

- The null-space form

$$(14) \qquad p(x, \mu) = -Z(Z^T H Z)^{-1} Z^T g,$$

where Z is an $n \times (n - m)$ matrix, with independent columns, such that $\bar{A}Z = 0$.

The second and third measures for the distance to the central path are $\|p(x, \mu)\|$ and $\|p(x, \mu)\|_{H(x,\mu)}$, where the latter is defined by

$$\|p(x, \mu)\|_{H(x,\mu)}^2 = p(x, \mu)^T H(x, \mu) p(x, \mu).$$

We note that because $H(x, \mu)$ is positive definite, $\|\cdot\|_{H(x,\mu)}$ defines a norm.

In the remainder of the paper we sometimes write δ and p instead of $\delta(x, \mu)$ and $p(x, \mu)$ for brevity. We remark that although all of the three measures are used in the analysis, only $\|p\|_H$ is used in the algorithm.

We work with the null-space form for p, because it facilitates the analysis so much. In the analysis we also assume that $Z^T Z = I$, hence Z is orthonormal. In this case we have the following well-known properties:

Property 1. ZZ^T is the projection onto the null-space of $\bar{A}$.

Property 2. $\|Z\lambda\| = \|\lambda\|$ for any λ.

Property 3. $\|Z^T x\| \leq \|x\|$ for any x, with equality if x is in the null-space of $\bar{A}$.

The following lemma shows that there is a close relationship between the three measures.

Lemma 2. *For given x and μ, $\|p\|^2 \leq \|p\|_H^2 = -p^T g \leq \delta^2$.*

Proof. Using the null-space form (14) it follows that $ZZ^T g = -ZZ^T HZZ^T p$. Property 1 and the fact that $\bar{A}p = 0$ imply $ZZ^T p = p$. Consequently, we may write

$$-p^T g = -p^T ZZ^T g = (ZZ^T p)^T HZZ^T p = p^T H p = p^T\left(I + \frac{1}{\mu}\bar{Q}\right)p \geq \|p\|^2.$$

So, it remains to prove the last inequality of the lemma. Using the definition of p it follows that

$$\begin{aligned}
-g^T p &= g^T Z\left(I + \frac{1}{\mu} Z^T \bar{Q} Z\right)^{-1} Z^T g \\
&\leq \|Z^T g\|^2 \left\|\left(I + \frac{1}{\mu} Z^T \bar{Q} Z\right)^{-1}\right\| \\
&\leq \|Z^T g\|^2,
\end{aligned}$$

where the last inequality follows because the eigenvalues of $(I + (1/\mu)Z^T \bar{Q} Z)^{-1}$ are all positive and less than or equal to 1. Moreover,

$$Z^{\mathrm{T}}g = Z^{\mathrm{T}}\left(\frac{\bar{c}+\bar{Q}e}{\mu} - e\right) = Z^{\mathrm{T}}\left(\frac{\bar{c}+\bar{Q}e-\bar{A}^{\mathrm{T}}y}{\mu} - e\right),$$

where the last equality holds for any y, because $\bar{A}Z = 0$. Putting y equal to $y(x, \mu)$, it follows from Property 3 that $\|Z^{\mathrm{T}}g\| \le \delta$. Hence it follows that $-g^{\mathrm{T}}p \le \delta^2$. This proves the lemma. □

Now we prove some fundamental lemmas for nearly centered points. The following lemma shows quadratic convergence in the vicinity of the central path for all three measures.

LEMMA 3. *If* $\|p(x, \mu)\| < 1$, *then* $x^* = x + Xp(x, \mu)$ *is a strictly feasible point for* (QP). *Moreover,*

$$\|p(x^*, \mu)\| \le \|p(x^*, \mu)\|_{H(x^*,\mu)} \le \delta(x^*, \mu) \le \|p(x, \mu)\|^2 \le \|p(x, \mu)\|^2_{H(x,\mu)} \le \delta(x, \mu)^2.$$

PROOF. It is easy to see that $Ax^* = b$. Moreover, $x^* = X(e + p) > 0$, because $\|p\| < 1$. This proves the first part of the lemma.

Using the definition of $\delta(x^*, \mu)$ we have

$$\delta(x^*, \mu) = \min_y \left\| \frac{X(I+P)}{\mu}(c + QX(e+p) - A^{\mathrm{T}}y) - e \right\| = \min_y \left\| \frac{I+P}{\mu}(\bar{c} + \bar{Q}(e+p) - \bar{A}^{\mathrm{T}}y) - e \right\|. \tag{15}$$

Now using $y = \mu\bar{y}$ in (15), where $\bar{y}$ is defined by (12), it follows that

$$\delta(x^*, \mu) \le \|(I+P)(e-p) - e\| = \|Pp\| \le \|p\|^2.$$

This proves the middle inequality of the lemma. The rest follows immediately from Lemma 2. □

Note that a somewhat weaker quadratic convergence result is obtained in [24]. To be more precise, they obtain in a completely different analysis the following result:

$$\|p(x^*, \mu)\|_{H(x^*,\mu)} \le \frac{\|p\|_H^2}{1 - \|p\|_H}.$$

From this result we have that convergence only occurs if $\|p\|_H < \frac{1}{2}$.

LEMMA 4. *If* $\|p\|_H < 1$, *then*

$$f(x, \mu) - f(x(\mu), \mu) \leq \frac{\|p\|_H^2}{1 - \|p\|_H^2}.$$

PROOF. The barrier function f is convex for fixed μ, whence

$$f(x, \mu) - f(x + Xp, \mu) \leq -g^{\mathrm{T}}p = \|p\|_H^2, \tag{16}$$

where the equality follows from Lemma 2. Now let $x^0 := x$ and let $x^0, x^1, x^2, \ldots$ denote the sequence of points obtained by repeating Newton steps, starting at x^0. Then we may write, using Lemma 3,

$$f(x, \mu) - f(x(\mu), \mu) = \sum_{i=0}^{\infty} (f(x^i, \mu) - f(x^{i+1}, \mu)) \leq \sum_{i=0}^{\infty} \|p\|_H^{2^{i+1}} \leq \frac{\|p\|_H^2}{1 - \|p\|_H^2}. \quad \square$$

LEMMA 5. *If* $\|p\|_H < 1$, *then*

$$|q(x) - q(x(\mu))| \leq \frac{\|p\|_H(1 + \|p\|_H)}{1 - \|p\|_H} \mu\sqrt{n}.$$

PROOF. Since $q(x)$ is convex, we have

$$\nabla q(x)^{\mathrm{T}} Xp \leq q(x + Xp) - q(x) \leq \nabla q(x + Xp)^{\mathrm{T}} Xp. \tag{17}$$

For the left-hand side expression we can derive the following lower bound, using that $\nabla q(x) = c + Qx = \mu X^{-1}g + \mu X^{-1}e$:

$$\nabla q(x)^{\mathrm{T}} Xp = \mu g^{\mathrm{T}}p + \mu e^{\mathrm{T}}p \geq -\mu\|p\|_H^2 - \mu\|p\|_H\sqrt{n} \geq -\|p\|_H(1 + \|p\|_H)\mu\sqrt{n}, \tag{18}$$

where the first inequality follows from Lemma 2. Now, using (12), we derive an upper bound for the right-hand side expression in (17):

$$\begin{aligned} \nabla q(x + Xp)^{\mathrm{T}} Xp &= \bar{c}^{\mathrm{T}}p + p^{\mathrm{T}}\bar{Q}(e + p) = \bar{c}^{\mathrm{T}}p + p^{\mathrm{T}}(\mu\bar{A}^{\mathrm{T}}\bar{y} + \mu e - \bar{c} - \mu p) \\ &= \mu e^{\mathrm{T}}p - \mu\|p\|^2 \leq \mu e^{\mathrm{T}}p \leq \mu\|p\|_H\sqrt{n}. \end{aligned} \tag{19}$$

Consequently, substitution of (18) and (19) into (17) yields

$$|q(x) - q(x + Xp)| \leq \|p\|_H(1 + \|p\|_H)\mu\sqrt{n}.$$

The remainder of the proof follows by considering a sequence of Newton steps initiated at $x^0 := x$, as in the proof of Lemma 4. $\square$

4. The Algorithm. In our long-step algorithm a linesearch along the Newton direction is done until the iterate is sufficiently close to the current center. After that, the barrier parameter is reduced, and the process starts again.

In the next section we show that a linesearch along the Newton direction reduces the barrier function value by a constant if $\|p\|_H \geq \frac{1}{2}$. This enables us to derive an upper bound for the number of steps between reductions of μ.

If $\|p\|_H < \frac{1}{2}$ we can give bounds on $f(x, \mu) - f(x(\mu), \mu)$, $|q(x) - q(x(\mu))|$ and $c^T x - z^*$ by Lemmas 4 and 5.

Algorithm
Input:
μ_0 is the initial barrier value, $\mu_0 \leq 2^{O(L)}$;
t is an accuracy parameter, $t = O(L)$;
θ is the reduction parameter, $0 < \theta < 1$;
x^0 is a given interior feasible point such that $\|p(x^0, \mu_0)\|_{H(x^0, \mu_0)} \leq \frac{1}{2}$.

begin
  $x := x_0$; $\mu := \mu_0$;
  while $\mu > 2^{-t}$ **do**
  begin (outer step)
    while $\|p\|_H \geq \frac{1}{2}$ **do**
    begin (inner step)
      $\tilde{\alpha} := \arg\min_{\alpha > 0}\{f(x + \alpha X p, \mu): x + \alpha X p > 0\}$
      $x := x + \tilde{\alpha} X p$
    end (inner step)
    $\mu := (1 - \theta)\mu$;
  end (outer step)
end.

For finding the initial point that satisfies the input assumptions of the algorithm we refer the reader to, e.g., [28]. Later the centering assumption

$$\|p(x^0, \mu_0)\|_{H(x^0, \mu_0)} \leq \tfrac{1}{2}$$

is relaxed.

5. Convergence Analysis of the Algorithm, In this section we derive upper bounds for the total number of outer and inner iterations.

THEOREM 1. *After at most $K = O(L/\theta)$ outer iterations, the algorithm ends up with a primal solution such that $q(x) - z^* \leq 2^{-O(L)}$.*

PROOF. This is an easy consequence of (5) and Lemma 4. □

We note, that this final primal solution can be rounded to an optimal solution for (QP) in $O(n^3)$ arithmetic operations. (See, e.g., [25].)

The following lemma is needed to derive an upper bound for the number of inner iterations in each outer iteration. It states that a sufficient decrease in the value of the barrier function can be obtained by taking a step along the Newton direction. The same result is also obtained in [24] in a much more complicated way.

LEMMA 6. *Let* $\bar{\alpha} := (1 + \|p\|_H)^{-1}$. *Then*

$$\Delta f := f(x, \mu) - f(x + \bar{\alpha} X p, \mu) \geq \|p\|_H - \ln(1 + \|p\|_H).$$

PROOF. We write down the MacLaurin series for f with respect to α:

$$f(x + \alpha X p, \mu) = f(x, \mu) + \alpha g^{\mathrm{T}} p + \tfrac{1}{2}\alpha^2 p^{\mathrm{T}} H p + \sum_{k=3}^{\infty} t_k,$$

where t_k denotes the k-order term in the MacLaurin series. Since

$$t_k = \frac{(-\alpha)^k}{k} \sum_{i=1}^{n} p_i^k,$$

we find, by using Lemma 2,

$$|t_k| \leq \frac{\alpha^k}{k} \sum_{i=1}^{k} |p_i|^k \leq \frac{\alpha^k}{k} \left(\sum_{i=1}^{n} |p_i|^2 \right)^{k/2} \leq \frac{\alpha^k}{k} \|p\|^k \leq \frac{\alpha^k}{k} \|p\|_H^k.$$

Using Lemma 2, we have, for the linear and quadratic term in the MacLaurin series,

$$\alpha p^{\mathrm{T}} g + \tfrac{1}{2} a^2 p^{\mathrm{T}} H p = (\tfrac{1}{2}\alpha^2 - \alpha)\|p\|_H^2.$$

So we find

$$\begin{aligned} f(x + \alpha X p, \mu) &\leq f(x, \mu) + (\tfrac{1}{2}\alpha^2 - \alpha)\|p\|_H^2 + \sum_{k=3}^{\infty} \frac{\alpha^k}{k} \|p\|_H^k \\ &= f(x, \mu) - \alpha\|p\|_H^2 - \ln(1 - \alpha\|p\|_H) - \alpha\|p\|_H. \end{aligned}$$

Hence

$$\Delta f \geq \alpha(\|p\|_H^2 + \|p\|_H) + \ln(1 - \alpha\|p\|_H). \tag{20}$$

The right-hand side is maximal if $\alpha = \bar{\alpha} = (1 + \|p\|_H)^{-1}$. Substitution of this value finally gives

$$\Delta f \geq \|p\|_H - \ln(1 + \|p\|_H).$$

This proves the lemma. □

THEOREM 2. *Each outer iteration requires at most*

$$\frac{11\theta}{1-\theta}(\theta n + 3\sqrt{n}) + \tfrac{11}{3}$$

inner iterations.

PROOF. This proof is a generalization of Gonzaga's [9] proof for the linear case. Let us consider the $(k+1)$st outer iteration. The starting point is then (x^k, μ_k), with $\|p(x^k, \mu_{k-1})\|_{H(x^k, \mu_{k-1})} < \frac{1}{2}$. Let N denote the number of inner iterations. For each inner iteration we know, according to Lemma 6, that the decrease in the potential function value is at least

$$\Delta f \geq \tfrac{1}{2} - \ln(1 + \tfrac{1}{2}) > \tfrac{1}{11}.$$

Following the N inner iterations, we have x^{k+1} with $\|p(x^{k+1}, \mu_k)\|_{H(x^{k+1}, \mu_k)} < \frac{1}{2}$. So we have

$$f(x^{k+1}, \mu_k) \leq f(x^k, \mu_k) - \tfrac{1}{11}N.$$

Equivalently,

$$\tfrac{1}{11}N \leq f(x^k, \mu_k) - f(x^{k+1}, \mu_k). \tag{21}$$

Now we derive an upper bound for the right-hand side. The definition of $f(x, \mu)$, $x > 0$, implies that

$$\begin{aligned} f(x, \mu_k) &= f(x, \mu_{k-1}) + \frac{q(x)}{\mu_k} - \frac{q(x)}{\mu_{k-1}} \\ &= f(x, \mu_{k-1}) + \frac{q(x)}{\mu_{k-1}}\left(\frac{1}{1-\theta} - 1\right) \\ &= f(x, \mu_{k-1}) + \frac{\theta}{1-\theta}\frac{q(x)}{\mu_{k-1}}. \end{aligned}$$

Using this we obtain

(22)

$$f(x^k, \mu_k) - f(x^{k+1}, \mu_k) = f(x^k, \mu_{k-1}) - f(x^{k+1}, \mu_{k-1}) + \frac{\theta}{1-\theta} \frac{1}{\mu_{k-1}} (q(x^k) - q(x^{k+1})).$$

Because x^k and x^{k+1} are approximately centered with respect to $x(\mu_{k-1})$ and $x(\mu_k)$, respectively, using Lemma 5 for the first and Lemma 1 for the second inequality we find

$$\begin{aligned}
q(x^k) - q(x^{k+1}) &\leq q(x(\mu_{k-1})) + \tfrac{3}{2}\mu_{k-1}\sqrt{n} - q(x(\mu_k)) + \tfrac{3}{2}\mu_k\sqrt{n} \\
&= q(x(\mu_{k-1})) - q(x(\mu_k)) + \tfrac{3}{2}(2-\theta)\mu_{k-1}\sqrt{n} \\
&\leq (q(x(\mu_{k-1})) - d(x(\mu_{k-1}), y(\mu_{k-1}))) \\
&\quad - (q(x(\mu_k)) - d(x(\mu_k), y(\mu_k))) + 3\mu_{k-1}\sqrt{n} \\
&= \mu_{k-1}n - \mu_k n + 3\mu_{k-1}\sqrt{n} \\
&= \mu_{k-1}(\theta n + 3\sqrt{n}).
\end{aligned}$$

Secondly, using Lemma 4 (with $\|p\|_H = \frac{1}{2}$), and the fact that $x(\mu_{k-1})$ minimizes $f(x(\mu_{k-1}), \mu_{k-1})$, we obtain

$$\begin{aligned}
f(x^k, \mu_{k-1}) - f(x^{k+1}, \mu_{k-1}) &= f(x^k, \mu_{k-1}) - f(x(\mu_{k-1}), \mu_{k-1}) \\
&\quad + f(x(\mu_{k-1}), \mu_{k-1}) - f(x^{k+1}, \mu_{k-1}) \\
&\leq f(x^k, \mu_{k-1}) - f(x(\mu_{k-1}), \mu_{k-1}) \\
&\leq \tfrac{1}{3}.
\end{aligned}$$

Hence, substitution of the last two inequalities into (22) yields

$$f(x^k, \mu_k) - f(x^{k+1}, \mu_k) \leq \frac{\theta}{1-\theta}(\theta n + 3\sqrt{n}) + \tfrac{1}{3}.$$

Substitution of this inequality into (21) yields the lemma. □

Combining Theorems 1 and 2, the total number of iterations turns out to be given by

(23)
$$\left[\frac{11}{1-\theta}(\theta n + 3\sqrt{n}) + \frac{11}{3\theta}\right] O(L).$$

This makes clear that:

- If we take $\theta = \Omega(1/\sqrt{n})$, then $O(\sqrt{n}L)$ iterations are needed.
- If we take $\theta = \Omega(1)$, then $O(nL)$ iterations are needed.

6. Concluding Remarks

6.1. Computing the Newton Direction. Even though our analysis is based on the null-space form for the Newton direction, in practice either the null-space or the row-space form can be used. It is obvious that the null-space form is more efficient when the number of linear constraints is relatively large compared with the number of variables. The row-space form is efficient when the number of linear constraints is small compared with the number of variables.

In our analysis we assumed that Z is orthonormal. However, the search direction does not change if Z is any basis for the nullspace of $\bar{A}$. So, in practice we do not have to do all the work to find an orthonormal Z on each iteration. For example, if a $\bar{Z}$ is found such that $A\bar{Z} = 0$, then $Z = X^{-1}\bar{Z}$ satisfies $AXZ = 0$. We refer the reader to [7] for the numerical aspects.

6.2. Obtaining Dual Feasible Solutions. At the end of each sequence of inner iterations we have a primal feasible x such that $\|p\|_H \leq 1$. The following lemma shows that a dual feasible solution can be obtained by performing an additional full Newton step, and projection.

LEMMA 7. *Let $x^* = x + Xp(x, \mu)$. If $\|p(x, \mu)\|_{H(x,\mu)} \leq 1$, then $\delta := \delta(x^*, \mu) \leq 1$ and $y := y(x^*, \mu)$ is dual feasible. Moreover, the duality gap satisfies*

$$\mu(n - \delta\sqrt{n}) \leq q(x^*) - d(x^*, y) \leq \mu(n + \delta\sqrt{n}).$$

PROOF. By Lemma 3 we have $\delta(x^*, \mu) \leq \|p(x, \mu)\|^2_{H(x,\mu)} \leq 1$. By the definition of $s(x, \mu) = c + Qx - A^{\mathrm{T}}y(x, \mu)$ we have

$$\delta(x^*, \mu) = \left\| \frac{X^*s(x^*, \mu)}{\mu} - e \right\| \leq 1.$$

This implies $s(x^*, \mu) \geq 0$, so $y(x^*, \mu)$ is dual feasible. Moreover,

$$\left| \frac{(x^*)^{\mathrm{T}}s(x^*, \mu)}{\mu} - n \right| = \left| e^{\mathrm{T}}\left(\frac{X^*s(x^*, \mu)}{\mu} - e \right) \right| \leq \|e\| \left\| \frac{X^*s(x^*, \mu)}{\mu} - e \right\| = \delta\sqrt{n}.$$

Consequently, using that $(x^*)^{\mathrm{T}}s(x^*, \mu) = q(x^*) - d(x^*, y)$,

$$\mu(n - \delta\sqrt{n}) \leq q(x^*) - d(x^*, y) \leq \mu(n + \delta\sqrt{n}).$$

□

6.3. Small-Step Path-Following Methods. Small-step path-following methods start at a nearly centered iterate and after the parameter is reduced by a small factor, a unit Newton step is taken. The reduction parameter is sufficiently small, such that the new iterate is again nearly centered with respect to the new center. Small-step barrier methods for convex quadratic programming have been given by Ye [28], Goldfarb and Liu [8], and Ben Daya and Shetty [1]. The following lemma shows that if θ is small, then we obtain such a small-step path-following method.

LEMMA 8. *Let* $x^* := x + Xp(x, \mu)$ *and* $\mu^* := (1-\theta)\mu$, *where* $\theta = 1/10\sqrt{n}$. *If* $\delta(x, \mu) \le \frac{1}{2}$, *then* $\delta(x^*, \mu^*) \le \frac{1}{2}$.

PROOF. Due to the definition of our measure we have

$$\begin{aligned}\delta(x, \mu^*) &= \left\| \frac{Xs(x, \mu^*)}{\mu^*} - e \right\| \\ &\le \left\| \frac{Xs(x, \mu)}{\mu^*} - e \right\| \\ &= \left\| \frac{1}{1-\theta}\left(\frac{Xs(x, \mu)}{\mu} - e\right) + \left(\frac{1}{1-\theta} - 1\right)e \right\| \\ &\le \frac{1}{1-\theta}\left(\delta(x, \mu) + \theta\sqrt{n}\right) \\ &\le \frac{1}{1-\frac{1}{10}}\left(\frac{1}{2} + \frac{1}{10}\right) \\ &= \frac{2}{3}.\end{aligned}$$

Now we can apply the quadratic convergence result (Lemma 3)

$$\delta(x^*, \mu^*) \le \delta(x, \mu^*)^2 \le \frac{4}{9} < \frac{1}{2}. \qquad \square$$

6.4. Relaxing the Initial Centering Condition. The initial centering condition $\|p(x^0, \mu_0)\|_{H(x^0, \mu_0)} \le \frac{1}{2}$ can be relaxed to

$$f(x^0, \mu_0) - f(x(\mu_0), \mu_0) \le O(\sqrt{n}L)$$

if $\theta = \Omega(1/\sqrt{n})$ is used, and to

$$f(x^0, \mu_0) - f(x(\mu_0), \mu_0) \le O(nL)$$

if $\theta = \Omega(1)$ is used. This holds because of Lemma 6. For the last case the assumption is equivalent with the assumption $x_j^0 \ge 2^{-L}$ for all j. This can be easily verified.

Since $x(\mu_0)$ is primal feasible, it can be written as a convex combination of basic feasible solutions. The coordinates x_j of each basic feasible solution satisfy $x_j \leq 2^L$. Moreover, $q(x^0) - q(x(\mu_0)) \leq 2^{O(L)}$. Therefore

$$f(x^0, \mu_0) - f(x(\mu_0), \mu_0) = \frac{q(x^0) - q(x(\mu_0))}{\mu_0} - \sum_{j=1}^{n} \ln x_j^0 + \sum_{j=1}^{n} \ln x_j(\mu_0) \leq O(nL).$$

6.5. Results for the LP Case. It is worthwhile to look at the results for the LP case, for which $Q = 0$. In this case the projected Newton direction (14) reduces to $p = -ZZ^{\mathrm{T}}g$, which coincides with the scaled projected gradient direction. It is easy to verify that the three measures $\|p\|$, $\|p\|_H$, and δ are exactly the same (i.e., in Lemma 2 equalities hold instead of inequalities). Hence, the resulting algorithm and results are the same as in [4].

6.6. Barrier Methods for LCP. It is well known (see, e.g., [23]) that the Linear Complementarity Problem (LCP)

$$\text{(LCP)} \qquad \begin{cases} y = Mx + q, & x, y \geq 0, \\ x^{\mathrm{T}}y = 0, \end{cases}$$

is completely equivalent to the following quadratic programming problem:

$$\text{(QP)} \qquad \begin{cases} \min x^{\mathrm{T}}y \\ y = Mx + q, & x, y \geq 0. \end{cases}$$

If M is positive semidefinite, then this problem is equivalent to a convex quadratic programming problem, since $x^{\mathrm{T}}y = \frac{1}{2}x^{\mathrm{T}}Qx + q^{\mathrm{T}}x$, where $Q = M + M^{\mathrm{T}}$. Consequently, the algorithm proposed in this paper can also be applied to positive semidefinite LCPs.

6.7. Comparison with Nesterov and Nemirovsky's Results. In Chapter 5 of their monograph [24], Nesterov and Nemirovsky analysed long-step barrier methods for QP with linear inequality constraints. Their analysis is totally different from ours: it is not based on changes in the barrier-function value, for example. On the one hand their analysis is more general, but on the other hand it is also very complicated. From their analysis it can be extracted that the total number of iterations is at most $O((1/\theta + n^4\theta^7)L \ln n)$. Note that the iteration bound (23) is better than this one. Our bound is much better if we deal with real long-step algorithms (i.e., θ is large). For example:

- If we take $\theta = \Omega(1/\sqrt{n})$, then Nesterov and Nemirovsky require $O(\sqrt{n}L \ln n)$ iterations. The difference with our $O(\sqrt{n}L)$ iteration bound is not so significant in this case.
- If we take $\theta = \Omega(1)$, then Nesterov and Nemirovsky require $O(n^4 L \ln n)$ iterations. This bound is much worse than our $O(nL)$ iteration bound.

Acknowledgments. Thanks are due to Todd who simplified the proof of Lemma 3.

References

[1] Ben Daya, M., and Shetty, C. M. (1988), Polynomial Barrier Function Algorithms for Convex Quadratic Programming, Research Report No. J88-5, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

[2] Den Hertog, D., Roos, C., and Terlaky, T. (1992), A Large-Step Analytic Center Method for a Class of Smooth Convex Programming Problems, *SIAM Journal on Optimization*, **2**, 55–70.

[3] Den Hertog, D., Roos, C., and Terlaky, T. (1992), On the Classical Logarithmic Barrier Method for a Class of Smooth Convex Programming Problems, *Journal on Optimization: Theory and Applications*, **73**, 1–25.

[4] Den Hertog, D., Roos, C., and Vial, J.-Ph. (1992), A Complexity Reduction for the Long-Step Path-Following Algorithm for Linear Programming, *SIAM Journal on Optimization*, **2**, 71–87.

[5] Dikin, I. I. (1967), Iterative Solution of Problems of Linear and Quadratic Programming, *Doklady Akademiia Nauk SSSR*, **174**, 747–748.

[6] Fiacco, A. V., and McCormick, G. P. (1968), *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*, Wiley, New York.

[7] Gill, P. E., Murray, W., and Wright, M. H. (1988), *Practical Optimization*, Academic Press, San Diego, CA.

[8] Goldfarb, D., and Liu, S. (1990), An $O(n^3L)$ Primal Interior Point Algorithm for Convex Quadratic Programming, *Mathematical Programming*, **49**, 325–340.

[9] Gonzaga, C. C. (1991), Large-Step Path-Following Methods for Linear Programming: Barrier Function Method, *SIAM Journal on Optimization*, **1**, 268–279.

[10] Jarre, F. (1990/91), On the Convergence of the Method of Analytic Centers when Applied to Convex Quadratic Programs, *Mathematical Programming*, **49**, 341–358.

[11] Jarre, F. (1989), The Method of Analytic Centers for Smooth Convex Programs, Dissertation, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Würzburg.

[12] Kapoor, S., and Vaidya, P. (1986), Fast Algorithms for Convex Quadratic Programming and Multicommodity Flows, *Proceedings of the 18th Annual ACM Symposium on the Theory of Computing*, pp. 147–159.

[13] Kapoor, S., and Vaidya, P. M. (1988), An Extension of Karmarkar's Interior Point Method to Convex Quadratic Programming, Technical Report, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL.

[14] Karmarkar, N. (1984), A New Polynomial-Time Algorithm for Linear Programming, *Combinatorica*, **4**, 373–395.

[15] Kojima, M., Megiddo, N., Noma, T., and Yoshise, A. (1991), *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Computer Science, Vol. 538, Springer-Verlag, Berlin.

[16] Kojima, M., Mizuno, S., and Yoshise, A. (1989), A Polynomial Time Algorithm for a Class of Linear Complementarity Problems, *Mathematical Programming*, **44**, 1–26.

[17] Kojima, M., Mizuno, S., and Yoshise, A. (1991), An $O(\sqrt{n}L)$ Iteration Potential Reduction Algorithm for Linear Complementarity Problems, *Mathematical Programming*, **50**, 331–342.

[18] Mehrotra, S., and Sun, J. (1990), *An Interior Point Algorithm for Solving Smooth Convex Programs Based on Newton's Method*, Contemporary Mathematics, Vol. 114. American Mathematical Society, Providence, RI, pp. 265–284.

[19] Mehrotra, S., and Sun, J. (1991), On Computing the Center of a Quadratically Constrained Set, *Mathematical Programming*, **50**, 81–89.

[20] Mizuno, S. (1990), An $O(n^3L)$ Algorithm Using a Sequence for a Linear Complementarity Problem, *Journal of the Operations Research Society of Japan*, **33**, 66–75.

[21] Monteiro, R. D. C., and Adler, I. (1989), Interior Path-Following Primal–Dual Algorithms, Part II: Convex Quadratic Programming, *Mathematical Programming*, **44**, 43–66.

[22] Monteiro, R. D. C., Adler, I., and Resende, M. G. C. (1990), A Polynomial-Time Primal–Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and Its Power Series Extension, *Mathematics of Operations Research*, **15**, 191–214.

[23] Murty, K. G. (1988), *Linear Complementarity, Linear and Nonlinear Programming*, Heldermann-Verlag, Berlin.

[24] Nesterov, Y. E., and Nemirovsky, A. S. (1989), Self-Concordant Functions and Polynomial Time Methods in Convex Programming, Book Preprint, CEMI, USSR Academy of Sciences, Moscow.

[25] Papadimitriou, C. R., and Steiglitz, K. (1982), *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, NJ.

[26] Renegar, J., and Shub, M. (1988) Simplified Complexity Analysis for Newton LP Methods, Technical Report No. 807, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.

[27] Roos, C., and Vial, J.-Ph. (1992), A Polynomial Method of Approximate Centers for Linear Programming, *Mathematical Programming*, **54**, 295–305.

[28] Ye, Y. (1987), Further Development on the Interior Algorithm for Convex Quadratic Programming, Engineering-Economic Systems Department, Stanford University, Stanford, CA.

[29] Ye, Y. (1987), Interior Algorithms for Linear, Quandratic and Linearly Constrained Convex Programming, Ph.D. Dissertation, Engineering-Economic Systems Department, Stanford University, Stanford, CA.

[30] Ye, Y., and Pardalos, P. (1991), A Class of Linear Complementarity Problems Solvable in Polynomial Time, *Linear Algebra and Its Applications*, **152**, 3–17.

[31] Ye, Y., and Tse, E. (1986), A Polynomial-Time Algorithm for Convex Quadratic Programming, Engineering-Economic Systems Department, Stanford University, Stanford, CA.