SIAM J. OPTIMIZATION
Vol. 2, No. 1, pp. 71–87, February 1992

A COMPLEXITY REDUCTION FOR THE LONG-STEP PATH-FOLLOWING ALGORITHM FOR LINEAR PROGRAMMING*

D. DEN HERTOG†, C. ROOS†, AND J.-PH. VIAL‡

Abstract. A modification of previously published long-step path-following algorithms for the solution of the linear programming problem is presented. This modification uses the simple Goldstein–Armijo rule. A $\sqrt{n}$ reduction in the complexity bound is obtained, while a linesearch may still be done. Depending on the updating scheme for the barrier parameter, the resulting complexity bounds are $O(n^3L)$ or $O(n^{3.5}L)$.

Key words. linear programming, interior point method, logarithmic barrier function, polynomial algorithm, Goldstein–Armijo rule

AMS(MOS) subject classification. 90C05

1. Introduction. The original $O(n^{3.5}L)$ complexity bound of short-step path-following methods was reduced to $O(n^3L)$ by Vaidya [15]; Gonzaga [6]; Kojima, Mizuno, and Yoshise [9]; and Monteiro and Adler [11]. This reduction was achieved by using Karmarkar's [8] partial updating scheme. Their partial updating analysis is based on steps of a fixed, short length, which fits into short-step methods in a natural way. In Mizuno and Todd [10] a partial updating analysis for an "adaptive-step" path-following algorithm is given.

In Roos and Vial [14] a long-step path-following algorithm is proposed, which is in fact a natural implementation of the classical logarithmic barrier function approach. The number of reductions of the barrier parameter is O(L). Each reduction is followed by a series of inner steps, aiming at getting close to the analytic center associated with the current value of the penalty parameter. It was proved that at most O(nL) inner steps are needed. This means that the total complexity is $O(n^4L)$.

This result was also obtained independently by Gonzaga [7] in a more general approach. He also showed that if the barrier parameter is reduced by a factor $1 - (\nu/\sqrt{n})$, $\nu > 0$, then at most $O(\sqrt{n}L)$ reductions and at most $O(1)$ inner steps are needed. So, the total complexity of this variant is $O(n^{3.5}L)$.

In this paper we show that, using a Goldstein–Armijo rule to safeguard the linesearches of the barrier function, a $\sqrt{n}$ reduction in the complexity bounds can be obtained for both versions. As mentioned above, the partial updating analysis in [15], [6], [9], and [11] is based on steps of a short, fixed length, and so it cannot be used in long-step algorithms. The Goldstein–Armijo rule was introduced in the complexity analysis for Karmarkar's [8] projective algorithm by Anstreicher [1]. Anstreicher and Bosch [2] used the rule to improve the complexity bound for Ye [16] and Freund's [4] affine potential reduction algorithm.

Some new aspects are used in the analysis. We will use quadratic convergence in the neighbourhood of the central path to prove some properties of nearly centered points. This also enables us to improve Gonzaga's [7] results. Also, the reduction in

* Received by the editors July 16, 1990; accepted for publication (in revised form) March 27, 1991. This paper was first presented at the Second Asilomar Workshop on Progress in Mathematical Programming, Yale University, New Haven, Connecticut 06520.

† Department of Mathematics and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands.

‡ Department of Commercial and Industrial Economics, University of Geneva, 2 rue de Candolle, CH-1211 Geneva 4, Switzerland. This author's research was partially supported by the Fonds National Suisse de la Recherche Scientifique, under research grant # 1 467 0 86.

the barrier function value after an inner step is proved in a more natural way by using a Taylor expansion.

The paper is organized as follows. In §2 we prove some properties of (nearly) centered points. Then, in §3 we describe our algorithm and in §4 we prove that the algorithm reduces the complexity bound by a factor $\sqrt{n}$.

Notation. As far as notations are concerned, e shall denote the vector of all ones. 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.

2. Properties near the central path. We consider the linear programming problem:

$$\text{(P)} \qquad \min \left\{ c^T x : Ax = b, x \geq 0 \right\}.$$

Here 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 for (P) is:

$$\text{(D)} \qquad \max \left\{ b^T y : A^T y + s = c, s \geq 0 \right\}.$$

Without loss of generality, we assume that all the coefficients are integers. We shall denote by L the length of the input data of (P).

We make the standard assumption that the feasible set of (P) is bounded and has a nonempty relative interior. In order to simplify the analysis we shall also assume that A has full rank, though this assumption is not essential.

We consider the primal logarithmic barrier

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

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

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

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:

$$\begin{aligned} A^T y + s &= c, \qquad & s &\geq 0, \\ Ax &= b, & x &\geq 0, \\ Xs &= \mu e, \end{aligned} \tag{2}$$

where y and s are m- and n-dimensional vectors, respectively. It is well known that the necessary and sufficient first-order optimality conditions for the minimum of the dual logarithmic barrier function are also (2).

Let us denote the unique solution of this system by $(x(\mu), y(\mu), s(\mu))$. The primal and dual central path is defined as the solution set $x(\mu)$ and $y(\mu)$, respectively, for

$\mu > 0$. It is well known that the duality gap in $(x(\mu), y(\mu), s(\mu))$ satisfies $x(\mu)^T s(\mu) = n\mu$. Hence, if $\mu \to 0$, then $x(\mu)$ and $y(\mu)$ 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. These results also follow from Fiacco and McCormick [3]. We will give another simple proof.

LEMMA 2.1. *The objective $c^T x(\mu)$ of the primal problem* (P) *is monotonically decreasing and the objective $b^T y(\mu)$ of the dual problem* (D) *is monotonically increasing if μ decreases.*

Proof. Using the fact that $x(\mu)$ and $y(\mu)$ satisfy (2) and taking derivatives with respect to μ we obtain

$$\begin{aligned} A^T y' + s' &= 0, \\ Ax' &= 0, \\ Xs' + Sx' &= e, \end{aligned} \tag{3}$$

where the prime denotes the derivative with respect to μ. Now, using the relations of (2) and (3), we find

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

where the last equality follows because $(x')^T s' = -(Ax')^T y' = 0$. This proves the first part of the lemma.

To prove the second part of the lemma, we multiply the last equality of (3) by AS^{-1}:

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

which reduces to $AX^2 s' = b$. Taking the inner product with y' results in

$$b^T y' = (y')^T AX^2 s' = (A^T y')^T X^2 s' = -(s')^T X^2 s' \leq 0.$$

This proves the second part of the lemma. □

Roos and Vial [13] introduced the following measure of the distance of an interior feasible point to the central path:

$$\delta(x, \mu) := \min_{y,s} \left\{ \left\| \frac{Xs}{\mu} - e \right\| : A^T y + s = c \right\}. \tag{4}$$

The unique solution of the minimization problem in the definition of $\delta(x, \mu)$ is denoted by $(y(x, \mu), s(x, \mu))$. It can easily be verified that

$$x = x(\mu) \iff \delta(x, \mu) = 0 \implies s(x, \mu) = s(\mu).$$

The next lemma states that there is a close relationship between this measure and the projected Newton direction $p(x, \mu)$, which is obtained from (cf., e.g., [5])

$$\begin{pmatrix} X^{-2} & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \frac{y}{\mu} \end{pmatrix} = \begin{pmatrix} \frac{c}{\mu} - X^{-1}e \\ 0 \end{pmatrix}. \tag{5}$$

LEMMA 2.2. *For given x and μ, $\delta(x,\mu) = \|X^{-1}p(x,\mu)\|$.*
Proof. From (5) we can derive that $p(x,\mu) = Xq$, where

$$q = e - \frac{Xs}{\mu} \tag{6}$$

with

$$s = c - A^T y \tag{7}$$

and

$$y = (AX^2A^T)^{-1}AX(Xc - \mu e). \tag{8}$$

It can easily be verified that $s = s(x,\mu)$. Thus the lemma is proved. □

We note that a closed-form solution for $p(x,\mu)$ is given by

$$p(x,\mu) = -XP_{AX}\left(\frac{Xc}{\mu} - e\right), \tag{9}$$

where P_{AX} denotes the orthogonal projection on the null space of the matrix AX. Consequently, the projected Newton direction and the scaled projected gradient direction associated with f coincide. In the following we will write p instead of $p(x,\mu)$.

Now we will prove some fundamental lemmas for nearly centered points.

LEMMA 2.3. *If $\delta := \delta(x,\mu) \leq 1$, then $y := y(x,\mu)$ is dual-feasible. Moreover,*

$$\mu(n - \delta\sqrt{n}) \leq c^T x - b^T y \leq \mu(n + \delta\sqrt{n}).$$

Proof. By the definition of $s(x,\mu)$ we have

$$\left\|\frac{Xs(x,\mu)}{\mu} - e\right\| \leq 1.$$

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

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

Consequently, since $x^T s(x,\mu) = c^T x - b^T y$,

$$\mu(n - \delta\sqrt{n}) \leq c^T x - b^T y \leq \mu(n + \delta\sqrt{n}).$$ □

LEMMA 2.4. *If $\delta(x,\mu) < 1$, then $x^* = x + p$ is a strictly feasible point for* (P). *Moreover,*

$$\delta(x^*,\mu) \leq \delta(x,\mu)^2.$$

Proof. In the proof we make use of the vector t defined by

$$t = \frac{Xs(x,\mu)}{\mu}.$$

Note that

$$x^* = x + p = x + X(e - t) = 2x - Xt.$$

From $\delta(x, \mu) < 1$ we deduce that $\|t - e\| < 1$. Hence

$$2e - t > 0.$$

As a consequence one has, since $x > 0$,

$$x^* = 2x - Xt = X(2e - t) > 0.$$

So x^* is strictly feasible, because $Ax^* = Ax + Ap = b$.

The definition of $s(x^*, \mu)$ implies the following:

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

Using that $x^* = 2x - Xt$ we find

$$X^* X^{-1} t - e = (2X - XT) X^{-1} t - e = 2t - Tt - e = (E - T)(t - e).$$

Hence

$$\delta(x^*, \mu) \leq \max_i |1 - t_i| \|t - e\| \leq \delta(x, \mu)^2. \qquad \square$$

LEMMA 2.5. *If* $\delta := \delta(x, \mu) < 1$, *then*

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

Proof. The barrier function f is convex in x, whence

$$f(x + p, \mu) \geq f(x, \mu) + p^T \nabla f(x, \mu).$$

Now using (9) and $AXX^{-1}p = Ap = 0$,

$$\begin{aligned} p^T \nabla f(x, \mu) &= (X^{-1}p)^T X \nabla f(x, \mu) \\ &= (X^{-1}p)^T P_{AX} (X \nabla f(x, \mu)) \\ &= -(X^{-1}p)^T X^{-1} p \\ &= -\delta^2, \end{aligned} \tag{10}$$

where the last equality follows from Lemma 2.2. Substitution gives

$$f(x + p, \mu) \geq f(x, \mu) - \delta^2,$$

or equivalently,

$$f(x, \mu) - f(x + p, \mu) \leq \delta^2. \tag{11}$$

Now let $x^0 := x$ and let $x^0, x^1, x^2, \cdots$ denote the sequence of points obtained by repeating Newton steps, starting at x^0. By Lemma 2.4 we have

$$\delta(x^i, \mu) \leq \delta(x^0, \mu)^{2^i} = \delta^{2^i}. \tag{12}$$

So, using (11), we may write

$$\begin{aligned} f(x,\mu) - f(x(\mu),\mu) &= \sum_{i=0}^{\infty} \left(f(x^i,\mu) - f(x^{i+1},\mu) \right) \\ &\leq \sum_{i=0}^{\infty} \delta(x^i,\mu)^2 \\ &\leq \sum_{i=0}^{\infty} \delta^{2^{i+1}} \\ &\leq \frac{\delta^2}{1-\delta^2}. \end{aligned}$$

□

LEMMA 2.6. *If* $\delta := \delta(x,\mu) < 1$, *then*

$$|c^T x - c^T x(\mu)| \leq \frac{\delta(1+\delta)}{1-\delta} \mu\sqrt{n}.$$

Proof. From (10) we have $p^T \nabla f(x,\mu) = -\delta^2$. On the other hand,

$$\begin{aligned} p^T \nabla f(x,\mu) &= p^T \left(\frac{c}{\mu} - X^{-1} e \right) \\ &= \frac{c^T p}{\mu} - e^T X^{-1} p. \end{aligned}$$

So we have

$$\frac{c^T p}{\mu} = -\delta^2 + e^T X^{-1} p$$

or

$$c^T p = \mu(-\delta^2 + e^T X^{-1} p).$$

Using the Cauchy–Schwarz inequality, we obtain

$$|e^T X^{-1} p| \leq \|X^{-1} p\| \|e\| = \delta\sqrt{n},$$

where the last equality follows from Lemma 2.2. From this we deduce that

$$|c^T p| \leq \mu(\delta^2 + \delta\sqrt{n}) = \delta\left(1 + \frac{\delta}{\sqrt{n}}\right) \mu\sqrt{n} \leq \delta(1+\delta)\,\mu\sqrt{n}. \tag{13}$$

Again, let $x^0 := x$ and let $x^0, x^1, x^2, \cdots$ denote the sequence of points obtained by repeating Newton steps, starting at x^0. Then we have

$$\begin{aligned} |c^T x - c^T x(\mu)| &= \left| \sum_{i=0}^{\infty} \left(c^T x^i - c^T x^{i+1} \right) \right| \\ &\leq \sum_{i=0}^{\infty} |c^T p(x^i,\mu)| \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{i=0}^{\infty} \delta(x^i,\mu)(1+\delta(x^i,\mu))\mu\sqrt{n} \\
&\leq \sum_{i=0}^{\infty} \delta^{2^i}(1+\delta^{2^i})\mu\sqrt{n} \\
&\leq (1+\delta)\mu\sqrt{n}\sum_{i=0}^{\infty} \delta^{2^i} \\
&\leq \frac{\delta(1+\delta)}{1-\delta}\mu\sqrt{n},
\end{aligned}
$$

where the second inequality follows from (13) and the third inequality from (12). □

In [7] results similar to those in Lemmas 2.5 and 2.6 have been obtained in a different way for more centered x, namely, $\delta(x,\mu) \leq 0.1$. Our results hold for $\delta(x,\mu) < 1$. Moreover, for $\delta(x,\mu) \leq 0.1$, our bounds are tighter.

3. The revised long-step algorithm. Long-step barrier methods work as follows: fix μ, do linesearches along Newton directions until the iterate is in the neighbourhood of the current center, then reduce the barrier parameter, and repeat this process. Hence, at each iteration of these methods, one has to solve the linear system (5). Essentially this means that the $(m+n)\times(m+n)$ coefficient matrix of this system, denoted M,

$$
\begin{pmatrix} X^{-2} & A^T \\ A & 0 \end{pmatrix}
$$

has to be inverted. Hence, assuming that $m = O(n)$, at each iteration $O(n^3)$ arithmetic operations are needed. The matrices in two successive iterations differ only due to changes in X. Now consider the hypothetical case when only one entry of x changes. Then the new coefficient matrix M' differs from M only by a rank-one matrix. This makes it clear that we can write

$$
M' = M + uv^T,
$$

where u and v are suitable vectors. With the help of the Sherman–Morrison formula,

$$
(M+uv^T)^{-1} = M^{-1} - \frac{M^{-1}uv^TM^{-1}}{1+v^TM^{-1}u},
$$

the inverse of M' can be calculated from the inverse of M in only $O(n^2)$ arithmetic operations. If we require an exact solution of the system of equations we will generally need to make n such rank-one modifications. Therefore, $O(n^3)$ arithmetic operations will be needed at each iteration.

However, assume that instead of solving system (5) we solve

$$
(14) \qquad \begin{pmatrix} \tilde{X}^{-2} & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} -\tilde{p} \\ \frac{\tilde{y}}{\mu} \end{pmatrix} = \begin{pmatrix} \frac{c}{\mu} - X^{-1}e \\ 0 \end{pmatrix}, \qquad \tilde{y} \in \mathbb{R}^m,
$$

where $\tilde{X}$ is a working matrix closely related to X. Actually, the diagonal term $\tilde{x}_j$ of $\tilde{X}$ is updated during the inner iteration only if $\tilde{x}_j$ differs too much from x_j. If a limited number of components of $\tilde{x}$ are updated at a given iteration, a reduced computational cost can be achieved using the Sherman–Morrison formula. Of course one does not obtain the exact projected Newton direction p, but an approximation $\tilde{p}$ of it.

The purpose of this paper is to show that by performing a safeguarded linesearch along $\tilde{p}$, one can achieve the double goal of enforcing a significant decrease of the barrier function at each iteration, while maintaining a relatively small number of updates in the components of $\tilde{x}$, thereby achieving a computational saving in solving (14).

In order to work out these ideas we introduce the diagonal matrix D, with diagonal element d_j, defined by

$$\tilde{X} = XD.$$

Let $\rho > 1$ be some fixed number. The algorithm is designed so as to maintain the inequality

$$\frac{1}{\rho} \leq d_i \leq \rho, \qquad 1 \leq i \leq n. \tag{15}$$

Karmarkar [8] already used approximate solutions and partial updating to reduce the complexity bound for his algorithm. Using these approximate solutions for X, we will show that on the average only $\sqrt{n}$ rank-one modifications are needed, without increasing the complexity bound for the required number of iterations. This can be reached by submitting the linesearch to a Goldstein–Armijo condition.

To measure the distance to the central path, we shall now use a slightly different metric. We define

$$\tilde{\delta}(x,\mu) := \min_{y,s} \left\{ \left\| D\left(\frac{Xs}{\mu} - e\right) \right\| : A^T y + s = c \right\}. \tag{16}$$

Again, there is a close relationship between this measure and the approximate Newton direction $\tilde{p}$. It can easily be verified that

$$\tilde{\delta}(x,\mu) = \|\tilde{X}^{-1}\tilde{p}\|.$$

A closed-form solution for $\tilde{p}$ is

$$\tilde{p} = -\tilde{X} P_{A\tilde{X}} D \left(\frac{Xc}{\mu} - e\right).$$

It is clear from the definition that $\tilde{\delta}(x,\mu) = 0$ if and only if $x = x(\mu)$. In other words, we will have

$$\delta(x,\mu) = 0 \Longleftrightarrow \tilde{\delta}(x,\mu) = 0.$$

It is easy to verify that

$$\frac{1}{\rho}\delta(x,\mu) \leq \tilde{\delta}(x,\mu) \leq \rho\delta(x,\mu). \tag{17}$$

Consequently, if $\tilde{\delta}(x,\mu) \leq \frac{1}{\rho}$, then we have $\delta(x,\mu) \leq 1$, and then the lemmas proved in the previous section hold.

The Goldstein–Armijo condition can be formulated as follows:

$$\frac{\Delta f}{\alpha} \geq -\zeta \frac{df(x + \alpha\tilde{p},\mu)}{d\alpha}|_{\alpha=0}, \tag{18}$$

where Δf is the change in the barrier function value and $0 < \zeta < 1$. This condition is a well-known rule in nonlinear programming. It permits significant decreases of $f(x, \mu)$, but prevents excessively large steps. Note that we have

$$\text{(19)} \qquad \frac{df(x + \alpha\tilde{p}, \mu)}{d\alpha}|_{\alpha=0} = \left(\frac{c}{\mu} - X^{-1}e\right)^T \tilde{p} = -\|\tilde{X}^{-1}\tilde{p}\|^2 = -\tilde{\delta}(x, \mu)^2.$$

We will now describe the revised algorithm.

Revised long-step algorithm

Input:
μ_0 is the initial barrier value, $\mu_0 \leq 2^L$;
θ is the reduction parameter, $0 < \theta < 1$;
ρ is the coordinate update parameter, $\rho > 1$;
ζ is the Goldstein–Armijo factor, $\zeta \leq \frac{1}{2}$;
x^0 is a given interior feasible point such that $\tilde{\delta}(x^0, \mu_0) \leq \frac{1}{2\rho}$;

begin
  $x := x^0$; $\tilde{x} := x^0$; $\mu := \mu_0$;
  while $x^T s(x, \mu) > 2^{-L}$ **do**
  begin (outer step)
    while $\tilde{\delta}(x, \mu) > \frac{1}{2\rho}$ **do**
    begin (inner step)
      $D := \tilde{X}X^{-1}$
      $\tilde{\alpha} := \arg\min_{\alpha>0} \left\{f(x + \alpha\tilde{p}, \mu) : x + \alpha\tilde{p} > 0,\ \frac{\Delta f}{\alpha} \geq \zeta\tilde{\delta}(x, \mu)^2\right\}$
      $x := x + \tilde{\alpha}\tilde{p}$
      for $j := 1$ to n do if $(\tilde{x}_j / x_j) \notin (\frac{1}{\rho}, \rho)$ then $\tilde{x}_j := x_j$
    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., Renegar [12].

4. Convergence analysis of the revised long-step algorithm. We first give an upper bound for the total number of outer and inner iterations. Finally we derive an upper bound for the total number of coordinate updates of $\tilde{x}$.

Henceforth we shall denote $\{x^j\}$, $j = 0, 1, 2, \cdots$, the sequence of inner iterates and $\{\mu_k\}$, $k = 0, 1, 2, \cdots$, the sequence of parameter values during the successive outer iterations. Suppose that x^j is the current iterate when μ_k is calculated. Then set $p_k = j$. Take $p_0 = 0$. Then for any $j > 0$ there is a k such that $p_k < j \leq p_{k+1}$, and the value of μ used in the calculation of x^j is $\mu_k = (1 - \theta)^k \mu_0$.

THEOREM 4.1. *After at most $K = O(\frac{L}{\theta})$ outer iterations, the algorithm ends up with a primal and a dual solution such that $x^T s \leq 2^{-L}$.*

Proof. Since $\tilde{\delta}(x, \mu) \leq \frac{1}{2\rho}$ implies $\delta(x, \mu) \leq \frac{1}{2}$, we can derive an upper bound for the duality gap after K outer iterations from Lemma 2.3:

$$x^T s \leq \mu_K \left(n + \frac{1}{2}\sqrt{n}\right),$$

where $\mu_K = (1-\theta)^K \mu_0$. This means that $x^T s \le 2^{-L}$ certainly holds if

$$(1-\theta)^K \mu_0 \left(n + \frac{1}{2}\sqrt{n}\right) \le 2^{-L}.$$

Taking logarithms we obtain

$$K \ge \frac{L + \ln(n + \frac{1}{2}\sqrt{n}) + \ln \mu_0}{-\ln(1-\theta)}.$$

Since we have assumed that $\mu_0 \le 2^L$, and since $\theta \le -\ln(1-\theta)$, this certainly holds if $K = O(\frac{L}{\theta})$. □

The following two lemmas are needed to derive an upper bound for the number of inner iterations in each outer iteration. The first lemma estimates the difference in barrier function value between the starting point and end point of an outer iteration. The proof is in essence due to Gonzaga [7]. The second lemma states that a sufficient decrease in the barrier function value can be obtained by taking a step along the direction $\tilde{p}$. Moreover, it shows that for any $\zeta \le \frac{1}{2}$, the Goldstein–Armijo rule (18) can be enforced with the default value

$$\overline{\alpha} = \frac{1}{\rho(\tilde{\delta} + \rho)}.$$

Thus the algorithm is well defined.

LEMMA 4.2. *One has*

$$f(x^{p_k}, \mu_k) - f(x^{p_{k+1}}, \mu_k) \le \frac{\theta}{1-\theta} \left(\theta n + 3\sqrt{n}\right) + \frac{1}{3}.$$

Proof. The definition of $f(x,\mu)$, $x > 0$, implies that

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

Using this we obtain

$$\begin{aligned} f(x^{p_k},\mu_k) - f(x^{p_{k+1}},\mu_k) = f(x^{p_k},\mu_{k-1}) - f(x^{p_{k+1}},\mu_{k-1}) \\ + \frac{\theta}{1-\theta} \frac{1}{\mu_{k-1}} (c^T x^{p_k} - c^T x^{p_{k+1}}). \end{aligned} \tag{20}$$

First note that because x^{p_k} and $x^{p_{k+1}}$ are approximately centered with respect to $x(\mu_{k-1})$ and $x(\mu_k)$, respectively, using Lemma 2.6 and $\delta = \frac{1}{2}$ for the first and Lemma 2.1 for the second inequality, we find

$$\begin{aligned} c^T x^{p_k} - c^T x^{p_{k+1}} &\le c^T x(\mu_{k-1}) + \frac{3}{2}\mu_{k-1}\sqrt{n} - c^T x(\mu_k) + \frac{3}{2}\mu_k \sqrt{n} \\ &= c^T x(\mu_{k-1}) - c^T x(\mu_k) + \frac{3}{2}(2-\theta)\mu_{k-1}\sqrt{n} \end{aligned}$$

$$\begin{aligned}
&\leq \left(c^T x(\mu_{k-1}) - b^T y(\mu_{k-1})\right) \\
&\quad - \left(c^T x(\mu_k) - b^T y(\mu_k)\right) + 3\mu_{k-1}\sqrt{n} \\
&= \mu_{k-1} n - \mu_k n + 3\mu_{k-1}\sqrt{n} \\
&= \theta \mu_{k-1} n + 3\mu_{k-1}\sqrt{n} \\
&= \mu_{k-1}\left(\theta n + 3\sqrt{n}\right).
\end{aligned}$$

Second, using the fact that $x(\mu_{k-1})$ minimizes $f(x(\mu_{k-1}), \mu_{k-1})$ and using Lemma 2.5 and $\delta = \frac{1}{2}$, we obtain

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

Hence, substitution of the last two inequalities into (20) yields

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

This proves the lemma. □

The following lemma will be used in Lemma 4.4.

LEMMA 4.3. *For $v \geq 0$ we have:*

$$\ln(1+v) \leq v - \frac{v^2}{2(1+v)}.$$

Proof. First note that $-\ln(1+v) = \ln(1 - \frac{v}{1+v})$. Now using Karmarkar's [8] well-known inequality, we have for $v \geq 0$

$$\ln\left(1 - \frac{v}{1+v}\right) \geq -\frac{v}{1+v} - \frac{1}{2}\frac{\left(\frac{v}{1+v}\right)^2}{1 - \frac{v}{1+v}} = -\frac{v}{1+v} - \frac{v^2}{2(1+v)}.$$

This means that

$$\ln(1+v) \leq \frac{v}{1+v} + \frac{v^2}{2(1+v)} = v - \frac{v^2}{2(1+v)}.$$ □

LEMMA 4.4. *Let $\tilde{\delta} := \tilde{\delta}(x, \mu)$, $\overline{\alpha} := [\rho(\tilde{\delta} + \rho)]^{-1}$. Then*

$$\bar{\Delta} f := f(x, \mu) - f(x + \overline{\alpha}\tilde{p}, \mu) \geq \frac{\tilde{\delta}}{\rho} - \ln\left(1 + \frac{\tilde{\delta}}{\rho}\right).$$

Moreover,

$$\frac{\bar{\Delta} f}{\overline{\alpha}} \geq \zeta \tilde{\delta}^2 \quad \textit{for } \zeta \leq \frac{1}{2}.$$

Proof. We write down the Taylor expansion for f:

$$f(x+\alpha\tilde{p},\mu) = f(x,\mu) + \alpha\tilde{p}^T\nabla f(x,\mu) + \frac{1}{2}\alpha^2\tilde{p}^T\nabla^2 f(x,\mu)\tilde{p} + \sum_{k=3}^{\infty} t_k,$$

where t_k again denotes the kth-order term in the Taylor expansion.

Using the fact that

$$t_k = \frac{(-\alpha)^k}{k}\sum_{i=1}^{n} x_i^{-k}\tilde{p}_i^k,$$

we derive by the definition of $\tilde{x}$ and $\tilde{\delta}$,

$$|t_k| \le \frac{\alpha^k}{k}\sum_{i=1}^{n}|x_i^{-1}\tilde{p}_i|^k = \frac{\alpha^k}{k}\sum_{i=1}^{n}|d_i\tilde{x}_i^{-1}\tilde{p}_i|^k \le \frac{\alpha^k}{k}\left(\sum_{i=1}^{n}|d_i\tilde{x}_i^{-1}\tilde{p}_i|^2\right)^{k/2} = \frac{\alpha^k}{k}\rho^k\tilde{\delta}^k.$$

Further,

$$\tilde{p}^T\nabla^2 f(x,\mu)\tilde{p} = \tilde{p}^T X^{-2}\tilde{p} = \|D\tilde{X}^{-1}\tilde{p}\|^2 \le \rho^2\tilde{\delta}^2,$$

and, using the fact that $A\tilde{X}\tilde{X}^{-1}\tilde{p} = A\tilde{p} = 0$,

$$\begin{aligned}\tilde{p}^T\nabla f(x,\mu) &= (\tilde{X}^{-1}\tilde{p})^T\tilde{X}\nabla f(x,\mu)\\ &= (\tilde{X}^{-1}\tilde{p})^T P_{A\tilde{X}}(\tilde{X}\nabla f(x,\mu))\\ &= -(\tilde{X}^{-1}\tilde{p})^T\tilde{X}^{-1}\tilde{p}\\ &= -\tilde{\delta}^2.\end{aligned}$$

So if $\alpha\rho\tilde{\delta} < 1$, we find

$$\begin{aligned}f(x+\alpha\tilde{p},\mu) &\le f(x,\mu) - \alpha\tilde{\delta}^2 + \frac{1}{2}\alpha^2\rho^2\tilde{\delta}^2 + \sum_{k=3}^{\infty}\frac{\alpha^k}{k}\rho^k\tilde{\delta}^k\\ &= f(x,\mu) - \alpha\tilde{\delta}^2 - \ln(1-\alpha\rho\tilde{\delta}) - \alpha\rho\tilde{\delta}.\end{aligned}$$

Hence

$$f(x,\mu) - f(x+\alpha\tilde{p},\mu) \ge \alpha\tilde{\delta}^2 + \alpha\rho\tilde{\delta} + \ln(1-\alpha\rho\tilde{\delta}).$$

The right-hand side is maximal if $\alpha = \overline{\alpha} = [\rho(\tilde{\delta}+\rho)]^{-1}$. Note that $\overline{\alpha}\rho\tilde{\delta} < 1$. Substitution of this value finally gives

$$\bar{\Delta} f \ge \frac{\tilde{\delta}^2}{\rho(\tilde{\delta}+\rho)} + \frac{\rho\tilde{\delta}}{\rho(\tilde{\delta}+\rho)} + \ln\left(1 - \frac{\rho\tilde{\delta}}{\rho(\tilde{\delta}+\rho)}\right) = \frac{\tilde{\delta}}{\rho} - \ln\left(1+\frac{\tilde{\delta}}{\rho}\right). \tag{21}$$

This proves the first part of the lemma. The second part follows immediately from (21) and Lemma 4.3:

$$\bar{\Delta} f \ge \frac{\tilde{\delta}}{\rho} - \ln\left(1+\frac{\tilde{\delta}}{\rho}\right) \ge \frac{\tilde{\delta}}{\rho} - \frac{\tilde{\delta}}{\rho} + \frac{\tilde{\delta}^2/\rho^2}{2(1+\frac{\tilde{\delta}}{\rho})} = \frac{\tilde{\delta}^2}{2\rho(\tilde{\delta}+\rho)} = \frac{1}{2}\overline{\alpha}\tilde{\delta}^2. \tag{22}$$

□

THEOREM 4.5. *The number of inner iterations for each outer iteration, denoted by P, is bounded by*

$$P \leq 12\frac{\theta\rho^4}{1-\theta}\left(\theta n + 3\sqrt{n}\right) + 4\rho^4.$$

Proof. Let us consider the $(k+1)$st outer iteration. Let P denote the number of inner iterations. For each inner iteration we know, by the definition of $\tilde{\alpha}$ and (22), that the decrease in the barrier function value is larger than

$$\frac{\tilde{\delta}^2}{2\rho(\tilde{\delta}+\rho)}.$$

Since this expression is an increasing function of $\tilde{\delta}$, and since during each iteration $\tilde{\delta} \geq \frac{1}{2\rho}$, we have

$$\frac{\tilde{\delta}^2}{2\rho(\tilde{\delta}+\rho)} \geq \frac{1}{12\rho^4}.$$

Consequently, we have

$$f(x^{p_{k+1}}, \mu_k) \leq f(x^{p_k}, \mu_k) - \frac{1}{12\rho^4}P.$$

Equivalently,

$$\frac{1}{12\rho^4}P \leq f(x^{p_k}, \mu_k) - f(x^{p_{k+1}}, \mu_k).$$

Now using Lemma 4.2, the theorem follows. □

Consequently, using an additional Goldstein–Armijo rule and approximate solutions does not influence the order of the total number of outer and inner iterations.

The last theorem will give an upper bound for the total number of coordinate updates in $\tilde{x}$. For the proof of this theorem we make use of some results obtained by Anstreicher [1]. The following lemma will be used in the theorem.

LEMMA 4.6. *Let $w \in \mathbb{R}$, $0 < w < 1$, and $v \in \mathbb{R}$, $v \geq w$. Then*

$$|\ln v| \leq \frac{|1-v||\ln w|}{1-w}.$$

Proof. Defining

$$f(z) := \begin{cases} \frac{\ln z}{z-1} & \text{if } z \neq 1, \\ 1 & \text{if } z = 1, \end{cases}$$

it is easy to see that $f(z)$ is monotonically decreasing and positive for $z \in (0, \infty)$. Hence

$$\left|\frac{\ln v}{v-1}\right| \leq \left|\frac{\ln w}{w-1}\right| = \frac{|\ln w|}{1-w}.$$

This implies the lemma. □

THEOREM 4.7. *The total number M of coordinate updates of $\tilde{x}$ up to the last inner iteration N is bounded by*

$$M \leq \frac{2\rho^3\sqrt{n}}{\zeta(\rho-1)}\left(\frac{\theta n + 3\sqrt{n}}{1-\theta} + \frac{1}{3\theta}\right) O(L).$$

Proof. Let k_1 be an iteration at which an update of $\tilde{x}_i$ is performed. Let $k_2 > k_1$ be the first iteration at which $\tilde{x}_i$ is updated again. Then we have

$$\begin{aligned}\prod_{j=k_1+1}^{k_2} \max\left(\frac{x_i^j}{x_i^{j-1}}, \frac{x_i^{j-1}}{x_i^j}\right) &\geq \max\left(\prod_{j=k_1+1}^{k_2} \frac{x_i^j}{x_i^{j-1}}, \prod_{j=k_1+1}^{k_2} \frac{x_i^{j-1}}{x_i^j}\right) \\ &= \max\left(\frac{x_i^{k_2}}{x_i^{k_1}}, \frac{x_i^{k_1}}{x_i^{k_2}}\right) \\ &\geq \rho.\end{aligned}$$

Taking logarithms and defining

$$r_i^j := 1 + \tilde{\alpha}_j (x_i^j)^{-1} \tilde{p}_i^j = \frac{x_i^{j+1}}{x_i^j},$$

we obtain

$$\ln \rho \leq \sum_{j=k_1}^{k_2-1} |\ln r_i^j|. \tag{23}$$

Let

$$\hat{r}_i^j = \max\left\{r_i^j, \frac{1}{\rho}\right\}.$$

Inequality (23) can be sharpened to

$$\ln \rho \leq \sum_{j=k_1}^{k_2-1} |\ln \hat{r}_i^j|. \tag{24}$$

To prove (24), first assume that for some ℓ, $k_1 \leq \ell \leq k_2 - 1$, $r_i^j < \frac{1}{\rho}$. Then

$$\ln \rho = |\ln \hat{r}_i^\ell| \leq \sum_{j=k_1}^{k_2-1} |\ln \hat{r}_i^j|.$$

Otherwise, $\hat{r}_i^j = r_i^j$, $k_1 \leq j \leq k_2 - 1$, and (24) holds because of (23). Hence (24) has been proved.

We deduce from (24) a bound on the total number m_i of updates of coordinate i of $\tilde{x}$:

$$m_i \ln \rho \leq \sum_{j=0}^{N-1} |\ln \hat{r}_i^j|.$$

Consequently the total number of coordinate updates is bounded by

$$M \ln \rho \leq \sum_{j=0}^{N-1} \sum_{i=1}^{n} |\ln \hat{r}_i^j|. \tag{25}$$

In view of Lemma 4.6, with $v = \hat{r}_i^k$ and $w = \frac{1}{\rho}$,

$$|\ln \hat{r}_i^j| \leq \frac{\ln \rho}{1 - \frac{1}{\rho}} |1 - \hat{r}_i^j|. \tag{26}$$

Since $\hat{r}_i^j = r_i^j$ if $r_i^j \geq \frac{1}{\rho}$, and $\hat{r}_i^j > r_i^j$ if $r_i^j < \frac{1}{\rho}$, we always have

$$|1 - \hat{r}_i^j| \leq |1 - r_i^j| = \tilde{\alpha}_j |(x_i^j)^{-1} \tilde{p}_i^j|. \tag{27}$$

Substitution of (26) and (27) into (25) gives

$$M \leq \frac{\rho}{\rho - 1} \sum_{j=0}^{N-1} \tilde{\alpha}_j \sum_{i=1}^{n} |(x_i^j)^{-1} \tilde{p}_i^j|.$$

From the inequality between the l_1 and l_2 norms,

$$\sum_{i=1}^{n} |(x_i^j)^{-1} \tilde{p}_i^j| \leq \sqrt{n} \|(X^j)^{-1} \tilde{p}^j\| \leq \rho \sqrt{n} \|(\tilde{X}^j)^{-1} \tilde{p}^j\|.$$

Hence

$$M \leq \frac{\rho^2 \sqrt{n}}{\rho - 1} \sum_{j=0}^{N-1} \tilde{\alpha}_j \|(\tilde{X}^j)^{-1} \tilde{p}^j\|. \tag{28}$$

Since the Goldstein–Armijo condition is satisfied at each inner iteration, for any j and k such that $p_k < j \leq p_{k+1}$ (we will write $k(j)$ instead of k to denote its dependence on j),

$$\tilde{\alpha}_j \leq \frac{f(x^j, \mu_{k(j)}) - f(x^{j+1}, \mu_{k(j)})}{\zeta \|(\tilde{X}^j)^{-1} \tilde{p}^j\|^2}. \tag{29}$$

Substituting this into (28) we obtain

$$M \leq \frac{\rho^2 \sqrt{n}}{\zeta(\rho - 1)} \sum_{j=0}^{N-1} \frac{f(x^j, \mu_{k(j)}) - f(x^{j+1}, \mu_{k(j)})}{\|(\tilde{X}^j)^{-1} \tilde{p}^j\|}.$$

Since $\|(\tilde{X}^j)^{-1} \tilde{p}^j\| \geq \frac{1}{2\rho}$, this implies

$$\begin{aligned} M &\leq \frac{2\rho^3 \sqrt{n}}{\zeta(\rho - 1)} \sum_{j=0}^{N-1} (f(x^j, \mu_{k(j)}) - f(x^{j+1}, \mu_{k(j)})) \\ &= \frac{2\rho^3 \sqrt{n}}{\zeta(\rho - 1)} \sum_{k=0}^{K-1} (f(x^{p_k}, \mu_k) - f(x^{p_{k+1}}, \mu_k)). \end{aligned}$$

Now using Theorem 4.1 and Lemma 4.2 we obtain

$$M \leq \frac{2\rho^3\sqrt{n}}{\zeta(\rho-1)}\left(\frac{\theta n + 3\sqrt{n}}{1-\theta} + \frac{1}{3\theta}\right)O(L). \qquad \square$$

Theorems 4.1 and 4.5 imply that N, the total number of inner iterations needed by the algorithm, is bounded by

$$N \leq \left(12\frac{\rho^4}{1-\theta}\ (\theta n + 3\sqrt{n}) + 4\frac{\rho^4}{\theta}\right)O(L).$$

The total number of arithmetic operations in each iteration, aside from the work due to coordinate updates, is $O(n^2)$. The same amount of work must be done for one coordinate update. Consequently, the total number of arithmetic operations needed by the algorithm is $(N + M)O(n^2)$.

COROLLARY 4.8.

- *If $0 < \theta < 1$, independent of n and L, then the total number of iterations is bounded by $O(nL)$ and the total number of coordinate updates by $O(n^{1.5}L)$. Consequently, the total complexity is $O(n^{3.5}L)$.*
- *If $\theta = \nu/\sqrt{n}$, $\nu > 0$ and independent of n and L, then the total number of iterations is bounded by $O(\sqrt{n}L)$ and the total number of coordinate updates by $O(nL)$. Consequently, the total complexity is $O(n^3L)$.*

Acknowledgments. We would like to thank Kurt Anstreicher for the valuable suggestions for the proof of Theorem 4.7 that he made during a visit to Delft, from May 14–24, 1990. Thanks are also due to the two referees for their comments.

REFERENCES

[1] K. M. ANSTREICHER (1991), *A standard form variant, and safeguarded linesearch, for the modified Karmarkar algorithm*, Math. Programming, 47, pp. 337–351.

[2] K. M. ANSTREICHER AND R. A. BOSCH (1988), *Long steps in an $O(n^3L)$ algorithm for linear programming*, Preprint, Yale School of Organization and Management, New Haven, CT.

[3] A. V. FIACCO AND G. P. MCCORMICK (1968), *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, New York.

[4] R. M. FREUND (1988), *Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function*, Preprint, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

[5] P. E. GILL, W. MURRAY, M. A. SAUNDERS, J. A. TOMLIN, AND M. H. WRIGHT (1986), *On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method*, Math. Programming, 36, pp. 183–209.

[6] C. C. GONZAGA (1989), *An algorithm for solving linear programming problems in $O(n^3L)$ operations*, in Progress in Mathematical Programming, Interior Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, pp. 1–28.

[7] ——— (1989), *Large-steps path-following methods for linear programming, Part* I: *Barrier function method*, Report ES–210/89, Department of Systems Engineering and Computer Sciences, COPPE–Federal University of Rio de Janeiro, Rio de Janeiro, Brasil; also in SIAM J. Optimization, 1 (1991), pp. 268–279.

[8] N. KARMARKAR (1984), *A new polynomial-time algorithm for linear programming*, Combinatorica, 4, pp. 373–395.

[9] M. KOJIMA, S. MIZUNO, AND A. YOSHISE (1989), *A polynomial-time algorithm for a class of linear complementarity problems*, Math. Programming, 44, pp. 1–26.

[10] S. MIZUNO AND M. J. TODD (1989), *An $O(n^3L)$ long step path following algorithm for a linear complementarity problem*, Res. Report 23, Tokyo Institute of Technology, Tokyo, Japan.

[11] R. D. C. MONTEIRO AND I. ADLER (1989), *Interior path following primal-dual algorithms, Part II: Linear programming*, Math. Programming, 44, pp. 27–41.

[12] J. RENEGAR (1988), *A polynomial-time algorithm, based on Newton's method, for linear programming*, Math. Programming, 40, pp. 59–93.

[13] C. ROOS AND J.-PH. VIAL (1988), *A polynomial method of approximate centers for linear programming*, Math. Programming, to appear.

[14] ——— (1990), *Long steps with the logarithmic penalty barrier function in linear programming*, in Economic Decision-Making: Games, Economics and Optimization, dedicated to Jacques H. Drèze; J. Gabszevwicz, J.-F. Richard, and L. Wolsey, eds., Elsevier Science Publisher B.V., Amsterdam, pp. 433–441.

[15] P. M. VAIDYA (1990), *An algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ arithmetic operations*, Math. Programming, 47, pp. 175–201.

[16] Y. YE (1991), *An $O(n^3L)$ potential reduction algorithm for linear programming*, Math. Programming, 50, pp. 239–258.