RAIRO. Recherche opérationnelle

D. Den Hertog
C. Roos
T. Terlaky

Inverse barrier methods for linear programming

RAIRO. Recherche opérationnelle, tome 28, nº 2 (1994), p. 135-163
<http://www.numdam.org/item?id=RO_1994__28_2_135_0>

Recherche opérationnelle/Operations Research
(vol. 28, n° 2, 1994, p. 135 à 163)

INVERSE BARRIER METHODS FOR LINEAR PROGRAMMING (*)

by D. DEN HERTOG [1], C. ROOS [1] and T. TERLAKY [1]

Communicated by Pierre TOLLA

Abstract. – *In the recent interior point methods for linear programming much attention has been given to the logarithmic barrier method. In this paper we will analyse the class of inverse barrier methods for linear programming, in which the barrier is* $\sum x_i^{-r}$, *where* $r > 0$ *is the rank of the barrier.*

There are many similarities with the logarithmic barrier method. The minima of an inverse barrier function for different values of the barrier parameter define a "central path" dependent on r, called the r-path of the problem. For $r \downarrow 0$ *this path coincides with the central path determined by the logarithmic barrier function. We introduce a metric to measure the distance of a feasible point to a point on the path. We prove that in a certain region around a point on the path the Newton process converges quadratically. Moreover, outside this region, taking a step into the Newton direction decreases the barrier function value at least with a constant.*

We will derive upper bounds for the total number of iterations needed to obtain an ϵ*-optimal solution. Unfortunately, these bounds are not polynomial in the input length. Only if the rank r goes to zero we get a polynomiality result, but then we are actually working with the logarithmic barrier method.*

Keywords: Linear programming, interior point method, logarithmic barrier function, inverse barrier function.

Résumé. – *Dans les méthodes récentes par points intérieurs, l'attention a surtout porté sur la méthode de barrière logarithmique. Nous analysons dans cet article la classe des méthodes de barrière inverse, où la barrière est* $\sum x_i^{-r}$, *où* $r > 0$ *est le rang de la barrière.*

Il y a beaucoup de similitudes avec la méthode de barrière logarithmique. Les minima d'une fonction de barrière inverse pour les différentes valeurs du paramètre de barrière définissent un « chemin central » dépendant de r, appelé le r-chemin du problème. Pour $r \downarrow 0$ *ce chemin coïncide avec le chemin central défini par la fonction de barrière logarithmique ; nous introduisons un* ϵ *métrique pour mesurer la distance d'un point réalisable à un point du chemin. Nous montrons que dans une certaine région autour d'un point la méthode Newton possède la convergence quadratique. En outre, en dehors de cette région, un pas dans la direction de Newton décroît la fonction de barrière d'au moins une constante.*

(*) Received October 1991.

[1] Faculty of Technical Mathematics and Informatics, Mekelweg 4, 2628 CD Delft, Netherlands.
This work is completed under the support of a reasearch grant of Shell.

The third author is on leave from the Eötvös University, Budapest and partially supported by OTKA No. 2116.

Recherche opérationnelle/Operations Research, 0399-0559/94/02/$ 4.00

Nous donnons des bornes supérieures du nombre total d'itérations nécessaires pour obtenir une solution optimale à ϵ près. Malheureusement, ces bornes ne sont pas polynomiales par rapport à la longueur des données. C'est seulement si le rang r tend vers zéro que nous obtenons un résultat polynomial, mais alors nous travaillons en réalité avec la méthode de barrière logarithmique.

Mots clés : Programmation linéaire, méthode intérieure, fonction de barrière logarithmique, fonction de barrière inverse.

1. INTRODUCTION

Karmarkar's [11] projective method has initiated the fast developing field of interior point methods for linear programming. Since then many different interior point methods have been proposed, which appear to have several similarities if one analyses them more carefully.

A well-known interior point method is the classical logarithmic barrier function method, proposed by Frisch [7] and further developed by Fiacco and McCormick [5]. In this method the nonnegativity constraints $x_i \geq 0$ of the linear programming problem are replaced by an additional term $-\mu \sum \ln x_i$ in the object function. Now letting the barrier parameter μ go to zero, one can prove convergence. Moreover, in the case of linear programming polynomiality can be proved. See Gonzaga [8] and [9], Roos and Vial [18] and [19], and Den Hertog, Roos and Vial [3].

Besides this logarithmic barrier method there are other well-known barrier methods. Carroll [1] proposed the socalled inverse barrier method, in which the barrier function is $\sum x_i^{-1}$. This method was further developed by Fiacco and McCormick [5] and implemented by McCormick *et al.* [13] in the SUMT-2 and SUMT-3 codes. Kowalik [12] described the quadratic inverse barrier method, in which the barrier function is $\sum x_i^{-2}$. Fletcher and McCann [6] elaborated this idea.

More recently, Ericsson [4] studied the socalled entropy barrier method. Polyak [16] studied some modifications of classical barrier methods. But both didn't obtain upper bounds for the number of iterations: only convergence has been proved.

In the recent papers on interior point methods for linear programming only the logarithmic barrier method has been studied thoroughly. So the natural question arises if it is possible to construct effective methods based on other barrier functions. Is it also possible to prove polynomiality for such a method?

In this paper we will analyse the classical inverse barrier function method, which uses $\sum x_i^{-r}$ $(r > 0)$ as a barrier. The approach resembles the logarithmic barrier approach in Den Hertog, Roos and Vial [3]. The minima of the barrier function again form a path, called the r-path, which is of course different form the "central path" determined by the logarithmic barrier function. As in the logarithmic barrier case we define a metric to measure the distance of a feasible point to this path, and we will prove that within some region the Newton method is quadratically convergent. Outside this region we can prove that taking a step in the Newton direction results into a decrease in the barrier function value. Of course, the present Newton direction is different from the Newton direction for the logarithmic barrier function.

Unfortunately, we are not able to prove polynomiality. Under some assumptions we will prove that the total number of iterations to obtain an ϵ-optimal solution is bounded by

$$O\left(\sqrt{n}\left(\frac{n}{\epsilon}\right)^{r/2}\ln\left(\frac{n}{\epsilon}\right)\right) \quad \text{or} \quad O\left(n\left(\frac{n}{\epsilon}\right)^{r}\ln\left(\frac{n}{\epsilon}\right)\right)$$

dependent on the updating of the barrier parameter. In each iteration a linear system has to be solved. If r goes to zero then we will show that the method reduces to the logarithmic barrier method, which has the complexity bounds

$$O\left(\sqrt{n}\ln\left(\frac{n}{\epsilon}\right)\right) \quad \text{and} \quad O\left(n\ln\left(\frac{n}{\epsilon}\right)\right)$$

(*see* Gonzaga [9] and Den Hertog, Roos and Vial [3]).

The paper is organized as follows. In Section 2 we do the preliminary work. Then, in Section 3 we prove some properties of nearly centered points and we derive a lower bound for the decrease in the barrier function value after a step in the Newton direction. In Section 4 we will state our algorithms and derive upper bounds for the total number of iterations. Finally, in Section 5 we end up with some concluding remarks.

Notation: 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. The identity matrix is denots by I and the all-one vector by e. Finally $\|x\|$ denotes the l_2 norm and $\|x\|_\infty$ the l_∞ norm of x.

2. INVERSE BARRIER FUNCTION AND r-PATHS

Consider the primal linear programming problem:

$$(P) \quad \min\{c^T x : A x = b,\ x \geq 0\}.$$

Here A is an $m \times n$ matrix, 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:

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

We assume that the feasible set of (*P*) is bounded and has a nonempty relative interior. In fact we make the stronger assumption that all primal feasible variables are bounded form above by 1. It is easy to see that this can be accomplished by scaling the original problem. (*See* Appendix 3.) This assumption is trivially fulfilled if one of the equality constraints $A x = b$ is the "simplex constraint" $e^T x = 1$, as Karmarkar [11] assumed in his paper. In order to simplify the analysis we shall also assume that A has full rank, though this assumption is not essential.

Since the primal feasible region is bounded, the dual feasible region is unbounded. (*See* Appendix 4.) But we will show that the dual slack variables which we need in the analysis lie in a bounded region. (*See* Appendix 2.) Hence, by scaling, we can assume that those variables are bounded from above by 1 too. (*See* Appendix 3.)

To the primal problem we associate the following inverse barrier function

$$f_r(x, \mu) = \frac{c^T x}{\mu} + \frac{1}{r} \sum_{i=1}^{n} \frac{1}{x_i^r}. \tag{1}$$

We call r the rank of the barrier function.

The first and second order derivatives of this function are:

$$g := \nabla f_r(x, \mu) = \frac{c}{\mu} - X^{-r-1} e,$$
$$H := \nabla^2 f_r(x, \mu) = (r+1) X^{-r-2}.$$

Hence it follows that $f_r(x, \mu)$ is strictly convex on its relative interior. This means that there exists a unique minimum. The Karush-Kuhn-Tucker conditions for this minimum are

$$\left.\begin{array}{rl} A^T y + s = c, & s \geq 0, \\ A x = b, & x \geq 0, \\ X^{r+1} s = \mu e. & \end{array}\right\} \tag{2}$$

The unique solution of this system of equations is denoted by $x(\mu)$ and $(y(\mu), s(\mu))$. The set of solutions for $\mu > 0$ is called the r-path. Note that the 0-path is exactly the central path which has been analysed by many researchers during the last years (e. g. Megiddo [14]). The minimum of $f_r(x, \mu)$ for $\mu = \infty$ is called the r-center of the polytope. Hence, the r-path starts at the r-center.

It is straightforward to generalize the inverse barrier function $f_r(x, \mu)$ for the dual problem. In Figure 1 we have drawn several r-paths for 2-dimensional dual problems. Figure 1.a shows several r-paths for the following problem:

$$\begin{aligned} \max -y_1 + \frac{5}{2}\, y_2 \\ -1 \leq y_1,\, y_2 \leq 1 \\ y_1 - y_2 \leq 1 \\ -\frac{3}{5}\, y_1 + \frac{4}{5}\, y_2 \leq \frac{3}{5}. \end{aligned}$$

Figure 1.b shows r-paths for the same problem, with the objective replaced by $-3\, y_1 + 4.1\, y_2$.

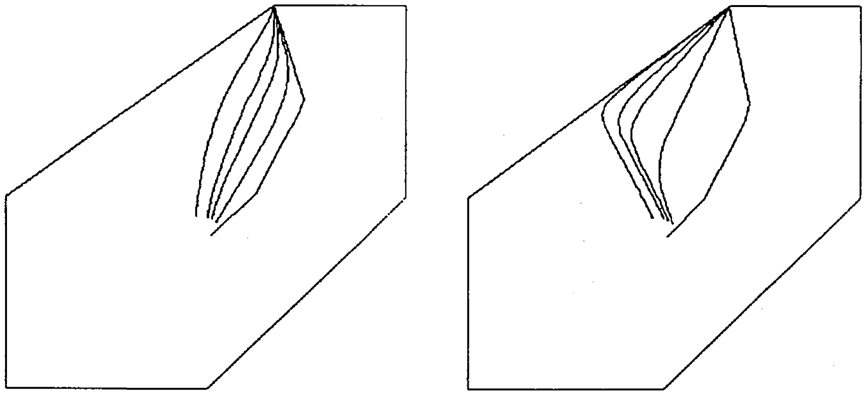
Note that, loosely speaking, coming into the neigborhood of the boundary is more penalized if the rank is large. This effect can indeed be seen in Figure 1. In Figure 1.b, for example, a large part of the 0-path is very close to one of the constraints, whereas the higher order paths approach the optimum more from the interior! So, using higher order barriers might be an advantage for (almost) degenerate problems.

Note that the "limit path" (the path for $r \to \infty$) is piecewise linear. Moreover, this limit path is only dependent on the constraints and the optimum. For example, in Figure 1.a and 1.b the 50-path for both problems (which only differ in the objective) are almost the same.

Without going into details we note that (under nondegeneracy) this piecewise linear path is in fact Huard's [10] path of centers for his special "quasi distance" function. This path was recently further analysed by White [21]. Moreover, if the columns of A have length 1, then this path also coincides with the "locus of centers" studied by Tamura *et al.* [20]. They proved that following this piecewise linear path corresponds to a primal simplex method.

The next lemma deals with monotonicity of the objective along the r-path.

LEMMA 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*



a b

Figure 1. — *r*-Paths for $r = 0, 1, 2, 5$, and 50, from the left to the right respectively.

increasing if μ decreases. Moreover, if $c^T x$ is not constant on the primal feasible region, then $c^T x(\mu)$ is strictly decreasing, and if $b \neq 0$ then $b^T y(\mu)$ is strictly increasing.

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

$$\left.\begin{aligned} A^T y' + s' &= 0, \\ A x' &= 0, \\ X^{r+1} s' + (r+1) X^r S x' &= e, \end{aligned}\right\} \tag{3}$$

where primes refer to derivatives with respect to μ. Now, using the relations of (2) and (3), we find

$$\begin{aligned} c^T x' &= (x')^T (s + A^T y) = (x')^T s = e^T (S x') \\ &= (X^{r+1} s' + (r+1) X^r S x')^T S x' \\ &= \mu (x')^T s' + (r+1)(x')^T S^2 X^r x' = (r+1)(x')^T S^2 X^r x' \geq 0, \end{aligned}$$

where the last equality follows because $(x')^T s' = -(Ax')^T y' = 0$. This proves that $c^T x(\mu)$ is decreasing along the path. Moreover, since $S^2 X^r$ is positive definite, $c^T x' = 0$ if and only if $x' = 0$. If $x' = 0$ then from (3) we have $X^{r+1} s' = e$ or $\mu s' = s$. This means that

$$c = A^T y + s = A^T y + \mu s' = A^T (y - \mu y').$$

So, in this case c is in the row space of A, which means that $c^T x$ is constant on the feasible region. Consequently, $c^T x(\mu)$ is strictly decreasing if $c^T x$ is not constant on the feasible region.

Now we multiply the last equality of (3) by $AS^{-1} X^{-r}$:

$$AS^{-1} X s' + (r+1) A x' = AS^{-1} X^{-r} e,$$

which reduces to $AX^{r+2} s' = b$. Taking the inner product with y' results into

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

This proves that $b^T y(\mu)$ is increasing along the path. Moreover, since X^{r+2} is positive definite, $b^T y' = 0$ if an only if $s' = 0$. If $s' = 0$ then from (3) we have $(r+1) X^r Sx' = e$ or $\mu (r+1) x' = x$. This means that

$$b = A x = \mu (r+1) A x' = 0.$$

Consequently, $b^T y(\mu)$ is strictly increasing along the path if $b \neq 0$. This completes the proof of the lemma. □

3. PROPERTIES NEAR THE r-PATH

We introduce the following measure for the distance of an interior feasible point to the r-path:

$$\delta(x, \mu) := \min_{y, s} \left\{ \left\| X^{-r/2} \left(\frac{X^{r+1} s}{\mu} - e \right) \right\| : A^T y + s = c \right\}. \qquad (4)$$

This measure will appear to be an appropriate one. 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

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

The next lemma states that there is a close relationship between this measure and the projected Newton direction $p(x, \mu)$ with respect to the barrier function, which is obtained from

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \frac{y}{\mu} \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}, \quad y \in \mathbb{R}^m. \qquad (5)$$

A closed-form formula for $p(x, \mu)$ is given by

$$p(x, \mu) = -H^{-1/2} P_{AH^{-1/2}} H^{-1/2} g, \tag{6}$$

where $P_{AH^{-1/2}}$ denotes the orthogonal projection onto the null space of the matrix $AH^{-1/2}$.

LEMMA 2: *For given x and* μ, $\delta(x, \mu) = \sqrt{r+1}\, \|H^{1/2} p(x, \mu)\|$.

Proof: From (6) we have

$$\begin{aligned} p(x, \mu) = &- H^{-1/2} (I - H^{-1/2} A^T (AH^{-1} A^T)^{-1} AH^{-1/2}) \\ &\times H^{-1/2} \left(\frac{c}{\mu} - X^{-r-1}\right) e. \end{aligned}$$

Now it is easy to derive that $\sqrt{r+1}\, p(x, \mu)$ can be written as $H^{-1/2} q$, where

$$q = X^{-r/2} \left(e - \frac{X^{r+1} s}{\mu}\right) \tag{7}$$

with

$$s = c - A^T y \tag{8}$$

and

$$y = \mu (AH^{-1} A^T)^{-1} AH^{-1} g. \tag{9}$$

On the other hand it can be verified that

$$s(x, \mu) = c - \mu A^T (AH^{-1} A^T)^{-1} AH^{-1} g.$$

Consequently, $s = s(x, \mu)$ and

$$\|q\| = \left\| X^{-r/2} \left(e - \frac{X^{r+1} s(x, \mu)}{\mu}\right)\right\| = \delta(x, \mu).$$

This proves the lemma. □

Note that $\|H^{1/2} p\| = \sqrt{p^T H p} := \|p\|_H$. This measure ("the Hessian norm of the Newton direction") already appeared in many other papers (e.g. Den Hertog *et al.* [2, 3]).

The Newton direction (6) is different from search directions used by other interior point methods: instead of the matrix $AX^2 A^T$ it uses the matrix $AX^{r+2} A^T$, so the scaling is different. In the sequel of the paper we will write p instead of $p(x, \mu)$.

Now we will prove some fundamental lemmas, which will be used in the following section to obtain upper bounds for the total number of iterations.

LEMMA 3: *If* $\delta := \delta(x, \mu) \leq 1$ *then* $y := y(x, \mu)$ *is dual feasible.*

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

$$\left\| X^{-r/2} \left(\frac{X^{r+1} s(x, \mu)}{\mu} - e \right) \right\| \leq 1. \tag{10}$$

Because x is feasible we have (due to our assumption) $x_i \leq 1$. Using this in (10) we get

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

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

Note that (7) gives an explicit expression for the Newton direction:

$$p = -H^{-1} \left(\frac{s}{\mu} - X^{-r-1} e \right) = -\frac{1}{r+1} \left(\frac{X^{r+2} s}{\mu} - x \right),$$

where $s = s(x, \mu)$. Hence the new point $\hat{x}$ after a full Newton step is given by

$$\hat{x} = x + p = \frac{r+2}{r+1} x - \frac{1}{r+1} \frac{X^{r+2} s}{\mu}.$$

Now we state our quadratic convergence result.

LEMMA 4: *If* $\delta(x, \mu) \leq 2/3$ *then* $\hat{x}$ *is strictly feasible. Moreover,* $\delta(\hat{x}, \mu) \leq \delta(x, \mu)^2$.

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

$$t = \frac{X^{r+1} s(x, \mu)}{\mu} - e.$$

Note that

$$\hat{x} = x - \frac{1}{r+1} Xt = X \left(e - \frac{1}{r+1} t \right). \tag{11}$$

From $\delta(x, \mu) \leq 2/3$ we deduce that $-2/3 \leq t_i \leq 2/3$. So $\hat{x}$ is strictly feasible.

The definition of $s(\hat{x}, \mu)$ implies the following:

$$\begin{aligned}\delta(\hat{x}, \mu) &= \left\| \hat{X}^{-r/2} \left(\frac{\hat{X}^{r+1}\, s(\hat{x}, \mu)}{\mu} - e \right) \right\| \\ &\leq \left\| \hat{X}^{-r/2} \left(\frac{\hat{X}^{r+1}\, s(x, \mu)}{\mu} - e \right) \right\|.\end{aligned}$$

Using (11) we find

$$\begin{aligned}\frac{\hat{X}^{r+1}\, s(x, \mu)}{\mu} - e &= \frac{(X - (1/(r+1))\, XT)^{r+1}\, s(x, \mu)}{\mu} - e \\ &= \left(I - \frac{1}{r+1}\, T \right)^{r+1} (t + e) - e.\end{aligned}$$

Let us call the right hand side vector $q(t)$. So

$$\begin{aligned}\delta(\hat{x}, \mu) &\leq \|\hat{X}^{-r/2}\, q(t)\| \\ &= \|\hat{X}^{-r/2}\, X^r\, X^{-r}\, T^2\, T^{-2}\, q(t)\| \\ &\leq \|\hat{X}^{-r/2}\, X^r\, T^{-2}\, q(t)\|_\infty \|X^{-r/2}\, t\|^2. \end{aligned} \tag{12}$$

Note that

$$\|X^{-r/2}\, t\|^2 = \delta(x, \mu)^2. \tag{13}$$

The proof will be completed by showing that

$$\|\hat{X}^{-r/2}\, X^r\, T^{-2}\, q(t)\|_\infty \leq 1. \tag{14}$$

Note that

$$\begin{aligned}&\|\hat{X}^{-r/2}\, X^r\, T^{-2}\, q(t)\|_\infty \\ &\quad \leq \|(\hat{X}^{-1}\, X)^{r/2}\, T^{-2}\, q(t)\|_\infty \\ &\quad = \max_i \left| \frac{(1 - (1/(r+1))\, t_i)^{r+1}\, (t_i + 1) - 1}{t_i^2\, (1 - (1/(r+1))\, t_i)^{r/2}} \right|. \end{aligned} \tag{15}$$

Remember that $-2/3 \leq t_i \leq 2/3$. Now we continue by evaluating (15). First note that $(1-(1/(r+1))\,t_i)^{r+1}\,(t_i+1)$ is less than 1 (it has its maximum for $t_i = 0$). Moreover, using Lagrange's remainder formula with respect to t_i we thus have

$$0 \leq 1-\left(1-\frac{1}{r+1}\,t_i\right)^{r+1}(t_i+1) = 1-\left(1-t_i+\gamma^2\,\frac{r}{r+1}\,t_i^2\right)(t_i+1),$$

where $0 < \gamma < 1$. This means that

$$1-\left(1-\frac{1}{r+1}\,t_i\right)^{r+1}(t_i+1) \leq 1-(1-t_i)(t_i+1) = t_i^2.$$

Now substituting this into the right hand side of (15), we get

$$\left|\frac{(1-(1/(r+1))\,t_i)^{r+1}\,(t_i+1)-1}{t_i^2\,(1-(1/(r+1))\,t_i)^{r/2}}\right| \leq \frac{1}{(1-(1/(r+1))\,t_i)^{r/2}},$$

which implies (14) if $t_i \leq 0$. Now we will prove that inequality (14) also holds for $0 \leq t_i \leq 2/3$. Note that it suffices to show that

$$\eta\,(t_i) := 1-\left(1-\frac{t_i}{r+1}\right)^{r+1}(t_i+1)-t_i^2\left(1-\frac{t_i}{r+1}\right)^{r/2} \leq 0,$$

for $0 \leq t_i \leq 2/3$. After some algebraic manipulations we obtain the derivative:

$$\begin{aligned}\eta'\,(t_i) =& \frac{r+2}{r+1}\left(1-\frac{t_i}{r+1}\right)^{(r/2)-1}\\ &\times t_i\left(\left(1-\frac{t_i}{r+1}\right)^{(r/2)+1}-\frac{4+4r-(4+r)\,t_i}{2\,(r+2)}\right).\end{aligned}$$

Because $0 \leq 1-(t_i/(r+1)) \leq 1$ we have

$$\begin{aligned}\eta'\,(t_i) &\leq \frac{r+2}{r+1}\left(1-\frac{t_i}{r+1}\right)^{(r/2)-1} t_i\left(1-\frac{t_i}{r+1}-\frac{4+4r-(4+r)\,t_i}{2\,(r+2)}\right)\\ &= \frac{r\,(r+3)}{2\,(r+1)^2}\left(1-\frac{t_i}{r+1}\right)^{(r/2)-1} t_i\left(t_i-\frac{2\,(r+1)}{r+3}\right)\\ &\leq 0, \end{aligned} \tag{16}$$

for $0 \leq t_i \leq 2/3$. Since $\eta(0) = 0$ we obtain $\eta(t_i) \leq 0$ for $0 \leq t_i \leq 2/3$. Consequently, inequality (14) holds for $t_i \leq 2/3$. Finally, substituting (13) and (14) into (12) gives

$$\delta(\hat{x}, \mu) \leq \delta(x, \mu)^2. \quad \square$$

Note that if $r \geq 1$ then inequality (16) holds for $t_i \leq 1$. Hence, if the rank of the barrier is equal to or larger than 1, then the quadratic convergence region is even wider: $\delta(x, \mu) \leq 1$.

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 following theorem shows that if the reduction factor is sufficiently small, then the iterate is again nearly centered with respect to the new center. We define $B_r(x)$ as the barrier term $\sum_{i=1}^{n} 1/x_i^r$. Note that $B_r(x) = \|X^{-r/2} e\|^2$.

THEOREM 1: *Let* $\delta(x, \mu) \leq 1/2$ *and* $\hat{\mu} := (1 - \theta)\mu$, *where* $\theta = 1/(9\sqrt{B_r(x)})$, *and* $\hat{x}$ *is defined as before, then* $\delta(\hat{x}, \hat{\mu}) \leq 1/2$.

Proof: Due to the definition of our measure we have

$$\begin{aligned}
\delta(x, \hat{\mu}) &= \left\| \frac{X^{(r/2)+1} s(x, \hat{\mu})}{\hat{\mu}} - X^{-r/2} e \right\| \\
&\leq \left\| \frac{X^{(r/2)+1} s(x, \mu)}{\hat{\mu}} - X^{-r/2} e \right\| \\
&= \left\| \frac{1}{1-\theta} \left(\frac{X^{(r/2)+1} s(x, \mu)}{\mu} - X^{-r/2} e \right) + \left(\frac{1}{1-\theta} - 1 \right) X^{-r/2} e \right\| \\
&\leq \frac{1}{1-\theta} \left(\delta(x, \mu) + \theta \sqrt{B_r(x)} \right) \\
&\leq \frac{1}{1-(1/9)} \left(\frac{1}{2} + \frac{1}{9} \right) = \frac{11}{16}.
\end{aligned}$$

Now we can apply the quadratic convergence result (Lemma 4)

$$\delta(\hat{x}, \hat{\mu}) \leq \delta(x, \hat{\mu})^2 \leq \left(\frac{11}{16} \right)^2 < \frac{1}{2}. \quad \square$$

The following lemma gives an upper bound for the difference in barrier function value in a nearly centered point x and $x(\mu)$.

LEMMA 5: *If* $\delta := \delta(x, \mu) < 2/3$ *then*

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

Proof: The barrier function f_r is conveQx in x, whence

$$f_r(x+p, \mu) \geq f_r(x, \mu) + p^T g.$$

Now using that $p = -H^{-1/2} P_{AH^{-1/2}} H^{-1/2} g$,

$$\begin{aligned} p^T g &= \left(H^{-1/2} g\right)^T H^{1/2} p \\ &= -\left(H^{-1/2} g\right)^T P_{AH^{-1/2}} H^{1/2} g \\ &= -\left(H^{-1/2} p\right)^T H^{1/2} p \\ &= -\frac{\delta^2}{r+1}, \end{aligned} \tag{17}$$

where the last equality follows from Lemma 2. Substitution gives

$$f_r(x+p, \mu) \geq f_r(x, \mu) - \frac{\delta^2}{r+1},$$

or equivalently

$$f_r(x, \mu) - f_r(x+p, \mu) \leq \frac{\delta^2}{r+1}.$$

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. Due to Lemma 4 we have

$$\delta(x^i, \mu)^2 \leq \delta(x^{i-1}, \mu)^4 \leq \ldots \leq \delta^{2^{i+1}}. \tag{18}$$

Now we may write

$$\begin{aligned} f_r(x, \mu) - f_r(x(\mu), \mu) &= \sum_{i=0}^{\infty} (f_r(x^i, \mu) - f_r(x^{i+1}, \mu)) \\ &\leq \sum_{i=0}^{\infty} \frac{\delta(x^i, \mu)^2}{r+1} \\ &\leq \frac{1}{r+1} \sum_{i=0}^{\infty} (\delta^2)^{2^i} \\ &\leq \frac{1}{r+1} \cdot \frac{\delta^2}{1-\delta^2}. \quad \square \end{aligned}$$

The following lemma shows that taking a certain step in the Newton direction gives a reduction in the barrier function value. We will use this result for the long-step path-following method.

LEMMA 6: *For*

$$\overline{\alpha} = \frac{r+1}{\delta}\,[1-(1+\delta)^{-1/(r+1)}]$$

we have

$$\Delta f_r := f_r\,(x,\,\mu) - f_r\,(x+\overline{\alpha}\,p,\,\mu) \geq \delta + \frac{r+1}{r}\,(1-(1+\delta)^{r/(r+1)}) \geq 0.$$

Proof: We write down the Taylor expansion for f_r:

$$f_r\,(x+\alpha\,p,\,\mu) = f_r\,(x,\,\mu) + \alpha\,p^T\,g + \sum_{k=2}^{\infty} t_k,$$

where t_k denotes the k-th order term in the Taylor expansion. From equation (17) we have $p^T\,g = -\delta^2/(r+1)$. Using that, for $k \geq 2$,

$$t_k = (-\alpha)^k\,\frac{(r+k-1)\cdots(r+1)}{k!}\,\sum_{i=1}^{n} x_i^{-r-k}\,p_i^k,$$

we derive

$$\begin{aligned}
|t_k| &\leq \alpha^k\,\frac{(r+k-1)\cdots(r+1)}{k!}\,\sum_{i=1}^{n}\left(\frac{|p_i|}{x_i^{(r/2)+1}}\right)^k \\
&\leq \alpha^k\,\frac{(r+k-1)\cdots(r+1)}{k!}\,\left[\sum_{i=1}^{n}\left(\frac{p_i}{x_i^{(r/2)+1}}\right)^2\right]^{k/2} \\
&= \left(\frac{\alpha\,\delta}{r+1}\right)^k\,\frac{(r+k-1)\cdots(r+1)}{k!}.
\end{aligned}$$

So we find

$$\begin{aligned}
f_r\,(x+\alpha\,p,\,\mu) &\leq f_r\,(x,\,\mu) - \alpha\,\frac{1}{r+1}\,\delta^2 \\
&\quad + \frac{1}{r}\sum_{k=2}^{n}\left(\frac{\alpha\,\delta}{r+1}\right)^k\,\frac{(r+k-1)\cdots(r+1)\,r}{k!} \\
&= f_r\,(x,\,\mu) - \alpha\,\frac{1}{r+1}\,\delta^2 - \frac{1}{r} - \frac{\alpha\,\delta}{r+1} + \frac{1}{r\,(1-(\alpha\,\delta/r+1))^r},
\end{aligned}$$

where the last equality holds if $\alpha\,\delta < r + 1$. Hence

$$f_r(x,\,\mu) - f_r(x + \alpha p,\,\mu) \geq \alpha \frac{1}{r+1}\delta^2 + \frac{1}{r} + \frac{\alpha\,\delta}{r+1} - \frac{1}{r\,(1 - (\alpha\,\delta/(r+1)))^r}.$$

The right hand side is maximal for

$$\overline{\alpha} = \frac{r+1}{\delta}\,[1 - (1+\delta)^{-1/(r+1)}].$$

Note that $\overline{\alpha}\,\delta < r + 1$. Substitution of this value finally gives

$$\Delta f_r \geq \delta + \frac{r+1}{r}\,(1 - (1+\delta)^{r/(r+1)}). \tag{19}$$

The right hand side is zero for $\delta = 0$ and is an increasing function of δ. Consequently

$$\delta + \frac{r+1}{r}\,(1 - (1+\delta)^{r/(r+1)}) \geq 0.$$

This proves the lemma. □

The following lemma gives a lower and upper bound for the duality gap in a nearly centered point.

LEMMA 7: *If* $\delta\,(x,\,\mu) \leq 1$ *then*

$$\mu\left(B_r\,(x) - \delta\sqrt{B_r\,(x)}\right) \leq c^T\,x - b^T\,y\,(x,\,\mu) \leq \mu\left(B_r\,(x) + \delta\sqrt{B_r\,(x)}\right).$$

Proof: We have,

$$\begin{aligned}
&\left|\frac{x^T\,s\,(x,\mu)}{\mu} - \|X^{-(r/2)}\,e\|^2\right| \\
&= \left|\left(X^{-r/2}\,e\right)^T\left(\frac{X^{(r/2)+1}\,s\,(x,\,\mu)}{\mu} - X^{-(r/2)}\,e\right)\right| \\
&\leq \|X^{-(r/2)}\,e\|\left\|\frac{X^{(r/2)+1}\,s\,(x,\,\mu)}{\mu} - X^{-(r/2)}\,e\right\| = \sqrt{B_r\,(x)\,\delta.}
\end{aligned}$$

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

$$\mu\left(B_r(x) - \delta\sqrt{B_r(x)}\right) \leq c^T x - b^T y(x,\mu)$$
$$\leq \mu\left(B_r(x) + \delta\sqrt{B_r(x)}\right). \quad \square$$

In Appendix 2 it is shown that if x is approximately centered with respect to $\mu \leq \mu_0 < \infty$, then $s(x, \mu)$ lies in a bounded region. Consequently, we can scale the dual problem (*D*) such that $s_i(x, \mu) \leq 1$, $i = 1, \ldots, n$, if $\delta(x, \mu) \leq 1$. (*See* Appendix 3.) This is used in the following lemma, which shows that the barrier term $B_r(x) = \sum_{i=1}^{n} 1/x_i^r$ is bounded in the neighborhood of the r-path. This result will be needed in the last two lemmas of this section.

LEMMA 8: *Let* $\delta := \delta(x, \mu) \leq 1$. *Then*

$$B_r(x) = \|X^{-r/2} e\|^2 \leq n\left[\frac{1}{\mu(1-\delta)}\right]^{r/(r+1)}.$$

Proof: Recall from (4) that $\delta = \|X^{-r/2}((X^{r+1} s(x, \mu)/\mu) - e)\|$. Let $s_1, \ldots, s_n$ denote the coordinates of $s(x, \mu)$. Then it follows that

$$x_i^{r+1} s_i \geq \mu(1 - x_i^{r/2}\delta) \geq \mu(1-\delta)$$

Consequently, since $s_i \leq 1$,

$$x_i \geq \left[\frac{\mu(1-\delta)}{s_i}\right]^{1/(r+1)} \geq [\mu(1-\delta)]^{1/(r+1)}.$$

From this it follows that

$$\|X^{-(r/2)} e\|^2 \leq n\left[\frac{1}{\mu(1-\delta)}\right]^{r/(r+1)} \quad \square$$

The following lemma gives an upper bound for the difference in object function value in a nearly centered point x and $x(\mu)$.

LEMMA 9: *If* $\delta := \delta(x, \mu) < 1$, *then*

$$|c^T x - c^T x(\mu)| \leq \frac{\mu\,\delta}{(r+1)(1-\delta)} \left(1 + \sqrt{n}\left[\frac{1}{\mu(1-\delta)}\right]^{r/(2(r+1))}\right).$$

Proof: From (17) we have $p^T g = -\delta^2/(r+1)$. On the other hand

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

So we have

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

or

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

Using Cauchy-Schwarz's inequality, we obtain

$$|e^T X^{-r-1} p| \leq \|X^{-r/2-1} p\| \, \|X^{-(r/2)} e\| = \frac{\delta}{r+1} \sqrt{B_r(x)},$$

where the last equality follows from Lemma 2. From this we deduce that

$$\begin{aligned} |c^T p| &\leq \mu \left(\frac{\delta^2}{r+1} + \frac{\delta}{r+1} \sqrt{B_r(x)}\right) \\ &\leq \frac{\mu\delta}{r+1} \left(1 + \sqrt{n}\left[\frac{1}{\mu(1-\delta)}\right]^{r/(2(r+1))}\right). \end{aligned}$$

Again, 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. From (18) we have $\delta(x^i, \mu) \leq \delta^{2^i}$. Consequently, we have

$$
\begin{aligned}
|c^T x - c^T x(\mu)| &= \left|\sum_{i=0}^{\infty} (c^T x^i - c^T x^{i+1})\right| \\
&\leq \sum_{i=0}^{\infty} \left| (c^T p(x^i, \mu) \right| \\
&\leq \sum_{i=0}^{\infty} \frac{\mu\,\delta(x^i, \mu)}{r+1} \left(1+\sqrt{n}\left[\frac{1}{\mu(1-\delta(x^i, \mu))}\right]^{r/(2(r+1))}\right) \\
&\leq \sum_{i=0}^{\infty} \frac{\mu}{r+1}\,\delta^{2^i} \left(1+\sqrt{n}\left[\frac{1}{\mu(1-\delta)}\right]^{r/(2(r+1))}\right) \\
&\leq \frac{\mu\,\delta}{(r+1)(1-\delta)} \left(1+\sqrt{n}\left[\frac{1}{\mu(1-\delta)}\right]^{r/(2(r+1))}\right). \quad \square
\end{aligned}
$$

4. PATH-FOLLOWING ALGORITHMS

4.1. Short Steps

The analysis in the previous section suggest the following short-step algorithm.

ALGORITHM 1:

Step 0: $x := x^0$, $\mu := \mu_0 \leq (n/\varepsilon)^{r+1}$, where x^0 and μ_0 satisfy $\delta(x^0, \mu_0) \leq 1/2$;

Step 1: if $x^T s \leq \epsilon$ then STOP;

Step 2: $\mu := (1 - (1/(9\sqrt{B_r(x)})))\,\mu$;

Step 3: $x := x + p$;

Step 4: go to Step 1.

There are several papers which give transformations to obtain initial points which are (approximately) centered with respect to the 0–path. See e. g. Renegar [17], Monteiro and Adler [15] and Gonzaga [8]. It is straightforward to generalize these transformations for r-paths.

THEOREM 2: *Algorithm 1 stops after at most*

$$O\left(\sqrt{n}\left(\frac{n}{\epsilon}\right)^{r/2}\ln\frac{n}{\epsilon}\right)$$

iterations.

Proof: Let μ_k denote the barrier parameter and x^k the iterate in the k-th iteration. Theorem 1 implies that $\delta := \delta(x, \mu_k) \leq 1/2$, where $x := x^k$. Hence, Lemma 8 gives that

$$B_r(x) \leq n\left(\frac{2}{\mu_k}\right)^{r/(r+1)}.$$

Now, using Lemma 7 we obtain

$$x^T s \leq \frac{3}{2}\mu_k B_r(x) \leq \frac{3}{2}\mu_k n\left(\frac{2}{\mu_k}\right)^{r/(r+1)} < 3n\mu_k^{1/(r+1)}.$$

The algorithm stops if $x^T s \leq \epsilon$. So it certainly has stopped if

$$3n\mu_k^{1/(r+1)} \leq \epsilon,$$

or

$$\mu_k \leq \left(\frac{\epsilon}{3n}\right)^{r+1}. \tag{20}$$

Hence, during the execution of the algorithm, we have that $\mu_k > (\epsilon/3n)^{r+1}$, whence

$$\theta = \frac{1}{9\sqrt{B_r(x)}} \geq \frac{1}{9\sqrt{n}}\left(\frac{\mu_k}{2}\right)^{r/(2(r+1))} \geq \frac{1}{13\sqrt{n}}\left(\frac{\epsilon}{3n}\right)^{r/2}.$$

Now note that the algorithm certainly stops if

$$\mu_k = (1-\theta)^k \mu_0 \leq \left(\frac{\epsilon}{3n}\right)^{r+1}.$$

Taking logarithms, and using that $-\ln(1-\theta) \geq \theta$, we obtain that

$$k \geq \frac{1}{\theta}\left(\ln\left(\frac{3n}{\epsilon}\right)^{r+1} + \ln\mu_0\right). \tag{21}$$

Now, using the assumption for μ_0 in Step 0, we derive that after

$$O\left(\sqrt{n}\left(\frac{n}{\epsilon}\right)^{r/2}\ln\frac{n}{\epsilon}\right)$$

iterations the algorithm certainly stops. □

This upper bound for the total number of iterations is not polynomial in the input length of the data for $r > 0$. It also makes clear that if the rank of the barrier r is small then the upper bound is better. If $r = 0$ then we obtain the well-known complexity bound $O(\sqrt{n}\,|\ln \epsilon|)$, which is polynomial in the input size. It is easy to verify that $r = 0$ corresponds to the logarithmic barrier approach.

4.2. Long Steps

Now we state the long-step algorithm.

ALGORITHM 2:

Step 0: x^0, $\mu := \mu_0 \leq (n/\varepsilon)^{r+1}$, where x^0 and μ_0 satisfy $\delta(x^0, \mu_0) \leq 1/2$;

Step 1: if $x^T s \leq \epsilon$ then STOP;

Step 2: $\delta(x, \mu) \leq 1/2$ then go to Step 5;

Step 3: $x := x + \tilde{\alpha}\, p$, where $\tilde{\alpha}$ minimizes $f_r(x + \alpha\, p, \mu)$;

Step 4: go to Step 2;

Step 5: $\mu := 1/2\, \mu$

Step 6: go to Step 1.

THEOREM 3: *Algorithm 2 stops after at most*

$$O\left(n\left(\frac{n}{\epsilon}\right)^{r}\ln\frac{n}{\epsilon}\right)$$

iterations.

Proof: We denote the barrier parameter value in an arbitrary outer iteration by $\overline{\overline{\mu}}$, while the parameter value in the previous outer iteration is denoted by $\overline{\mu}$. The iterate at the beginning of the outer iteration is denoted by x. Hence x is centered with respect to $x(\overline{\mu})$ and $\overline{\overline{\mu}} = (1/2)\,\overline{\mu}$. Note that because of Lemma 6 during each inner iteration the decrease in the barrier function value is at least

$$\Delta = \frac{1}{2} + \frac{r+1}{r}\left(1 - \left(\frac{3}{2}\right)^{r/(r+1)}\right) > 0.$$

Now let P denotes the number of inner iterations during one outer iteration. Then we have

$$P\Delta \leq f_r(x, \overline{\overline{\mu}}) - f_r(x(\overline{\overline{\mu}}), \overline{\overline{\mu}}). \tag{22}$$

Let us call the right hand side of (22) $\Phi(x, \overline{\overline{\mu}})$. According to the mean value theorem there is a $\hat{\mu} \in (\overline{\overline{\mu}}, \overline{\mu})$ such that

$$\Phi(x, \overline{\overline{\mu}}) = \Phi(x, \overline{\mu}) + \left.\frac{d\Phi(x, \mu)}{d\mu}\right|_{\mu=\hat{\mu}} (\overline{\overline{\mu}} - \overline{\mu}). \tag{23}$$

Let us now look at $d\Phi(x, \mu)/d\mu$. We have

$$\frac{df_r(x, \mu)}{d\mu} = -\frac{c^T x}{\mu^2},$$

and

$$\begin{aligned}\frac{df_r(x(\mu), \mu)}{d\mu} &= -\frac{c^T x(\mu)}{\mu^2} + \frac{c^T x'}{\mu} - \sum_{i=1}^{n} \frac{x_i'}{x_i(\mu)^{r+1}} \\ &= -\frac{c^T x(\mu)}{\mu^2} + \frac{c^T x'}{\mu} - \frac{s^T x'}{\mu} \\ &= -\frac{c^T x(\mu)}{\mu^2},\end{aligned}$$

where the two last equations follow from (2) and (3). So

$$-\left.\frac{d\Phi(x, \mu)}{d\mu}\right|_{\mu=\hat{\mu}} = \left.\frac{c^T x - c^T x(\mu)}{\mu^2}\right|_{\mu=\hat{\mu}} \leq \frac{|c^T x - c^T x(\overline{\overline{\mu}})|}{\overline{\overline{\mu}}^2},$$

where the last inequality follows from the fact that $\overline{\overline{\mu}} < \hat{\mu}$ and from Lemma 1. Substituting this into (23) gives

$$\begin{aligned}\Phi(x, \overline{\overline{\mu}}) &\leq \Phi(x, \overline{\mu}) + \frac{|c^T x - c^T x(\overline{\overline{\mu}})|}{\overline{\overline{\mu}}^2} (\overline{\mu} - \overline{\overline{\mu}}) \\ &\leq \Phi(x, \overline{\mu}) + \left(\frac{|c^T x - c^T x(\overline{\mu})|}{\overline{\overline{\mu}}} \right. \\ &\quad \left. + \frac{|c^T x(\overline{\mu}) - c^T x(\overline{\overline{\mu}})|}{\overline{\overline{\mu}}}\right) \frac{(\overline{\mu} - \overline{\overline{\mu}})}{\overline{\overline{\mu}}}.\end{aligned} \tag{24}$$

Because x is centered wit respect to $\overline{\mu}$ we have due to Lemma 5 and $\delta(x, \overline{\mu}) \leq 1/2$,

$$\Phi(x, \overline{\mu}) \leq \frac{1}{r+1} \cdot \frac{1/4}{1-(1/4)} < 1.$$

Now note that due to Lemma 9 and $\delta(x, \overline{\mu}) \leq 1/2$

$$|c^T x - c^T x(\overline{\mu})| \leq \frac{\overline{\mu}}{r+1}\left(1 + \sqrt{n}\left(\frac{2}{\overline{\mu}}\right)^{r/(2(r+1))}\right).$$

Moreover,

$$\begin{aligned}|c^T x(\overline{\mu}) - c^T x(\overline{\overline{\mu}})| &\leq c^T x(\overline{\mu}) - b^T y(x(\overline{\mu}), \overline{\mu}) \\ &\leq \overline{\mu} B_r(x(\overline{\mu})) \\ &\leq \overline{\mu} n \left(\frac{1}{\overline{\mu}}\right)^{r/(r+1)} \\ &= n \overline{\mu}^{1/(r+1)},\end{aligned}$$

where the second inequality follows form Lemma 7 and the third inequality from Lemma 8. Plugging all these upper bounds in (24) gives

$$\begin{aligned}\Phi(x, \overline{\overline{\mu}}) &\leq 1 + \frac{1}{\overline{\overline{\mu}}}\left(\overline{\mu}\left(1 + \sqrt{n}\left(\frac{2}{\overline{\mu}}\right)^{r/(2(r+1))}\right) + n\,\overline{\mu}^{1/(r+1)}\right)\frac{(\overline{\mu} - \overline{\overline{\mu}})}{\overline{\overline{\mu}}} \\ &\leq 1 + 2\left(\left(1 + \sqrt{n}\left(\frac{2}{\overline{\mu}}\right)^{r/(2(r+1))}\right) + n\left(\frac{1}{\overline{\mu}}\right)^{r/(r+1)}\right) \\ &\leq 3 + 5\,n\left(\frac{1}{\overline{\mu}}\right)^{r/(r+1)} \qquad (25)\end{aligned}$$

From (20) it follows that the algorithm certainly stops if the barrier parameter is smaller than $\mu_k \leq (\epsilon/3\,n)^{r+1}$. Consequently from (25) we have

$$\Phi(x, \overline{\overline{\mu}}) \leq 3 + 5\,n\left(\frac{3\,n}{\epsilon}\right)^r.$$

Hence combining this with (22) gives that at most $(1/\Delta)\,(3+5\,n\,(3\,n/\epsilon)^r)$ Newton steps are needed to return to the vicinity of the r-path. According to (21) at most $O(\ln(n/\epsilon))$ outer iterations (reductions of the barrier parameter) are needed. Hence after at most $O(n\,(n/\epsilon)^r \ln(n/\epsilon))$ iterations the algorithm ends up with an ϵ-optimal solutions. □

Again for $r = 0$ the result is exactly the same as obtained in many papers for long-step logarithmic barrier methods. See e. g. Gonzaga [9], Roos and Vial [19], and Den Hertog, Roos and Vial [3].

5. CONCLUDING REMARKS

We have developed path-following methods which use an inverse barrier function. Under some assumptions we also derived upper bounds for the total number of iterations. These bounds don't give polynomiality for $r > 0$.

It is easy to verify that $r = 0$ corresponds to the logarithmic barrier method and the results in this paper for $r = 0$ are similar to those obtained in [3]. Lemma 6 for example gives a lower bounds for the reduction in the barrier function value after a step in the Newton direction. For $r = 0$ we have (using L'Hôpital's rule):

$$\begin{aligned}\Delta f_0 &\geq \lim_{r\downarrow 0}\left[\delta + \frac{r+1}{r}\left(1-(1+\delta)^{r/(r+1)}\right)\right]\\ &= \delta + \lim_{r\downarrow 0}\frac{1-(1+\delta)^{r/(r+1)}}{r/(r+1)}\\ &= \delta - \lim_{r\downarrow 0}\frac{1/(r+1)^2\,(1+\delta)^{r/(r+1)}\ln(1+\delta)}{1/(r+1)^2}\\ &= \delta - \ln(1+\delta),\end{aligned}$$

which is well-known for the logarithmic barrier method.

The derived upper bounds for the total number of iterations are increasing in r. So, from the theoretical point of view, the logarithmic barrier method ($r = 0$) seems to have the best performance.

In this paper we assumed that the primal feasible region is bounded and has a nonempty interior. This assumption is necessary for the analysis of the long-step algorithm. It is easy to verify that for the short-step algorithm it is sufficient to assume that both the primal and the dual have a nonempty interior.

Note that $x(\mu)$ is not necessarily differentiable in $\mu = 0$ when $r > 0$. For the analysis given in this paper, this property was not necessary. But sometimes (e. g. for extrapolation techniques) differentiability in $\mu = 0$ is required. This can easily be accomplished by raising μ to the power $r+1$ in the barrier function (1). This has been proved by Fiacco and McCormick [5].

ACKNOWLEDGEMENTS

Thanks are due to Gert Dijkhuizen, who pointed out some characteristics of the expression (15), by using graphical software.

APPENDIX

A.1. INTRODUCTION

The purpose of this Appendix is to present some elementary properties of the primal and dual feasible region, hidden in the literature. Also the effect of scaling will be discussed.

We introduce the following notations: $\mathcal{F}_p$ and $\mathcal{F}_d$ are the primal and dual feasible regions respectively; $\mathcal{F}_p^0$ and $\mathcal{F}_d^0$ are the corresponding (relative) interiors; $\mathcal{F}_p^*$ and $\mathcal{F}_d^*$ are the corresponding optimal sets. We also introduce the following primal/dual level sets:

$$\mathcal{F}_p(\sigma) := \{x \mid x \in \mathcal{F}_p,\ c^T x \leq \sigma\},$$
$$\mathcal{F}_d(\sigma) := \{s \mid (y,\ s) \in \mathcal{F}_d,\ b^T y \geq \sigma\},$$

and

$$\mathcal{F}(\sigma) := \{(x,\ s) \mid x \in \mathcal{F}_p,\ (y,\ s) \in \mathcal{F}_d,\ c^T x - b^T y \leq \sigma\}.$$

We shall assume that both $\mathcal{F}_p^0$ and $\mathcal{F}_d^0$ are nonempty. Note that these assumptions are less restrictive than the assumptions made in Section 2. From these assumptions it is well-known from duality theory that both $\mathcal{F}_p^*$ and $\mathcal{F}_d^*$ are nonempty.

A.2. BOUNDEDNESS OF $\mathcal{F}_p(\sigma)$, $\mathcal{F}_d(\sigma)$ AND $\mathcal{F}(\sigma)$

Megiddo [14] showed that $\mathcal{F}_p(\sigma)$ and $\mathcal{F}_d(\sigma)$ are bounded; in particular $\mathcal{F}_p^*$ and $\mathcal{F}_d^*$ are bounded. His proof uses the logarithmic barrier function. We give another (more simple) proof, which uses Farkas' lemma.

LEMMA 10:

$$\mathcal{F}_p(\sigma)\ \textit{is bounded} \Leftrightarrow \mathcal{F}_d^0 \neq \emptyset$$

and

$$\mathcal{F}_d(\sigma)\ \textit{is bounded} \Leftrightarrow \mathcal{F}_p^0 \neq \emptyset.$$

Proof: It is easy to see that

$$\mathcal{F}_d(\sigma) \text{ is bounded} \Leftrightarrow \{x \mid A\,x = 0,\ c^T x \leq 0,\ e^T x = 1,\ x \geq 0\} = \emptyset.$$

Using Farkas' lemma this is equivalent with

$$\{(y,\ \eta,\ \nu) \mid A^T y - \eta\, c + \nu\, e \leq 0,\ \eta \geq 0,\ \nu > 0\} \neq \emptyset.$$

Evidentally, this is equivalent with

$$\{(y,\ \eta,\ \nu) \mid A^T y + s = \eta\, c,\ s > 0,\ \eta \geq 0\} \neq \emptyset.$$

Now there are two cases:

- $\eta > 0$, in which case $(y/\eta,\ s/\eta)$ is dual feasible, and hence $\mathcal{F}_d^0 \neq \emptyset$;
- $\eta = 0$, in which case we have a solution for the set $\{A^T y + s = 0,\ s > 0\}$.

This means that if $(y,\ s)$ is in this set, then for arbitrary y_0 and λ sufficiently large we have that $(y_0 + \lambda y,\ c - A^T y_0 + \lambda s)$ is an interior dual solution. Consequently, $\mathcal{F}_d^0 \neq \emptyset$.

On the other hand, if $\mathcal{F}_d^0 \neq \emptyset$ then

$$\{(y,\ \eta,\ \nu) \mid A^T y + s = \eta\, c,\ s > 0,\ \eta \geq 0\} \neq \emptyset,$$

because it has a solution for $\eta = 1$. So, finally we conclude that

$$\mathcal{F}_p(\sigma) \text{ is bounded} \Leftrightarrow \mathcal{F}_d^0 \neq \emptyset.$$

In the same way it can be proved that

$$\mathcal{F}_d(\sigma) \text{ is bounded} \Leftrightarrow \mathcal{F}_p^0 \neq \emptyset. \quad \square$$

LEMMA 11: *$\mathcal{F}(\sigma)$ is bounded.*

Proof: Let z^* denote the optimal value of problem (P). If $(x,\ s) \in \mathcal{F}(\sigma)$ then $c^T x - z^* \leq \sigma$ and $z^* - b^T y \leq \sigma$. So, it is obvious that

$$\mathcal{F}(\sigma) \subseteq \mathcal{F}_p(z^* + \sigma) \times \mathcal{F}_d(z^* - \sigma).$$

From Lemma 10 it follows that both $\mathcal{F}_p(z^* + \sigma)$ and $\mathcal{F}_p(z^* - \sigma)$ are bounded. Hence $\mathcal{F}(\sigma)$ is bounded. $\square$

A.2. BOUNDEDNESS OF $s_i(x,\ \mu)$

The following lemma shows that if x is approximately centered with respect to $\mu \leq \mu_0 < \infty$, then $s(x,\ \mu)$ lies in a bounded region.

LEMMA 12: *Let $\delta(x, \mu) \leq 1/2$ where $\mu \leq \mu_0 < \infty$. Then $s_i(x, \mu)$; $i = 1, \ldots, n$, are bounded from above.*

Proof: A consequence of Lemma 7 is that

$$c^T x - b^T y(x, \mu) \leq 2 \mu B_r(x).$$

From Lemma 5 we derive that $f_r(x, \mu) - f_r(x(\mu), \mu) \leq 1/3$. Thus if follows that

$$\frac{1}{r} B_r(x) \leq \frac{1}{3} - \frac{c^T x}{\mu} + \frac{c^T x(\mu)}{\mu} + \frac{1}{r} B_r(x(\mu)).$$

Combining these two inequalities gives

$$c^T x - b^T y(x, \mu) \leq 2 \left(\frac{r}{3} \mu - r c^T x + r c^T x(\mu) + \mu B_r(x(\mu))\right).$$

Now note that

$$c^T x(\mu) - c^T x \leq c^T x(\mu) - b^T y(\mu) =: \sigma(\mu)$$

and from (2) it also follows that $\mu B_r(x(\mu)) = \sigma(\mu)$. Consequently, it follows that

$$c^T x - b^T y(x, \mu) \leq 2 \left(\frac{r}{3} \mu + (r+1) \sigma(\mu)\right).$$

From Lemma 1 we derive that the gap $\sigma(\mu)$ is decreasing along the r-path. Hence we have that

$$(x, s(x, \mu)) \in \mathcal{F}\left(\frac{2r}{3} \mu + 2(r+1) \sigma(\mu)\right) \subseteq \mathcal{F}\left(\frac{2r}{3} \mu_0 + 2(r+1) \sigma(\mu_0)\right),$$

which was shown to be bounded in Lemma 11. □

A.3. SYMMETRICAL SCALING OF THE PROBLEM

From Appendix 2 we have that all $s(x, \mu)$ for $\mu \leq \mu_0 < \infty$ lie in a bounded region. Since the primal feasible region is bounded too, there exists a κ such that $x_i \leq \kappa$, for all primal feasible solutions, and $s_i(x, \mu) \leq \kappa$, for all $\mu \leq \mu_0$.

Now consider the following symmetrical scaling of the primal and dual problem

$$(\tilde{P}) \quad \min\{\tilde{c}^T \tilde{x} : \tilde{A}\,\tilde{x} = \tilde{b},\ \tilde{x} \geq 0\},$$

and

$$(\tilde{D}) \quad \max\{\tilde{b}^T \tilde{y} :\ \tilde{A}^T \tilde{y} + \tilde{s} = \tilde{c},\ \tilde{s} \geq 0\}.$$

where

$$\tilde{c} = c/\kappa, \quad \tilde{b} = b/\kappa, \quad \tilde{A} = A.$$

It is easy to verify that if x is primal feasible then $\tilde{x}_i = x_i/\kappa \leq 1$ is feasible for $(\tilde{P})$. By definition we have

$$s\,(x,\ \mu) = c - A^T\,(AH^{-1}\,A^T)^{-1}\,AH^{-1}\,(c - \mu\,X^{-r-1}\,e).$$

Consequently

$$\begin{aligned}\tilde{s}\,(\tilde{x},\ \tilde{\mu}) &= \tilde{c} - \tilde{A}\,(\tilde{A}\tilde{H}^{-1}\,\tilde{A}^T)^{-1}\,\tilde{A}\tilde{H}^{-1}\,(\tilde{c} - \tilde{\mu}\,\tilde{X}^{-r-1}\,e)\\ &= c/\kappa - A^T\,(AH^{-1}\,A^T)^{-1}\,AH^{-1}\,(c/\kappa - \tilde{\mu}\kappa^{r+1}\,X^{-r-1}\,e)\\ &= \frac{s\,(x,\ \tilde{\mu}\kappa^{r+2})}{\kappa}.\end{aligned}$$

Hence

$$\tilde{s}\left(\tilde{x},\ \frac{\mu}{\kappa^{r+2}}\right) = \frac{s\,(x,\ \mu)}{\kappa} \leq 1.$$

A.4. EITHER THE PRIMAL OR THE DUAL INTERIOR IS UNBOUNDED

LEMMA 13: *Either $\mathcal{F}_p^0$ or $\mathcal{F}_d^0$ is unbounded.*

Proof: Suppose $\mathcal{F}_p^0$ is bounded. Then

$$\{x | A\,x = 0,\ e^T\,x = 1,\ x \geq 0\} = \emptyset.$$

As a consequence of Farkas' lemma we have that

$$\{(y,\ \eta) | A^T\,y + \eta\,e \leq 0,\ \eta > 0\} \neq \emptyset.$$

This means that

$$\{(y,\ s)|A^T\,y+s=0,\ s>0\}\neq\emptyset.$$

Consequently, $\mathcal{F}_d^0$ is unbounded. In the same way it can be proved that if $\mathcal{F}_d^0$ is bounded, then $\mathcal{F}_p^0$ is unbounded. □

REFERENCES

1. C. W. CARROLL, The Created Response Surface Technique for Optimizing Nonlinear Restrained Systems, *Operations Research*, 1961, *9*, pp. 169-184.
2. D. DEN HERTOG, C. ROOS and T. TERLAKY, A potential Reduction Variant of Renegar's Short-Step Path-Following Method for Linear Programming, *Linear Algebra and Its Applications*, 1991, *68*, pp. 43-68.
3. D. DEN HERTOG, C. ROOS and J.-Ph. VIAL, A $\sqrt{n}$ Complexity Reduction for Long Step Path-following Methods, *SIAM Journal on Optimization*, 1992, *2*, pp. 71-87.
4. J. R. ERIKSSON, *An Iterative Primal-Dual Algorithm for Linear Programming*, Report LiTH-MAT-R-1985-10, 1985, Department of Mathematics, Linköping University, Linköping, Sweden.
5. A. V. FIACCO and G. P. MCCORMICK, *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*, Wiley and Sons, New York, 1968.
6. R. FLETCHER and A. P. MCCANN, Acceleration Techniques for Nonlinear Programming, In *Optimization*, R. Fletcher ed., Academic Press, London, 1969, pp. 203-214.
7. R. FRISCH, *The Logarithmic Potential Method for Solving Linear Programming Problems*, Memorandum, University Institute of Economics, Oslo, 1955.
8. C. C. GONZAGA, An Algorithm for Solving Linear Programming Problems in $O(n^3 L)$ Operations, In *Progress in Mathematical Programming, Interior Point and Related Methods*, pp. 1-28, N. Megiddo ed., Springer Verlag, New York, 1989.
9. C. C. GONZAGA, Large-Steps Path-Following Methods for Linear Programming: Barrier Function Method, *SIAM Journal on Optimization*, 1991, *1*, pp. 268-279.
10. P. HUARD, Resolution of Mathematical Programming with Nonlinear Constraints by the Methods of Centres, In *Nonlinear Programming*, J. Abadie ed., North-Holland Publishing Company, Amsterdam, Holland, 1989, pp. 207-219.
11. N. KARMARKAR, A New Polynomial-Time Algorithm for Linear Programming, *Combinatorica, 4*, 1984, pp. 373-395.
12. J. KOWALIK, Nonlinear Programming Procedures and Design Optimization, *Acta Polyntech. Scand.*, 1966, *13*, Trondheim.
13. G. P. MCCORMICK, W. C. MYLANDER and A. V. FIACCO, Computer Program Implementing the Sequential Unconstrained Minimization Technique for Nonlinear Programming, Technical Paper RAC-TP-151, Research Analysis Corporation, McLean, 1965.
14. N. MEGIDDO, Pathways to the Optimal Set in Linear Programming, In *Progress in Mathematical Programming, Interior Point and Related Methods*, pp. 131-158, N. Megiddo ed., Springer Verlag, New York, 1989.
15. R. D. C. MONTEIRO and I. ADLER, Interior Path Following Prima-Dual Algorithms, Part I: Linear Programming, *Mathematical Programming*, 1989, *44*, pp. 27-41.
16. R. A. POLYAK, Modified Barrier Functions (theory and methods), *Mathematical Programming*, 1992, *54*, pp. 174-222.

17. J. Renegar, A Polynomial-Time Algorithm, Based on Newton's Method, for Linear Programming, *Mathematical Programming*, 1988, *40*, pp. 59-93.

18. C. Roos and J.-Ph. Vial, A Polynomial Method of Approximate Centers for Linear Programming, *Mathematical Programming*, 1992, *54*, pp. 295-305.

19. C. Roos and J.-Ph. Vial, Long Steps with the Logarithmic Penalty Barrier Function in Linear Programming, In *Economic Decision-Making: Games, Economics and Optimization*, dedicated to Jacques H. Drèze, edited by J. Gabszevwicz, J.-F. Richard and L. Wolsey, Elsevier Sciences Publisher B. V., 1989, pp. 433-441.

20. A. Tamura, H. Takehara, K. Fukuda, S. Fujishige and S. Kojima, A Dual Primal Simplex Methods for Linear Programming, *Journal of the Operations Research Society of Japan*, 1988, *31*, pp. 413-429.

21. D. J. White, *Linear Programming and Huard's Method of Centres, Working, Paper*, Universities of Manchester and Virginia, United Kingdom, 1989.