Pergamon

Comput. Maths. Math. Phys., Vol.35, No.6, pp. 673–686, 1995

0965-5425(95)00088-7

THE USE OF NEWTON'S METHOD FOR LINEAR PROGRAMMING[1]

YU.G. EVTUSHENKO, V.G. ZHADAN and A.P. CHERENKOV

Moscow

(*Revised 04 October 2002*)

Continuous and discrete versions of the barrier-Newton method for linear programming are considered. This primal-dual method is based on the use of Newton's method to find points in the direct and dual spaces which satisfy a consistent system of optimality conditions. The local and non-local properties of the method are investigated. In the discrete versions of the method, the steps used in the direct and dual spaces are different. When the steps are chosen by certain rules, the method converges at superlinear and quadratic rates. In one version of the method the steps are chosen from the condition of steepest descent, and a range of initial conditions for which not more than two iterations are required is identified.

INTRODUCTION

Newton's method is one of the most efficient means of solving systems of non-linear equations and optimization problems [1, 2, 3]. Numerous versions of the method intended for linear programming (LP) have appeared recently (see [4, 5, 6], for instance). There is a detailed review of these methods in [7]. Primal-dual algorithms for which Newton's method is used to solve a parametrized system of equations, the limiting form of which gives optimality conditions for the direct and dual problems, are of special interest [8]–[12]. These methods have both quite a high local rate of convergence and polynomial algorithms. Another primal-dual method is described in this paper, based on solving a system of equations which set the optimality conditions in an LP problem, namely complementary slackness and accessibility (cf. [13]).

If a transformation of spaces is used to avoid having to stipulate that the variables are nonnegative [14, 15] and Newton's method is used to find points satisfying the Kuhn–Tucker conditions, a whole family of different methods is obtained. These have been examined for the general problem of nonlinear programming in [15, 16], and for LP in [17].

The present paper investigates methods of a special class, corresponding to component-wise transformations of spaces which involve the right-hand sides of the systems of ordinary differential equations describing the method, being multiplied by diagonal matrices which act as barriers and do not permit the trajectories to intersect the boundaries of positive orthants either in the original space or in the space of dual complementary variables. Unlike [17], the main focus here is the choice of steps in the direct and dual spaces which, as in [18], might be different. Versions of the method with small steps and with steps close to one or chosen by solving auxiliary optimization problems are considered separately.

[1] *Zh. Vȳchisl. Mat. Mat. Fiz.* Vol.35, No.6, pp. 850–866, 1995.

Financial support for this research was provided by the Russian Foundation for Basic Research (Project Nos 93-012-450 and 94-01-01379).

1. STATEMENT OF THE PROBLEM, BASIC IDEAS OF THE METHOD

Let $x = [x^1, \dots, x^n]$, $u = [u^1, \dots, u^m]$ be vectors from the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We will consider the direct and dual LP problems given in the form

$$\min_{x \in X} c^\top x, \qquad X = \{x \in \mathbb{R}^n : b - Ax = 0_m, \quad x \geq 0_n\}, \tag{1.1}$$

$$\max_{u \in U} b^\top u, \qquad U = \{u \in \mathbb{R}^m : v = c - A^\top u \geq 0_n\}. \tag{1.2}$$

Here and below A is an $m \times n$ matrix in which $m < n$, 0_i is the zero i-dimensional vector, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. The symbol 0_{ij} will denote the zero $i \times j$ matrix.

We will consider the following sets:

$$\begin{array}{lcllcl} \mathbb{R}^n_+ & = & \{x \in \mathbb{R}^n : x \geq 0_n\}, & \mathbb{R}^n_{++} & = & \{x \in \mathbb{R}^n : x > 0_n\}, \\ \operatorname{int} U & = & \{u \in \mathbb{R}^m : v = c - A^\top u > 0_n\}, & \operatorname{ri} X & = & \{x \in \mathbb{R}^n_{++} : Ax = b\}. \end{array}$$

It will be assumed everywhere that the rank of the matrix A is equal to m, the sets $\operatorname{int} U$ and $\operatorname{ri} X$ are non-empty, and problem (1.1) has a unique solution x_*, which is not degenerate. Then the dual problem (1.2) also has a unique solution u_* and is non-degenerate, the vector x_* has m non-zero components, and the vector $v_* = c - A^\top u_*$ has m zero components. Also, the conditions of complementary slackness $x^j_* v^j_* = 0$, $1 \leq j \leq n$, and strict complementary slackness are satisfied, that is, from $x^j_* = 0$ it follows that $v^j_* > 0$.

Linear and nonlinear programming problems were solved by Newton's method in [16, 17]. An entire family of numerical methods was obtained as a result. We shall confine our consideration here to just one numerical scheme, in which the iterations are constructed from the formula

$$W(x_k, u_k, \lambda_k) \begin{bmatrix} x_{k+1} - x_k \\ u_{k+1} - u_k \end{bmatrix} = - \begin{bmatrix} \alpha_k D(x_k) v_k \\ \tau_k (A x_k - b) \end{bmatrix}. \tag{1.3}$$

The subscript k here is the iteration number, $D(z)$ is a diagonal matrix in which the vector z lies on the principal diagonal, α_k, τ_k, λ_k are certain positive coefficients, $\lambda_k = \alpha_k / \tau_k$, the n-dimensional vector v has the form $v = v(u) = c - A^\top u$, and W is a square matrix of order $n + m$,

$$W(x, u, \lambda) = \begin{bmatrix} \lambda D(v) & -D(x) A^\top \\ A & 0_{mm} \end{bmatrix}.$$

At points where the vectors x_k and v_k have only non-zero components, system (1.3) can be written in the form

$$x_{k+1} = D(x_k)[e_n + \tau_k(\eta_k - e_n)], \qquad u_{k+1} = u_k + \alpha_k \mu_k, \tag{1.4}$$

where e_n is an n-dimensional vector, all of whose components are equal to one, and the vectors $\eta \in \mathbb{R}^n$, $\mu \in \mathbb{R}^m$ are determined from the formulae

$$\eta(x, u) = D^{-1}(v) A^\top \mu(x, u), \qquad \mu(x, u) = M^{-1}(x, u) b. \tag{1.5}$$

Here $M(x, u) = A D(x) D^{-1}(v) A^\top$ is the Gram matrix.

Using the relation $v = c - A^\top u$ between vectors u and v, we can rewrite (1.4) in variables x and v:

$$x_{k+1} = D(x_k)[e_n + \tau_k(\eta_k - e_n)], \qquad v_{k+1} = D(v_k)(e_n - \alpha_k \eta_k). \tag{1.6}$$

We now introduce the matrix $\Lambda \in \mathbb{R}^{d \times n}$, where $d = n - m$ is the defect of the matrix A. The columns of the matrix $\Lambda^\top$ form a basis of the null space of A, that is, $A \Lambda^\top = 0_{md}$. We then consider the set $V = \{v \in \mathbb{R}^n : \Lambda(v - c) = 0_d\}$. If $v \in V$, the vector $v - c$ will lie in the space of the rows of the matrix A and, for the vector v, there is an m-dimensional vector u such that $v = c - A^\top u$. The second relation of (1.6) implies that $\Lambda(v_{k+1} - c) = \Lambda(v_k - c) - \alpha_k \Lambda A^\top \mu_k = \Lambda(v_k - c)$. Thus, if $v_0 \in V$, $v_k \in V$ for all k and relations (1.4) are equivalent to (1.6). The iterations can be performed either in the space of x and u, or in the space of x and v.

Lemma 1. *Let x_* and u_* be non-degenerate solutions of problems* (1.1) *and* (1.2). *Then the matrix $W(x_*, u_*, \lambda)$ is non-degenerate.*

Proof. Without loss of generality, we can assume that the first m components of the vector x_* are non-zero. Then the vectors x_*, v_* and the matrices A and W can be represented in the form

$$x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \qquad v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \qquad x_*^B > 0_m, \quad x_*^N = 0_d, \quad v_*^B = 0_m, \quad v_*^N > 0_d, \tag{1.7}$$

$$A = [B \mid N], \qquad W(x_*, u_*, \lambda) = \begin{bmatrix} 0_{mm} & 0_{md} & -D(x_*^B)B^\top \\ 0_{dm} & \lambda D(v_*^N) & 0_{dm} \\ B & N & 0_{mm} \end{bmatrix}. \tag{1.8}$$

Here $B \in \mathbb{R}^{m \times m}$, $N \in \mathbb{R}^{m \times d}$.

To prove this, it is sufficient to show that the following system of homogeneous algebraic equations has only a zero solution:

$$D(x_*^B)B^\top \bar{u} = 0_m, \qquad \lambda D(v_*^N)\bar{x}^N = 0_d, \qquad B\bar{x}^B + N\bar{x}^N = 0_m,$$

where $\bar{x}^B \in \mathbb{R}^m$, $\bar{x}^N \in \mathbb{R}^d$, $\bar{u} \in \mathbb{R}^m$. But this is obvious, since B is a non-degenerate matrix. $\square$

Lemma 2. *For any $x \in \mathbb{R}^n_{++}$, $u \in \operatorname{int} U$, $\lambda \in \mathbb{R}^1_{++}$ the matrix $W(x, u, \lambda)$ is non-degenerate.*

Since the matrix $M(x, u)$ is non-degenerate on the sets under consideration, Lemma 2 follows from Frobenius' formula for the inverse matrix:

$$W^{-1} = \begin{bmatrix} \lambda^{-1}D^{-1}(v)[I_n - D(x)A^\top M^{-1}AD^{-1}(v) & D^{-1}(v)D(x)A^\top M^{-1} \\ -M^{-1}AD^{-1}(v) & \lambda M^{-1} \end{bmatrix}.$$

Here and below I_s is the $s \times s$ unit matrix.

We now combine the vectors x and u into one symbol, putting $z^\top = [x^\top, u^\top] \in \mathbb{R}^{n+m}$. According to Lemma 2, the right-hand sides of relations (1.6) are uniquely defined if all the components of the vectors x and $v(u)$ are strictly positive. Vectors z for which $z \in \mathbb{R}^n_{++}$, $u \in \operatorname{int} U$ are called interior. Versions of method (1.3) in which z remains an interior vector on all iterations will be called interior point methods.

Now consider the scalar $\Phi_k = x_k^\top v_k$. It follows from (1.4), (1.6) that

$$Ax_{k+1} - b = (1 - \tau_k)(Ax_k - b), \tag{1.9}$$

$$\Phi_{k+1} = (1 - \tau_k)\Phi_k + (\tau_k - \alpha_k)\mu_k^\top Ax_k + \alpha_k \tau_k \mu_k^\top (Ax_k - b). \tag{1.10}$$

Let $x_k > 0_n$, $v_k > 0_n$. Then in order to guarantee that the vectors x_{k+1} and v_{k+1} shall be non-negative, the steps α_k, τ_k must satisfy the conditions $e_n \geq a_k \eta_k$, $e_n \geq \tau(e_n - \eta_k)$. It is easy to see that these conditions apply if

$$\alpha_k \leq \alpha_k^* = \frac{1}{[\eta_k^*]_+}, \qquad 0 < \tau_k \leq \tau_k^* = \frac{1}{[1 - \eta_*^k]_+}, \tag{1.11}$$

where $[\alpha]_+ = \max[0, \alpha]$, η_k^* and η_*^k are, respectively, the maximum and minimum components of the vector η_k. We shall assume that $\alpha_k^* = +\infty$ if $\eta_k^* \leq 0$, and $\tau_k^* = +\infty$ if $\eta_*^k \geq 1$.

The numbers τ_k^* and α_k^* determine the largest possible steps with respect to the direct and dual variables along a Newtonian direction for which all the components of the vectors x and v remain non-negative on the kth iteration.

Lemma 3. *Let $x_k > 0_n$, $v_k > 0_n$, $v_k \in V$ and $b \neq 0_m$. Then $\eta_k^* > 0$. If, in addition, $x_k \in X$, the set X is bounded and the vector c does not belong to the space of rows of the matrix A, then $\eta_*^k < 1$.*

Proof. Suppose the contrary: let $\eta_k^* \leq 0$. Then (1.5) implies that $A^\top \mu_k \leq 0_n$, and according to (1.4) – (1.6), $b^\top u_{k+1} - b^\top u_k = \alpha_k b^\top M^{-1}(x_k, u_k) b > 0$, $v_{k+1} > 0_n$. Thus, moving in the Newtonian direction, as $\alpha_k \to \infty$ all the components of the vector v_{k+1} are non-negative and the objective function of the dual problem tends to infinity. This contradicts the existence of a bounded solution of problem (1.2). Thus $\eta_k^* > 0$ and, therefore, the maximum step α_k^* is bounded.

Let $x_k \in X$. Then according to (1.9) the point $x_{k+1} \in X$. Assuming, on the contrary, that $\eta_*^k \geq 1$, if $\eta_k^* > 1$, there should be a jth component of the vector η_k such that $\eta_k^j > 1$. Hence, by (1.4), $x_{k+1}^j \to \infty$ as $\tau_k \to \infty$. But this is impossible, since X is bounded. If $\eta_*^k = \eta_k^* = 1$, $\alpha_k^* = 1$ and when $\alpha_k = 1$ it follows from (1.6) that $v_{k=1} = 0_n$, $c = A^\top u_k$, contradicting the conditions of the lemma. Hence $\eta_*^k < 1$. The lemma is proved. $\square$

A full description of the numerical methods (1.3) and (1.5) requires rules for choosing the steps α_k and τ_k. We will describe methods of three types, with different rules for selecting the steps.

1. Steps α and τ are fixed and sufficiently small. In that case process (1.3) is similar to the continuous version of the method, considered in the next section.

2. Steps α and τ are close to unity. This method is similar to Newton's method.

3. At each iteration, the steps α_k and τ_k are chosen from the solution of auxiliary optimization problems. These can be called methods of steepest descent.

Some versions of the methods of these classes will be discussed below.

2. METHODS OF THE FIRST CLASS

In [17], method (1.3) was obtained from the continuous version in which the solution of problems (1.1) and (1.2) was reduced to finding limit points (as $t \to \infty$) of solutions of the following system of ordinary differential equations:

$$\lambda D(v)\frac{dx}{dt} - D(x)A^\top \frac{du}{dt} = -\alpha D(x)v, \tag{2.1}$$

$$A\frac{dx}{dt} = -\tau(Ax - b). \tag{2.2}$$

Here $x(t, z_0)$, $u(t, z_0)$ are solutions of the Cauchy problem (2.1), (2.2) with the initial-conditions vector $z_0^\top = [x_0^\top, u_0^\top]$.

Let $\alpha > 0$, $\tau > 0$ and $\lambda = 1$. Then method (2.1), (2.2 is locally convergent. The following theorem was proved in [17].

Theorem 1. *Let x_* and u_* be isolated non-degenerate solutions of problems* (1.1) *and* (1.2). *Then the pair $[x_*, u_*]$ is an asymptotically stable position of equilibrium for system* (2.1), (2.2).

The discrete version of method (1.3) *converges locally at least linearly for fixed parameters* λ_k, α_k *and* τ_k *such that* $\lambda_k = 1$, $0 < \alpha_k < 2$ *and* $0 < \tau_k < 2$.

System (2.1), (2.2) has $n + m$ first integrals:

$$D^\lambda(x(t, z_0))\, v(t, z_0) = e^{-\alpha t} D^\lambda(x_0) v_0, \tag{2.3}$$
$$Ax(t, z_0) - b = e^{-\tau t}(Ax_0 - b). \tag{2.4}$$

It follows that the components of the vectors x and v do not change sign on trajectories of system (2.1), (2.2). Thus if we take $x_0 > 0_n$, $v_0 > 0_n$ and the paths (2.1), (2.2) are bounded and continuable as $t \to \infty$, the Kuhn–Tucker optimality conditions for problem (1.1) will be satisfied at the limit points x_*, v_*:

$$Ax_* = b, \qquad x_* \geq 0_n, \qquad v_* \geq 0_n, \qquad D(x_*)v_* = 0_n.$$

Thus problems (1.1) and (1.2) can be solved either by finding the limit points of the solution of Cauchy's problem for system (2.1), (2.2), or by finding the limits of the solution of the non-linear system (2.3), (2.4) as $t \to \infty$. The first technique will be used below. If all the components of the vectors x and v are non-zero, system (2.1), (2.2) can be solved for the derivatives and we will obtain the equivalent system

$$\frac{dx}{dt} = \tau D(x)[\eta(x, u) - e_n], \qquad \frac{du}{dt} = \alpha\mu(x, u). \tag{2.5}$$

The right-hand sides of system (2.5) are defined at all points for which $z^\top(t, z_0) = [x^\top(t, z_0), u^\top(t, z_0)]$ is an interior vector. However, the right-hand sides are not defined at the point $z_*^\top = [x_*^\top, u_*^\top]$. We will show that if as $z(t, z_0)$ approaches z_* it remains an interior point, the vector functions $\mu(x(t, z_0), u(t, z_0))$, and $\eta(x(t, z_0), u(t, z_0))$ have finite limits. As in (1.7), (1.8) we will represent the vector η in the form $\eta^\top = [(\eta^B)^\top, (\eta^N)^\top]$, $\eta^B \in \mathbb{R}^m$, $\eta^N \in \mathbb{R}^d$.

Lemma 4. *Let the points* x_* *and* u_* *be non-degenerate solutions of problems* (1.1) *and* (1.2), *respectively, where the point* x_* *can be represented in the form* (1.7), (1.8). *Then for any* $x > 0$ *and* $u \in \text{int}U$

$$\eta^B = e_m - D^{-1}(x_*^B)\delta x^B + s_1(\delta z), \tag{2.6}$$
$$\eta^N = D^{-1}(v_*^N)\delta v^N + s_2(\delta z), \tag{2.7}$$
$$\mu = -\delta u + s_3(\delta z), \tag{2.8}$$

where $\delta x = x - x_*$, $\delta u = u - u_*$, $\delta z^\top = [\delta x^\top, \delta u^\top]$, $\delta v = -A^\top \delta u$; $\|s_i(\delta z)\| = o(\|\delta z\|)$, $i = 1, 2, 3$; *the vectors* δx *and* δv *are separated into components* δx^B, δx^N *and* δv^B, δv^N, *respectively, as in* (1.7), (1.8).

Proof. After separation of the vector x_*, the matrix $M(x, u)$ can be represented in the form

$$M(x, u) = M^B(x, u) + M^N(x, u). \tag{2.9}$$

Here

$$M^B(x, u) = BD(x^B)D^{-1}(v^B(u))B^\top, \qquad M^N(x, u) = ND(x^N)D^{-1}(v^N(u))N^\top. \tag{2.10}$$

If $M^B(x, u)$ is a non-degenerate matrix, then apart from (2.9), we have

$$M(x, u) = M^B(x, u)\{I_m + [M^B(x, u)]^{-1}M^N(x, u)\}.$$

Thus

$$M^{-1} = [I_m + (M^B)^{-1}M^N]^{-1}(M^B)^{-1} = [M^B(x,u)]^{-1} + S(x,u), \tag{2.11}$$

where $S = -(M^B)^{-1}M^N[I_m - (M^B)^{-1}M^N + \ldots](M^B)^{-1}$.

On the basis of (2.10) we have

$$[M^B(x,u)]^{-1} = (B^\top)^{-1}D(v^B(u))D^{-1}(x^B)B^{-1}. \tag{2.12}$$

Since, by virtue of the fact that the solutions of the direct and dual problems are non-degenerate, $x_*^N = 0_n$, $v_*^B = 0_m$, $v_*^N > 0_d$, it follows from (2.11) and (2.12) that $[M(x_*, u_*)]^{-1} = [M^B(x_*, u_*)]^{-1} = 0_{mm}$.

We have the obvious representations

$$b = Bx_*^B = Bx^B - B\delta x^B, \tag{2.13}$$

$$\delta v^N = -N^\top \delta u = N^\top (B^\top)^{-1}\delta v^B, \qquad \|S(x,u)\| = o(\|\delta z\|). \tag{2.14}$$

Substituting (2.11) and (2.13) into (1.5), we obtain

$$\eta(x,u) = D^{-1}(v(u))A^\top M^{-1}(x,u)b = \eta_1(x,u) + s(\delta z). \tag{2.15}$$

Here $\eta_1(x,u) = D^{-1}(v(u))A^\top (M^B)^{-1}B(x^B - \delta x^B)$, $\|s(\delta z)\| = o(\|\delta z\|)$.

Using relation (2.12) and the first equation of (2.14), we obtain (2.8) and the following formulae:

$$\begin{aligned}
\eta_1^B &= D^{-1}(v^B)B^\top (B^\top)^{-1}D(v^B)D^{-1}(x^B)B^{-1}B(x^B - \delta x^B) = \\
&= D^{-1}(x^B)(x^B - \delta x^B) = e_m - D^{-1}(x_*^B)\delta x_*^B + \theta^B(\delta z), \\
\eta_1^N &= D^{-1}(v^N)N^\top (B^\top)^{-1}D(v^B)D^{-1}(x^B)B^{-1}B(x^B - \delta x^B) = \\
&= D^{-1}(x^N)N^\top (B^\top)^{-1}v^B + \theta^N(\delta z) = D^{-1}(v^N)N^\top (B^\top)\delta v^B + \theta^N(\delta z) = \\
&= D^{-1}(v_*^N)\delta v^N + \theta_1^N(\delta z),
\end{aligned}$$

where $\|\theta^B(\delta z)\| = o(\|\delta z\|)$, $\|\theta_1^N(\delta z)\| = o(\|\delta z\|)$. Since $v_*^B = 0$, equations (2.6) and (2.7) follow from this and (2.15). This proves the lemma. □

The lemma explains why the vector $\eta(x,u)$ is often referred to as an indicator vector (see [19], for example). All the components of the vector $\eta(x_*, u_*)$ comprise zeros and ones, the basic components of the vector x_* corresponding to ones in the vector $\eta(x_*, u_*)$. Zero components of x_* correspond to zeros of $\eta(x_*, u_*)$.

We now consider the Lyapunov function $F(x,u)$ and Lebesgue set Ω_0 given by the formulae

$$F(x,u) = \|D^\lambda(x)(c - A^\top u)\| + \|Ax - b\|,$$
$$\Omega_0 = \{[x,u] : F(x,u) \le F(x_0,u_0),\ \ x \ge 0_n,\ \ v \ge 0_n,\ \ b^\top u_0 \le b^\top u\}.$$

Theorem 2. *Let $[x_*, u_*]$ be a pair of non-degenerate solutions of problems* (1.1) *and* (1.2). *Suppose, further, that the Lebesgue set Ω_0 is bounded. Then for any interior initial conditions vector z_0 we have the following properties:*

(1) the matrix $M(x(t,z_0), u(t,z_0))$ is non-degenerate for any $t \ge 0$,

(2) $z(t,z_0) \in \Omega_0$ and $v(t,z_0) \in V$ for all $t \ge 0$,

(3) the objective function of the dual problem $b^\top u$ increases monotonely on trajectories of system (2.5),

(4) the pair of solutions $[x(t,z_0), u(t,z_0)]$ *of system* (2.5) *is bounded, is continuable and converges to the pair* $[x_*, u_*]$ *as* $t \to \infty$.

We can illustrate the properties of the method by the simplest example in which $n = 2$, $m = 1$, $A = [1,1]$, $b = 1$, $c^\top = [-2,1]$. Obviously, $x_*^\top = [1,0]$, $u_* = -2$, $v_*^\top = [0,3]$, $f_* = c^\top x_* = -2$.

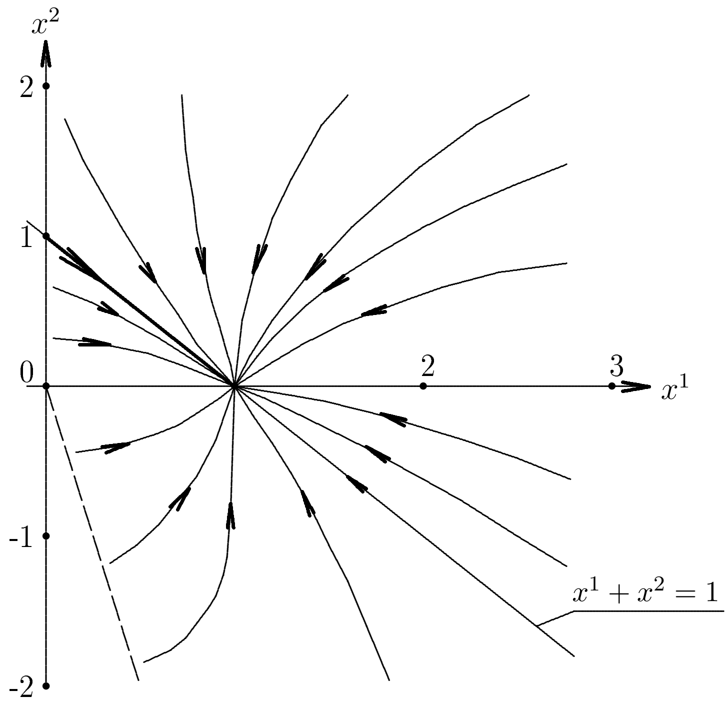

Fig. 1

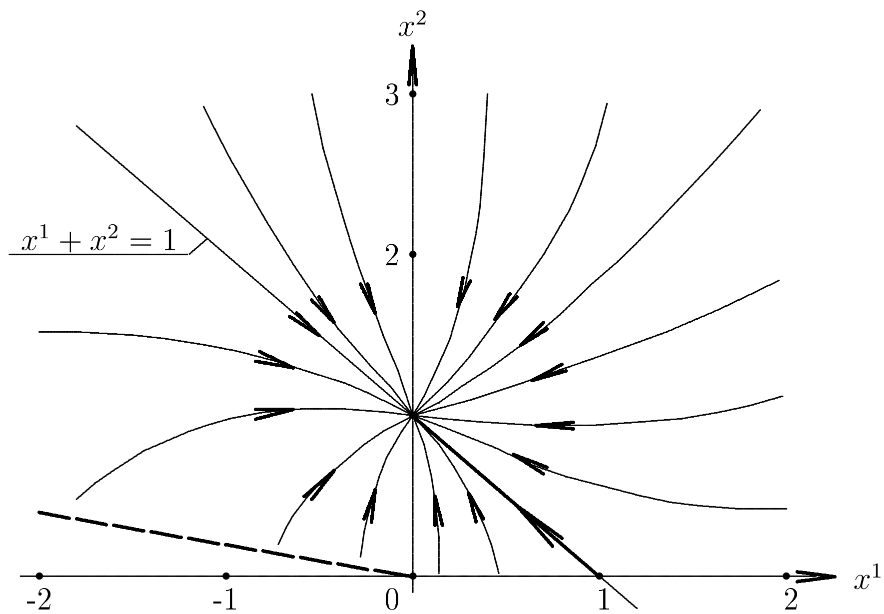

Fig. 2

Figure 1 shows the phase trajectories of system (2.1), (2.2) in the x^1, x^2 plane. We have taken $u_0 = -3$ as the initial vector of the dual variables, in which case $v_0 = [1,4]$. All trajectories starting from points with strictly positive components will converge to the point

x_*. The dashed line denotes the ray $x^1 v_0^1 + x^2 v_0^2 = 0$, $x^1 \geq 0$, $x^2 \leq 0$, corresponding to the set of points for which the matrix $M(x, u_0)$ is degenerate. The quantities x and v vary during the iterations, and the ray rotates, approaching the vertical axis as $t \to \infty$. It is clear from the figure that there is convergence towards x_* for any x_0 above that ray.

Figure 2 shows the phase portrait in the case where $u_0 = 2$ and $v_0 = [-4, -1]$. All the trajectories starting from $\mathbb{R}^2_{++}$ converge to the point $x^* = [0, 1]$, which is a solution of the problem of finding the maximum of the function $c^\top x$ on X. In that case, the point x_* is unstable, while x^* is an attractor.

Integrating system (2.1), (2.2) by Euler's method, we obtain method (1.4). The iterative process (1.4) will be similar to the process described by the system of differential equations (2.1), (2.2) if the steps α_k and τ_k are small enough. The properties of this version of the methods have been investigated by G.V. Smirnov, who has shown, in particular, that the discrete version of the method for small α_k and τ_k involves a polynomial. At the same time, it is obvious that method (1.4) is more effective if steps α_k and τ_k are sufficiently large. These versions of the method will be analyzed later.

3. LOCAL PROPERTIES OF THE METHOD

It follows from formulae (1.5) that the vector $D(x)\eta$ is a solution of the system of linear algebraic equations $AD(x)\eta = b$. If $x = x_*$, then the components of the vector η corresponding to basic components of the vector x_* will be one, and those corresponding to non-basic components will be zero. Therefore, $\eta_* = 0$, $\eta^* = 1$, $\alpha^* = \tau^* = 1$. Thus, the largest admissible steps giving non-negative x and v in the neighborhood of a solution are close to one. We will consider the simplest way of choosing α_k and τ_k. Assume that

$$\alpha_k = (1 - \varrho_k)\alpha_k^*, \qquad \tau_k = (1 - \varrho_k)\tau_k^*, \tag{3.1}$$

where $0 < \varrho_k < 1$. Three rules for choosing ϱ_k are:

$$0 < \varrho_k < 1, \qquad \lim_{k \to \infty} \varrho_k = 0, \tag{3.2}$$

$$\varrho_k = \max\left[\varkappa \Phi_k, 1 - \delta\right], \qquad 0 < \varkappa, \quad 0 < \delta < 1, \tag{3.3}$$

$$\varrho_k = \frac{\varkappa \Phi_k}{1 + \varkappa \Phi_k}, \qquad \varkappa > 0. \tag{3.4}$$

For all three methods, it is guaranteed that $0 < 1 - \varrho_k < 1$. Thus if $x_k > 0_n$, $v_k > 0_n$, these vectors are also strictly positive on the next iteration.

Theorem 3. *Under the assumptions of Lemma 4, let steps α_k and τ_k be chosen according to (3.1) – (3.4) at each iteration. Then method (1.4) converges locally to $[x_*, u_*]$ superlinearly at least, that is,*

$$\overline{\lim_{k \to \infty}} \frac{|x_{k+1} - x_*|}{|x_k - x_*|} = 0, \qquad \overline{\lim_{k \to \infty}} \frac{|u_{k+1}^j - u_*^j|}{|u_k^j - u_*^j|} = 0, \qquad 1 \leq i \leq n, \quad 1 \leq j \leq m. \tag{3.5}$$

$$\lim_{k \to \infty} \alpha_k^* = \lim_{k \to \infty} \tau_k^* = 1, \qquad \lim_{k \to \infty} \mu_k = 0_m, \quad \lim_{k \to \infty} \eta_k^* = 1, \quad \lim_{k \to \infty} \eta_*^k = 0. \tag{3.6}$$

If rules (3.3) or (3.4) are used the convergence is quadratic.

Proof. Let $\Delta x_k = x_k - x_*$, $\Delta v_k = v_k - v_*$, $\Delta u_k = u_k - u_*$, $\Delta z_k = [\Delta x_k, \Delta u_k]$. Then method (1.4) can be written in the form

$$\begin{aligned}\Delta x_{k+1} &= [I_n - \tau_k D(e_n - \eta_k)]\Delta x_k - \tau_k D(x_*)(e_n - \eta_k), && (3.7)\\ \Delta v_{k+1} &= [I_n - \alpha_k D(\eta_k)]\Delta v_k - \alpha D(v_*)\eta_k, && (3.8)\\ \Delta u_{k+1} &= \Delta u_k + \alpha_k \mu_k. && (3.9)\end{aligned}$$

Suppose that the pair of vectors $[x_k, u_k]$ is close to $[x_*, u_*]$ and the norm of the vector Δz_k is a small quantity of order ε. Suppose also, to fix our ideas, that the basis at the point x_* comprises the first m columns of the matrix A, that is, representation (1.7), (1.8) is being used. Then it follows from (2.6) and (2.7) that

$$\begin{aligned}\eta_k^B &= e_m - D^{-1}(x_*^B)\Delta x_k^B + s_1(\Delta z_k), && (3.10)\\ \eta_k^N &= D^{-1}(v_*^N)\Delta v_k^N + s_2(\Delta z_k), && (3.11)\\ \mu_k &= -\Delta u_k + s_3(\Delta z_k), && (3.12)\end{aligned}$$

where $\|s_i(\Delta z_k)\| = O(\varepsilon^2),\ i = 1, 2, 3.$

Substituting these relations into the right-hand sides of (3.7) – (3.9), we obtain

$$\begin{aligned}\Delta x_{k+1} &= (1 - \tau_k)\Delta x_k + \theta_1(\Delta z_k), && (3.13)\\ \Delta v_{k+1} &= (1 - \alpha_k)\Delta v_k + \theta_2(\Delta z_k), && (3.14)\\ \Delta u_{k+1} &= (1 - \alpha_k)\Delta u_k + \theta_3(\Delta z_k). && (3.15)\end{aligned}$$

Here $\|\theta_i(\Delta z_k)\| = O(\varepsilon^2),\ i = 1, 2, 3$; α_k and τ_k are determined from conditions (3.1). Under these assumptions we have

$$\begin{aligned}\frac{1}{\alpha_k^*} &= \eta_k^* = 2 - \min_{1\le i\le m}\left[\frac{x_k^i}{x_*^i}\right] + O(\varepsilon)^2,\\ \frac{1}{\tau_k^*} &= \eta_*^k = 2 - \min_{1\le i\le n}\left[\frac{v_k^i}{v_*^i}\right] + O(\varepsilon)^2.\end{aligned}$$

It follows that

$$\alpha_k^* = 1 + O(\varepsilon), \qquad \tau_k^* = 1 + O(\varepsilon). \tag{3.16}$$

We now substitute these formulae and (3.1) into the right-hand side of (3.13) – (3.15), and after some reduction obtain

$$\Delta x_{k+1} = \varrho_k \Delta x_k + \tilde{\theta}_1(\Delta z_k), \qquad \Delta u_{k+1} = \varrho_k \Delta u_k + \tilde{\theta}_2(\Delta z_k),$$

where $\|\tilde{\theta}_i(\Delta z_k)\| = O(\varepsilon^2),\ i = 1, 2$. Then using (3.2) we arrive at (3.5). Equations (3.6) follow from (3.10) – (3.12) and (3.16).

If rules (3.3) or (3.4) are used, near a solution

$$\varrho_k = \gamma\left(\sum_{i=1}^{m} x_*^i \Delta v_k^i + \sum_{i=m+1}^{n} v_*^i \Delta x_k^i\right) + O(\varepsilon^2).$$

From this and (3.13) – (3.15) we conclude that $\|\Delta z_{k+1}\| = O(\varepsilon^2)$, that is, the convergence is quadratic. This proves the theorem. □

The possibility in some cases of solving problems (1.1) and (1.2) in a finite number of steps is an important feature of method (1.4). Let us consider this property. We introduce index sets depending on the vectors x and v:

$$\sigma(x) = \{1 \le i \le n : \mathrm{e}^i = 0\}, \qquad \sigma(v) = \{1 \le i \le n : v^i = 0\}.$$

If all the components of x and v are non-zero, then $\sigma(x) = \emptyset$, $\sigma(v) = \emptyset$.

Let '$_k$ denote the set of initial pairs $[x_0, u_0]$ such that algorithm (1.3) for $\tau_k = \alpha_k = 1$ gives a solution of both problems (1.1) and (1.2) after k iterations. We will define the sets

$$\begin{aligned} \Omega_1 &= \{[x,u] : x = x_*, \ \ \sigma(x) \cap \sigma(v) = \emptyset\}, \\ \Omega_2 &= \{[x,u] : u = u_*, \ \ \sigma(x) \cap \sigma(v) = \emptyset\}, \\ \Omega_3 &= \{[x,u] : \sigma(x) = \sigma(x_*), \ \ \sigma(x) \cap \sigma(v) = \emptyset\}. \end{aligned}$$

Theorem 4. *Suppose that problems* (1.1) *and* (1.2) *have non-degenerate solutions x_* and u_*. Suppose also that the parameters in method* (1.3) *are chosen as follows: $\lambda_k = \tau_k = \alpha_k = 1$; then $\Omega_1 \subseteq T_1$, $\Omega_2 \subseteq T_1$, $\Omega_3 \subseteq T_2$.*

Proof. When $\lambda_k = \tau_k = \alpha_k = 1$ method (1.3) can be simplified to the form

$$D(v_k)x_{k+1} - D(x_k)A^\top u_{k+1} = -D(x_k)A^\top u_k, \qquad Ax_{k+1} = b.$$

Thus if the pair $[x_0, u_0] \in T_1$, we must have

$$D(v_0)x_* = D(x_0)A^\top(u_* - u_0). \tag{3.17}$$

It is easy to show that any pair of initial conditions $[e_0, u_0]$ from Ω_1 or Ω_2 satisfies (3.17).

Let us prove the last statement, that $\Omega_3 \subseteq T_2$. Suppose that there is a representation similar to (1.7), (1.8) for a point x_0, that is $x_0^B \neq 0_m$, $x_0^N = 0_d$. Then

$$D(v_0^B)x_1^B - D(x_0^B)B^\top u_1 = -D(x_0^B)B^\top u_0, \qquad D(v_0^N)x_1^N = 0_d, \qquad Bx_1^B + Nx_1^N = b.$$

All the components of the vector v_0^B are non-zero, and, therefore,

$$x_1^N = 0_d, \quad x_1^B = B^{-1}b = x_*^B, \quad u_1 = u_0 + (B^\top)^{-1}D^{-1}(x_0^B)D(v_0^B)x_*^B.$$

Thus, the exact vector $x = x_*$ is obtained in one step.

Letting $k = 1$ we find

$$D(v_1^B)x_2^B - D(x_1^B)B^\top u_2 = -D(x_1^B)B^\top u_1, \qquad D(v_1^N)x_2^N = 0_d, \qquad Bx_2^B + Nx_2^N = b.$$

The solution of this system is obvious:

$$\begin{aligned} x_2^B &= x_*^B, \quad x_2^N = 0_d, \quad u_2 = (B^\top)^{-1}c^B, \quad v_2^B = 0_m, \\ v_2^N &= c^N - N^\top(B^\top)^{-1}c^B = v_*^N > 0_d, \end{aligned}$$

hence the exact solution of both problems has been found. This proves the theorem. □

The result just obtained is global in character. For example, if $[x_0, u_0] \in \Omega_2$, the vector x_0 can even have negative components.

4. THE INTERIOR POINT METHOD WITH STEEPEST DESCENT

We will determine the rules for choosing steps α_k and τ_k at an interior point $z_k^\top = [x_k^\top, v_k^\top]$ from the solution of certain auxiliary problems. In order for the point $z_{k+1}^\top$ to be an interior point, we must have

$$0 \leq \alpha_k \leq \omega\alpha_k^* = \bar{\alpha}_k, \qquad 0 \leq \tau_k \leq \omega\tau_k^* = \bar{\tau}_k, \tag{4.1}$$

where $0 < \omega < 1$. For simplicity, we will assume that α_k^* and τ_k^* are finite. From (4.1) and (1.11) we obtain the inequality

$$\frac{1}{\bar{\alpha}_k} + \frac{1}{\bar{\tau}_k} \geq \frac{1}{\omega}. \tag{4.2}$$

Under condition (4.1), the steps α_k and τ_k are best chosen so as to minimize Φ_{k+1} and the norm of the discrepancy $\|Ax_{k+1} - b\|$, determined from formulae (1.10) and (1.9), as much as possible. Omitting the subscript k in order to simplify matters, and introducing the notation $L = \mu^\top Ax$, $M = \mu^\top b$, $N = ||Ax - b||$, we arrive at the two criteria:

$$\varphi(\alpha, \tau) = \Phi - L\alpha + (L - \Phi)\tau + (L - M)\alpha\tau, \qquad \varphi_2(\tau) = N|1 - \tau|.$$

We combine these into one by means of their linear convolution. The auxiliary problem thus obtained is: to find

$$\varphi^0(\bar{\alpha}, \bar{\tau}) = \min_{0 \leq \alpha \leq \bar{\alpha},\ 0 \leq \tau \leq \bar{\tau}} \varphi(\alpha, \tau), \tag{4.3}$$

$$\varphi(\alpha, \tau) = \varphi_1(\alpha, \tau) + \varphi_2(\tau). \tag{4.4}$$

Denote the solution of problem (4.3) by α^0, τ^0. We will assume that $b \neq 0_m$, so that $M > 0$. In addition, we will use the fact that $\varphi_1(\alpha, \tau) > 0$ on the set under consideration.

In the general case, the function $\varphi(\alpha, \tau)$ is piecewise-bilinear. Thus at least one of the extremal points will be the vertex of one of the rectangular regions where it is bilinear.

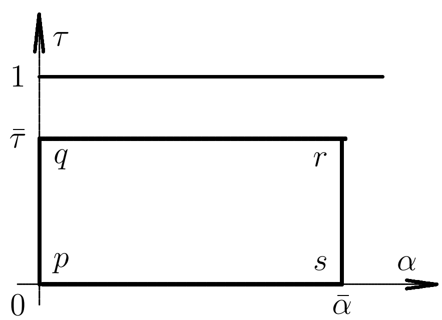

Fig. 3

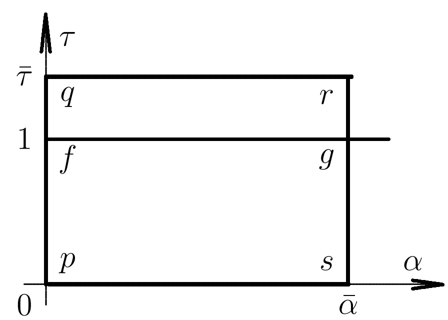

Fig. 4

If $\bar{\tau} < 1$, the function $\varphi(\alpha, \tau)$ is bilinear and can be expressed by the formula

$$\varphi(\alpha, \tau) = \Phi + N - L\alpha + (L - \Phi - N)\tau + (L - M)\alpha\tau, \tag{4.5}$$

its range of definition being the rectangle $pqrs$ (Fig. 3). We now write the values of $\varphi(\alpha, \tau)$ at its vertices:

$$\varphi_p = \Phi + N, \qquad \varphi_q = (\Phi + N)(1 - \bar{\tau}) + L\bar{\tau}, \tag{4.6}$$

$$\varphi_r = (\Phi + N)(1 - \bar{\tau}) + L(\bar{\tau} - \bar{\alpha}) + (L - M)\bar{\alpha}\bar{\tau}, \qquad \varphi_s = \Phi + N - L\bar{\alpha}. \tag{4.7}$$

Since $\varphi_s < \varphi_p$ for $L > 0$ and $\varphi_q < \varphi_p$ if $L \leq 0$, $\min \varphi$ cannot be reached at the point p. The form of the solution will depend on the behavior of the function φ on the segments qr and rs. Since $\partial\varphi(\alpha, \bar{\tau})/\partial\alpha = -L(1 - \bar{\tau}) - M\bar{\tau}$, $\partial\varphi(\bar{\alpha}, \tau)/\partial\tau = -(\Phi + N + M\bar{\alpha}) + L(1 + \bar{\alpha})$, we find the solution of the problem. It is given in Table 1, the values $\varphi^0(\bar{\alpha}, \bar{\tau})$ being found using formulae (4.6), (4.7).

If $\bar{\tau} \geq 1$ and $\tau \leq 1$, the function $\varphi(\alpha, \tau)$ has the form (4.5), and if $\tau \geq 1$

$$\varphi(\alpha, \tau) = \Phi - N - L\alpha + (L - \Phi + N)\tau + (L - M)\alpha\tau.$$

The function $\varphi(\alpha, \tau)$ is piecewise-bilinear. The subregions where the function is bilinear are the rectangles $pfgs$ and $fqrg$ (Fig. 4). We find the values of $\varphi(\alpha, \tau)$ at their vertices:

$$\varphi_p = \Phi + N, \qquad \varphi_f = L, \qquad \varphi_q = (\Phi - N)(1 - \bar{\tau}) + L\bar{\tau}, \tag{4.8}$$

$$\varphi_r = (\Phi - N + L\bar{\alpha})(\bar{\tau} - 1) + (L - M\bar{\alpha})\bar{\tau}, \quad \varphi_g = L - M\bar{\alpha}, \quad \varphi_s = \Phi + N - L\bar{\alpha}, \tag{4.9}$$

Table 1

<table>
<tr><td></td><td>$(\partial\varphi)/(\partial\alpha)|_{\tau=\bar{\tau}} > 0$</td><td>$(\partial\varphi)/(\partial\alpha)|_{\tau=\bar{\tau}} = 0$</td><td>$(\partial\varphi)/(\partial\alpha)|_{\tau=\bar{\tau}} < 0$</td></tr>
<tr><td>$(\partial\varphi)/(\partial\tau)|_{\alpha=\bar{\alpha}} < 0$</td><td>point q</td><td>segment qr</td><td>point r</td></tr>
<tr><td>$(\partial\varphi)/(\partial\tau)|_{\alpha=\bar{\alpha}} = 0$</td><td colspan="2" rowspan="2">iz not realized</td><td>segment rs</td></tr>
<tr><td>$(\partial\varphi)/(\partial\tau)|_{\alpha=\bar{\alpha}} > 0$</td><td>point s</td></tr>
</table>

The vertex p cannot be the point of a minimum of $\varphi(\alpha, \tau)$ since, in view of the relation $\varphi_1(\alpha, \tau) > 0$, we must have $L > 0$ and, therefore, $\varphi_s < \varphi_p$. The minimum of $\varphi(\alpha, \tau)$ cannot be reached at the vertex f either, since $\varphi_g < \varphi_f$. Analyzing the behavior of $\varphi(\alpha, \tau)$ on the sides qr and rs, using the relations

$$\begin{aligned}
\left(\frac{\partial\varphi}{\partial\alpha}\right)_{\tau=\bar{\tau}} &= L(\bar{\tau} - 1) - M\bar{\tau}, \\
\left(\frac{\partial\varphi}{\partial\tau}\right)_{\substack{\alpha=\bar{\alpha} \\ \tau>1}} &= L(1 + \bar{\alpha}) + N - (\Phi + M\bar{\alpha}) \geq \left(\frac{\partial\varphi}{\partial\tau}\right)_{\substack{\alpha=\bar{\alpha} \\ \tau<1}} = \\
&= L(1 + \bar{\alpha}) - (\Phi + N + M\bar{\alpha}),
\end{aligned}$$

various possible versions of the solution are obtained. These are given in Table 2, the values $\varphi^0(\bar{\alpha}, \bar{\tau})$ being found from (4.8), (4.9).

Thus, if the optimum strategies we have obtained for choosing steps α_k and τ_k are followed, in every case the step α_k can be taken equal to 0 or $\bar{\alpha}$. In the case when $\bar{\tau} \geq 1$, apart from the extreme values 0 and $\bar{\tau}$, the step τ_k can also take the intermediate value 1. With steps α_k and τ_k chosen from the solution of the auxiliary problem (4.4), the interior point method (1.4) is called *the method of steepest descent.*

We will now find bounds for the quantity $\varphi^0(\bar{\alpha}, \bar{\tau})$ on the assumption that

$$||\eta_k||_\infty = \max_{1 \leq i \leq n} |\eta_k^i| \leq C.$$

In that case, according to (1.11) and (4.1), we have

$$\bar{\alpha} \geq \frac{\omega}{C}, \qquad \bar{\tau} \geq \frac{\omega}{1 + C}. \tag{4.10}$$

We will find

$$\psi^* = \sup_{\bar{\alpha} \geq \omega/C,\ \bar{\tau} \geq \omega/(1+C)} \varphi^0(\bar{\alpha}, \bar{\tau}).$$

We use the fact that at any admissible point of $[\alpha, \tau]$ the value of $\varphi^0(\bar{\alpha}, \bar{\tau})$ should not be larger than the value $\varphi(\alpha, \tau)$.

In the case when $\bar{\tau} \geq 1$, we take s and f as such points:

$$\varphi^0(\bar{\alpha}, \bar{\tau}) \leq \min\{\varphi_s, \varphi_f\} \leq \max_L \min\{L, \Phi + N - \bar{\alpha}L\} = \frac{\Phi + N}{1 + \bar{\alpha}}. \tag{4.11}$$

Table 2

	$(\partial\varphi)/(\partial\alpha)\vert_{\tau=\bar{\tau}} > 0$	$(\partial\varphi)/(\partial\alpha)\vert_{\tau=\bar{\tau}} = 0$	$(\partial\varphi)/(\partial\alpha)\vert_{\tau=\bar{\tau}} < 0$
$(\partial\varphi)/(\partial\tau)\vert_{\alpha=\bar{\alpha},\tau>1} < 0$ $\left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} < 0 = \left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	point q	segment qr open polygon qrg	point r segment rg
$\left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} = 0 = \left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	is not realized	is not realized	segment rs
$\left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} < 0 < \left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	arg min $\{\varphi_q, \varphi_g\}$	point g	point g
$\left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} = 0 < \left.\frac{\partial\varphi}{\partial\tau}\right\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	arg min $\{\varphi_q, \varphi_{sg}\}$	segment sg	segment sg
$0 < (\partial\varphi)/(\partial\tau)\vert_{\alpha=\bar{\alpha},\tau<1}$	arg min $\{\varphi_q, \varphi_s\}$	point s	point s

In the case when $\bar{\tau} < 1$ we turn to points q, r and s. Since $M > 0$, we have $\varphi_r < < (\Phi + N)(1 - \bar{\tau}) + (\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau})L$. Thus

$$\varphi^0(\bar{\alpha}, \bar{\tau}) \leq \min\{(\Phi + N)(1 - \bar{\tau}) + L\bar{\tau}, (\Phi + N)(1 - \bar{\tau}) + (\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau})L, \Phi + N - \bar{\alpha}L\} \leq$$
$$\leq (\Phi + N)(1 - \bar{\tau}) + \max_L \min\{\bar{\tau}L, (\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau})L, (\Phi + N)\bar{\tau} - \bar{\alpha}L\}. \tag{4.12}$$

It follows from $\bar{\tau} \leq 1$ that $\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau} \leq \bar{\tau}$. Thus if $\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau} \geq 0$, the maximum with respect to L is found from the condition $(\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau})L = (\Phi + N)\bar{\tau} - \bar{\alpha}L$ and is reached at the point $L = (\Phi + N)/(1 + \bar{\alpha})$, as shown in Fig. 5. Substituting the given value of L into the right-hand side of (4.12), we arrive at the conclusion that inequality (4.11) still applies under this assumption. The case $\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau} < 0$ corresponds to the situation shown in Fig. 6. The maximum is determined from the condition $\bar{\tau}L = (\bar{\tau} - \bar{\alpha} + \bar{\alpha}\bar{\tau})L$ and is reached at zero. Instead of (4.11), we have

$$\varphi^0(\bar{\alpha}, \bar{\tau}) \leq (\Phi + N)(1 - \bar{\tau}). \tag{4.13}$$

It follows from (4.1), (4.11) and (4.13) that

$$\varphi^* \leq \nu(\Phi + N), \qquad \nu = \max\left\{\max_{\bar{\alpha}\geq\omega/C} \frac{1}{1 + \bar{\alpha}}, \max_{\bar{\tau}\geq\omega/(1+C)} (1 - \bar{\tau})\right\} = 1 - \frac{\omega}{1 + C}. \tag{4.14}$$

This maximum is unattainable and unimprovable. That $(\varphi^0(\bar{\alpha}, \bar{\tau}) < \varphi^*)$ is unattainable follows from the previous argument, and that the estimate cannot be improved is clear from the following example: putting $\Phi = N = 1$, $L = 0$, $M \to 0$, $\omega = 0.8$, $C = 0.6$, $\alpha^* = 1.875$, $\tau^* = 0.625$ and checking that conditions (4.2), (4.10) and $\bar{\tau} < 1$ are satisfied, and that the point r is an optimum, using (4.6), (4.7) and (4.14) we find $\varphi^0 = 1 - 3M/4$, $\nu(\Phi + N) = 1$, that is, $\lim_{M\to 0} \varphi^0 = \nu(\Phi + N)$.

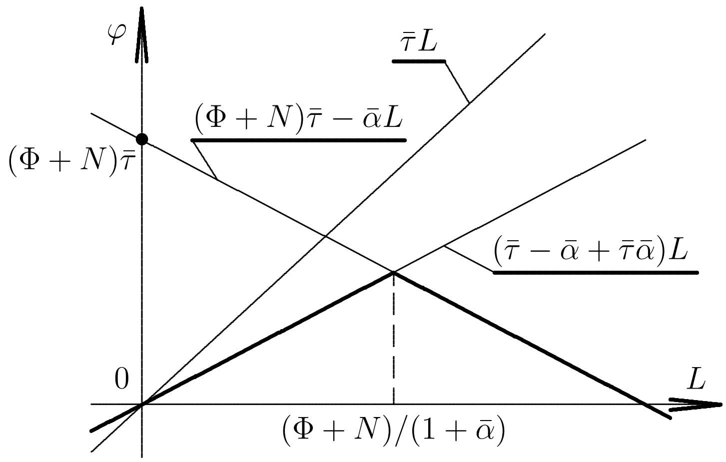

Fig. 5

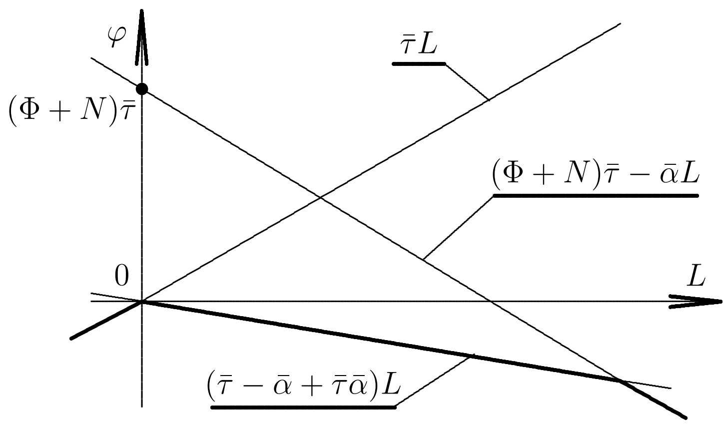

Fig. 6

On the basis of (4.14), using the definition of φ^*, we obtain

$$\Phi_{k+1} + ||Ax_{k+1} - b|| \leq \nu(\Phi_k + ||Ax_k - b||).$$

This inequality enables us to estimate the sufficient number of steps for the point $[x_k, u_k]$ to be in some neighborhood of the solution of problems (1.1) and (1.2).

Theorem 5. *Let* $z_0 = [x_0, u_0]$ *be an interior point, and let the sequences* $\{x_k\}$ *and* $\{u_k\}$ *generated by the method of steepest descent be such that* $\|\eta_k\|_\infty \leq C$ *for all* k*. Then for any* $\varepsilon > 0$*, the function* $\Phi(x,u) = v^\top(u)x + \|Ax - b\|$ *will become less than* ε *after not more than*

$$K = \left\lceil \frac{1+C}{\omega} \ln \frac{\Phi(x_0, u_0)}{\varepsilon} \right\rceil$$

iterations, where $\lceil a \rceil$ *is the smallest integer approaching the number* a *from above.*

The solution of the example of Section 2 found by the method of steepest descent is shown in Fig. 7. The parameter ω was taken equal to 0.9. Note that as ω approaches one, the number of iterations needed to solve the problem with prescribed accuracy decreases, reaching two for some initial points.

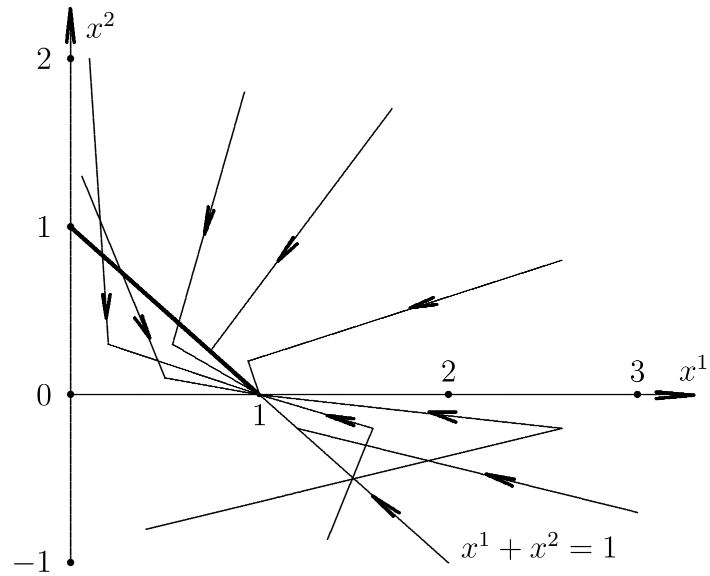

Fig. 7

REFERENCES

1. ORTEGA J.M. and RHEINBOLDT W.C., *The Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York, 1970.

2. VASIL'YEV F.P., *The Numerical Solution of Extremal Problems.* Nauka, Moscow, 1980.

3. YEVTUSHENKO Yu.G., *Methods of Solving Extremal Problems and their Application in Optimization Systems.* Nauka, Moscow, 1982.

4. IRI M. and IMAI H., A multiplicative barrier function method for linear programming. *Algorithmica* 455–482, 1986.

5. De GHELLINK G. and VIAL J.-P., A polynomial Newton method for linear programming. *Algorithmica* 425–453, 1986.

6. RENEGAR J., A polynomial-time algorithm, based on Newton's method for linear programming. *Math. Program.* **40**, 55–94, 1988.

7. GONZAGA C.C., Path-following methods for linear programming. *SIAM Rev.* **34**, 167–224, 1992.

8. KOJIMA M., M1ZUNO S. and YOSHISE A., A primal-dual interior point method for linear programming. In : MEGIDDO N. (Ed.), *Progress Math. Program. Interior Point and Relative Methods*, Ch. 2. Springer Verlag, Berlin, 1989.

9. MONTEIRO R.C. and ADLER I., Interior path-following primal-dual algorithm. Part I: Linear programming. *Math. Program.* **44**, 43–66, 1989.

10. McSHANE K., MONMA C. and SHANNO D., An implementation of a primal-dual interior point method for linear programming. *ORSA J. Comp.* 70–83, 1989.

11. YE Y., TAPIA R. and ZHANG Y., A superlinear convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Tech. Report TR91-22. Rice Univ. Houston, Texas, 1991.

12. JANSEN B., ROOS C., TERLAKY T. and VIAI J.-P., Primal-dual target-following algorithms for linear programming, Tech. Report 93–107, Faculty of Tech. Math and Informatics, TU Delft, 1993.

13. YEREMIN I.I. and ASTAF'YEV N.N., *Introduction to the Theory of Linear and Convex Programming*. Nauka, Moscow, 1976.

14. EVTUSHENKO Yu.G. and ZHADAN V.G., The space transformation technique in mathematical programming. In: KALL P. (Ed.), *System Modelling and Optimization. Lecture Notes in Control and Information Science*, 180. *Proc. 15th IFIP Conference*. Springer, Berlin, 1991.

15. EVTUSHENKO Yu.G. and ZHADAN V.G., Stable barrier-projection and barrier-Newton methods in nonlinear programming. *Optimization Methods and Software* **3**, 237–256, 1994.

16. YEVTUSHENKO Yu.G. and ZHADAN V.G., Barrier-projection and barrier-Newton optimization method (the non-linear programming case). In: *Communications in Computational Mathematics*. VTs AN S.S.S.R., Moscow, 1991.

17. YEVTUSHENKO Yu.G. and ZHADAN V.G., Barrier-projection and barrier-Newton optimization methods (the linear programming ease). In: *Communications in Computational Mathematics*. VTs RAN, Moscow, 1992.

18. YE Y., Line search in potential reduction algorithms for linear programming. Dept. Management Sci, University of Iowa, Iowa City, IA, 1989.

19. EL-BAKLY A.S., TAPIA R.A. and ZHANG Y., A study of indicators for identifying zero variables in interior-point methods. Tech. Report TR91–15, Rice Univ., Houston, Texas, 1991.