SIAM J. CONTROL AND OPTIMIZATION
Vol. 22, No. 6, November 1984

010

TWO-METRIC PROJECTION METHODS FOR CONSTRAINED OPTIMIZATION*

ELI M. GAFNI† AND DIMITRI P. BERTSEKAS‡

Abstract. This paper is concerned with the problem $\min\{f(x)|x\in X\}$ where X is a convex subset of a linear space H, and f is a smooth real-valued function on H. We propose the class of methods $x_{k+1}=P(x_k-\alpha_k g_k)$, where P denotes projection on X with respect to a Hilbert space norm $\|\cdot\|$, g_k denotes the Frechet derivative of f at x_k with respect to another Hilbert space norm $\|\cdot\|_k$ on H, and α_k is a positive scalar stepsize. We thus remove an important restriction in the original proposal of Goldstein [1] and Levitin and Poljak [2], where the norms $\|\cdot\|$ and $\|\cdot\|_k$ must be the same. It is therefore possible to match the norm $\|\cdot\|$ with the structure of X so that the projection operation is simplified while at the same time reserving the option to choose $\|\cdot\|_k$ on the basis of approximations to the Hessian of f so as to attain a typically superlinear rate of convergence. The resulting methods are particularly attractive for large-scale problems with specially structured constraint sets such as optimal control and nonlinear multi-commodity network flow problems. The latter class of problems is discussed in some detail.

Key words. constrained optimization, gradient projection, convergence analysis, multicommodity flow problems, large-scale optimization

1. Introduction.

Projection methods stemming from the original proposal of Goldstein [1], and Levitin and Poljak [2] are often very useful for solving the problem

$$\begin{aligned}&\text{minimize } f(x)\\ &\text{subject to } x\in X\end{aligned}\tag{1}$$

where $f: H\to R$ and X is a convex subset of a linear space H. They take the form

$$x_{k+1}=P_k(x_k-\alpha_k g_k)\tag{2}$$

where α_k is a positive scalar stepsize, $P_k(\cdot)$ denotes projection on X with respect to some Hilbert space norm $\|\cdot\|_k$ on H and g_k denotes the Frechet derivative of f with respect to $\|\cdot\|_k$, i.e., g_k is the vector in H satisfying

$$f(x)=f(x_k)+\langle g_k, x-x_k\rangle_k+o(\|x-x_k\|_k),\tag{3}$$

where $\langle\cdot,\cdot\rangle_k$ denotes the inner product corresponding to $\|\cdot\|_k$.

As an example let $H=R^n$, and B_k be an $n\times n$ positive definite symmetric matrix. Consider the inner product and norm corresponding to B_k

$$\langle x, y\rangle_k=x'B_k y,\qquad \|x\|_k=(\langle x, x\rangle_k)^{1/2}\quad \forall x, y\in H,\tag{4}$$

where all vectors above are considered to be column vectors and prime denotes transposition. With respect to this norm we have (cf. (3))

$$g_k=B_k^{-1}\nabla f(x_k),\tag{5}$$

* Received by the editors August 8, 1982, and in revised form July 27, 1983.

† This work was supported by the National Science Foundation under grant NSF/ECS 79-20834. Computer Science Department, University of California, Los Angeles, California 90024.

‡ Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

where $\nabla f(x_k)$ is the vector of first partial derivatives of f

$$\nabla f(x_k) = \begin{bmatrix} \dfrac{\partial f(x_k)}{\partial x^1} \\ \vdots \\ \dfrac{\partial f(x_k)}{\partial x^n} \end{bmatrix}. \tag{6}$$

When problem (1) is unconstrained ($X = H$), iteration (2) takes the familiar form

$$x_{k+1} = x_k - \alpha_k B_k^{-1} \nabla f(x_k).$$

Otherwise the vector

$$x_{k+1} = P_k(x_k - \alpha_k g_k)$$

is the solution of the problem

$$\begin{aligned} &\text{minimize } \|x - x_k + \alpha_k g_k\|_k^2 \\ &\text{subject to } x \in X. \end{aligned}$$

A straightforward computation using (4) and (5) shows that the problem above is equivalent to the problem

$$\begin{aligned} &\text{minimize } \nabla f(x_k)'(x - x_k) + \frac{1}{2\alpha_k}(x - x_k)' B_k (x - x_k) \\ &\text{subject to } x \in X. \end{aligned} \tag{7}$$

When X is a polyhedral set and B_k is a quasi-Newton approximation of the Hessian of f, the resulting method is closely related to recursive quadratic programming methods which currently enjoy a great deal of popularity (e.g., Garcia-Palomares [3], Gill et al. [4]).

It is generally recognized that in order for the methods above to be effective it is essential that the computational overhead for solving the quadratic programming problem (7) should not be excessive. For large-scale problems this overhead can be greatly reduced if the matrix B_k is chosen in a way that matches the structure of the constraint set. For example if X is the Cartesian product $\prod_{i=1}^m X_i$ of m simpler sets X_i, the matrix B_k can be chosen to be block diagonal with one block corresponding to each set X_i, in which case the projection problem (7) decomposes naturally. Unfortunately, such a choice of B_k precludes the possibility of superlinear convergence of the algorithm, which typically cannot be achieved unless B_k is chosen to be a suitable approximation of the Hessian matrix of f [3], [5].

The purpose of this paper is to propose projection methods of the form

$$x_{k+1} = P(x_k - \alpha_k g_k) \tag{8}$$

where the norms $\|\cdot\|$ *and* $\|\cdot\|_k$ *corresponding to the projection and the differentiation operators respectively can be different.* This allows the option to choose $\|\cdot\|$ to match the structure of X, thereby making the projection operation computationally efficient, while reserving the option to choose $\|\cdot\|_k$ on the basis of second derivatives of f thereby making the algorithm capable of superlinear convergence. When $H = R^n$, the projection norm $\|\cdot\|$ is the standard Euclidean norm

$$\|x\| = (x'x)^{1/2} = |x|, \tag{9}$$

and the derivative norm $\|\cdot\|_k$ is specified by an $n \times n$ positive definite symmetric matrix B_k

$$\|x\|_k = (x' B_k x)^{1/2}, \tag{10}$$

the vector x_{k+1} of (8) is obtained by solving the quadratic programming subproblem

$$\begin{aligned} &\text{minimize } g_k'(x - x_k) + \frac{1}{2\alpha_k}|x - x_k|^2 \\ &\text{subject to } x \in X \end{aligned} \tag{11}$$

where

$$g_k = B_k^{-1} \nabla f(x_k). \tag{12}$$

The quadratic programming problem (11) may be very easy to solve if X has special structure. As an example consider the case of an orthant constraint

$$X = \{x | 0 \leqq x^i, i = 1, \cdots, n\}. \tag{13}$$

Then, the iteration takes the form

$$x_{k+1} = [x_k - \alpha_k B_k^{-1} \nabla f(x_k)]^+ \tag{14}$$

where for any vector $v \in R^n$ with coordinates v^i, $i = 1, \cdots, n$ we denote by v^+ the vector with coordinates

$$(v^i)^+ = \max\{0, v^i\}.$$

Iteration (14) was first proposed in Bertsekas [6], and served as the starting point for the present paper. It was originally developed for use in a practical application reported in [18]. The computational overhead involved in (14) is much smaller than the one involved in solving the corresponding quadratic program (7) particularly for problems of large dimension. Indeed large optimal control problems have been solved using (14) (see [6]) that, in our view, would be impossible to solve by setting up the corresponding quadratic programming (7) and using standard pivoting techniques. Similarly (14) holds an important advantage over active set methods [4] where only one constraint is allowed to enter the active set at each iteration. Such methods require at least as many iterations as the number of active constraints at the optimal solution which are not active at the starting vector, and are in our view a poor choice for problems of very large dimension.

An important point is that it is not true in general that for an arbitrary positive definite choice B_k, iteration (14) is a descent iteration (in the sense that if x_k is not a critical point, then for α_k sufficiently small we have $f(x_{k+1}) < f(x_k)$). Indeed this is the main difficulty in constructing two-metric extensions of the Goldstein–Levitin–Poljak method. It was shown, however, in [6] (see also [19]) that if B_k is chosen to be partially diagonal with respect to a suitable subset of coordinates, then (14) becomes a descent iteration. We give a nontrivial extension of this result in the next section (Proposition 1). The construction of the "scaled gradient" g_k satisfying the descent condition

$$\langle g_k, \nabla f(x_k) \rangle > 0 \tag{15}$$

is based on a decomposition of the negative gradient into two orthogonal components by projection on an appropriate pair of cones that are dual to each other. One of the two components is then "scaled" by multiplication with a positive definite self-adjoint operator (which may incorporate second derivative information) and added to the first

component to yield g_k. The method of construction is such that g_k, in addition to (15), also satisfies

$$f[P(x_k - \alpha g_k)] < f(x_k)$$

for all α in an interval $(0, \bar{\alpha}_k]$, $\bar{\alpha}_k > 0$.

Section 3 describes the main algorithm and proves its convergence. While other stepsize rules are possible, we restrict attention to an Armijo-like stepsize rule for selecting α_k on the arc

$$\{z | z = P(x_k - \alpha g_k),\ \alpha > 0\}$$

which is patterned after similar rules proposed in Bertsekas [6], [7]. Variations of the basic algorithm are considered in § 5, while in § 4 we consider rate of convergence aspects of algorithm (8), (11), (12) as applied to finite dimensional problems. We show that the descent direction g_k can be constructed on the basis of second derivatives of f so that the method has a typically superlinear rate of convergence. Here we restrict attention to Newton-like versions of the algorithm. Quasi–Newton, and approximate Newton implementations based on successive overrelaxation or conjugate gradient methods are also possible. A superlinearly convergent conjugate gradient-based implementation is applied to a large-scale multicommodity flow problem in the last section of the paper.

While the algorithm is stated and analyzed in general terms, we pay special attention to the case where X is a finite dimensional polyhedral set with a decomposable structure since we believe that this is the case where the algorithm of this paper is most likely to find application.

2. The algorithmic map and its descent properties. Consider the problem

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } x \in X \end{aligned} \tag{16}$$

where f is a real-valued function on a Hilbert space H, and X is a nonempty, closed, convex subset of H. The inner product and norm on H will be denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ respectively. We say that two vectors $x, y \in H$ are orthogonal if $\langle x, y \rangle = 0$. For any $z \in H$ we denote by $P(z)$ the unique projection of z on X, i.e.,

$$P(z) = \arg\min \{\|x - z\| \mid x \in X\}. \tag{17}$$

We assume that f is continuously Frechet differentiable on H. The Frechet derivative at a vector $x \in H$ will be denoted by $\nabla f(x)$. It is the unique vector in H satisfying

$$f(z) = f(x) + \langle \nabla f(x), z - x \rangle + o(\|z - x\|)$$

where $o(\|z - x\|)/\|z - x\| \to 0$ as $z \to x$. We say that a vector $x^* \in X$ is *critical* with respect to problem (16) if

$$\langle \nabla f(x^*), x - x^* \rangle \geqq 0 \quad \forall x \in X, \tag{18}$$

or equivalently, if $x^* = P[x^* - \nabla f(x^*)]$.

It will be convenient for our purposes to represent the set X as an intersection of half spaces

$$X = \{x | \langle a_i, x \rangle \leqq b_i,\ \forall i \in I\}, \tag{19}$$

where I is a, possibly infinite, index set and, for each $i \in I$, a_i is a nonzero vector in H and b_i is a scalar. For each closed convex set X there exists at least one such

representation. We will assume that the set I is nonempty—the case where I is empty corresponds to an unconstrained problem which is not the subject of this paper. *Our algorithm will be defined in terms of a specific collection* $\{(a_i, b_i) | i \in I\}$ *satisfying* (19) *which will be assumed given.* This is not an important restriction for many problems of interest including, of course, the case where X is a polyhedron in R^n.

We now describe the algorithmic mapping on which our method is based. For a given vector $x \in X$ we will define an arc of points $\{x(\alpha) | \alpha \geqq 0\}$ which depends on an index set $I_x \subset I$ and an operator D_x which will be described further shortly. The index set I_x is required to satisfy

$$I_x \supset \{i \in I | \langle a_i, x \rangle \geqq b_i - \varepsilon \|a_i\|\} \tag{20}$$

where ε is some positive scalar. Let C_x be the cone defined by

$$C_x = \{z | \langle a_i, z \rangle \leqq 0, \forall i \in I_x\} \tag{21}$$

and C_x^+ be the dual cone of C_x

$$C_x^+ = \{z | \langle y, z \rangle \leqq 0, \forall y \in C_x\}. \tag{22}$$

For orientation purposes we mention that if X is a polyhedral subset of R^n (or more generally if the index set I is finite), and ε is sufficiently small, then I_x can consist of the indexes of the active constraints at x, i.e., we may take $I_x = \{i | \langle a_i, x \rangle = b_i, i \in I\}$. In that case C_x is the cone of feasible directions at x, while C_x^+ is the cone generated by the vectors a_i corresponding to the active constraints at x. More generally C_x is a (possibly empty) subset of the set of feasible directions at x, and for any $\Delta x \in C_x$ with $\|\Delta x\| \leqq \varepsilon$ the vector $x + \Delta x$ belongs to X.

Let d_x be the projection of $[-\nabla f(x)]$ on C_x, i.e.,

$$d_x = \arg\min \{\|z + \nabla f(x)\| | z \in C_x\}. \tag{23}$$

Define

$$d_x^+ = -[\nabla f(x) + d_x]. \tag{24}$$

It can be easily seen that the vectors d_x and d_x^+ are orthogonal and that d_x^+ is the projection of $[-\nabla f(x)]$ on C_x^+, i.e.,

$$d_x^+ = \arg\min \{\|z + \nabla f(x)\| | z \in C_x^+\}. \tag{25}$$

Note that if the norm $\|\cdot\|$ on H is such that projection on the set X is relatively simple, then typically the same is true for the projection (23), required to compute d_x and d_x^+.

Let Γ_x be the subspace spanned by the elements of C_x which are orthogonal to d_x^+, i.e.,

$$\Gamma_x = \text{span}\{C_x \cap \{z | \langle z, d_x^+ \rangle = 0\}\}. \tag{26}$$

Note that

$$d_x \in \Gamma_x \tag{27}$$

since d_x belongs to C_x and is orthogonal to d_x^+. Let $D_x : \Gamma_x \to \Gamma_x$ be a positive definite self-adjoint operator mapping Γ_x into itself. Consider the projection $\tilde{d}_x$ of $D_x d_x$ on the closed cone $C_x \cap \{z | \langle z, d_x^+ \rangle = 0\}$, i.e.

$$\tilde{d}_x = \arg\min \{\|z - D_x d_x\| | z \in C_x, \langle z, d_x^+ \rangle = 0\}. \tag{28}$$

Consider also the direction vector

$$g = -(d_x^+ + \tilde{d}_x). \tag{29}$$

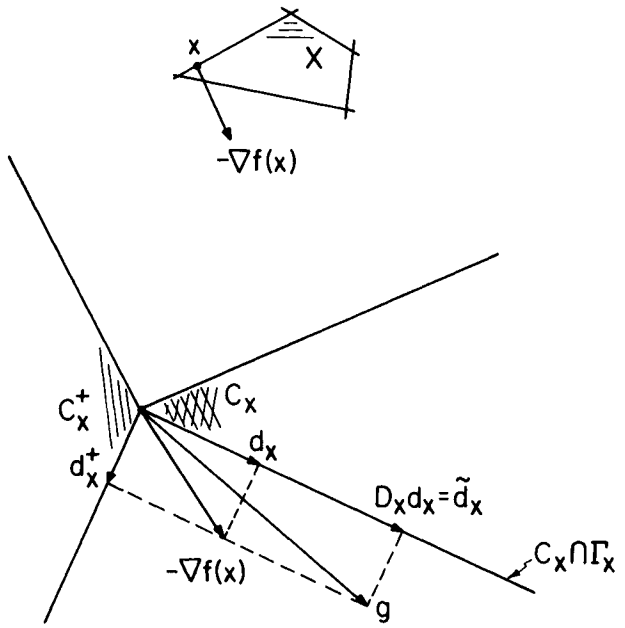

FIG. 1. *A case where both C_x and C_x^+ have nonempty interior in R^2 and $-\nabla f$ lies outside C_x.*

Given x, I_x, and D_x, our algorithm chooses the next iterate along the arc

$$x(\alpha) = P(x - \alpha g), \qquad \alpha \geqq 0. \tag{30}$$

The stepsize α will be chosen by an Armijo-like stepsize rule that will be described in the next section.

The process by means of which the direction g is obtained is illustrated in Figs. 1–4. The crucial fact that will be shown in Proposition 1 below is that, if x is not critical, then for sufficiently small $\alpha > 0$ we have $f[x(\alpha)] < f(x)$, i.e., by moving along the arc $x(\alpha)$ of (30) we can decrease the value of the objective. Furthermore we have $\langle \nabla f(x), g \rangle > 0$ which means that g can be viewed as a "scaled" gradient, i.e., the product of $\nabla f(x)$ with a positive definite self-adjoint operator. We now demonstrate the process of calculating the direction g for some interesting specially structured constraint sets.

Example 1. Let $H = R^n$, $\langle x, y \rangle = x'y$, and X be the positive orthant

$$X = \{z | x^i \geqq 0,\ i = 1, \cdots, n\}.$$

Then X consists of the intersection of the n halfspaces $\{x | x^i \geqq 0\}$ $i = 1, \cdots, n$ and is of the form (19). The set I_x must contain all indices i such that $0 \leqq x^i \leqq \varepsilon$ (cf. (20)). The cones C_x and C_x^+ are given by

$$C_x = \{z | z^i \geqq 0,\ \forall i \in I_x\}, \qquad C_x^+ = \{z | z^i \leqq 0,\ \forall i \in I_x,\ z^j = 0,\ \forall j \notin I_x\}.$$

The vector d_x, and d_x^+ (cf. (23), (24)) have coordinates given by

$$d_x^i = \begin{cases} -\dfrac{\partial f(x)}{\partial x^i} & \text{if } i \notin \hat{I}_x, \\ 0 & \text{if } i \in \hat{I}_x, \end{cases} \qquad d_x^{i+} = \begin{cases} 0 & \text{if } i \notin \hat{I}_x, \\ -\dfrac{\partial f(x)}{\partial x^i} & \text{if } i \in \hat{I}_x, \end{cases}$$

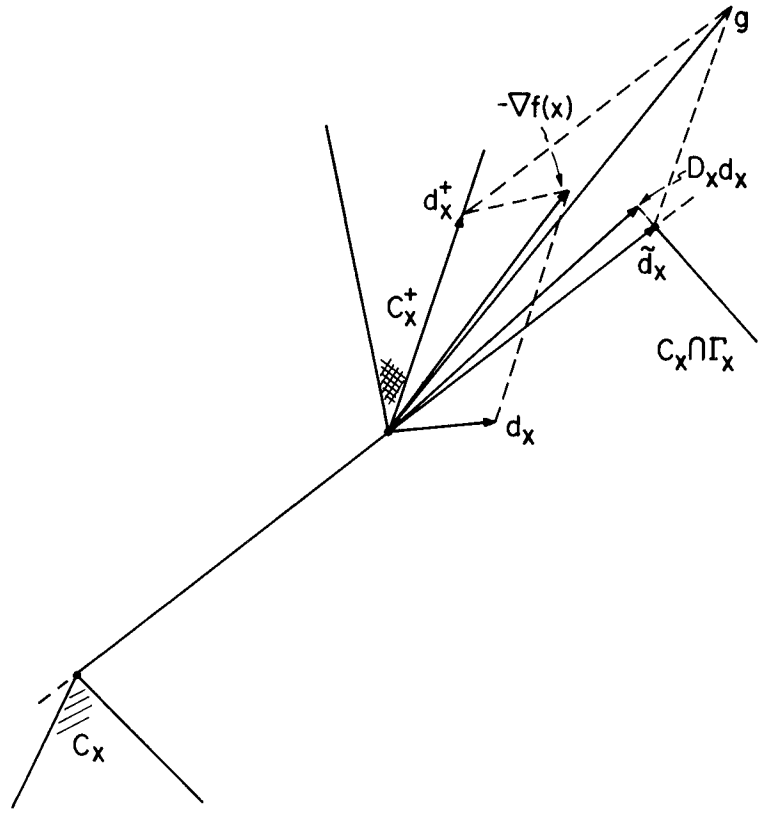

FIG. 2. *Obtaining g for a case where C_x^+ lies on a two-dimensional manifold in R^3.*

where

$$\hat{I}_x = \left\{ i \,\middle|\, i \in I_x \text{ and } \frac{\partial f(x)}{\partial x^i} > 0 \right\}.$$

If $\hat{I}_x$ is empty then $\Gamma_x = R^n$ and we have $d_x = -\nabla f(x)$, $d_x^+ = 0$. In this case $g = -D_x d_x = D_x \nabla f(x)$ where D_x is any $n \times n$ positive definite symmetric matrix. If I_x is not empty, by rearranging indices if necessary assume that for some integer p with $0 \leqq p \leqq n-1$ we have $\hat{I}_x = \{p+1, \cdots, n\}$. Partition $\nabla f(x)$ as

$$\nabla f(x) = \begin{bmatrix} \tilde{w} \\ \hat{w} \end{bmatrix}$$

where $\tilde{w} \in R^p$ and $w \in R^{n-p}$. The vector g is given by

$$g = \begin{bmatrix} (D_x \tilde{w})^{\#} \\ \hat{w} \end{bmatrix}$$

where D_x is a $p \times p$ positive definite symmetric matrix, $(D_x \tilde{w})^{\#}$ denotes projection of $D_x \tilde{w}$ on C_x, i.e., $(D_x \tilde{w})^{\#}$ is obtained from $D_x \tilde{w}$ by setting to zero those coordinates of $D_x \tilde{w}$ which are negative and whose indices belong to I_x.

Example 2. Let $H = R^n$, and X be the unit simplex

$$X = \left\{ x \,\middle|\, \sum_{i=1}^{n} x^i = 1, x^i \geqq 0, i = 1, \cdots, n \right\}. \tag{31}$$

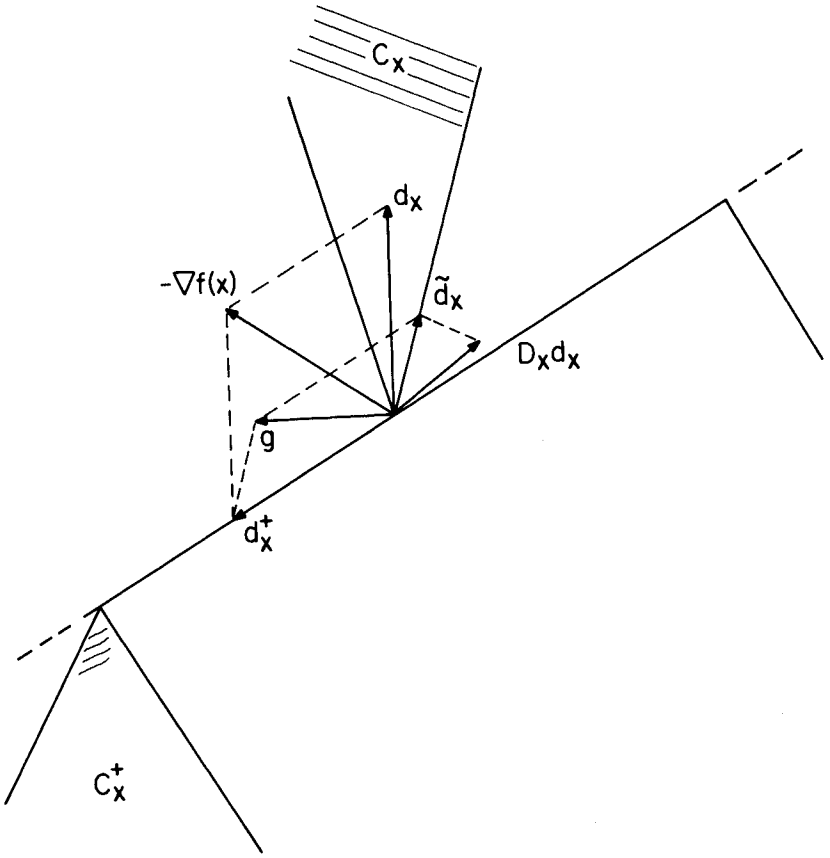

FIG. 3. *Obtaining g for a case where C_x lies on a two-dimensional manifold in R^3.*

Suppose the inner product on R^n is taken to be

$$\langle x, y\rangle = \sum_{i=1}^{n} s^i x^i y^i \tag{32}$$

where s^i, $i = 1, \cdots, n$ are some positive scalars. Let $\hat{I}_x$ be a set of indices including those indices i such that $0 \leqq x^i \leqq \varepsilon/\sqrt{s^i}$. Then the cone C_x can be taken to be

$$C_x = \left\{ z \,\middle|\, \sum_{i=1}^{n} z^i = 0,\ z^i \geqq 0,\ \forall i \in \hat{I}_x \right\}. \tag{33}$$

The vector d_x is obtained as the solution of the projection problem

$$\begin{aligned} &\text{minimize } \frac{1}{2} \sum_{i=1}^{n} s^i \left[z^i + \frac{1}{s^i} \frac{\partial f(x)}{\partial x^i} \right]^2 \\ &\text{subject to } \sum_{i=1}^{n} z^i = 0, \quad z^i \geqq 0, \quad i \in \hat{I}_x. \end{aligned} \tag{34}$$

The solution of this problem is very simple. By introducing a Lagrange multiplier λ for the equality constraint $\sum_{i=1}^{n} z^i = 0$, we obtain that λ is the solution of the piecewise

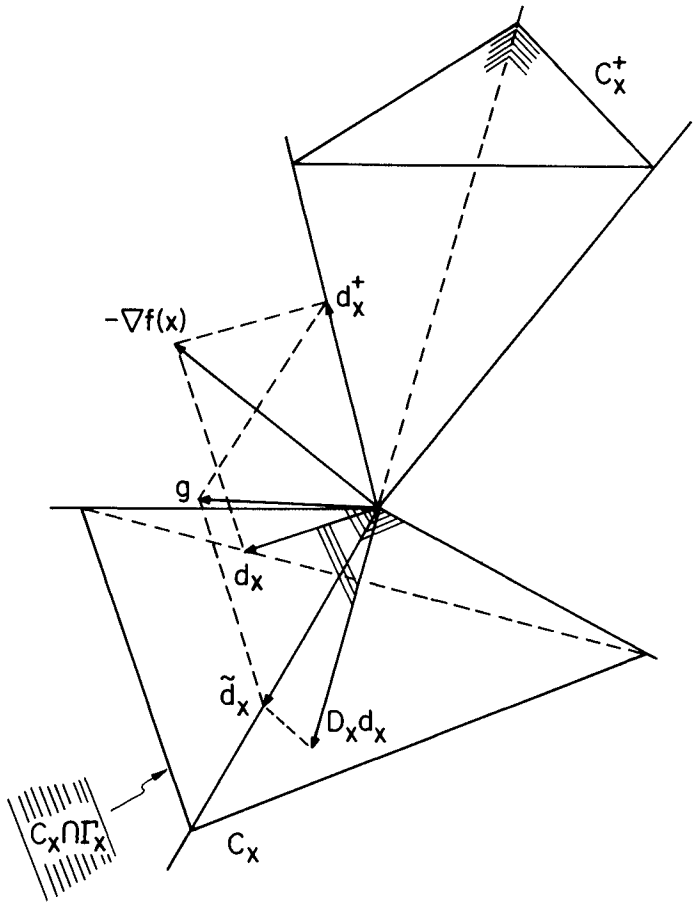

FIG. 4. *Obtaining g for a case where both C_x^+ and C_x have nonempty interior in R^3.*

linear equation

$$\sum_{i\in\hat{I}_x}\frac{1}{s^i}\left[\lambda-\frac{\partial f(x)}{\partial x^i}\right]^+ + \sum_{i\notin\hat{I}_x}\frac{1}{s^i}\left[\lambda-\frac{\partial f(x)}{\partial x^i}\right]=0. \tag{35}$$

This equation can be solved by the well-known method of sorting the breakpoints $\partial f(x)/\partial x^i$, $i\in\hat{I}_x$ in decreasing order, and testing the values of the left side at the breakpoints until two successive values bracket zero. Once λ is obtained, the coordinates of d_x are given by

$$d_x^i=\begin{cases}\dfrac{1}{s^i}\left[\lambda-\dfrac{\partial f(x)}{\partial x^i}\right]^+ & \text{if } i\in\hat{I}_x,\\[2ex] \dfrac{1}{s^i}\left[\lambda-\dfrac{\partial f(x)}{\partial x^i}\right] & \text{if } i\notin\hat{I}_x.\end{cases} \tag{36}$$

The vector d_x^+ is then obtained from the equation

$$d_x^+=-[\nabla f(x)+d_x].$$

Let

$$\tilde{I}_x=\left\{i\,\middle|\, i\in\hat{I}_x \text{ and } \lambda<\frac{\partial f(x)}{\partial x^i}\right\}. \tag{37}$$

It is easily verified that the subspace Γ_x is given by

$$\Gamma_x = \left\{ z \,\middle|\, \sum_{i=1}^{n} z^i = 0,\ z^i = 0,\ \forall i \in \tilde{I}_x \right\}. \tag{38}$$

The vector $\tilde{d}_x$ is obtained as the solution of the simple projection problem.

$$\begin{aligned} &\text{minimize } \frac{1}{2} \sum_{i=1}^{n} s^i [z^i - (D_x d_x)^i]^2 \\ &\text{subject to } \sum_{i=1}^{n} z^i = 0,\ z^i \geqq 0 \ \ \forall i \in \hat{I}_x, \quad z^j = 0 \ \forall j \in \tilde{I}_x \end{aligned} \tag{39}$$

where $(D_x d_x)^i$ is the ith coordinate of the vector $D_x d_x$ obtained by multiplying d_x with an $n \times n$ symmetric matrix D_x which maps Γ_x into Γ_x and is positive definite on Γ_x. We will comment further on the choice of D_x in the last section of the paper. The vector g is given now by $g = -(\tilde{d}_x + d_x^+)$. Note that the solution of both projection problems (34) and (39), as well as the problem of projection on the simplex X of (31) is greatly simplified by the choice of the "diagonal" metric specified by (32).

Proposition 1 below is the main result regarding the algorithmic map specified by (20)–(24), (28)–(30). For its proof we will need the following lemma, the proof of which is given in Appendix A.

LEMMA 1. *Let Ω be a closed convex subset of a Hilbert space H, and let $P_\Omega(\cdot)$ denote projection on Ω. For every $x \in \Omega$ and $z \in H$:*

a) *The function $h:(0, \infty) \to R$ defined by*

$$h(\alpha) = \frac{\| P_\Omega(x + \alpha z) - x \|}{\alpha} \quad \forall \alpha > 0$$

is monotonically nonincreasing.

b) *If y is any direction of recession of Ω (i.e., $(x + \alpha y) \in \Omega$ for all $\alpha \geqq 0$), then*

$$\langle y, x + z \rangle \leqq \langle y, P_\Omega(x + z) \rangle. \tag{40}$$

PROPOSITION 1. *For $x \in X$, let $\varepsilon > 0$ and I_x satisfy* (20), *and let $D_x : \Gamma_x \to \Gamma_x$ be a positive definite self-adjoint operator on the subspace Γ_x defined by* (21)–(26). *Consider the arc $\{x(\alpha) | \alpha \geqq 0\}$ defined by* (23), (24), (28)–(30).

a) *If x is critical, then*

$$x(\alpha) = x \quad \forall \alpha \geqq 0.$$

b) *If x is not critical, then*

$$\langle \nabla f(x), g \rangle > 0, \tag{41}$$

and

$$\langle \nabla f(x), x - x(\alpha) \rangle \geqq \alpha \langle d_x, D_x d_x \rangle + \frac{1}{\alpha} \| x(\alpha) - (x + \alpha \tilde{d}_x) \|^2 > 0 \quad \forall \alpha \in \left(0, \frac{\varepsilon}{\| g \|}\right). \tag{42}$$

Furthermore there exists $\bar{\alpha} > 0$ such that

$$f(x) > f[x(\alpha)] \quad \forall \alpha \in (0, \bar{\alpha}]. \tag{43}$$

Proof. a) It is easily seen that for every $z \in C_x$ we have

$$\left(x + \frac{\varepsilon}{\| z \|} z \right) \in X \tag{44}$$

in view of the definitions (19)–(21). Since x is critical, we have $\langle \nabla f(x), y-x\rangle \geqq 0$ for all $y \in X$. Therefore using (44) we have

$$\langle \nabla f(x), z\rangle \geqq 0 \quad \forall z \in C_x. \tag{45}$$

From the definitions of C_x^+, d_x and d_x^+ (cf. (21)–(24)) and (45) it follows that

$$-\nabla f(x) \in C_x^+$$

and

$$d_x^+ = -\nabla f(x), \qquad d_x = 0.$$

Using (28)–(30), we obtain $x(\alpha) = P[x - \alpha \nabla f(x)]$. Since x is critical, we have that $x = P[x - \alpha \nabla f(x)]$ for all $\alpha \geqq 0$ and the conclusion follows.

b) We have by using the facts $\nabla f(x) = -(d_x + d_x^+)$ and $\langle \tilde{d}_x, d_x^+\rangle = 0$

$$\langle \tilde{d}_x, \nabla f(x)\rangle = -\langle \tilde{d}_x, d_x + d_x^+\rangle = -\langle \tilde{d}_x, d_x\rangle. \tag{46}$$

Now $\tilde{d}_x$ is the projection of $D_x d_x$ on the cone $C_x \cap \{z \mid \langle z, d_x^+\rangle = 0\}$, d_x belongs to this cone and therefore is a direction of recession. Using Lemma 1b), it follows that

$$\langle d_x, \tilde{d}_x\rangle \geqq \langle d_x, D_x d_x\rangle. \tag{47}$$

Combining (46) and (47), we obtain

$$\langle \tilde{d}_x, \nabla f(x)\rangle \leqq -\langle d_x, D_x d_x\rangle \leqq 0 \tag{48}$$

where the second inequality is strict if and only if $d_x \neq 0$. Also d_x^+ is the projection of $-\nabla f(x)$ on C_x^+, so

$$\langle d_x^+, \nabla f(x)\rangle \leqq 0 \tag{49}$$

with strict inequality if and only if $d_x^+ \neq 0$. Combining (48) and (49) and using the fact $g = -(d_x^+ + \tilde{d}_x)$, we obtain

$$\langle g, \nabla f(x)\rangle \geqq 0 \tag{50}$$

with equality if and only if $d_x = 0$ and $d_x^+ = 0$, or, equivalently $\nabla f(x) = 0$. Since x is not critical, we must have $\nabla f(x) \neq 0$, so strict inequality holds in (50) and (41) is proved.

Take any $\alpha \in (0, \varepsilon / \|g\|)$. Since projection on a closed convex set is a nonexpansive operator (see e.g. [8] or use the Cauchy–Schwarz inequality to strengthen (B.16) in Appendix B), we have

$$\|x(\alpha) - x\| \leqq \|x - \alpha g - x\| = \alpha \|g\| < \varepsilon. \tag{51}$$

Therefore we have

$$\langle a_i, x\rangle < b_i - \varepsilon \|a_i\| < b_i - \langle a_i, x(\alpha) - x\rangle \quad \forall i \notin I_x$$

and as a result

$$\langle a_i, x(\alpha)\rangle < b_i \quad \forall i \notin I_x.$$

It follows that $x(\alpha)$ is also the projection of the vector $x - \alpha g$ on the set $\Omega_x \supset X$ given by

$$\Omega_x = \{z \mid \langle a_i, z\rangle \leqq b_i, i \in I_x\},$$

i.e.,

$$x(\alpha) = \arg\min \{\|z - (x - \alpha g)\| \mid z \in \Omega_x\}. \tag{52}$$

Now the vector d_x is easily seen to be a direction of recession of the set Ω_x, so by Lemma 1b) we have

$$\langle d_x, x(\alpha)\rangle \geqq \langle d_x, x-\alpha g\rangle = \langle d_x, x+\alpha d_x^+ + \alpha \tilde{d}_x\rangle.$$

Since $\langle d_x, d_x^+\rangle = 0$, the relation above is written by using also (47)

$$-\langle d_x, x-x(\alpha)\rangle \geqq \alpha \langle d_x, D_x d_x\rangle. \tag{53}$$

In view of the fact $\tilde{d}_x \in C_x$ we have $(x+\alpha\tilde{d}_x)\in\Omega_x$, and since $x(\alpha)$ is the projection on Ω_x of $(x+\alpha d_x^+ + \alpha\tilde{d}_x)$ (cf. (52)), we have

$$\langle x+\alpha d_x^+ + \alpha\tilde{d}_x - x(\alpha), x+\alpha\tilde{d}_x - x(\alpha)\rangle \leqq 0.$$

Equivalently, using the fact $\langle d_x^+, \tilde{d}_x\rangle = 0$,

$$-\langle d_x^+, x-x(\alpha)\rangle \geqq \frac{\|x(\alpha)-(x+\alpha\tilde{d}_x)\|^2}{\alpha}. \tag{54}$$

By combining (53) and (54) and using the fact $\nabla f(x) = -(d_x + d_x^+)$, we obtain

$$\langle \nabla f(x), x-x(\alpha)\rangle \geqq \alpha\langle d_x, D_x d_x\rangle + \frac{\|x(\alpha)-(x+\alpha\tilde{d}_x)\|^2}{\alpha} \tag{55}$$

which is the left inequality in (42). To show that the right side of (55) cannot be zero, note that if it were, then we would have both $d_x = 0$ (implying $\tilde{d}_x = 0$, $x(\alpha) = P(x-\alpha\nabla f(x))$ and $x(\alpha) = x+\alpha\tilde{d}_x$ (implying $P(x-\alpha\nabla f(x)) = x$). Since x is not critical, we arrive at a contradiction. Therefore the right inequality in (42) is also proved.

By using the mean value theorem, we have

$$f(x)-f[x(\alpha)] = \langle \nabla f(x), x-x(\alpha)\rangle + \langle \nabla f(\zeta_\alpha) - \nabla f(x), x-x(\alpha)\rangle \tag{56}$$

where ζ_α lies on the line segment joining x and $x(\alpha)$. Using (55) and (56), we obtain for all $\alpha \in (0, \varepsilon/\|g\|)$

$$\begin{aligned}\frac{1}{\alpha}\{f(x)-f[x(\alpha)]\} \geqq \langle d_x, D_x d_x\rangle + \frac{\|x(\alpha)-(x+\alpha\tilde{d}_x)\|^2}{\alpha^2} \\ + \left\langle \nabla f(\zeta_\alpha) - \nabla f(x), \frac{x-x(\alpha)}{\alpha}\right\rangle.\end{aligned} \tag{57}$$

Using (51) and the Cauchy–Schwarz inequality, we see that

$$\left\langle \nabla f(\zeta_\alpha) - \nabla f(x), \frac{x-x(\alpha)}{\alpha}\right\rangle \geqq -\|\nabla f(\zeta_\alpha) - \nabla f(x)\| \cdot \|g\|. \tag{58}$$

Since $\|\nabla f(\zeta_\alpha) - \nabla f(x)\| \to 0$ as $\alpha \to 0$, we see from (57) and (58) that if $d_x \neq 0$ then for all positive but sufficiently small α we have $f(x) > f[x(\alpha)]$. If $d_x = 0$ then $\tilde{d}_x = 0$ and using Lemma 1a)

$$\frac{\|x(\alpha)-(x+\alpha\tilde{d}_x)\|^2}{\alpha^2} = \frac{\|x(\alpha)-x\|^2}{\alpha^2} \geqq \|x(1)-x\|^2 \quad \forall \alpha \in (0, 1]. \tag{59}$$

From (57), (58) and (59) we see again that when $d_x = 0$, then for all positive but sufficiently small α we have $f(x) > f[x(\alpha)]$. Therefore, there exists $\bar{\alpha} > 0$ such that (43) holds in both cases where $d_x = 0$ and $d_x \neq 0$. Q.E.D.

3. Convergence analysis. The previous section has shown how a vector $x \in X$, a scalar $\varepsilon > 0$, an index set I_x satisfying

$$I_x \supset \{i \in I | \langle a_i, x \rangle \geqq b_i - \varepsilon \| a_i \| \},$$

and a positive definite self-adjoint operator $D_x : \Gamma_x \to \Gamma_x$ where Γ_x is the subspace defined by (21)–(26), uniquely define an arc of points $x(\alpha) \in X$, $\alpha \geqq 0$ where

$$x(\alpha) = P(x - \alpha g), \qquad \alpha \geqq 0$$

and g is defined via (23), (24), (28)–(30). Furthermore for each $x \in X$ which is not critical, Proposition 1b) shows that by choosing α sufficiently small, we can obtain a point of lower cost on this arc. Therefore any procedure that, for any given $x \in X$, chooses I_x, ε, and D_x satisfying the above requirements, coupled with a rule for selecting a point of lower cost on the corresponding arc $x(\alpha)$ leads to a descent algorithm. There is a large variety of possibilities along these lines but we will focus attention on the following broad class of methods:

We assume that we are given a continuous function $\varepsilon : X \to R$ such that

$$\varepsilon(x) \geqq 0 \quad \forall x \in X, \tag{60}$$

$$\varepsilon(x) = 0 \Rightarrow x \text{ is critical} \tag{61}$$

(for example $\varepsilon(x) = \min \{\varepsilon, \| x - P[x - \nabla f(x)] \|\}$ where $\varepsilon > 0$ is a given constant). We are also given scalars $\beta \in (0, 1)$, $\sigma \in (0, 1/2)$, $\lambda_1 > 0$ and $\lambda_2 > 0$ with $\lambda_1 \leqq \lambda_2$.

At the beginning of the kth iteration of the algorithm we have a vector $x_k \in X$. If x_k is critical, we set $x_{k+1} = x_k$. Else we obtain the next vector x_{k+1} as follows:

Step 1. Choose an index set $I_k \subset I$ satisfying

$$I_k \supset \{i \in I | \langle a_i, x_k \rangle \geqq b_i - \varepsilon(x_k) \| a_i \| \}, \tag{62}$$

and compute

$$d_k = \arg\min \{ \| z + \nabla f(x_k) \| \, | \, z \in C_k \}, \tag{63}$$

$$d_k^+ = -[\nabla f(x_k) + d_k] \tag{64}$$

where

$$C_k = \{ z | \langle a_i, z \rangle \leqq 0, \ i \in I_k \}. \tag{65}$$

Step 2. Choose a positive definite self-adjoint operator $D_k : \Gamma_k \to \Gamma_k$, where

$$\Gamma_k = \text{span} \{ C_k \cap \{ z | \langle z, d_k^+ \rangle = 0 \} \}, \tag{66}$$

and D_k satisfies

$$\| D_k \| \leqq \lambda_2 \quad \text{and} \quad \lambda_1 \| z \|^2 \leqq \langle z, D_k z \rangle \quad \forall z \in \Gamma_k. \tag{67}$$

Compute $\tilde{d}_k$ given by

$$\tilde{d}_k = \arg\min \{ \| z - D_k d_k \| \, | \, z \in C_k, \langle z, d_k^+ \rangle = 0 \}. \tag{68}$$

Define

$$g_k = -(d_k^+ + \tilde{d}_k) \tag{69}$$

and

$$x_k(\alpha) = P(x_k - \alpha g_k) \quad \forall \alpha \geqq 0. \tag{70}$$

Step 3. Set

$$x_{k+1} = x_k(\alpha_k) \tag{71}$$

where

$$\alpha_k = \beta^{m_k} \tag{72}$$

and m_k is the first nonnegative integer m satisfying

$$f(x_k) - f[x_k(\beta^m)] \geqq \sigma \left\{ \beta^m \langle d_k, D_k d_k \rangle + \frac{\|x_k(\beta^m) - (x_k + \beta^m \tilde{d}_k)\|^2}{\beta^m} \right\}. \tag{73}$$

Proposition 1b) shows that x_{k+1} is well defined via the stepsize rule (71)–(73) in the sense that m_k is a (finite) integer and furthermore

$$f(x_k) > f(x_{k+1})$$

for all k for which x_k is not critical. The following proposition is our main convergence result.

PROPOSITION 2. *Every limit point of a sequence* $\{x_k\}$ *generated by the algorithm above is a critical point.*

Proof. Let $\{x_k\}_K$ be a subsequence of $\{x_k\}$ converging to a point $\bar{x}$ which is not critical. We will arrive at a contradiction. Since $\{\alpha_k\}$ is bounded, we assume without loss of generality that

$$\lim_{\substack{k\to\infty \\ k\in K}} \alpha_k = \bar{\alpha}$$

where $\bar{\alpha} \in [0, 1]$. Since $\{f(x_k)\}$ decreases monotonically to $f(\bar{x})$, it follows from the form of the stepsize rule that

$$\lim_{\substack{k\to\infty \\ k\in K}} \alpha_k \langle d_k, D_x d_x \rangle = 0, \tag{74}$$

$$\lim_{\substack{k\to\infty \\ k\in K}} \frac{\|x_k(\alpha_k) - (x_k + \alpha_k \tilde{d}_k)\|^2}{\alpha_k} = 0. \tag{75}$$

We consider two cases:

Case 1 ($\bar{\alpha} > 0$). It follows from (74) and the fact $\langle d_k, D_k d_k \rangle \geqq \lambda_1 \|d_k\|^2$ (cf. (67)) that $\lim_{k\to\infty, k\in K} d_k = 0$, and therefore also

$$\lim_{\substack{k\to\infty \\ k\in K}} \tilde{d}_k = 0, \qquad \lim_{\substack{k\to\infty \\ k\in K}} d_k^+ = -\nabla f(\bar{x}).$$

By taking the limit as $k \to \infty$, $k \in K$, in the equation $x_k(\alpha_k) = P(x_k + \alpha_k d_k^+ + \alpha_k \tilde{d}_k)$, using the continuity of the P operator, we obtain

$$\lim_{\substack{k\to\infty \\ k\in K}} x_k(\alpha_k) = P[\bar{x} - \bar{\alpha} \nabla f(\bar{x})].$$

Therefore (75) yields

$$\bar{x} = P[\bar{x} - \bar{\alpha} \nabla f(\bar{x})].$$

Since $\bar{\alpha} > 0$ this implies that $\bar{x}$ is critical, thereby contradicting our earlier assumption.

Case 2 ($\bar{\alpha}=0$). It follows that for all $k\in K$ which are sufficiently large

$$f(x_k)-f\left[x_k\left(\frac{\alpha_k}{\beta}\right)\right]<\sigma\left\{\frac{\alpha_k}{\beta}\langle d_k, D_k d_k\rangle+\frac{\|x_k(\alpha_k/\beta)-(x_k+(\alpha_k/\beta)\tilde{d}_k)\|^2}{\alpha_k/\beta}\right\}, \tag{76}$$

i.e., the test (73) of the stepsize rule will be failed at least once for all $k\in K$ sufficiently large.

Since $g_k=-(d_k^+ + \tilde{d}_k)$, $\langle d_k^+, \tilde{d}_k\rangle=0$, we have

$$\|g_k\|^2=\|d_k^+\|^2+\|\tilde{d}_k\|^2. \tag{77}$$

Since $\tilde{d}_k$ is the projection of $D_k d_k$ on $C_x\cap\{z|\langle z, d_k^+\rangle=0\}$, we must have $\|\tilde{d}_k\|\leqq\|D_k d_k\|$ and, using (67), $\|\tilde{d}_k\|\leqq\lambda_2\|d_k\|$. Therefore from (77) and the fact $\|d_k^+\|\leqq\|\nabla f(x_k)\|$, $\|d_k\|\leqq\|\nabla f(x_k)\|$ we obtain

$$\|g_k\|^2\leqq(1+\lambda_2^2)\|\nabla f(x_k)\|^2.$$

It follows that

$$\limsup_{\substack{k\to\infty\\ k\in K}}\|g_k\|<\infty. \tag{78}$$

We also have

$$\lim_{\substack{k\to\infty\\ k\in K}}\varepsilon(x_k)=\varepsilon(\bar{x})>0. \tag{79}$$

It follows from (78), (79) and the fact $\bar{\alpha}=0$ that for all $k\in K$ sufficiently large $\alpha_k/\beta\in(0, \varepsilon(x_k)/\|g_k\|)$ and therefore using Proposition 1b) (cf. (42)), we obtain

$$\left\langle\nabla f(x_k), x_k-x_k\left(\frac{\alpha_k}{\beta}\right)\right\rangle\geqq\frac{\alpha_k}{\beta}\langle d_k, D_k d_k\rangle+\frac{\|x_k(\alpha_k/\beta)-(x_k+(\alpha_k/\beta)\tilde{d}_k)\|^2}{\alpha_k/\beta}. \tag{80}$$

Using the mean value theorem, we have

$$f(x_k)-f\left[x_k\left(\frac{\alpha_k}{\beta}\right)\right]=\left\langle\nabla f(x_k), x_k-x_k\left(\frac{\alpha_k}{\beta}\right)\right\rangle+\left\langle\nabla f(\zeta_k)-\nabla f(x_k), x_k-x_k\left(\frac{\alpha_k}{\beta}\right)\right\rangle \tag{81}$$

where ζ_k lies on the line segment connecting x_k and $x_k(\alpha_k/\beta)$. From (76), (80), and (81) we obtain for all $k\in K$ sufficiently large

$$(1-\sigma)\left\{\langle d_k, D_k d_k\rangle+\frac{\|x_k(\alpha_k/\beta)-(x_k+(\alpha_k/\beta)\tilde{d}_k)\|^2}{(\alpha_k/\beta)^2}\right\}$$

$$\leqq\left\langle\nabla f(x_k)-\nabla f(\zeta_k), \frac{x_k-x_k(\alpha_k/\beta)}{\alpha_k/\beta}\right\rangle. \tag{82}$$

Since (cf. (51), (78)) we have

$$\limsup_{\substack{k\to\infty\\ k\in K}}\frac{\|x_k-x_k(\alpha_k/\beta)\|}{\alpha_k/\beta}\leqq\limsup_{\substack{k\to\infty\\ k\in K}}\|g_k\|<\infty$$

and

$$\lim_{\substack{k\to\infty\\ k\in K}}\|\nabla f(x_k)-\nabla f(\zeta_k)\|=0,$$

it follows that the right side of (82) tends to zero as $k \to \infty$, $k \in K$. Therefore so does the left side which implies that

$$\lim_{\substack{k\to\infty \\ k\in K}} d_k = 0, \qquad \lim_{\substack{k\to\infty \\ k\in K}} \tilde{d}_k = 0 \tag{83}$$

and

$$\lim_{\substack{k\to\infty \\ k\in K}} \frac{\|x_k(\alpha_k/\beta) - (x_k + (\alpha_k/\beta)\tilde{d}_k)\|^2}{(\alpha_k/\beta)^2} = 0. \tag{84}$$

Since it follows from (79) and (83) that there exists $\bar{k}$ such that

$$x_k + \frac{\alpha_k}{\beta}\tilde{d}_k \in X \quad \forall k \geqq \bar{k},$$

we obtain using Lemma 1a)

$$\frac{\|x_k(\alpha_k/\beta) - (x_k + (\alpha_k/\beta)\tilde{d}_k)\|^2}{(\alpha_k/\beta)^2} \geqq \left\| P\left[\left(x_k + \frac{\alpha_k}{\beta}\tilde{d}_k\right) + d_k^+\right] - \left(x_k + \frac{\alpha_k}{\beta}\tilde{d}_k\right)\right\|^2. \tag{85}$$

From (84) and (85) it follows that

$$\lim_{\substack{k\to\infty \\ k\in K}} \left\| P\left[\left(x_k + \frac{\alpha_k}{\beta}\tilde{d}_k\right) - (\nabla f(x_k) + d_k)\right] - \left(x_k + \frac{\alpha_k}{\beta}\tilde{d}_k\right)\right\|^2 = 0.$$

Using (83), we obtain

$$\|P[\bar{x} - \nabla f(\bar{x})] - \bar{x}\| = 0,$$

which contradicts the assumption that $\bar{x}$ is not critical. Q.E.D.

We mention that some of the requirements on the sequences $\{\varepsilon(x_k)\}$ and $\{D_k\}$ can be relaxed without affecting the result of Proposition 2. In place of continuity of $\varepsilon(\cdot)$ and assumption (67) it is sufficient to require that if $\{x_k\}_K$ is a subsequence converging to a noncritical point $\bar{x}$, then

$$\liminf_{\substack{k\to\infty \\ k\in K}} \varepsilon(x_k) > 0,$$

$$\liminf_{\substack{k\to\infty \\ k\in K}} \inf \{\langle z, D_k z\rangle \mid \|z\| = 1,\ z \in \Gamma_k\} > 0$$

$$\limsup_{\substack{k\to\infty \\ k\in K}} \|D_k\| < \infty.$$

This can be verified by inspection of the proof of Proposition 2.

A practically important generalization of the algorithm results if we allow the norm on the Hilbert space H to change from one iteration to the next. By this we mean that at each iteration k a new inner product $\langle \cdot, \cdot \rangle_k$ and corresponding norm $\|\cdot\|_k$ on H are considered. The statement of the algorithm and corresponding assumptions must be modified as follows:

a) The gradient $\nabla f(x_k)$ will be with respect to the current inner product $\langle \cdot, \cdot \rangle_k$ (cf. (3)).

b) The projection defining d_k, d_k^+, $\tilde{d}_k$ and the arc $x_k(\cdot)$ should be with respect to the current norm $\|\cdot\|_k$.

c) The assumptions on I_k, and D_k, and the stepsize rule should be restated in terms of the current inner product and norm.

There is no difficulty in reworking the proof of Proposition 2 for this generalized version of the algorithm provided we assume that all the norms $\|\cdot\|_k$, $k=0,1,\cdots$ are "equivalent" to the original norm $\|\cdot\|$ on H in the sense that for some $m>0$ and $M>0$ we have

$$m\|z\| \leqq \|z\|_k \leqq M\|z\| \quad \forall z \in H, \quad k=0,1,\cdots.$$

Naturally the norms $\|\cdot\|_k$ should be such that projection on X with respect to any one of them is relatively easy, for otherwise the purpose of the methodology of this paper is defeated. The motivation for considering a different inner product at each iteration stems from the fact that it is often desirable in nonlinear programming algorithms to introduce iteration-dependent scaling on the optimization variables. This is sometimes referred to as "preconditioning." The use of the operator D_k fulfills that need to a great extent but while this operator scales the component d_x of the negative gradient, it does not affect at all the second component d_x^+. The role of an iteration-dependent norm can be understood by considering situations where the index set I_k is so large that the cone C_k is empty. In this case $d_k^+ = -\nabla f(x_k)$, $\tilde{d}_k = 0$ and the kth iteration reduces to an iteration of the original Goldstein–Levitin–Poljak method, for which practical experience shows that simple, for example diagonal, scaling at each iteration can sometimes result in spectacular computational savings.

4. Rate of convergence. In this section we will analyze the rate of convergence of algorithm (62)–(73) for the case where X is polyhedral and H is finite dimensional. An important property of the Goldstein–Levitin–Poljak method (cf. [7]) is that if it generates a sequence $\{x_k\}$ converging to a strict local minimum $\bar{x}$ satisfying certain sufficiency conditions (compare with [7]), then after some index $\bar{k}$ the vectors x_k lie on the manifold of active constraints at $\bar{x}$, i.e., $x_k \in \bar{x} + N_{\bar{x}}$ where

$$N_{\bar{x}} = \{z \mid \langle a_i, z\rangle = 0, \forall i \in A_{\bar{x}}\} \tag{86}$$

and where

$$A_{\bar{x}} = \{i \mid i \in I, \langle a_i, \bar{x}\rangle = b_i\}. \tag{87}$$

Our algorithm preserves this important characteristic. Indeed, we will see that, under mild assumptions, our algorithm "identifies" the set of active constraints at the limit point in a finite number of iterations, and subsequently reduces to an unconstrained optimization method on this subspace. This brings to bear the rate of convergence results available from unconstrained optimization.

The rate of convergence analysis will be carried out under the following assumptions:

(A) H is finite dimensional, X is polyhedral, f is continuously Frechet differentiable, and ∇f is Lipschitz continuous on bounded sets, i.e., for every bounded set there exists $L>0$ such that for every x and y in X we have

$$\|\nabla f(x) - \nabla f(y)\| \leqq L\|x-y\|. \tag{88}$$

(B) $\bar{x}$ is a strict local minimum and there exists $\delta > 0$ such that

$$P(y) \in \bar{x} + N_{\bar{x}} \quad \forall y \text{ such that } \|\bar{x} - \nabla f(\bar{x}) - y\| \leqq \delta. \tag{89}$$

(C) The function $\varepsilon(x)$ in the algorithm has the form

$$\varepsilon(x) = \min\{\varepsilon, \|x - P[x - \nabla f(x)]\|\}, \tag{90}$$

where $\varepsilon > 0$ is a given scalar. Furthermore the set I_k in the algorithm is chosen to be (cf. (62))

$$I_k = \{i \in I | \langle a_i, x_k \rangle \geqq b_i - \varepsilon(x_k) \| a_i \| \}. \tag{91}$$

The Lipschitz condition (88) is satisfied in particular if f is twice continuously differentiable. Condition (89) is a weakened version of an often employed regularity and strict complementarity assumption which requires that the set of vectors $\{a_i | i \in A_{\bar{x}}\}$ is linearly independent and all Lagrange multipliers corresponding to the active constraints are strictly positive. The form (90) for $\varepsilon(x)$ is required for technical purposes in our subsequent proof. The reader can verify that there are other forms of $\varepsilon(x)$ that are equally suitable. Finally the choice (91) for the set I_k is natural and is ordinarily the one that is best for algorithmic purposes.

The following proposition allows us to transfer rate of convergence results from unconstrained minimization to algorithm (62)–(73).

PROPOSITION 3. *Let $\bar{x}$ be a limit point of the sequence $\{x_k\}$ generated by iteration* (62)–(73), *and let Assumptions* (A)–(C) *hold. Then*

$$\lim_{k \to \infty} x_k = \bar{x} \tag{92}$$

and there exists $\bar{k}$ such that for all $k \geqq \bar{k}$ we have

$$x_k \in \bar{x} + N_{\bar{x}}, \tag{93}$$

$$\Gamma_k = \operatorname{span} \{C_k \cap \{z | \langle z, d_k^+ \rangle = 0\}\} = N_{\bar{x}}, \tag{94}$$

$$d_k = \arg \min \{\| \nabla f(x_k) + z \| | z \in N_{\bar{x}}\}, \tag{95}$$

$$x_{k+1} = x_k + \alpha_k D_k d_k, \tag{96}$$

where $\alpha_k = \beta^{m_k}$ and m_k is the first nonnegative integer m for which

$$f(x_k) - f[x_k(\beta^m)] \geqq \sigma \beta^m \langle d_k, D_k d_k \rangle. \tag{97}$$

The proof of Proposition 3 is given in Appendix B. From (96) and (97) we see that eventually the method reduces to an unconstrained minimization method on the manifold $\bar{x} + N_{\bar{x}}$. The proposition shows that if the matrix D_k is chosen so that for all k sufficiently large it is equal to the inverse Hessian of f restricted on the manifold $\bar{x} + N_{\bar{x}}$, then the method essentially reduces to the unconstrained Newton method and attains a superlinear rate of convergence.

5. Algorithmic variations. Many variations on iteration (62)–(73) are possible. One of them, changing the metric on the Hilbert space H from iteration to iteration, was discussed at the end of § 3. In this section we discuss other variations. These will include the use, in various cases, of a pseudometric on H instead of a metric, variations on the step size rules and finally variations on the various projections in (62)–(73). We will state the variations without a convergence proof. In each case, the reworking of the proofs of §§ 2–3 to show that the variation is valid, poses no difficulty.

Singular transformation of variables through a pseudometric. Here we address the case where X is not a solid body in H, i.e., for some linear manifold M we have $X \subset M \neq H$. In this case we observe that (42) is the only place where a metric as opposed to a pseudometric is needed. Noticing that if $X \subset M$, then all quantities in (42) belong to M, one can conclude that all that is necessary is to have a metric on M. This leads us to consider the use of pseudometric on H provided it induces a metric on M. Furthermore, we can change the pseudometric on H from iteration to

iteration, as we can change the metric, provided that the metrics induced on M are equivalent in the sense described in § 3. In some cases the introduction of a pseudometric serves to facilitate the projection further (see [17, Chap. 4]).

Stepsize rules. The Armijo-like rule (73) can be viewed as a combination of the Armijo rule used in unconstrained minimization [9], and an Armijo-like rule for constrained optimization proposed by Bertsekas in [7, cf. eq. (12)]. Corresponding to an alternate suggestion made there [7, cf. eq. (22)], we can replace (73) by

$$f(x_k)-f(x_k(\beta^m)) \geqq \sigma\{\beta^m\langle d_k, D_k d_k\rangle + \langle \nabla f(x_k), (x_k+\beta^m \tilde{d}_k)-x_k(\beta^m)\rangle\}. \tag{98}$$

Also, a variation of the Goldstein stepsize rule [9] can be employed, in which $\sigma < 0.5$ and α is chosen such that

$$\begin{aligned}(1-\sigma)\{\alpha\langle d_k, D_k d_k\rangle + \langle \nabla f(x_k), (x_k+\alpha\tilde{d}_k)-x_k(\alpha)\rangle\}\\ \geqq f(x_k)-f(x_k(\alpha))\\ \geqq \sigma\{\alpha\langle d_k, D_k d_k\rangle + \langle \nabla f(x_k), (x_k+\alpha\tilde{d}_k)-x_k(\alpha)\rangle\}.\end{aligned} \tag{99}$$

The rule (99) is the counterpart of (98). The reader can easily construct the counterpart to (73).

Variations on the projections. There is one central observation in the paper, namely, the projections of $D_k d_k$ and d_k^+ on any closed convex set for which d_k is a direction of recession, result in descent directions. By employing different sets with this property, variations on the algorithm result since different directions may be obtained and different arcs may be searched.

The first variation is to replace C_k in (68) by $(\Omega_k - x_k)$, i.e.

$$\tilde{d}_k = \arg\min\{\|z - D_k d_k\| \mid z \in \Omega_k - x_k, \langle z, d_k^+\rangle = 0\} \tag{100}$$

where

$$\Omega_k = \{z \mid \langle a_i, z\rangle \leqq b_i, \forall i \in I_k\}.$$

Evidently

$$\Omega_k - x_k \supset C_k$$

and as a result d_k is a direction of recession of $\Omega_k - x_k$, which implies that $\tilde{d}_k$ defined by (100) is a descent direction.

Interestingly, this variation gives rise to a variation in the stepsize search. Since the set $\{z \mid z \in C_k, \langle z, d_k^+\rangle = 0\}$ is a cone, the vector $\tilde{d}_k$ of (68) satisfies

$$\alpha\tilde{d}_k = \arg\min\{\|\alpha D_k d_k - z\| \mid z \in C_k, \langle z, d_k^+\rangle = 0\}.$$

Thus, (70) can be interpreted as

$$x_k(\alpha) = P[x_k + \alpha d_k^+ + q_k(\alpha)]$$

where

$$q_k(\alpha) = \arg\min\{\|\alpha D_k d_k - z\| \mid z \in C_k, \langle z, d_k^+\rangle = 0\}.$$

When C_k is replaced by $\Omega_k - x_k$, a new algorithm results by searching along the arc

$$x_k(\alpha) = P[x_k + \alpha d_k^+ + \tilde{q}_k(\alpha)]$$

where

$$\tilde{q}_k(\alpha) = \arg\min\{\|\alpha D_k d_k - z\| \mid z \in \Omega_k - x_k, \langle z, d_k^+\rangle = 0\}.$$

Indeed, the particular algorithm suggested in [6] can be considered to be an implementation of the last variation for an orthant constraint.

6. Multicommodity network flow problems. In this last section we apply algorithm (62)–(73) to a classical nonlinear multicommodity network flow problem and present some computational results. In vew of the typically very large number of variables and constraints of this problem, active set methods of the type presented in [4] are in our view entirely unsuitable.

We consider a network consisting of N nodes, $1, 2, \cdots, N$, and a set of directed links denoted by $\mathscr{L}$. We assume that the network is connected in the sense that for any two nodes m, n, there is a directed path from m to n. We are given a set W of ordered node pairs referred to as origin-destination (or OD) pairs. For each OD pair $w \in W$, we are given a set of directed paths P_w that begin at the origin node and terminate at the destination node. For each $w \in W$ we are also given a positive scalar r_w referred to as the input of OD pair w. This input must be optimally divided among the paths in P_w so as to minimize a certain objective function.

For every path $p \in P_w$ corresponding to an OD pair $w \in W$ we denote by x^p the flow travelling on p. These flows must satisfy

$$\sum_{p \in P_w} x^p = r_w \quad \forall w \in W, \tag{101}$$

$$x^p \geqq 0 \quad \forall p \in P_w,\ w \in W. \tag{102}$$

Equations (101), (102) define the constraint set of the optimization problem—a Cartesian product of simplices.

In Example 2 we discussed the application of our method to the case of a simplex constraint. It is not difficult to see that if we take a "diagonal" metric on the space, the multicommodity flow problem decomposes in the sense explained below.

Let x denote the vector of variables x^p, $p \in P_w$, $w \in W$, and let x^w denote the vector of variables x^p, $p \in P_w$. Let $C_x(x^w)$ and $\Gamma_x(x^w)$ denote the cone and subspace, respectively, in $R^{|w|}$, generated at x, when all variables aside from those in x^w are considered fixed and $\varepsilon = \varepsilon(x)$. Then

$$C_x = \prod_{w \in W} C_x(x^w), \quad \nabla f(x) = (\cdots, \nabla_{xw} f(x), \cdots), \quad \Gamma_x = \prod_{w \in W} \Gamma_x(x^w).$$

Thus all projections decompose and therefore in many respects the multicommodity flow problem is not different from the problem with a single simplex constraint. The only points where the "interaction" among the simplices appears is in computing ε_k, and in computing $D_k d_k$.

To every set of path flows $\{x^p | p \in P_w,\ w \in W\}$ satisfying (101), (102) there corresponds a flow f^a for every link $a \in \mathscr{L}$. It is defined by the relation

$$f^a = \sum_{w \in W} \sum_{p \in P_w} 1_p(a) x^p \quad \forall a \in \mathscr{L} \tag{103}$$

where $1_p(a) = 1$ if the path p contains the link a and $1_p(a) = 0$ otherwise. If we denote by f the vector of link flows, we can write relation (103) as

$$f = Ex \tag{104}$$

where E is the arc-chain matrix of the network.

For each link $a \in \mathscr{L}$ we are given a convex, twice continuously differentiable scalar function $D_a(f^a)$ with strictly positive second derivative for all $f^a \geqq 0$. The objective

function is given by

$$D(f)=\sum_{a\in\mathscr{L}} D_a(f^a). \tag{105}$$

By using (104), we can write the problem in terms of the path flow variables x^p as

$$\text{minimize } J(x)=D(Ex)$$

subject to:

$$\sum_{p\in P_w} x^p = r_w \quad \forall w\in W,$$

$$x^p \geqq 0 \quad \forall p\in P_w,\ w\in W.$$

In communication network applications the function D may express, for example, average delay per message [10], [11] or a flow control objective [12], while in transportation networks it may arise via a user or system optimization principle formulation [13], [14], [15]. We concentrate on the separable form of D given by (105), although what follows admits an extension to the nonseparable case.

A Newton-like method will be obtained if we chose $D_k d_k$ so that $x_k + D_k d_k$ is the minimum of the quadratic approximation to f on $x_k + \Gamma_k$. For this we must find $A\bar{v}$ where $\bar{v}$ solves

$$\underset{v}{\text{minimize}}\ \langle \nabla J(x_k), Av\rangle + \tfrac{1}{2}\langle Av, \nabla^2 J(x_k)Av\rangle \tag{106}$$

and where A is a matrix such that its columns are linearly independent and span Γ_k.

The particular structure of the objective function (105) gives rise to a Hessian matrix which makes the solution of (106) relatively easy to obtain. Indeed, using (105) we can rewrite (106) as

$$\underset{v}{\text{minimize}}\ \langle E'\nabla D(f_k), Av\rangle + \tfrac{1}{2}\langle Av, E'\nabla^2 D(f_k)EAv\rangle, \tag{107}$$

where $f_k = Ex_k$ and prime denotes transposition. A key fact (described in detail in Bertsekas and Gafni [16]) is that problem (107), in light of $\nabla^2 D(f_k)$ being diagonal, can be solved by the Conjugate Gradient (C-G) method using graph type operations without explicitly storing the matrix

$$A'E'\nabla^2 D(f_k)EA.$$

Note that a solution to (107) exists since $E'\nabla D(f_k)$ is in the range of the nonnegative definite matrix $E'\nabla^2 D(f_k)E$.

Computational results. A version of the algorithm was run on an example of the multicommodity flow problem. The network is shown in Fig. 5. Each OD pair was restricted to use only two prespecified paths. This reduced the programming load significantly, yet captured the essence of the algorithm. It is conjectured that the results we obtained are representative of the behavior of the algorithm when applied to more complex multicommodity flow problems.

The algorithm was operated in three modes distinguished by the other rules according to which the C-G method was stopped. In the first mode (denoted by Newton) the C-G iteration was run to the exact solution of problem (107). In the second mode, (denoted by approximate Newton) the C-G iteration was run until its residual was reduced by a factor of $\frac{1}{8}$ over the starting residual (this factor was chosen on a heuristic basis). Finally, in the third mode the C-G method was allowed to perform only one

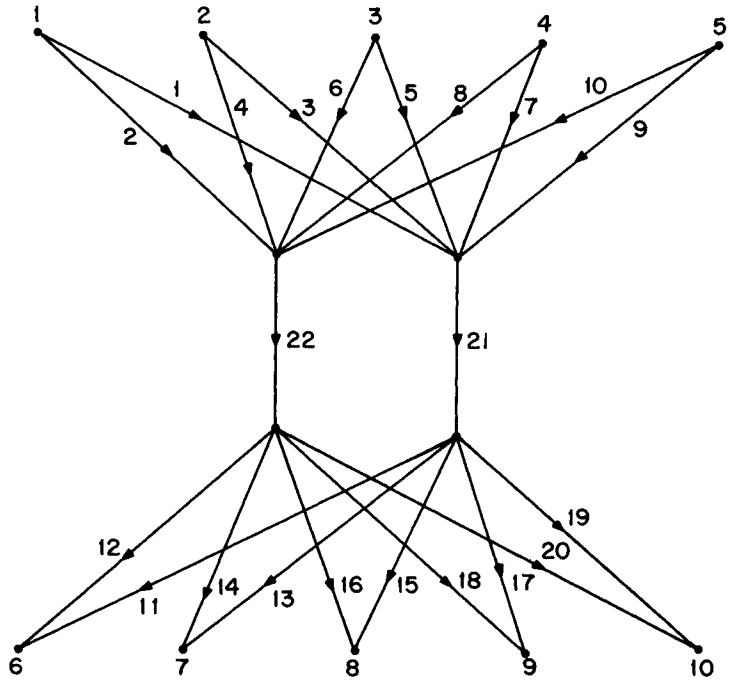

FIG. 5. *The network; initially all flows traverse link* 21.

step (denoted by 1-step—this results in a diagonally scaled version of the original Goldstein–Levitin–Poljak method). In all these modes, in addition to their particular stopping rule, the C-G method was stopped whenever for any OD pair w the flow on the path with the smallest partial derivative of cost became negative. Each time this happened, the last point in the sequence of points generated by the C-G method subiteration was connected by a line to the point preceding it. The point on the line at which the particular path flow became zero was taken as the result of the C-G iteration. We used different values ε_k for different OD pairs, according to a variation of (60) (with $\varepsilon = 0.2$).

We used two types of objective functions. The first is

$$D_a(f^a) = \frac{f^a}{C_a - f^a} \quad \forall a \in \mathcal{L}$$

where C_a is a given positive scalar expressing the "capacity" of link a. This function is typically used to express queueing delay in communication networks. The second type was taken to be quadratic. We used two sets of inputs, one to simulate heavy loading and one to simulate light loading. For each combination of cost function and input we present the results corresponding to the three versions in Table 1.

Our main observation from the results of Table 3 as well as additional experimentation with multicommodity flow problems is that in the early iterations the 1-step method makes almost as much progress as the other two more sophisticated methods but tends to slow down considerably after reaching the vicinity of the optimum. Also the approximate Newton method does almost as well as Newton's method in terms of number of iterations. However the computational overhead per iteration for Newton's method is considerably larger. This is reflected in the results of Table 3 which show

TABLE 1
Capacities.

	16.5	20	7.5	10	10
	15	5	9	7.5	7.5
$M =$	3	15	6	8	3
	10	6	10	10	14
	50	35	x	x	x

$C_a = m_{ij},\ i = \frac{a}{5} + 1.\ j = a - 5(i-1)$

TABLE 2
Low input. High input = low input $\times 1.75$.

origin \ destination	6	7	8	9	10
1	0.5	1	1.5	2	2.5
2	1	1	1	1	1
3	0.5	0.5	1.5	1.5	3.5
4	0.25	0.25	2	0.25	0.25
5	0.75	0.75	0.75	0	0

TABLE 3

	Initial value	Final value	No. of iterations	Total no. of C-G subiterations
Low load				
Nonquadratic objective	$1.600616 \cdot 10^6$			
Newton		8.743550	16	29
Approximate Newton		8.758665	16	16
1-step		8.758665	16	16
Quadratic objective	$1.866326 \cdot 10^1$			
Newton		7.255231	5	17
Approximate Newton		7.255231	7	13
1-step		7.255231	12	12
High load				
Nonquadratic objective	$9.759996 \cdot 10^6$			
Newton		$3.737092 \cdot 10^1$	14	117
Approximate Newton		$3.737745 \cdot 10^1$	15	30
1-step		$3.747400 \cdot 10^1$	15	15
Quadratic objective	$9.759996 \cdot 10^6$			
Newton		$1.521299 \cdot 10^1$	5	24
Approximate Newton		$1.521299 \cdot 10^1$	13	27
1-step		$1.521301 \cdot 10^1$	16	16

in three cases out of four a larger number of conjugate gradient subiterations for Newton's method. Throughout our computational experiments (see also [17]) the approximate Newton method based on conjugate gradient subiterations has performed very well and, together with its variations, is in our view the most powerful class of methods available at present for nonlinear multicommodity network flow problems.

Appendix A.

Proof of Lemma 1. a) Fix $x \in X$, $z \in H$ and $\gamma > 1$. Denote

$$a = x + z, \qquad b = x + \gamma z. \tag{A.1}$$

Let $\bar{a}$ and $\bar{b}$ be the projections on X of a and b respectively. It will suffice to show that

$$\|\bar{b} - x\| \leqq \gamma \|\bar{a} - x\|. \tag{A.2}$$

If $\bar{a} = x$ then clearly $\bar{b} = x$, so (A.2) holds. Also if $a \in X$ then $\bar{a} = a = x + z$ so (A.2) becomes $\|\bar{b} - x\| \leqq \gamma \|z\| = \|b - x\|$ which again holds by the contraction property of the projection. Finally if $\bar{a} = \bar{b}$ then (A.2) also holds. Therefore it will suffice to show (A.2) in the case where $\bar{a} \neq \bar{b}$, $\bar{a} \neq x$, $\bar{b} \neq x$, $\bar{a} \notin X$, $\bar{b} \notin X$ shown in Fig. (A.1).

Let H_a and H_b be the two hyperplanes that are orthogonal to $(\bar{b} - \bar{a})$ and pass through $\bar{a}$ and $\bar{b}$ respectively. Since $\langle \bar{b} - \bar{a}, b - \bar{b} \rangle \geqq 0$ and $\langle \bar{b} - \bar{a}, a - \bar{a} \rangle \leqq 0$, we have that neither a nor b lie strictly between the two hyperplanes H_a and H_b. Furthermore x lies on the same side of H_a as a, and $x \notin H_a$. Denote the intersections of the line $\{x + \alpha(b - x) | \alpha \in R\}$ with H_a and H_b by s_a and s_b respectively. Denote the intersection of the line $\{x + \alpha(\bar{a} - x) | \alpha \in R\}$ with H_b by w. We have

$$\begin{aligned} \gamma = \frac{\|b - x\|}{\|a - x\|} &\geqq \frac{\|s_b - x\|}{\|s_a - x\|} = \frac{\|w - x\|}{\|\bar{a} - x\|} = \frac{\|w - \bar{a}\| + \|\bar{a} - x\|}{\|\bar{a} - x\|} \\ &\geqq \frac{\|\bar{b} - \bar{a}\| + \|\bar{a} - x\|}{\|\bar{a} - x\|} \geqq \frac{\|\bar{b} - x\|}{\|\bar{a} - x\|} \end{aligned} \tag{A.3}$$

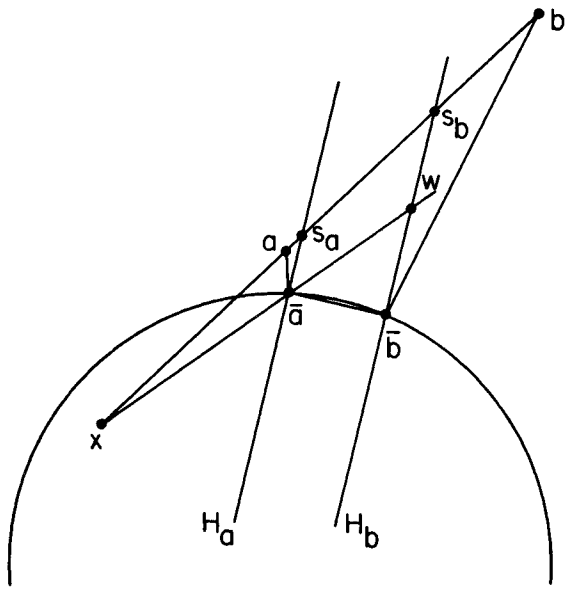

FIG. A.1

where the third equality is by similarity of triangles, the next to last inequality follows from the orthogonality relation $\langle w-\bar{b}, \bar{b}-\bar{a}\rangle = 0$, and the last inequality is obtained from the triangle inequality. From (A.3) we obtain (A.2) which was to be proved.

b) Since y is a direction of recession of Ω, we have

$$P_\Omega(x+z)+y \in \Omega. \tag{A.4}$$

Thus by definition of projection on a closed convex set

$$\langle (x+z)-P_\Omega(x+z), (P_\Omega(x+z)+y)-P_\Omega(x+z)\rangle \leqq 0 \tag{A.5}$$

or equivalently

$$\langle (x+z)-P_\Omega(x+z), y\rangle \leqq 0,$$

and (40) follows. Q.E.D.

Appendix B. We develop the main arguments for the proof of Proposition 3 through a sequence of lemmas. In what follows we use the word "eventually" to mean "there exists $\bar{k}$ such that for all $k \geqq \bar{k}$," where $\bar{k}$ may be different for each case.

LEMMA B.1. *Under the conditions of Proposition* 3, $\lim_{k\to\infty} x_k = \bar{x}$ *and eventually*

$$I_k = A_{\bar{x}}. \tag{B.1}$$

Proof. By relation (73), since $\bar{x}$ is a limit point and the algorithm decreases the value of the objective function at each iteration, we have

$$\lim_{k\to\infty} \|x_{k+1}-x_k\| = 0,$$

which implies, again by the descent property and the fact that $\bar{x}$ is a strict local minimum

$$\lim_{k\to\infty} x_k = \bar{x}. \tag{B.2}$$

Therefore from (90)

$$\lim_{k\to\infty} \varepsilon(x_k) = \varepsilon(\bar{x}) = 0. \tag{B.3}$$

Since the set I is finite, it follows from (87), (91) and (B.3) that eventually

$$I_k \subset A_{\bar{x}}. \tag{B.4}$$

To show the reverse inclusion we must show that eventually

$$\langle a_i, x_k\rangle \geqq b_i - \varepsilon(x_k)\|a_i\| \quad \forall i \in A_{\bar{x}}. \tag{B.5}$$

By the Cauchy–Schwarz inequality, (B.3) and (90) we have eventually

$$\varepsilon(x_k)\|a_i\| = \|x_k - P[x_k - \nabla f(x_k)]\| \cdot \|a_i\| \geqq \langle P[x_k - \nabla f(x_k)] - x_k, a_i\rangle.$$

Therefore in order to show (B.5) it suffices to show that eventually

$$\langle a_i, P[x_k - \nabla f(x_k)]\rangle = b_i \quad \forall i \in A_{\bar{x}}$$

or equivalently

$$P[x_k - \nabla f(x_k)] \varepsilon \bar{x} + N_{\bar{x}}.$$

Since $x_k \to \bar{x}$ this follows from Assumption (B). Q.E.D.

LEMMA B.2. *Under the conditions of Proposition* 3 *for each* $\bar{\alpha} \in (0, 1]$, *eventually we have*

$$x_k(\alpha) \in \bar{x} + N_{\bar{x}} \quad \forall \alpha \in [\bar{\alpha}, 1]. \tag{B.6}$$

Proof. From Lemma B.1 we have $x_k \to \bar{x}$ and eventually $C_k = \bar{C}$ where

$$\bar{C} = \{z \mid \langle z, a_i \rangle \leqq 0, \forall i \in A_{\bar{x}}\}. \tag{B.7}$$

Since the projection of $-\nabla f(\bar{x})$ on $\bar{C}$ is the zero vector and d_k is eventually the projection of $-\nabla f(x_k)$ on $\bar{C}$ it follows that

$$\lim_{k\to\infty} d_k = 0. \tag{B.8}$$

Since $\tilde{d}_k$ is the projection of $D_k d_k$ on a subset of C_k, and $\{\|D_k\|\}$ is bounded above (cf. (67), (68)), it follows that

$$\lim_{k\to\infty} \tilde{d}_k = 0. \tag{B.9}$$

Since $-\nabla f(x_k) = d_k^+ + d_k$ and $g_k = -(d_k^+ + \tilde{d}_k)$

$$\lim_{k\to\infty} g_k = \nabla f(\bar{x}). \tag{B.10}$$

A simple argument shows that Assumption (B) implies that for all $\alpha \in [0, 1]$

$$P(y) \in \bar{x} + N_{\bar{x}} \quad \forall y \text{ such that } \|\bar{x} - \alpha \nabla f(\bar{x}) - y\| \leqq \alpha\delta. \tag{B.11}$$

For any $\bar{\alpha} \in (0, 1]$, equation (B.10) shows that we have eventually

$$\|\bar{x} - \alpha \nabla f(\bar{x}) - (x_k - \alpha g_k)\| \leqq \alpha\delta \quad \forall \alpha \in [\bar{\alpha}, 1].$$

Therefore from (B.11) we have eventually

$$x_k(\alpha) = P(x_k - \alpha g_k) \in \bar{x} + N_{\bar{x}} \quad \forall \alpha \in [\bar{\alpha}, 1]. \quad \text{Q.E.D.}$$

LEMMA B.3. *Under the conditions of Proposition* 3

$$\liminf_{k\to\infty} \alpha_k > 0.$$

Proof. From Lemma B.1 we have eventually $I_k = A_{\bar{x}}$ and $x_k \to \bar{x}$, while from (B.8) we have $\|g_k\| \to \|\nabla f(\bar{x})\|$. Therefore from Proposition 1b) [cf. (42)] it follows that there exists $\hat{\alpha} > 0$ such that eventually

$$\langle \nabla f(x_k), x_k - x_k(\alpha) \rangle \geqq \alpha \langle d_k, D_k d_k \rangle + \frac{1}{\alpha} \|x_k(\alpha) - (x_k + \alpha \tilde{d}_k)\|^2 \quad \forall \alpha \in (0, \hat{\alpha}].$$

Using this relation, we get that eventually

$$\begin{aligned}
f(x_k) - f[x_k(\alpha)] &\geqq \langle \nabla f(x_k), x_k - x_k(\alpha) \rangle - \frac{L}{2} \|x_k(\alpha) - x_k\|^2 \\
&\geqq \alpha \langle d_k, D_k d_k \rangle + \frac{1}{\alpha} \|x_k(\alpha) - (x_k + \alpha \tilde{d}_k)\|^2 - \frac{L}{2} \|x_k(\alpha) - x_k\|^2 \\
&\geqq \alpha \langle d_k, D_k d_k \rangle + \frac{1}{\alpha} \|x_k(\alpha) - (x_k + \alpha \tilde{d}_k)\|^2 \\
&\quad - L\|\alpha \tilde{d}_k\|^2 - L\|x_k(\alpha) - (x_k + \alpha \tilde{d}_k)\|^2 \\
&\geqq \alpha(1 - \alpha L \lambda_2) \langle d_k, D_k d_k \rangle + \left(\frac{1}{\alpha} - L\right) \|x_k(\alpha) - (x_k + \alpha \tilde{d}_k)\|^2
\end{aligned}$$

where the third inequality follows from

$$\|x + y\|^2 \leqq 2\|x\|^2 + 2\|y\|^2,$$

the last inequality follows from (67) and L is a Lipschitz constant that corresponds to any nonempty bounded neighborhood of $\bar{x}$. Taking any $\bar{\alpha} > 0$ satisfying

$$\bar{\alpha} \leqq \hat{\alpha}, \quad 1 - \bar{\alpha} L \lambda_2 > \sigma, \quad \bar{\alpha}\left(\frac{1}{\bar{\alpha}} - L\right) > \sigma$$

we obtain, using (73) that

$$\liminf_{k \to \infty} \alpha_k > \bar{\alpha}$$

and the Lemma is proved. Q.E.D.

Proof of Proposition 3. The fact $\lim_{k\to\infty} x_k = \bar{x}$ is part of Lemma B.1, while (93) follows from Lemmas B.2 and B.3.

In order to show (94) we note that from Lemma B.1 and (B.8) we have eventually

$$C_k = \bar{C}, \qquad C_k^+ = \bar{C}^+ \tag{B.12}$$

and

$$\lim_{k \to \infty} d_k^+ = -\nabla f(\bar{x}). \tag{B.13}$$

Equation (B.13) implies that eventually assumption (B) holds with d_k^+ replacing $-\nabla f(\bar{x})$ and $\delta/2$ replacing δ. Therefore for all $i \in A_{\bar{x}}$ and $\rho_i > 0$ such that $\|\rho_i a_i\| < \delta/2$ we have

$$P(\bar{x} + d_k^+ \pm \rho_i a_i) \in \bar{x} + N_{\bar{x}}, \tag{B.14}$$

$$P(\bar{x} + d_k^+) \in \bar{x} + N_{\bar{x}}. \tag{B.15}$$

For any $z_1, z_2 \in H$ we have from a general property of projection on X

$$\langle z_1 - P(z_1), P(z_2) - P(z_1) \rangle \leqq 0,$$

$$\langle z_2 - P(z_2), P(z_1) - P(z_2) \rangle \leqq 0.$$

By adding these two inequalities, we obtain

$$\|P(z_1) - P(z_2)\|^2 \leqq \langle z_1 - z_2, P(z_1) - P(z_2) \rangle \quad \forall z_1, z_2 \in H. \tag{B.16}$$

By applying (B.16), we obtain

$$\|P(\bar{x} + d_k^+ \pm \rho_i a_i) - P(\bar{x} + d_k^+)\|^2 \leqq \langle \pm \rho_i a_i, P(\bar{x} + d_k^+ \pm \rho_i a_i) - P(\bar{x} + d_k^+) \rangle. \tag{B.17}$$

Since $\langle a_i, z \rangle = 0$ for all $z \in N_{\bar{x}}$, $i \in A_{\bar{x}}$ it follows from (B.14), (B.15) that the right side of (B.17) is zero and therefore eventually

$$P(\bar{x} + d_k^+ \pm \rho_i a_i) = P(\bar{x} + d_k^+) \quad \forall i \in A_{\bar{x}}.$$

Since from (B.12) we have eventually $d_k^+ \in \bar{C}^+$, it follows that $P(\bar{x} + d_k^+) = \bar{x}$ and therefore also

$$P(\bar{x} + d_k^+ \pm \rho_i a_i) = \bar{x} \quad \forall i \in A_{\bar{x}}.$$

Hence eventually

$$d_k^+ \pm \rho_i a_i \in \bar{C}^+ \quad \forall i \in A_{\bar{x}}$$

which implies that

$$\langle d_k^+ \pm \rho_i a_i, y \rangle \leqq 0 \quad \forall y \in \bar{C}, i \in A_{\bar{x}}. \tag{B.18}$$

Let

$$y \in \{z | z \in C_k, \langle z, d_k^+ \rangle = 0\}.$$

From (B.12) and (B.18) we have eventually

$$\langle a_i, y\rangle = 0 \quad \forall i \in A_{\bar{x}},$$

or equivalently $y \in N_{\bar{x}}$. Hence eventually

$$N_{\bar{x}} \supset \{z \mid z \in C_k, \langle z, d_k^+\rangle = 0\}$$

and it follows that

$$\text{span } N_{\bar{x}} = N_{\bar{x}} \supset \text{span}\{C_k \cap \{z \mid \langle z, d_k^+\rangle = 0\}\} = \Gamma_k.$$

To show the reverse inclusion, note that if $y \in N_{\bar{x}}$ then by Assumption (B) and (B.12) we have eventually

$$\langle y, d_k^+\rangle = 0.$$

Since $N_{\bar{x}} \subset \bar{C}$ and eventually $C_k = \bar{C}$, it follows that eventually $y \in C_k \cap \{z \mid \langle z, d_k^+\rangle = 0\}$ and a fortiori $y \in \text{span}\{C_k \cap \{z \mid \langle z, d_k^+\rangle = 0\}\} = \Gamma_k$. Therefore eventually

$$N_{\bar{x}} \subset \Gamma_k$$

and the proof of (94) is complete.

Since d_k is the projection of $-\nabla f(x_k)$ on $C_k \cap \{z \mid \langle z, d_k^+\rangle = 0\}$, equation (95) follows easily from (94).

Also from (93) and (94) we have eventually $x_k \in \bar{x} + N_{\bar{x}}$, $d_k \in N_{\bar{x}}$, $\tilde{d}_k \in N_{\bar{x}}$ and d_k^+ is orthogonal to $N_{\bar{x}}$, while by Lemma B.2 the vector x_{k+1} is the projection of $x_k + \alpha_k(\tilde{d}_k + d_k^+)$ on $\bar{x} + N_{\bar{x}}$. Therefore (96) and (97) follow. Q.E.D.

REFERENCES

[1] A. A. GOLDSTEIN, *Convex programming in Hilbert space*, Bull. Amer. Math. Soc., 70 (1964), pp. 709–710.

[2] E. S. LEVITIN AND B. T. POLJAK, *Constrained minimization problems*, USSR Comp. Math. Phys., 6 (1966), pp. 1–50.

[3] U. M. GARCIA-PALOMARES, *Superlinearly convergent algorithms for linearly constrained optimization*, in Nonlinear Programming 2, O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds., Academic Press, New York, 1975, pp. 101–121.

[4] P. E. GILL, W. MURRAY AND M. H. WRIGHT, *Practical Optimization*, Academic Press, New York, 1981.

[5] J. C. DUNN, *Global and asymptotic rate of convergence estimates for a class of projected gradient processes*, this Journal, 18 (1981), pp. 659–674.

[6] D. P. BERTSEKAS, *Projected Newton methods for optimization problems with simple constraints*, this Journal, 20 (1982), pp. 221–246.

[7] ———, *On the Goldstein–Levitin–Poljak gradient projection method*, Proc. 1974 IEEE Conf. on Decision and Control, Pheonix, Az., pp. 47–52; IEEE Trans. Automat. Control, AC-20 (1976), pp. 174–184.

[8] J.-J. MOREAU, *Convexity and duality*, in Functional Analysis and Optimization, E. R. Caianiello, ed., Academic Press, New York, 1966.

[9] E. POLAK, *Computational Methods in Optimization: A Unified Approach*, Academic Press, New York, 1971.

[10] R. G. GALLAGER, *A minium delay routing algorithm using distributed computation*, IEEE Trans. Comm., COM-25 (1977), pp. 73–85.

[11] D. P. BERTSEKAS, E. M. GAFNI AND R. G. GALLAGER, *Second derivative algorithms for minimum delay distributed routing in networks*, Report LIDS-R-1082, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, May 1979; IEEE Trans. Comm., COM-32 (1984), to appear.

[12] R. G. GALLAGER AND S. J. GOLESTAANI, *Flow control and routing algorithms for data networks*, Proc. Fifth International Conference on Computer Communication (ICCC-80), Atlanta, GA, Nov. 1980, pp. 779–784.

[13] D. P. BERTSEKAS AND E. M. GAFNI, *Projection methods for variational inequalities with application to the traffic assignment problem*, in Math. Prog. Study, D. C. Sorensen and R. J.-B. Wets, eds., North-Holland, Amsterdam, 1982, pp. 139–159.

[14] S. DAFERMOS, *Traffic equilibrium and variational inequalities*, Transportation Sci., 14 (1980), pp. 42–54.
[15] H. Z. AASHTIANI AND T. L. MAGNANTI, *Equilibria on a congested transportation network*, SIAM J. Alg. Disc. Meth., 2 (1981), pp. 213–226.
[16] D. P. BERTSEKAS AND E. M. GAFNI, *Projected Newton methods and optimization of multicommodity flows*, LIDS Rep. P–1140, Massachusetts Institute of Technology, Cambridge, 1981, IEEE Trans. Automat. Control, AC-28 (1983), pp. 1090–1096.
[17] E. M. GAFNI, *The integration of routing and flow-control for voice and data in a computer communication network*, PhD. dissertation, Dept. Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Aug. 1982.
[18] D. P. BERTSEKAS, G. S. LAUER, N. R. SANDELL, JR AND T. A. POSBERGH, *Optimal short-time scheduling of large-scale power systems*, IEEE Trans. Automat Control, AC-28 (1983), pp. 1–11.
[19] D. P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.