Structural Optimization 10, 209–216 © Springer-Verlag 1995

Structural optimization using the Newton modified barrier method

N.S. Khot

Structures Division, Flight Dynamics Directorate, Wright-Patterson Air Force Base, OH 45433-6553, USA

R.A. Polyak

Department of Operations Research & Engineering, George Mason University, Fairfax, VA 22030, USA

R. Schneur

PTCG, Burlington, MASS 01803, USA

L. Berke

NASA Lewis Research Center, Cleveland, OH 44135, USA

Abstract The Newton Modified Barrier Method (NMBM) is applied to structural optimization problems with large a number of design variables and constraints. This nonlinear mathematical programming algorithm was based on the Modified Barrier Function (MBF) theory and the Newton method for unconstrained optimization. The distinctive feature of the NMBM method is the rate of convergence that is due to the fact that the design remains in the Newton area after each Lagrange multiplier update. This convergence characteristic is illustrated by application to structural problems with a varying number of design variables and constraints. The results are compared with those obtained by optimality criteria (OC) methods and by the ASTROS program.

1 Introduction

The problems facing designers today tend to be complex and often novel with limited past experience to guide the designer to satisfy conflicting requirements. The computer automated analysis and simulation capabilities have reached a very high level of sophistication and power to provide detailed information about the response behaviour of the design. These capabilities, however, often provide little information as to how to change a complex design to improve its behaviour or its cost measure, or both. New automated optimization approaches are needed to provide guidance for improving the preliminary and for fine tuning the final designs.

The use of finite element methods in structural analysis leads to a large number of size variables. It is often convenient to treat each element size as an independent variable instead of developing a variable linking procedure to reduce the number of variables, or to promote smoothness of member size variations. Such reductions of problem size are, however, always advantageous from a computational point of view.

The derivations of OC methods (see Berke and Khot 1987; Khot 1981) are associated with structural energy theorems that are involved in stiffness related optimization problems. These energy theorems assure separability and the associated simplifications. The number of needed iterations depends not on the number of design variables but on low sensitivity of internal forces to changes in member sizes. These OC methods have been shown to be extremely efficient but only for special cases of structural constraints. With the increasing importance of multidisciplinary design requirements specialized algorithms are becoming less attractive for the emerging complexities.

With the increasing speed of computers, a major part of the importance of computational efficiency needed for optimization iterations is diminishing. In such a computational environment, developments turn towards robust optimization methods based on sound mathematical foundations.

The NMBM method applied here initially only to structural problems promises to satisfy these requirements. Its performance is compared here with that of OC and the ASTROS programs for problems that satisfy the conditions of the applicability of OC methods. Separability and the linear objective function are not necessary conditions for the proposed method, but are believed to improve convergence behaviour. Work is in progress to compare this approach to other algorithms in the multidisciplinary problem setting using COMETBOARD, a testbed for optimization algorithms (Gendy *et al.* 1994).

The NMBM is a numerical realization of the general modified barrier method (MBM) for constrained optimization (Polyak 1992a). The barrier function (MBF) used in this method possesses the best properties of the classical Lagrangian and classical barrier functions (CBF) (see Fiacco and McCormick 1968), and at the same time eliminates their main drawbacks. The MBF retains the best properties of the augmented Lagrangian (see Powell 1969; Rockafellar 1973; Bertsekas 1982) and also develops new important qualities. The properties of MBF near the primal and dual solution make it possible to avoid illconditioning of the CBF's Hessian. This makes the numerical process more stable and improves the rate of convergence. The application of MBF-type methods for solving truss topology design problems (Ben-Tal *et al.* 1992) leads us to believe that it has good potential for also solving large-scale structural optimization problems.

The next sections give the formulation of the structural optimization problem using the NMBM approach, followed by presentations of MBF, MBM, NMBM in turn. In the last section, three truss structures using NMBM are presented. The convergence characteristics are compared with solutions obtained by ASTROS (Neill and Herendeen 1993) and by an OC method.

2 Structural optimization problem

The structural optimization problem, which we are going to consider, can be stated as follows:

$$W^* = \min W = \sum_{i=1}^{n} \rho_i \ell_i x_i \,, \tag{1}$$

subject to constraints

$$g_j(x) = c_j - \bar{c}_j = \sum_{i=1}^{n} Q_{ij} x_i^{-1} - \bar{c}_j \le 0, \quad j = 1, \ldots, m\,, \tag{2}$$

and

$$x_i \ge 0, \quad i = 1, \ldots, m\,. \tag{3}$$

The x_i are the design variables, ρ_i is the specific weight, and ℓ_i is a function of the geometry of the elements. Parameters Q_{ji} depend on the nature of the constraints and are a function of the sensitivity, and c_j and $\bar{c}_j$ are the actual and limited values of the j-th constraints.

Before we start our analysis, let us rewrite the problem (1)-(3) in a more convenient manner for further consideration.

We introduce the reciprocal variables $y_i = x_i^{-1}, i = 1, \ldots, n$ and the vector $\mathbf{y} = (y_\ell, \ldots, y_n)$. We also set $r_i = \rho_i \ell_i$, so that the objective function (1) can be rewritten as follows:

$$W = \sum_{i=1}^{n} r_i y_i^{-1},$$

and for the constraints (2) we have

$$g_j(y) = \bar{c}_j - \sum_{i=1}^{n} Q_{ji} y_i \ge 0, \quad j = 1, \ldots, m\,.$$

The objective function

$$W = \sum_{i=1}^{n} r_i y_i^{-1}$$

is a barrier type function; therefore, if the initial approximation $y^0 = (y_1^0, \ldots, y_n^0)$ is inside the feasible set, i.e. all $y_i^0 > 0$, then the minimization procedure will keep the variables positive.

Let $\mathbf{Q} = \| Q_{ji} \|, j = 1, \ldots, m, \; i = 1, \ldots, n$, be an $m \times n$ matrix with elements Q_{ji}; $\mathbf{r} = (r_\ell, \ldots, r_n)$ and $\mathbf{y}^{-1} = (y_1^{-1}, \ldots, y_n^{-1})$ the n-dimensional vectors, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ an m-dimensional vector of Lagrange multipliers, and $\mathbf{g}(y) = (\bar{c}_1 - \sum_{i=1}^{n} Q_{\ell i} y_i, \ldots, \bar{c}_m - \sum_{i=1}^{n} Q_{mi} y_i) = \bar{c} - \mathbf{Q}\mathbf{y}$ an m-dimensional vector function. Then the problem (1)-(3) can be written as follows:

$$W^* = f(y^*) = \min f(y) = \min(r, y^{-1}) = \sum_{i=1}^{n} r_i y_i^{-1}\,, \tag{4}$$

subject to the constraints

$$\mathbf{g}(\mathbf{y}) = \bar{c} - \mathbf{Q}\mathbf{y} \ge 0\,. \tag{5}$$

For structural design problems, the feasible set that is defined by (5), namely,

$$\Omega = \{\mathbf{y} : g_j(\mathbf{y}) \ge 0, \; j = 1, \ldots, m\}\,,$$

has a nonempty interior; i.e. there exists $\mathbf{y}^0 > 0$ such that

$$g_j(\mathbf{y}^0) > 0, \quad j = 1, \ldots, m\,. \tag{6}$$

Therefore, the Karush-Kuhn-Tucker (KKT) conditions hold true.

Let $\mathbf{R} = [\text{diag}\; r_i]_{i=\ell}^{n}$ be a diagonal matrix with entries $r_i, i = 1, \ldots, n$ and let $L(\mathbf{y}, \boldsymbol{\lambda})$ be the Lagrangian for the problem (4)-(5)

$$L(\mathbf{y}, \boldsymbol{\lambda}) = f(\mathbf{y}) - \sum_{j=1}^{m} \lambda_j g_j(y) = (\mathbf{r}, \mathbf{y}^{-1}) - (\boldsymbol{\lambda}, \bar{c} - \mathbf{Q}\mathbf{y})\,,$$

then, for any KKT pair (y, λ), we have,

$$\nabla_{\mathbf{y}} L(\mathbf{y}, \boldsymbol{\lambda}) = -\mathbf{R}\mathbf{y}^{-2} + \boldsymbol{\lambda}\mathbf{Q} = 0\,, \tag{7}$$

$$\bar{c} - \mathbf{Q}\mathbf{y} \ge 0, (\boldsymbol{\lambda}, \bar{c} - \mathbf{Q}\mathbf{y}) = 0, \boldsymbol{\lambda} \ge 0\,. \tag{8}$$

The modified barrier method, which we describe below, generates primal and dual sequences that converge to a KKT pair $(\mathbf{y}^*, \boldsymbol{\lambda}^*)$.

Due to the nature of the structural design problem, the feasible set Ω is bounded. Therefore, the objective function $f(\mathbf{y})$ is strongly convex on Ω. Thus the solution $\mathbf{y}^*$ of the problem (4)-(5) is unique, while the set of Lagrange multipliers $\boldsymbol{\lambda}$ that satisfies (7) and (8) is bounded due to the Slater condition (6).

In the next section, we will introduce the modified barrier function for the structural optimization problem (4)-(5).

3 Modified barrier function

Let $k > 0$ be a scaling parameter. The constraints (5) can be rewritten in the following form:

$$k^{-1} \ln[k g_j(\mathbf{y}) + 1] \ge 0, \quad j = 1, \ldots, m\,. \tag{9}$$

We shall assume that $\ln t = -\infty$ for $t \le 0$.

The modified barrier function is a classical Lagrangian for the equivalent problem (4) and (9)

$$F(y, \lambda, k) = \sum_{i=1}^{n} r_i y_i^{-1} - k^{-1} \sum_{j=1}^{m} \lambda_j \ln[k g_j(y) + 1]\,. \tag{10}$$

The MBF theory and the corresponding methods for nonlinear and linear optimization have been developed by Polyak (1992a, b). For the sake of completeness, we shall specify the main MBF properties for the structural optimization problem (4)-(5).

Let $\mathbf{Y} = [\text{diag}\; y_i]_{i=1}^{n}$ be a diagonal matrix in n-dimensional space with entries y_i and $\Lambda = [\text{diag}\; \lambda_j]_{i=1}^{m}$ be a diagonal matrix in m-dimensional space with entries λ_j.

For any KKT pair $(\mathbf{y}^*, \boldsymbol{\lambda}^*)$ that satisfies (7) and (8), and any $k > 0$, the following properties are true:

P1 $\bullet\; F(\mathbf{y}^*, \boldsymbol{\lambda}^*, k) = f_0(\mathbf{y}^*) = \sum_{i=1}^{n} r_i (y_i^*)^{-1}\,,$

P2 $\bullet\; \nabla_{\mathbf{y}} F(\mathbf{y}^*, \boldsymbol{\lambda}^*, k) = \nabla_{\mathbf{y}} L(\mathbf{y}^*, \boldsymbol{\lambda}^*) = -\mathbf{R}(\mathbf{y}^*)^{-2} + \boldsymbol{\lambda}^*\mathbf{Q} = 0$

P3 $\bullet\; \nabla^2_{\mathbf{yy}} F(\mathbf{y}^*, \boldsymbol{\lambda}^*, k) = 2\mathbf{R}(\mathbf{Y}^*)^{-3} + k\mathbf{Q}^T \Lambda^* \mathbf{Q}\,,$

where $(\mathbf{y}^*)^{-2} = [(y_1^*)^{-2}, \ldots, (y_n^*)^{-2}]$ is a vector in $\mathbf{R}^n$ and $(\mathbf{Y}^*)^{-3} = [\mathrm{diag}(y_i^*)^{-3}]_{i=1}^n$ is a diagonal matrix with entries $(y_i^*)^{-3}$.

The properties P1-P3 show that in contrast to CBF, the MBF as well as its first and second derivatives exist at the primal solution and possess important properties in the neighbourhood of the primal-dual solution.

The optimal value of the objective function coincides with the MBF value for any KKT pair $(\mathbf{y}^*, \lambda^*)$ and $\mathbf{y}^*$ is a unique minimizer of $F(\mathbf{y}, \lambda^*, k)$ for any $k > 0$.

Taking into account that during the minimization procedure the vector $\mathbf{y}$ is positive, the MBF $F(\mathbf{y}, \lambda, k)$ is strongly convex and smooth in an extended feasible set, $\Omega_k = \{\mathbf{y} : g_j(\mathbf{y}) \geq -k^{-1}, i = 1, \ldots, m\}$.

The MBF properties enable us to develop a method, which converges to the primal-dual solution for any positive parameter $k > 0$ (see Jensen and Polyak 1994). This means that in contrast to the interior point methods, that are based on classical barrier functions (see Gonzaga 1992), the MBF approach provides convergence without increasing the barrier parameter unboundedly.

Therefore, the condition number of the MBF Hessian is stable from the beginning to the end of the process. It contributes to both the rate of convergence and numerical stability. It also makes more efficient use of Newton's method for unconstrained minimization $F(\mathbf{y}, \lambda, k)$ in $\mathbf{y}$, which is the key element in the numerical realization of the MBF method that we are going to describe in the next section.

4 Modified barrier methods

The MBF method consists of unconstrained minimization $F(\mathbf{y}, \lambda, k)$ in $\mathbf{y}$ and updating the vector of Lagrange multipliers λ. The barrier parameter $k > 0$ can be fixed or one can change it after every Lagrange multiplier update. We will now describe two versions of the MBF method, starting with the permanent parameter version.

We start with positive vector $\mathbf{y}^0$, which belongs to the extended feasible set, i.e.

$$g_j(\mathbf{y}^0) \geq -k^{-1}, \quad j = 1, \ldots, m. \tag{11}$$

It is worth mentioning that for any positive vector $\mathbf{y}^0$, one can find such $k > 0$ that inequalities (11) will be satisfied.

Thus, finding the initial primal approximation does not require extra computational work. As far as the initial vector of Lagrange multipliers is concerned, we can choose any positive vector $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)$.

Let $k > 0$ be fixed, $\mathbf{y} \in \Omega_k$, and $\lambda^0 = (1, \ldots, 1) \in \mathbf{R}^n$. Assume that the approximations $\mathbf{y}^s$ and λ^s have been found already. Then the next approximations for $\mathbf{y}^{s+1}$ and λ^{s+1} are found by the formulae

$$\mathbf{y}^{s+1} : F(\mathbf{y}^{s+\ell}, \lambda^s, k) = \min\{F(\mathbf{y}, \lambda^s, k) \mid \mathbf{y} \in \mathbf{R}^n\}, \tag{12}$$

and

$$\lambda^{s+1} : \lambda_j^{s+1} = \lambda_j^s[k g_j(\mathbf{y}^{s+1}) + 1]^{-1}, \quad j = 1, \ldots, m. \tag{13}$$

From (11) and (12), we have

$$\nabla_{\mathbf{y}} F(\mathbf{y}^{s+1}, \lambda^s, k) = \nabla_{\mathbf{y}} L(\mathbf{y}^{s+1}, \lambda^{s+1}) = 0, \tag{14}$$

i.e. the unconstrained minimum of the MBF $F(\mathbf{y}, \lambda^s, k)$ coincides with the unconstrained minimum of the classical Lagrangian $L(\mathbf{y}, \lambda^{s+1})$ in $\mathbf{y}$.

It was proven (see Jensen and Polyak 1992) that if Ω is bounded and (6) holds true, then the method (12)-(13) is executable, i.e. the unconstrained minimizer $\mathbf{y}^s$ exists for any $k > 0$. In the case of the structural optimization problem (4)-(5), the primal minimizer y^s is always unique.

It was also proven that the method (12)-(13) converges, i.e. $\lim_{s\to\infty} f(\mathbf{y}^s) = f(\mathbf{y}^*)$ whether or not the objective functions and the constraint functions are strongly convex.

Moreover, the next estimation,

$$\mid f(\mathbf{y}^s) - f(\mathbf{y}^*) \mid = (k\alpha_s)^{-1/2}, \tag{15}$$

holds true for any fixed $k > 0$ and $\lim_{s\to\infty} \alpha_s = 0$. For a structural design problem, the estimation (15) can be improved. The objective function $f(y)$ for the structural design problem (4)-(5) is strongly convex; therefore, if λ^* is unique, then (due to Polyak 1992a) the MBF method (12)-(13) converges with a linear rate. In other words, for any fixed and large enough $k > 0$ there exists $0 < q < 1$ such that the following estimation holds true:

$$\max\{\| \mathbf{y}^{s+1} - \mathbf{y}^* \|, \| \lambda^{s+1} - \lambda^* \|\} \leq$$

$$(c/k) \| \lambda^s - \lambda^* \| = q \| \lambda^s - \lambda^* \|, \tag{16}$$

and $q \to 0$ if $k \to \infty$. Therefore by increasing the barrier parameter k step by step, one can improve the rate of convergence.

Let us consider the second version of the MBF method

$$\mathbf{y}^{s+1} : F(\mathbf{y}^{s+1}, \lambda^s, k) = \min\{F(\mathbf{y}, \lambda^s, k_s) \mid \mathbf{y} \in \mathbf{R}^n\}, \tag{17}$$

$$\lambda^{s+1} : \lambda_j^{s+1} = \lambda_j^s[k_s g_j(\mathbf{y}^{s+1}) + 1]^{-1}, \quad j = 1, \ldots, m. \tag{18}$$

If λ^* is unique, then the following estimation is true:

$$\max\{\| \mathbf{y}^s - \mathbf{y}^* \|, \| \lambda^s - \lambda^* \|\} \leq q_1 \cdot \ldots \cdot q_s,$$

with $q_s \to 0$ if $k_s \to \infty$.

Before we turn to the numerical realization of the MBF method, we would like to emphasize that for every $s \geq 1$,

$$\nabla_{\mathbf{y}} L(\mathbf{y}^s, \lambda^s) = 0. \tag{19}$$

Also, $g_i(\mathbf{y}^s) \to 0$ for the active constraints, i.e. for $j \in J = \{j : g_j(\mathbf{y}^*) = 0\}$.

For the passive constraints $j : g_j(\mathbf{y}^*) \geq \sigma > 0$, therefore due to the formula for the Lagrange multipliers update, we have

$$\lambda_j^s = \lambda_j^{s-1}[k g_j(\mathbf{y}^s) + 1]^{-1} \leq \lambda^s(k\sigma + 1)^{-1}, \tag{20}$$

i.e. $\lambda_j^s \leq \lambda_j^0(k\sigma + 1)^{-s} \to 0, j \bar{\epsilon} J$.

We can see that, in contrast to the augmented Lagrangian method (see Powell 1969; Rockafellar 1973; Bertsekas 1982), the MBF does not require any special care to keep the Lagrange multipliers non-negative. It happens automatically because of (13) or (18) for Lagrange multiplier update.

Due to (14) and the formula for Lagrange multiplier update, the sequence $\{(\mathbf{y}^s, \lambda^s)\}_{s=0}^{\infty}$ converges to a KKT pair $(\mathbf{y}^*, \lambda^*)$. To estimate the "distance" between the current approximation $(\mathbf{y}, \lambda)$ and the solution $(\mathbf{y}^*, \lambda^*)$ we introduce the following merit function:

$$v(\mathbf{w}, k) = v(\mathbf{y}, \lambda, k) =$$

$$\max\{G(\mathbf{y}), \| \nabla_{\mathbf{y}} L(\mathbf{y}, \lambda) \|, \sum_{j=1}^{m} \lambda_j \mid g_j(\mathbf{y}) \mid\},$$

where $G(\mathbf{y}) = -\{\min g_j(\mathbf{y}) \mid 1 \leq j \leq< m\}$.

It is easy to see that for any feasible vector y, any positive Lagrange multiplier vector $\lambda = (\lambda_1, \ldots, \lambda_m)$ and any $k > 0$ the merit function $v(w, k)$ is non-negative.

It is also clear that

$$v(\mathbf{y}, \boldsymbol{\lambda}, k) = 0 \tag{21}$$

is equivalent to (7), so that

$$v(\mathbf{y}, \boldsymbol{\lambda}, k) = 0 \longleftrightarrow \mathbf{y} = \mathbf{y}^*, \boldsymbol{\lambda} = \boldsymbol{\lambda}^*,$$

for any $k > 0$.

Also, there exists $M > 0$ such that

$$v(s(\mathbf{w}^s, k) \equiv v(\mathbf{y}^s, \boldsymbol{\lambda}^s, k) =$$

$$v(\mathbf{y}^s, \boldsymbol{\lambda}^s, k) - v(\mathbf{y}^*, \boldsymbol{\lambda}^*, k) \le M \parallel \boldsymbol{\lambda}^s - \boldsymbol{\lambda}^* \parallel .$$

Therefore for any given $0 < \gamma < 1$ there exists $k_\gamma > 0$ such that

$$v(\mathbf{w}^s, k) \equiv v(\mathbf{y}^s, \boldsymbol{\lambda}^s, k) \le \gamma^s, \tag{22}$$

for any $k \ge k_\gamma$.

We shall use the merit function $v(y, \lambda, k)$ and the estimation (22) for the numerical realization of the MBF method.

5 Newton modified barrier method

The modified barrier methods (12)-(13) or (17)-(18) require solving the unconstrained optimization problem (12) or (17) at every step, which in turn, requires an infinite number of arithmetic operations. Therefore, to realize the MBF method numerically, we must find a way to replace the vector $\mathbf{y}^s$ in (13) by such an approximation $\hat{y}^s$, which can be found in finite number of arithmetic operations and retains the estimation (16). Such a way has been indicated by Polyak (1992a, see p. 205).

In this paper, to find an approximation for $\mathbf{y}^s$ we will use the Newton method up to the point when the norm in the Newton direction becomes smaller than a given small enough number $\varepsilon > 0$, which is defined by the accuracy which we finally want to achieve.

The Newton method is particularly efficient for structural optimization due to the following three facts.

1. The MBF $F(y, \lambda, k)$ is smooth and strongly convex for $y \in \Omega_k$.
2. It is easy to compute the MBF Hessian $\nabla^2_{yy} F(y, \lambda, k)$.
3. The numerical linear algebra for solving normal systems of equations to find the Newton directions is well-developed.

To describe the Newton MB method we introduce the matrix $\Delta(\mathbf{y}, k) = \{\operatorname{diag}[kg_j(\mathbf{y})+1]\}_{j=1}^m$ with entries $kg_j(y)+1$. Let us also remember that

$$\mathbf{R} = [\operatorname{diag} r_i]_{i=1}^m, \mathbf{Y} = [\operatorname{diag} y_i]_{i=1}^n, \Lambda = [\operatorname{diag} \lambda_j]_{j=1}^m .$$

For the gradient, we obtain

$$\nabla_{\mathbf{y}} F(\mathbf{y}, \boldsymbol{\lambda}, k) = -\mathbf{R}\mathbf{y}^{-2} - \boldsymbol{\lambda}\Delta^{-1}(\mathbf{y}, k)\mathbf{Q}, \tag{23}$$

while the Hessian is given by

$$\nabla^2_{\mathbf{yy}} F(\mathbf{y}, \boldsymbol{\lambda}, k) = 2\mathbf{R}\mathbf{y}^{-3} + k\mathbf{Q}^T \Lambda \Delta^{-2}(\mathbf{y}, k)\mathbf{Q} . \tag{24}$$

Now we will describe the NMBM. The NMBM has two basic parts. First, we find the approximation $\hat{y}^{s+1}$ for the primal minimizer $\mathbf{y}^{s+1}$. To find the approximation we will use the Newton method for minimizing the function $F(\mathbf{y}, \boldsymbol{\lambda}^s, k)$ in y when both the vector of Lagrange multipliers $\boldsymbol{\lambda}^s$ and the parameter $k > 0$ are fixed.

Second, we update the Lagrange multipliers by using the approximation $\hat{\mathbf{y}}^{s+1}$ instead of $\mathbf{y}^{s+1}$ in (13).

After every Lagrange multiplier update, we check the merit function $v(w, k)$, which is supposed to converge to zero at an *a priori* given rate $0 < \gamma < 1$. If the barrier parameter $k > 0$ is chosen properly, i.e. $k \ge k_\gamma > 0$, then every Lagrange multiplier update will decrease the function $v(w, k)$ by a given factor $0 < \gamma < 1$, i.e.

$$v(w^{s+1}, k) \le v(w^s, k)\gamma . \tag{25}$$

If the inequality (25) does not hold, it means that the barrier parameter is not large enough and we must increase it.

The most costly operation in the NMBM is solving the system of equations

$$\nabla^2_{\mathbf{yy}} F(\mathbf{y}, \boldsymbol{\lambda}^s, k)d = -\nabla_{\mathbf{y}} F(\mathbf{y}, \boldsymbol{\lambda}^s, k), \tag{26}$$

to find the Newton direction $d \equiv d(\mathbf{y}, \boldsymbol{\lambda}^s, k)$. To guarantee convergence from any starting point $\mathbf{y}^0 \in \Omega$, we must use the Newton method with a step size.

To find a proper step size, we will use the relaxation inequality,

$$F(\mathbf{y} + t\mathbf{d}, \boldsymbol{\lambda}, k) - F(\mathbf{y}, \boldsymbol{\lambda}, k) \le (1/3)t[\nabla_{\mathbf{y}} F(\mathbf{y}, \boldsymbol{\lambda}, k), \mathbf{d}] , \tag{27}$$

starting with $t = 1$.

If (27) is fulfilled for $t = 1$, then we have a "pure" Newton step; otherwise, we check (27) for $t := (t/2)$ and continue to do so while the inequality (27) will be fulfilled.

Due to the properties of the MBF $F(\mathbf{y}, \boldsymbol{\lambda}, k)$, for any fixed positive vector $\boldsymbol{\lambda} > 0$ and $k > 0$, it will come to a point where the inequality (27) will be fulfilled for $t = 1$. At this point, we have reached the Newton area and, from this point on, we will use only "pure" Newton steps. Therefore, the minimization procedure will converge to the primal minimizer

$$\mathbf{y}(\boldsymbol{\lambda}, k) = \operatorname{argmin}\{F(\mathbf{y}, \boldsymbol{\lambda}, k) \mid x \in \mathbf{R}^n\}$$

quadratically. The stopping criterion is given by the formula

$$\parallel \mathbf{d} \parallel_{\nabla^2_{\mathbf{yy}} F(\mathbf{y}, \boldsymbol{\lambda}, k)} = \parallel \mathbf{d} \parallel_{\nabla^2_{\mathbf{yy}} F(\cdot)} \le \varepsilon , \tag{28}$$

where $\varepsilon > 0$ is the small enough number defined by the accuracy of the approximation to the solution y^* we want to achieve. Now, we describe the Newton modified barrier method (see flow chart).

Let ε be the accuracy which we want to achieve for the final approximation $\hat{y}$ to the solution y^*. We consider a monotonically increasing sequence $\{k_s\}_{s=0}^{\infty} : \lim k_s = +\infty$.

We start with $\mathbf{y}^0 \in \Omega_k, \lambda_0 = (1, \ldots, 1) \in \mathbf{R}^m, k = k(0) = k_0, \ell(0) = 1$, and $0 < \gamma \le 0.5$. Now assume that $y^s, \lambda^s, k(s), \ell(s)$ have been found already. To find the next approximation, one must perform the following steps.

0. Start with $\mathbf{y} := \mathbf{y}^s$.
1. Set $\boldsymbol{\lambda} := \boldsymbol{\lambda}^s, k := k(s), \ell := \ell(s)$.
2. Find the Newton direction $\mathbf{d} = \mathbf{d}(\mathbf{y}, \boldsymbol{\lambda}, k)$ by solving the system
$$\nabla^2_{\mathbf{yy}} F(\mathbf{y}, \boldsymbol{\lambda}, k)d = -\nabla_{\mathbf{y}} F(\mathbf{y}, \boldsymbol{\lambda}, k) \tag{29}$$
and set $t = 1$.
3. If $\mathbf{y} + t\mathbf{d} \in \Omega_k$, then go to 4; if $\mathbf{y} + t\mathbf{d} \notin \Omega_k$, then $t := (t/2)$ and go to 3.
4. Check the inequalities (27). If they are not fulfilled, set $t := (t/2)$ and check (27) again; if (27) is fulfilled and $t = 1$, go to 5; if (27) is fulfilled but $0 < t < 1$, then set $y := y + td$ and go to 2.
5. If $\parallel d \parallel_{\nabla_{yy} F(\cdot)} \le \varepsilon$, go to 6; otherwise, go to 2.

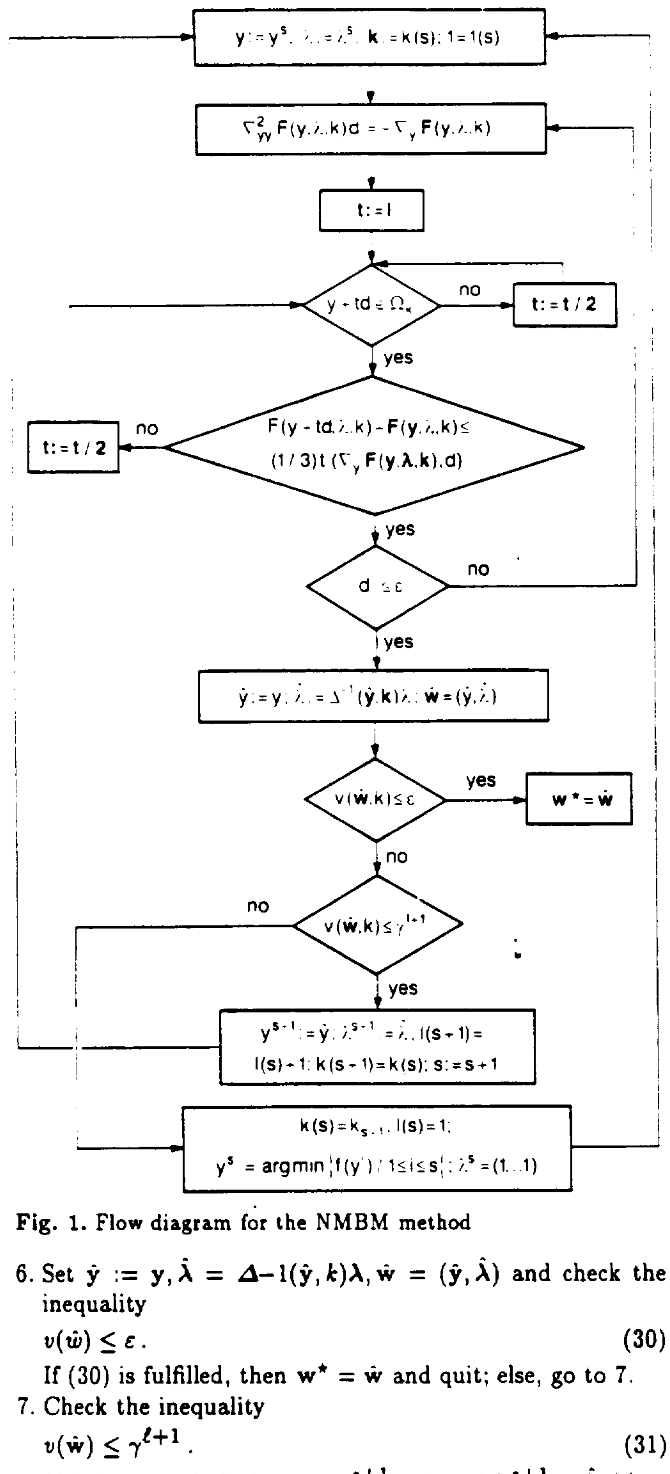

Fig. 1. Flow diagram for the NMBM method

6. Set $\hat{\mathbf{y}} := \mathbf{y}, \hat{\boldsymbol{\lambda}} = \boldsymbol{\Delta}{-}1(\hat{\mathbf{y}}, k)\boldsymbol{\lambda}, \hat{\mathbf{w}} = (\hat{\mathbf{y}}, \hat{\boldsymbol{\lambda}})$ and check the inequality

$$v(\hat{w}) \leq \varepsilon\,. \tag{30}$$

If (30) is fulfilled, then $\mathbf{w}^* = \hat{\mathbf{w}}$ and quit; else, go to 7.

7. Check the inequality

$$v(\hat{\mathbf{w}}) \leq \gamma^{\ell+1}\,. \tag{31}$$

If (31) is satisfied, then set $\mathbf{y}^{s+1} = \hat{\mathbf{y}}$ and $\boldsymbol{\lambda}^{s+1} = \hat{\boldsymbol{\lambda}}, k(s+1) = k(s), \ell(s+1) = \ell(s)+1, s+1 := s$, and go to 1. If (31) is not satisfied, then set $\mathbf{y}^{s+1} = \operatorname{argmin}\{f(\mathbf{y}^i) \mid i = 1,\ldots,s+1\}, \boldsymbol{\lambda}^{s+1} = e = (1,\ldots,1), k(s+1) = k_{s+1}, \ell(s+1) = \ell(s), s+1 := s$, and go to 1.

For any $0 < \gamma < 1$ there is a number s_0 such that for $s \geq s_0$, the barrier parameter will not change and the step-size $t = 1$; i.e. we will perform only the "pure" Newton method and the primal minimizer will stay in the Newton area after every Lagrange multiplier update. Such a point we call a *hot-start*.

From this point on, every Lagrange multiplier update requires at most $O(\ln\ln\varepsilon^{-1})$ Newton steps, and every Lagrange multiplier update shrinks the distance between the current approximation and the primal-dual solution by a factor $0 < \gamma < 1$. This follows from the basic properties of MBF (see Polyak 1992a) and Newton's method (see Smale 1986).

The number s_0 depends on the condition of the constrained optimization problem (see Polyak 1992a, b), which in turn depends on the "measure" of nondegeneracy of the problem (4)-(5). The value s_0 can be decreased by increasing the barrier parameter $k > 0$. Moreover, for any nondegenerate constrained optimization problem and any fixed $0 < \gamma \leq 0.5$, there exists $k_\gamma > 0$ such that for any $k \geq k_\gamma$, one can achieve the *hot-start* with $s_0 = 1$.

However, if $k > 0$ is large enough it will take extra computational work to reach the *hot-start* at the first stage, where the Lagrange multipliers are fixed and one must change the barrier parameter $k > 0$. To reach the *hot-start*, one can use shifted barrier function

$$F(x,e,k) = \sum_{i=1}^{n} r_i y_i^{-1} + k^{-1}\sum \ln[kg_j(y)+1]\,.$$

In the case of the structural optimization problem, the shifted barrier functions possess the so-called self-concordant properties (see Nesterov and Nemirovsky 1994). Therefore from a "warm-start" for a fixed $k > 0$, i.e. having an approximation in the Newton area for $\mathbf{y}(k) = \operatorname{argmin}\{F(\mathbf{y},\mathbf{e},k) \mid \mathbf{y} \in \mathbf{R}^n\}$, one can improve the current approximation in $[1-O(m^{-0.5})]$ time ($m > n$) by performing one Newton step and updating the barrier parameter k to $k[1-O(m^{-0.5})]^{-1}$ (see Gonzaga 1992; Nesterov and Nemirovsky 1994). All this means that to reach the hot-start, one must perform $O(\sqrt{m}\ln k_\gamma)$ Newton steps.

Beginning from the hot-start, we do not need to change the barrier parameter k from step to step; instead, we update the Lagrange multipliers after at most $O(\ln\ln\varepsilon^{-1})$ Newton steps by minimizing $F(y,\lambda,k)$ in y under fixed $k > k_\gamma$. Every Lagrange multiplier update leads to an improvement of the current approximation by a factor $0 < \gamma \leq 0.5$.

In other words, from the hot-start on, one must perform at most $O(\ln\ln\varepsilon^{-1})$ Newton steps instead of $O(\sqrt{m})$ Newton steps in the interior point methods to improve the current approximation by the same factor. Moreover, from some point on, the number of Newton steps between two successive Lagrange multiplier update decreases after every update. This is very important for structural optimization because solving problem (4)-(5) is only one part of the procedure. The second part, which is very time-consuming, is structural analysis. This analysis starts with the current approximation for y, which we use to recompute the matrix Q and then to solve the problem (4)-(5) again. The existence of the hot-start makes the optimization part of the procedure substantially easier. In the next section, we will show some numerical results.

6 Examples

The algorithms based on the NMBM to obtain a minimum weight structure were used to optimize three truss structures. These structures have been used before for studying the characteristics of the optimality criteria algorithm (see Berke and

Khot 1987; Morris 1982; and others). In this investigation, the structures were also designed by using the ASTROS program and another program based on an optimality criterion (OC) approach.

In the program using an optimality criterion approach, a linear recursive relation based on the reciprocal design variables was used to update the design variables, and a linear set of simultaneous equations was used to calculate the Lagrange multipliers associated with the active constraints. The three programs were run on three different computer systems and locations and, as such, an exact comparison of the computational time was not feasible. The main objective was instead to study the convergence characteristics of the methods and the various associated programs.

The analysis package for finite element analysis and evaluation of the sensitivities for the programs based on the optimality criteria and the NMB method were identical. The design surfaces near the optimum for all the problems were nearly flat and consequently convergence near the optimum was very slow. The initial designs for all the structures were feasible and were obtained by analysing the structure with cross-sectional areas equal to 1.0 in^2 for all the elements, then scaling the design to satisfy the constraints. The elastic modulus was 10^7 lbs/in^2 and the density of the material used to calculate the weight of the structure was 0.1 lbs^3/in^3. The cross-sectional areas of all the elements were treated as design variables.

6.1 Problem 1 - 42-element truss

The structure shown in Fig. 2 was subjected to +40 kips at nodes 5 and 6, and +20 kips at nodes 11 and 12, respectively, in the vertical direction. The displacements at nodes 6 and 12 were limited to 10 in and 12 in, respectively, and the minimum size constraint was 0.1 in^2. The initial sizes of all the elements were 39.746 in^2 and the weight of the structure was 57111.67 lbs. The iteration history for the three methods is given in Fig. 3. The optimum design weight was 17425.72 lbs.

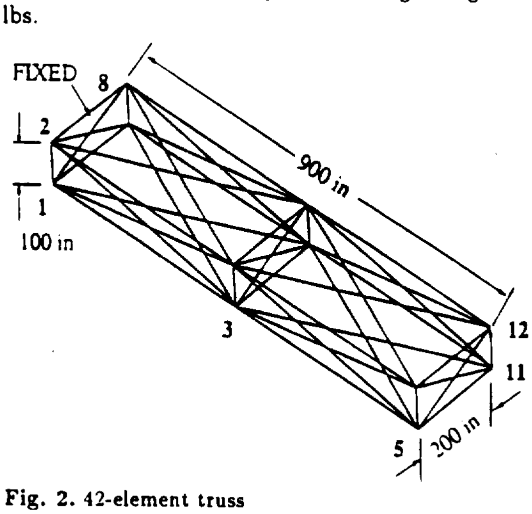

Fig. 2. 42-element truss

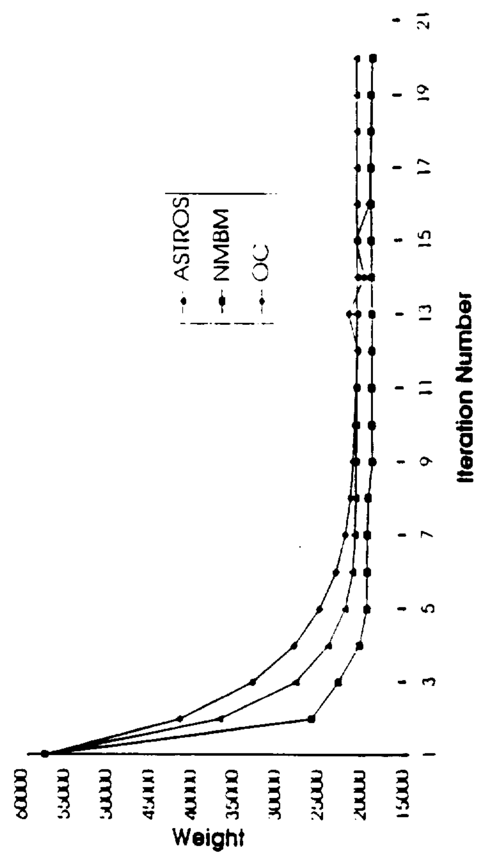

Fig. 3. Iteration history for 42-element truss

6.2 Problem 2 - 148-element truss

The truss shown in Fig. 4 was subjected to 40 kips loads applied in the $x, y,$ and z-directions at nodes 7 and 8. Similarly, a load of 20 kips was applied at nodes 31 and 32 in the $x, y,$ and z-directions. The displacements at nodes 8 and 32 were restricted to 10 in and 20 in, respectively, while the minimum size constraint was 0.1 in^2. The iteration history is given in Fig. 5. The initial areas of all the elements were 8.07 in^2, and the weight of the structure was 29474.29 lbs. The optimum design weight was 10499.41 lbs.

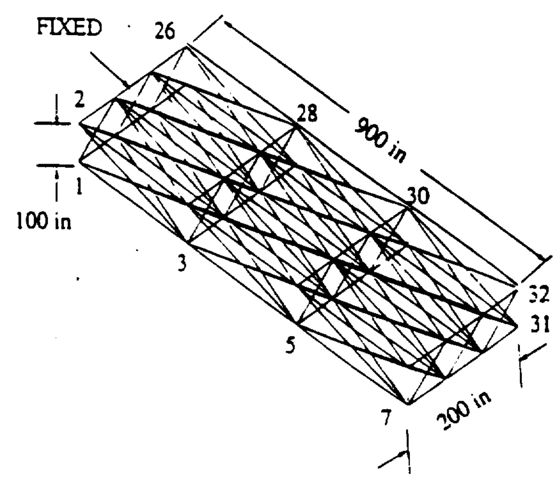

Fig. 4. 148-element truss

6.3 Problem 3 - 721-element truss

The geometry of the structure is shown Fig. 6. It has 182 nodes and 721 elements. The length, width and height of the box-beam are 600 in, 240 in and 30 in, respectively. The

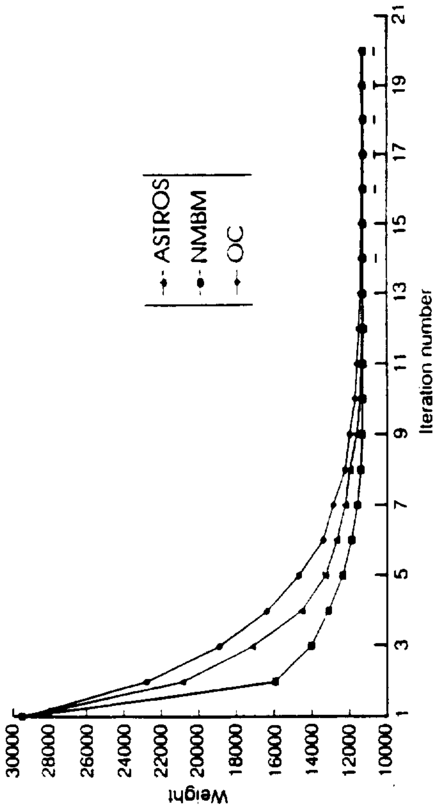

Fig. 5. Iteration history for 148-element truss

box-beam has six bays in the transverse direction and twelve bays in the longitudinal direction.

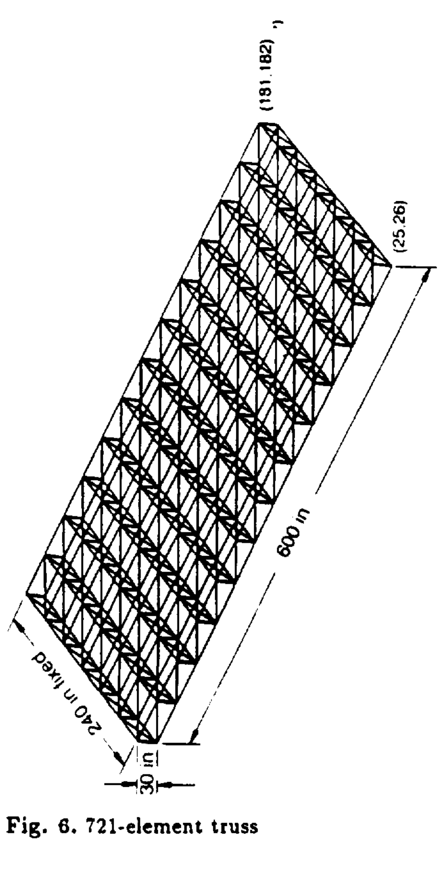

Fig. 6. 721-element truss

A load of 40000 lbs was applied in the three coordinate directions at nodes 25 and 26. At nodes 181 and 182, a load of 20000 lbs was applied in the three coordinate directions. The minimum size constraint was 1 in^2.

The initial area of the members was equal to 32.323 in^2, which was obtained by scaling the structure to satisfy the displacement constraints. The initial weight of the structure was 114384 lbs. The iteration history for this problem is shown in Fig. 7. The ASTROS program was terminated after twenty iterations because of the excessive computational time needed for each additional iteration. The optimum design for this structure had a weight 20624.2 lbs.

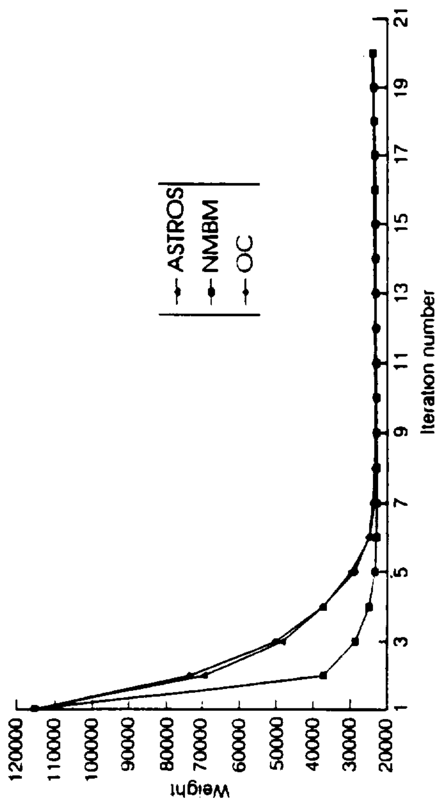

Fig. 7. Iteration history for 721-element truss

The main difference between the three structures was the number of design variables. The objective in this study was to study how the number of design variables influences the convergence characteristics. Comparing the results for all the three problems, it is seen that the initial rate of decrease in weight with each iteration was substantially higher for the NMBM algorithm than for the other methods. This was true for all three problems in spite of the differences in the number of design variables. This indicates that the NMB method makes maximum use of the information provided by the analysis and sensitivity subroutines. It should be remarked that the computational effort for these structural optimization problems, satisfying separability conditions, is the least by far for the OC method. The NMBM pays for its quick convergence with many internal iterations within an outer iteration. ASTROS uses Taylor series approximations for internal iterations that are not needed at all for OC. For all three methods, each outer iteration is defined here as just one full structural analysis.

7 Summary

The Newton modified barrier method, based on rigorous mathematical derivation, was used to solve three structural optimization test problems. The varying number of size variables for the three probllms was used to study the effect of the number of variables on the number of iterations needed to achieve similar levels of convergence. This dependence was shown to be weak for the NMBM, similar to this well-known characteristic of the basic OC methods. The same problems were solved by the ASTROS program with MicroDot as the optimizer, and also by an OC-based program. The NMBM is shown to have superior convergence properties, reducing the objective function at a much faster and more uniform rate, thus indicating superior robustness. This fast converging robust approach has promising potential for solving structural sizing problems with large numbers of variables within the emerging multidisciplinary optimization (MDO) problem setting.

References

Ben-Tal, A.; Yuzefovich, I.; Zibulevsky, M. 1992: Penalty/barrier multiplier methods of minimax and constrained smooth convex program. *Research Report 9/92, Optimization Laboratory, Technion, Haifa, Israel*

Berke, L.; Khot, N.S. 1987: Structural optimization using optimality criteria. In: Mota Soares, C.A. (ed.) *Computer aided optimal design: structural and mechanical systems*, pp. 271-311. Berlin, Heidelberg, New York: Springer

Bertsekas, D. 1982: *Constrained optimization and Lagrange multiplier methods*. New York: Academic Press

Fiacco, A.; McCormick, G. 1968: *Nonlinear programming: sequential unconstrained minimization techniques*. New York: John Wiley and Sons

Received Feb. 9, 1995

Gendy, A.; Patnaik, S.; Hopkins, D.; Berke, L. 1994: Design optimization of solar power module components of space station Fredom using COMETBOARDS. *Proc. 5th AIAA/NASA/USAF/ISSMO Symp. on Multidisciplinary Analysis and Optimization*, Part 2, pp. 881-891

Gonzaga, C. 1992: Path-following methods for linear programming. *SIAM Rev.* **34**, 167-224

Jensen, D.; Polyak, R. 1994: The convergence of a modified barrier method for convex programming. *IBM J. Res. Develop.* **38**, 307-321

Khot, N.S. 1981: Algorithms based on optimality criteria to design minimum weight structure. *Eng. Opt.* **5**, 73-90

Morris, A.J. (ed.) 1982: *Foundations of structural optimization: a unified approach*. New York: John Wiley and Sons

Neill, D.J.; Herendeen, D. 1993: ASTROS enhancements, volume 1: ASTROS user's manual. *Technical Report WL-TR-93-3025*

Nesterov, Y.; Nemirovsky, A. 1994: *Interior point polynomial methods in convex programming: theory and algorithms*. Philadelphia: SIAM

Polyak, R. 1992a: Modified barrier functions (theory and methods). *Math. Prog.* **54**, 177-222

Polyak, R. 1992b: Modified barrier functions in linear programming. *Research Report, IBM T.J. Watson Research Center, New York*

Powell, M.J. 1969: A method for nonlinear constraints in minimization problems. In: *Optimization*, pp. 283-298. New York: Academic Press

Rockafellar, R.T. 1973: Dual approach to solving nonlinear programming problems by unconstrained minimization. *Math. Prog.* **5**, 354-379

Smale, S. 1986: Newton's method estimates from data at one point. In: Ewing, R.E. *et al.* (eds.) *The merging of disciplines in pure, applied and computational mathematics*, pp. 185-196. New York: Springer