Computers & Structures Vol. 33, No. 2, pp. 375–384, 1989
Printed in Great Britain.

0045-7949/89 $3.00 + 0.00

ISOLATED GLOBAL OPTIMALITY IN TRUSS SIZING PROBLEMS

L. M. C. Simões

Departamento de Engenharia Civil, Faculdade de Ciencias e Tecnologia, Universidade de Coimbra, Coimbra, Portugal

(*Received* 23 *September* 1988)

Abstract—A new test is described to verify the existence of an isolated local constrained minimum and to ensure that it is a global minimum for a nonconvex program such as the least-volume design of elastic trusses. Illustrative examples covering most of the situations found in engineering practice are given.

1. INTRODUCTION

Many efficient computer codes have been proposed for minimum volume/weight design with continuous sizing variables. These methods can be divided into two categories: the optimality criteria and the mathematical programming approach. The first category is based on the minimization of the Lagrangian and the second on the application of nonlinear parameter optimization methods. (The literature is extensive, and has been reviewed in [1].) These methods need some starting point which essentially prescribes both computational effort and optimization performance. As, in most of these problems, the solution domain is nonconvex and has many subminima, the algorithms tend to converge to a minimum near the starting point. Therefore, the possibility of multiple optima and the risk of achieving a sub-optimal (locally optimum) solution still remain. Combinatorial optimization methods [2] can be expected to overcome this problem although they are computationally expensive.

Many theories are available to characterize the local minima. Perhaps the most suitable theoretical development for our purposes is that of McCormick [3]. Unfortunately, it happens to be rather abstruse. It has been quoted to reflect its full scope, but successful applications of the theory depend on problem-specific simplifications.

In this paper a condition is given which, if satisfied, guarantees that a local optimum is the global optimum for the optimal sizing of trusses. Before meaningful results applicable to this problem are given, some conditions and theorems will be stated.

2. CHARACTERIZATION OF LOCAL SOLUTIONS

In nonlinear programming problems a local minimum of the objective function in the constraint set is often obtained. The question that naturally arises is whether this minimum is the global minimum. For convex programming problems a local minimum is also the global minimum. Without convexity it is difficult to say much about global properties of the solution. Many theories are available, however, to characterize local minima. The approach in local theory is to suppose that the mathematical programming problem

$$\min f(x_i) \tag{1a}$$

such that

$$g_j(x_i) = 0; \quad j = 1, \ldots \beta; \quad i = 1, \ldots n \tag{1b}$$

$$g_j(x_i) \geqslant 0; \quad j = \beta + 1, \ldots Q \tag{1c}$$

has a local extremum at point x^* and then find conditions among $f(x)$ and $g(x)$ that must hold at this point. In this way, many points in the constrained region can be eliminated as candidates for a relative minimum. Such conditions are therefore considered necessary. In some problems it is possible to obtain a set of conditions guaranteeing that a point yields a local minimum. Conditions of this kind will be termed 'sufficient'. This paper describes a test which ensures that an isolated local constrained minimum ('isolated' meaning it is not a limit point of local minima) is a global minimum for a nonconvex program. Before meaningful results may be given for eqn (1), some conditions are given, which must be satisfied by the objective function (1a) and the constraints (1b) and (1c).

2.1 *Regularity assumption*

Definition [4]. A feasible point $\mathbf{x}$ is called a regular point of the domain if $f(x)$ is differentiable at $\mathbf{x}$ and if the gradients $\nabla_x g_j(\mathbf{x})$ for only those $j > \beta$ with $g_j(\mathbf{x}) = 0$ (active inequality constraints) and all $\nabla_x g_j(\mathbf{x})$ for $j = 1, \ldots \beta$ (equality constraints) are linearly independent.

This definition is equivalent to the requirement that the $q \times n$ matrix of the active constraints' derivatives

$C(\mathbf{x})'$ has full rank, where the first β rows are the derivatives $g_j(\mathbf{x})$ for $j = 1, \dots \beta$ and the remaining rows refer to $j > \beta$ with $g_j(\mathbf{x}) = 0$ $(j = 1, \dots q)$.

2.2 Karush–Kuhn–Tucker (K–K–T) necessary conditions

Theorem [4]. Let the functions $f(x)$ and $g_j(x)$, for $j = 1, \dots q$, where the first β constraints and the remaining $q - \beta$ are the equality constraints and the active inequality constraints, respectively, be differentiable and let $\mathbf{x}$ be a feasible regular point. For $\mathbf{x}$ to be a local minimum of eqn (1) it is necessary that there exist multiplier vectors $\gamma \in \mathbb{R}\beta$, $\mu \in \mathbb{R}q - \beta$ such that

$$\boldsymbol{\mu}_j \geqslant 0; \quad \text{for } j = \beta + 1, \dots q \tag{2a}$$

$$\gamma_j g_j(\mathbf{x}) = 0; \quad \text{for } j = 1, \dots \beta \tag{2b}$$

$$\nabla_{xi}\mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu}) = 0; \quad \text{for } i = 1, \dots n, \tag{2c}$$

where

$$\mathscr{L}(x, \gamma, \mu) = f(x) - \Sigma_{j=1,\beta}\gamma_j g_j(x) - \Sigma_{j=\beta+1,q}\mu_j g_j(x)$$

is termed the Lagrangian.

2.3 Implicit function theorem

A result of matrix calculus concerning the solvability of the systems of equations defined by the K–K–T necessary conditions is required. Defining the function,

$$\phi(z) = \begin{bmatrix} \nabla_x \mathscr{L}(x, \gamma, \mu) \\ -C(x) \end{bmatrix}, \tag{3}$$

where $z' = [x'\gamma'\mu']$ and $C(x)$ is a q-vector whose elements are the active constraints at the regular point x, to satisfy the K–K–T conditions at a local minimum it is necessary to have

$$\phi(\mathbf{z}) = 0 \quad \text{and} \quad v_j \geqslant 0 \quad \text{for } \beta \leqslant j \leqslant q. \tag{4}$$

Theorem [5]. If at $z = \mathbf{z}$ there is a solution $\phi(\mathbf{z}) = \boldsymbol{\eta} = 0$ and the $(n + q) \times (n + q)$ derivative matrix of $\phi(z)$ has an inverse at every point z, then there is a continuously differentiable solution to the equations $\eta = \phi(z)$ in the neighbourhood of $\mathbf{z}$.

This theorem allows investigation of the effect of variations of z on the solution $\eta = \phi(z)$, and in particular on the objective function value $f(x)$, caused by variations in x.

2.4 Second-order sufficiency conditions for local optimality

Theorem [4]. Let $f(x)$ and $g_j(x)$, $j = 1, \dots q$, where the first β constraints and the remaining $q - \beta$ are the equality constraints and the active inequality constraints, respectively, have two continuous derivatives. Let $\mathbf{x}$ be a feasible regular point satisfying the K–K–T necessary conditions for eqn (1).

For every $v \neq 0$ in $\mathbb{R}^n$ such that $\nabla_x g_j(\mathbf{x})v = 0$, for $j = 1, \dots \beta$, and $\nabla_x g_j(\mathbf{x})v = 0$, for $j = \beta + 1, \dots q$ for each $j > \beta$ with $\boldsymbol{\mu}_j > 0$, let

$$v'\nabla_x^2\mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu})v > 0, \tag{5}$$

where $\nabla_x^2\mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu})$ is the matrix representing the second derivative of the Lagrangian. Then $\mathbf{x}$ is an isolated local minimum.

The following development will allow us to express eqn (5) in more general terms. Suppose $N(x)$ denotes a continuously differentiable $n \times (n - q)$ matrix function giving the null space of the active constraint derivative matrix $C(x)'$. A necessary and sufficient condition for $C(x)'v = 0$ is that $v = N(x)u$ for some u.

It is therefore possible to express the condition (5) so as to have the matrix

$$H(\mathbf{x}, \gamma, \boldsymbol{\mu}) = N(\mathbf{x})'\nabla_x^2\mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu})N(\mathbf{x}) \tag{6}$$

positive definite.

Also associated with $C(x)'$ is a generalized inverse $C(x)'_{\#}$. Note that when the rank of $C(x)'$ is q, $C(x)'_{\#}$ is given by

$$C(x)'\# = C(x)'[C(x)'C(x)']^{-1}, \tag{7}$$

and $C(x)'C(x)'_{\#} = I_q$. Further, it is assumed that $N(x)$ and $C(x)'_{\#}$ are related in that

$$[I_n - C(x)'_{\#}C(x)'] = N(x)W(x) \tag{8}$$

for some matrix $W(x)$.

2.5 Isolated global optimality

Theorem. Define for $x \in \mathbb{R}^n$,

$$0 \leqslant t \leqslant 1 \quad \text{and} \quad 0 \leqslant s \leqslant 1 \tag{9a}$$

$$y(t) = \mathbf{x}(1 - t) + xt \tag{9b}$$

$$\mathbf{y}(s, t) = \mathbf{x}(1 - s) + [\mathbf{x}(1 - t) + xt]s. \tag{9c}$$

Let $f(x)$ and $g_j(x)$, $j = 1, \dots q$, where the first β equalities and the remaining $q - \beta$ are the equality constraints and the active inequality constraints, respectively, have two continuous derivatives, and let x^* be a feasible point satisfying both the K–K–T necessary conditions and the second-order sufficiency conditions. Assume further that

(i) $\displaystyle\int_0^1 C[y(t)]'\,\mathrm{d}t$ has rank q.

(ii) $$v' \int_0^1 \int_0^1 \nabla^2_{xx} \mathscr{L}[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}s\,\mathrm{d}tv > 0$$

for all $v \neq 0$, where

$$\int_0^1 C[y(t)]'\,\mathrm{d}tv = 0,$$

then

$$\phi' = \begin{bmatrix} \int_0^1 \int_0^1 \nabla^2_x \mathscr{L}[\mathbf{y}(s,t),\boldsymbol{\gamma},\boldsymbol{\mu}]\,\mathrm{d}s\,\mathrm{d}t & -\int_0^1 C[y(t)]''\mathrm{d}t \\ -\int_0^1 C[y(t)]''\mathrm{d}t & 0 \end{bmatrix} \tag{10}$$

has an inverse. Denote this by

$$\phi'^{-1} = \begin{bmatrix} A & B' \\ B & D \end{bmatrix}, \tag{11}$$

where A is $n \times n$ and B' is $n \times q$.

(iii) The quantity

$$-\int_0^1 f[y(t)]\,\mathrm{d}tB' > 0 \tag{12}$$

component by component.

Then $f(x) > f(\mathbf{x})$ for any feasible x. If (i)–(iii) hold for all $x \in \mathbb{R}^n$, then $\mathbf{x}$ is the unique global minimum of f/g.

Proof. Consider the following [recall $\nabla\mathscr{L}(\mathbf{x},\gamma,\boldsymbol{\mu}) = 0$, $C(\mathbf{x}) = 0$]:

$$\begin{bmatrix} \int_0^1 \nabla_x \mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t \\ -C(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \int_0^1 \int_0^1 \nabla^2_x \mathscr{L}[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}s\,\mathrm{d}t & -\int_0^1 C[y(t)]''\,\mathrm{d}t \\ -\int_0^1 C[y(t)]''\,\mathrm{d}t & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\gamma \\ \Delta\mu \end{bmatrix},$$

where

$$\Delta x = (x - \mathbf{x}),\ \Delta\gamma = (\gamma - \gamma) = \Delta\boldsymbol{\mu} = (\boldsymbol{\mu} - \boldsymbol{\mu}) = 0.$$

The existence of the above inverse is guaranteed by the theorem hypothesis. It can be written explicitly after the following notational considerations are given. Let N be any $n \times (n-q)$ matrix giving the null space of $P = \int_0^1 C[y(t)]''\,\mathrm{d}t$. Let

$$H = \int_0^1 \int_0^1 \nabla^2_x \mathscr{L}[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}s\,\mathrm{d}t.$$

The explicit inverse, which is invariant to the choice of the particular null space matrix N and generalized-inverse $P_{\#}$ of P is then

$$\begin{bmatrix} H & -P' \\ -P & 0 \end{bmatrix}^{-1} = \begin{bmatrix} N(N'HN)^{-1}N' & -[I - N(N'HN)^{-1}N'H]P_{\#} \\ -P'_{\#}[I - HN(N'HN)^{-1}N'] & -P'_{\#}[H - HN(N'HN)^{-1}N'H]P_{\#} \end{bmatrix}.$$

The difference between the objective function values at x and $\mathbf{x}$ may be found by integrating the gradients of the objective function along the incremental path Δx.

$$\begin{aligned} f(x) - f(\mathbf{x}) &= \int_0^1 f[y(t)]'\,\mathrm{d}t\,\Delta x \\ &= \int_0^1 f[y(t)]'\,\mathrm{d}tN[y(t)]H[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]^{-1} \\ &\quad \times N[y(t)]' \int_0^1 \nabla_x \mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t \\ &\quad - \int_0^1 f[y(t)]'\,\mathrm{d}tP'C(x). \end{aligned}$$

Now

$$\int_0^1 C[y(t)]'\,\mathrm{d}tN[y(t)] = 0,$$

so that

$$\begin{aligned} \int_0^1 f[y(t)'\,\mathrm{d}tN[y(t)] &= \int_0^1 \{f[y(t)]' - [\gamma'\mu']C[y(t)]'\} \\ &\quad \times \mathrm{d}tN[y(t)] \\ &= \int_0^1 \nabla_x \mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}tN[y(t)]. \end{aligned}$$

Finally,

$$\begin{aligned} f(x) - f(\mathbf{x}) &= \int_0^1 \nabla_x \mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t'N[y(t)] \\ &\quad \times H[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]^{-1}N[y(t)]' \\ &\quad \times \int_0^1 \nabla_x \mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t \\ &\quad - \int_0^1 f[y(t)]'\,\mathrm{d}tP'C(x). \end{aligned}$$

As $x \in R^n$, $C(x) \geqslant 0$. Thus both terms on the right-hand side above are non-negative. If

$$\int_0^1 \nabla_x \mathscr{L}[y(t), \gamma, \boldsymbol{\mu}]\, \mathrm{d}t' N \neq 0,$$

the first term above is positive. The second term is positive if $C_j(x) > 0$ for one of $j = 1, \ldots q$. The only way both terms can be equal to zero is if

$$\int_0^1 \nabla_x \mathscr{L}[y(t), \gamma, \boldsymbol{\mu}]\, \mathrm{d}t' N = 0$$

and $C(x) = 0$. If both these conditions hold, it is easy to show that $\Delta x = x - \mathbf{x} = 0$. Thus $f(x) > f(\mathbf{x})$ when $x \neq \mathbf{x}$. Q.E.D.

3. MINIMUM VOLUME TRUSS DESIGN

Under the assumption that the cost of each section is proportional to its area, it is desirable to find the minimum volume of an elastic truss,

$$\min l^T a, \tag{13}$$

subject to bounds on nodal displacements, allowable stresses, member areas and buckling constraints. The topology of the truss is assumed to be given and this is achieved by imposing lower bounds on the member areas. It is also assumed that there will be no geometrical effects to be taken into account. The equilibrium equations in the nodal-stiffness format involve m design variables (cross-sections) and β state variables (nodal displacements):

$$Kd = \lambda, \tag{14}$$

where λ is the β vector of the nodal loads (β is the degree of kinematic freedom) and d is the vector of nodal displacements. K is a square matrix function linearly dependent on the design variables, derived from

$$K = AkA', \tag{15}$$

where A is the direction cosine $\beta \times m$ matrix of rank β and k is an m-diagonal matrix whose elements are member stiffnesses.

The equilibrium eqns (14) can also be written

$$An = \lambda, \tag{16}$$

where n is the m vector of member forces and

$$n = Sa, \tag{17}$$

where S is a diagonal matrix whose elements are the member stresses.

The bounds on member areas, nodal displacements and member stresses can be written, respectively,

$$a_l \leqslant a \leqslant a_u \Rightarrow a \geqslant a_l \tag{18a}$$

$$-a \geqslant -a_u \tag{18b}$$

$$d_l \leqslant d \leqslant d_u \Rightarrow d \geqslant d_l \tag{19a}$$

$$-d \geqslant -d_u \tag{19b}$$

$$s_l \leqslant s = LA'd \leqslant s_u \Rightarrow LA'd \geqslant s_l \tag{20a}$$

$$-LA'd \geqslant s_u. \tag{20b}$$

L is an m-diagonal matrix whose elements are the Young's modulus divided by the member lengths.

As the equality constraints are nonlinear, this constrained optimization is nonlinear or a mathematical program. The least volume design of elastic trusses has $n = m + \beta$ variables $x' = [a'd']$ and it is intended to minimize eqn (13), subject to the equilibrium constraints (14) and the linear inequalities (18)–(20) that can be seen as an instance of eqns (1b) and (1c), respectively.

3.1 Regularity assumption

The objective function (13) and the constraints (18)–(20) are linear and therefore differentiable. The only nonlinearity stems from the equilibrium eqns (14), which are also differentiable functions. The local optimum that is going to be tested for isolated global optimality is obtained by using convex programming techniques. Therefore it satisfies the regularity assumption, the K–K–T necessary conditions and second-order conditions for isolated local optimality. In the truss optimization problem only the equilibrium constraints are nonlinear, and are therefore liable to cause nonconvex behaviour To test the assumptions of the theorem given in Sec. 2.5 it is necessary to work out the properties of the points that belong to the domain of the truss sizing problem.

By operating in the curved manifold defined by the equality constraints (14), the active constraint vector $C(a, d)$ has β elements. To form its derivative matrix $C'(a, d)$, it is necessary to evaluate

$$C(a, d)' = [\nabla_a C(a, d) \nabla_d C(a, d)], \tag{21}$$

assuming that the columns of $C'(a, d)$ are numbered so that the first m columns would correspond to derivatives with respect to the design variables a and the remaining β are associated with the derivatives with respect to the state variables d

$$C(a, d)' = [AS\ K]. \tag{22}$$

Premultiplying eqn (22) by K^{-1} yields

$$K^{-1}C(a,d)' = [K^{-1}AS\ I_\beta]. \tag{23}$$

Therefore $C(a, d)'$ has rank β for any point (a, d) satisfying the equilibrium equations, and any such point is a regular point. The rank β of $C(a, d)'$ could also be reached from the matrix produce AS, because A has rank β and S is an $m \times m$ diagonal matrix.

A matrix function giving the null space of the $\beta \times (m + \beta)$ matrix $C'(a, d)$ is a matrix $N(a, d)$ such that

$$C'(a,d)N(a,d) = 0_{\beta \times m}, \tag{24}$$

or

$$N(a,d) = \begin{bmatrix} I_m \\ -K^{-1}AS \end{bmatrix}. \tag{25}$$

$N(a, d)$ is a choice of all solutions satisfying these homogeneous equations and has rank m.

By including an active inequality corresponding to the lower area bound i, $C'(a, d)$ would be increased by one row from the matrix previously defined. In this row the element of the column i will be unity and all other elements will be zero. The rank of $C'(a, d)$ is increased by one, as this new row is linearly independent of the rows of the basis defined by the K matrix:

$$C(a,d)' = \begin{bmatrix} AS & K \\ 0\ldots 0\,1\,0\ldots 0 & 0\ldots 0 \end{bmatrix}. \tag{26}$$

The matrix that gives its null space can be derived from $N(a, d)$ by suppressing the ith column:

$$C(a,d)'N(a,d) = 0, \tag{27}$$

or

$$\begin{bmatrix} A^iS^i & K \\ 0\ldots 0\,1\,0\ldots 0 & 0\ldots 0 \end{bmatrix} \times \begin{bmatrix} I_{i-1} & 0_{i-1\times n-i} \\ 0 & \ldots\ 0 \\ 0_{n-i\times i-1} & I_{m-i} \\ \multicolumn{2}{c}{K^{-1}A^iS^i} \end{bmatrix} = 0_{\beta+1,m}, \tag{28}$$

where A^i and S^i are obtained by deleting column i in A and row i and column i in S.

A similar argument is valid when the bounds on the stresses and/or nodal displacements are active constraints. The new rows would be linearly independent of the matrix AS that has rank β and the rank of $C(a, d)'$ would then become β plus the number of independent active stress/displacements. It is usually considered in the design of a truss that this structure is able to sustain alternative loading conditions. Each loading condition is associated with β equilibrium relations, giving a set of nodal displacements d by fixing the design variables a. If the alternative loading conditions are l, the number of variables for the least volume design becomes $n = m + l \times \beta$ and the number of bilinear equilibrium equations is $l \times \beta$. Suppose there are two alternative loading conditions. In the curved manifold defined by the equilibrium equations, the active constraint derivative matrix becomes

$$C(a,d)' = \begin{bmatrix} AS^{(1)} & K & 0 \\ AS^{(2)} & 0 & K \end{bmatrix}, \tag{29}$$

where $S^{(1)}$ and $S^{(2)}$ are diagonal matrices whose elements are the member stresses corresponding to loading conditions (1) and (2), respectively, and its rank is $2 \times \beta$.

Another requirement sometimes imposed concerns variable linking. This is done to reduce the number of cross-sections with different sizes. The active constraint derivative matrix preserves its rank (unless the number of different design variables $m < \beta$) although the number of columns is reduced, because the columns corresponding to the same design variable in the matrix product AS are added.

3.2 Second-order sufficiency conditions of the reduced Hessian matrix

The only nonlinearity stems from the equilibrium eqns (14), which are also differentiable functions. It is therefore necessary to understand the influence of these equations on the behaviour of the mathematical program to verify the applicability of the theory.

The local optimum that is going to be tested for isolated global optimality is obtained by using convex programming techniques, satisfying the second-order conditions for isolated local optimality. It remains to be seen if the points lying on the equality constrained subspace also satisfy these conditions. The $(m + \beta) \times (m + \beta)$ second-derivative matrix of the Lagrangian is a copositive matrix function of the values assumed by the Lagrange multipliers:

$$\nabla^2_{a,d}\mathscr{L}[(a,d),\gamma,\mu] = \begin{bmatrix} 0 & -PA^t \\ -AP & 0 \end{bmatrix}, \tag{30}$$

where P is an $m \times m$ diagonal matrix whose elements p_i are a linear combination of the β Lagrange multipliers corresponding to the bilinear equilibrium equations γ and is independent of the multipliers related to the remaining linear constraints of the problem $\boldsymbol{\mu}$:

$$P = LA^t\gamma. \tag{31}$$

It is well known that a copositive matrix has a set of symmetric eigenvalues. It is required that the matrix $H[(a,d),\gamma,\boldsymbol{\mu}]$ should be positive definite, to satisfy assumption (ii) of the theorem for isolated global optimality. Therefore it is necessary to establish whether the points lying on the equality constrained manifold make $H[(a,d),\gamma,\boldsymbol{\mu}]$ positive definite:

$$H[(a,d),\gamma,\boldsymbol{\mu}]$$

$$= N(a,d)^t\,\nabla^2_{a,d}\mathscr{L}[(a,d),\gamma,\boldsymbol{\mu}]N(a,d) \qquad (32a)$$

$$= [I - SA^tK^{-1}]\begin{bmatrix} 0 & -PA^t \\ -AP & 0 \end{bmatrix} \times \begin{bmatrix} I \\ -K^{-1}AS \end{bmatrix} \qquad (32b)$$

$$= SA^tK^{-1}AP + PA^tK^{-1}AS. \qquad (32c)$$

$H[(a,d),\gamma,\mu]$ is a symmetric matrix given by the sum of two matrices, where the second is the transpose of the first. The stiffness matrix K is a $\beta \times \beta$ symmetric positive definite matrix of rank β. Thus the same holds true for K^{-1}.

At a local minimum $[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]$ it is easy to check the second-order sufficiency conditions and the positive definitiveness of H.

By enforcing the K–K–T necessary conditions for local optimality related to the design variables, a relationship is obtained, linking the elements of the diagonal matrices S and P:

$$\nabla_a\mathscr{L}[(a,d),\gamma,\mu] = 0.$$

Thus,

$$l - \mu_a^l + \mu_a^u - SA^T\gamma = 0$$

$$\Rightarrow SA^T\gamma = l - \mu_a^l + \mu_a^u, \qquad (33)$$

where l is the m-vector of member lengths and $\mu_a^l, \mu_a^u \geqslant 0$ are the Lagrange multipliers corresponding to the constraints associated with the lower and upper bounds on the design variables.

From eqns (31) and (33) we have

$$P_i p_i = l - \mu_{ia}^l/l_i + \mu_{ia}^u/l_i. \qquad (34)$$

Suppose any $l - \mu_{ia}^l/l_i < 0$, i.e. at least one of the Lagrange multipliers corresponding to active lower bounds on the area exceeds the member length. A necessary condition for the positive definitiveness of a real symmetric matrix states that all diagonal elements must be positive. As h_{ii} is negative, this condition is not met. In this case, the manifold defined by the equilibrium constraints makes H indefinite. It is therefore necessary to investigate the effect caused by the insertion of any active area bound inequality in the active constraint vector $C(a,d)$. Using $N(a,d)$ given by (28), $H[(a,d),\gamma,\mu]$ becomes an $(m-1)\times(m-1)$ matrix given by

$$H[(a,d),\gamma,\mu] = S^iA^{it}K^{-1}A^iP^i + P^iA^{it}K^{-1}A^iS^i, \qquad (35)$$

where P^i is an $(m-1)\times(m-1)$ diagonal matrix derived from P by suppressing its row i and column i. In H the row and column corresponding to the index of the active area constraint vanish. Therefore the negative eigenvalue due to the negative diagonal element should no longer be considered. If more than one lower bound on the areas is active, they can be equally inserted in the active constraint derivative matrix. The final matrix H is obtained from the matrix derived in the equality constraint manifold by deleting all the rows and columns corresponding to indices of active areas.

Assuming as in Sec. 2.5 a fixed set of Lagrange multipliers, the matrices P and S would remain unchanged, i.e. if the stresses in matrix S are fixed and the design variables are allowed to vary, the matrix H is positive definite because it results from a similarity transformation on K^{-1}.

3.3 Second-order estimate of the multipliers

From the proof of the theorem given in Sec. 2.5 the difference $f(x) - f(\mathbf{x})$ considering the paths defined by $y(t)$ and $\mathbf{y}(s,t)$ is made up of two contributions:

$$f(x) - f(\mathbf{x}) = \int_0^1 \nabla_x\mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t^t N[y(t)] \times H[\mathbf{y}(s,t),\gamma,\boldsymbol{\mu}]^{-1}N[y(t)]^t \times \int_0^1 \nabla_x\mathscr{L}[y(t),\gamma,\boldsymbol{\mu}]\,\mathrm{d}t - \int_0^1 f[y(t)]^t_x\,\mathrm{d}t V^R X(z). \qquad (36)$$

The first contribution is a step from x to x_0 which drives the Lagrangian derivative to 0. The second contribution, from x_0 to $\mathbf{x}$, seeks to satisfy the active constraints violated in the first step. In terms of the inverse function approach, x_0 consists of the first n components of $\phi[0^Tg(x)^T]^{-1t}$. The exact difference can be computed by integrating from $[x^t\gamma^t\mu^t]$ to $[x_0^t\gamma^t\mu^t]$ along $\phi^{-1}[t\nabla_x\mathscr{L}(x,\gamma,\mu)^tg(x)^t]^t$. The other term is obtained by integrating from $[x_0^t\gamma^t\mu^t]$ to $[\mathbf{x}^t\boldsymbol{\gamma}^t\boldsymbol{\mu}^t]$ along the curve $\phi[0^tg(x)^t]^{-1t}$.

There are two estimates of the generalized Lagrange multipliers associated with a local minimum. The first is

$$[\gamma' \mu'] = f(\mathbf{x})'^{\prime} C(\mathbf{x})_{\#}. \quad (37)$$

This is used in the computation of the Lagrange Hessian matrix. The second estimate is

$$f(\mathbf{x})'^{\prime}\{I - N(\mathbf{x})[N(\mathbf{x})' H(\mathbf{x}, \gamma, \boldsymbol{\mu}) N(\mathbf{x})]$$
$$- 1 N(\mathbf{x})' \nabla_x^2 \mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu})\} C(\mathbf{x})_{\#}. \quad (38)$$

This estimate determines whether or not a constraint is to be considered binding at the local solution. The importance of this second-order estimate was first pointed out in [6]. It is really the estimate of the multipliers (38) at the minimum of f in the subspace $C(x) = C(\mathbf{x})$.

This result is used here in a different context to prove the theorem for isolated global optimality. Its third assumption requires that the quantity given by the product

$$\int f'[y(t)] B' \,\mathrm{d}t \quad (39)$$

should be positive term by term. This results in a q vector given by the product of the constant vector

$$\int f'[y(t)] \,\mathrm{d}t = [l' 0'] \quad (40)$$

times the matrix

$$-[I - S(\mathbf{x}) H(\mathbf{x}, \gamma, \boldsymbol{\mu})^{-1}$$
$$\times S(\mathbf{x})' \nabla_x^2 \mathscr{L}(\mathbf{x}, \gamma, \boldsymbol{\mu})] C'(\mathbf{x})_{\#}. \quad (41)$$

We note that the first β components can always be taken to be positive, as they are related to equality constraints. Therefore the test corresponding to the third assumption of the theorem for isolated global optimality must be performed only on the active lower area bounds.

4. APPLICATIONS

The following examples are referred to in the literature as a basis for comparison of the efficiencies of several algorithms. They will be used here to make statements about the validity of the solutions found by a convex programming technique.

4.1 Three-bar truss subject to a single loading condition

Using the nodal-stiffness description, the minimum volume design of the three-bar truss represented in

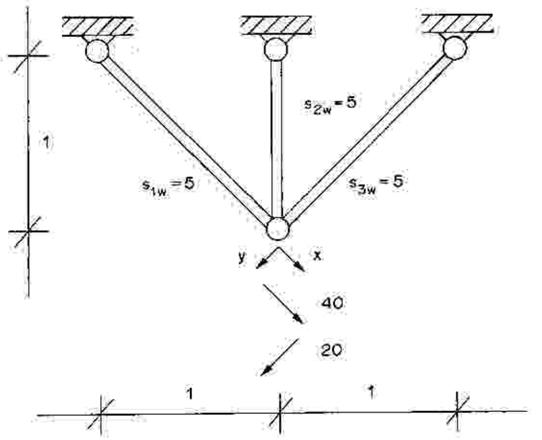

Fig. 1.

Fig. 1 undergoing two alternative loading conditions, subject to stress constraints and allowing continuous variation of the area variables, can be stated as follows:

$$\min\ (2)^{1/2} a_1 + a_2 + (2)^{1/2} a_3 \quad (42a)$$

$$\text{such that } (2)^{1/2}/2 a_1 d_1 + 1/2 a_2 (d_1 + d_2) = 40 \quad (42b)$$

$$1/2 a_2 (d_1 + d_2) + (2)^{1/2}/2 a_3 d_2 = 20 \quad (42c)$$

$$0 \leqslant s_1 = (2)^{1/2}/2 d_1 \leqslant 5 \quad (42d)$$

$$0 \leqslant s_2 = (2)^{1/2}/2 (d_1 + d_2) \leqslant 5 \quad (42e)$$

$$-5 \leqslant s_3 = (2)^{1/2}/2 d_2 \leqslant 0 \quad (42f)$$

$$a_1^l = 0.1 \leqslant a_1 \leqslant a_1^u = 11; \quad 0.1 \leqslant a_2 \leqslant 4;$$
$$0.1 \leqslant a_3 \leqslant 5. \quad (42g)$$

Optimum solution:

a_1	a_2	a_3	s_1	s_2	s_3	d_1	d_2
4.0	5.657	0.1	5.0	5.0	0.	7.071	0

The Lagrange multiplier corresponding to the lower bound on a_3, $\mu_{3a}{}^l$ is $(2)^{1/2}$.

If all active constraints were to be considered, the rank of the active constraint derivative matrix would be n and the matrix giving the null space of $C(a, d)'$ does not exist. The second-order estimate of the Lagrange multipliers would coincide with the first-order estimate and the third assumption in the theorem for global optimality is meaningless.

As the nonlinearity results from the equilibrium equations, it is necessary to understand the influence of the latter in the behaviour of this mathematical program. If only the equilibrium constraints are included in the active set, the reduced Hessian matrix H becomes singular because the μ_{3a}^l equals the corresponding member length and the inverse of H does not exist. Again, the third assumption of the theorem cannot be checked.

It is therefore necessary to include the lower area bound corresponding to a_3 in the active set. H becomes a 2×2 positive definite matrix. It is only necessary to verify the third component of the vector given by the product

$$-[I' \, 0']\{I - S(\mathbf{a},\mathbf{d})H(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]^{-1} \times S(\mathbf{a},\mathbf{d})'\nabla_x^2\mathscr{L}[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]\} C'(\mathbf{a},\mathbf{d})_{\#} \quad (43)$$

because the first two components are related to equality constraints and they can always be made positive. The minimum found is unique because the third component is positive.

4.2 *Three-bar truss subject to two alternative loading conditions*

The least-volume solution for the truss represented in Fig. 1 subjected to two alternative loading conditions, lower limits on the cross-sectional areas (0.1) and the same stress limits in all bars (± 5.0) is

a_1	a_2	a_3	s_1^1	s_2^2	s_3^1	s_1^2	s_2^2	s_3^3	d_1^1	d_2^1	d_1^2	d_2^2
7.0	2.1	2.8	5.0	3.2	−1.8	−0.9	4.9	5.0	7.1	−2.5	−1.2	7.1

Here the lower bounds on the member areas are not active. As, for multiple loading conditions, the nature of the matrix H remains unaltered,

$$H[(a,d),\gamma,\boldsymbol{\mu}] = \sum_{k=1,2} [S^{(k)}A^TK^{-1}AP^{(k)} + P^{(k)}A^Tk^{-1}AS^{(k)}], \quad (44)$$

we may infer that the optimum found is unique.

4.3 *Ten-bar truss*

The ten-bar truss represented in Fig. 2 is subjected to a single loading condition. If the member stresses in all bars are limited to ± 2.5 and the lower bounds on all cross-sectional areas are 0.1, the problem yields an optimal solution of 44.2, corresponding to

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
7.9	0.1	8.1	3.9	0.1	0.1	5.8	5.5	3.7	0.1
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
2.5	2.5	−2.5	−2.5	0.0	2.5	2.5	−2.5	3.7	−2.5
d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8		
5.0	−20.0	−5.0	−22.5	2.5	−7.5	−2.5	−7.5		

The Lagrange multipliers corresponding to the lower bounds on the variables are

$$\mu_{2a}^l = 0.6, \quad \mu_{5a}^l = 1.0, \quad \mu_{6a}^l = 0.6, \quad \mu_{10a}^l = 0.2.$$

As the Lagrange multipliers corresponding to the lower bounds on the areas do not exceed the corresponding member length, this can be regarded as a special case of positive semidefinitiveness in $H[(a,d),\gamma,\mu]$ and the locus of points satisfying the equilibrium equations define a convex domain. Thus the minimum found is unique. The active lower bounds could also be incorporated in the active constraint matrix $C(a,d)$, satisfying the third assumption of the theorem for isolated global optimality.

4.4 *Ten-bar truss with one member stronger*

It is possible to have an indefinite matrix H in the subspace defined by the equilibrium equations and the local optimum to be unique in the problem domain. When the stress limit on bar 9 is increased to 3.75 an optimal solution of 41.6 is found:

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
7.9	0.1	8.1	3.9	0.1	0.1	5.7	5.6	5.6	0.1
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
2.5	1.6	−2.5	−2.5	0.0	1.6	2.5	−2.5	2.5	−2.2
d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8		
4.0	−18.5	−5.0	−20.0	2.5	−7.5	−2.5	−7.5		

The Lagrange multipliers corresponding to the lower bounds on the variables are

$$\mu_{2a}^l = 0.0, \quad \mu_{5a}^l = 1.0, \quad \mu_{6a}^l = 1.7.$$

The matrix H in the subspace defined by the equilibrium relations would have one negative eigenvalue corresponding to a negative diagonal element in the second row of this real symmetric matrix. To check whether this minimum is unique it is necessary to include the active constraints related to the lower bounds on a_2, a_3 and a_6. The rank of the active constraint derivative matrix $C(a,d)'$ is increased by three. To verify the third assumption of the theorem for global optimality, it is necessary to verify the three

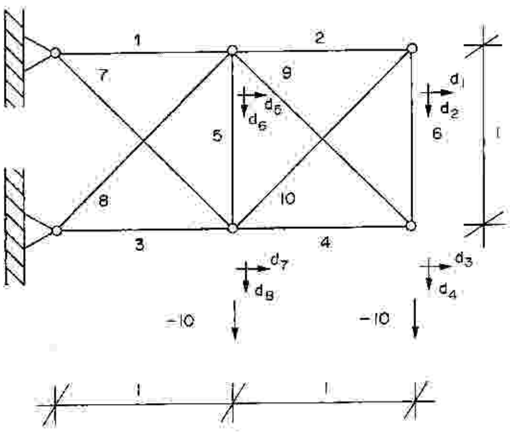

Fig. 2.

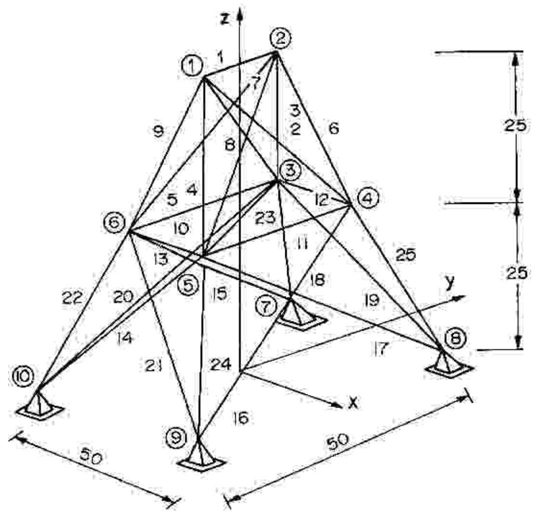

Fig. 3.

components corresponding to a_2^l, a_3^l and a_6^l of the vector given by the product (43),

$$-[I'0']\{I - S(\mathbf{a},\mathbf{d})H[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]^{-1} \times S(\mathbf{a},\mathbf{d})'\nabla_x^2\mathscr{L}[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]\}C'(\mathbf{a},\mathbf{d})_{\#},$$

because the first eight components are related to equality constraints and can always be made positive. The minimum found is unique because all components are positive.

4.5 Ten-bar truss with maximum displacements prescribed

If the absolute value of the displacement is limited to ± 3.5 fully nonconvex behaviour would occur. Two solutions arise corresponding to the *OF* values 219.9 and 223.3:

It is necessary to verify the corresponding components of the vector given by the product (43),

$$-[I'0']\{I - S(\mathbf{a},\mathbf{d})H[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]^{-1} \times S(\mathbf{a},\mathbf{d})'\nabla_x^2\mathscr{L}[(\mathbf{a},\mathbf{d}),\gamma,\boldsymbol{\mu}]\}C'(\mathbf{a},\mathbf{d})_{\#}.$$

The first eight components are related to equality constraints and they can always be made positive. The components related to a_2, a_5 and a_{10}, respectively, are negative in the two cases presented. This appears to be a necessary condition for multiple optimality.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
48.7	0.1	35.6	24.1	0.1	1.2	9.4	34.3	34.1	0.1
s_1	a_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
0.4	−0.3	−0.6	−0.4	2.2	−0.0	2.9	−0.8	0.8	0.7
d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8		
0.2	−3.5	−1.0	−3.5	0.4	−1.3	−0.6	−3.5		

The Lagrange multipliers corresponding to lower bounds on the areas for the lower local solution are

$$\mu_{2a}^l = 4.9, \quad \mu_{5a}^l = 9.9, \quad \mu_{10a}^l = 12.5.$$

The second local optimum is associated with

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
48.7	0.1	38.1	23.3	0.1	0.1	13.7	33.1	33.0	0.1
s_1	a_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
0.4	0.0	−0.5	−0.4	1.3	0.0	2.0	−0.9	0.9	0.0
d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8		
0.4	−3.5	−1.0	−3.5	0.4	−1.3	−0.5	−2.6		

$$v_{2a}^l = 1.0, \quad \mu_{5a}^l = 5.1, \quad \mu_{6a}^l = 1.0, \quad \mu_{10a}^l = 1.4.$$

4.6 Twenty-five-bar tower

A tower, whose configuration is shown in Fig. 3, is subjected to two alternative loading conditions. By imposing the four-fold symmetry on the structure explicitly, the 25 members may have their cross-sectional areas linked to eight design variables. The structure is designed against stiffness requirements in that the nodal displacements are not permitted to exceed 0.35 and the topology is fixed by imposing a minimum limit on the cross-sectional areas of 0.01.

The test for isolated global optimality is successful and we may infer that the optimum obtained in this case is also the unique optimum.

Load condition	Node	Direction of load x	y	z
1	1	1.0	10.0	−5.0
	2	0.0	10.0	−5.0
	3	0.5	0.0	0.0
	4	0.5	0.0	0.0
2	1	0.0	10.0	−5.0
	2	−1.0	10.0	−5.0
	4	−0.5	0.0	0.0
	5	−0.5	0.0	0.0

Member areas for the optimum design are

1	2,3,4,5	6,7,8,9	10,11	12,13	14,15,16,17	18,19,20,21	22,23,24,25
0.01	2.043	3.003	0.01	0.01	0.683	1.623	2.672

5. CONCLUSIONS

There is no simple rule for classifying structures as convex or nonconvex with respect to optimization: they may be of either type depending on the numerical constants involved. The boundaries of the feasible domain are formed by the stress and displacement constraints. In statically determinate structures these would be hyperplanes in an appropriately chosen set of co-ordinates, whereas in indeterminate problems the boundaries become slightly curved, and the curvatures may be of the nonconvex type. This has been found in a ten-bar truss with prescribed maximum displacements. A condition is given here, which, if satisfied, guarantees that a local optimum is the global optimum for the optimal sizing of trusses. It has been shown that if no lower bounds on the areas are active constraints of an isolated local minimum given by convex programming, this point is an isolated global minimum for the least-volume design of trusses. If the constraints associated with lower bounds are active a test consisting of obtaining second-order estimates of the Lagrange multipliers must be performed. If each element of the resulting vector is non-negative, then the local minimum is also the isolated global solution. If this is not the case, the problem may possess multiple optima.

REFERENCES

1. V. B. Venkayya, Structure optimization: a review and some recommendations. *Int. J. Numer. Meth. Engng* **13,** 203–228 (1978).
2. L. M. C. Simões, Search for the global optimum in structural optimization problems. *Engng Opt.* **11,** 49–67 (1987).
3. G. P. McCormick, Locating an isolated global minimizer of a constrained nonconvex program. *Math. Oper. Res.* **5,** 435–443 (1980).
4. W. I. Zangwill, *Nonlinear Programming: A Unified Approach.* Prentice–Hall, Englewood Cliffs, NJ (1969).
5. C. Goffman, *Calculus of Several Variables.* Harper & Row, New York (1965).
6. P. E. Gill and W. Murray, The computation of Lagrange multiplier estimates for constrained minimization. NPL NAC Rep. 77 (1977).