Computer Assisted Mechanics and Engineering Sciences, 4: 283–293, 1997.

Comparison of two staggered schemes for optimization with a critical point constraint

Krzysztof Wiśniewski, Ewa Turska and Michał Kleiber
Institute of Fundamental Technological Research, Polish Academy of Sciences
Świętokrzyska 21, 00-049 Warsaw, Poland

(Received May 12, 1997)

Two staggered solution schemes for a minimum volume optimization problem with a critical point constraint are considered. Scheme 1 leads to optimization at a critical (maximum) point, while Scheme 2 results in optimization at a maximum load. The reduced optimization problems for each of the schemes are different, and the derivatives for them must be defined consistently with the step preceding optimization. Using an example of a simple nonlinear two-bar truss and performing a rigorous analysis of roots of the equilibrium equation and of their limits, we show that properties of the derivative of displacements at the critical load and the derivative of critical displacements are very different.

Then the methods of calculating various design derivatives are described and both solution schemes are tested on the truss example. Conclusions are related to accuracy and rate of convergence of both schemes, as well as to their sensitivity to inaccuracies characteristic for large scale numerical implementations.

1. INTRODUCTION

Complexity of equilibrium paths for nonlinear equilibrium equations, which may possess extremum and bifurcation points, makes optimization a challenging task. Essential for designing is the question of the stability of the solution, and it is usually formulated as a requirement that the minimum should be located either on a stable branch, e.g. [11], or, at most, at the first critical point, e.g. [4, 5, 6, 16]. The latter requirement is of interest in the present paper, and it is used for a minimum volume problem.

The optimization problem is solved using staggered (bordering) solution schemes, which proved to be useful in many contexts, e.g. [13]. A variety of staggered schemes can be devised, and these which are used in the present paper may be characterized as follows:

- the problem is split, and solved in several steps, with solution of a reduced optimization problem being only one of them,
- the reduced optimization problem is formulated in such a way that the state variable is always excluded from the minimization,
- steps are repeated until convergence is reached.

Different staggered schemes result in different minimization tasks actually being solved, and we present and discuss the following schemes:

1. a critical point is found, then, a reduced optimization problem is solved at this point. The idea of this scheme is given in [4] for the maximum critical load problem, and we extend it to the minimum volume problem.
2. a solution of the equilibrium equation for the maximum load is found, then, a reduced optimization problem is solved at this load level. This scheme has not yet been described in the literature.

These two schemes require different derivatives of the objective function and constraints, which must be defined consistently with the algorithmic steps actually being performed. Therefore, if we consider displacements, a clear distinction between the derivative of displacements at the critical load and the derivative of critical displacements must be made. Using the example of a simple nonlinear two-bar truss we show that their properties are astonishingly different. Both derivatives are obtained by performing a rigorous analysis of roots of the equilibrium equation and of properties of their limits. It must be stressed that only the derivatives which are consistent with the algorithmic setup yield correct solution of the optimization problem, and therefore they cannot be mistaken.

Subsequently, methods of calculating various design derivatives characteristic for the schemes of interest are described. Scheme 1 requires a design derivative of the critical load, while Scheme 2 necessitates a design derivative of the determinant of the tangent matrix, and the design derivative of the displacements.

Using the example of the nonlinear truss we tested both solution schemes, and formulated conclusions related to:

- accuracy and rate of convergence of both schemes,
- sensitivity of Scheme 1 to inaccuracies in determination of the critical point,
- sensitivity of Scheme 2 to inaccuracies in solution of the equilibrium equation.

These inaccuracies are characteristic for large scale numerical implementations of both schemes.

Finally, we note in passing that calculation of the derivatives with respect to the design variables is a subject of the design sensitivity analysis, and it has already beeen performed for a wide class of materially and kinematically nonlinear problems, see the overview given e.g. in [7] and [8]. These derivatives can be used to modify the structure in order to achieve a required performance. The present paper indicates that when the design parameters are sought as a solution of an optimization problem then the question of derivatives is more complicated. They depend on the solution scheme, and, accordingly, for new solution schemes new types of the design derivatives need to be calculated.

2. DEFINITION OF SOLUTION SCHEMES

In this section we define two staggered solution schemes for the minimum volume problem with a critical point constraint. But first, we specify the equilibrium equation in the following way

$$\mathbf{r}(\mathbf{b},\mathbf{z},\mu) \equiv \mathbf{f}(\mathbf{b},\mathbf{z}) - \mu\mathbf{p} = \mathbf{0} \tag{1}$$

where $\mathbf{b}$ is the design variable (e.g. cross-sectional area, Young modulus, etc.), $\mathbf{z}$ is the state variable (e.g. displacement, rotation, etc.), μ is the load factor, and $\mathbf{p}$ is the fixed external load. The internal force is denoted by $\mathbf{f}$. The critical point is defined as a solution of the following set of equation

$$\left.\begin{aligned} \mathbf{r}(\mathbf{b},\mathbf{z},\mu) &= \mathbf{0} \\ \det \mathbf{K}(\mathbf{b},\mathbf{z}) &= 0 \end{aligned}\right\} \tag{2}$$

for a fixed $\mathbf{b}$, and will be denoted as $\{\mathbf{z}_{\text{cr}}, \mu_{\text{cr}}\}$. Here $\mathbf{K} \equiv \partial\mathbf{r}/\partial\mathbf{z}$ is the tangent operator.

Let us consider the minimum volume problem formulated as follows:

$$\min_{(\mathbf{b},\mathbf{z},\mu)} V(\mathbf{b}) \tag{3}$$

subject to:

$$\begin{aligned} &\mathbf{r}(\mathbf{b},\mathbf{z},\mu) = \mathbf{0} && \text{— equilibrium equation (nonlinear),} \\ &\det \mathbf{K}(\mathbf{b},\mathbf{z}) = 0 && \text{— determinant constraint,} \\ &g_\mu = \mu^{\max} - \mu = 0 && \text{— load constraint.} \end{aligned}$$

Here, V is a volume of the structure, and $\mu^{\max}$ denotes the prescribed maximum value of the load factor. The above problem may be solved using the following two-step staggered schemes.

Scheme 1: minimization at the critical point

Step 1. For given $\mathbf{b}$ solve

$$\left.\begin{aligned} \mathbf{r}(\mathbf{b},\mathbf{z},\mu) &= \mathbf{0} \\ \det \mathbf{K}(\mathbf{b},\mathbf{z}) &= 0 \end{aligned}\right\} \tag{4}$$

to find a critical point $\{\mathbf{z}_{\mathrm{cr}}, \mu_{\mathrm{cr}}\}$. This step implies $\mathbf{z}_{\mathrm{cr}} = \mathbf{z}_{\mathrm{cr}}(\mathbf{b})$ and $\mu_{\mathrm{cr}} = \mu_{\mathrm{cr}}(\mathbf{b})$.

Step 2. Solve

$$\min_{(\mathbf{b})} V(\mathbf{b}) \tag{5}$$

subject to $g_\mu = \mu^{\max} - \mu_{\mathrm{cr}}(\mathbf{b}) = 0$ to update $\mathbf{b}$. Note that g_μ is a nonlinear function of $\mathbf{b}$.

These steps are repeated until convergence is achieved. To provide derivatives of the objective function and constraints the design sensitivities $\mathrm{d}V/\mathrm{d}\mathbf{b}$ and $\mathrm{d}g_\mu/\mathrm{d}\mathbf{b} = -\mathrm{d}\mu_{\mathrm{cr}}/\mathrm{d}\mathbf{b}$ must be calculated at the critical point.

Scheme 2: minimization at the maximum load

Step 1. For given $\mathbf{b}$ and $\mu^{\max}$ solve

$$\mathbf{r}(\mathbf{b},\mathbf{z},\mu^{\max}) = \mathbf{0} \tag{6}$$

to find $\mathbf{z}^{eq}$. This step implies $\mathbf{z}^{eq} = \mathbf{z}(\mathbf{b},\mu^{\max})$.

Step 2. Solve

$$\min_{(\mathbf{b})} V(\mathbf{b}) \tag{7}$$

subject to $\det \mathbf{K}(\mathbf{b},\mathbf{z}^{eq}) = 0$ to update $\mathbf{b}$.

These steps are repeated until convergence is achieved. To provide derivatives of the objective function and constraints the design derivatives $\mathrm{d}V/\mathrm{d}\mathbf{b}$ and $\mathrm{d}\det \mathbf{K}/\mathrm{d}\mathbf{b}$ must be calculated at the maximum load.

The main advantage of the above defined staggered schemes consists in separating $\mathbf{b}$ and $\mathbf{z}$, and solving the minimization problem in terms of $\mathbf{b}$ only.

3. Design derivatives of displacement and critical displacement

Below we define two design derivatives related to displacements, which correspond with two staggered schemes specified in the previous section, and discuss the differences between them.

Let us define the design derivative of critical displacement, which occurs in Scheme 1, as follows

$$\frac{\mathrm{d}\mathbf{z}_{\mathrm{cr}}}{\mathrm{d}\mathbf{b}} \equiv \lim_{\Delta\mathbf{b}\to 0} \frac{\mathbf{z}_{\mathrm{cr}}(\mathbf{b}+\Delta\mathbf{b}) - \mathbf{z}_{\mathrm{cr}}(\mathbf{b})}{\Delta\mathbf{b}} \tag{8}$$

Besides, let us define the design derivative of displacement for Scheme 2 as

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{b}} \equiv \lim_{\Delta\mathbf{b}\to 0} \frac{\mathbf{z}(\mathbf{b}+\Delta\mathbf{b},\mu) - \mathbf{z}(\mathbf{b},\mu)}{\Delta\mathbf{b}} \tag{9}$$

which for $\mu = \mu^{cr}$ takes the following form

$$\left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{b}}\right)_{cr} = \lim_{\Delta\mathbf{b}\to 0} \frac{\mathbf{z}(\mathbf{b}+\Delta\mathbf{b},\mu_{cr}) - \mathbf{z}(\mathbf{b},\mu_{cr})}{\Delta\mathbf{b}} \tag{10}$$

Note that values of $\mathbf{z}(\mathbf{b},\mu_{cr})$ and $\mathbf{z}_{cr}(\mathbf{b})$ are equal but their interpretation is different, as the first one results from the equilibrium equation (1), while the second from the critical point equations (2). Therefore, properties of derivatives (8) and (10) are different, and a clear distinction between them must be made. In sequel we demonstrate for a truss that properties of the two derivatives can be surprisingly different.

Let us consider a non-symmetrical two-bar truss introduced in [5], see Fig. 1, of hight H, cross-sectional areas of the bars A_1 and A_2, spans of the bars L_1 and L_2, and a Young modulus E. The truss is loaded by force P, and the vertical displacement of the loaded node is denoted by u. Throughout this paper the following units are used: [in] for length and displacement, [in^2] for area, [in^3] for volume, [lb] for force, [psi] for modulus of elasticity.

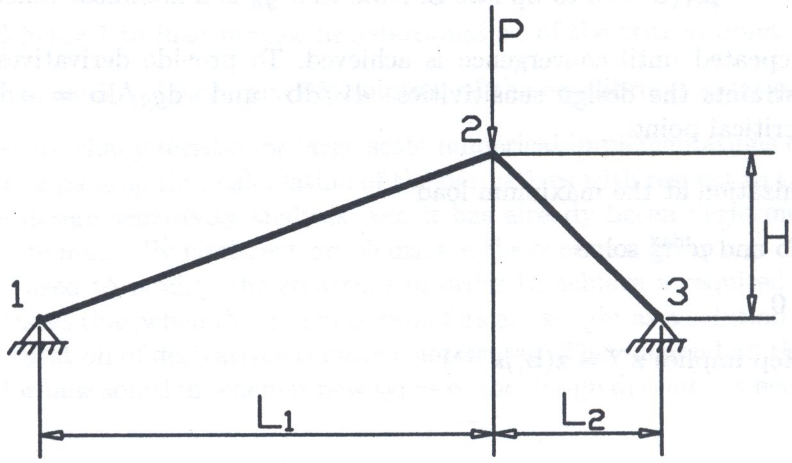

Fig. 1. Unsymmetric two-bar truss

If we use the engineering (right stretch) strain measure $\epsilon = (l - l_0)/l_0$, where l and l_0 is the actual and initial length of a bar, and assume that $(H/L_1)^2 \ll 1$ then the equilibrium equation for the whole truss can be written as follows

$$r = C\left(\tfrac{1}{2}z^3 - \tfrac{3}{2}\alpha z^2 + \alpha^2 z\right) - P = 0\,, \tag{11}$$

where

$$C = eE\gamma(\gamma+1)^2\,, \quad e = \frac{A_1 A_2}{A_1 + A_2\gamma}, \quad \gamma = \frac{L_1}{L_2}, \quad z = \frac{u}{L_1}, \quad \alpha = \frac{H}{L_1}\,. \tag{12}$$

Comparing the above equation with (1) we identify z as a state variable, A_1 as a design variable, and P stands for $\mu\mathbf{p}$.

The critical point, Eq. (2), here defined by $r = 0$ and $\mathrm{d}P/\mathrm{d}z = 0$, is as follows

$$\left.\begin{aligned} p_{cr} &= Cg\,, \quad \text{where} \quad g = \frac{\alpha^3}{3\sqrt{3}} \\ z_{cr} &= \alpha\left(1 - \frac{1}{\sqrt{3}}\right) \end{aligned}\right\} \tag{13}$$

where $p_{cr} = P_{cr}/(EA_2)$. It can be verified that in our case the critical point is a maximum (limit) point.

For the data $E = 10^7$, $L_1 = 200$, $L_2 = 50$, $H = 2.5$, $A_1 = A_2 = 0.79975$ we obtain $P_{cr} = p_{cr}EA_2 = 60.12187$ and $u_{cr} = z_{cr}L_1 = 1.05662$.

To compute the design derivatives we have to establish the dependence $z = z(A_1)$ by checking roots of equilibrium equation (11) for the perturbed design variable $A_1 + \Delta A_1$. It can be done for the critical load p_{cr} because then roots can be conveniently related to z_{cr}, and besides we are particularly interested in the design derivative of displacement at a critical point. A complete analysis of roots of the perturbed equilibrium equation and properties of their limits for $\Delta A_1 \to 0$ is presented in [15].

Derivative $(\mathrm{dz/db})_{cr}$ for truss

The design derivative of displacement at (A_1, p_{cr}) is defined according to (10) as follows

$$\left(\frac{\mathrm{d}z}{\mathrm{d}A_1}\right)_{cr} \equiv \lim_{\Delta A_1 \to 0} \frac{z(A_1 + \Delta A_1, p_{cr}) - z(A_1, p_{cr})}{\Delta A_1} \tag{14}$$

where $z(A_1, p_{cr}) = z_{cr}$, and z_{cr} is given by Eq. (13). To specify $z(A_1 + \Delta A_1, p_{cr})$ we have to choose roots from the vicinity of z_{cr}, i.e. those z^* which converge to z_{cr} for $\Delta A_1 \to 0$. Equation (11) can be given a canonical form $y^3 + 3py + 2q = 0$, with the determinant defined as $D = q^2 + p^3$, see [15] for details. Selecting the roots we reject roots no. 1 for all values of D, and roots no. 2 and 3 for $D < 0$, as they are imaginary. Hence only roots no. 2 and 3 for $D \leq 0$ ($\Delta A_1 \geq 0$) are considered, which are

$$\begin{aligned} z_2^* &= \alpha\left[1 - \frac{1}{\sqrt{3}}(\cos\tfrac{1}{3}\varphi + \sqrt{3}\sin\tfrac{1}{3}\varphi)\right], \\ z_3^* &= \alpha\left[1 - \frac{1}{\sqrt{3}}(\cos\tfrac{1}{3}\varphi - \sqrt{3}\sin\tfrac{1}{3}\varphi)\right], \end{aligned} \tag{15}$$

where $\cos\varphi = 1 - \lambda$, and $\lambda = \gamma A_2 \Delta A_1/(A_1 + A_2\gamma)(A_1 + \Delta A_1)$. For them we have

$$\frac{z_{2,3}^* - z_{cr}}{\Delta A_1} = -\frac{\alpha}{\sqrt{3}} \frac{(\cos\frac{1}{3}\sqrt{2\lambda} \pm \sqrt{3}\sin\frac{1}{3}\sqrt{2\lambda} - 1)}{\Delta A_1} \tag{16}$$

Let us split the above expression and consider limits for particular components. Firstly, $\cos\frac{1}{3}\sqrt{2\lambda} - 1 = -2\sin^2\frac{1}{6}\sqrt{2\lambda}$, and we may write

$$\frac{\sin^2\frac{1}{6}\sqrt{2\lambda}}{\Delta A_1} = \frac{\sin\frac{1}{6}\sqrt{2\lambda}}{\frac{1}{6}\sqrt{2\lambda}} \frac{\sin\frac{1}{6}\sqrt{2\lambda}}{\frac{1}{6}\sqrt{2\lambda}} \frac{\frac{1}{36}2\lambda}{\Delta A_1} \xrightarrow{\Delta A_1 \to 0} \frac{1}{18}\frac{p_1}{p_2} \tag{17}$$

because

$$\frac{\sin\frac{1}{6}\sqrt{2\lambda}}{\frac{1}{6}\sqrt{2\lambda}} \xrightarrow{\Delta A_1 \to 0} 1 \quad \text{and} \quad \frac{\lambda}{\Delta A_1} = \frac{p_1 \Delta A_1}{(p_2 + p_3\Delta A_1)\Delta A_1} \xrightarrow{\Delta A_1 \to 0} \frac{p_1}{p_2}, \tag{18}$$

where $\lambda = \dfrac{p_1 \Delta A_1}{(p_2 + p_3 \Delta A_1)}$. Hence, the limit of the above component is finite. Secondly,

$$\frac{\sin\frac{1}{3}\sqrt{2\lambda}}{\Delta A_1} = \frac{\sin\frac{1}{3}\sqrt{2\lambda}}{\frac{1}{3}\sqrt{2\lambda}} \frac{\frac{1}{3}\sqrt{2\lambda}}{\Delta A_1} \xrightarrow{\Delta A_1 \to 0} \pm\infty \tag{19}$$

because

$$\frac{\sin\frac{1}{3}\sqrt{2\lambda}}{\frac{1}{3}\sqrt{2\lambda}} \xrightarrow{\Delta A_1 \to 0} 1 \tag{20}$$

$$\frac{\sqrt{2\lambda}}{\Delta A_1} = \sqrt{\frac{2p_1\Delta A_1}{(p_2+p_3\Delta A_1)(\Delta A_1)^2}} = \sqrt{\frac{2p_1}{(p_2+p_3\Delta A_1)\Delta A_1}} \xrightarrow{\Delta A_1\to 0} \pm\infty$$

We conclude that at the critical point $(dz/dA_1)_{cr} \to -\infty$ for root no. 2, and $(dz/dA_1)_{cr} \to +\infty$ for root no. 3 as $\Delta A_1^{-1/2}$ for $\Delta A_1 \to 0$. The above conclusion may also be reached in another way shown in [15].

In Fig. 2 the solution of equilibrium equation, $r = 0$, and the design derivative of displacement dz/dA_1, are shown. The design derivative of displacement is calculated according to

$$\frac{dz}{dA_1} = -K^{-1}\frac{\partial f}{\partial A_1}, \qquad K \equiv \frac{\partial f}{\partial z} = C\frac{da}{dz} \tag{21}$$

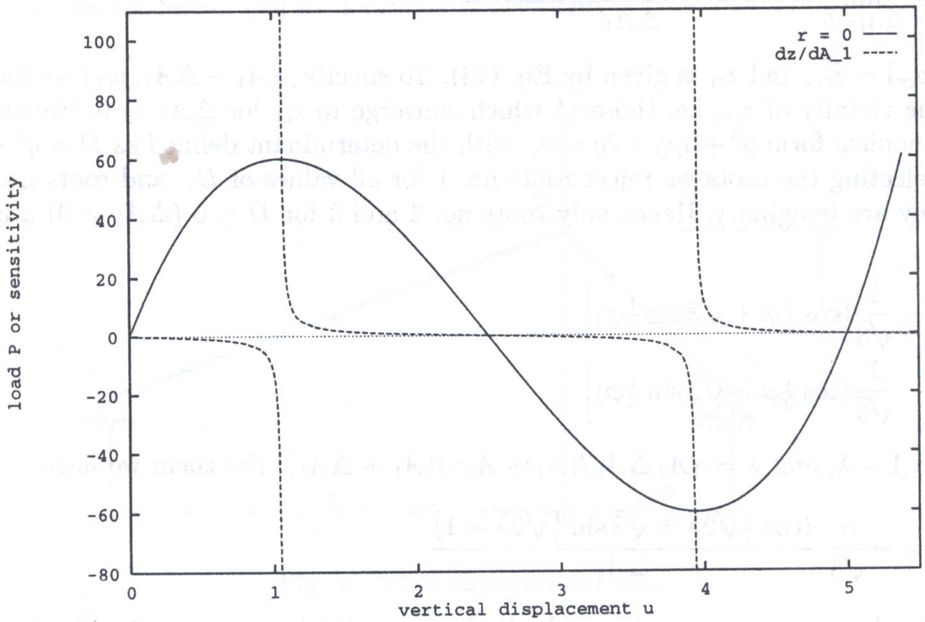

Fig. 2. Equilibrium path and design derivatives of displacement for Scheme 2.

The following data was assumed: $E = 10^7$, $L_1 = 200$, $L_2 = 50$, $H = 2.5$, $A_1 = A_2 = 0.79975$. It is evident from this figure that at extremum points the design derivative of displacement grows unbounded. For the initial part of the equilibrium curve (loads lower than p_{cr}) [10] also provides numerical results, which illustrate this unbounded growth.

Derivative dz_{cr}/db for truss

The design derivative of critical displacement at A_1 is defined as follows

$$\frac{dz_{cr}}{dA_1} \equiv \lim_{\Delta A_1\to 0} \frac{z_{cr}(A_1+\Delta A_1) - z_{cr}(A_1)}{\Delta A_1} \tag{22}$$

where $z_{cr}(A_1) = z_{cr}$, and z_{cr} is given by Eq. (13). As z_{cr} does not depend explicitly on A_1, i.e. $z_{cr}(A_1+\Delta A_1) = z_{cr}(A_1) = \alpha(1 - 1/\sqrt{3})$, hence

$$\frac{dz_{cr}}{dA_1} = 0 \tag{23}$$

Because z_{cr} does not depend on A_1 thus the critical points for different A_1 lie on the same line, $z = z_{cr} = 1.05662$.

4. DESIGN DERIVATIVE OF CRITICAL LOAD

Step 1 of Scheme 1 implies that we have $\mu_{cr} = \mu(\mathbf{b})$. The design derivative of the critical load is defined analogously as the design derivative of the critical displacement, i.e.

$$\frac{d\mu_{cr}}{d\mathbf{b}} \equiv \lim_{\Delta\mathbf{b}\to 0} \frac{\mu_{cr}(\mathbf{b}+\Delta\mathbf{b}) - \mu_{cr}(\mathbf{b})}{\Delta\mathbf{b}} \tag{24}$$

Two methods for calculating $d\mu/d\mathbf{b}$, which have been proposed till now, are described below. In the sequel we omit subscript *cr* to simplify the notation, but all equations are written at the critical point.

Method 1. The residual as well as the determinant of the tangent matrix depend on the design variable $\mathbf{b}$, i.e. $\mathbf{r} = \mathbf{r}(\mathbf{b}, \mathbf{z}(\mathbf{b}), \mu(\mathbf{b}))$, and $d = \det \mathbf{K}(\mathbf{b}, \mathbf{z}(\mathbf{b}))$. Thus we may linearize the set of equations (2) near a design $\mathbf{b}$

$$\left.\begin{aligned} \mathbf{r}(\mathbf{b}+\Delta\mathbf{b}) &= \mathbf{r}(\mathbf{b}) + \frac{d\mathbf{r}}{d\mathbf{b}}\Delta\mathbf{b} = \mathbf{0} \\ d(\mathbf{b}+\Delta\mathbf{b}) &= d(\mathbf{b}) + \frac{dd}{d\mathbf{b}}\Delta\mathbf{b} = 0 \end{aligned}\right\} \tag{25}$$

At a critical point $\mathbf{r}(\mathbf{b}) = \mathbf{0}$ and $d(\mathbf{b}) = 0$, and we may write $(d\mathbf{r}/d\mathbf{b})\Delta\mathbf{b} = \mathbf{0}$ and $(dd/d\mathbf{b}) \cdot \Delta\mathbf{b} = 0$. These equations are satisfied for all $\Delta\mathbf{b}$ if $\det(d\mathbf{r}/d\mathbf{b}) = 0$ and $dd/d\mathbf{b} = \mathbf{0}$. For a case of a single design variable, i.e. when $dim\ \mathbf{b} = 1$, which we consider below, these conditions simplify to $d\mathbf{r}/d\mathbf{b} = \mathbf{0}$ and $dd/d\mathbf{b} = 0$. Then,

$$\begin{aligned} \frac{d\mathbf{r}}{d\mathbf{b}} &= \frac{\partial\mathbf{r}}{\partial\mathbf{b}} + \frac{\partial\mathbf{r}}{\partial\mathbf{z}}\frac{d\mathbf{z}}{d\mathbf{b}} + \frac{\partial\mathbf{r}}{\partial\mu}\frac{d\mu}{d\mathbf{b}} = \mathbf{0} \\ \frac{d\det\mathbf{K}}{d\mathbf{b}} &= \frac{\partial\det\mathbf{K}}{\partial\mathbf{b}} + \frac{\partial\det\mathbf{K}}{\partial\mathbf{z}}\frac{d\mathbf{z}}{d\mathbf{b}} = \mathbf{0} \end{aligned} \tag{26}$$

Note that the derivative $d\mathbf{z}/d\mathbf{b}$ is defined by Eq. (8), not Eq. (10) ! It has been suggested in [4] that this set of equations can be solved for $d\mathbf{z}/d\mathbf{b}$ and $d\mu/d\mathbf{b}$.

Method 2. An alternative approach is based on the critical state condition expressed as $\mathbf{v}^T\,\mathbf{K} = \mathbf{0}$, where $\mathbf{v}$ is the left eigenvector. (For symmetrical $\mathbf{K}$ left and right eigenvectors are identical.) For a residual $\mathbf{r} = \mathbf{f}(\mathbf{b}, \mathbf{z}) - \mu\mathbf{p}$, the first of Eq. (26) can be rewritten as

$$\frac{d\mathbf{r}}{d\mathbf{b}} = \frac{\partial\mathbf{f}}{\partial\mathbf{b}} + \frac{\partial\mathbf{f}}{\partial\mathbf{z}}\frac{d\mathbf{z}}{d\mathbf{b}} - \frac{d\mu}{d\mathbf{b}}\,\mathbf{p} = \mathbf{0} \tag{27}$$

It has been proposed in [16], and [3] that if we pre-multiply this equation by the eigenvector $\mathbf{v}$ then the term with $d\mathbf{z}/d\mathbf{b}$ vanishes due to the fact that $\partial\mathbf{f}/\partial\mathbf{z} = \mathbf{K}$ and $\mathbf{v}^T\mathbf{K} = \mathbf{0}$ by definition of the critical state. Then,

$$\frac{d\mu}{d\mathbf{b}} = \frac{\mathbf{v}^T\,\partial\mathbf{f}/\partial\mathbf{b}}{\mathbf{v}^T\mathbf{p}} \tag{28}$$

Note that this approach fails if $d\mathbf{z}_{cr}/d\mathbf{b}$ grows to infinity faster than $\mathbf{v}^T\mathbf{K}$ decreases to zero.

The above two methods were tested for the truss, what is described in [15] in detail, and they yielded

$$\frac{dp_{cr}}{dA_1} = \frac{de}{dA_1}\,E\,\gamma(\gamma+1)^2\,g \tag{29}$$

where g is defined in (13). Actually, as this is a one-dimensional example, the use of neither of the two methods was necessary. The formula for dp_{cr}/dA_1 can be obtained directly from the first of Eq. (13).

5. Design derivative of determinant

Step 1 of Scheme 2 implies that we have $\mathbf{z} = \mathbf{z}(\mathbf{b}, \mu^{\max})$, and a design derivative of the determinant can be defined as follows

$$\frac{\mathrm{d}\det\mathbf{K}}{\mathrm{d}\mathbf{b}} \equiv \lim_{\Delta\mathbf{b}\to 0} \frac{\det\mathbf{K}(\mathbf{b}+\Delta\mathbf{b}, \mathbf{z}(\mathbf{b}+\Delta\mathbf{b}, \mu^{\max})) - \det\mathbf{K}(\mathbf{b}, \mathbf{z}(\mathbf{b}, \mu^{\max}))}{\Delta\mathbf{b}} \tag{30}$$

where $\mu^{\max}$ is fixed. Hence, the derivative can be expressed as

$$\frac{\mathrm{d}\det\mathbf{K}}{\mathrm{d}\mathbf{b}} = \frac{\partial\det\mathbf{K}}{\partial\mathbf{b}} + \frac{\partial\det\mathbf{K}}{\partial\mathbf{z}}\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{b}} \tag{31}$$

where formulas for $\partial\det\mathbf{K}/\partial\mathbf{z}$, $\partial\det\mathbf{K}/\partial\mathbf{b}$, and $\mathrm{d}\mathbf{z}/\mathrm{d}\mathbf{b}$ are given below.

Calculation of $\partial\det\mathbf{K}/\partial\mathbf{b}$

Let us define a design derivative of the determinant as follows

$$\frac{\partial\det\mathbf{K}}{\partial\mathbf{b}} \equiv \lim_{\Delta\mathbf{b}\to 0} \frac{\det\mathbf{K}(\mathbf{b}+\Delta\mathbf{b}, \mathbf{z}) - \det\mathbf{K}(\mathbf{b}, \mathbf{z})}{\Delta\mathbf{b}} \tag{32}$$

where $\mathbf{z}$ is fixed. It can be shown that $\partial\det\mathbf{K}/\partial\mathbf{b}$ may be evaluated using the following formula

$$\frac{\partial\det\mathbf{K}}{\partial\mathbf{b}} = \operatorname{adj}\mathbf{K} : \nabla_{\mathbf{b}}\mathbf{K} \qquad \text{or} \qquad \frac{\partial\det\mathbf{K}}{\partial b} = \operatorname{tr}\left\{\operatorname{adj}\mathbf{K}\left(\frac{\partial\mathbf{K}}{\partial b}\right)^{\mathrm{T}}\right\} \tag{33}$$

where a colon denotes contraction over two indices, and $\nabla_{\mathbf{b}}\mathbf{K} \equiv \partial\mathbf{K}/\partial\mathbf{b}$. The adjunct matrix $\operatorname{adj}\mathbf{K}$ for a nonsingular $\mathbf{K}$ can be calculated as $\operatorname{adj}\mathbf{K} = \det\mathbf{K}\,\mathbf{K}^{-1}$. At the limit point, where $\det\mathbf{K} = 0$, we must use a definition of $\operatorname{adj}\mathbf{K}$ as a matrix of cofactors, i.e. $\operatorname{adj}\mathbf{K} \equiv [D_{ij}]^{\mathrm{T}}$, where $D_{ij} = (-1)^{i+j}\, d_{ij}$. The minor d_{ij} is defined as a determinant of matrix $\mathbf{K}$ with the i-th row and the j-th column removed, i.e. $d_{ij} \equiv \det\mathbf{K}^{ij}$.

Calculation of $\partial\det\mathbf{K}/\partial\mathbf{z}$

Let us define a derivative of the determinant with respect to the state variable as follows

$$\frac{\partial\det\mathbf{K}}{\partial\mathbf{z}} \equiv \lim_{\Delta\mathbf{z}\to 0} \frac{\det\mathbf{K}(\mathbf{b}, \mathbf{z}+\Delta\mathbf{z}) - \det\mathbf{K}(\mathbf{b}, \mathbf{z})}{\Delta\mathbf{z}} \tag{34}$$

where $\mathbf{b}$ is fixed. A formula for $\partial\det\mathbf{K}/\partial\mathbf{b}$ is analogous to (33),

$$\frac{\partial\det\mathbf{K}}{\partial\mathbf{z}} = \operatorname{adj}\mathbf{K} : \nabla_{\mathbf{z}}\mathbf{K} \qquad \text{or} \qquad \frac{\partial\det\mathbf{K}}{\partial\mathbf{z}_i} = \operatorname{tr}\left\{\operatorname{adj}\mathbf{K}\left(\frac{\partial\mathbf{K}}{\partial\mathbf{z}_i}\right)^{\mathrm{T}}\right\} \tag{35}$$

where $\nabla_{\mathbf{z}}\mathbf{K} \equiv \partial\mathbf{K}/\partial\mathbf{z}$. The gradient of the tangent matrix $\nabla_{\mathbf{z}}\mathbf{K}$ does not exist only at bifurcation points, and can be calculated either analytically, or using the finite difference formula $\nabla_{\mathbf{z}_i}\mathbf{K} \approx [\mathbf{K}(\mathbf{b}, \mathbf{z}+\Delta\mathbf{z}_i) - \mathbf{K}(\mathbf{b}, \mathbf{z})]/\Delta\mathbf{z}_i$. Both require evaluation of $\operatorname{adj}\mathbf{K} : \nabla_{\mathbf{z}}\mathbf{K} = \det\mathbf{K}\,\operatorname{tr}(\mathbf{K}^{-1}\cdot\nabla_{\mathbf{z}}\mathbf{K})$, what cannot be done on the element level. Alternatively, $\partial\det\mathbf{K}/\partial\mathbf{z}$ may be computed using the finite difference method

$$\frac{\partial\det\mathbf{K}}{\partial\mathbf{z}_i} \approx \frac{\det\mathbf{K}(\mathbf{b}, \mathbf{z}+\Delta\mathbf{z}_i) - \det\mathbf{K}(\mathbf{b}, \mathbf{z})}{\Delta\mathbf{z}_i} \tag{36}$$

but this method is prohibitively expensive for large systems.

Calculation of dz/db

Let us recall the definition (9) of the design derivative of the displacement. On differentiation of the equilibrium equation $\mathbf{r}(\mathbf{b}, \mathbf{z}(\mathbf{b})) = \mathbf{0}$ for a configuration in equilibrium, we obtain

$$\frac{\partial \mathbf{r}}{\partial \mathbf{b}} + \frac{\partial \mathbf{r}}{\partial \mathbf{z}} \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{b}} = \mathbf{0} \tag{37}$$

from which we calculate $\mathrm{d}\mathbf{z}/\mathrm{d}\mathbf{b} = -\mathbf{K}^{-1} (\partial \mathbf{r}/\partial \mathbf{b})$.

For the truss the design derivative of the determinant can be expressed as follows

$$\frac{\mathrm{d} \det K}{\mathrm{d} A_1} = \frac{\partial \det K}{\partial A_1} + \frac{\partial \det K}{\partial z} \frac{\mathrm{d}z}{\mathrm{d}A_1}, \tag{38}$$

where

$$\frac{\partial \det K}{\partial A_1} = \frac{\mathrm{d}e}{\mathrm{d}A_1} E\gamma(\gamma+1)^2 \frac{\mathrm{d}a}{\mathrm{d}z}, \qquad \frac{\mathrm{d}e}{\mathrm{d}A_1} = \frac{A_2^2\gamma}{(A_1 + A_2\gamma)^2}, \qquad \frac{\mathrm{d}a}{\mathrm{d}z} = \frac{3}{2}z^2 - 3\alpha z + \alpha^2$$

$$\frac{\partial \det K}{\partial z} = C\,\frac{\mathrm{d}^2 a}{\mathrm{d}z^2}, \qquad \frac{\mathrm{d}^2 a}{\mathrm{d}z^2} = 3(z - \alpha) \tag{39}$$

$$\frac{\mathrm{d}z}{\mathrm{d}A_1} = -K^{-1}\,\frac{\mathrm{d}r}{\mathrm{d}A_1}, \qquad \frac{\mathrm{d}r}{\mathrm{d}A_1} = \frac{\mathrm{d}e}{\mathrm{d}A_1} E\gamma(\gamma+1)^2\, a$$

6. NUMERICAL EXAMPLES

Below, the staggered schemes defined in Section 2 are specialized for the two-bar truss example.

Scheme 1: optimization at the critical point

Step 1. Solve

$$\left.\begin{aligned} r(A_1, z, p) &= 0 \\ \det K(A_1, z) &= 0 \end{aligned}\right\} \tag{40}$$

to find a critical point $\{z_{\mathrm{cr}}, p_{\mathrm{cr}}\}$,

Step 2. Solve

$$\min_{A_1} V \tag{41}$$

subject to $g_\mu = p^{\max} - p_{\mathrm{cr}}(A_1) = 0$ to update A_1.

Scheme 2: optimization at the maximum load

Step 1. Solve

$$r(A_1, z, p) = 0 \tag{42}$$

to find a solution z^{eq},

Step 2. Solve

$$\min_{A_1} V \tag{43}$$

subject to $\det K(A_1, z^{eq}) = 0$ to update A_1.

Numerical calculations were performed for $E = 10^7$, $L_1 = 200$, $L_2 = 50$, $H = 2.5$. The parameters are defined as follows, $V = l_1 A_1 + l_2 A_2$, $dV/\mathrm{d}A_1 = l_1$, $\mathrm{d}g_\mu/\mathrm{d}A_1 = -\mathrm{d}p_{\rm cr}/\mathrm{d}A_1$, where $l_1 = \sqrt{L_1^2 + H^2}$, $l_2 = \sqrt{L_2^2 + H^2}$. The cross-section area of bar no. 2 is fixed, and calculated as $A_2 = V_0/(l_1 + l_2) = 0.79975$. $p^{\max}$ was assumed to be equal to $p_{\rm cr} = 60.12187$, as can be obtained from Eq. (13) for $A_1 = A_2$.

For Scheme 1 the critical point is given by Eq. (13), and $\mathrm{d}p_{\rm cr}/\mathrm{d}A_1$ is given by Eq. (29). For Scheme 2 $\det K = K$ is given by Eq. (21) and $\mathrm{d}\det K/\mathrm{d}A_1$ by Eq. (39).

The Sequential Quadratic Programming procedure of the ADS optimization program, [14], was used. Results of calculations are presented in Table 1, and compared with results obtained via analytical formulas. They are of good accuracy, showing that the formulation is correct. As indicates the number of function evaluations, gradient evaluations, and iterations Scheme 1 converges faster, but Scheme 2 also does, despite the infinity of the design derivative of the displacement at the critical point. Besides, the schemes are analyzed with respect their sensitivity to characteristic inaccuracies, i.e. for Scheme 1 in determination of the critical point, and for Scheme 2 in solving the equilibrium equation. Therefore, the critical load or the solution of the equilibrium equation, were initially perturbed by ±5%, and the error was halved in each iteration. The results presented in Table 1 indicate that the assumed error was never fatal, and convergence was always achieved. As A_1 for the +5% error and Scheme 2 is the most inaccurate one thus Scheme 2 seems to be more prone to inaccuracies than Scheme 1.

Table 1. Solution of the minimum volume problem. Sequential Quadratic Programming

method	error %	A_1	$P_{\rm cr}$	V	no.func.eval.	no.grad.eval.	no.iter.
scheme 1	0	0.80085	60.18800	200.22	34	12	48
scheme 1	+5	0.80076	60.18268	200.20	35	12	49
scheme 1	–5	0.80092	60.19234	200.23	45	13	60
scheme 2	0	0.80327	60.12187	200.70	80	20	102
scheme 2	+5	0.83879	60.12187	207.80	80	20	102
scheme 2	–5	0.80474	60.12187	200.99	79	20	101
analytical	0	0.79975	60.12187	200.00	—	—	—

SQP: ADS with istrat=8, iopt=4, ioned =7.

7. Conclusions

a) Two solution schemes for the minimum volume optimization problem with the critical point constraint were formulated. Scheme 1 leads to optimization at a critical point, while Scheme 2 results in optimization at the maximum load. Scheme 1 is a version of a scheme proposed in [4] for the maximum critical load problem, while Scheme 2 has not yet been described in the literature.

It was pointed out that the design derivatives must be specified consistently with the solution scheme used. This implies specific relations between μ, $\mathbf{z}$ and $\mathbf{b}$. Therefore, a clear distinction must be made between the design derivative of critical displacement and the design derivative of displacement at a critical point, as they can have completely different properties. For the example of the truss the design derivative of the displacement in the vicinity of a critical point grows infinitely, like $\Delta A_1^{-1/2}$ as $\Delta A_1 \to 0$, and it does not exist at the critical point. On the other hand, the design derivative of critical displacement exists at the critical point and is equal to zero.

b) The methods of calculations the design derivatives for both schemes were described and tested. For the example of the truss both schemes worked well and yielded accurate results. As indicates the number of function evaluations, gradient evaluations, and iterations Scheme 1 converges faster, but Scheme 2 also does, despite the infinity of design derivative of displacement at the critical point. It was also checked that both schemes are rather insensitive to inaccuracies, Scheme 1 to inaccuracies in determination of the critical load, and Scheme 2 in solution of the equilibrium equation. The inaccuracies only slowed down the rate of convergence but were not fatal.

ACKNOWLEDGMENT

This research was partly supported by the Polish Committee for Scientific Research under grant No. 253/T07/97/12 and 7T07A01810.

REFERENCES

[1] J.S. Arora, C.C. Wu. Design sensitivity analysis and optimization of nonlinear structures. In: C.A. Mota Soares, ed., *Computer Aided Optimal Design: Structural and Mechanical Systems. NATO ASI Series*, **27**: 589–603, Springer-Verlag, 1987.

[2] K.K. Choi, J.L.T. Santos. Design sensitivity analysis of nonlinear structural systems. Part I. Theory. *Int. J. Num. Meth. Engng*, **24**: 2039–2055, 1987.

[3] R.T. Haftka. Semi-analytical static nonlinear structural sensitivity analysis. *AIAA J.*, **31**(7): 1307–1312, 1993.

[4] M.P. Kamat. Optimization of shallow arches against instability using sensitivity derivatives. *Finite Elements in Analysis and Design*, **3**: 277–284, 1987.

[5] M.P. Kamat, N.S. Khot, V.B. Venkayya. Optimization of shallow trusses against limit point instability. *AIAA J.*, **22**(3): 403–408, March 1984.

[6] N.S. Khot, M.P. Kamat. Minimum weight design of truss structures with geometric nonlinear behavior. *AIAA J.*, **23**(1): 139–144, January 1985.

[7] M. Kleiber, T.D. Hien. Parameter sensitivity of inelastic buckling and post-buckling response. *Comput. Methods Appl. Mech. Engng.*, accepted for publication, 1997.

[8] M. Kleiber, T.D. Hien, E. Postek. Incremental finite element sensitivity analysis for nonlinear mechanics applications. *Int. J. Num. Meth. Engng*, **37**: 3291–3308, 1994.

[9] Z. Mroz, R.T. Haftka. Design sensitivity analysis of non-linear structures in regular and critical states. In: *"Mathematical Methods in Computer Aided Optimal Design". Lecture Notes from Advanced TEMPUS Course, July 12-16, 1993*, 1–32. Faculty of Mathematics and Physics, The Charles University, Prague, 1993.

[10] J.C. Park, K.K. Choi. Design sensitivity analysis of critical load factor for nonlinear structural systems. *Computers & Structures*, **36**(5): 823–838.

[11] U.T. Ringertz. Numerical methods for optimization of nonlinear shell structures. *Structural Optimization*, **4**: 193–198, 1992.

[12] J.L.T. Santos, K.K. Choi. Design sensitivity analysis of nonlinear structural systems. Part II: Numerical methods. *Int. J. Num. Meth. Engng*, **26**: 2097–2114, 1988.

[13] E. Turska, K. Wisniewski, B.A. Schrefler. Error propagation of staggered solution procedures for transient problems. *Comput. Methods Appl. Mech. Engng.*, **114**: 17–188, 1993.

[14] G.N. Vanderplaats. *ADS — A Fortran program for automated design synthesis. Version 2.01. Users Manual.* January 1987.

[15] K. Wisniewski, J.L.T. Santos. On design derivatives for optimization with a critical point constraint. *Structural Optimization*, **11**: 120–127, 1996.

[16] C.C. Wu, J.S. Arora. Design sensitivity analysis and optimization of nonlinear response using incremental procedure. *AIAA J.*, **25**: 1118–1125, 1987.