A BOUND THEOREM IN EIGENVALUES AND
ITS PRACTICAL, APPLICATIONS

Bruce M. Irons*
University of Wales
Swansea

Gabe Treharne†
Ove Arup and Partners
London

A familiar but undervalued theorem is presented, with applications. The theorem is that for positive-definite systems, the eigenvalues are bounded by the extreme element eigenvalues, and also by the extreme eigenvalues of infinitesimal elements. Familiar applications are to the relative stiffnesses of different formulations of the same element, to a roundoff criterion, and to the stability of elastoplastic solutions. Less familiar applications are to the noise dissipation and possible instability of first order marching solutions, and to the instability of creep solutions.

NOMENCLATURE

A, B, M, K are square symmetric matrices.

x, v, w are eigenvectors.

a_i, b_i are contributions to x^TAx and x^TBx from element i: $b_i > 0$.

λ, γ, Γ are eigenvalues.

ω is the over-relaxation factor.

D, D_p are modulus matrices relating stress and strain.

INTRODUCTION

The theorem discussed here gives bounds for eigenvalues in a form especially suited to finite element problems. It is already known, (Reference 9), it is indeed trivial, but it is not as familiar as it should be. Its use in large-scale computing often enables us to comment on the result by means of a very small calculation. Its use in teaching helps us to emphasize the most important aspects of eigenvalues, and gives students more physical insight. The purpose of this paper is to urge its inclusion in finite element courses.

*Senior Lecturer.
†Engineer.

It is fruitful to define an eigenvalue as the stationary value of a Rayleigh quotient. If $\mathbf{Ax} = \lambda\mathbf{Bx}$, then the Rayleigh quotient for a slightly perturbated vector $x + dx$ is

$$\frac{(x+dx)^T A(x+dx)}{(x+dx)^T B(x+dx)} = \frac{x^T Ax + x^T Adx + dx^T Ax + dx^T Adx}{x^T Bx + x^T Bdx + dx^T Bx + dx^T Bdx}$$

$$= \frac{\lambda\,(x^T Bx + x^T Bdx + dx^T Bx) + dx^T Adx}{(x^T Bx + x^T Bdx + dx^T Bx) + dx^T Bdx}$$

which equals λ to second order in $\mathbf{dx}$. The quadratic forms $\mathbf{x^T Ax}$ and $\mathbf{x^T Bx}$ are energies, in general. Thus an eigenvector $\mathbf{x}$ is usually a set of displacements which give the ratio of two energies a stationary value, which is the corresponding eigenvalue λ. The Rayleigh quotient cannot lie outside the range λ_{min} to λ_{max}, because $\mathbf{x^T Ax}/\mathbf{x^T Bx}$ is a well-behaved function of the $\mathbf{x}$.

THE BOUND THEOREM FOR ELEMENT EIGENVALUES

Consider $\lambda = (a_1 + a_2 + \ldots + a_n) / (b_1 + b_2 + \ldots + b_n)$, where all the b_i are positive. If every a_i/b_i is greater than m and less than M, therefore, $mb_i < a_i < Mb_i$. Thus

$$\frac{m(b_1+b_2+ \ldots +b_n)}{(b_1+b_2+ \ldots +b_n)} < \frac{a_1+a_2+ \ldots +a_n}{b_1+b_2+ \ldots +b_n} = \lambda < \frac{M(b_1+b_2+ \ldots +b_n)}{(b_1+b_2+ \ldots +b_n)}$$

or $m < \lambda < M$. Here a_i denotes the contribution to $\mathbf{x^T Ax}$ from element i and b_i denotes the contribution to $\mathbf{x^T Bx}$ from element i. But a_i/b_i is itself bounded by the lowest and highest eigenvalues of element i, namely λ_{min}^i and λ_{max}^i. Thus we have the theorem:

$$\lambda_{min}^i \text{ (over all elements)} < \lambda_{min} \text{ (for structure)}$$

$$< \lambda_{max} \text{ (for structure)} < \lambda_{max}^i \text{ (over all elements)} \qquad (1)$$

For example, the highest frequency of an idealized structure is less than the highest frequency of its smallest element.

THE BOUND THEOREM FOR INFINITESIMAL ELEMENTS

We now consider a_i and b_i as the contributions to $\mathbf{x^T Ax}$ and $\mathbf{x^T Bx}$ of an infinitesimal element i. The extreme eigenvalues of all such infinitesimal elements within a finite element bound the extreme eigenvalues of that element. Further, the extreme eigenvalues of all the infinitesimal elements in a structure bound the extreme eigenvalues of that structure. This is the most useful form of the theorem.

APPLICATION TO A ROUNDOFF CRITERION

A useful criterion of roundoff sensitivity in a Gaussian reduction is the "Diagonal energy" (the sum of the N diagonal contributions to the strain energy) divided by the N^2 terms of the strain energy (Reference 1):

$$\frac{U_d}{U} = \frac{\Sigma_i \, K_{ii} x_i^2}{\Sigma_i \Sigma_j K_{ij} x_i x_j} \qquad (i, \, j = 1 \text{ to } N) \qquad (2)$$

The bandwidth also affects the situation, although to a smaller extent. Now if u_d/u for each element is bounded — as it is if the element matrices are positive definite — then U_d/U is also bounded regardless of how many elements are assembled. A long cylinder with radial motions only is a case in point, or a beam on an elastic foundation: so is the problem of least-squares surface fitting: so is the case, usually, in marching solutions. Unfortunately, most element stiffness matrices are semidefinite.

RELATIVE STIFFNESS OF TWO ELEMENT FORMULATIONS

Suppose we have two formulations of the same problem, e.g. one by potential energy, the other by complementary energy. The first gives assembled stiffness K_1 and the second K_2. If

$$x^T K_1 x \; < \; x^T K_2 x \qquad (3)$$

for all x, we say that the first formulation is less stiff than the second (Reference 2). Extending our ideas of positive-definiteness, we write $K_1 < K_2$. This condition follows if $k_1 < k_2$ for all elements, i.e. if all the eigenvalues from $k_1 x = \lambda k_2 x$ are less than 1, for all elements. The situation is complicated by two facts:

      1. For certain x, for example the rigid body motions, the element forces $k_1 x$ and $k_2 x$ are zero. These are ignored, as $a_i = b_i = 0$.

      2. Certain other x — for example, the constant strain conditions — give $k_1 x = k_2 x$. Indeed, if both formulations are to converge, the unit values of λ must appear.

ROUGH ESTIMATE OF THE CRITICAL STEP—LENGTH
IN TEMPERATURE TRANSIENTS

According to Leech (Reference 3) the critical step-length is about 1/6 of the smallest period in dynamic problems. Small elements have high frequencies. Therefore the expense of a marching solution increases as the mesh is refined, not only because there are more variables, but also because the time step decreases. It is useful to be able to cost this effect in advance.

In a temperature transient, we have $C \tau = - Q \tau$ where τ = temperatures, C = conductivities, and Q = heat capacities. The eigen-situation here is

$$\mathbf{C}\tau = -\lambda\mathbf{Q}\tau \tag{4}$$

so that $\qquad \tau(t) = e^{-\lambda t}\,\tau(0) \tag{5}$

That is, all the temperatures decrease together exponentially, at the same rate. To estimate the biggest λ, we consider the smallest element, we guess the temperature pattern that is most rapidly evanescent as in Figure 1, and we compute λ as a Rayleigh quotient, evidently a very small calculation.

When trapezoidal integration is used, the solution should be stable. What uninformed workers take for instability is usually noise decaying very slowly — although with nonlinearities this can cause instability. Many workers (e.g. C. Parekh, C. Taylor, B. Irons at Swansea, K. Fullard at C.E.G.B. Berkeley, Gloucestershire) have independently proposed re-start techniques to eliminate noise. For example, one steps from $\tau(0)$ to $\tau(t)$ normally, then re-starts at $\tau^*(t/2) = \frac{1}{2}[\tau(0)+\tau(t)]$ having eliminated most of the noise. Figure 2 shows that the spectral effectiveness of this, the simplest algorithm, is surprisingly good. Occasionally without re-starts a solution diverges slowly (in say 200 steps). Presumably roundoff generates more noise per step than is dissipated (Reference 8).

THE CONVERGENCE OF THE ITERATIVE PROCESS IN AN
INCREMENT OF PLASTIC DEFORMATION

Following Gallagher and Marcal (References 4, 5) we take $\mathbf{K}$ as the elastic stiffness and $\mathbf{K_p}$ as the plastic incremental stiffness. During the iteration we compute the residual forces $\mathbf{R}$ which, added to the external loads, would exactly balance the internal loads due to the incorrect state of stress in the structure (Reference 11). We then add to the current deflections $\boldsymbol{\delta}$ of the structure a correction, giving:

$$\boldsymbol{\delta}^* = \boldsymbol{\delta} + \omega\,\mathbf{K}^{-1}\mathbf{R} \tag{6}$$

where ω is the over-relaxation factor. If we are fortunate, and $\mathbf{K_p}^{-1} = \mathbf{K}^{-1}$ exactly, then $\boldsymbol{\delta}^*$ is correct. In general, however, there will again be a nonzero residual:

$$\mathbf{R}^* = \mathbf{K_p} \text{ multiplied by error in } \boldsymbol{\delta}^*$$

$$= (\mathbf{I} - \omega\mathbf{K_p}\mathbf{K}^{-1})\,\mathbf{R}$$

Similarly, $\quad \mathbf{R}^{**\cdots*} = (\mathbf{I} - \omega\mathbf{K_p}\mathbf{K}^{-1})^n\mathbf{R} \tag{7}$

If $\mathbf{R}$ iterates towards zero, the process converges. With this technique we can avoid creating and inverting $\mathbf{K_p}$. But we must consider the conditions for convergence if we are to use the method confidently. We consider first the eigenvectors $\mathbf{v}_i$ such that $\mathbf{K_p}\mathbf{v}_i = \lambda_i\mathbf{K}\mathbf{v}_i$, but it is more convenient to work with the force perturbation, $\mathbf{w}_i = \mathbf{K}\mathbf{v}_i$. For Equation (7) we now have

$$(\mathbf{I} - \omega\mathbf{K_p}\mathbf{K}^{-1})\mathbf{w}_i = (1 - \omega\lambda_i)\mathbf{w}_i \tag{8}$$

which is a scalar multiplier. If Equation (7) is to converge, then, the factor must lie between -1 and 1, giving
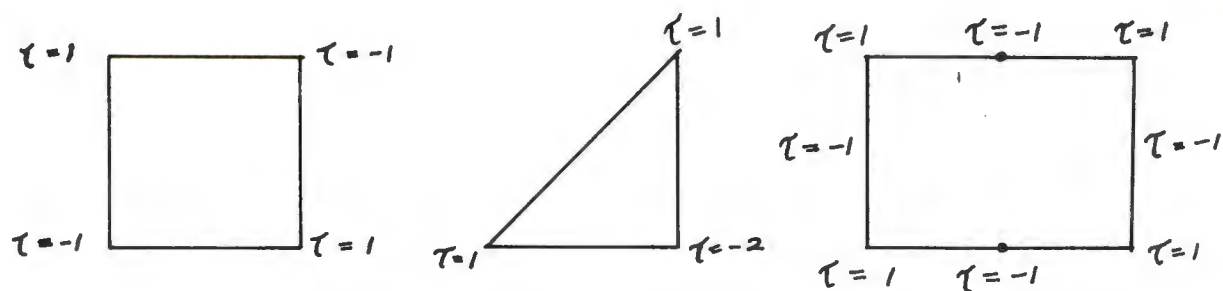
$$0 < \omega\lambda_i < 2 \tag{9}$$

Figure 1. Intelligent Guesswork in Approximating the Noisiest Eigenvector

$$x^*_{\frac{1}{2}} = \tfrac{1}{2}(x_0 + x_1)$$

$$x^*_1 = \tfrac{1}{4}\left(x_0 + 2x_1 + x_2\right)$$

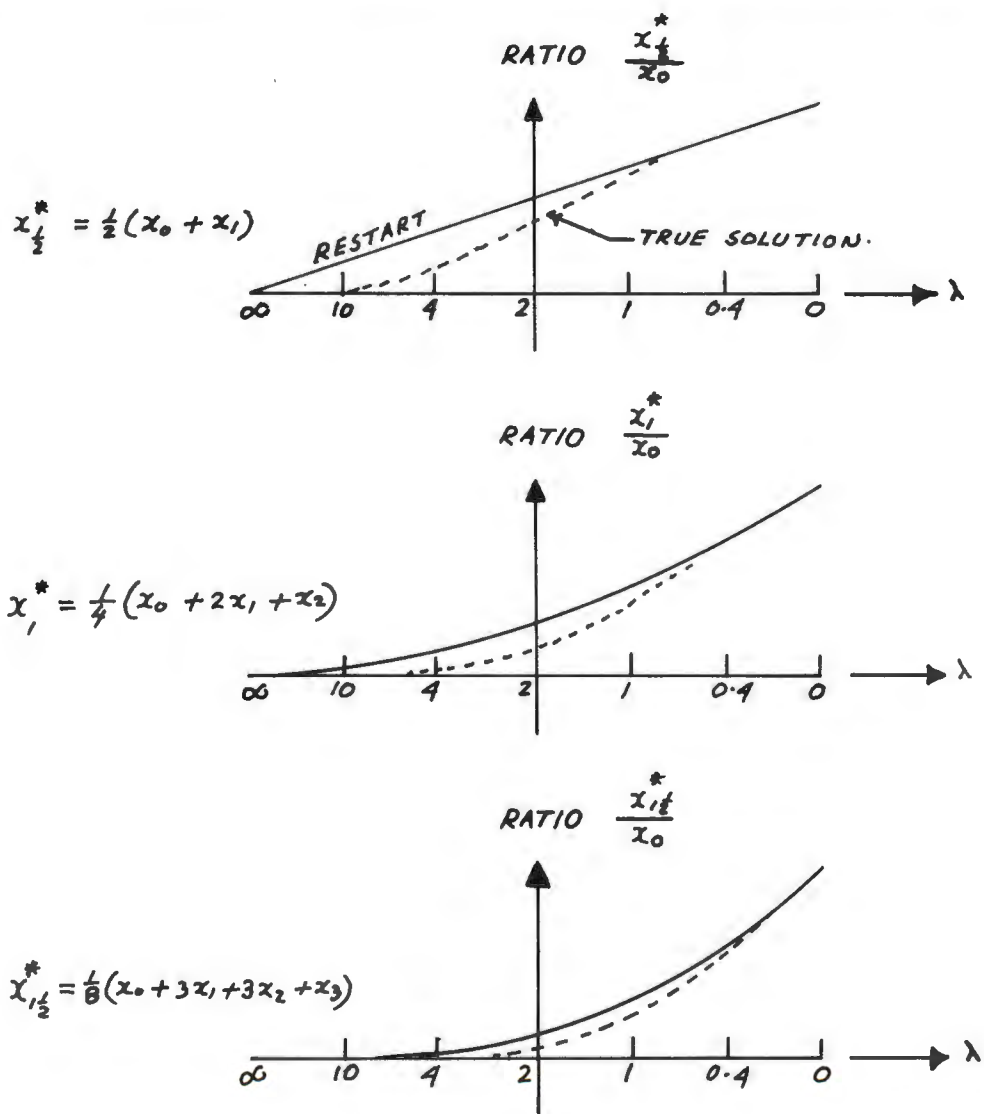$$x^*_{1\frac{1}{2}} = \tfrac{1}{8}(x_0 + 3x_1 + 3x_2 + x_3)$$

Figure 2. Performance of Different Noise-Rejecting Algorithms

As before, the λ_i are bounded by the extreme eigenvalues for individual elements. Presumably, $\lambda_i = 1$ for unyielded elements and $\lambda_i < 1$ for yielded elements. A more useful but wider bound depends only on the properties of an infinitesimal region of material. We choose, not too carefully, a stress increment which appears to give the smallest purely elastic strain energy or the largest purely plastic strain energy. We then take the ratio as an adequate bound on λ.

This is a nontrivial result (Reference 7). A numerical process may converge well in simple cases (e.g. Jacobi relaxation) but fail utterly in problems of any real size. But in the case of plasticity it appears that, if the process converges for a single constant-stress element, then it converges for a large structure. This conclusion holds generally. If an iterative or marching algorithm always succeeds with an infinitesimal element, then it should succeed with a large structure.

Professor G. Maier of Milan has remarked (private communication) that although this conclusion is universally true, the assumptions are unrealistic. Between iterations a small region can switch from elastic to plastic or vice versa, so that the w_i change radically; however the bounds on λ_i remain. As observed by Nayak (Reference 10) this behaviour is not rare, and indeed iteration with tangential stiffnesses is hazardous because the switching can become chaotic. However, we can prove that if Equation (9) is satisfied the perturbation

$$p = \Sigma a_i w_i$$

tends towards zero even if the a_i and w_i change from one iteration to the next.

We consider

$$p^T K^{-1} p = \Sigma a_i^2 w_i^T K^{-1} w_i$$

because the w_i are orthogonal. Each iteration effectively multiplies $w_i^T K^{-1} w_i$ by the factor $F_i = (1 - \omega \lambda_i)^2$ giving a modified $p^T K^{-1} p$:

$$\Sigma a_i^2 F_i w_i^T K^{-1} w_i \leq F_{max} p^T K^{-1} p$$

Therefore if $F_i \leq F_{max} < 1$ for all iterations, i.e. if Equation (9) is satisfied, $p^T K^{-1} p$ tends towards zero and the iteration converges.

We note from Equation (9) that a given value of ω secures convergence for $0 < \lambda < 2/\omega$, and that for convergence $0 < \omega < 2/\lambda_{max}$. One would intuitively expect under-relaxation to be safer in all practical cases; Equation (9) confirms this.

Further, we can now choose a near-optimum value of ω. In the present case, λ_{min} is known from the stress-strain at the worst point in the structure and $\lambda_{max} = 1$ for the elastic regions. Thus the value of ω that minimizes the greatest absolute value of $(1 - \omega \lambda)$, the convergence ratio, is given by

$$1 - \omega \lambda_{min} = -(1 - \omega \lambda_{max})$$

$$\text{or} \quad \omega = 2/(\lambda_{min} + \lambda_{max}) \tag{10}$$

250

NONSYMMETRIC MATRICES

If the problem generates a nonsymmetric matrix K_p, D. J. Naylor has suggested that $K = \frac{1}{2}(K_p + K_p^T)$ might be shown to give convergence. The theorem gives no support here, as the eigenvalues can be complex. It is not known how successful such approaches would be. But the loss of many useful theorems must discourage the most intrepid worker from voluntarily involving himself with nonsymmetric matrices.

STABILITY IN A MARCHING SOLUTION OF
CREEP—TIME PROBLEMS

This is another case that depends directly on the stress-strain laws rather than on some postulated behaviour of the smallest finite element. The problem is technically important, because most workers prefer the simplest form of integration (in which the stress is assumed constant over each time interval) and the intervals must be short to avoid instability: longer intervals would usually give sufficient accuracy, even with this crude type of integration, except perhaps in the early time-steps (Reference 6).

The effective stress:

$$\bar{\sigma}^2 = \frac{3}{2} \sigma_{ij} \sigma'_{ij} \text{ where } \sigma'_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}$$

$$= \frac{1}{2}[(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 + 6\tau_{xy}^2 + \ldots] \qquad (11)$$

is associated with an effective strain in the neighborhood of some time t:

$$\bar{\epsilon} = \bar{\epsilon}\,(\bar{\sigma},\, t)$$

$$\text{thus} \quad d\,\epsilon_x^c = \frac{d\bar{\epsilon}}{\bar{\sigma}}(\sigma_x - \frac{1}{2}\sigma_y - \frac{1}{2}\sigma_z) \text{ etc.} \qquad (12)$$

$$d\gamma_{xy}^c = \frac{d\bar{\epsilon}}{\bar{\sigma}}\,(3\tau_{xy}) \text{ etc.} \qquad (13)$$

as in the Prandtl—Reuss equations. Thus at the start of an interval δ we have the stresses $\boldsymbol{\sigma}$ and we have the elastic stiffness $\mathbf{K}$. We can also postulate a creep stiffness $(\delta t)^{-1}K_c$, although this concept is not pursued for the moment. As before we could consider eigenvectors of displacements or of perturbating forces, but in this case it is more profitable to consider vectors of perturbating stresses. We thus consider an infinitesimal element of material, whose stresses are perturbated from the true solution, so that the stress and the strain perturbations relax exponentially together: the most critical condition is the case giving the fastest rate of decay.

This is essentially the same concept as used above, but we now modify it slightly to take account of the integration process. Consider the problem in one degree of freedom, $y = -\lambda y$, solved using steps of δt, over which y is regarded as constant:

$$y_{i+1} = y_i - \lambda y_i \delta t$$

251

Leaving aside questions of accuracy, the condition for stability is

$$0 < \lambda \delta t < 2$$

Returning to the problem of creep flow, we consider a problem in plane strain ($\epsilon_z = 0$) giving both elastic and creep strains:

$$\text{elastic} \quad \epsilon_z = \sigma_z / E$$

$$\text{creep} \quad \epsilon_z = \sigma_z \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}$$

But the total $\epsilon_z = 0$ at all times, so that

$$\frac{\sigma_z}{E} = -\sigma_z \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}$$

$$\sigma_z = -\lambda \sigma_z \quad \text{where} \quad \lambda = \frac{E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}}{}$$

Substituting λ in Equation (14)

$$E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t} \delta t < 2$$

Thus we derive the limiting step length as required:

$$\delta t < \frac{2}{E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}} \tag{15}$$

THE THREE—DIMENSIONAL CASE IN PLASTICITY AND IN CREEP

In the case of plasticity it is reasonably straightforward to express the result in a general three-dimensional form:

$$\mathbf{D} \, \epsilon = \lambda \, \mathbf{D_p} \epsilon$$

where $\mathbf{D}$ is the elastic modulus matrix and $\mathbf{D_p}$ is the plastic modulus matrix. This gives bounds for the λ of Equation (10).

The case of creep is more difficult. We postulate an interacting set of infinitesimal elements, with no external load involved in the perturbation of the stresses on these elements. Thus we have

$$\delta \epsilon_e = \mathbf{D}^{-1} \delta \sigma$$

$$(\delta \epsilon_c)_{kl} = \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t} \frac{\partial \bar{\sigma}}{\partial \sigma_{ij}} \delta \sigma_{ij} \frac{\partial \bar{\sigma}}{\partial \sigma_{kl}}$$

$$= \mathbf{D}_c^{-1} \, \delta \sigma$$

252

The total strain rate is

$$\delta \epsilon = D_c^{-1} \, \delta \sigma + D^{-1} \, \delta \sigma \tag{17}$$

If we assume that the given perturbation decays according to exp $(-\lambda t)$:

$$\delta \epsilon = [D_c^{-1} - \lambda D^{-1}] \, \delta \sigma \tag{18}$$

We can now construct an energy of perturbations:

$$\tfrac{1}{2} \int \delta \sigma^T \delta \epsilon \, d(\text{vol}) = \tfrac{1}{2} \int \delta \sigma^T \, [D_c^{-1} - \lambda D^{-1}] \, \delta \sigma \, d(\text{vol}) \tag{19}$$

whose derivative with respect to any relevant nodal variable gives an unbalanced force. This is zero, and defines the eigenvalues which control the convergence.

As before, these eigenvalues are bounded by the extreme eigenvalues of the infinitesimal elements:

$$D_c^{-1} \, \delta \sigma = \lambda D^{-1} \, \delta \sigma \tag{20}$$

$$\text{or } D \, \delta \epsilon = \lambda D_c \, \delta \epsilon \tag{21}$$

and hence $\delta t < 2/\lambda_{max}$.

In the isotropic case the criterion was used automatically in the programs of Reference 6. Instability was successfully avoided, and the indications were that the judgments were somewhat pessimistic throughout, but not excessively so.

CONCLUSION

We have discussed the applications of an easily taught and easily remembered theorem. Some of its applications are exceptionally useful. Therefore the theorem should be taught in finite element courses.

REFERENCES

1. Irons, B. M., "Roundoff Criteria in Direct Stiffness Solutions," AIAA Journal, Vol. 6, pp 1308–1312, July 1968.

2. Khanna, J., and Hooley, R. F., "Comparison and Evaluation of Stiffness Matrices", AIAA Journal, Vol. 4, pp 2105–2111, December 1966.

3. Leech, J. W., Pao-Tan Hsu, and Mack, E. W., "Stability of a Finite-Difference Method for Solving Matrix Equations", AIAA Journal, Vol. 3, pp 2172–2173, November 1965.

4. Dupuis, G. A., Pfaffinger, D. D., and Marcal, P. V., "Effective Use of The Incremental Stiffness Matrices in Nonlinear Geometric Analysis", IUTAM Conference on High-Speed Computing of Elastic Structures, University of Liege, Belgium, August 1970.

5. Gallagher, R. H., and Marcal, P. V., (to be published).

6. Treharne, G., "Applications of The Finite Element Method to The Stress Analysis of Materials Subject to Creep", Ph.D. Dissertation, University of Wales, 1971.

7. Argyris, J. H., and Scharpf, D. W., "Methods of Elastoplastic Analysis", Symposium on Finite Element Techniques, University of Stuttgart, June 1969.

8. Kan, D. K. Y., "Solutions Techniques for Large Finite Element Problems Concerning Creep and Temperature Transients in Gas Turbine Blades", Ph.D. Dissertation, University of Wales, 1971, (to be submitted).

9. Tong, P., paper presented at University of Waterloo in conference on finite elements, March 1971.

10. Nayak, G. C., "Plasticity and Large Deformation Problems by the Finite Element Method", Ph.D. Dissertation, University of Wales (Swansea), 1971.

11. Zienkiewicz, O. C., "The Finite Element Method in Engineering Science", McGraw-Hill, 1971.