

On the Solution of
Nonlinear Finite Element Equations

by

Klaus-Jürgen Bathe and Anil Chaudhary
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Eduardo N. Dvorkin
ADINA Engineering, Inc.
Watertown, MA 02172

Miloš Kojić
University of Kragujevac
Kragujevac, Yugoslavia

Abstract

A brief overview of some procedures for the solution of nonlinear finite element equations is given. The iterative schemes we use for the solution of the complete finite element equations are summarized, and the key ideas of a new method for the stress integration in plasticity and creep are briefly presented.

1. Introduction

It is well recognized that the procedures used for the solution of the nonlinear finite element equations are a most important ingredient of a computer program for nonlinear finite element analysis. The equation solution techniques should be as reliable and effective as possible, and indeed the state of these procedures in the computer program frequently decides whether a specific finite element model can or cannot be solved.

For the above reason, much research effort has gone into the development of efficient and reliable equation solution procedures, and significant advances have been made during the recent years; however, with the successful application of nonlinear analysis methods to various problem areas, the demand to solve ever increasingly complex problems has increased — quite naturally and to our delight — which in turn led to identify where important new improvements in equation solution methods are necessary.

In this brief overview we consider the static analysis of structures but the constitutive relations can be time-dependent. In this case, using the notation of ref. [1], the equations to be solved are

$${}^{t+\Delta t}\underline{R} - {}^{t+\Delta t}\underline{F} = \underline{0} \quad (1)$$

where ${}^{t+\Delta t}\underline{R}$ is a vector of the externally applied nodal point forces and ${}^{t+\Delta t}\underline{F}$ is a vector of nodal point forces equivalent in the virtual work sense to the element stresses. The left superscript $t+\Delta t$ denotes that we consider the configuration at the time (load) step $t+\Delta t$. The superscript $t+\Delta t$ is only a variable to denote the load level if the constitutive relations are time-independent, but denotes actual time otherwise (when creep or visco-plastic material models are used in the finite element model).

In the following discussion we assume that the externally applied loads are deformation independent. If the external loading varies as a function of the deformations, the load vector must be updated during the iterations given in Section 2, and there is the question of whether to include a nonsymmetric stiffness matrix contribution. We have found that, in practice, it is frequently more effective to not include the nonsymmetric part in the tangent stiffness matrix.

Using the displacement-based finite element method, the basic unknowns in Eq. (1) are the nodal point displacements at time $t+\Delta t$, which we denote ${}^{t+\Delta t}\underline{U}$. Given the displacements of the element nodal points, we evaluate the element strains, ${}^{t+\Delta t}\underline{\epsilon}$, and stresses, ${}^{t+\Delta t}\underline{\sigma}$, and calculate the vector of nodal forces ${}^{t+\Delta t}\underline{F}$. In the discussion we need to consider only one element, because the element contributions are added by the usual procedures of the direct stiffness method.

Since the relations in Eq. (1) are nonlinear, we need to iterate for the solution ${}^{t+\Delta t}\underline{U}$, and at the end of iteration $(i-1)$, the displacement vector is ${}^{t+\Delta t}\underline{U}^{(i-1)}$, giving the element strains ${}^{t+\Delta t}\underline{\epsilon}^{(i-1)}$, stresses ${}^{t+\Delta t}\underline{\sigma}^{(i-1)}$ and the force vector ${}^{t+\Delta t}\underline{F}^{(i-1)}$,

$${}^{t+\Delta t}\underline{F}^{(i-1)} = \int_V \underline{B}^T {}^{t+\Delta t}\underline{\sigma}^{(i-1)} dV \quad (2)$$

where \underline{B} is the strain-displacement matrix and V is the volume of the finite element. In Eq. (2) we consider a materially-nonlinear-only analysis, because the matrix \underline{B} and scalar V are constant. In a geometrically nonlinear analysis, the total or updated Lagrangian formulation would be employed and the appropriate strain-displacement matrix and volume integrations are used [1], but the same basic iterative procedures to solve Eq. (1) are employed.

Considering the above steps of solution, there are two major ingredients that are the basis and represent the iter-

ation; namely, first, the evaluation of the stresses from the given displacements and, second, the evaluation of the next displacement increments.

The effective evaluation of the stresses $\underline{\sigma}^{(i-1)}$ from the given displacements $\underline{U}^{(i-1)}$ is important and can provide difficulties in particular when creep effects that evolve over long time spans are considered. In this case an implicit integration scheme must be employed for the stress integration which leads to iteration for the stresses at each integration point of a finite element. Since convergence must be reached, it is necessary to select the size of the time step Δt - or the subincrementation within the time step - such as to assure iteration convergence in the stresses at each integration point [2]. In effect, therefore, the radius of convergence in the stress iterations at the element integration points has a pronounced effect on the selection of the time (or load) steps that can be employed in the solution.

The evaluation of the next nodal point displacement increment $\Delta \underline{U}^{(i)}$ should of course be performed in such a way as to obtain rapid and cost-effective convergence to the solution for the complete element system. This is achieved using an appropriate coefficient matrix, updating procedures for this matrix, relaxation techniques, line searches or other methods [1,3].

Finally, there is the question of measuring convergence, where basically the displacements, out-of-balance forces and the strain energy can be used.

Our objective in this talk is to give a brief overview of some recent developments we have pursued for the more effective solution of nonlinear finite element equations, and to summarize some experiences with these procedures. In Section 2 of the paper we discuss the methods that pertain to the iteration for the nodal point displacements, for which we use the Newton and BFGS methods - with or without line searches - and an automatic step-by-step solution scheme [4]. The theory for these techniques has been published extensively before; therefore, only a very brief summary is given here.

In Section 3 we briefly present a new algorithm for the stress iteration at the element integration points [5]. The method shows much promise because the iteration converges effectively and removes much of the restriction on the size of the load or time steps encountered with earlier algorithms.

Finally, in Section 4 we summarize the convergence criteria which we are using in analyses. Sample solutions with the techniques presented here have been published in the references and will be presented at the conference [4,5,9].

2. Iterations for the Displacements

The basic equations used in the Newton iterations and the BFGS method for the displacement solution are

$$\underline{K} \Delta \underline{U} = \underline{t} + \Delta t \underline{R} - \underline{t} + \Delta t \underline{F}^{(i-1)}; \quad i = 1, 2, \dots \quad (3)$$

where $\Delta \bar{U} = \Delta U^{(i)}$, $t+\Delta t_{\underline{U}}^{(i)} = t+\Delta t_{\underline{U}}^{(i-1)} + \Delta U^{(i)}$, unless convergence (as defined later) has not been reached. In this case a line search is performed by iterating with $k=1, \dots, \ell$ as follows,

$$\begin{aligned} \Delta \bar{U}^T \left(t+\Delta t_{\underline{R}} - t+\Delta t_{\underline{F}(k)}^{(i)} \right) \\ \leq \text{stol} * \Delta \bar{U}^T \left(t+\Delta t_{\underline{R}} - t+\Delta t_{\underline{F}(i-1)} \right) \end{aligned} \quad (4)$$

To evaluate $t+\Delta t_{\underline{F}(k)}^{(i)}$ in Eq. (4) we use the displacement vector

$$t+\Delta t_{\underline{U}(k)}^{(i)} = t+\Delta t_{\underline{U}}^{(i-1)} + \beta_k \Delta \bar{U}; \quad \Delta \underline{U}^{(i)} = \beta_\ell \Delta \bar{U} \quad (5)$$

In Eqs. (3) to (4) we use the following notation:

- $\tau_{\underline{K}}$ = tangent stiffness matrix, different for the various iteration schemes,
- ℓ = number of line searches performed in iteration (i),
- stol = convergence tolerance on the line searches,
- β_k = line search parameter in k'th line search.

We are using three different schemes based on Eq. (3):

Method 1

The modified Newton iteration, in which a new stiffness matrix is computed at the beginning of each load step; hence $\tau_{\underline{K}} = {}^t_{\underline{K}}$. Line searches can be performed using this technique [1].

Method 2

The BFGS method in which a new coefficient matrix is computed in each iteration (i), but is evaluated using matrix updates of rank two; here $\tau_{\underline{K}} = {}^t_{\underline{K}}^{*(i-1)}$. Line searches are also always used in this iteration [1,6].

Method 3

The full Newton method in which a new coefficient matrix is evaluated in each iteration; hence $\tau_{\underline{K}} = {}^{t+\Delta t}_{\underline{K}}^{(i-1)}$. This method can be employed with or without line searches.

By an additional approach the solution of the finite element equations can be obtained using an automatic load stepping algorithm [3], which is based on the concepts proposed by Wempner [7] and Riks [8].

Method 4

The automatic procedure uses the equations

$$\tau_{\underline{K}} \Delta \underline{U}^{(i)} = t+\Delta t \lambda^{(i)} \underline{R} - t+\Delta t \underline{F}^{(i-1)} \quad (6)$$

$$t+\Delta t \underline{U}^{(i)} = t+\Delta t \underline{U}^{(i-1)} + \Delta \underline{U}^{(i)} \quad (7)$$

$$f\left(t+\Delta t \underline{U}^{(i)}, t+\Delta t \lambda^{(i)}\right) = 0 \quad (8)$$

where \underline{R} is a vector (constant) representing a load distribution which is scaled by the scalar $t+\Delta t \lambda^{(i)}$, and Eq. (8) represents a constraint between the load and displacement magnitudes. As constraint equations we are using either the spherical arc length constraint or a constraint on the increment of external work [4].

The practical advantage in using this method is that the analyst does not need to specify load increments (as in Methods 1 to 3), and the procedure can also calculate the post-collapse response of a structural model (the loads are decreasing while the displacements are still increasing).

3. Iteration for Integration Point Stresses

With plasticity and creep effects in the analysis a major difficulty lies in the accurate integration of the stresses. This stress integration must be performed at each integration point of the finite elements. The importance of an efficient stress integration is readily realized by considering some of the practical requirements.

For a mesh of 1000 three-dimensional elements evaluated with 3x3x3 Gauss integration, there are 27,000 integration points. If 100 time steps are used to evaluate the response and an average of 3 to 4 iterations for the displacements are performed per step, then a total of about 10^7 integration point stress evaluations need be performed!

Each of these stress computations involves the evaluation of the following integral,

$$t+\Delta t \underline{\sigma}^{(i)} = t \underline{\sigma} + \int_{t_e}^{t+\Delta t \underline{e}^{(i)}} \underline{C}^{EP} d\underline{e} \quad (9)$$

where it should be noted that the integration is always performed from the accepted equilibrium configuration at time t [1]. It is most important to integrate in each iteration anew from time t in order to avoid the accumulation of solution errors.

Considering the integration in Eq. (9), some major difficulties are:

- In plasticity, the stress state must remain within or on the yield surface. This requires tight constraints on the possible variation of the stresses; namely, the numerical procedure has to assure that the stresses do not "grow out" of the yield surface.
- In creep, an algorithm need be used that - when long

time periods must be considered — allows relatively large time steps. The α -method, discussed for example in ref. [2], allows in principle relatively large time steps Δt when $\alpha \geq \frac{1}{2}$. However, in practice, the iteration for the stresses in the implicit time integration does not converge unless small enough time steps are employed. The difficulty is somewhat alleviated by using subincrementation within each time step, but even with this approach the possible time step size is rather small when measured on the total time span that need be considered.

The algorithm we have developed shows much promise with respect to the above difficulties, because the iteration in the implicit time integration of the integration point stresses converges very effectively, which means that the time step size and the subincrementation can be chosen based on accuracy considerations only. The details of our algorithm are presented in ref. [5], and in the following we only give a brief account of the procedure.

A basic equation of thermo-plasticity is, considering isotropic hardening, and denoting tensor components by a curl (\curvearrowright)

$$\underline{\dot{e}}^P = {}^t\Lambda \underline{\dot{S}} \quad (10)$$

where the $\underline{\dot{e}}^P$ are the time derivatives of the plastic strains, the $\underline{\dot{S}}$ are the deviatoric stresses and ${}^t\Lambda$ is a scalar,

$${}^t\Lambda = \frac{3}{2} \frac{\underline{\dot{e}}^P}{\underline{\dot{\sigma}}_y} \left(\frac{\partial \underline{\dot{e}}^P}{\partial \underline{\dot{\sigma}}_y} {}^t\dot{\sigma}_y + \frac{\partial \underline{\dot{e}}^P}{\partial \underline{\dot{\theta}}} {}^t\dot{\theta} \right) \quad (11)$$

where ${}^t\sigma_y$ is the instantaneous yield stress at time t , $\underline{\dot{e}}^P$ is the effective plastic strain and ${}^t\theta$ is the temperature. With the yield stress defined as a function of the effective plastic strain and temperature, Eq. (11) can directly be used to evaluate ${}^t\Lambda$ [5].

Considering the classical description of creep, the inelastic strains are given by

$$\underline{\dot{e}}^C = {}^t\gamma \underline{\dot{S}} \quad (12)$$

where

$${}^t\gamma = \frac{3}{2} \frac{\underline{\dot{e}}^C}{\underline{\dot{\sigma}}} \quad (13)$$

and $\underline{\dot{e}}^C$ and $\underline{\dot{\sigma}}$ are the effective creep strain and the effective stress, respectively.

For the numerical solution, we rewrite Eqs. (10) to (13) in the form

$$\underline{\Delta e}^P = \Delta t \tau_\Lambda \underline{\tau}_S \quad (14)$$

$$\tau_\Lambda \Delta t = \frac{3}{2} \frac{\underline{\Delta e}^P}{\tau_\sigma} \quad (15)$$

$$\underline{\Delta e}^C = \Delta t \tau_\gamma \underline{\tau}_S \quad (16)$$

$$\tau_\gamma \Delta t = \frac{3}{2} \frac{\underline{\Delta e}^C}{\tau_\sigma} \quad (17)$$

where the superscript τ denotes that weighted values using the α -method are employed,

$$\underline{\tau}_S = (1-\alpha) \underline{t}_S + \alpha \underline{t}^{+\Delta t}_S \quad (18)$$

and so on. In practice, we want to use $\alpha \geq \frac{1}{2}$.

The deviatoric stresses at time $t+\Delta t$ and iteration (i) are

$$\underline{t}^{+\Delta t}_S(i) = \frac{1}{2} \underline{t}^{+\Delta t}_G \left(\underline{t}^{+\Delta t}_* (i) - \underline{t}^{+\Delta t}_P(i) - \underline{t}^{+\Delta t}_C(i) \right) \quad (19)$$

where

$$\underline{t}^{+\Delta t}_G = \frac{\underline{t}^{+\Delta t}_E}{2(1 + \underline{t}^{+\Delta t}_\nu)} \quad (20)$$

and $\underline{t}^{+\Delta t}_E$, $\underline{t}^{+\Delta t}_\nu$ are the Young's modulus and Poisson's ratio at time $t+\Delta t$ (that is at the temperature $t^{+\Delta t}_\theta$). In Eq. (19) $\underline{t}^{+\Delta t}_*$ stores the total deviatoric strains. Use of Eqs. (14) to (19) yields

$$\begin{aligned} \underline{t}^{+\Delta t}_S(i) = & \frac{\underline{t}^{+\Delta t}_E}{1 + \underline{t}^{+\Delta t}_\nu + \alpha \Delta t \underline{t}^{+\Delta t}_E (\tau_\Lambda(i) + \tau_\gamma(i))} \left(\underline{t}^{+\Delta t}_* (i) \right. \\ & \left. - \Delta t(1-\alpha) (\tau_\Lambda(i) + \tau_\gamma(i)) \underline{t}_S - \underline{t}^P - \underline{t}^C \right) \quad (21) \end{aligned}$$

and the mean stress is obtained from

$$\underline{t}^{+\Delta t}_{\sigma_m}(i) = \frac{\underline{t}^{+\Delta t}_E}{1-2\underline{t}^{+\Delta t}_\nu} \left(\underline{t}^{+\Delta t}_{e_m}(i) - \underline{t}^{+\Delta t}_{e^{TH}} \right) \quad (22)$$

where $\underline{t}^{+\Delta t}_{e_m}(i)$ is the mean strain and $\underline{t}^{+\Delta t}_{e^{TH}}$ is the thermal strain.

With the displacements $t+\Delta t \underline{U}^{(i)}$ given, it is possible to iterate on Eq. (21), using $\alpha \geq \frac{1}{2}$ until the stresses used to evaluate $\tau_{\Lambda}^{(i)}$ and $\tau_{\gamma}^{(i)}$ are very close to the stresses evaluated in Eq. (21). However, this iteration requires that the time (or load) step be small enough for convergence of the iteration. If a Newton-Raphson type of iteration is employed, experience shows that in creep problems the time step Δt may have to be unduly small.

The basic idea in the development of our effective stress function algorithm is to iterate on a single unknown variable — the effective stress — rather than on the six unknown stresses in Eq. (21). The governing equation is obtained by taking the scalar product of both sides of Eq. (19), which gives

$$\left(t+\Delta t \underline{\sigma}^{(i)}\right)^2 = 9\left(t+\Delta t \underline{G}\right)^2 \left(t+\Delta t \underline{e}^*E(i)\right)^2 \quad (23)$$

where $t+\Delta t \underline{e}^*E(i)$ denotes the elastic deviatoric effective strain at time $t+\Delta t$ and iteration (i),

$$t+\Delta t \underline{e}^*E(i) = \sqrt{\frac{2}{3} \left(t+\Delta t \underline{e}^*E(i)\right) \left(t+\Delta t \underline{e}^*E(i)\right)} \quad (24)$$

Since the right-hand-side of Eq. (23) depends only on the effective stress $t+\Delta t \underline{\sigma}^{(i)}$, our problem has reduced to finding the effective stress which corresponds to the zero of the effective stress function,

$$f\left(t+\Delta t \underline{\sigma}^{(i)}\right) = \left(t+\Delta t \underline{\sigma}^{(i)}\right)^2 - 9\left(t+\Delta t \underline{G}\right)^2 \left(t+\Delta t \underline{e}^*E(i)\right)^2 \quad (25)$$

To solve for the zero of the effective stress function, ie. the unknown effective stress $t+\Delta t \underline{\sigma}^{(i)}$, we substitute for $t+\Delta t \underline{e}^*E(i)$ from Eq. (21) and use a simple but very stable bisection algorithm. Formally, the iteration is continued until with $k=1,2,3,\dots,\ell$

$$f\left(t+\Delta t \underline{\sigma}^{(i)}_{(k)}\right) = \left(t+\Delta t \underline{\sigma}^{(i)}_{(k)}\right)^2 - 9\left(t+\Delta t \underline{G}\right)^2 \left(t+\Delta t \underline{e}^*E(i)_{(k)}\right)^2 \quad (26)$$

$$\left|f\left(t+\Delta t \underline{\sigma}^{(i)}_{(\ell)}\right)\right| \leq \epsilon \quad (27)$$

where ϵ is a convergence tolerance.

Once the effective stress $t+\Delta t \underline{\sigma}^{(i)}$ is known (assumed equal to $t+\Delta t \underline{\sigma}^{(i)}_{(\ell)}$ where ℓ denotes the last iteration), Eq. (21) is used to calculate $t+\Delta t \underline{S}^{(i)}$ and Eq. (22) gives $t+\Delta t \underline{\sigma}_m^{(i)}$.

The importance of the algorithm lies in that only one

unknown variable, the effective stress, need be solved for in the stress point iteration, which allows very large time steps without difficulties of iteration convergence. Hence, with this algorithm the step Δt need only be selected based on accuracy considerations. The algorithm can of course also be employed with subincrementation [5].

4. Convergence criteria

A particular difficulty lies in establishing convergence criteria for the iterative solution of Eq. (1) that are generally applicable in the sense that

- the iteration will be terminated as soon as a reasonable accuracy has been achieved, but not earlier;
- the convergence tolerances to achieve a desired accuracy can be chosen by the analyst.

There is much difficulty in satisfying the above objectives in a general analysis program because of the variety of nonlinearities and response histories that may be modeled. Based on our experience, we have chosen the following convergence criteria to measure the iterative process,

energy tolerance

$$\frac{\Delta \underline{U}^{(i)T} \left(\underline{t} + \Delta t \underline{R} - \underline{t} + \Delta t \underline{F}^{(i-1)} \right)}{\Delta \underline{U}^{(1)T} \left(\underline{t} + \Delta t \underline{R} - \underline{t} \underline{F} \right)} \leq \text{ETOL} \quad (28)$$

force tolerance

$$\frac{\| \underline{t} + \Delta t \underline{R} - \underline{t} + \Delta t \underline{F}^{(i-1)} \|_2^F}{\text{RNORM}} \leq \text{RTOL} \quad (29)$$

where $\| \dots \|_2^F$ denotes the Euclidian norm and only the nodal point forces are included,

moment tolerance

$$\frac{\| \underline{t} + \Delta t \underline{R} - \underline{t} + \Delta t \underline{F}^{(i-1)} \|_2^M}{\text{RMNORM}} \leq \text{RTOL} \quad (30)$$

where only the nodal point moments are included. In Eqs. (28) to (30) ETOL and RTOL are energy and force/moment convergence tolerances, and RNORM and RMNORM are a reference force and a reference moment, respectively. All of these measures have to be chosen by the analyst, and we can give the following brief recommendations:

- It is frequently only necessary to use the energy convergence tolerance, provided ETOL is small enough, say

$ETOL = 10^{-4}$. Note that the convergence check in Eq. (28) includes the effects of the forces, moments, translational displacements and rotations.

- The force and moment convergence tolerances are by-passed when only the energy convergence tolerance is employed. However, it is advisable to always calculate and print the values of $\| \underline{R}^{t+\Delta t} - \underline{R}^{t+\Delta t_{(i-1)}} \|_2^F$ and $\| \underline{R}^{t+\Delta t} - \underline{R}^{t+\Delta t_{(i-1)}} \|_2^M$ so that the analyst can check whether, when convergence has been reached using only Eq. (28), the magnitudes of the out-of-balance forces and moments are actually small enough. To assure that the out-of-balance forces and moments are smaller than a certain value the force and moment tolerances in Eqs. (29) and (30) are employed with RTOL typically equal to 10^{-2} and appropriate values for RNORM and RMNORM. The choice of RNORM and RMNORM clearly depends on the desired accuracy in the analysis, but typically RNORM would be equal to the maximum of the applied load to the structure.

5. Concluding Remarks

The solution of the finite element equations in geometric and material nonlinear analysis encompasses a number of difficult aspects that still require much research effort to obtain increasingly more effective and more automatic procedures — these difficulties lie in the load/displacement incrementation and iteration on the coupled finite element equations for the nonlinear response and the calculation of critical points, in the computation of the stresses at the element numerical integration points and in choosing appropriate convergence tolerances. Surely, an experienced analyst can already solve very complex nonlinear problems. However, a major research challenge lies in providing significantly more automatic solution procedures that can be employed in CAD/CAM software by a large community of design engineers.

References

- [1] K.J. Bathe, Finite Element Procedures in Engineering Analysis, Prentice-Hall, Inc., 1982.
- [2] M.D. Snyder and K.J. Bathe, "A Solution Procedure for Thermo-Elastic-Plastic and Creep Problems", J. Nuclear Eng. and Design, Vol. 64, pp. 49-80, 1981.
- [3] P.G. Bergan, G. Horrigmoe, B. Kråkeland and T.H. Søreide, "Solution Techniques for Nonlinear Finite Element Problems", Int. J. Num. Meth. in Eng., Vol. 12, pp. 1677-1696, 1978.
- [4] K.J. Bathe and E.N. Dvorkin, "On the Automatic Solution of Nonlinear Finite Element Equations", J. Computers and Structures, Vol. 17, No. 5/6, pp. 871-880, 1983.
- [5] M. Kojić and K.J. Bathe, "The Effective Stress Function Algorithm for the Solution of Elasto-Plasticity and Creep Problems", to be published.
- [6] K.J. Bathe and A.P. Cimento, "Some Practical Procedures for the Solution of Nonlinear Finite Element Equations", J. Comp. Meth. Appl. Mech. and Eng., Vol. 22, pp. 59-65, 1980.
- [7] G.A. Wempner, "Discrete Approximation Related to Nonlinear Theories of Solids", Int. J. Solids and Struct., Vol. 7, pp. 1581-1599, 1971.
- [8] E. Riks, "An Incremental Approach to the Solution of Snapping and Buckling Problems", Int. J. Solids and Structures, Vol. 15, pp. 529-551, 1979.
- [9] "ADINA System Verification Manual", ADINA Engineering Inc., Watertown, Ma, and ADINA Engineering AB, Västerås, Sweden.