

SOLUTION OF NONLINEAR PROBLEMS
BY A REDUCED RANK MODIFICATION OF THE ITERATION MATRIX

Alberto Cardona
Sergio R. Idelsohn
Grupo de Tecnología Mecánica
Instituto de Desarrollo Tecnológico para la Industria
Química (CONICET-UNL)
Guemes 3450 - 3000 Santa Fe - Argentina

RESUMEN

Se presenta un algoritmo para calcular la solución de problemas no lineales haciendo uso de un conocimiento previamente adquirido acerca de un subespacio en el que se encuentra principalmente contenida la solución. Esta información es usada para hacer una modificación parcial de la matriz de iteración, modificación que se realiza en cada paso del ciclo iterativo en las direcciones de base del subespacio. El método es aquí aplicado para calcular la respuesta transitoria en problemas no lineales de conducción de calor.

ABSTRACT

An algorithm to compute the solution of nonlinear problems is presented. It makes use of a previously acquired knowledge about a subspace in which the solution is supposed to be mainly contained. This information is used in order to make a partial modification of the iteration matrix. The modification is made at some steps of the iterative cycle, and it updates the matrix in the directions which are basis of the subspace. The method is here tested by computing the transient nonlinear response in thermal conduction problems.

INTRODUCTION

Reduction methods have proved to be an efficient tool to reduce the computational costs when numerically solving nonlinear statics and dynamics problems. Some recent publications report applications of them to structural and thermal problems [1-7]. These methods are based on seeking the solution in a subspace of the original discretization space, and since they reduce the order of the system to be solved, the computer cost to get the solution is highly reduced. The base of the subspace is computed from the original discretization and so, the advantageous property of easily handling complex geometries and/or boundary conditions, is here retained. There are different proposals for basis vectors, which are chosen by considering the problem to be studied.

The approximate solution is computed by nullifying the projection of the residue onto the subspace. Those components that lie in normal directions to this space, can not be eliminated since the method does not see them. Normally these components have a negligible magnitude and so, the obtained solution approaches with high accuracy the true solution. When their magnitudes reach important values, the search sub space is changed in order to eliminate them.

Numerical trials have shown that a problem-independent bound to the latter components that points when the basis vectors should be changed can not be established [5-7]. This measure changes considerably with the problem under consideration. Also, a great value for this tolerance does not indicate if the true solution is well approximated or not. So, for these reasons, the method loses reliability.

We are now proposing a new method which eliminates this inconvenience and also makes use of the information given by the basis vectors. This method is based on making a reduced rank modification of the tangent matrix. The proposed method of modification changes the iteration matrix so that at each iteration it updates all components belonging to the subspace, as the true tangent matrix does. The other components of the unknown vector are updated as if a previous tangent matrix is employed. That is, the new iteration matrix is modified only in the directions pointed out by the basis vectors, where the most important changes in the response are developed. The computational cost of this method is shown to be of the same order as that of a reduction method.

In this paper the method is tested by solving nonstationary nonlinear thermal conduction problems. Comparisons are made with the solution obtained by a reduction method and with the solution obtained by solving the full system of equations by Newton's method.

GOVERNING EQUATIONS

Let us consider a body Ω enclosed by a surface Γ with unit normal n which is subdivided into a prescribed T-surface Γ_T and a prescribed flux surface Γ_q . The governing equations are:

$$\begin{aligned} -\nabla q + s &= \rho c \dot{T} && \text{in } \Omega \\ T &= T^* && \text{on } \Gamma_T \end{aligned}$$

$$\begin{aligned} \mathbf{q} \cdot \mathbf{n} &= \mathbf{q}_n^* && \text{on } \Gamma_q \\ T &= T_0 && \text{in } \Omega \text{ at } t=0 \end{aligned} \quad (1)$$

where T = temperature ; s = source per unit volume ; q = heat flux ;
 ρ = density ; c = specific heat ; and a superposed dot designates the
time (t) derivative.

Fourier's law is assumed for the heat law:

$$\mathbf{q} = -k \nabla T \quad (2)$$

where k is the thermal conductivity matrix.

Performing the finite element discretization process and applying
the standard weighted residuals Galerkin procedure, equations (1-2)
can be recast into the following discrete system of equations:

$$\underline{\mathbf{G}}(\underline{\mathbf{T}}) + \underline{\mathbf{C}}(\underline{\mathbf{T}}) \dot{\underline{\mathbf{T}}} = \underline{\mathbf{F}} \quad (3)$$

in which

$$\begin{aligned} G_i(\underline{\mathbf{T}}) &= \int_{\Omega} \nabla N_i \cdot k(T) \nabla N_j \cdot T_j \, d\Omega \\ C_{ij}(\underline{\mathbf{T}}) &= \int_{\Omega} N_i (\rho c) N_j \, d\Omega \\ F_i &= \int_{\Omega} N_i s \, d\Omega + \int_{\Gamma_q} N_i q_n^* \, d\Gamma \end{aligned} \quad (4)$$

The term $\underline{\mathbf{T}}$ represents the vector of nodal temperatures; N_i are
the element shape functions; $\underline{\mathbf{C}}$ is the global heat capacity matrix;
 $\underline{\mathbf{F}}$ is the global "loads" vector and $\underline{\mathbf{G}}$ is the global internal loads
vector which, for linear heat conduction, can be written as

$$\underline{\mathbf{G}}(\underline{\mathbf{T}}) = \underline{\mathbf{K}} \underline{\mathbf{T}} \quad (5)$$

$\underline{\mathbf{K}}$ denotes the global "linear stiffness" matrix.

The system of nonlinear ordinary differential equations (3) needs
to be time-integrated. This integration is performed numerically by
appealing to the generalized trapezoidal rule.

By assuming a linear variation of the temperature field in a dis-
crete time interval Δt , i.e. from time t_n to time t_{n+1} , and by
imposing that (3) be satisfied at time $t_{n+\alpha} = t_n + \alpha \Delta t$, we have

$$\underline{\mathbf{G}}(T_{n+\alpha}) + \frac{1}{\alpha \Delta t} \underline{\mathbf{C}}_{n+\alpha} T_{n+\alpha} = \underline{\mathbf{F}}_{n+\alpha} + \frac{1}{\alpha \Delta t} \underline{\mathbf{C}}_{n+\alpha} T_n \quad (\alpha \in (0,1]) \quad (6)$$

Solving (6) for $T_{n+\alpha}$, the temperature at time t_{n+1} is given by
the following expression:

$$\underline{T}_{n+1} = \frac{1}{\alpha} \underline{T}_{n+\alpha} + \left(1 - \frac{1}{\alpha}\right) \underline{T}_n \quad (7)$$

This scheme is unconditionally stable for $\alpha \geq 0.5$, for both linear and nonlinear cases [8].

SOLUTION BY A REDUCTION METHOD

Reduction methods are based on generating a set of vectors that forms the basis of a subspace. The system response is assumed to be completely contained in this subspace. Also, eq. (3) is projected onto this subspace in order to get a reduced system of ordinary differential equations that afterwards is numerically time-integrated to compute the response.

After projecting the nonlinear system of differential equations onto the subspace spanned by the basis vectors, the following reduced system of equations is obtained:

$$\overline{\underline{G}}(\underline{y}) + \overline{\underline{C}} \dot{\underline{y}} = \overline{\underline{F}} \quad (8)$$

where

$$\begin{aligned} \overline{\underline{G}} &= \underline{\Psi}^T \underline{G} \\ \overline{\underline{C}} &= \underline{\Psi}^T \underline{C} \underline{\Psi} \\ \overline{\underline{F}} &= \underline{\Psi}^T \underline{F} \end{aligned} \quad (9)$$

and $\underline{\Psi}$ notes for the set of basis vectors.

This system of ordinary differential equations should be time-integrated to compute the response. We employ the trapezoidal rule for this purpose, and to solve the resulting system of nonlinear algebraic equations, we use Newton's method.

There exists many proposals to compute the basis vectors. Clearly, the success of the method depends strongly on the choice of the set of vectors. In the following, we will resume an algorithm for generating a set of basis vectors to solve nonstationary nonlinear thermal problems presented in [7].

The basis vectors are computed by using a modification of the Lanczos' sequence of vectors [3,4,9]. The algorithm for computing them is implemented by employing an equivalent tangent stiffness matrix:

$$K_{ij} = \int_{\Omega} \nabla N_i k_{equiv} \nabla N_j d\Omega \quad (10)$$

Note that this matrix is obtained by omitting the nonsymmetrical component of the true tangent stiffness matrix. An equivalent capacity matrix is computed as:

$$C_{ij} = \int_{\Omega} N_i(\rho c)_{equiv} N_j d\Omega \quad (11)$$

The algorithm for the generation of the basis vectors is written as:

$$\begin{aligned} \underline{K} \bar{\Psi}_{i+1} &= \underline{C} \Psi_i \\ \alpha_i &= \Psi_i^T \underline{C} \bar{\Psi}_{i+1} \\ \bar{\Psi}_{i+1} &= \bar{\Psi}_{i+1} - \alpha_i \Psi_i - \beta_i \Psi_{i-1} \\ \beta_{i+1} &= (\bar{\Psi}_{i+1}^T \underline{C} \bar{\Psi}_{i+1})^{1/2} \\ \bar{\Psi}_{i+1} &= (1/\beta_{i+1}) \bar{\Psi}_{i+1} \end{aligned} \quad (12)$$

This simple form of the Lanczos' algorithm is subjected to a rapid loss of orthogonality with respect to the first computed vectors. In order to prevent this error, a full reorthogonalization is performed periodically. Various algorithms exist that indicate when this correction should be performed [3].

In order to solve thermal transient problems, we propose to employ as a first vector an approximation to the increment of nodal temperatures computed at time $\alpha \Delta t$, that is, the vector computed by solving

$$[\underline{K} + (1/\alpha \Delta t) \underline{C}] \bar{\Psi}_1 = \underline{F} - (1/\alpha \Delta t) \underline{C} \underline{T}_0 \quad (13)$$

This vector is afterwards \underline{C} -normalized and introduced into the algorithm (12) to yield a set of vectors. After that, we take a unit constant vector, which is \underline{C} -orthonormalized and added to this set to form the basis. Note that, for many cases, the latter vector represents the temperature status at $t=\infty$; then, especially for these cases, a large participation factor can be expected for it. When computing the first basis vector we take as equivalent properties k_{equiv} and c_{equiv} the mean material properties in the range of temperatures we are working. Then, the other basis vectors are computed by using the material properties given by the temperature distribution at time $\alpha \Delta t$.

The system response is now expanded as:

$$\underline{T} = \underline{T}_0 + \sum_{i=1}^R \Psi_i y_i = \underline{T}_0 + \underline{\Psi} \underline{y} \quad (14)$$

Whenever the nonlinear terms play an important role, the basis cannot express the temperature increment for a long period of time. The system properties change with the temperature status and so the basis vectors should change accordingly. We can rewrite (14) giving:

$$\underline{T} = \underline{T}_0 + \sum_{i=1}^R \psi_i(z_i) z_i \quad (15)$$

We now assume that the temperature increments can be developed in a Taylor series as follows:

$$\underline{T} = \underline{T}_0 + \left. \frac{\partial \underline{T}}{\partial z_i} \right|_0 z_i + \left. \frac{\partial^2 \underline{T}}{\partial z_i \partial z_j} \right|_0 \frac{z_i z_j}{2} + \dots \quad (16)$$

where repeated indexes imply summation from 1 to R. From (15), the derivatives of the nodal temperatures at \underline{T}_0 are:

$$\left. \frac{\partial \underline{T}}{\partial z_i} \right|_0 = \psi_i(\underline{T}_0) \quad (17)$$

$$\left. \frac{\partial^2 \underline{T}}{\partial z_i \partial z_j} \right|_0 = \left(\frac{\partial \psi_i}{\partial z_j} + \frac{\partial \psi_j}{\partial z_i} \right) (\underline{T}_0)$$

Then, we express the temperature increments as the linear combination of the Lanczos' vectors and their derivatives, both evaluated at \underline{T}_0 :

$$\underline{T} = L \left\{ \psi_1(\underline{T}_0), \left(\frac{\partial \psi_1}{\partial z_j} + \frac{\partial \psi_j}{\partial z_1} \right) (\underline{T}_0), \dots \right\} = \underline{\Psi} \underline{y} \quad (18)$$

Note that the entire process can be seen as a generalized Taylor series with "free" coefficients \underline{y} .

The derivatives of the basis vectors can be obtained after differentiating equation (12-a) and C-orthonormalizing. This procedure leads to the following expression, in which an approximation to the derivatives is computed:

$$\underline{K} \frac{\partial \psi_i}{\partial z_j} = - \frac{\partial \underline{K}}{\partial z_j} \psi_i \quad (19)$$

Approximate derivatives of the tangent stiffness matrix are computed by employing material properties obtained with a linear fit model of the temperature-dependent actual material laws. This approximation gives better results than the differentiation of the true material laws because the linear fit eliminates small localized variations of the material laws that affect the behaviour predicted by the basis vector derivatives.

SOLUTION BY REDUCED RANK MODIFICATION OF THE ITERATION MATRIX

When projecting the system of equations (6) onto the subspace spanned by the basis vectors, a small error is introduced. Due to this error, we have no means to determine, after convergence, if the method

has given a good enough approximation to the true solution. Moreover, bounds for the error norm have proved to be highly problem dependent.

We are now proposing an iterative method that seeks the solution in the whole space of discretization, but at the same time makes use of the information offered by the basis vectors. Since the solution is looked for in the full space, we will be able in theory to find a solution for which the error norm reaches a desired small value.

In order to compute the response, we should solve eq. (6). We will rewrite this equation as forming an N-dimensional system of nonlinear equations:

$$\underline{F}(\underline{x}) = 0 \quad (20)$$

A general iterative method to find a root for this system may be written as [10]:

$$\underline{x}^{k+1} = \underline{x}^k - \underline{A}_k^{-1} \underline{F}(\underline{x}^k) \quad (21)$$

In Newton's method, the iteration matrix \underline{A}_k is chosen to be the Gateaux-derivative of \underline{F} at \underline{x}^k :

$$\underline{x}^{k+1} = \underline{x}^k - \underline{F}'(\underline{x}^k)^{-1} \underline{F}(\underline{x}^k) \quad (22)$$

Many other choices of \underline{A}_k are also available leading to different methods; however, Newton's is the only one for which, under certain conditions on \underline{F} , an estimate of the form

$$\|\underline{x}^{k+1} - \underline{x}^*\| \leq C \|\underline{x}^k - \underline{x}^*\|^2 \quad (23)$$

holds, provided \underline{x}^k is sufficiently closed to a solution \underline{x}^* . This property of "quadratic convergence" is not obtained by other methods and so, for this reason, Newton's method becomes particularly attractive. However, its use in practice is expensive because at each step it requires the solution of a linear system which, especially in problems arising from partial differential equations, can reach several thousands of equations and unknowns.

Let us suppose we have formed and factorized the tangent matrix $\underline{F}'(\underline{x}^0) = \underline{A}_0$ for some given configuration \underline{x}^0 . Then, there is a family of techniques that search the solution to (20) employing an iteration matrix that is formed by modifying \underline{A}_0 (Quasi-Newton methods belong to these kind of techniques) [10]. We will propose a method of this type, in which a reduced rank actualization of the iteration matrix \underline{A}_0 is made. That is, the difference $\underline{A}_k - \underline{A}_0$ now is of rank R. Any NxN matrix of rank R may be written in the form $\underline{U}\underline{V}^T$, where \underline{U} and \underline{V} are NxR matrices of rank R:

$$\underline{A}_k = \underline{A}_0 + \underline{U}_{k-1} \underline{V}_{k-1}^T \quad (21)$$

Now, the hypothesis we are working with is that the temperature increment is mainly contained in the subspace spanned by the basis vectors $\underline{\Psi}$:

$$\Delta \underline{T} = \underline{\Psi} \underline{y} + \underline{\xi} \quad (25)$$

being $\|\underline{\xi}\| \ll \|\underline{\Psi} \underline{y}\|$ and where $\underline{\xi}$ is any vector \underline{C} -orthogonal to all elements of the $\underline{\Psi}$ -subspace. Then, we will ask to the iteration matrix \underline{A}_k to satisfy the following requirements:

$$\begin{aligned} \underline{A}_k \underline{\Psi} \underline{y} &= \underline{F}'(\underline{x}^k) \underline{\Psi} \underline{y} \quad \text{for all } \underline{y} \\ \underline{A}_k \underline{\xi} &= \underline{A}_0 \underline{\xi} \end{aligned} \quad (26)$$

A matrix of the form (24) that verifies these requirements can be built as follows:

$$\underline{A}_k = \underline{A}_0 + (\Delta \underline{A}_k \underline{\Psi}) (\underline{\Psi}^T \underline{C}) \quad (27)$$

where:

$$\Delta \underline{A}_k = \underline{F}'(\underline{x}^k) - \underline{A}_0 \quad (28)$$

Note that, due to the \underline{C} -orthogonality of the basis vectors $\underline{\Psi}$, this matrix verifies eqs. (26). Also, its projection onto the subspace spanned by the basis vectors coincides with the projection of the true tangent iteration matrix:

$$\underline{\Psi}^T \underline{A}_k \underline{\Psi} = \underline{\Psi}^T \underline{F}'(\underline{x}^k) \underline{\Psi} = \overline{\underline{F}'(\underline{x}^k)} \quad (29)$$

To solve eq. (21), we need to compute the inverse of \underline{A}_k . This can be done by using the Sherman-Morrison formula [10]:

$$\underline{A}_k^{-1} = \underline{A}_0^{-1} - \underline{A}_0^{-1} \underline{U} (\underline{I} + \underline{V}^T \underline{A}_0^{-1} \underline{U})^{-1} \underline{V}^T \underline{A}_0^{-1} \quad (30)$$

In fact, this inverse is not computed explicitly; instead, the following iterative algorithm is used to get the solution:

When computing and factorizing \underline{A}_0 :

$$\underline{\Psi}^T \underline{A}_0 = \underline{\Psi}^T \underline{C} \underline{A}_0^{-1}$$

At iteration k : i)

$$\underline{\Psi}^T \underline{\Delta A}_k = \Delta \underline{A}_k \underline{\Psi}$$

$$\text{ii) } \underline{B}_k = (\underline{I} + \underline{\Psi}^T \underline{A}_0^{-1} \underline{\Psi}^T \underline{\Delta A}_k)^{-1}$$

$$\text{iii) } \underline{b} = \underline{B}_k \underline{\Psi}^T \underline{A}_0^{-1} \underline{F}(\underline{x}^k) \quad (31)$$

$$\text{iv) } \underline{c} = \underline{\Psi}^T \underline{\Delta A}_k \underline{b}$$

$$\text{v) } \underline{r} = \underline{F}(\underline{x}^k) - \underline{c}$$

$$\text{vi) } \underline{x}^{k+1} = \underline{x}^k - \underline{A}_0^{-1} \underline{r}$$

If $\| \underline{F}(\underline{x}^{k+1}) \| < \text{TOL } 1 \rightarrow$ end of iterations

If $\| \underline{F}(\underline{x}^{k+1}) \| < \text{TOL } 2 \rightarrow k = k+1 ; \text{GOTO Step (iii).}$

ELSE $\rightarrow k = k+1 ; \text{GOTO Step (i).}$

At each iteration, the costlier steps are (i) and (ii). The first one involves the computation of the changes in the tangent iteration matrix passing by all finite elements. In the second step, the matrix product $\underline{\Psi}_0^T \underline{\Psi}_\Delta \underline{A}_k$ is performed. When the norm of the residue reaches a value beneath a specified tolerance TOL2, these steps are avoided and the iterations proceed from steps (iii) to (vi) until the desired tolerance TOL1 is attained.

NUMERICAL EXAMPLE

To test the method under consideration, the following example was treated in which the transient nonlinear response of a bar submitted to a suddenly applied flux was computed. The bar displayed in Fig. 1 was modeled by using 10 equally spaced linear finite elements giving 11 degrees of freedom. Its temperature dependent material properties are displayed in Fig. 2.

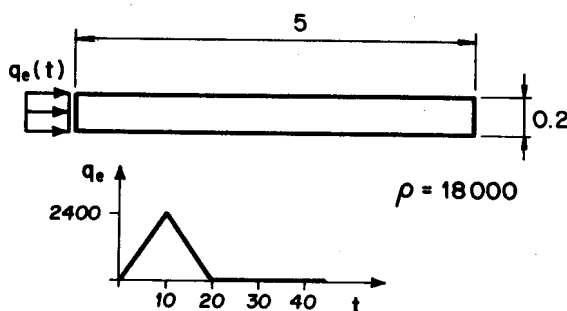


Fig. 1. Test example

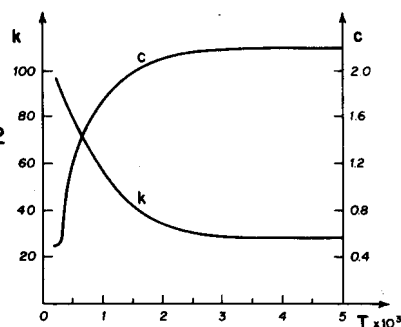


Fig. 2. Temperature-dependent material property

The transient response of the bar was computed by three different methods: (i) a full system solution with standard Newton's iteration; (ii) a full system solution with a reduced rank iteration scheme; (iii) a reduced system solution. In all cases, a time step of 2.5 was employed. Both the reduced rank and the reduced system solutions were computed by using four basis vectors evaluated as indicated in (12-13), and no derivatives were employed. Fig. 3 displays the responses evaluated by the three methods at different times. The full system and the reduced rank solutions are completely superposed, while the reduced system solution gives a highly accurate approximation to them.

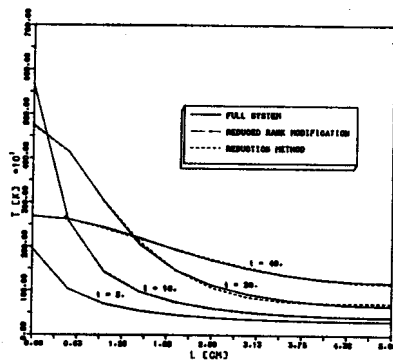


Fig. 3. Response for different times

Fig. 4 shows the evolution in time of the residual norm, for both the reduction method and the reduced rank solutions. We can see that, in spite of giving a quite well approximation to the true response, the reduced system solution maintains a high residue indicating an unbalance in flux between elements.

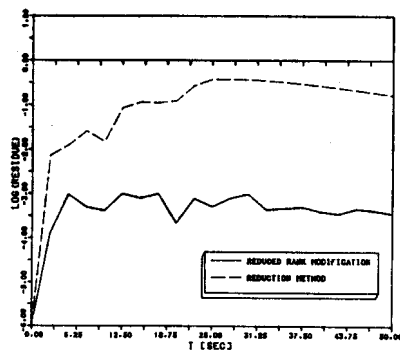


Fig. 4. Evolution of the residual norm (log plot)

The reduced rank method eliminates the latter inconvenience of the reduction method. However, the convergence of the iterative method was affected in this case (a line-search procedure was implemented to improve the convergence properties). Fig. 5 shows the number of iterations per step for the three different solutions. We can see that with the reduced rank method the number of iterations was appreciably increased. When also the derivative of the first basis vector was included in the basis, the number of iterations decreased. This fact confirms that the convergence properties are now influenced by the quality of the basis.

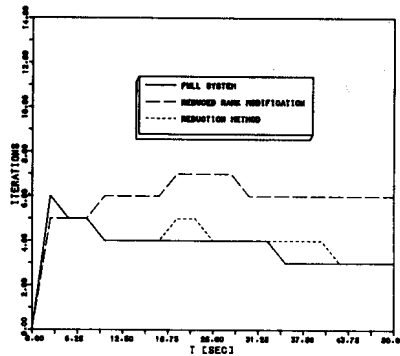


Fig. 5. Number of iterations per step

When dealing with more complex situations (i.e. 3-D cases) we have encountered severe convergence problems but the method needs to be tested more completely to give a final conclusion on its performance.

CONCLUDING REMARKS

An algorithm to compute the solution of nonlinear problems was presented. It makes use of a previously acquired knowledge about the subspace in which the solution is supposed to be mainly contained. The method is highly related with reduction methods, and it employs the same procedures to compute the reduced subspace basis. Now, in this new method, this information is used in order to make a partial modification of the iteration matrix. The modification is made at some steps of the iterative cycle, and it updates the matrix in the directions which are basis of the subspace. So, an iterative cycle which searches the solution in the whole space of discretization is formed, and it is then possible to compute the exact solution to the problem in consideration.

The method was here tested by computing the transient nonlinear response in thermal conduction problems. Some convergence difficulties were evidenced, and more testing should be done to get a final conclusion on its performance.

REFERENCES

- [1] A.K. Noor, "Recent Advances in Reduction Methods for Nonlinear Problems", Internat. J. Comput. and Struct., Vol. 13, 1981, pp. 31-44.
- [2] K.J. Bathe and S. Gracewski, "On Nonlinear Dynamic Analysis using Substructuring and Mode Superposition", Internat. J. Comput. and Struct., Vol. 13, 1981, pp. 699-707.
- [3] B. Nour-Omid and R.W. Clough, "Dynamic Analysis of Structures using Lanczos' Coordinates", Earthquake Eng. Struc. Dyn., Vol. 12, 1984, pp. 565-577.

- [4] S.R. Idelsohn and A. Cardona, "A Load-Dependent Basis for Reduced Nonlinear Structural Dynamics", Proc. Symp. Advances and Trends Struct. and Dynamics, Washington, October 1984, to appear in Internat. J. Comput. and Struct., 1985.
- [5] S.R. Idelsohn and A. Cardona, "A Reduction Method for Nonlinear Structural Dynamic Analysis", Comp. Meth. Appl. Mech. Eng., Vol. 49, 1985, pp. 253-279.
- [6] A.K. Noor, C.D. Balch and M.A. Shibut, "Reduction Methods for Nonlinear Steady-State Thermal Analysis", Internat. J. Numer. Meth. Eng., Vol. 20, 1984, pp. 1323-1348.
- [7] A. Cardona and S.R. Idelsohn, "Solution of Nonlinear Thermal Transient Problems by a Reduction Method", GTM-INTEC, Santa Fe, 1985.
- [8] T.J.R. Hughes, "Unconditionally Stable Algorithms for Nonlinear Heat Conduction", Comp. Meth. Appl. Mech. Eng., Vol. 10, 1977, pp. 135-139.
- [9] C. Lanczos, "An Iterative Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators", J. Res. Nat. Bur. Standards sect. B, Vol. 45, 1950, pp. 255-282.
- [10] Ortega J.M. and Rheinboldt W.C., "Iterative Solution of Nonlinear Equations in Several Variables", Academic Press, New York, 1970.