

An Improved Iterative Optimization Technique for the Leftmost Eigenpairs of Large Symmetric Matrices

GIUSEPPE GAMBOLATI, GIORGIO PINI, AND FLAVIO SARTORETTO

*Istituto di Matematica Applicata,
Via Belzoni, 7, 35131 Padua, Italy*

Received June 2, 1986; revised November 20, 1986

An accelerated optimization technique combined with a stepwise deflation procedure is presented for the efficient evaluation of the p ($p \leq 20$) leftmost eigenvalues and eigenvectors of finite element symmetric positive definite (p.d.) matrices of very large size. The optimization is performed on the Rayleigh quotient of the deflated matrices by the aid of a conjugate gradient (CG) scheme effectively preconditioned with the incomplete Cholesky factorization. No "a priori" estimate of acceleration parameters is required. Numerical experiments on large arbitrarily sparse problems taken from the engineering finite elements (f.e.) practice show a very fast convergence rate for any value of p within the explored interval and particularly so for the minimal eigenpair. In this case the number of iterations needed to achieve an accurate solution may be as much as 2 orders of magnitude smaller than the problem size. Several results concerning the spectral behavior of the CG preconditioning matrices are also given and discussed. © 1988 Academic Press, Inc.

1. INTRODUCTION

The efficient treatment of large numerical models requires computational techniques which are not yet safely established since the computer technology is still in progress and the size of the problems that can be addressed is continuously growing. We present here an efficient technique for estimating the smallest eigenpairs of large sparse matrices. Most of the methods available for the solution of the eigenproblem

$$Av = \lambda v,$$

where A is a sparse symmetric matrix, have been reviewed by Parlett [1]. A vast class of techniques for a matrix A which is also positive definite rely on the optimization of the Rayleigh quotient $q(x)$:

$$q(A, x) = q(x) = \frac{x^T Ax}{x^T x} \tag{1}$$

It is well known that ratio (1) is stationary if x is an eigenvector of A and $q(x)$ takes on the maximal and minimal value for the highest and lowest eigenvectors, respectively.

The minimization of (1) by a gradient (steepest descent) scheme was first described by Hestenes and Karush [2, 3], see also Hestenes [4]. Unfortunately the steepest descent converges very slowly if the matrix is not small. A more reliable algorithm, referred to as "coordinate relaxation," was subsequently studied by Faddeev and Faddeeva [5] and reanalyzed by several other authors (Nesbet [6]; Bender and Shavitt [7]; Shavitt *et al.* [8]; Falk [9]). Nisbet [10], Ruhe [11], and Schwarz [12] suggested an improvement of the method with the aid of an acceleration parameter ω in complete analogy to SOR as applied to the solution of linear sets of equations arising from the finite difference integration of elliptic boundary value problems. However, the above authors failed to provide a general theory for the preliminary assessment of the best over-relaxation factor which is therefore to be determined empirically for any given matrix A .

An important modification of the gradient technique for the optimization of (1) is represented by the method of conjugate gradients (CG) originally developed by Hestenes and Stiefel [13] for the solution of linear symmetric p.d. systems. Along this line are the works by Bradbury and Fletcher [14], Fried [15, 16], Gerardin [17], and Ruhe [18]. In a recent paper Papadrakakis [19] has combined the CG scheme with the symmetric successive coordinate over-relaxation but a serious deficiency of this approach is again the need for improving the convergence through an optimal acceleration factor which is problem dependent and generally unknown "a priori."

Iterative algorithms based on convergent splittings have been analyzed by Ruhe [20] and also proposed by Evans and Shanehchi [21] who suggested a more efficient splitting choice. It is to be noted that in these works too, the convergence is to be properly speeded up by the use of a relaxation parameter which is difficult to estimate for general matrices.

All these well-established techniques become computationally expensive for large size problems and particularly so if a number of the smallest eigenpairs is sought. According to Sameh and Wisniewski [22] "no efficient methods for simultaneously obtaining several eigenvalues and eigenvectors is available."

Actually iterative approaches for the simultaneous computation of the leftmost eigenpairs based on either trace minimization (Sameh and Wisniewski [22]) or multiple Rayleigh quotient optimization (Longsine and McCormick [23]; Schwarz [24]) have lately been developed. Unfortunately they are not easy to implement practically in a computer code and their numerical behavior has been explored mainly with small and unrealistic sample problems.

In recent years a modification of the method of conjugate gradients (MCG modified conjugate gradients) has significantly improved the performance of this scheme as applied to the solution of large sparse p.d. systems (Kershaw [25]; Concus *et al.* [26]; Meijerink and van der Vorst [27]; Axelsson [28]; Gambolati [29, 30]; Manteuffel [31]; Gambolati and Volpi [32]). The number M of iterations required to achieve a relative accuracy ε is related to the square root of the spectral condition number ξ of the MCG iteration matrix $E = AK^{-1}$, K^{-1} being the CG preconditioning matrix. It may be shown (Axelsson [33]) that:

$$M = \text{int} \left(\frac{1}{2} \sqrt{\xi} \ln \frac{2}{\varepsilon} + 1 \right)$$

where $\xi(E) = \lambda_1(E)/\lambda_N(E)$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ being the ordered real positive eigenvalues of E . The previous equation holds for values of ξ not too close to 1 and may be safely used for $\xi > 3$ or 4.

For any given problem the behavior of ξ is dependent on the preconditioning matrix K^{-1} . Many preconditioners belong to the class of the incomplete Cholesky factorization (Eisenstat [34]; Nour-Omid [35]; Ajiz and Jennings [36]; Jackson and Robinson [37]). For f.e. systems arising from the numerical integration of subsurface flow and structural equations the incomplete decomposition, referred to as ICCG(0) by Meijerink and van der Vorst [27], has proven extraordinarily effective (Gambolati and Perdon [38]).

Attempts have also been made to extend the MCG scheme to nonsymmetric matrices (Gambolati [39]; Axelsson [40]) but the results have not been so successful.

The idea underlying the present work is to apply a similar preconditioning procedure for the CG optimization of the Rayleigh quotient (1) together with a shifting deflation technique to assess the p (with $15 \leq p \leq 20$) leftmost eigenpairs of A . Early results from the eigenanalysis of very large matrices (Gambolati and Perdon [41], Perdon and Gambolati [42]) emphasize the promising features of the accelerated conjugate gradients.

In this paper the algorithm for the evaluation of the p leftmost eigenpairs by the stepwise deflating MCG approach is first outlined. The iterative deflation is conceptually similar to that employed by Shavitt *et al.* [8], Falk [9], and Schwarz [12]. Schwarz [24] raised the concern that such a procedure would not work satisfactorily but it will be shown that this is not true in a finite element context.

The high numerical performance of the combined scheme is demonstrated for some large size problems related to the f.e. integration of flow and elasticity equations. An extensive analysis of the behaviour of $\xi(E)$ during the deflation process is also supplied. It is shown empirically that the number of iterations required for an accurate solution is smaller when the extreme eigenvalue is to be assessed and increases as the position of the desired eigenvalue moves rightward within the spectral interval. However for all the values of p explored in the present analysis ($15 \leq p \leq 20$) the convergence proved to be good. For any of the given examples the structure of the matrix, the distribution of the p leftmost eigenvalues and the MCG convergence profiles are presented and discussed.

2. ITERATIVE OPTIMIZATION OF RAYLEIGH QUOTIENTS

2.1. *The Accelerated Conjugate Gradients for the Minimal Eigenvalue and Eigenvector*

Denote by $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ the real positive eigenvalues of A and by v_1, v_2, \dots, v_N the corresponding eigenvectors. In the engineering practice it is well known that the smallest eigenpairs correspond to the relevant modes of oscillation of the system being numerically analyzed while the highest have very little physical meaning as they are severely influenced by the truncation errors. Therefore we focus our attention on the p leftmost stationary values of (1), starting from the minimal eigenpairs, i.e., λ_N and v_N .

Let us indicate by $g(x)$ and $H(x)$ the gradient and the Hessian of the Rayleigh quotient, respectively. We readily obtain

$$g(x) = q'(x) = \frac{2}{x^T x} [Ax - q(x)x] \quad (2)$$

$$H(x) = q''(x) = \frac{2}{x^T x} \{A - q(x)I - [g(x)]x^T - x[g(x)]^T\}. \quad (3)$$

The rate of convergence of an optimization algorithm for the computation of λ_N is dependent on the spectral condition number $\xi(H)$ of $H(v_n)$, namely on the ratio between the largest and the smallest (different from zero) eigenvalue of the Hessian evaluated for $x = v_N$ (see Ruhe [18]). From (3) we have immediately:

$$\xi(H(v_N)) = \frac{\lambda_1 - \lambda_N}{\lambda_{N-1} - \lambda_N}. \quad (4)$$

The traditional CG recursive equations to minimize (1) may be found in several papers, see, for instance, Bradbury and Fletcher [14], Ruhe [18], and Schwarz [24]. To derive the MCG scheme we proceed as follows. Set

$$y = Xx, \quad (5)$$

where X is an auxiliary symmetric matrix and replace (5) in (1). We may write

$$q_1(y) = \frac{y^T X^{-1} A X^{-1} y}{y^T X^{-1} X^{-1} y} = \frac{y^T G y}{y^T K^{-1} y}, \quad (6)$$

where $G = X^{-1} A X^{-1}$ and $K^{-1} = X^{-1} X^{-1}$. We note that G and K^{-1} are both symmetric p.d. and Eq. (6) may be regarded as the Rayleigh quotient of the generalized eigenproblem $Gy = \lambda K^{-1}y$. If we minimize the right-hand side of Eq. (6) by the CG scheme (e.g., Ruhe [18]) and then restore the original variable x through Eq. (5) we obtain the following MCG relationships

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k p_k \\ p_k &= K^{-1} g_k + \beta_{k-1} p_{k-1} \\ g_k &= 2 \frac{Ax_k - q(x_k)x_k}{x_k^T x_k} \end{aligned} \quad (7)$$

$$\beta_{k-1} = -\frac{p_{k-1}^T AK^{-1}g_k}{p_{k-1}^T Ap_{k-1}}$$

$$p_0 = K^{-1}g_0,$$

where

$$\alpha_k = 0.5(nd - mb + \sqrt{\Delta})/(bc - ad)$$

$$a = p_k^T Ax_k \quad b = p_k^T Ap_k$$

$$c = p_k^T x_k \quad d = p_k^T p_k$$

$$m = x_k^T x_k \quad n = x_k^T Ax_k$$

$$\Delta = (nd - mb)^2 - 4(bc - ad)(ma - nc).$$

We have

$$q(x_k) \rightarrow \lambda_N \quad x_k \rightarrow v_N$$

K^{-1} is the preconditioning matrix of the MCG scheme (7). If $K=I$, Eqs. (7) turn into the usual CG iteration. The initial vector x_0 is arbitrarily chosen. In the leftmost stationary point y_N (for which $q_1(y_N) = \lambda_N$) the Hessian of $q_1(y)$ takes on the expression

$$H(y_N) = \frac{2}{y_N^T K^{-1} y_N} [G - \lambda_N K^{-1}]. \quad (8)$$

We observe that $G - \lambda_N K^{-1}$ is similar to $(A - \lambda_N I) K^{-1}$.

If we assume that K^{-1} equals A^{-1} , i.e., the inverse of A , we get the following spectral condition number for H :

$$\xi(H) = \frac{\lambda_1 - \lambda_N}{\lambda_{N-1} - \lambda_N} \frac{\lambda_{N-1}}{\lambda_1} \quad (9)$$

For large finite difference or finite element matrices arising from the engineering practice usually

$$\lambda_{N-1} \ll \lambda_1$$

and hence ratio (9) is much smaller than (4). Consequently the MCG algorithm with $K^{-1} = A^{-1}$ would converge much faster than the traditional CG scheme.

Actually A^{-1} is unknown and its calculation would be expensive and require much storage. We may use instead an approximation to A^{-1} as is efficiently provided by one of the several incomplete Cholesky decompositions of A (see

[25–27, 36]). As was already mentioned in the Introduction an excellent choice for the solution of linear systems in a f.e. context turned out to be

$$K^{-1} = (LL^T)^{-1}, \quad (10)$$

where L is the incomplete triangular factor of A having the same sparsity pattern as A . Experience has shown (Gambolati and Perdon [38]) that $\xi(A(LL^T)^{-1})$ is several orders of magnitude smaller than $\xi(A)$ and therefore (10) may be viewed as a good approximation to A^{-1} . It follows that the modified conjugate gradients (7) with a preconditioning matrix equal to (10) are indeed expected to converge faster than the CG scheme.

2.2. The Iterative Stepwise Procedure for the Next Higher Stationary Values

Define the new matrix A_1 as

$$A_1 = A + (\alpha_1 - \lambda_N) v_N v_N^T, \quad (11)$$

where λ_N and v_N have just been computed as described in the previous section. If the scalar α_1 is chosen so as to be larger than λ_{N-1} and possibly (but not necessarily) smaller than λ_{N-2} , i.e.,

$$\lambda_{N-1} < \alpha_1 < \lambda_{N-2} \quad (12)$$

the ordered eigenpairs of A_1 turn out to be:

$$\begin{aligned} &\lambda_{N-1}, \alpha_1, \lambda_{N-2}, \dots, \lambda_1 \\ &v_{N-1}, v_N, v_{N-2}, \dots, v_1. \end{aligned}$$

To prove the previous statement one should recall that the vectors v_1, v_2, \dots, v_N form an orthonormal set.

The leftmost eigenvalue and eigenvector of A_1 are now λ_{N-1} and v_{N-1} and may be computed by the MCG procedure given in Section 2.1, after replacing A by A_1 and with the aid of the same preconditioning matrix $(LL^T)^{-1}$. In the numerical models of large hydrodynamic systems or mechanical structures $\lambda_N \ll \lambda_1$, the characteristic values exhibit to some extent a uniform distribution between the extreme bounds and the separation is usually not too bad. Thus it may easily occur that λ_{N-1} is quite close to λ_N and yet well separated from it, from a numerical point of view. Parameter α_1 can be taken so that the shift $\alpha_1 - \lambda_N$ is a very small quantity compared to the length of the spectrum of A and $\xi(A_1 K^{-1})$ is not appreciably larger than $\xi(A K^{-1})$. Equation (12) may be hard to satisfy in practice and we may take α_1 in order that λ_N and v_N are not shifted too far beyond λ_{N-1} and v_{N-1} and $(LL^T)^{-1}$ still represents an acceptable estimate of the inverse of A_1 . (On the other hand, α_1 should not be close to λ_{N-1} to prevent the retardation of the convergence of the algorithm to v_{N-1} .) In practical computations the

appropriate α_1 is not known a priori and has to be determined iteratively. We form the matrix

$$A_1^{(i)} = A + (\alpha_1^{(i)} - \lambda_N) v_N v_N^T, \quad i = 0, 1, \dots, \quad (13)$$

where we set

$$\alpha_1^{(0)} = \beta \lambda_N, \quad \alpha_1^{(i)} = \beta \alpha_1^{(i-1)}, \quad \beta > 1.$$

The overall procedure works as follows. Choose $\beta > 1$. If $\alpha_1^{(0)}$ is smaller than λ_{N-1} increase $\alpha_1^{(i)}$ according to the criterion given above until the MCG scheme (Eqs. (7)) applied on $A_1^{(i)}$ converges to an eigenvalue which is different from $\alpha_1^{(i)}$. This will be λ_{N-1} . Continue the iterations to ensure a good estimate for v_{N-1} as well. Remember that the convergence to the minimal eigenvalue is twice the convergence of the corresponding eigenvector (Schwarz [12]; Ruhe [18]) and that a good evaluation for v_{N-1} is required to guarantee an effective deflation procedure. It is to be noted that the numerical trials performed to obtain a value for $\alpha_1^{(i)}$ greater than λ_{N-1} are not computationally expensive if one starts from an initial guessed solution $x_0 = v_N$. When $\alpha_1^{(i)} < \lambda_{N-1}$ the MCG iteration gives $q(A_1^{(i)}, x_k) \rightarrow \alpha_1^{(i)}$ and $x_k \rightarrow v_N$, but as soon as $\alpha_1^{(i)} > \lambda_{N-1}$ the differences $d_{1,k}^{(i)} = |q(A_1^{(i)}, x_k) - \alpha_1^{(i)}|/\alpha_1^{(i)}$ and $r_{1,k}^{(i)} = |A_1^{(i)} x_k - q(A_1^{(i)}, x_k) x_k|/|x_k|$ tend to increase with k . A test is defined for $d_{1,k}^{(i)}$. If the magnitude of $d_{1,k}^{(i)}$ exceeds a specified tolerance T_d then this means that $q(A_1^{(i)}, x_k)$ and x_k are converging towards $\lambda_{N-1} \neq \alpha_1^{(i)}$ and v_{N-1} . To improve the convergence the iteration is restarted with an arbitrary new vector w (different from v_N) and is completed when a sufficiently small $r_{1,k}^{(i)}$ value is obtained. If $d_{1,k}^{(i)}$ does not appreciably increase after a preselected number I_f of iterations we increase $\alpha_1^{(i)}$ and start again the MCG scheme with $x_0 = v_N$. It is interesting to notice that in principle the conjugate directions could not move from the stationary point v_N . In practice, however, due to roundoff errors after an initial almost static behavior, $d_{1,k}^{(i)}$ grows with k and thus we know that the stationary point v_N is no longer the minimal value of the Rayleigh quotient of A_1 .

Making use of the already computed eigenpairs the shifting deflation procedure given above is readily extended to the assessment of λ_{N-j} and v_{N-j} through the matrix $A_j^{(i)}$:

$$A_j^{(i)} = A + (\alpha_j^{(i)} - \lambda_N) v_N v_N^T + (\alpha_j^{(i)} - \lambda_{N-1}) v_{N-1} v_{N-1}^T + \dots + (\alpha_j^{(i)} - \lambda_{N-j+1}) v_{N-j+1} v_{N-j+1}^T \quad j = 1, 2, \dots, p; i = 1, 2, \dots, \quad (14)$$

where now we have to look for a scalar $\alpha_j^{(i)}$ satisfying inequality (15):

$$\alpha_j^{(i)} = \beta \alpha_j^{(i-1)} > \lambda_{N-j} \quad (15)$$

$$\alpha_j^{(0)} = \beta \lambda_{N-j+1}, \quad \beta > 1.$$

It is worth noting that the shifts in Eq. (14) accumulate the leftmost $(N-j+1)$ eigenpairs of A in a unique stationary point where the Rayleigh quotient of $A_j^{(i)}$

takes on the value $\alpha_j^{(i)}$ corresponding to the distinct eigenvectors $v_N, v_{N-1}, \dots, v_{N-j+1}$. Multiple eigenvalues close to the minimal one may greatly slow down the convergence rate and make it more difficult to recognize whether $\alpha_j^{(i)}$ has overcome λ_{N-j} . However the variable shifts defined by Eq. (14) produce the smallest modification of A consistently with the deflation approach, thus enhancing the probability for K^{-1} to be still an acceptable approximation to $(A_j^{(i)})^{-1}$. Actually the numerical results given in the next section shows that $\xi(A_j^{(i)}K^{-1})$ increases slowly with j , at least for not too large j -values thus showing that $(LL^T)^{-1}$ remains an effective MCG preconditioning matrix as the deflation proceeds. To prevent retardation of the convergence $\alpha_j^{(i)}$ should not be too close to λ_{N-j} . An optimal $\alpha_j^{(i)}$ realizes a satisfactory trade-off between the two opposite needs for having $A_j^{(i)}$ close to A (small $\alpha_j^{(i)}$) and the multiple stationary point $v_N, v_{N-1}, \dots, v_{N-j+1}$ far from v_{N-j} (large $\alpha_j^{(i)}$).

One last consideration bears mention. Matrices $A_j^{(i)}$ are not sparse, consequently the stepwise procedure described above is to be performed implicitly by saving in core memory all the eigenpairs currently found.

3. NUMERICAL RESULTS

The performance of the deflation—MCG scheme presented in this paper has been analyzed for the computation of the p (with $15 \leq p \leq 20$) leftmost eigenvalues and eigenvectors of 4 large sparse symmetric p.d. matrices whose size N ranges from 156 to 2304. They arise from the finite element integration of the flow equation in 3D multiaquifer systems and of the elasticity equations of layered 3D structures subject to vertical and horizontal land subsidence. The order of the problems is 156, 812, 1802, and 2304. The overall number of nonzero A coefficients is 996, 5458, 24468, and 18712, respectively, to which a sparsity percentage of 95.9%, 99.2%, 99.2%, and 99.6% corresponds. The structure of the matrices may be seen in Figs. 1, 2, 3, and 4. Note that the general pattern is that of a banded structure with a somewhat irregular distribution of the nonzero terms within the band.

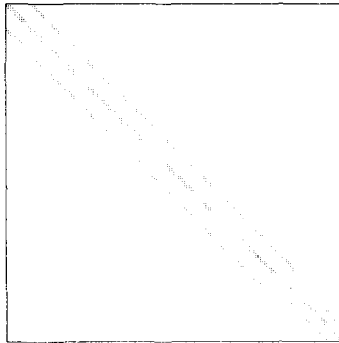


FIG. 1. Structure of the f.e. matrix for the test problem with $N=156$ (integration of the steady diffusion equation in a 3D axisymmetric system).

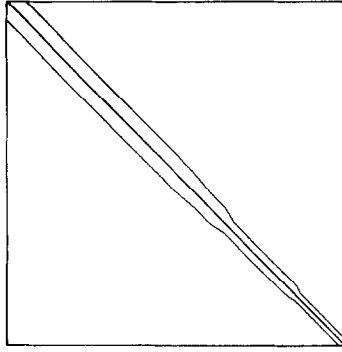


FIG. 2. Structure of the f.e. matrix for the test problem with $N = 812$ (integration of the unsteady diffusion equation in a 3D axisymmetric system).

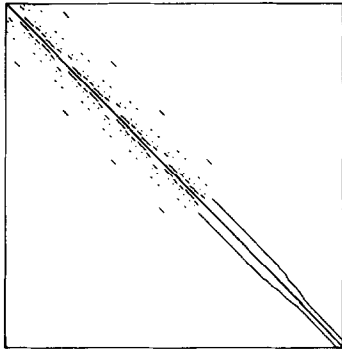


FIG. 3. Structure of the f.e. matrix for the test problem with $N = 1802$ (integration of the elastic equilibrium equations in a 3D axisymmetric system).

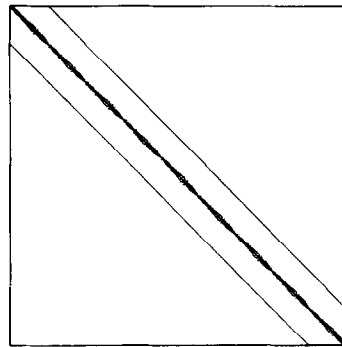


FIG. 4. Structure of the f.e. matrix for the test problem with $N = 2304$ (integration of the steady flow equation in a 3D multiaquifer system).

The convergence is monitored by the use of two quantities

$$e_{k,j} = \left| \frac{q_j(x_k) - \lambda_{N-j}}{\lambda_{N-j}} \right| \quad (16)$$

$$r_{k,j} = |A_j x_k - q_j(x_k) x_k| / |x_k|, \quad (17)$$

where $q_j(x_k) = x_k^T A_j x_k / x_k^T x_k$, $e_{k,j}$ is the relative error of the k th approximation $q_j(x_k)$ to λ_{N-j} and $r_{k,j}$ is the residual for the estimate x_k of v_{N-j} . In Eq. (16) λ_{N-j} is the $(N-j)$ th characteristic root computed to the double precision (16 decimal digits) machine accuracy, while in formula (17) $|\cdot|$ stands for the Euclidean norm.

Since the eigenpairs already determined are currently employed by the shifting process and the convergence rate of the eigenvalue approximation is twice that of the eigenvector approximation, each iterative MCG cycle is completed when $e_{k,j}$ and $r_{k,j}$ have achieved a rather small value.

Figures 5, 6, 7, and 8 show the behavior of $e_{k,j}$ vs the number of iterations k for several j values and the matrices with $N = 156, 812, 1802,$ and 2304 , respectively. Note that $j = 0$ means the minimal eigenvalue of the original matrix A . The dis-

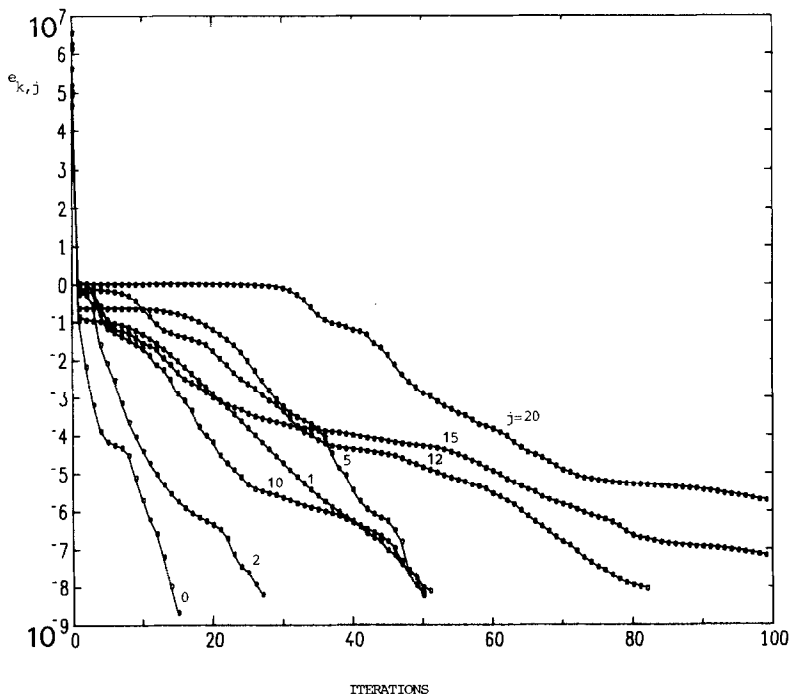


FIG. 5. Relative errors vs number of iterations in the computation of the leftmost $p + 1$ (with $p = 20$) eigenvalues of the f.e. matrix with $N = 156$.

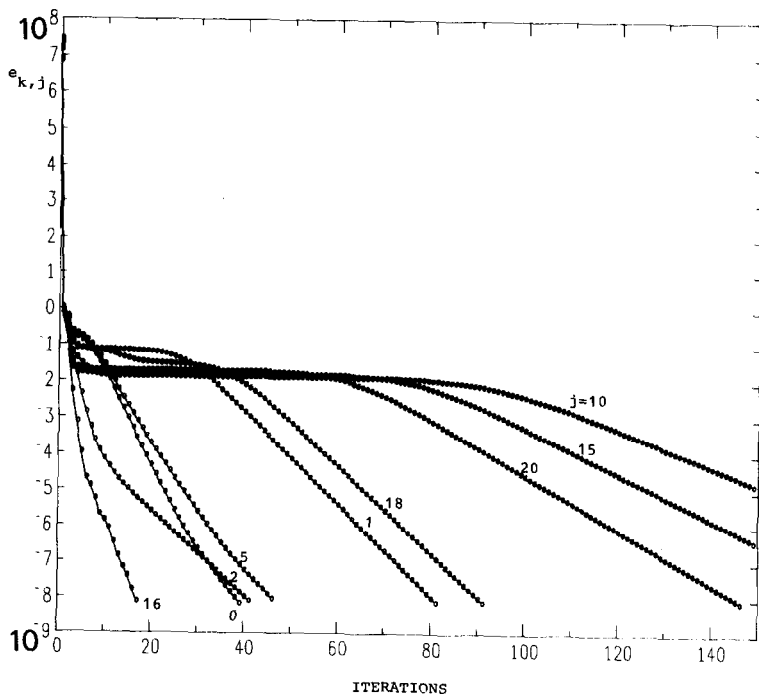


FIG. 6. The same as Fig. 5 for the f.e. matrix with $N = 812$ ($p = 20$).

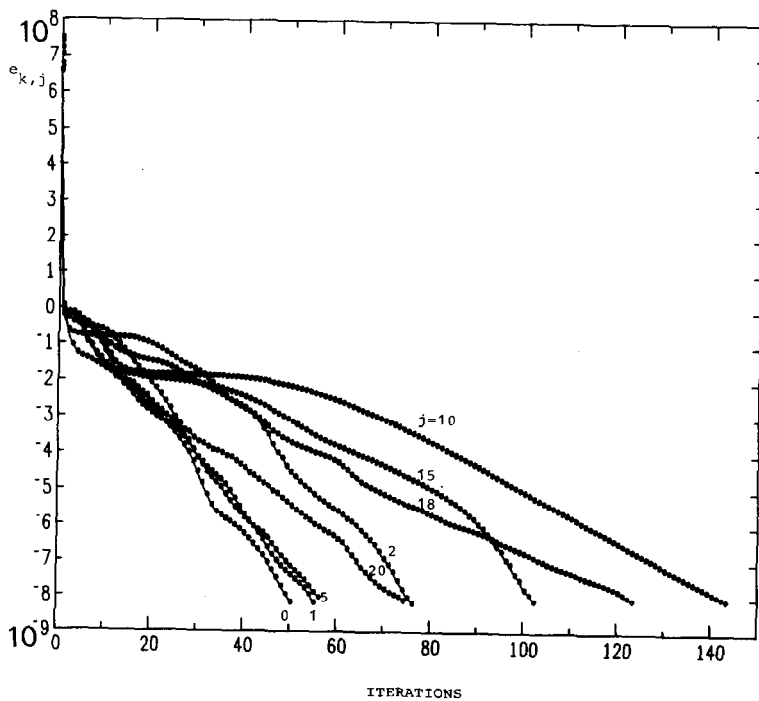


FIG. 7. The same as Fig. 5 for the f.e. matrix with $N = 1802$ ($p = 20$).

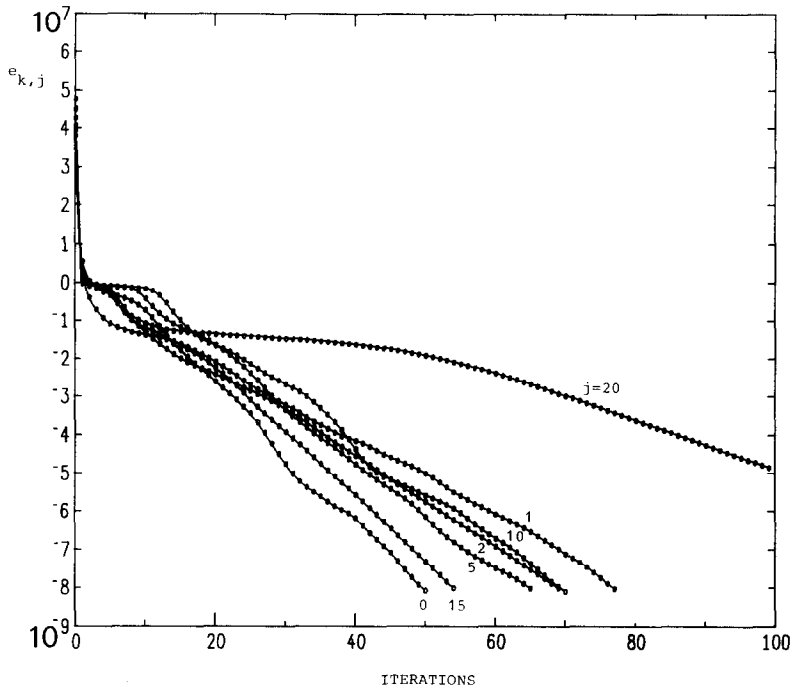


FIG. 8. The same as Fig. 5 for the f.e. matrix with $N = 2304$ ($p = 20$).

tribution of the smallest eigenvalues of A thus found is given in Table I which also reports the β -values operatively adopted to construct $\alpha_j^{(i)}$ according to criterion (15). We note that $\alpha_j^{(0)}$ was sufficient for almost all eigenpairs and that $i = 1$ was necessary in very few cases. The convergence rate is sensitive to β as is shown in Figs. 9, 10, and 11, obtained with $\beta = 1.25$, 1.7, and 5, respectively, for the matrix with $N = 156$. (A similar result is given in Fig. 5 for $\beta = 2.2$.)

The best convergence is that displayed by Fig. 10. Figure 9 points out that, if β is too small, there may be the accumulation of the shifted eigenpairs in the vicinity of the wanted solution with a very slow initial convergence (e.g., $j = 5, 10, 15$). Figure 11 points out that, if β is too large, we may have a slow convergence for the largest j 's, where $(LL^T)^{-1}$ no longer constitutes an effective preconditioning matrix for the CG scheme.

Table II gives some representative results about the computational burden for the determination of an appropriate $\alpha_j^{(i)}$ value for various β 's as compared to the overall number of iterations needed to find the desired eigenvalue λ_{N-j} ($N = 156$, $j = 1, \dots, 5, 7, 10, 15, 20$). We have set the number of trial iterations $I_j = 10$, the tolerance $T_d = 10^{-2}$, and the restart vector $w = (1, \dots, 1)^T$.

Table II shows that a satisfactory value for β , which is problem dependent, is to be determined experimentally. Fortunately the procedure appears to be quite robust

TABLE I
Distribution of the 21 Lowest Eigenvalues of the F.E. Matrices Analyzed by the Stepwise Deflation-MCG Procedure

N	β	λ_N	λ_{N-1}	λ_{N-2}	λ_{N-3}	λ_{N-4}	λ_{N-5}	λ_{N-7}	λ_{N-10}	λ_{N-15}	λ_{N-20}
156	2.2	1.06×10^{-8}	2.09×10^{-8}	2.93×10^{-8}	3.49×10^{-8}	4.88×10^{-8}	8.73×10^{-8}	1.38×10^{-7}	2.39×10^{-7}	4.33×10^{-7}	8.04×10^{-7}
812	1.9	1.69×10^{-8}	1.85×10^{-8}	1.99×10^{-8}	2.12×10^{-8}	2.22×10^{-8}	2.35×10^{-8}	2.64×10^{-8}	2.80×10^{-8}	3.19×10^{-8}	6.41×10^{-8}
1802	1.8	1.70×10^3	2.18×10^3	3.38×10^3	3.88×10^3	4.93×10^3	5.12×10^3	7.11×10^3	9.67×10^3	1.26×10^4	1.53×10^4
2304	2	2.45×10^{-7}	4.33×10^{-7}	4.89×10^{-7}	5.88×10^{-7}	7.28×10^{-7}	7.96×10^{-7}	9.53×10^{-7}	1.25×10^{-6}	1.72×10^{-6}	2.36×10^{-6}

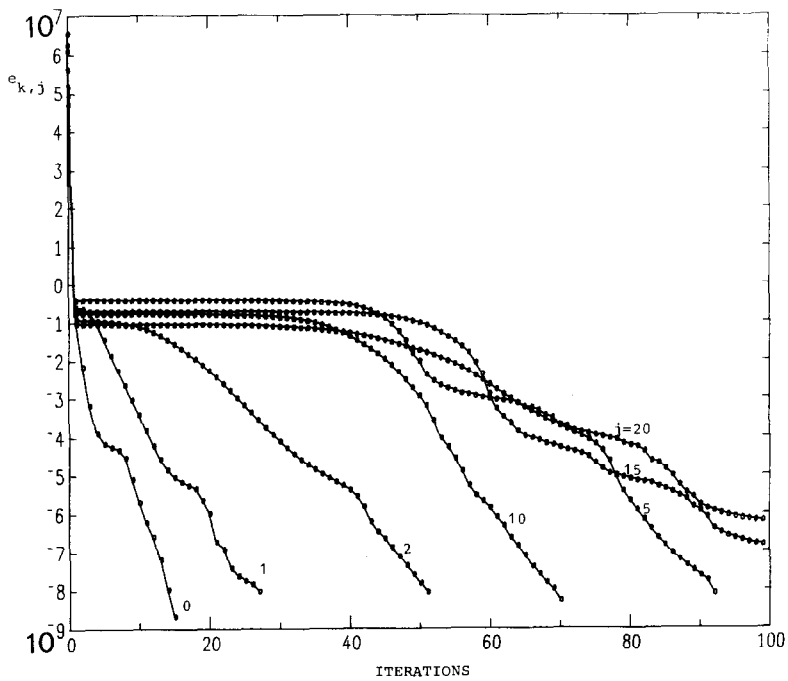


FIG. 9. Convergence profiles for several of the 20 leftmost eigenvalues of the f.e. matrix with $N = 156$ for $\beta = 1.25$.

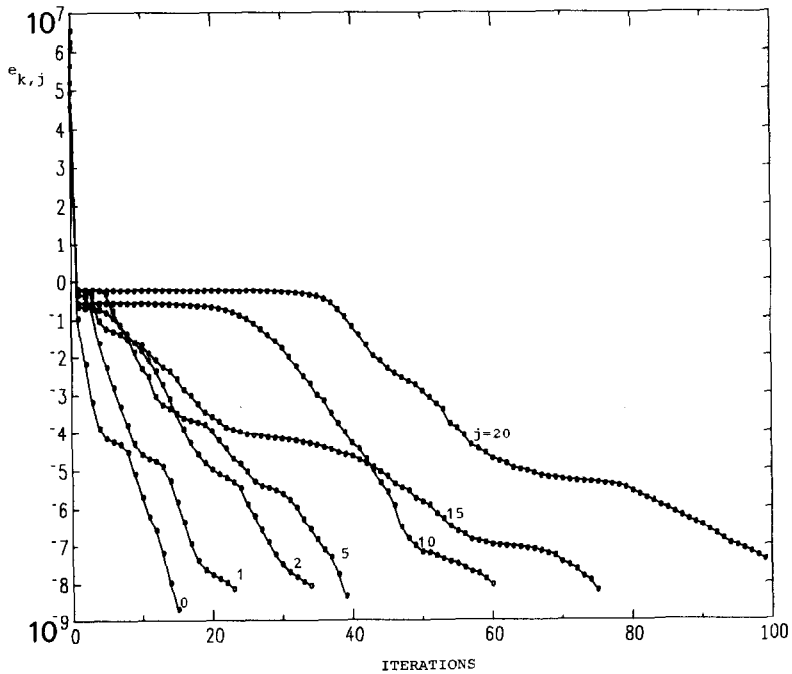


FIG. 10. Convergence profiles for several of the 20 leftmost eigenvalues of the f.e. matrix with $N = 156$ for $\beta = 1.7$.

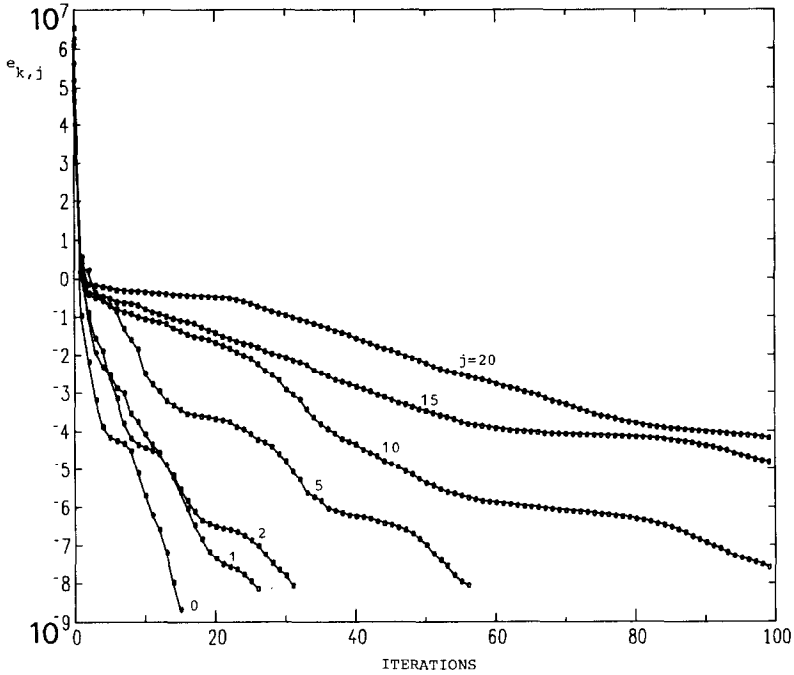


FIG. 11. Convergence profiles for several of the 20 leftmost eigenvalues of the f.e. matrix with $N = 156$ for $\beta = 5$.

even if the intermediate β values are characterized by a fairly smaller number of cumulative iterations (last column of Table II). A few experiments may suggest good indications as to the right β . For instance $\beta = 2$ proved appropriate for all the finite element matrices analyzed in the present paper.

TABLE II

Number of Iterations a Which Proved Necessary to Recognize That an Appropriate Value for $\alpha_j^{(i)}$ Was Obtained and Total Number of Iterations t Needed Both to Find the Right $\alpha_j^{(i)}$ and to Compute the Eigenvalue λ_{N-j} for Several β and j ($N = 156$)

β		λ_{N-1}	λ_{N-2}	λ_{N-3}	λ_{N-4}	λ_{N-5}	λ_{N-7}	λ_{N-10}	λ_{N-15}	λ_{N-20}	Σ
1.25	a	42	37	37	35	48	39	25	23	45	—
	t	69	88	69	90	140	80	95	135	173	939
1.70	a	24	16	15	13	15	13	12	3	7	—
	t	47	50	44	52	54	49	72	78	124	570
2.20	a	12	10	10	13	10	3	3	2	6	—
	t	62	37	33	40	60	51	54	117	125	579
5.00	a	2	3	3	3	2	2	2	2	2	—
	t	28	34	50	47	58	144	110	178	202	851

The most favorable convergence conditions would theoretically occur if $\alpha_j^{(0)} = \lambda_{N-j}$, as in this case we would have the accumulation right in λ_{N-j} and A_j as close as possible to A . It may be interesting to derive "a posteriori" the corresponding profiles. These are shown in Fig. 12 for $N = 156$ and are actually better than the equivalent profiles of Figs. 9, 10, and 11. To obtain the results of Fig. 12, $\alpha_j^{(0)} = \lambda_{N-j}$ has been used along with the eigenvectors $v_N, v_{N-1}, \dots, v_{N-j+1}$ already computed.

In this limiting case the behavior of the spectral condition number of $E_j = A_j K^{-1}$ is provided in Fig. 13. Note that generally $\xi(E_j)$ increases with the deflation level j , after an initial nearly constant behavior. This is an evidence that $(LL^T)^{-1}$ is still a good approximation to A_j^{-1} for several j values. For $N = 2304$ $\xi(E_j)$ is practically constant up to $p = 20$ and for $N = 1802$ it grows very slowly. It may therefore be concluded that the preconditioning discussed in the present paper may be expected to be effective for several steps of the deflation process.

Now we elaborate a little more on the implementation of scheme (14), on the automatic determination of the right $\alpha_j^{(i)}$ and on the effectiveness of our preconditioning procedure.

Figure 14 gives some possible practical behaviors of $r_{k,j}^{(i)}$ and is drawn for $j = 2$, $\beta = 1.25$, and $N = 156$ (for a cross reference see also Fig. 9 and Table II).

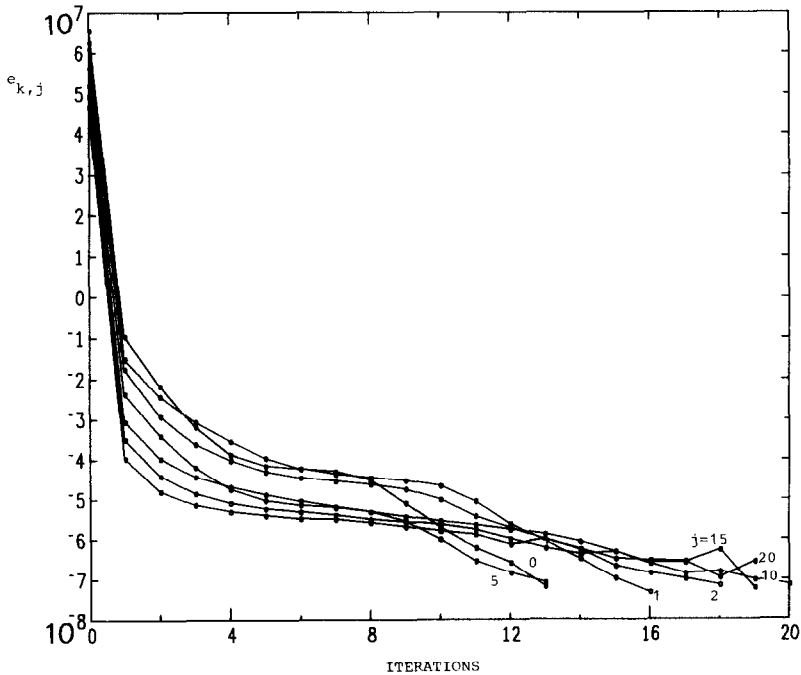


FIG. 12. Convergence profiles for several of the 20 leftmost eigenvalues of the f.e. matrix with $N = 156$ when $\alpha_j^{(0)} = \lambda_{N-j}$ is assumed ($j = 0, 1, \dots, 20$).

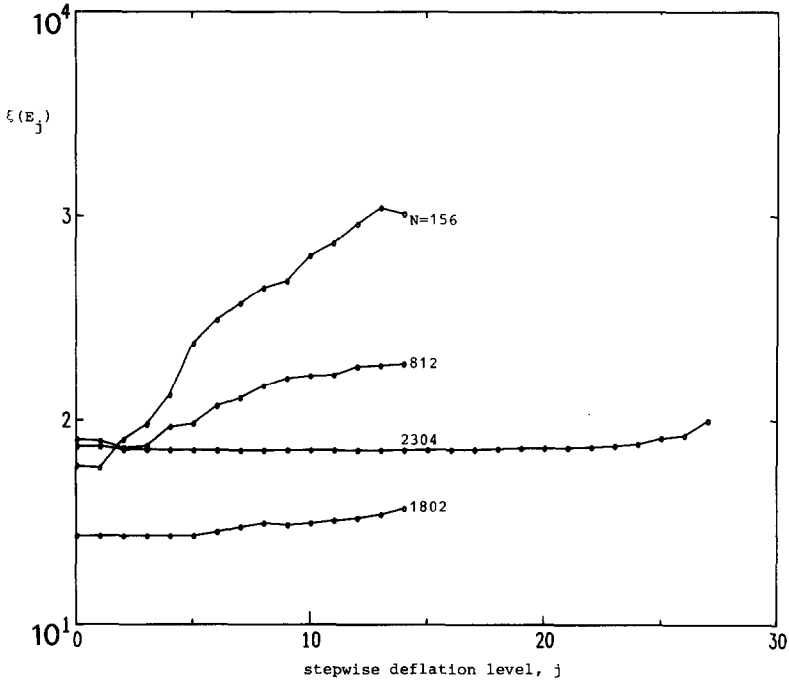


FIG. 13. Behavior of the spectral condition number of the matrices $E_j = A_j K^{-1}$ when $\alpha_j^{(0)} = \lambda_{N-j}$ is assumed ($j=0, 1, \dots, 20$). In real calculations $\xi(E_j)$ is equal or greater than the values given in this figure.

If we set $x_0 = v_{N-j+1}$, as long as $\alpha_j^{(i)} < \lambda_{N-j}$ the residual (17) remains approximately constant with the iteration k and typically exhibits small oscillations (profile (a) of Fig. 14).

When $\alpha_j^{(i)}$ becomes larger than λ_{N-j} , $r_{k,j}^{(i)}$ behaves like profile (b) of Fig. 14, where three distinct segments may be recognized: initially nearly constant and then an increase followed by a progressive decrease when x_k converges to v_{N-j} (in principle $r_{k,j}^{(i)}$ should always be zero as $\alpha_j^{(i)}$ and v_{N-j+1} identify a stationary point; in practice due to roundoff errors the behavior displayed by profile (b) occurs and the iterates x_k moves progressively far from v_{N-j+1} to converge eventually to v_{N-j}). As soon as we note that $r_{k,j}^{(i)}$ increases (second type segment of profile (b) in Fig. 14) the iteration is restarted with a different choice for x_0 (e.g., x_0 equal to $(1, \dots, 1)^T$). Then the overall convergence is accelerated as shown by profile (c) of Fig. 14.

Finally we provide some documentary evidence about the advantage of preconditioning the CG scheme. Table III gives the spectral condition number ξ of the Hessian (8) when K^{-1} is set equal to I (identity matrix), $(LL^T)^{-1}$ (L being the incomplete Cholesky factor of A), and A^{-1} , for the computation of the minimal eigenpair λ_N and v_N . We remind that the choice $K = I$ is equivalent to applying the ordinary CG iteration. It may be observed that $\xi(H)$ decreases drastically when CG is preconditioned with $K^{-1} = (LL^T)^{-1}$. A further decrease occurs if K^{-1} equals

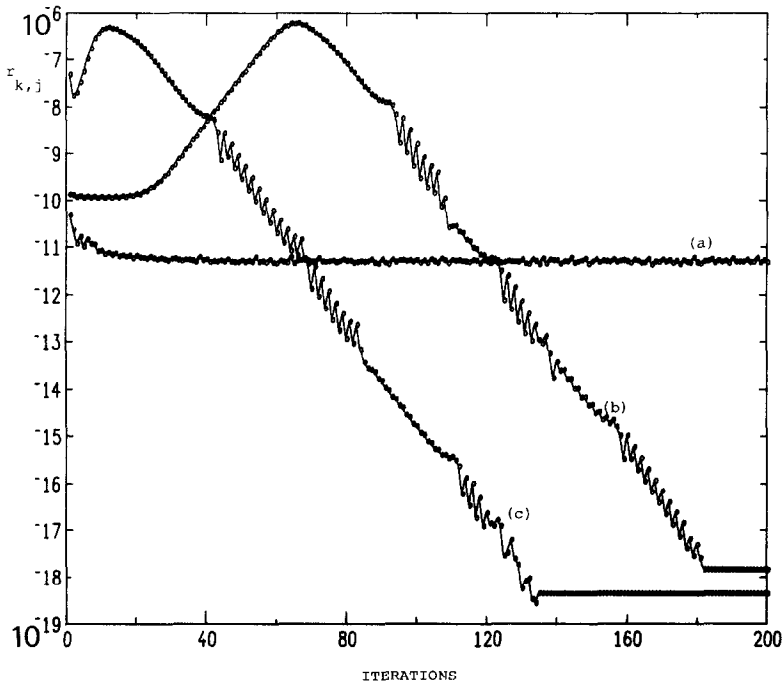


FIG. 14. Behavior of the residual $r_{k,j}^{(l)}$ vs the number of iterations in the search for the $(N-j)$ th eigenpair of the f.e. matrix with $N=156$: $x_0 = v_{N-j+1}$ and $\alpha_j^{(0)} < \lambda_{N-j}$ (profile a); $x_0 = v_{N-j+1}$ and $\alpha_j^{(1)} > \lambda_{N-j}$ (profile b); $x_0 = (1, \dots, 1)^T$ and $\alpha_j^{(1)} > \lambda_{N-j}$ (profile c). The values for j and β are 2 and 1.25, respectively.

A^{-1} . Table III shows that the incomplete Cholesky factorization represents a quite effective preconditioning procedure for f.e. eigenproblems and is expected to enhance greatly the performance of the traditional conjugate gradients. It also gives a satisfactory account of the accelerated convergence profiles reported in Figs. 5, 6, 7, and 8.

TABLE III

Spectral Condition Number of the Hessian (8) for Different Preconditioning Matrices K^{-1} in the Computation of the Minimal Eigenpair λ_N and v_N

N	$\xi(H)$		
	$K^{-1} = I$	$K^{-1} = (LL^T)^{-1}$	$K^{-1} = A^{-1}$
156	0.97×10^8	0.60×10^2	0.20×10^1
812	0.19×10^{12}	0.45×10^2	0.12×10^2
1802	0.21×10^{10}	0.69×10^2	0.45×10^1
2304	0.53×10^7	0.11×10^3	0.23×10^1

4. CONCLUSIONS

The following points are worth summarizing:

(1) An efficient stepwise deflation procedure and an accelerated conjugate gradient technique have been developed and applied to the numerical computation of the p ($p \leq 20$) leftmost eigenpairs of symmetric p.d. matrices arising from finite element problems of structural mechanics and subsurface hydrodynamics.

(2) At the j th step of the deflation process, the eigenpairs already evaluated are shifted in a unique stationary point to the right but not too close to the new minimal eigensolution λ_{N-j} and v_{N-j} of the deflated matrix A_j . To ensure an effective CG preconditioning the shift $\alpha_j^{(i)}$ should not occur far beyond the desired value λ_{N-j} .

(3) The selection of an appropriate parameter β for the iterative updating of the values for the shift $\alpha_j^{(i)}$ should not prove a difficult task. However, some preliminary trials are recommended in order to get a very rough acquaintance with the actual values and distribution of the leftmost characteristic roots of the specific eigenproblem to be solved.

(4) The real convergence rate of the MCG scheme to the minimal stationary point of a specific Rayleigh quotient $q_j(x)$ depends on the shifting value $\alpha_j^{(i)}$ and also on the separation between λ_{N-j} and the next higher eigenvalues. If the separation is not too bad and $\alpha_j^{(i)}$ is far enough from λ_{N-j} , the MCG iterations should increase with the deflation level j , namely with the parallel increase of the spectral condition number $\xi(A_j K^{-1})$. However, the increase is not necessarily monotonic with j depending ultimately on the actual distribution of the eigenvalues close to λ_{N-j} . The f.e. experiments have shown that $\xi(A_j K^{-1})$ grows slowly with j , i.e., $(LL^T)^{-1}$ appears to be a satisfactory preconditioning matrix for the CG algorithm applied to the computation of the smallest stationary point of $q_j(x)$, $j=0, \dots, p$, with $p \leq 20$.

(5) It might be worth comparing the computational cost of the stepwise iterative approach developed in the present paper with alternative procedures based on the simultaneous evaluation of the p leftmost eigenpairs (Longsine and McCormick [2, 3]; Sameh and Wisniewski [22]; Schwarz [24]). Further investigations in this direction are currently under way at the University of Padua.

REFERENCES

1. B. N. PARLETT, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980), p. 348.
2. M. R. HESTENES AND W. KARUSH, *J. Res. Nat. Bur. Stand.* **47**, 45 (1951).
3. M. R. HESTENES AND W. KARUSH, *J. Res. Nat. Bur. Stand.* **47**, 471 (1951).
4. M. R. HESTENES, *Appl. Math. Series, Nat. Bur. Stand.* **29**, 89 (1953).
5. D. K. FADDEEV AND V. N. FADDEEVA, *Computational Methods of Linear Algebra* (Freeman, San Francisco, 1963).
6. R. K. NESBET, *J. Chem. Phys.* **43**, 311 (1965).

7. C. F. BENDER AND I. SHAVITT, *J. Comput. Phys.* **6**, 146 (1970).
8. I. SHAVITT, C. F. BENDER, A. PIPANO, AND R. P. HOSTENY, *J. Comput. Phys.* **11**, 90 (1973).
9. S. FALK, *Z. Angew. Math. Mech.* **53**, 73 (1973).
10. R. M. NISBET, *J. Comput. Phys.* **10**, 614 (1972).
11. A. RUHE, *Math. Comput.* **28**, 695 (1974).
12. H. R. SCHWARZ, *Comput. Methods Appl. Mech. Eng.* **3**, 11 (1974).
13. M. R. HESTENES AND E. STIEFEL, *J. Res. Nat. Bur. Stand.* **49**, 409 (1952).
14. W. W. BRADBURY AND R. FLETCHER, *Numer. Math.* **9**, 259 (1966).
15. I. FRIED, *AIAA J.* **7**, 739 (1969).
16. I. FRIED, *J. Sound Vibrat.* **20**, 333 (1972).
17. M. GERARDIN, *J. Sound Vibrat.* **19**, 319 (1971).
18. A. RUHE, *Sparse Matrix Techniques*, edited by V. A. Barker (Springer-Verlag, New York, 1977), p. 130.
19. M. PAPADRAKAKIS, *Int. J. Numer. Methods Eng.* **20**, 2283 (1984).
20. A. RUHE, *J. Comput. Phys.* **19**, 110 (1975).
21. D. J. EVANS AND J. SHANEHCHI, *Comput. Methods App. Mech. Eng.* **31**, 251 (1982).
22. A. H. SAMEH AND J. A. WISNIEWSKI, *SIAM J. Numer. Anal.* **19**, 1243 (1982).
23. D. E. LONGSINE AND S. F. MCCORMICK, *Linear Algebra Appl.* **34**, 195 (1980).
24. H. R. SCHWARZ, *Numerical Integration of Differential Equations and Large Linear Systems*, edited by J. Hinze (Springer-Verlag, Berlin, 1982), p. 384.
25. D. S. KERSHAW, *J. Comput. Phys.* **26**, 43 (1978).
26. P. CONCUS, G. H. GOLUB, AND D. P. O'LEARY, *Sparse Matrix Computations*, edited by J. R. Bunch and D. J. Rose (Academic Press, New York, 1976), p. 309.
27. J. A. MEIJERINK AND H. A. VAN DER VORST, *Math. Comput.* **31**, 148 (1977).
28. O. AXELSSON, *Sparse Matrix Techniques*, edited by V. A. Barker (Springer-Verlag, New York, 1977), p. 1.
29. G. GAMBOLATI, *Int. J. Numer. Methods Eng.* **15**, 661 (1980).
30. G. GAMBOLATI, *IVth Int. Conf. Finite Elements Water Resources* (University of Mississippi, 1980), p. 215.
31. T. A. MANTEUFFEL, *Math. Comput.* **24**, 473 (1980).
32. G. GAMBOLATI AND G. VOLPI, *Information Processing 80* (North-Holland, Amsterdam, 1980), p. 729.
33. O. AXELSSON, *Comput. Methods Appl. Mech. Eng.* **9**, 123 (1976).
34. S. C. EISENSTAT, *SIAM J. Sci. Stat. Comput.* **2**, 1 (1981).
35. B. NOUR-OMID, *Proc. Int. Conf. on "Innovative Methods for Nonlinear Problems,"* edited by W. K. Lin, T. Belytschko, and K. C. Park (Pineridge, Swansea, UK, 1984), p. 17.
36. M. A. AJIZ AND A. JENNINGS, *Int. J. Numer. Methods Eng.* **20**, 949 (1984).
37. C. P. JACKSON AND P. C. ROBINSON, *Int. J. Numer. Methods Eng.* **21**, 1315 (1985).
38. G. GAMBOLATI AND A. M. PERDON, *Fundamentals of Transport Phenomena in Porous Media*, edited by J. Bear and M. J. Corapcioglu (Nijhoff, Dordrecht, 1984), p. 953.
39. G. GAMBOLATI, *Adv. Water Resour.* **2**, 123 (1979).
40. O. AXELSSON, *Linear Algebra Appl.* **29**, 1 (1980).
41. G. GAMBOLATI AND A. M. PERDON, *Comput. Methods App. Mech. Eng.* **41**, 1 (1983).
42. A. M. PERDON AND G. GAMBOLATI, *Comput. Methods App. Mech. Eng.* **56**, 251 (1986).