JOURNAL OF COMPUTATIONAL PHYSICS 60, 315–337 (1985)

Preconditioned Minimal Residual Methods for Chebyshev Spectral Calculations*

CLAUDIO CANUTO AND ALFIO QUARTERONI

Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Virginia 23665, and Istituto di Analisi Numerica del CNR, 27100 Pavia, Italy

The problem of preconditioning the pseudospectral Chebyshev approximation of an elliptic operator is considered. The numerical sensitivity to variations of the coefficients of the operator are investigated for two classes of preconditioning matrices: one arising from finite differences, the other from finite elements. The preconditioned system is solved by a conjugate gradient type method, and by a DuFort–Frankel method with dynamical parameters. The methods are compared on some test problems with the Richardson method [13] and with the minimal residual Richardson method [21].

1. INTRODUCTION

Spectral methods are today widely used in the numerical solution of a variety of boundary value problems (e.g., [9, 14, 16]). Their better accuracy makes them preferable to the more classical finite difference or finite element methods in many situations, especially for smooth problems (see [14, Chaps. 3 and 8]).

Improving the computational efficiency of spectral methods is one of the main purposes of the current investigation on these methods. In this direction an important aspect is the definition of spectral approximations in "complex" geometries: different techniques of mapping and patching have been proposed to reduce the computational domain to a simple domain, such as a square or a cube. Another relevant aspect is to find efficient techniques of solution of the algebraic systems arising from spectral methods. Actually, spectral approximations of elliptic boundary value problems lead to full and very ill-conditioned matrices. In the special case of constant coefficient operators, efficient direct methods which exploit the orthogonality of the spectral basis have been proposed [9, 10, 22]. For nonconstant coefficient problems, direct methods are impractical, because of the typically large size of the systems to be soved. Instead, considerable attention has been devoted after Orszag's paper [13] to the simultaneous use of iterative methods and preconditioning techniques.

* Research was supported by the National Aeronautics and Space Administration under NASA Contract NAS1-17070 while the authors were in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665 and by the Istituto di Analisi Numerica del CNR, Pavia, Italy.

0021-9991/85 $3.00

Actually, variable coefficient operators can be evaluated in a spectral way using fast Fourier transform techniques. Therefore, iterative methods of solution which require the evaluation of the residual at each iteration are particularly well suited for spectral approximations.

In the present paper, we present and discuss the results of a number of numerical tests on the iterative solution of preconditioned systems arising from Chebyshev approximations. The first part is devoted to the analysis of the preconditioning of spectral matrices. The sensitivity to variations of the coefficients, to leading and lower order terms is investigated. Besides the standard finite difference matrix proposed in [13], we consider a finite element matrix, which essentially retains the same preconditioning properties, being moreover symmetric. In both cases, an incomplete factorization satisfying the row-sum agreement criterion is used.

Some iterative methods are considered next. The aim of our analysis is to investigate their behavior and performances when applied to the solution of algebraic systems arising from approximations of spectral type. Therefore, they will be compared with other iterative methods already widely used for the same systems. A preconditioned conjugate gradient method (which has been recently used in fluid dynamics and transonic flow calculations via finite elements, see [6] and the references therein) was found to be rather slow on the tested problems, although it may be very robust in more complicated situations. The DuFort–Frankel method (first applied by Gottlieb *et al.* [7, 8] to spectral calulations and here considered as a 2-parameter preconditioned iterative method) yields good results when the optimal parameters are used. In order to overcome the difficulty of finding such parameters, we propose a modified version of the DuFort–Frankel method, devised according to a "minimal residual" strategy. The new method, compared with other iterative techniques in the literature, was the fastest in terms of speed of convergence.

In the paper these methods are tested on linear elliptic problems. However, the interest of this investigation goes far beyond such kind of test problems. Indeed, more complicated problems, such as the Navier–Stokes equations, are reduced in most of the current spectral approximations to a sequence of linear elliptic problems, whose resolution can be carried out by the methods proposed here.

Iterative methods coupled with suitable preconditioning techniques can dramatically reduce the gap in computational efficiency between spectral methods and more conventional methods such as finite differences or finite elements. The extra cost required by spectral methods is, however, more than compensated by the superior accuracy they produce. This means that the accuracy required in most applications is attained by spectral methods with a considerably lower number of unknowns, hence with globally less computational costs.

Part of this work has been made while the authors were visiting the Institute for Computer Applications in Science and Engineering (ICASE). The numerical results reported here were obtained on the Honeywell 6040 at the University of Pavia. Programs were written in double precision. The eigenvalues of Section 2 were obtained by EISPACK routines.

2. The Preconditioning of Spectral Matrices

Let L be a smooth second-order elliptic partial differential operator over the interval $\Omega^1 = (-1, 1)$ or the square $\Omega^2 = (-1, 1)^2$. We consider homogeneous Dirichlet boundary conditions for L, i.e., functions on which L acts will be assumed to vanish identically on the boundary. L_{sp} will denote the Chebyshev pseudospectral approximation of L of order N. This means that the approximate solution is a polynomial of degree N and derivatives are computed after interpolating the function by a polynomial of degree N at the Chebyshev nodes ($\{x_j = \cos(\pi j/N)\}$, $j = 0,..., N$ if $\Omega = \Omega^1$; $\{x_i, x_j\}$, $0 \leqslant i, j \leqslant N$ if $\Omega = \Omega^2$). For instance, if

$$Lu = -(a(x)u_x)_x + b(x)u,$$

with $a(x) \geqslant a_0 > 0$, $b(x) \geqslant 0$, then,

$$L_{sp}u = -\frac{d}{dx} I_N(au_x) + I_N(bu),$$

where $I_N(v)$ is the algebraic polynomial of degree N which interpolates v at the Chebyshev nodes x_j, $j = 0,..., N$. Since u is a polynomial of degree N, vanishing at the boundary, it is uniquely determined through its values at the interior Chebyshev points. These are the unknowns we are solving for. Thus the linear mapping

$$\{u(x_j)\}_{j=1,...,N-1} \to \{(L_{sp}u)(x_i)\}_{i=1,...,N-1}$$

defines a matrix of order $(N-1)$, which we still denote by L_{sp}. We identify L_{sp} with the matrix which maps the set of values of a polynomial u at the interior Chebyshev nodes into the set of values of the spectral approximation of Lu at the same nodes. The evaluation of $L_{sp}u$ can be done efficiently by computing each derivative at the Chebyshev nodes via fast Fourier transform methods (see, e.g., [9]).

It is known that L_{sp} has a full structure. Moreover, its condition number is $O(N^4)$ [9]. These are considered negative aspects of spectral Chebyshev approximations versus finite difference and finite element methods. However, a tremendous improvement in the computational efficiency of spectral methods comes from the observation that L_{sp} can be easily approximated by a sparse matrix A, such that the condition number of the matrix $A^{-1}L_{sp}$ is close to 1 (see Orszag [13]). Recall that the rate of convergence of an iterative method increases as the condition number of the matrix is closer to 1. Throughout the paper we take the ratio $\kappa = \kappa(M) = |\lambda_{max}|/|\lambda_{min}|$ as a measure of the condition number of the matrix M, and we refer to it improperly as the condition number even when M is not symmetric.

In the following, A will denote any matrix having these properties, and it will be called a preconditioning matrix. A is assumed to be related to some discretization of the operator L, usually by finite differences or finite elements. Sometimes we shall relate A to some other elliptic operator $\mathscr{L}$, with the same principal part as L.

In one space dimension, the simplest way of building a preconditioning matrix is to use non-equally spaced finite differences at the Chebyshev nodes. The resulting matrix is tridiagonal, and it can be factorized in $O(N)$ operations. If $Lu = -u_{xx}$, the corresponding preconditioning matrix is given by $A = \{a_{ij}\}$, where

$$a_{jj} = \frac{2}{h_j h_{j-1}}; \qquad a_{j,j-1} = \frac{-2}{h_{j-1}(h_j + h_{j-1})};$$
$$a_{j,j+1} = \frac{-2}{h_j(h_j + h_{j-1})}; \qquad h_j = x_j - x_{j+1}. \tag{2.1}$$

In Table I, the operator $Lu = -u_{xx}$ is considered. The smallest and the largest eigenvalue $\lambda_{\min}$ and $\lambda_{\max}$, and the ratio $\kappa = \lambda_{\max}/\lambda_{\min}$ are reported for both the matrices L_{sp} and $A^{-1}L_{sp}$. As shown in [22], the largest eigenvalue of L_{sp} grows like N^4, while the eigenvalues of the preconditioned matrix $A^{-1}L_{sp}$ lie in the interval $[1.0, \pi^2/4]$. The spectrum of $A^{-1}L_{sp}$ exhibits a similar behavior even if the elliptic operator L contains lower order terms (see [13]).

The preconditioning properties of the matrix A seem to be rather insensitive to the lower order terms of L. Table II shows that the condition number $\kappa(A^{-1}L_{sp})$ is kept small when A is just the finite difference approximation of the second-order term of L, and the lower order terms are not prevailing. This implies that the preconditioning matrix can be kept fixed in solving nonlinear problems in which the lower order terms only change during the iterations. In all the cases considered below, the smallest eigenvalue $\lambda_{\min}$ is close to 1, and it converges to 1 from above as N increases.

When the magnitude of the lower order terms is exceedingly large, the condition number of $A^{-1}L_{sp}$ deteriorates. However, the spectrum is still uniformly bounded in N, as shown in Table III.

A family of variable coefficient operators $Lu = -(\alpha u_x)_x$, with $0 < \alpha_0 \leqslant \alpha(x) \leqslant \alpha_1$, is considered in Table IV. The eigenvalues of the matrix $A^{-1}L_{sp}$ are bounded

TABLE I

$Lu = -u_{xx}$; Au = Finite Differences at Chebyshev Points for Lu

	L_{sp}			$A^{-1}L_{sp}$		
N	$\lambda_{\min}$	$\lambda_{\max}$	κ	$\lambda_{\min}$	$\lambda_{\max}$	κ
4	2.46	0.20 *E*2	0.80 *E*1	1.0	1.75	1.75
8	2.47	0.21 *E*3	0.87 *E*2	1.0	2.13	2.13
16	2.47	0.32 *E*4	0.13 *E*4	1.0	2.30	2.30
32	2.47	0.50 *E*5	0.20 *E*5	1.0	2.38	2.38
64	2.47	0.80 *E*6	0.32 *E*6	1.0	2.43	2.43
128	2.47	0.13 *E*8	0.52 *E*7	1.0	2.45	2.45

TABLE II

Condition Number $\kappa(A^{-1}L_{sp})$

N	$\delta = 0.0$ $\gamma = 1.0$	$\delta = 0.0$ $\gamma = 10.0$	$\delta = 1.0$ $\gamma = 0.0$	$\delta = 10.0$ $\gamma = 0.0$	$\delta = 10.0$ $\gamma = 10.0$
4	1.38	2.25	1.45	2.29	1.61
8	1.87	3.45	1.90	2.01	1.96
16	2.16	4.30	2.15	2.28	2.41
32	2.32	4.73	2.31	2.47	2.82
64	2.40	4.92	2.38	2.85	3.25
128	2.44	4.99	2.43	3.08	3.54

Note. $Lu = -u_{xx} + \delta u_x + \gamma u$; Au = finite differences for $\mathscr{L}u = -u_{xx}$.

TABLE III

Condition Number $\kappa(A^{-1}L_{sp})$

N	$\delta = 0.0$ $\gamma = 100.0$	$\delta = 100.0$ $\gamma = 0.0$	$\delta = 1000.0$ $\gamma = 0.0$
4	3.95	21.76	217.43
8	17.70	19.55	195.39
16	28.90	18.08	181.05
32	35.89	19.18	177.83
64	39.17	21.55	176.24
128	40.57	24.32	204.05

Note. $Lu = -u_{xx} + \delta u_x + \gamma u$; Au = finite differences for $\mathscr{L}u = -u_{xx}$.

TABLE IV

$Lu = -((1 + 10^{v}x^2)u_x)_x$; Au = Finite Differences for Lu

	$v = 0$		$v = 1$		$v = 2$	
N	λ_{min}	κ	λ_{min}	κ	λ_{min}	κ
4	1.04	2.49	1.11	5.03	1.13	7.09
8	1.01	3.04	1.10	7.11	1.01	14.77
16	1.00	3.27	1.00	7.70	1.00	21.65
32	1.00	3.38	1.00	7.94	1.00	24.03
64	1.00	3.43	1.00	8.06	1.00	24.44
128	1.00	3.46	1.00	8.12	1.00	24.61

independently of N, although the bound is larger than for the constant coefficient operator. The condition number κ is close to the one in Table I when a moderate perturbation is applied, otherwise it grows slowly and linearly with the total variation of α.

When the coefficient α depends itself on the solution u (as in the full potential equation) one would not change the preconditioning matrix at each iteration, in order to save factorization time. This situation is simulated to a certain extent in Table V. The effects of preconditioning the spectral matrix of a variable coefficient operator by a constant coefficient operator matrix are reported. The spectrum of $A^{-1}L_{sp}$ is bounded independently of N. κ is comparable with the one of Table IV when the perturbation is moderate, but it becomes noticeably worse when the distance between the preconditioning and the spectral operators increases. In this case, if the factorization is carried out in a number of operations of the order of the number of unknowns (in one dimension or with an incomplete factorization), the worsening of the condition number may not be balanced by the saving in factorization time (unless the computation of the entries of A is particularly expensive).

The matrices A considered so far arise from a finite difference approximation of the operator $\mathscr{L}$ at the Chebyshev points. Even if $\mathscr{L}$ is formally self-adjoint, A is not symmetric, nor is L_{sp}. Actually A splits up as $A = D \cdot \tilde{A}$, with D diagonal and $\tilde{A}$ symmetric. Some iterative techniques require the symmetry of the preconditioning matrix (see Sect. 3). This can be accomplished by discretizing a suitable variational formulation of the elliptic operator via finite elements as follows.

If $\mathscr{L}u = -(\alpha u_x)_x$, the bilinear form associated to $\mathscr{L}$ is

$$a(u, v) = \int_{\Omega^1} \alpha u_x (v\omega)_x \, dx, \tag{2.2}$$

where $\omega(x) = (1 - x^2)^{-1/2}$. The form $a(u, v)$ is continuous and coercive on the

TABLE V

$Lu = -((1 + 10^v x^2)u_x)_x$; Au = Finite Differences for $\mathscr{L}u = -u_{xx}$

	$v = 0$		$v = 1$	
N	λ_{min}	κ	λ_{mon}	κ
4	1.74	2.23	8.51	3.70
8	1.48	3.07	4.70	7.61
16	1.27	3.79	2.71	12.79
32	1.16	4.27	1.83	18.46
64	1.10	4.60	1.43	23.29

weighted Sobolev space $H^1_{0,\omega}(\Omega^1)$ (cf. [2]), but it is not symmetric. However the "reduced" form

$$\tilde{a}(u, v) = \int_{\Omega^1} \alpha u_x v_x \omega \, dx \tag{2.3}$$

is still coercive and continuous on $H^1_{0,\omega}(\Omega^1)$, and trivially symmetric. Assuming u and v continuous and piecewise linear between contiguous Chebyshev knots, we associate a matrix $A = \{a_{ij}\}$ to (2.3) by setting

$$a_{ij} = \tilde{a}(\phi_i, \phi_j), \tag{2.4}$$

where ϕ_k is continuous piecewise linear and $\phi_k(x_l) = \delta_{kl}$. For instance, if $\mathcal{L}u = -u_{xx}$, we have after dropping the common factor π/N:

$$a_{jj} = \frac{1}{h_j^2} + \frac{1}{h_{l-1}^2}; \qquad a_{j,j-1} = \frac{-1}{h_{j-1}^2}; \qquad a_{j,j+1} = \frac{-1}{h_j^2} \tag{2.5}$$

(compare with (2.1)). The spectrum of A behaves like the spectrum of the corresponding finite difference matrix, and the preconditioning properties are only slightly worse, as shown in Table VI.

Up to now we considered 1-dimensional problems. In two dimensions one can still use a finite difference or finite element matrix, say B, in the preconditioning. The corresponding results are similar to those in one dimension. However, the exact "inversion" of such a matrix is more expensive, since the factors in its LU decomposition have a bandwidth of order $O(N)$ instead of $O(1)$. In order to overcome this drawback, different techniques of incomplete factorization have been successfully proposed (cf., e.g., [11, 12, 4]). The idea is to replace the exact factors L and U by some approximations $\tilde{L}$ and $\tilde{U}$, which retain a very sparse structure. $\tilde{L}$ and $\tilde{U}$ are computed by incomplete steps of Gaussian elimination, under the condition that certain quantities depending on the product $\tilde{L}\tilde{U}$ agree with the

TABLE VI

$Lu = -u_{xx}$; Au = Finite Elements at the Chebyshev Points for Lu

N	$\lambda_{\min}(A^{-1}L_{sp})$	$\lambda_{\max}(A^{-1}L_{sp})$	$\kappa(A^{-1}L_{sp})$
4	1.25	3.29	2.63
8	1.16	4.10	3.55
16	1.13	4.53	3.99
32	1.13	4.74	4.19
64	1.13	4.84	4.29
128	1.13	4.89	4.33

TABLE VII

$Lu = -\alpha u_{xx} - u_{yy}$; Au = Incompletely Factorized Finite Difference Matrix for LU

	$\alpha = 1.0$		$\alpha = 10.0$		$\alpha = 100.0$	
N	$\lambda_{\min}$	κ	$\lambda_{\min}$	κ	$\lambda_{\min}$	κ
4	1.08	1.72	1.01	1.75	1.00	1.76
8	1.06	2.72	1.03	2.43	1.01	2.13
16	1.04	4.06	1.03	5.34	1.01	3.27

corresponding quantities for B. The matrix $A = \tilde{L}\tilde{U}$ is then used in the preconditioning.

In our computations, the incomplete factorization was done according to row-sums agreement condition (see [4]). Namely, let $b^{(0)}$ and $b^{(k)}$ denote the diagonal and the off-diagonals of a mth order matrix B, i.e., $b^{(k)} = \{b_{i,i+k} \mid 1 \leqslant i, i+k \leqslant m\}$. If B is the five point finite difference matrix for a second-order operator at the Chebyshev points in the square, then only $b^{(0)}$, $b^{(\pm 1)}$, and $b^{(\pm N)}$ are not identically zero. The incomplete factors $\tilde{L}$ and $\tilde{U}$ of B have $\tilde{l}^{(0)}$, $\tilde{l}^{(-1)}$, $\tilde{l}^{(-N)}$ and $\tilde{u}^{(0)}$, $\tilde{u}^{(1)}$, $\tilde{u}^{(N)}$ respectively as nonzero (off)-diagonals. $\tilde{u}^{(0)}$ is chosen to be $\equiv 1.0$, while the off-diagonal elements are recursively determined by the condition that $a^{(\pm 1)} \equiv b^{(\pm 1)}$ and $a^{(\pm N)} \equiv b^{(\pm N)}$, where $A = \tilde{L}\tilde{U}$. Finally, $\tilde{l}^{(0)}$ is such that the sum of each row in A equals the corresponding sum in B. Thus we have, wherever the indices are between 1 and m:

$$\tilde{u}_{i,i+1} = b_{i,i+1}$$

$$\tilde{u}_{i,i+N} = b_{i,i+N}$$

$$\tilde{l}_{i,i-N}\tilde{u}_{i-N,i-N} = b_{i,i-N}$$

$$\tilde{l}_{i,i-1}\tilde{u}_{i-1,i-1} = b_{i,i-1}$$

$$\tilde{l}_{i,i-1}\tilde{u}_{i-1,i} + \tilde{l}_{i,i-N}\tilde{u}_{i-N,i} + \tilde{u}_{i,i} = b_{i,i} - \tilde{l}_{i,i-N}\tilde{u}_{i-N,i-N+1} - \tilde{l}_{i,i-1}\tilde{u}_{i-1,i-N+1}.$$

TABLE VIII

$Lu = -(\alpha u_x)_x - (\beta u_y)_y$; $\alpha = \beta = 1 + 10x^2y^2$;
Au = Incompletely Factorized Finite Difference Matrix for Lu

N	$\lambda_{\min}$	κ
4	1.09	3.29
8	1.08	4.92
16	1.04	9.33

TABLE IX

Lu as in Table VIII
Au = Incompletely Factorized Finite Difference for $\mathcal{L}u = -u_{xx} - u_{yy}$

N	$\lambda_{\min}$	κ
4	1.48	6.89
8	1.44	10.94
16	1.23	19.90

In Tables VII–IX, we list some results about the preconditioning by an incompletely factorized finite difference matrix (for other results see [21]). Table VII refers to a constant coefficients operator. The different ratios between the coefficients of u_{xx} and u_{yy} are supposed to mimic the effect of the stretching of coordinates in a mapping process. The spectral matrix of a variable coefficients operator was preconditioned by the finite differences representation of the same operator (Table VIII), or by that of a constant coefficient operator (Table IX).

Unlike the case of complete factorization, the condition number grows linearly with the number of unknowns. However, it ranges within moderate bounds (except when a different operator is used in the preconditioning). This gives evidence to the convenience of using incompletely factorized preconditioning matrices in spectral calculations. As for finite difference methods (see [12]), better results can be achieved, with slightly more computational effort, by a higher order incomplete factorization in which $\tilde{L}$ and $\tilde{U}$ have one more nonzero off-diagonal.

3. A Conjugate Gradient Method

Even if the differential operator L is self-adjoint, the matrix arising from a Chebyshev spectral approximation is not symmetric. Thus, one can apply the standard conjugate gradient method (CG) to the normal equations of the preconditioned system. Alternatively, one can use CG-type methods for nonsymmetric systems, like those proposed by Vinsome [15], Young and Jea [19], Axelsson [1], or those by Concus and Golub [3], and Widlund [20]: the methods of the first class may require the storage of back steps of the solution (however, see Wong [21] for an application of a truncated version of [1] to spectral calculations), while the methods of the second class require that the symmetric part of the system be easily solvable.

In the previous section it has been pointed out that the spectral matrix can be preconditioned using a symmetric positive definite matrix, connected with some finite element approximation of the elliptic operator. This suggests the use of the following preconditioned version of the CG method: Minimize

$$J(u) = r^T A^{-1} r \qquad r = L_{\text{sp}} u - f \tag{3.1}$$

by CG iterations in $\mathbb{R}^n$ equipped with the inner product

$$((u, v)) = u^T A v. \tag{3.2}$$

The corresponding algorithm is as follows.

Given $u^0 \in \mathbb{R}^n$, compute

$$z^0 = A^{-1}(f - L_{sp} u^0)$$
$$g^0 = A^{-1} L_{sp}^T z^0, \qquad w^0 = g^0.$$

Then set for $k = 0, 1, \ldots,$

$$u^{k+1} = u^k + \alpha^k w^k, \qquad \text{where} \quad \alpha^k = \frac{((z^k, A^{-1} L_{sp} w^k))}{((A^{-1} L_{sp} w^k, A^{-1} L_{sp} w^k))} \tag{3.3}$$
$$z^{k+1} = z^k - \alpha^k A^{-1} L_{sp} w^k$$
$$g^{k+1} = A^{-1} L_{sp}^T z^{k+1}$$
$$w^{k+1} = g^{k+1} + \gamma^{k+1} w^k, \qquad \text{where} \quad \gamma^{k+1} = \frac{((g^{k+1}, g^{k+1}))}{((g^k, g^k))}.$$

Note that g^{k+1} is the Frechet derivative of the functional $J(u)$ computed at $u = u^{k+1}$ with respect to the inner product (3.2). Algorithm (3.3) does not coincide with the classical conjugate gradient method applied to the normal equations. However, the Frechet derivative of J is $A^{-1} L_{sp}^T A^{-1} L_{sp}$, which shows that the method behaves qualitatively as a CG method for the normal equations. It is precisely the CG method for solving the normal equations associated with $(A^{-(1/2)} L_{sp} A^{-(1/2)})(A^{1/2} u) = A^{-(1/2)} f$.

It is not difficult to see that the computation of α^k and γ^{k+1} can be done in terms of Euclidean products, which do not require the multiplication by the preconditioning matrix A. The product $L_{sp}^T z^{k+1}$ can be executed through fast Fourier transforms, and the entries of the matrix L_{sp} need not to be computed. Actually, assume that $L_{sp} u = -[I_N(\alpha u_x)]_x$ is the Chebyshev pseudospectral approximation of $Lu = -(\alpha u_x)_x$, where $I_N w$ is the Nth degree polynomial interpolating w at the nodes x_j, $j = 0, \ldots, N$. We identify an Nth degree polynomial vanishing at $x = \pm 1$ with the vector of its values at x_j, $j = 1, \ldots, N-1$. Recall that

$$\int_{-1}^{1} u(x)\, v(x)\, \omega(x)\, dx$$
$$= \frac{\pi}{N} \sum_{j=1}^{N-1} u(x_j)\, v(x_j) + \frac{2\pi}{N} \{u(-1)\, v(-1) + u(1)\, v(1)\} \tag{3.4}$$

for any u, v such that $uv \in \mathbb{P}_{2N-1}$. Then

$$(u, L_{sp}^T v)_{\mathbb{R}^{N-1}} = (L_{sp} u, v)_{\mathbb{R}^{N-1}} = -\frac{N}{\pi} \int_{-1}^{1} [I_N(\alpha u_x)]_x v \omega\, dx.$$

Integration by parts and several applications of (3.4) yield

$$L_{sp}^T v = -(I_N z)_x + \beta z$$

where

$$z = \alpha(v_x + \beta v), \qquad \beta(x) = \frac{\omega_x}{\omega}(x) = \frac{x}{1-x^2}. \tag{3.5}$$

Given the values of v or z at the Chebyshev points, we compute their coefficients in the expansion in Chebyshev orthogonal polynomials of first kind, using the fast Fourier transform (FFT). The coefficients of the x derivative can then be easily computed by simple recursion formulae. Finally, the values of the derivative at the collocation nodes can be obtained by inverse FFT (see [9]). Similar expressions hold in two dimensions.

Algorithm (3.3) was used to compute the spectral solution for the test problems:

$$\begin{aligned} &Lu \equiv -(\alpha u_x)_x = f, \qquad -1 < x < 1, \\ &\alpha \equiv 1.0 \qquad \text{or} \qquad \alpha(x) = 1 + 10x^2 \\ &u(x) = \sin \pi x, \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} &Lu \equiv -(\alpha u_x)_x - (\alpha u_y)_y = f, \qquad -1 < x, y < 1, \\ &\alpha \equiv 1.0 \qquad \text{or} \qquad \alpha(x, y) = 1 + 10x^2y^2 \\ &u(x, y) = \sin \pi x \sin \pi y. \end{aligned} \tag{3.7}$$

In the Tables X and XI we report the minimum number NIT of iterations required to get $\text{RES} < 10^{-8}$, where the relative residual is defined by

$$\text{RES}^2 = \frac{(r, r)}{(f, f)}, \qquad r = f - L_{sp}u. \tag{3.8}$$

TABLE X

CG Method for Problem 3.6

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	1	0.31 *E*-17	0.18 *E*0	1	0.90 *E*-17	0.11 *E*-1
8	3	0.12 *E*-16	0.31 *E*-3	3	0.19 *E*-13	0.40 *E*-3
16	7	0.53 *E*-14	0.27 *E*-11	8	0.13 *E*-15	0.61 *E*-11
32	14	0.58 *E*-8	0.12 *E*-9	16	0.60 *E*-12	0.88 *E*-14
64	20	0.52 *E*-8	0.84 *E*-10	24	0.24 *E*-8	0.18 *E*-10
128	26	0.52 *E*-8	0.47 *E*-11	29	0.55 *E*-8	0.15 *E*-10

Note. Au = finite element matrix for Lu.

TABLE XI

CG Method for Problem 3.7

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2y^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	21	0.31 *E*-8	0.18 *E*0	32	0.88 *E*-8	0.10 *E*-1
8	44	0.75 *E*-8	0.31 *E*-3	72	0.99 *E*-8	0.15 *E*-3
16	80	0.99 *E*-8	0.89 *E*-9	70	0.13 *E*-2	0.56 *E*-4

Note. Au = incompletely factorized finite element matrix for Lu.

The initial guess was $u^0 \equiv 0$. ERR is the corresponding relative error on the solution

$$\text{ERR} = \frac{\|u_{\text{sp}} - u_{\text{exact}}\|}{\|u_{\text{exact}}\|}, \tag{3.9}$$

where $\|u\| = (u, u)^{1/2}$ is the discrete l^2-norm on the grid.

It is seen that the number of iterations NIT to match the stopping criterion $\text{RES} < 10^{-8}$ increases sublinearly in one dimension and linearly in two dimensions with the degree N of polynomials. This seems qualitatively in accordance with the behavior of the condition number of the matrix $A^{-1}L_{\text{sp}}$ (see Table VI). The slow convergence of the method in two dimensions is due to the fact that the convergence factor behaves like the one of a CG method for the normal equations. Moreover, the finite element matrix A has slightly worse preconditioning properties than its finite difference counterpart.

4. The DuFort–Frankel (DF) Method

The DuFort–Frankel method can be applied to the numerical solution of steady-state equations

$$Bu = g \tag{4.1}$$

(the eigenvalues of B having positive real parts) as an iterative procedure depending on two positive parameters δ and γ:

$$\frac{u^{k+1} - u^{k-1}}{2\delta} = g - Bu^k - \gamma(u^{k+1} - 2u^k + u^{k-1}). \tag{4.2}$$

This can be written as a one step method in the form

$$\begin{bmatrix} u^{k+1} \\ u^k \end{bmatrix} = G(B; \delta, \gamma) \begin{bmatrix} u^k \\ u^{k-1} \end{bmatrix} + \frac{1 + 2\delta\gamma}{2\delta} \begin{bmatrix} g \\ 0 \end{bmatrix}, \tag{4.3}$$

with proper definition of the matrix G.

The DF scheme is a stationary second-degree method of the type considered in [17, 18]. In connection with spectral methods it has been studied by Gottlieb and Gustafsson [7] and by Funaro [5]. If B has real strictly positive eigenvalues (the largest and the smallest eigenvalues being denoted by $\lambda_{\max}$ and $\lambda_{\min}$, respectively), then it is seen that the method is convergent if

$$\gamma > \gamma_{\text{LIM}} = \frac{\lambda_{\max}}{4}. \tag{4.4}$$

Moreover, Funaro [5] proves that the spectral radius $\rho(G)$ as a function of δ and γ has a curve of local minima (with respect to increments in the δ or in the γ direction separately) given by the branches of hyperbola

$$\gamma = \frac{1 + \delta^2 \lambda_{\max}^2}{4\delta^2 \lambda_{\max}} \quad \text{if} \quad \gamma < \frac{\lambda_{\min} + \lambda_{\max}}{4}, \tag{4.5}$$

$$\gamma = \frac{1 + \delta^2 \lambda_{\min}^2}{4\delta^2 \lambda_{\text{mib}}} \quad \text{if} \quad \gamma > \frac{\lambda_{\min} + \lambda_{\max}}{4}. \tag{4.6}$$

$\rho(G)$ attains its absolute minimum at the intersection of the two branches, i.e., at the "optimal parameters"

$$\delta^* = \frac{1}{\sqrt{\lambda_{\min} \lambda_{\max}}}, \qquad \gamma^* = \frac{\lambda_{\min} + \lambda_{\max}}{4}, \tag{4.7}$$

where

$$\rho(G)_{\text{opt}} = \rho^* = \frac{\sqrt{\lambda_{\max}/\lambda_{\min}} - 1}{\sqrt{\lambda_{\max}/\lambda_{\min}} + 1}. \tag{4.8}$$

The DF method with the optimal parameters (4.7) was applied to the soluion of the test problems (3.6)–(3.7) by a preconditioned spectral method, as shown in Tables XII and XIII. Hence, we set in (4.2) $Bu = A^{-1} L_{\text{sp}} u$ and $g = A^{-1} f$, where A is the finite difference matrix associated to L, incompletely factorized in two dimensions. One DF iteration requires one multiplication $z = L_{\text{sp}} u^k$ and one forward–backward substitution $Aw = z$. The optimal parameters were computed using the exact values of $\lambda_{\min}$ and $\lambda_{\max}$ obtained in the previous section; only for case $N = 32$ of Table 4.2 the optimal parameters were computed by an extrapolation procedure, described below. The initial guess was $u^0 \equiv 0$, while u^1 was computed by a step of

TABLE XII

DF Method with Optimal Parameters for Problem (3.6)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	9	0.33 *E*–8	0.18 *E*0	19	0.41 *E*–8	0.11 *E*–1
8	12	0.19 *E*–8	0.13 *E*–3	24	0.97 *E*–8	0.40 *E*–3
16	12	0.98 *E*–8	0.39 *E*–8	26	0.52 *E*–8	0.56 *E*–8
32	14	0.25 *E*–8	0.14 *E*–8	29	0.37 *E*–8	0.79 *E*–8
64	14	0.29 *E*–8	0.29 *E*–8	28	0.83 *E*–8	0.75 *E*–8
128	14	0.57 *E*–8	0.56 *E*–8	28	0.88 *E*–8	0.86 *E*–8

Note. Au = finite differences for Lu.

the modified Euler method for the preconditioned system. NIT, RES, and ERR are defined as in Section 3, Equations (3.8)–(3.9).

It is seen that the number of iterations needed to satisfy the stopping test is bounded as a function of N in the 1-dimensional tests, while it is linearly growing in two dimensions. This corresponds to the behavior of the condition number of the matrix $A^{-1}L_{sp}$, as reported in Tables IV and VIII.

Moreover, NIT is comparable with the one relative to the CG method in one dimension, and definitely smaller in two dimensions. Since one DF iteration is faster than one CG iteration (by a factor of 1.7 both in one dimension and in two dimensions), we conclude that the DF method with optimal parameters exhibits a globally better behavior than the CG method on the tested problems.

The speedup in the convergence due to the use of a preconditioning technique is particularly impressive for the DF method. This is suggested by formula (4.8), which shows the dependence of the optimal spectral radius on the condition num-

TABLE XIII

DF Method with Optimal Parameters for Problem (3.7)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2y^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	9	0.70 *E*–8	0.18 *E*0	15	0.20 *E*–8	0.10 *E*–1
8	14	0.97 *E*–8	0.13 *E*–3	20	0.84 *E*–8	0.15 *E*–3
16	20	0.56 *E*–8	0.11 *E*–8	30	0.98 *E*–8	0.64 *E*–8
32	59	0.86 *E*–8	0.48 *E*–8	49	0.82 *E*–8	0.92 *E*–9

Note. Au = incompletely factorized finite difference matrix for Lu.

TABLE XIV

DuFort–Frankel Method without Preconditioning

N	4	8	16	32	64
NIT	23	84	327	>400	$\gg 400$

ber of B. Table XIV reports the performance of the DF method with optimal parameters without preconditioning (i.e., $Bu = L_{sp}u$) for problem (3.6) with $\alpha \equiv 1$ (compare with Table XII).

The practical interest of formulae (4.7) relies on the explicit knowledge of λ_{min} and λ_{max}, which is rarely the case. Approximate values of δ^* and γ^* may be obtained in different ways, for instance, by estimates on the eigenvalues of B or by extrapolation of correct values of δ^* and γ^* computed on coarser grids. It was found that linear extrapolation on the parameters as functions of N may lead to negative values of γ^*. Instead, linear extrapolation on the ratios of contiguous values of the parameters gives accurate answers. As already mentioned, the case $N = 32$ in Table XIII was run with extrapolated "optimal" parameters.

Unfortunately, the method appears to be rather sensitive to the choice of parameters, especially around the curve of optimality. The qualitative behavior of $\rho(G)$ as a function of γ for fixed δ (or conversely) is similar to the one encountered in an SOR method. Table XV shows the values of NIT for problem (3.6) with $\alpha = 1 + 10x^2y^2$, $N = 32$, and finite differences preconditioning, as a function of γ and δ.

The previous considerations suggest that the DF method, although in principle very powerful, may be poorly efficient in applications if the user attempts to use a fixed value of the constants δ and γ in all the iterations.

However, it is possible to transform the DF method into a completely parameter-free iterative scheme following a "minimal residual' strategy which has been proven successful in connection with other iterative schemes. The parameters γ and δ are computed at each iteration in order to minimize the l^2-norm of the residual

TABLE XV

NIT as a Function of γ and δ

δ/δ^*					
4	$\gg 400$	193	438	$\gg 400$	$\gg 400$
2	>400	98	162	416	$\gg 400$
1	>400	49	97	183	392
$\frac{1}{2}$	>400	87	61	134	196
$\frac{1}{4}$	$\gg 400$	147	145	86	151
	γ_{LIM}/γ^*	1	2	4	8

γ/γ^*

$r = f - L_{sp}u$. Given u^k, r^k and u^{k-1}, r^{k-1} then u^{k+1} and r^{k+1} are defined according to (4.2) as

$$\begin{aligned} u^{k+1} &= c_1 A^{-1} r^k + c_2 u^k + c_3 u^{k-1} \\ r^{k+1} &= -c_1 L_{sp} A^{-1} r^k + c_2 r^k + c_3 r^{k-1}, \end{aligned} \tag{4.9}$$

where

$$c_1 = \frac{2\delta}{1+2\delta\gamma}, \qquad c_2 = \frac{4\delta\gamma}{1+2\delta\gamma}, \qquad c_3 = \frac{1-2\delta\gamma}{1+2\delta\gamma}. \tag{4.10}$$

(r^{k+1}, r^{k+1}) is minimized if one sets in (4.10),

$$\begin{aligned} \delta = \delta^k &= \frac{1}{2}\frac{(q, r^{k-1} - \alpha s)}{(q, q - \beta s)} \\ \gamma = \gamma^k &= \frac{1}{2\delta^k}\frac{(p, 2\delta^k q - r^{k-1})}{(p, s)}, \end{aligned} \tag{4.11}$$

where $p = r^k - r^{k-1}$, $q = L_{sp}A^{-1}r^k$, $s = r^k + p$, $\alpha = (p, r^{k-1})/(p, s)$, and $\beta = (p, q)/(p, s)$.

In order to derive (4.11), let us differentiate c_1, c_2, and c_3 with respect to δ and γ. We get the following identities:

$$\begin{aligned} \frac{dr^{k+1}}{d\gamma} &= 4\delta\psi(\gamma, \delta)\{\delta L_{sp}A^{-1}r^k + r^k - r^{k-1}\} \\ \frac{dr^{k+1}}{d\delta} &= -2\psi(\gamma, \delta)\{L_{sp}A^{-1}r^k - 2\gamma(r^k - r^{k-1})\}, \end{aligned}$$

where $\psi(\gamma, \delta) = 1/(1+2\gamma\delta)^2$ is a positive bounded function. The system

$$\frac{d}{d\gamma}(r^{k+1}, r^{k+1}) = 0$$

$$\frac{d}{d\delta}(r^{k+1}, r^{k+1}) = 0$$

can now be written as

$$\begin{aligned} (\delta L_{sp}A^{-1}r^k + r^k - r^{k-1}, r^{k+1}) &= 0 \\ (L_{sp}A^{-1}r^k - 2\gamma(r^k - r^{k-1}), r^{k+1}) &= 0. \end{aligned}$$

Suitable combinations of these equations show that

$$(r^k - r^{k-1}, r^{k+1}) = 0 \tag{4.12}$$

$$(L_{sp}A^{-1}r^k, r^{k+1}) = 0. \tag{4.13}$$

TABLE XVI

MRDF Method for Problem (3.6)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	1	0.11 *E*–17	0.18 *E*0	1	0.70 *E*–17	0.11 *E*–1
8	5	0.48 *E*–9	0.13 *E*–3	8	0.55 *E*–8	0.40 *E*–3
16	7	0.71 *E*–8	0.38 *E*–9	11	0.46 *E*–8	0.35 *E*–8
32	4	0.56 *E*–9	0.13 *E*–9	9	0.48 *E*–8	0.22 *E*–8
64	3	0.10 *E*–9	0.20 *E*–10	3	0.45 *E*–8	0.18 *E*–8
128	2	0.33 *E*–9	0.68 *E*–10	2	0.10 *E*–8	0.56 *E*–9

Note. Au = finite differences for Lu.

Using the second identity in (4.9) we get the second identity in (4.11) by (4.12). Now, using (4.9) and some more algebra we deduce from (4.13) the expression for δ in (4.11) and then the expression for γ in (4.11).

This algorithm can be called "minimal residual DuFort–Frankel" (MRDF) method. Although formally similar to Orthores [1], a variant of the conjugate gradient method proposed in [19], our method differs from it in the way the acceleration parameters are determined. One MRDF iteration requires one forward–backward substitution $Az = r^k$ and one multiplication $w = L_{sp}s$; moreover, r^{k-1} needs to be stored with u^{k-1}. Note that if $u^1 = u^0$ but $r^0 \neq 0$ the algorithm cannot converge. Hence u^1 should be chosen in such a way that $u^1 - u^0$ and r^0 be roughly comparable. For instance u^1 can be computed from u^0 by one step of the minimal residual Richardson method (see Sect. 5(b)).

Tables XVI and XVII are analogous to Tables XII and XIII, except that the MRDF method was used instead of the DF method with optimal parameters. In

TABLE XVII

MRDF Method for Problem (3.7)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2y^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	7	0.23 *E*–8	0.18 *E*0	10	0.84 *E*–8	0.10 *E*–1
8	13	0.47 *E*–8	0.13 *E*–3	20	0.77 *E*–8	0.15 *E*–3
16	19	0.63 *E*–8	0.80 *E*–9	29	0.90 *E*–8	0.79 *E*–8
32	36	0.50 *E*–8	0.96 *E*–9	46	0.78 *E*–8	0.86 *E*–9

Note. Au = incompletely factorized finite difference matrix for Lu.

cumstances. In two dimensions, the improvement of the performances is less impressive.

However, one must not forget that the main improvement of the MRDF over the DF method relies on the complete automatization in the choice of parameters.

5. Comparisons with Other Methods

The preconditioned CG and DF methods were compared with two other iterative techniques recently suggested for spectral calculations: the Richardson iteration proposed by Orszag [13], and the minimal residual Richardson method proposed by Wong [21]. We briefly review these techniques and we report for the sake of completeness their behavior on the test problems used throughout this paper.

(a) *Richardson Method* [13]

Given u^0, compute u^{k+1} from u^k by solving

$$Au^{k+1} = Au^k - \alpha(L_{sp}u^k - f), \tag{5.1}$$

where $0 < \alpha < 2/\lambda_{max}$, λ_{min} and λ_{max} being the smallest and the largest eigenvalue of $A^{-1}L_{sp}$. The optimal value of α,

$$\alpha_{opt} = \frac{2}{\lambda_{min} + \lambda_{max}}, \tag{5.2}$$

was computed exactly and used in the following tests, as shown in Tables XVIII and XIX. One iteration requires one multiplication $z = L_{sp}w$ and one forward–backward substitution $Ax = b$.

TABLE XVIII

Richardson Method for Problem (3.6)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	8	0.90 *E*–8	0.18 *E*0	33	0.63 *E*–8	0.11 *E*–1
8	17	0.81 *E*–8	0.13 *E*–3	62	0.95 *E*–8	0.40 *E*–3
16	20	0.77 *E*–8	0.54 *E*–8	71	0.82 *E*–8	0.68 *E*–8
32	21	0.70 *E*–8	0.64 *E*–8	73	0.96 *E*–8	0.88 *E*–8
64	22	0.70 *E*–8	0.41 *E*–8	74	0.97 *E*–8	0.94 *E*–8
128	22	0.42 *E*–8	0.51 *E*–8	75	0.87 *E*–8	0.86 *E*–8

Note. Au = finite differences for Lu.

TABLE XIX

Richardson Method for Problem (3.7)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2y^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	12	0.87 *E*-8	0.18 *E*0	23	0.84 *E*-8	0.10 *E*-1
8	24	0.71 *E*-8	0.13 *E*-3	45	0.69 *E*-8	0.15 *E*-3
16	39	0.92 *E*-8	0.92 *E*-8	90	0.99 *E*-8	0.21 *E*-8

Note. Au = incompletely factorized finite difference matrix for Lu.

(b) *Minimal Residual Richardson (MRR) Method* [21]

In the previous scheme, compute $\alpha = \alpha^k$ at each iteration in order to minimize the residual (r^{k+1}, r^{k+1}). Hence one gets

Given u^0, compute $r^0 = f - L_{sp}u^0$, $z^0 = A^{-1}r^0$, then set

$$\begin{aligned} u^{k+1} &= u^k + \alpha^k z^k \qquad \text{where} \quad \alpha^k = \frac{(r^k, L_{sp}z^k)}{(L_{sp}z^k, L_{sp}z^k)}, \\ r^{k+1} &= r^k - \alpha^k L_{sp}z^k \\ z^{k+1} &= A^{-1}r^{k+1}. \end{aligned} \tag{5.3}$$

The computational effort per iteration is comparable to that of the Richardson method, (see Table XXII). Note that this method is obtained from the previous one by the same strategy used in deriving the MRDF from the pure DF method.

TABLE XX

MRR Method for Problem (3.6)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	1	0.12 *E*-17	0.18 *E*0	1	0.5 *E*-17	0.11 *E*-1
8	10	0.32 *E*-8	0.13 *E*-3	13	0.22 *E*-10	0.40 *E*-3
16	8	0.78 *E*-8	0.29 *E*-9	13	0.76 *E*-8	0.28 *E*-8
32	5	0.56 *E*-9	0.19 *E*-11	10	0.62 *E*-8	0.37 *E*-8
64	4	0.14 *E*-9	0.12 *E*-10	4	0.58 *E*-8	0.15 *E*-8
128	3	0.34 *E*-9	0.48 *E*-10	3	0.16 *E*-8	0.12 *E*-8

Note. Au = finite differences for Lu.

TABLE XXI

MRR Method for Problem (3.7)

	$\alpha \equiv 1.0$			$\alpha = 1 + 10x^2y^2$		
N	NIT	RES	ERR	NIT	RES	ERR
4	9	0.98 *E*–8	0.18 *E*0	18	0.96 *E*–8	0.10 *E*–1
8	18	0.11 *E*–8	0.13 *E*–3	22	0.29 *E*–8	0.15 *E*–3
16	23	0.90 *E*–8	0.53 *E*–8	32	0.14 *E*–8	0.60 *E*–9
32	58	0.88 *E*–8	0.19 *E*–8	58	0.89 *E*–8	0.43 *E*–9

Note. Au = incompletely factorized finite difference matrix for Lu.

(c) *Comparisons*

The speed of convergence of the methods previously discussed was compared on the basis of the number of iterations and the CPU time. Two significant cases were considered.

Case 1. Problem (3.6) with $\alpha = 1 + 10x^2$, $N = 128$, i.e., 127 grid points in the interval $(-1, 1)$.

Case 2. Problem (3.7) with $\alpha = 1 + 10x^2y^2$, $N = 32$, i.e., 31×31 grid points in the square $(-1, 1)^2$.

Define for the sake of simplicity the following labels:

(A) Richardson method (5.1) with optimal parameter (5.2)

(B) Minimal residual Richardson method (5.3)

(C) Conjugate gradient method (3.3)

(D) DuFort–Frankel method (4.2) with optimal parameters (4.7)

(E) Minimal residual DuFort–Frankel method (4.9).

We used the standard finite difference (finite element for method (C)) preconditioning matrix on the spectral grid, incompletely factorized in Case 2 according to the method described in Section 2. The optimal parameters were computed with the exact values of $\lambda_{\min}$ and $\lambda_{\max}$. $u^0 \equiv 0$ was the initial guess.

The results in Figs. 1 and 2 are in a sense machine- and programmer-independent. The relative performances of the methods can be analyzed according to the global CPU-time consumption, using Table XXII. Hence, Figs. 1 and 2 also summarize the relative performances of the methods in terms of cost, except for method (C) which is roughly 1.7 times slower than the others.

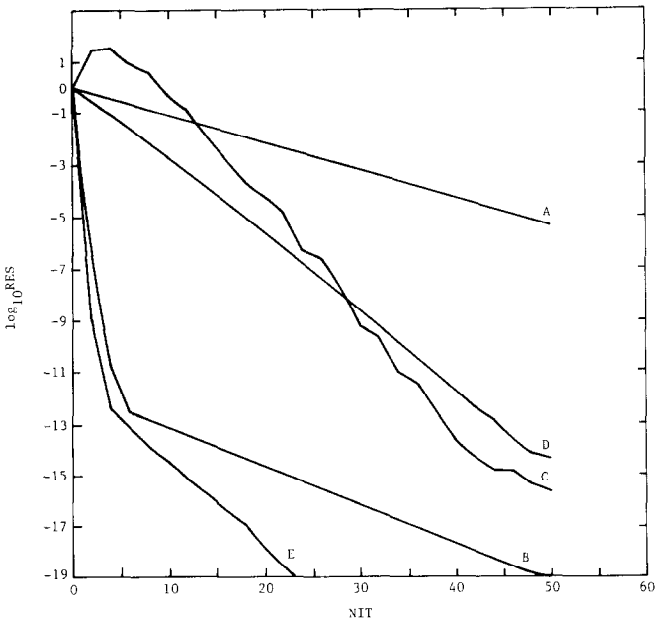

FIG. 1. Case 1: convergence histories vs number of iterations.

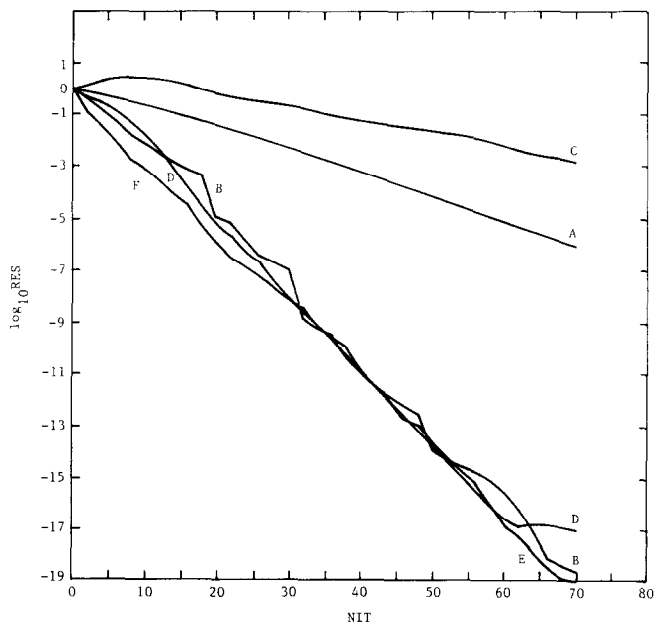

FIG. 2. Case 2: convergence histories vs number of iterations.

TABLE XXII

CPU-Time per Iteration in Hours

Method	A	B	C	D	E
Case 1	0.272 *E*–3	0.280 *E*–3	0.467 *E*–3	0.273 *E*–3	0.285 *E*–3
Case 2	0.128 *E*–2	0.128 *E*–2	0.217 *E*–2	0.127 *E*–2	0.130 *E*–2

COMMENTS

Globally, the results confirm the utility of preconditioning techniques in spectral calculations: few iterations are needed to reach the spectral accuracy, which corresponds in the test problems to a relative residual of 10^{-18}.

Methods (A) and (D) behave exactly like predicted by the theory: the error is reduced at each iteration by a factor $(\kappa-1)/(\kappa+1)$ for method (A), and $(\sqrt{\kappa}-1)/(\sqrt{\kappa}+1)$ for method (B) (κ is the condition number of the preconditioned matrix).

The conjugate gradient method gives contradictory answers in terms of speed of convergence: in one dimension the factor of reduction of the error is smaller than that for method (D), while in two dimensions it is comparable with that of method (A). In both cases, the method turns out to be not competitive in terms of computer time.

The "minimal residual" strategy is always winning over the "optimal parameter" strategy, also where the exact optimal parameters can be used. In particular, the MRR method is superior even to the Richardson method with Chebyshev acceleration, proposed in [13]. (According to [13, p. 86], the Chebyshev acceleration increases the speed of Richardson method by a factor of 2, although it requires the extra-storage of the vector u^{k-1}.)

The MRDF method requires the storage of u^{k-1} and r^{k-1}, being a 2-step method. However, the extra memory required results in a better accuracy, and the MRDF method appears in all cases the fastest method among those tested in this report.

CONCLUSIONS

The iterative methods proposed here have been compared with the most widely used iterative methods for solving systems arising from spectral approximations (see, e.g., the introductory paper by Gottlieb, Hussaini and Orszag in [16]).

Our investigation suggests that the use of conjugate gradient techniques together with spectral methods is quite delicate and will require further analysis; precisely, CG techniques which work well for finite difference or finite element approximations do not seem to be successfully transportable in a spectral context.

On the other hand, the second proposed method is always competitive with the best performing iterative method for spectral systems used so far. The improvement in speed of convergence of our method is more and more appreciable as the conditioning of the algebraic system gets better.

ACKNOWLEDGMENTS

The authors thank the editors for suggesting several improvements in the clarity of exposition and in the references to the existing literature.

REFERENCES

1. O. AXELSSON, "On Preconditioning and Convergence Acceleration in Sparse Matrix Problems," CERN Report 74–10, Geneva, Switzerland, 1974.
2. C. CANUTO AND A. QUARTERONI, *in* "Spectral Methods for Partial Differential Equations" (Robert G. Voigt, David Gottlieb, and M. Yousuff Hussaini, Eds.), pp. 55–78 SIAM, Philadelphia, 1984.
3. P. CONCUS AND G. G. GOLUB, *in* "Lecture Notes in Economics and Mathematical Systems" (R. Glowinski and J. L. Lions, Eds.), Vol. 134, pp. 56–65, Springer-Verlag, Berlin/Heidelberg/New York, 1976.
4. T. DUPONT, R. P. KENDALL, AND H. H. RACHFORD, JR., *SIAM J. Numer. Anal.*, **5** (1968), 559–573.
5. D. FUNARO, *RAIRO Anal. Numer.*, in press.
6. R. GLOWINSKI, B. MANTEL, J. PERIAUX, O. PIRONNEAU, AND G. POIRIER, *in* "Computing Methods in Applied Sciences and Engineering" (R. Glowinski and J. L. Lions, Eds.), pp. 445–487, North-Holland, Amsterdam/New York, 1980.
7. D. GOTTLIEB AND B. GUSTAFSSON, *SIAM J. Numer. Anal.* **13** (1976), 129–144.
8. D. GOTTLIEB AND L. LUSTMAN, *Comput. Fluids* **11**, No. 2, (1983), 107–120.
9. D. GOTTLIEB AND S. A. ORSZAG, "Numerical Analysis of Spectral Methods: Theory and
10. D. B. HAIDVOGEL AND T. A. ZANG, *J. Comput. Phys.* **30** (1979), 167–180.
11. J. A. MEIJERINK AND H. A. VAN DER VORST, *Math. Comp.* **31** (1977), 148–162.
12. J. A. MEIJERINK AND H. A. VAN DER VORST, *J. Comput. Phys.* **44** (1981), 134–155.
13. S. A. ORSZAG, *J. Comput. Phys.* **37** (1980), 70–92.
14. R. PEYRET AND T. D. TAYLOR, "Computational Methods for Fluid Flow," Springer, New York, 1982.
15. P. K. W. VINSOME, *in* "Proc. 4th Sympos. Reservoir Simulation," pp. 149–159, Society of Petroleum Engineers of AIME, New York, 1976.
16. R. G. VOIGT, D. GOTTLIEB, AND M. Y. HUSSAINI (Eds.), "Spectral Methods for Partial Differential Equations," SIAM, Philadelphia, 1984.
17. D. M. YOUNG, *J. Approx. Theory*, **5** (1971), 137–149.
18. D. M. YOUNG, "Iterative Solutions of Large Linear Systems," Academic Press, New York, 1971.
19. D. M. YOUNG AND K. C. JEA, *Linear Algebra Appl.* **34** (1980), 159–194.
20. O. WIDLUND, *SIAM J. Numer. Anal.* **15** (1978), 801–812.
21. T. A. ZANG, Y. S. WONG, AND M. Y. HUSSAINI, *J. Comput. Phys.* **48** (1982), 485–501.
22. P. HALDENWANG, G. LABROSSE, S. ABBOUDI AND M. DEVILLE, *J. Comput. Phys.* **55** (1984), 115–128.