Computers Math. Applic. Vol. 17, No. 7, pp. 1149–1159, 1989
Printed in Great Britain.

0097-4943/89 $3.00 + 0.00

THE ACCELERATION OF MATRIX POWER METHODS BY CYCLIC VARIATIONS OF THE SHIFT PARAMETER

I. J. D. CRAIG and A. D. SNEYD
Department of Mathematics, University of Waikato, Hamilton, New Zealand

(*Received* 18 *April* 1988)

Abstract—The traditional matrix power method converges very slowly when the dominant eigenvalues have weak relative separation. In the present paper we show that the convergence rate can be greatly accelerated by incorporating systematic cyclic displacements within the power iteration. Analytic convergence results, together with test-case numerical examples, confirm that computation times can be dramatically reduced, typically by factors of 20 and more. Indeed the cyclic shift technique works most effectively on those problems which are *most recalcitrant* to the traditional power method.

We also apply the technique to accelerating the simultaneous determination of several dominant eigenmodes by the block power method. Our results are shown to provide a simple strategy for isolating the dominant modes of self-adjoint differential operators, by explicit time-integration methods.

1. INTRODUCTION

The power method provides a simple iterative technique for determining the dominant eigenmodes of linear operators. Given an initial trial vector $\mathbf{v}^0$, the power iteration is defined by

$$\mathbf{v}^{n+1} = (A + pI)\mathbf{v}^n \tag{1}$$

where A is a real square matrix of order N, I the identity matrix, and p is an adjustable relaxation parameter (sometimes call the shift or displacement parameter). When $p = 0$ the iteration normally converges to the numerically dominant mode, i.e. to the eigenvector associated with the eigenvalue of greatest magnitude. A non-zero p is used either to speed up convergence to the dominant mode, or to shift convergence to the eigenvector at the opposite end of the spectrum.

As Wilkinson [1] remarks, the most important application of the power method probably lies in determining the extreme eigenmodes of large sparse matrices. Indeed, for very large systems such as may be derived by the finite-difference approximation of multi-dimensional partial differential equations, a power method is often all that can be used since factorization techniques usually require prohibitive overheads of computer storage. For this reason a great deal of attention has been focussed on improving and generalizing the simple power iteration (e.g. Refs [1–3]) especially on accelerating its typically slow rate of convergence. Of particular interest is the fact that *ad-hoc* variations of the shift parameter can markedly improve the convergence rate, but a systematic method of choosing such variations has not been developed. The purpose of our paper is to show that sequences of cyclic shifts can be determined *a priori* which so accelerate the convergence rate that the traditional limitations of the power method can be largely overcome. The cyclic-shift method works most effectively on problems where the leading eigenvalues are weakly separated i.e. on just those problems which are most recalcitrant to the conventional power method.

This work was motivated by a specific physical problem—determining the dominant physical eigenmodes of a self-adjoint linear differential operator associated with the stability of three-dimensional magnetic fields [4, 5]. A power iteration (1) was found to give poor convergence, and standard refinements—for instance the block power method—proved ineffective in reducing computation time. (For a typical calculation the matrix A could be of order 10^4.) It was noted however that Chodura and Schlüter [6] had shown heuristically, in a non-linear application, that cyclic time-steps may be effective in overcoming the "A" stability limit on the maximal time-step of explicit time-integration schemes. This technique, although compromised by efficiency-reducing

"safety-factors" in the original application, could be rigorously applied to the linear problem in hand, and gave an order of magnitude reduction in computation time [4]. In a more general context this same technique provides simply an optimal sequence of cyclic displacements for the power iteration (1).

The cyclic-shift technique is developed in Section 2 where it is illustrated by simple numerical examples. We show analytically and by numerical example that computation time can be reduced by factors of 20 or more. Various methods for implementing the technique are also considered. In Section 3 we discuss two further applications of the method: (i) cyclic acceleration of the block power method, in which several trial vectors are iterated simultaneously, and (ii) determining the dominant eigenmodes of self-adjoint differential operators. Our conclusions are summarized in Section 4.

2. THE CYCLIC POWER ITERATION

2.1. Introduction

To consider the problem in its simplest possible form we assume that the matrix A has a real complete eigensystem $\{\mathbf{v}_i, \lambda_i; i = 1, 2, \ldots, N\}$ with eigenvalues arranged so that

$$\lambda_1 > \lambda_2 \geqslant \lambda_3 \geqslant \cdots \geqslant \lambda_N. \tag{2}$$

We shall assume that $(\mathbf{v}_1, \lambda_1)$ is the mode of prime interest, and take $|\lambda_1| > |\lambda_N|$ unless otherwise stated. The numerical dominance of $|\lambda_1|$ is relaxed in Section 3.2 where we consider the problem of determining the dominant physical mode of a self-adjoint differential operator. In this case $|\lambda_N|$ dominates $|\lambda_1|$ typically by several orders of magnitude.

2.2. Convergence for fixed shifts

Under the assumptions mentioned above, an arbitrary vector $\mathbf{v}^0$ can be expanded in terms of the eigenvectors of A,

$$\mathbf{v}^0 = \sum_{i=1}^{N} \alpha_i \mathbf{v}_i. \tag{3}$$

Substituting the expansion for $\mathbf{v}^0$ into equation (1) we find that at the nth iteration,

$$\mathbf{v}^n = \sum_{i=1}^{N} \alpha_i (\lambda_i + p)^n \mathbf{v}_i, \tag{4}$$

which shows that the iteration tends to isolate the mode for which $|\lambda_i + p|$ is maximal. In particular if p is chosen so that

$$|\lambda_1 + p| > |\lambda_N + p| \tag{5}$$

the iteration converges to a multiple of the desired mode $\mathbf{v}_1$ (assuming that $\alpha_1 \neq 0$). In this case the asymptotic rate of convergence is given by

$$q = \max\left[\frac{|\lambda_2 + p|}{|\lambda_1 + p|}, \frac{|\lambda_N + p|}{|\lambda_1 + p|}\right], \tag{6}$$

and the best choice for p is, by symmetry,

$$p = -\tfrac{1}{2}(\lambda_2 + \lambda_N). \tag{7}$$

this yields the optimal fixed-shift convergence rate,

$$q = \left|\frac{\lambda_2 - \lambda_N}{2\lambda_1 - (\lambda_2 + \lambda_N)}\right|. \tag{8}$$

Note that convergence to the mode $(\mathbf{v}_N, \lambda_N)$ at the opposite end of the spectrum is assured by choosing p so that the inequaltiy (5) is reversed.

It is simple matter to estimate the dominant eigenvalue form the successive iterates $\mathbf{v}^n$, $\mathbf{v}^{n+1}$, since for n sufficiently large, the ratio of any pair of corresponding components is approx. $1:(\lambda_1 + p)$. A useful estimate is also provided by the Rayleigh quotient

$$\frac{\mathbf{v}^{n+1} \cdot \mathbf{v}^n}{\mathbf{v}^n \cdot \mathbf{v}^n} = \frac{(A + pI)\mathbf{v}^n \cdot \mathbf{v}^n}{\mathbf{v}^n \cdot \mathbf{v}^n} \approx \lambda_1 + p. \tag{9}$$

This latter method has the advantage of providing a faster convergence rate of q^2 when A is symmetric (e.g. Ref. [1]).

The limitations of the fixed-shift method can be illustrated by considering the symmetric tridiagonal matrix:

$$A(r) = \begin{bmatrix} 1-2r & r & 0 & \cdots & 0 & 0 \\ r & 1-2r & r & \cdots & 0 & 0 \\ 0 & r & 1-2r & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1-2r & r \\ 0 & 0 & 0 & \cdots & r & 1-2r \end{bmatrix} \tag{10}$$

which has eigenvalues given by (see Ref. [7])

$$\lambda_i = 1 - 4r \sin^2\left(\frac{i\pi}{2(N+1)}\right). \tag{11}$$

This is a typical second-order difference matrix, yet convergence of the power iterations is made difficult by the clustering of the spectrum around $\lambda_1 \approx 1$ for large N. For $N = 90$, $r = 0.4$ we have

$$\lambda_1 = 0.99952, \quad \lambda_2 = 0.99809, \quad \lambda_N = -0.59952, \tag{12}$$

and the optimal convergence ratio $q = 0.99821$. This is only a minor improvement on the $p = 0$ ratio, $q = \lambda_2/\lambda_1 = 0.99857$. In either case we anticipate that several thousand iterations will be required to isolate the dominant mode, as confirmed by the results in Table 1 (below).

2.3. Convergence for cyclic shifts

In the fixed-shift method the amplification factor for any mode over a single iteration is $\lambda + p$. Now consider a sequence of K cycles comprising the individual shifts $p_1, p_2, \ldots, p_K$. The amplification factor for any mode over one such cycle is

$$\gamma_K(\lambda) = \prod_{k=1}^{K} (\lambda + p_K), \tag{13}$$

and the object is to choose shifts p_k which will maximize the desired $|\gamma_K(\lambda)|$ at the expense of all others. Since γ_K is a polynomial of degree K in λ the problem is to construct a polynomial filter with the most favourable possible characteristics. The solution is provided (see [1]) by the Chebyshev polynomials $T_K(x)$, defined by

$$T_K(x) = \begin{cases} \cos K\theta, & x = \cos\theta, \quad 0 \leqslant x \leqslant 1; \\ \cosh K\theta, & x = \cosh\theta, \quad x \geqslant 1; \end{cases} \qquad T_K(-x) = (-1)^K T_K(x). \tag{14}$$

These have the useful feature that for $|x| \leqslant 1$ they oscillate between ± 1, while outside this interval they increase rapidly in magnitude with $|x|$ (Fig. 1). The zeros β_i of $T_K(x)$ all lie in the interval $-1 \leqslant x \leqslant 1$ and are given by

$$\beta_i = \cos[(2i-1)\pi/2K], \quad i = 1, 2, \ldots, K. \tag{15}$$

Our strategy is first to perform a linear transformation from λ to x so that the itnerval $\lambda_a \leqslant \lambda \leqslant \lambda_b$ is mapped onto $-1 \leqslant x \leqslant 1$. We call the interval $[\lambda_a, \lambda_b]$ the "killing interval" since it includes the eigenvalues which we desire to suppress in the course of the power iteration. Now the p_i are chosen

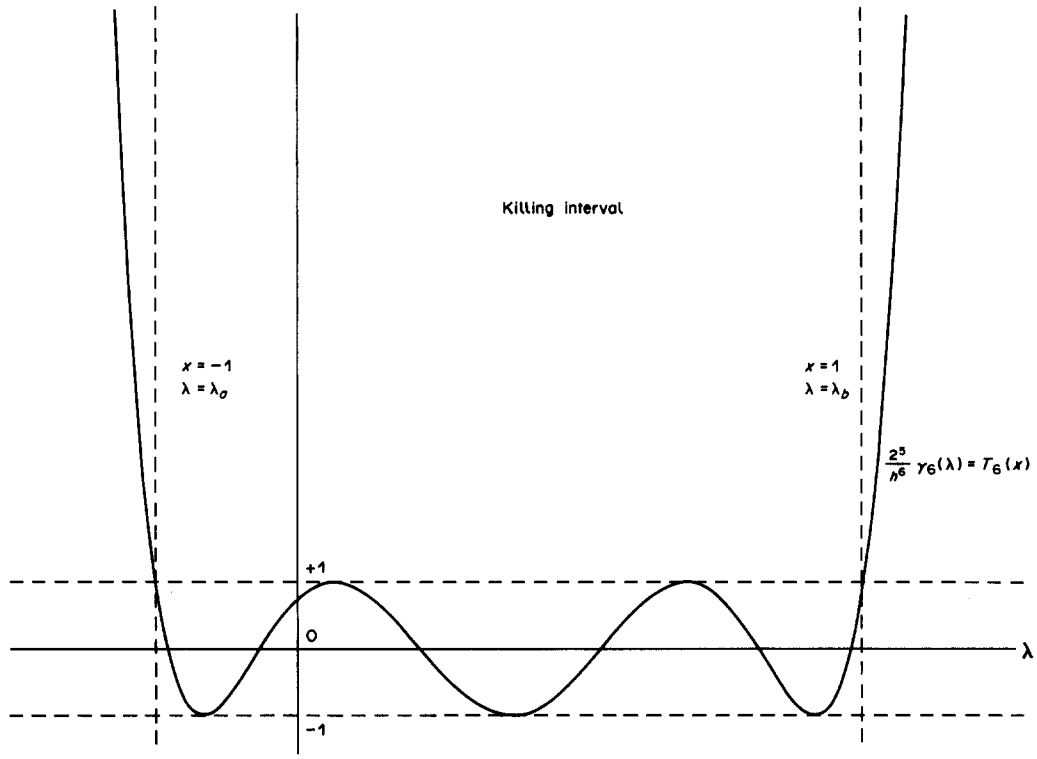

Fig. 1. Graph of Chebyshev polynomial $T_6(x)$.

so that the K zeros of $\gamma_K(\lambda)$ correspond to $x = \beta_i$ and consequently $\gamma_K(\lambda) = \text{constant} \times T_K(x)$. Specifically therefore we set

$$x = \frac{\lambda - c}{h}, \quad c = \tfrac{1}{2}(\lambda_a + \lambda_b), \quad h = \tfrac{1}{2}(\lambda_a - \lambda_b). \tag{16}$$

Then

$$\gamma_K(\lambda) = \prod_{i=1}^{K} (hx + c + p_i) = \frac{h^K}{2^{K-1} T_K(x)} \tag{17}$$

provided we choose

$$p_i = -\beta_i h - c. \tag{18}$$

The constant of proportionality in equation (17) is determined by noting that the coefficient of x^K in $T_K(x)$ is 2^{K-1}.

To ensure a growth factor of precisely $T_K((\lambda_i - c)/h)$ for each eigenmode, it can be seen from equation (17) that the trial vector should be multiplied by a factor $2^{K-1}/h^K$ after each sequence of cyclic shifts. However since it is only *relative* amplitudes which are important, this scaling is normally omitted.

Thus in order to find the largest eigenvalue λ_i we choose a killing interval $[\lambda_N, \lambda_2]$. The convergence rate per cycle is given by

$$q^K = \frac{\max\{|T_K(x_j)|, j = 2, 3, \ldots, N\}}{|T_K(x_1)|} \tag{19}$$

where $x_j = (\lambda_j - c)/h$ is the image of λ_j under the linear transformation (16). Because of our choice of killing interval each $|T_K(x_j)| \ll 1$ (with equality when $j = 2$ and $j = N$), so

$$q^K = \frac{1}{|T_K(x_1)|}. \tag{20}$$

The accelerated convergence can be illustrated by the example (12): for $K = 1, 10, 50$ we obtain convergence rates of 0.99821, 0.98322 and 0.95503 respectively. For example with $K = 50$ we obtain a 20-fold reduction in the number of iterations required (see Table 1.) If the Rayleigh quotient is used to estimate the eigenvalue, the above convergence factors are effectively squared.

2.4. Acceleration gain from cyclic shifts

After m Chebyshev iterations the error ϵ in our estimate of λ_i will be given by

$$\epsilon = \frac{c_0}{|T_K(x_1)|^m}, \tag{21}$$

where c_0 is a constant which depends on the initial eigenvalue guess. Writing $\epsilon = 10^{-w}$ and taking logs in equation (21) gives the following estimate for $n(K, w)$, the number of iterations necessary to achieve the specified accuracy:

$$n(K, w) = mK = \frac{Kw}{\log(|T_K(x_1)|)}. \tag{22}$$

(We have assumed that high accuracy is required and neglected $\log(c_0)$ in comparison with w.) If the Rayleigh quotient estimate is used, then for a symmetric matrix (or any operator with orthogonal eigenfunctions) the above estimate of n should be halved. The effectiveness of the Chebyshev cycle can be measured by the gain parameter $g(K, x_1)$, which we define by

$$g(K, x_1) = n(1, w)/n(K, w) = \frac{\log[|T_K(x_1)|]}{K \log(|x_1|)}. \tag{23}$$

Thus g is the ratio of the number of iterations required using the straightforward power method, to the number of iterations required using a Chebyshev cycle of order K. In the difficult case when λ_1 and λ_2 are almost equal,

$$x_1 = 1 + \delta, \quad 0 < \delta = \frac{2(\lambda_1 - \lambda_2)}{\lambda_2 - \lambda_N} \ll 1. \tag{24}$$

Then since $T'_K(1) = K^2$ we can approximate

$$g(K, x_1) \approx \frac{\log(1 + K^2\delta)}{K(\log(1 + \delta)} = K, \tag{25}$$

where the last approximation is valid only if $K^2\delta \ll 1$. For small K therefore the gain increases linearly with K (so the number of iterations necessary to achieve a specified accuracy falls off as K^{-1}), but for larger K it can be shown that g approaches the asymptotic limit

$$\lim_{K \to \infty} g(K, x_1) = \left(\frac{2}{\delta}\right)^{1/2}. \tag{26}$$

This behaviour is illustrated in Fig. 2 where g is plotted against K on a log–log scale, for various values of δ. It can be seen that the maximum achievable gain increases as δ decreases [cf. equation (26)], or in other words, the more difficult the problem, the more effective is the method of cyclic shifts. For a given δ there is no point in increasing K beyond a certain limit—for example when $\delta = 10^{-2}$, a Chebyshev cycle of order 30 will provide almost the maximum achievable gain.

In the above analysis we have assumed that the upper limit of the killing interval ($x = 1$) was chosen to coincide with λ_2. In practice the exact value of λ_2 will not be known, but may be estimated approximately. The best strategy will then be to place the end of the killing interval at an *underestimate* of λ_2, so that $x_2 = 1 + s$ say, and $x_1 = 1 + s + \delta$. (The parameter s is a measure of the error in the estimate of λ_2.) Now both λ_1 and λ_2 lie outside the killing interval, and equation (22) will be modified to

$$n(K, w) = \frac{Kw}{\log[T_K(1 + s + \delta)] - \log[T_K(1 + s)]}, \tag{27}$$

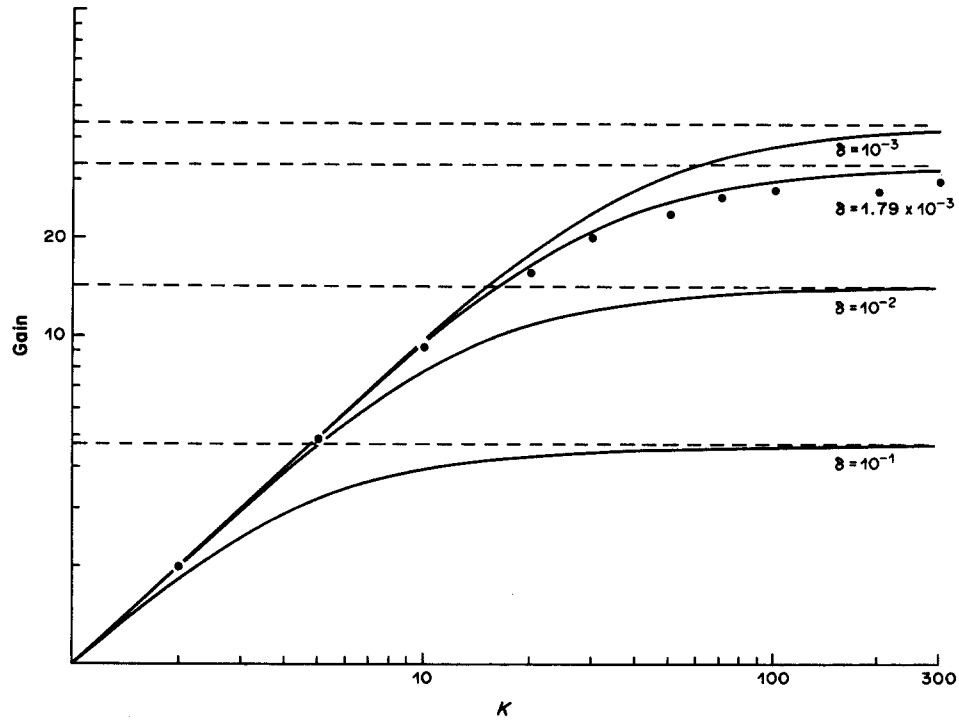

Fig. 2. Graphs of gain vs K for various values of δ. The dots correspond to the results of actual calculations summarized in Table 1. Note that each axis scale is logarithmic.

and the gain function will be given by

$$g(K, \delta, s) = \frac{\log[T_K(1+s+\delta)] - \log[T_K(1+s)]}{K[\log(1+s+\delta) - \log(1+s)]}. \tag{28}$$

For small K, $g \approx K$ as before, but the asymptotic maximum gain is reduced:

$$\lim_{K\to\infty} g(K, \delta, s) = \frac{\sqrt{2}}{\sqrt{s+\delta} + \sqrt{s}}. \tag{29}$$

Figure 3 shows graphs of g vs K for $\delta = 10^{-2}$ and various values of s. This clearly illustrates that the maximum gain decreases as s increases, and can be approached for smaller K.

2.5. Example calculations

Trial calculations were performed on the matrix $A(r)$ described in Section 2.2, with $N = 90$. The initial vector was chosen using a random number generator, and the iteration was halted when the error was reduced to at least 10^{-8}. The computations were performed using 16 decimal digit arithmetic, and a Rayleigh quotient was taken for the eigenvalue. For high values of K, there was a large decrease in error over each Chebyshev cycle, so the final error was often much smaller than 10^{-8}. The results of the trial calculations are listed in Table 1, and also plotted in Fig. 2. It can be seen that the achieved gains are generally quite close to the theoretical prediction, but fall off

Table 1. Gains obtained in trial calculations using the method of cyclic shifts

K	No. of cycles	Error w	Gain
1	6500	9.83	1.0
2	1500	9.00	1.98
5	300	11.04	4.87
10	80	11.19	9.25
20	20	9.43	15.6
30	12	10.87	19.97
50	5	8.91	23.57
70	5	14.05	26.54
100	3	12.66	27.90
200	1	8.38	27.72
300	1	13.51	29.79

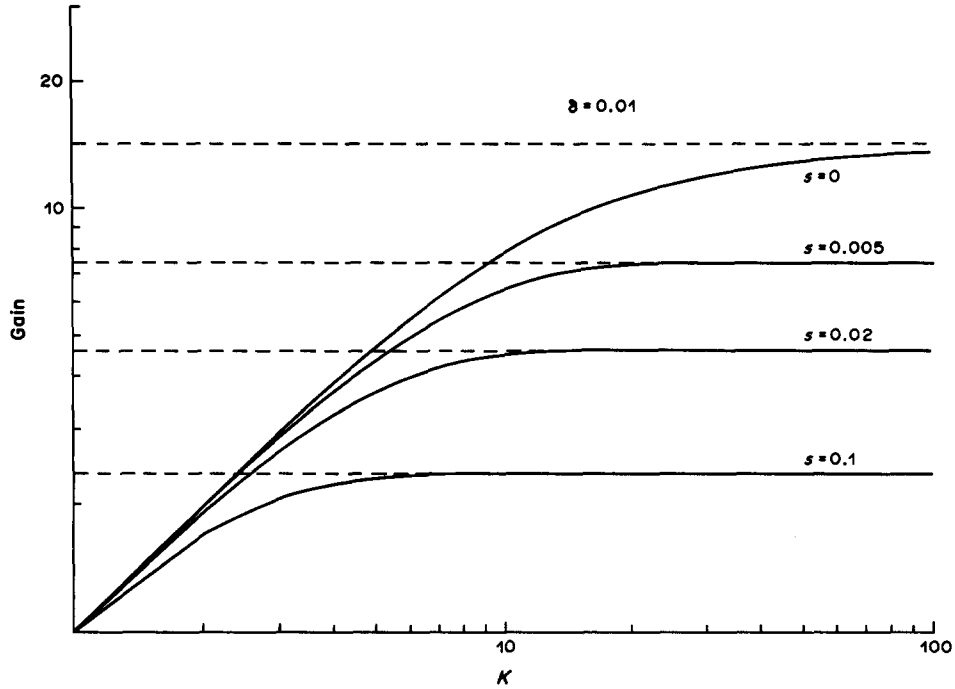

Fig. 3. Graphs of gain vs K for $\delta = 0.01$ and various values of s. (The two dominant eigenvalues both lie outside the killing interval; λ_1, λ_2 correspond to $x = 1 + s + \delta$ and $x = 1 + s$, respectively.)

slightly for large K. The maximum achieved gain was almost 30, compared with the theoretical maximum of 33.4 corresponding to $\delta = 1.79 \times 10^{-3}$.

2.6. Implementation of the cyclic-shift technique

The method of cyclic shifts requires that the shift factors $p_1, p_2, \ldots, p_k$ be computed from the zeros of the Chebyshev polynomial, but otherwise the implementation is identical to that of the traditional power method. We have found that for high K (say $K \geqslant 100$) some numerical precision can be lost in the computation of the mode amplification factors, i.e. in individual $T_K(x_i)$. In such cases numerical accuracy can be maintained by re-ordering the shift factors p_K so that large shifts are immediately followed by small ones (cf. Ref. [6]).

An alternative strategy, which avoids calculation of individual shifts, is provided by the iteration

$$\mathbf{v}^{k+1} = \frac{2(A - cI)}{h}\mathbf{v}^k - \mathbf{v}^{k-1}. \tag{30}$$

This has the effect of generating the Chebyshev polynomial growth factors $\gamma_k(\lambda)$ by virtue of the recurrence formula

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \tag{31}$$

and is theoretically identical to the cyclic-shift technique. (We note that this scheme is outlined briefly in Ref. [1] but its relation to variable displacements is not recognized.) This scheme has the advantage of giving numerical stability at high K, but has the drawback of requiring the storage of an extra trial vector. In situations where the maximum theoretical gain is approached for relatively small values of K—for example when the killing interval is not known exactly, as explained in Section 2.4—it may be better to use the cyclic-shift method to perform several complete Chebyshev cycles for a small K, than to use equation (30) and perform just one cycle for very large K. The cyclic-shift method is more suitable for generalizing to a block power iteration (see Section 3.1).

3. FURTHER APPLICATIONS

3.1. *Block power method*

In the block power method (e.g. Refs [2], [3]) several independent initial vectors are subjected simultaneously to the power iteration. An initial $N \times M$ matrix U of M trial vectors is selected (typically $N \gg M$), which, using the notation of Ref. [3], can be written in the form

$$U = QC, \quad \text{where} \quad Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N\} \tag{32}$$

is an $N \times N$ matrix composed of the eigenvectors, and C an $N \times M$ matrix of expansion coefficients. Then

$$AU = Q\Lambda C, \quad \text{where} \quad A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \tag{33}$$

is diagonal matrix containing the eigenvalues of A. Now Q is partitioned into Q_a and Q_b where Q_a is $N \times M$ and consists of just the first M eigenvectors, and Λ and C similarly partitioned. Thus equation (33) can be written as

$$AU = V = Q_a\Lambda_a C_a + Q_b\Lambda_b C_b \tag{34}$$

Since Λ_a contains the M dominant eigenvalues the first term of the right-hand side of equation (34) dominates the second. After each iteration of the form (34) an $M \times M$ *interaction matrix* is formed, whose eigenvectors can be used to construct an estimate P of C_a^{-1}. Then

$$W = VP \approx Q_a\Lambda_a + 0(\lambda_{M+1}) \tag{35}$$

provides an improved estimate of the first M eigenvectors, and is used as the U for the next iteration.

The block power method can be accelerated by replacing the single power iteration (34), with the sequence

$$V = \prod_{i=1}^{K} (A + p_i I)U = Q\Gamma C = Q_a\Gamma_a C_a + Q_b\Gamma_b C_b, \tag{36}$$

where (apart from an ignorable scaling factor)

$$\Gamma = \text{diag}(T_K(x_1), T_K(x_2), \ldots, T_K(x_N). \tag{37}$$

For the straightforward block power method the convergence rate for the jth eigenvector ($1 \leqslant j \leqslant M$) is approximately λ_i/λ_{M+1} (see Ref. [3]). When the cyclic iteration (36) is adopted this convergence rate is improved to $T_K(x_j)/(T_K(x_{M+1})$, which represents a gain approximately proportional to K, as for the power iteration. An additional saving is that the interaction matrix calculation is required only after every K steps. Table 2 gives results of some typical calculations

Table 2. Gains obtained in trial calculations using the block power method.

K	M	No. of iterations	Gain (approx)
1	1	11,150	1.0
1	2	3564	1.6
1	5	529	4.2
1	9	211	5.9
5	1	342	6.5
5	2	123	9.1
5	5	29	15.4
5	9	15	16.5
9	1	99	12.5
9	2	56	11.1
9	5	12	20.6
9	9	8	17.2
15	1	42	17.6
15	2	17	21.9
15	5	8	18.6
15	9	5	16.5
25	1	20	22.3
25	2	9	24.8
25	5	5	17.8
25	9	3	16.5

to illustrate the improvement which can be achieved in practice by the cyclic iteration. In the present example, increases in K tend to be more effective in reducing computation time than increases in M. Note also that the gain factors have not been normalized with respect to the accuracy achieved, as in Table 1; the iteration is simply terminated when successive iterates agree to within 10^{-12}, a procedure which tends to *underestimate* the gain. For $M = 1$ the iteration number is noticeably greater than the corresponding value in Table 1, presumably because the Rayleigh quotient estimate cannot be used (which means the iteration numbers are roughly doubled).

3.2. Determining dominant eigenfunctions of continuous operators by explicit time integration

The shifted power iteration is analogous to explicit time integration of an initial value problem and is well suited to investigating the eigenstructure of continuous differential operators. To illustrate the problem we consider as a simple prototype, the diffusion equation

$$u_t = L(u) = u_{xx}, \quad 0 \leqslant x \leqslant 1, \tag{38}$$

with boundary conditions $u(0) = u(1) = 0$ for all t. The eigenfunctions and eigenvalues of L are

$$\xi_i(x) = \sin(\pi i x), \qquad \lambda_i = -i^2\pi^2, \quad i = 1, 2, 3, \ldots, \tag{39}$$

so the dominant mode decays as $\exp(-\pi^2 t)$. Thus the time-evolution of an arbitrary initial disturbance—represented as a superposition of the $\xi_i(x)$—eventually isolates the fundamental mode.

To obtain an approximate numerical solution consider the classical explicit scheme

$$u_j^{n+1} = u_j^n + \Delta t(u_{j+1}^n - 2u_j^n + u_{j-1}^n)/(\Delta x^2), \tag{40}$$

where u_j^n is the approximation for $u(j\,\Delta x, n\,\Delta t)$, $\Delta x = 1/(N+1)$ is the grid size, and $u_0^N = u_{N+1}^n = 0$, by virtue of the boundary conditions. The vector representing the solution at time-step n can be conveniently written in the form,

$$\mathbf{u}^n = (I + A_L\,\Delta t)^n \mathbf{u}^0, \tag{41}$$

where A_L is the second-order difference matrix which approximates the continuous operator L. Accordingly the numerical problem is reduced to a power iteration with Δt playing a role analogous to the shift parameter p.

The convergence of the iteration (41) clearly depends on the eigenstructure of A_L, which can be found analytically (see e.g. Ref. [7]) as

$$(\xi_i)_j = \sin(ij\pi\,\Delta x); \quad \mu_i = -4\sin^2(\tfrac{1}{2}i\pi\,\Delta x)/(\Delta x^2), \tag{42}$$

which confirms that the vectors ξ provide discrete representations of the first N eigenfunctions. If $\mathbf{u}^0$ is expanded in terms of the ξ_i, equation (41) becomes

$$\mathbf{u}^n = \sum_{i=1}^{N} \alpha_i(1 + \mu_i\Delta t)^n \xi_i, \tag{43}$$

where α_i are the expansion coefficients. A necessary condition for convergence however is that each $|1 + \mu_i\Delta t| \leqslant 1$ which means in particular that

$$\Delta t \leqslant \frac{2}{|\mu_N|} \leqslant \tfrac{1}{2}(\Delta x^2). \tag{44}$$

This condition ensures that the "mesh-mode" ξ_N does not grow numerically. In fact the time-step for optimum convergence is smaller than the bound (44):

$$\Delta t = \frac{-2}{\lambda_2 + \lambda_N}. \tag{45}$$

This result is essentially (7) if we identify $\Delta t = 1/p$. It should be noted that the time-step (45) does not give an accurate description of the true time evolution since it speeds up the asymptotic convergence. Condition (44) implies very small time-steps at high resolution and so limits the efficiency of the simple time-integration scheme (41).

Now suppose that instead of one fixed time-step Δt we use a sequence $\{\Delta t_i\}$, $i = 1, 2, \ldots, K$, so that equation (43) is replaced by

$$\mathbf{u}^n = \sum_{i=1}^{N} \alpha_i [\gamma_K(\mu_i)]^n \xi_i, \quad \gamma_K(\mu) = \prod_{j=1}^{K} (1 + \mu \Delta t_i). \tag{46}$$

As before we choose a killing interval $[\mu_a, \mu_b]$ which is linearly transformed to $-1 \leqslant x \leqslant 1$. Setting

$$\Delta t_i = \frac{-1}{\beta_i h + c}, \quad \text{where} \quad c = \tfrac{1}{2}(\mu_a + \mu_b), \quad h = \tfrac{1}{2}(\mu_b - \mu_a), \tag{47}$$

then ensures that the zeros of $\gamma_K(\mu)$ coincide with those of $T_K(x)$, so that equation (46) becomes,

$$\mathbf{u}^n = \sum_{i=1}^{N} \alpha_i \left[\frac{T_K(x_i)}{T_K(-c/h)} \right]^n \xi_i. \tag{48}$$

[The constant of proportionality between $T_K(x)$ and $\gamma_K(\mu)$ is found by noting that $\gamma_k(0) = 1$.] Note that $K = 1$ with killing interval $[\mu_N, \mu_2]$ reproduces the optimum time-step result (45). For higher K however the dominant mode becomes more strongly amplified and the convergence rate improves, the gains being given by the formulae derived in Section 2.4

This simple example contains much of the essence of the eigensolution problem: the efficiency of the power iteration (41) is limited by the requirement of a small enough time-step to ensure numerical damping of the "mesh-mode" ξ_N, and worsens rapidly with increasing resolution. The standard way to overcome the problem, by adopting A-stable implicit techniques, is generally overruled by the need to store and invert the matrix A_L especially in multidimensional problems [4] (although sparse matrix techniques may be effective). In such cases the cyclic-step approach effectively circumvents the time-step restriction (46), and leads to dramatically improved convergence rates. In practice the killing interval has to be determined empirically, say by taking a conservative estimate of μ_2 (or μ_{M+1} in the block power method) and determining μ_N by a rapid preliminary calculation. Applications of this strategy for cases in which $\lambda_1/\lambda_N \approx 10^{-6}$ are given in Refs [4, 5].

4. CONCLUSIONS

The inclusion of cyclic shifts into a power iteration is easy to implement and involves negligible overheads in terms of computer storage and time. The analytic results of Section 2 show that the typically slow convergence rate of the simple power iteration can be greatly accelerated, with reductions in computation time varying inversely with relative separation of the leading eigenvalues. The technique therefore becomes more effective as the problem becomes more difficult. The results given in Section 2 provide a guide to selecting the optimum number of Chebyshev cycles (K) based on an *a priori* estimate of eigenvalue separation.

As mentioned in Section 3.2 these results impact considerably on the practical problem of determining the dominant eigenmodes of large sparse matrix systems occurring in the finite difference representation of linear differential operators. Practical experience with applications [4, 5] suggests that the cyclic-shift technique leads typically to an order-of-magnitude reduction in computation time.

It is also possible to accelerate the simultaneous determination of several dominant eigenmodes by incorporating cyclic shifts within the block power method (Section 3.1). Simultaneous iteration has the advantage of faster and more secure convergence and this may outweigh the storage overhead required for the additional trial vectors, especially in cases where the sub-dominant eigenmodes are of interest. The block power method provides the most general and flexible application of the cyclic-shift technique. Numerical results (Table 2) suggest that it is often more effective to increase the number of Chebyshev cycles K than the number of trail vectors M, although the optimum (K, M) choice is obviously sensitive to the details of the spectrum.

We have shown that a two-step iteration scheme mentioned in Ref. [1] is essentially equivalent to the method of cyclic shifts. This implementation may be preferable for very large K but has the disadvantage of requiring simultaneous storage of two trial vectors. Moreover in many situations

near-optimum convergence is attained for quite small K (see especially Fig. 3) in which case the cyclic-shift technique may be easier to implement; it is certainly more convenient to implement in the block power method.

REFERENCES

1. J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford (1965).
2. M Clint and A. Jennings, A simultaneous iteration method for the unsymmetric eigenvalue problem. *J. Inst. Math. Appl.* **8,** 111–121 (1970).
3. A. Jennings and W. J. Stewart, Simultaneous iteration for partial eigensolution of real matrices. *J. Inst. Math. Appl.* **15,** 351–361 (1975).
4. I. J. D. Craig, A. N. McClymont and A. D. Sneyd, General methods for determining the linear stability of coronal magnetic fields. *Ap. J.* (in press).
5. A. D. Sneyd and I. J. D. Craig, Linear MHD stability of infinitely long flux tubes. University of Waikato Mathematics Research Report No. 155 (1988). *Astrophys. Space Sci.* (in press).
6. R. Chodura and A. Schlüter, A 3D code for MHD equilibrium and stability. *J. Comput. Phys.* **41,** 68–88 (1981).
7. G. D. Smith, *Numerical Solution of Partial differential Equations: Finite Difference Methods*, 3rd edn. OUP, Oxford (1985).