SIAM J. SCI. COMPUT.
Vol. 22, No. 3, pp. 767–786
© 2000 Society for Industrial and Applied Mathematics

A CIRCULANT PRECONDITIONER FOR THE SYSTEMS OF LMF-BASED ODE CODES*

D. BERTACCINI†

Abstract. In this paper, a recently introduced block circulant preconditioner for the linear systems of the codes for ordinary differential equations (ODEs) is investigated. Most ODE codes based on implicit formulas, at each integration step, need the solution of one or more unsymmetric linear systems that are often large and sparse. Here, the boundary value methods, a class of implicit methods for the numerical integration of ODEs based on linear multistep formulas, are considered more in detail for initial value problems.

Theoretical and practical arguments are given to show that the block circulant preconditioner can give fast preconditioned iterations for various classes of differential problems. Moreover, the *P-circulants*, a recently introduced circulant approximation for unsymmetric Toeplitz matrices, are shown to be more suitable sometimes than other circulant matrices for the underlying block preconditioner.

Key words. circulant preconditioning, unsymmetric block (almost) Toeplitz linear systems, numerical solution of differential equations, boundary value methods, implicit linear multistep formulas

AMS subject classifications. 65F10, 65N22, 65L05, 15A18

PII. S1064827599353476

1. Introduction.

The aim of this paper is to study a new class of block preconditioners for the linear systems

$$Mx = b \tag{1.1}$$

arising in the codes for the numerical integration of ordinary differential equations (ODEs) based on linear multistep formulas (LMFs) (see, e.g., [18]). The solution of the large and sparse system of equations, arising at each integration step, is one of the crucial parts in a numerical integrator based on implicit formulas; see, e.g., [12, 4]. Here will be considered the linear systems arising in boundary value methods (BVMs) (see [3] and references therein) or that can be reduced to those with some transformation on (1.1). BVMs are a class of numerical methods based on LMFs solving initial and boundary value problems for ordinary differential equations.

The underlying block preconditioner is based on the circulant approximation of the (small rank perturbation of) band Toeplitz matrices arising in (1.1). To this end, a new type of circulant approximation for Toeplitz matrices introduced in [1], the *P-circulant matrices*, or *P-circulants*, can be effective. T. F. Chan's *optimal* [5] and Strang's circulant [23] will also be considered. Other techniques for band Toeplitz linear systems can be found, e.g., in [11, 6, 16], but they are effective only in the symmetric case while, in general, matrices for (1.1) are not symmetric.

An $n \times n$ matrix T is said to be *Toeplitz* if its entries are constant along its diagonals; an $n \times n$ matrix C is called *circulant* if it is a Toeplitz matrix with the

*Received by the editors March 12, 1999; accepted for publication (in revised form) April 15, 2000; published electronically August 31, 2000.

http://www.siam.org/journals/sisc/22-3/35347.html

†Università di Roma "La Sapienza," Istituto Matematico "G. Castelnuovo," P. le A. Moro 2, 00185 Roma, Italy (bertaccini@na-net.ornl.gov). This research was partially supported by the Italian Ministry of Scientific Research.

following pattern:

$$C = circ(c_0, \ldots, c_{n-1}) = \begin{pmatrix} c_0 & c_1 & \ldots & c_{n-1} \\ c_{n-1} & c_0 & & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \ldots & c_0 \end{pmatrix}. \tag{1.2}$$

In the last decade, there has been intensive work on preconditioners for Toeplitz matrices and their spectral properties; see, for instance, the survey [8], and the references therein, and [11, 5, 6, 7, 14, 16, 21, 23, 24].

Consider for simplicity the linear initial value problem (IVP)

$$\begin{cases} y'(t) &= f(t, y(t)) := J\, y(t) + g(t), \quad t \in (t_0, T], \\ y(t_0) &= \eta, \end{cases} \tag{1.3}$$

where $y(t), g(t) : \mathbb{R} \to \mathbb{R}^m$, $J \in \mathbb{R}^{m \times m}$, $\eta \in \mathbb{R}^m$. The matrix M (1.1) related to the underlying LMF-based code, in general, can be written as

$$M = A \otimes I_m - hB \otimes J, \tag{1.4}$$

where A, B are the matrices of the integration method, h is the integration stepsize, and $\otimes$ is the Kronecker product (see, e.g., [19, Chapter 12]). A and B are usually reducible to Toeplitz plus small rank perturbation pattern. M turns out to be large and sparse when the Jacobian matrix of the underlying system of ODEs (and/or A and B) is so. In that case, the solution of the linear system (1.1) via a direct method is often computationally expensive and not easily parallelizable; therefore the use of a preconditioned iterative method is preferable.

We stress that the proposed iterative technique can be naturally generalized to the case of d-level structures arising in d-dimensional partial differential equations (PDEs). In that case, the direct solvers become very expensive since they do not exploit properly the multiple band structure (see section 5, Example 2).

It is interesting to observe that P-circulants preserve important properties occurring in the matrices A, B in (1.4), such as

- invertibility,
- eigenvalues in the right half plane,
- almost the same sparsity pattern.

Moreover, our block preconditioner has interesting implementation potentialities, in both a scalar and a parallel computing environment.

Under appropriate hypotheses, the results obtained here can be extended to the case where the given ODE is nonlinear. Indeed, we can observe that the discrete nonlinear problem corresponding to the approximation by an LMF can be solved iteratively by considering, at each step, discrete problems such as (1.1). Notice that there are several important additional topics to take into account in the nonlinear case, such as the convergence of the modified Newton approach (see, e.g., [13]) and the mesh selection strategy.

The paper is organized as follows. In section 2 we recall some classes of BVMs. In section 3 we introduce the block circulant preconditioner and the circulant approximations. Section 4 contains some notes on the computational cost and on the convergence of preconditioned iterations. Finally, section 5 contains some numerical tests.

2. BVMs and their matrix form. Recently, the class of BVMs for differential equations has been introduced (see [3] and references therein). Such methods are based on LMFs. In order to briefly describe them, suppose for simplicity that we have the linear IVP (1.3). A BVM approximates the solution of (1.3) by means of a discrete boundary value problem (BVP). The latter is obtained by using a k-step linear multistep formula of order p over a uniform mesh $t_j = t_0 + j\,h,\ j = 0, \ldots, s$, $h = (T - t_0)/s$:

$$\sum_{i=-\nu}^{k-\nu} \alpha_{i+\nu} y_{n+i} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu} f_{n+i}, \quad n = \nu, \ldots, s-k+\nu. \tag{2.1}$$

As usual, y_n is the discrete approximation to $y(t_n)$, $f_n = f(t_n, y_n) \equiv J\,y_n + g_n$, $g_n = g(t_n)$, and the values

$$y_0, \ldots, y_{\nu-1}, \quad y_{s-k+\nu+1}, \ldots, y_s \tag{2.2}$$

are given. We observe that the IVP (1.3) provides only the initial value y_0. It is possible to avoid supplying the other conditions in (2.2) by coupling the *main method* (2.1) with other difference schemes of order p, called *additional methods*, which provide the set of equations

$$\sum_{i=0}^{k} \alpha_i^{(j)} y_i = h \sum_{i=0}^{k} \beta_i^{(j)} f_i, \quad j = 1, \ldots, \nu-1, \tag{2.3}$$

$$\sum_{i=0}^{k} \alpha_{k-i}^{(j)} y_{s-i} = h \sum_{i=0}^{k} \beta_{k-i}^{(j)} f_{s-i}, \quad j = s-k+\nu+1, \ldots, s, \tag{2.4}$$

independent of those in (2.1). For simplicity, such formulas are assumed to have the same number of steps as the main method. The equations (2.1), (2.3), and (2.4) define the use of a BVM on problem (1.3). The advantage in using BVMs, over the known LMFs requiring only initial conditions, derives from their stability properties.

It is useful to cast BVMs in matrix form. This is done by introducing the matrices $A, B \in \mathbb{R}^{(s+1)\times(s+1)}$

$$A = \begin{pmatrix} 1 & \cdots & 0 & & & & \\ \alpha_0^{(1)} & \cdots & \alpha_k^{(1)} & & & & \\ \vdots & \vdots & \vdots & & & & \\ \alpha_0^{(\nu-1)} & \cdots & \alpha_k^{(\nu-1)} & & & & \\ \alpha_0 & \cdots & \alpha_k & & & & \\ & \alpha_0 & \cdots & \alpha_k & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \alpha_0 & \cdots & \alpha_k & \\ & & & \alpha_0^{(s-k+\nu+1)} & \cdots & \alpha_k^{(s-k+\nu+1)} & \\ & & & \vdots & \vdots & \vdots & \\ & & & \alpha_0^{(s)} & \cdots & \alpha_k^{(s)} & \end{pmatrix}, \tag{2.5}$$

and B similarly but with β_j's instead of α_j's and all zeros in its first row. The discrete problem generated by the application of the BVM (2.1)–(2.4) to problem (1.3) is then

given by

$$M\,Y = e_1 \otimes \eta + h\,(B \otimes I)g, \tag{2.6}$$
$$e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^{s+1}, \quad Y = (y_0, \ldots, y_s)^T, \quad g = (g_0, \ldots, g_s)^T,$$

$$M = A \otimes I_m - h\,B \otimes J.$$

The matrix M in (2.6) turns out to be large and sparse when $s \gg k$ and/or J is large and sparse.

2.1. Some families of BVMs. Here we give the definitions of some families of BVMs (see [3] for details). All considered methods are consistent, i.e., they satisfy the conditions

$$\rho(1) = 0, \qquad \rho'(1) = \sigma(1),$$

where $\rho(z)$ and $\sigma(z)$ denote, as usual, the two characteristic polynomials associated with the given method, i.e.,

$$\rho(z) = z^\nu \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu}\, z^j, \quad \sigma(z) = z^\nu \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu}\, z^j. \tag{2.7}$$

The generalized BDF, or GBDF, are a generalization of the backward differentiation formulas (BDF) (see [13]). They can be written in the form

$$\sum_{i=-\nu}^{k-\nu} \alpha_{i+\nu}\, y_{n+i} = h\, f_n, \quad n = \nu, \ldots, s-k+\nu, \tag{2.8}$$

where the coefficients $\{\alpha_i\}$ are uniquely determined by imposing that the method has maximum order, i.e., k, for all $k \geq 1$, with $\nu = (k+2)/2$ if k is even and $\nu = (k+1)/2$ if k is odd. Such methods are well suited for stiff problems. (See [3, Chapter 5] for details.)

The generalized Adams methods (GAM) are a generalization of the Adams–Moulton methods (see [13]). They can be written in the form

$$y_{n+\nu} - y_{n+\nu-1} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu}\, f_{n+i}, \tag{2.9}$$

where the coefficients $\{\beta_i\}$ are uniquely determined by imposing that the method has maximum order, i.e., $k+1$, for all $k \geq 1$, with $\nu = k/2$ if k is even and $(k+1)/2$ if k is odd. When k is odd, they are called extended trapezoidal rules (ETR), because they share the same stability properties of the trapezoidal rule. Such methods turn out to be well suited for approximating either Hamiltonian problems or continuous BVPs. When k is even, GAM are well suited for stiff problems (see [3, Chapters 6 and 7] for details).

ETR_2 are another generalization of the trapezoidal rule belonging to the class of symmetric schemes:

$$\sum_{i=-\nu}^{\nu-1} \alpha_{i+\nu}\, y_{n+i} = \frac{h}{2}(f_n + f_{n-1}), \tag{2.10}$$

where $k = 2\nu - 1$ is odd and the coefficients $\{\alpha_i\}$ are uniquely determined by imposing that the method has maximum order, i.e., $k+1$, $k = 1, 3, 5, \ldots$. When $\nu = 1$, then $k = 1$ and the formulas (2.10), (2.9) become the trapezoidal rule. Indeed, all such formulas can be regarded as generalizations of this method, sharing the same stability properties. Such methods turn out to be well suited for approximating both Hamiltonian problems and continuous BVPs (see [3, Chapter 7] for details).

3. Circulant approximations for the block preconditioner. Let M in (2.6) be the small rank perturbation of a block Toeplitz matrix generated by a BVM as in the previous section. A way to solve such large, sparse linear systems is through an iterative method; see, e.g., [12, 4]. To accelerate convergence, a preconditioner P should be chosen to approximate the matrix M while keeping the system

$$P\,x = c$$

cheap enough to solve with respect to the unpreconditioned iterations.

In order to obtain the preconditioner, let us consider the following approximation of the matrix M:

$$P = \breve{A} \otimes I_m - h\,\breve{B} \otimes \hat{J}, \tag{3.1}$$

where $\hat{J}$ can be a suitable approximation of the Jacobian matrix of the ODE, or the Jacobian itself. $\breve{A}$, $\breve{B}$ are circulant matrices the entries of which are derived from the coefficients of the main method (2.1) as follows:

$$\begin{aligned} &\breve{A} = circ(\tilde{\alpha}_j), \qquad \tilde{\alpha}_j = c_{j,1}(s)\alpha_{j+\nu} + c_{j,2}(s)\alpha_{j+\nu-(s+1)}, \\ &\breve{B} = circ(\tilde{\beta}_j), \qquad \tilde{\beta}_j = c_{j,3}(s)\beta_{j+\nu} + c_{j,4}(s)\beta_{j+\nu-(s+1)}, \qquad j = 0, \ldots, s, \end{aligned} \tag{3.2}$$

where the $c_{j,i}(s)$, $i = 1, \ldots, 4$, $j = 0, \ldots, s$ are linear in j. It is understood that α_j (β_j) is zero for $j < 0$ or $j > k$ in (3.2). The coefficients $c_{i,j}(s)$ in (3.2) are chosen in such a way that $\breve{A}$, $\breve{B}$ are suitable approximations of A, B in (2.5), respectively.

The approximation of A, B with Chan's optimal circulant (see [5]) requires that

$$c_{j,1}(s) = c_{j,3}(s) = 1 - \frac{j}{s+1}, \qquad c_{j,2}(s) = c_{j,4}(s) = \frac{j}{s+1}, \qquad j = 0, \ldots, s, \tag{3.3}$$

while for Strang's natural circulant (see [23])

$$c_{j,1}(s) = c_{j,3}(s) = 1, \qquad j = 0, \ldots, \left\lfloor \frac{s+1}{2} \right\rfloor,$$

$$c_{j,2}(s) = c_{j,4}(s) = 1, \qquad j = \left\lfloor \frac{s+1}{2} \right\rfloor + 1, \ldots, s, \qquad c_{r,j}(s) = 0 \ \text{ otherwise.} \tag{3.4}$$

Let us observe that, for an $(s+1) \times (s+1)$ Toeplitz matrix T, Chan's circulant $C = C(T)$ is defined to be the minimizer of

$$\|T - C\|_F \tag{3.5}$$

over all $(s+1) \times (s+1)$ circulant matrices C; see [5]. ($\|\cdot\|_F$ is the Frobenius norm.) Consider, instead of (3.3), the following definition of the coefficients $c_{j,i}(s)$:

$$c_{j,1}(s) = c_{j,3}(s) = 1 + \frac{j}{s+1}, \qquad c_{j,2}(s) = c_{j,4}(s) = \frac{j}{s+1}, \qquad j = 0, \ldots, s. \tag{3.6}$$

We will call P-circulants the circulant matrices defined in (3.2), (3.6). The definitions (3.2), (3.3), (3.2), (3.4) and (3.2), (3.6) differ for a quantity that vanishes as s increases for banded matrices, and (3.5) is not minimized for P-circulants. Despite this, the P-circulant matrices $\breve{A}$, $\breve{B}$, defined as (3.2), (3.6), are nonsingular and appropriate approximations of A, B given by (2.5), as observed in [2].

3.1. Spectral properties of the block preconditioner. We observe that positive stability is a sufficient condition for the invertibility of a matrix. Recall that a square matrix A is said to be positive stable if its eigenvalues have positive real part (see, e.g., [15]). The matrices A, B of the BVMs we consider here are assumed to be positive stable. This assumption is a consequence of the stability properties of the formulas; see [3]. It is interesting to observe that the P-circulant approximations $\breve{A}$, $\breve{B}$ given by (3.2), (3.6) preserve positive stability for the methods described in section 2.1 (see [2]). Let $\breve{A} = circ(\tilde{\alpha}_j)$ be a $(s+1) \times (s+1)$ circulant matrix defined in (3.2), (3.6) from the first characteristic polynomial of the main method (2.1). The eigenvalues ϕ_j, $j = 0, \ldots, s$, of $\breve{A}$ can be written as linear combinations of the entries of its first row (see [10]):

$$\phi_l = \sum_{j=0}^{s} \tilde{\alpha}_j \epsilon^{jl}, \quad l = 0, \ldots, s, \quad \epsilon = e^{2\pi\,\mathbf{i}/(s+1)}, \tag{3.7}$$

where $\mathbf{i}$ is the imaginary unit. From (3.2), (3.6) one obtains

$$\phi_l = \sum_{j=0}^{s} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \epsilon^{jl} + \sum_{j=0}^{s} \left(\frac{j}{s+1} \alpha_{j+\nu-(s+1)}\right) \epsilon^{jl},$$

which can be restated as

$$\phi_l = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \epsilon^{jl}, \qquad l = 0, \ldots, s. \tag{3.8}$$

A similar expression holds for the eigenvalues of $\breve{B}$:

$$\psi_l = \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu} \left(1 + \frac{j}{s+1}\right) \epsilon^{jl}, \qquad l = 0, \ldots, s. \tag{3.9}$$

In [2] we have observed that, for some A-stable ($A_{\nu,k-\nu}$-stable, a generalization of A-stability for (2.1) if $k > \nu$; see [3]) BVMs, the $l2$-norms of the P-circulant matrices $\breve{A}$, $\breve{B}$ are uniformly bounded and the $l2$-norms of their inverses are bounded with respect to the number of steps k of the method (2.1), i.e.,

$$\|\breve{A}\|_2 \le c_1, \qquad \|\breve{B}\|_2 \le c_2, \qquad \|\breve{A}^{-1}\|_2 \le c_3 \cdot (s+1), \qquad \|\breve{B}^{-1}\|_2 \le c_4 \cdot (s+1), \tag{3.10}$$

where c_j, $j = 1, \ldots, 4$, are constants of the order of unity. Exceptions are the GBDF (they have $\breve{B} \equiv I$) and the ETR_2.

Notice that the circulant matrix $\breve{A}$ defined as the optimal circulant (3.2), (3.3), may become severely ill-conditioned as k increases (see [2]).

We can give a sample of sufficient conditions for the block preconditioner (3.1) to be positive stable. From here on, for simplicity, we will assume the Jacobian matrix to be diagonalizable.

PROPOSITION 3.1. *Suppose that the eigenvalues μ_r, $r = 1, \ldots, m$, of the approximation of the Jacobian matrix J of the given ODE are in the left half plane, i.e., they have nonpositive real part. Each of the following conditions is sufficient for the positive stability of the block preconditioner P given by* (3.1):

1. *$\breve{A}$, $\breve{B}$ are P-circulants and* (2.1) *belongs to GBDF;*
2. *$\breve{A}$, $\breve{B}$ are positive stable and $\breve{B}$ is Hermitian;*
3. *$\breve{A}$, $\breve{B}$ are positive stable and, for $j \neq 0, (s+1)/2$, we have*

$$|\mathrm{Im}(\mu_r)| < \left(\frac{\mathrm{Re}(\phi_j) - h\mathrm{Re}(\psi_j)\mathrm{Re}(\mu_r)}{h\,|\mathrm{Im}(\psi_j)|} \right), \ r = 1, \ldots, m, \tag{3.11}$$

where ϕ_j, ψ_j, $j = 0, \ldots, s$, are the eigenvalues of $\breve{A}$, $\breve{B}$, respectively.

Proof. If J is the Jacobian matrix,

$$V^{-1} J V = D = \mathrm{diag}(\mu_1, \ldots, \mu_m). \tag{3.12}$$

The circulant matrices $\breve{A}$ and $\breve{B}$ are simultaneously diagonalized by the unitary Fourier matrix F (see [10])

$$F = (F)_{j,r}\,, \qquad (F)_{j,r} = \frac{1}{\sqrt{s+1}} \epsilon^{jr}, \qquad \epsilon = e^{2\pi\,\mathbf{i}/(s+1)}, \qquad 0 \leq j, r \leq s, \tag{3.13}$$

i.e.,

$$F\,\breve{A}\,F^* = \Lambda_A = \mathrm{diag}(\phi_0, \ldots, \phi_s), \quad F\,\breve{B}\,F^* = \Lambda_B = \mathrm{diag}(\psi_0, \ldots, \psi_s).$$

We can write

$$\begin{aligned} P &= (F^* \otimes I_m)\ (\Lambda_A \otimes I_m - \Lambda_B \otimes h\,J)\,(F \otimes I_m) \\ &= (F^* \otimes V)\ (\Lambda_A \otimes I_m - \Lambda_B \otimes h\,D)\left(F \otimes V^{-1}\right), \end{aligned} \tag{3.14}$$

where Λ_A, Λ_B are nonsingular diagonal matrices of size $s+1$. Let

$$\breve{\Lambda} = \Lambda_A \otimes I_m - \Lambda_B \otimes h\,D, \tag{3.15}$$

$$\begin{aligned} \breve{\Lambda} = \mathrm{diag}\,(&\phi_0 - h\psi_0\mu_1, \ldots, \phi_0 - h\psi_0\mu_m, \ldots, \\ &\phi_s - h\psi_s\mu_1, \ldots, \phi_s - h\psi_s\mu_m)\,. \end{aligned} \tag{3.16}$$

$\breve{\Lambda}$ is diagonal and, if

$$\mathrm{Re}(\phi_j - h\psi_j\mu_r) > 0, \quad j = 0, \ldots, s, \quad r = 1, \ldots, m, \tag{3.17}$$

then $\breve{\Lambda}$ (and P) is also positive stable. If $\breve{A}$, $\breve{B}$ are P-circulants, or, more generally, for all block preconditioners whose matrices $\breve{A}$, $\breve{B}$ are positive stable and $\breve{B}$ is Hermitian, we have that (3.17) holds true. Indeed, by the hypotheses, we have

$$\mathrm{Re}(\mu_r) \leq 0, \qquad \mathrm{Re}(\phi_j), \mathrm{Re}(\psi_j) > 0$$

and $\mathrm{Im}(\psi_j) = 0$, $j = 0, \ldots, s$, $r = 1, \ldots, m$.

If $\breve{B}$ is not Hermitian, the condition (3.17) is equivalent to

$$\mathrm{Re}(\phi_j) + h(\mathrm{Im}(\psi_j)\mathrm{Im}(\mu_r) - \mathrm{Re}(\psi_j)\mathrm{Re}(\mu_r)) > 0, \tag{3.18}$$

which can be restated as in (3.11). ∎

Notice that P may not be positive stable for certain formulas, even if the Jacobian matrix J has the spectrum in the left half plane.

Let us observe that, if the eigenvalues μ_r, $r = 1, \ldots, m$, of the matrix J of the IVP are in the left half plane, then the block preconditioner's matrix P is nonsingular for all schemes whose matrices $\breve{A}$, $\breve{B}$ in (3.2) are such that

$$\mathrm{Re}\left(\frac{\phi_j}{\psi_j}\right), \quad j = 0, \ldots, s, \tag{3.19}$$

is positive. Indeed, P is singular if and only if there exists at least a couple j, r, $j \in \{0, \ldots, s\}$, $r \in \{1, \ldots, m\}$, such that (see Proposition 3.1 and (3.16))

$$\phi_j - h\psi_j \mu_r = 0.$$

This cannot happen if $\mathrm{Re}(\phi_j/\psi_j)$ is positive.

We have verified that this is the case of P-circulant matrices for the BVMs introduced in section 2.1. Notice that for GBDF, this is a consequence of positive stability of P-circulants.

PROPOSITION 3.2. *If* $\mathrm{Re}(\mu_r) < -\delta$, $r = 1, \ldots, m$, *and* $\delta > 0$, *the block preconditioner* P *given by* (3.1) *which uses Strang's approximation* (3.2), (3.4) *is invertible, regardless of the stepsize* $h > 0$, *for all* $A_{\nu,k-\nu}$*-stable methods such that their boundary locus* $\rho(z)/\sigma(z)$, $|z| = 1$, *is a regular Jordan curve.*

Proof. Indeed, under the above hypotheses, we have

$$\mathrm{Re}(\rho(z)/\sigma(z)) \geq 0, \qquad |z| = 1$$

(see [3, Theorem 4.7.2]). By observing that $\mathrm{Re}(\rho(\epsilon^j)/\sigma(\epsilon^j))$, $j = 0, \ldots, s$, $\epsilon = e^{2\pi \mathbf{i}/(s+1)}$, is the ratio (3.19) when Strang's approximation is in use, it follows that $\phi_j - h\psi_j\mu_r$ cannot be zero. ∎

For more details on the boundary locus of an LMF, see [18] or [3].

Strang's approximation (3.4) for our block preconditioner was considered also in [9].

4. Using the preconditioner.

4.1. The computational cost. We have that the typical computational cost of the block circulant preconditioner (3.1) for an iterative Krylov subspace method is of the order of

$$(c_1 ms \log s + c_2\, c_3 \chi_1(J))\, n + c_3 \chi_2(J) \tag{4.1}$$

floating point operations. The c_i, $i = 1, 2$, are constants of moderate size, $c_3 \leq (s+1)$, n is the number of iterations of the iterative method, and $\chi_i(J)$, $i = 1, 2$, are suitable functions of the size and the structure of J, as explained in the sequel.

From (3.14), we can write P as

$$P = (F^* \otimes I_m) G (F \otimes I_m). \tag{4.2}$$

G is a $(s+1)$-block diagonal matrix with $m \times m$ diagonal blocks,

$$G_j = \phi_j I_m - h\psi_j\, J, \quad j = 0, \ldots, s, \tag{4.3}$$

that, under suitable assumptions (see, e.g., the previous section) are nonsingular matrices. Thus, to solve the linear systems whose matrices are given by (3.1), we need to apply FFTs of length $s+1$ to suitable permutations of the vector $\tilde{v}$ (asymptotic cost: $O(ms\log(s))$ operations). After this, we have to solve the $s+1$ linear systems whose matrices are given by (4.3). Denote with $\chi_2(J)$ the number of flops required for the factorization of G_j and with $\chi_1(J)$ the cost of the related back-substitution. We have, as an example, $\chi_i(J) = O(m)$, $i = 1, 2$, if J has a few nonzero diagonal entries (this is the case of the examples of section 5), or $\chi_2(J) = O(m^3)$, $\chi_1(J) = O(m^2)$, if J is dense and unstructured, etc. The decompositions should be computed only once and then stored.

If J is cheap enough to be diagonalized, and the condition number of V is moderate, where $D = V^{-1}JV$, the solution of the linear systems (2.6) can be found by direct inversion. Unfortunately, such hypotheses are relatively infrequent, especially for nonlinear problems. This is why we have not used the direct inversion for Examples 1 and 2 in section 5 even if possible (Example 3 has a nondiagonalizable Jacobian).

Notice that sometimes we do not need to solve all $s+1$ linear systems (4.3). Indeed, *after* the stepsize h has been chosen, we can check if (for those $j = 0, \ldots, s$ such that $\phi_j \neq 0$), the following condition is satisfied:

$$h \left| \frac{\psi_j}{\phi_j} \right| \|J\| < 1, \tag{4.4}$$

where ψ_j, ϕ_j are the eigenvalues of $\breve{B}$, $\breve{A}$, respectively. For those j such that (4.4) holds true, it can be observed that

$$G_j^{-1} \approx \phi_j^{-1} \left(I_m + h\frac{\psi_j}{\phi_j} J + h^2 \frac{\psi_j^2}{\phi_j^2} J^2 + \cdots + h^r \frac{\psi_j^r}{\phi_j^r} J^r \right), \tag{4.5}$$

where $r \geq 1$ is an integer and $\|\cdot\|$ is any norm such that $\|I\| = 1$. We have experienced that, truncating (4.5) to the linear term, we have a cheap and effective approximation for G_j^{-1}.

The condition (4.4) has proved useful also for stiff (nonlinear) problems. Indeed, often there are several subintervals of integration where the condition (4.4) is satisfied for some values of $j \in \{0, \ldots, s\}$. As an example, this is the case of van der Pol's equation

$$\begin{cases} y_1' = y_2, \\ y_2' = -y_1 + \mu y_2(1 - y_1^2), \quad t \in [0, \mu], \\ y_1(0) = 2, \qquad y_2(0) = 0. \end{cases} \tag{4.6}$$

If μ is large, (4.6) is stiff (see, e.g., [13]). However, we have experienced (up to $\mu = 1000$) that the condition (4.4) is satisfied in several mesh intervals for most of the indexes j (except in the layer regions when μ is large). Notice that $R_j = |\psi_j/\phi_j|$ is less than 1 for the BVMs considered here for several j, $j \in \{0, \ldots, s\}$. See Figure 4.1.

The matrices G_j occur in conjugate pairs, and several other ideas are possible to exploit their special structure; see [13, chapter IV.8].

The computational cost of the underlying block circulant preconditioners can be reduced defining ad hoc methods for differential equations. Indeed, the coefficients α_j, β_j, $j = 0, \ldots, k$, for the LMF (2.1) can be chosen such that the following ratio is

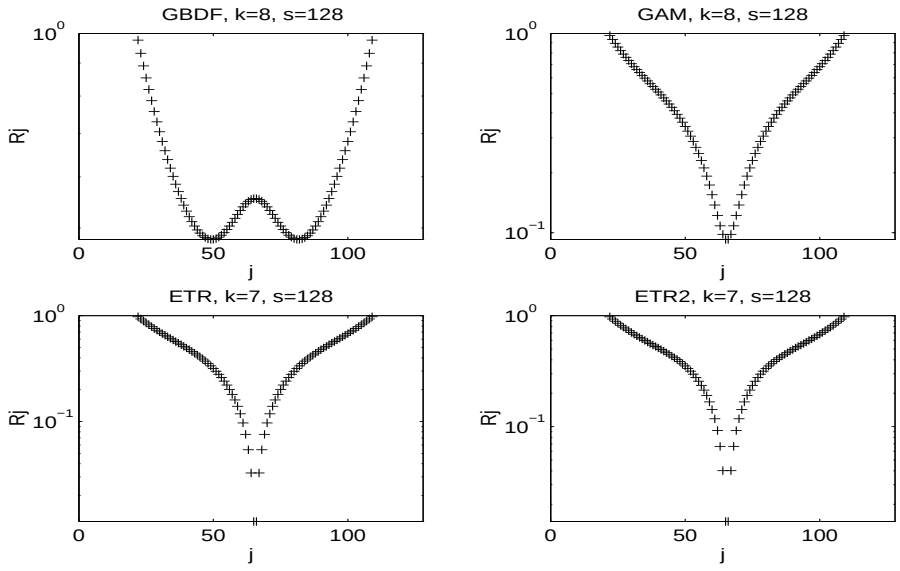

FIG. 4.1. $R_j = |\psi_j/\phi_j|$, $j = 0, \ldots, s$, *for* $k = 7, 8$ *(GBDF, GAM, ETR, ETR$_2$ formulas).*

kept constant:

$$\frac{\psi_j}{\phi_j} = \lambda, \ j = 0, \ldots, s,$$

where ϕ_j, ψ_j are, respectively, the eigenvalues of the P-circulant matrices $\breve{A}$, $\breve{B}$. Using those methods, we need only one factorization of an $m \times m$ matrix which has (almost) the same sparsity pattern of the Jacobian matrix of the underlying ODE. Notice that this process has some analogy with the diagonally implicit Runge–Kutta methods (see [13]).

An alternative approach to the direct solution of the linear systems (4.3) can be effective for particular problems. This is the case of the systems of time-dependent partial differential equations; see [14] and references therein. We will pursue a similar approach for our block preconditioner in a future work.

Finally, the block preconditioner is inherently parallel. The $s + 1$ linear systems (4.3) are independent and can be solved in parallel.

4.2. Convergence of the preconditioned iterations. We expect fast convergence of preconditioned iterations if the spectrum of the block preconditioned matrix is clustered around $(1, 0) \in \mathbb{C}$. To this end, notice that $P^{-1}M$, where P is the block P-circulant or the block Strang's preconditioner, can be written as the sum of the identity, a low rank and a small norm matrix.

THEOREM 4.1. *Let M be the matrix of the linear system* (2.6) *for the schemes described in section* 2.1, *and let P be its block P-circulant preconditioner as* (3.1). *Then, for fixed $\delta > 0$, there exist $C_\delta \geq 0$, $s_\delta \geq k$ such that, for all $s \geq s_\delta$ ($s+1$ is the size of A, B),*

$$P^{-1} M = I + M_\delta^{(1)} + M_\delta^{(2)}, \tag{4.7}$$

where $\operatorname{rank}(M_\delta^{(2)}) \le m\,[2(k+1)+C_\delta]$ *and* $\|M_\delta^{(1)}\|_2 \le \delta\, c_J$, *where* c_J *does not depend on* s.

If P *is defined as Strang's circulant,* $C_\delta = \|M_\delta^{(1)}\| = 0$.

Proof. Let $E = M - P$,

$$\begin{aligned} E &= (A - \breve{A}) \otimes I_m - h(B - \breve{B}) \otimes J \\ &= E_A \otimes I_m - h\, E_B \otimes J \\ &= (E_A^{(1)} + E_A^{(2)}) \otimes I_m - h(E_B^{(1)} + E_B^{(2)}) \otimes J. \end{aligned} \tag{4.8}$$

Here $E_A^{(1)}$ is an $(s+1)\times(s+1)$ Toeplitz matrix with bandwidth $m(k+1)$ defined from the coefficients α_j of the main method (2.1)

$$E_A^{(1)} = (e_{A,rl}^{(1)}), \quad e_{A,rl}^{(1)} = e_{A,r-l}^{(1)} = e_{A,j}^{(1)}, \quad j = -s, \dots, s,$$

and, for P-circulants and optimal circulants,

$$e_{A,j}^{(1)} = \begin{cases} \dfrac{-j}{s+1}\alpha_{j+\nu}, & j = -\nu, \dots, k-\nu \ \text{ (P-circulant)}, \\ \dfrac{|j|}{s+1}\alpha_{j+\nu}, & j = -\nu, \dots, k-\nu \ \text{ (optimal circulant)}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.9}$$

The entries of $E_B^{(1)}$ are defined similarly to (4.9), but with β_j instead of α_j. For Strang's natural circulant it is easy to check that $E_A^{(1)}$, $E_B^{(1)}$ are the null matrix; see (3.2), (3.4).

From the previous equations, we have

$$\begin{aligned} P^{-1}M = I + P^{-1}E &= I + P^{-1}(E_A^{(1)} \otimes I_m - h\, E_B^{(1)} \otimes J) \\ &\quad + P^{-1}(E_A^{(2)} \otimes I_m - h\, E_B^{(2)} \otimes J) = I + \tilde{M}^{(1)} + \tilde{M}^{(2)}, \end{aligned} \tag{4.10}$$

where $E_A^{(2)}$, $E_B^{(2)}$ are $(s+1)\times(s+1)$ matrices whose entries are nonzero at most in the following four corners:

1. $\nu \times (k+1)$ in the upper left;
2. $\nu \times \nu$, upper right;
3. $(k-\nu)\times(k+1)$, lower right;
4. $(k-\nu)\times(k-\nu)$, lower left.

From the above arguments, we have that

$$\operatorname{rank}(E_A^{(2)}) \le k+1, \qquad \operatorname{rank}(E_B^{(2)}) \le k+1,$$

and, if P^{-1} is well defined,

$$\operatorname{rank}(\tilde{M}^{(2)}) \le 2m\,(k+1),$$

independently on the size of $\breve{A}$, $\breve{B}$. Thus, if $\breve{A}$, $\breve{B}$ are the Strang's circulants as (3.2), (3.4), the thesis follows by setting $M_\delta^{(2)} = \tilde{M}^{(2)}$ and by observing that $C_\delta = \|M_\delta^{(1)}\| = 0$.

Let us consider $\tilde{M}^{(1)}$ when $\breve{A}$, $\breve{B}$ are P-circulant matrices. From (4.10) we have

$$\tilde{M}^{(1)} = \frac{1}{s+1} P^{-1} \left(\hat{E}_A^{(1)} \otimes I_m - h\, \hat{E}_B^{(1)} \otimes J \right), \tag{4.11}$$

where $\hat{E}_A^{(1)} = (s+1)E_A^{(1)}$, $\hat{E}_B^{(1)} = (s+1)E_B^{(1)}$ are banded Toeplitz matrices whose infinity norm is constant with respect to their dimension $s+1$.

We are interested in the behavior of $P^{-1}M$ for $s \to \infty$ (and then for $h \to 0$, $h = (T - t_0)/s$; see section 2.1). To this end, for the spectral properties of the P-circulant matrices $\breve{A}$, $\breve{B}$ (see [2] and section 3.1), $\tilde{M}^{(1)}$ in (4.10) can be written as the sum of a small rank and a small $l2$-norm matrix. Indeed, as in the proofs of [21, Theorems 3.2, 3.4], let us fix $\delta > 0$ and consider the cardinality C_δ of the set

$$\left\{ l \in \{0, \ldots, s\} : |\lambda_l(\breve{A})| \leq \frac{1}{\delta} \frac{1}{s+1} \right\},$$

where $\lambda_l(\breve{A})$ are the $s+1$ distinct eigenvalues of $\breve{A}$. For (3.8), $\lambda_l(\breve{A}) = \phi_k(x_l)$, $x_l = 2\pi l/(s+1)$, $l = 0, \ldots, s$, where

$$\phi_k(x) = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} \left(1 + \frac{j}{s+1} \right) e^{\mathbf{i}jx}. \tag{4.12}$$

The expression $\phi_k(x)$ is a trigonometric polynomial and, as observed in [2], for the formulas in (2.1),

$$\lim_{s\to\infty} \phi_k(x_0) = \lim_{s\to\infty} \phi_k(0) = 0,\ \phi_k'(x_0) \neq 0, \qquad \lim_{s\to\infty} \phi_k'(x_0) \neq 0, \tag{4.13}$$

i.e., x_0 is a simple zero for $\lim_{s\to\infty} \phi_k(x)$ and

$$\frac{c_1}{s+1} \leq |\phi_k(x)| \leq c_2, \qquad \mathrm{Re}\,(\phi_k(x)) > 0, \qquad x \in \mathbb{R}, \tag{4.14}$$

where c_1, c_2 are constants of the order of unity. The matrix $\breve{A}$ is normal (because it is circulant), thus $|\lambda_l(\breve{A})| = |\phi_k(x_l)|$ is a singular value. Moreover, for (4.13), (4.14), and (4.12), if l is small with respect to s,

$$|\phi_k(x_l)| \geq \hat{c} \frac{l}{s+1},$$

and, as a consequence,

$$C_\delta \leq \# \left\{ l : \hat{c} \frac{l}{s+1} \leq \frac{1}{\delta} \frac{1}{s+1} \right\} \leq \left\lceil \frac{1}{\delta\, \hat{c}} \right\rceil, \tag{4.15}$$

which does not depend on s since $\hat{c}$ is a constant.

Recalling that the circulant matrices are simultaneously diagonalized by the matrix F in (3.13) (see [10]), consider the following splitting:

$$\frac{1}{s+1} \breve{A}^{-1} = \hat{\Theta}_1 + \hat{\Theta}_2, \tag{4.16}$$

where the $s+1$ singular values of the circulant matrices $\hat{\Theta}_2$, $\hat{\Theta}_1$ are

$$\{a_0, \ldots, a_{C_\delta - 1},\ 0, \ldots, 0\}, \qquad \{0, \ldots, 0, a_{C_\delta}, \ldots, a_s\}, \tag{4.17}$$

respectively, and $\delta < a_j$, $j = 0, \ldots, C_\delta - 1$; $a_j < \delta$, $j = C_\delta, \ldots, s$, $a_j \geq a_i \geq 0$ if $0 \leq j < i \leq s$.

Thus, from (4.11) and (4.16) we can set $\tilde{M}^{(1)}$ as

$$\tilde{M}^{(1)} = \left(I_{s+1} \otimes I_m - h\breve{A}^{-1}\breve{B} \otimes J\right)^{-1} \left((\hat{\Theta}_1 + \hat{\Theta}_2) \otimes I_m\right) \left(\hat{E}_A^{(1)} \otimes I_m - h\hat{E}_B^{(1)} \otimes J\right),$$

i.e., we can split the above matrix as

$$\tilde{M}^{(1)} = M_\delta^{(1)} + \tilde{M}_\delta^{(2)},$$

where the products containing $\hat{\Theta}_2$ are collected in $\tilde{M}_\delta^{(2)}$, while those containing $\hat{\Theta}_1$ are in $M_\delta^{(1)}$. As a consequence of the above arguments and of (4.10), (4.11), fixed $\delta > 0$, we can find s_δ such that, for all $s \geq s_\delta$,

$$\tilde{M}^{(1)} = M_\delta^{(1)} + \tilde{M}_\delta^{(2)}, \quad \operatorname{rank}(\tilde{M}_\delta^{(2)}) \leq m\, C_\delta, \quad \|\hat{\Theta}_1 \hat{E}_A^{(1)}\|_2 \leq \delta\, c, \quad \|\hat{\Theta}_1 \hat{E}_B^{(1)}\|_2 \leq \delta\, c,$$

where c does not depend on s and usually is of the order of unity. If we define $M_\delta^{(2)} = \tilde{M}^{(2)} + \tilde{M}_\delta^{(2)}$, the first half of the thesis follows. To check that $M_\delta^{(1)}$ is a small norm matrix, we can transform $M_\delta^{(1)}$ by the matrices $(F \otimes V^{-1})$, $(F^* \otimes V)$, $J = VDV^{-1}$. (We suppose for simplicity that the Jacobian matrix J is diagonalizable.) Thus, if $N = (F \otimes V^{-1}) M_\delta^{(1)} (F^* \otimes V)$, we have

$$\begin{aligned} N &= \left(I_{s+1} \otimes I_m - h\Lambda_A^{-1}\Lambda_B \otimes D\right)^{-1} (F \otimes I_m) \\ &\quad \cdot (\hat{\Theta}_1 \otimes I_m) \left(\hat{E}_A^{(1)} \otimes I_m - h\hat{E}_B^{(1)} \otimes D\right) (F^* \otimes I_m) \\ &= \left(I_{s+1} \otimes I_m - h\Lambda_A^{-1}\Lambda_B \otimes D\right)^{-1} (\Lambda_1 F \otimes I_m) \\ &\quad \cdot \left(\hat{E}_A^{(1)} \otimes I_m - h\hat{E}_B^{(1)} \otimes D\right) (F^* \otimes I_m), \end{aligned} \tag{4.18}$$

where $\hat{\Theta}_1 = F^*\Lambda_1 F$, $\breve{A} = F^*\Lambda_A F$, $\breve{B} = F^*\Lambda_B F$, and Λ_1, Λ_A, Λ_B are diagonal matrices. From (4.18) and the above arguments we have the following bound:

$$\begin{aligned} \|N\|_2 &\leq \| \left(I_{s+1} \otimes I_m - h\Lambda_A^{-1}\Lambda_B \otimes D\right)^{-1} \|_2 \|F\hat{\Theta}_1 E_A^{(1)} F^* \otimes I_m\|_2 \\ &\quad + \max_r \|(I_{s+1} \otimes I_m - h\mu_r \Lambda_A^{-1}\Lambda_B)^{-1} (h\mu_r \Lambda_1 F\, \hat{E}_B^{(1)} F^*)\|_2 \\ &\leq \frac{\|F\hat{\Theta}_1 \hat{E}_A^{(1)} F^*\|}{\min_{j,r} \left|1 + h(-\mu_r)\dfrac{\psi_j}{\phi_j}\right|} + \frac{\|\Lambda_1 F \hat{E}_B^{(1)} F^*\|_2}{\min_{j,r} \left|\dfrac{1}{h(-\mu_r)} + \dfrac{\psi_j}{\phi_j}\right|} \\ &\leq \frac{\|\hat{\Theta}_1 \hat{E}_A^{(1)}\|_2}{c_{J,1}} + \frac{\|\hat{\Theta}_1 \hat{E}_B^{(1)}\|_2}{c_{J,2}} \leq \delta\, \hat{c}_J. \end{aligned} \tag{4.19}$$

Excluding the trivial case $\mu_r = 0$, we have that $\operatorname{Re}(1/(h(-\mu_r))) \geq 0$ if the Jacobian matrix J has eigenvalues whose real parts are nonpositive, and $c_{J,1}$, $c_{J,2}$ in the above expression can be bounded uniformly in s if $\operatorname{Re}(\psi_j/\phi_j) \geq \epsilon > 0$.We have verified that

this holds, e.g., for the methods of section 2.1. Thus, $\hat{c}_J$, c_J and then the bound for $\|M_\delta^{(1)}\|_2$ are independent from s.

Notice that when $h\,\mu_r$ is small, such as in the nonstiff case, we have simply that

$$\|M_\delta^{(1)}\|_2 \le c_2 \|\hat{\Theta}_1 \hat{E}_A^{(1)}\|_2 \le \delta\, c\, c_2,$$

where c_2 is a constant. □

Unfortunately, we cannot state a similar result for the block preconditioner (3.1) based on the optimal circulants because of the ill-conditioning arising when k is large (see [2]). Nonetheless, if (4.14) hold true, for fixed k suitably small, similar arguments to those used for P-circulants can be used to prove convergence.

Despite the fact that $\|M_\delta^{(1)}\|$ and C_δ are zero for P given by (3.1) using Strang's approximations, we have observed that such a preconditioner is often less suitable than others. Indeed, as an example, the Strang's circulants can be severely ill-conditioned or even singular, e.g., for problems such that some of the eigenvalues of their Jacobian matrix have zero or small modulus real part (or imaginary part large in absolute value with respect to the real part). Unfortunately, this is not infrequent, e.g., for stiff problems; see Example 1 and the remarks in the next section.

5. Numerical results. To show the effectiveness of our block preconditioner, we will integrate some test problems with the formulas (2.9), a generalization of Adams–Moulton methods. Such formulas have been found to be quite effective (see [17]).

Further numerical examples using other formulas can be found in [1].

We will compare the number of matrix-vector products needed to converge for Bi-CGSTAB [25] (Bi-CGSTAB(2) [22] for Example 1 because Bi-CGSTAB has shown much more erratic unpreconditioned convergence) and GMRES [20].

The initial guess for the iterative solvers is always zero. The stopping criterion is $\|r_k\| < 10^{-6}\|b\|$, r_k true residual. The stability properties of the LMF considered here (see [3]) allow an application without stepsize restriction. Hence, we will use a constant stepsize h. All calculations are done in Matlab.

In the columns labeled I_s, P_s, C_s and S_s in the tables below, we will give the number of matrix-vector products needed to the convergence of the unpreconditioned iterations, preconditioned iterations using P-circulants, and Chan's and Strang's approximations, respectively. A "−" means that convergence was not attained.

The column labeled "rfp" gives a rough estimate of the ratio of the floating points operations required for, respectively, the P-circulant preconditioned and the unpreconditioned GMRES.

Example 1. *Wave equation.*

Let us consider the wave equation

$$\begin{cases} u_{tt} - c u_{xx} = 0, \\ u(x,0) = g_1(x),\ u_t(x,0) = g_2(x), & x \in [0,\pi], \\ u(0,t) = u(\pi,t) = 0, & t \in [0,T]. \end{cases} \tag{5.1}$$

Approximating the operator $\partial^2/\partial x^2$ in (5.1) with centered differences and reducing the obtained IVP to the first order gives the system of $2N$ ODEs:

$$\begin{cases} y'(t) = H_{2N} y(t), & t \in [0,T], \\ y(0) = \eta, & \eta = (g_1(x_1)\ \cdots\ g_1(x_N) g_2(x_{N+1})\ \cdots\ g_2(x_{2N}))^T. \end{cases} \tag{5.2}$$

TABLE 5.1
Wave equation (5.1). *Number of matrix-vector products, GAM with* $k = 3$.

			GMRES				BiCGstab(2)				
N	s	h	I_s	P_s	C_s	S_s	I_s	P_s	C_s	S_s	rfp
20	8	$\pi/4$	298	31	32	31	>1000	40	44	64	.4e-1
50	8	$\pi/4$	776	34	35	45	-	45	52	81	.7e-3
100	8	$\pi/4$	>1000	34	37	71	-	48	52	109	.2e-3
20	16	$\pi/8$	311	36	38	40	609	45	49	104	.5e-1
50	16	$\pi/8$	>1000	42	46	62	-	52	60	206	.6e-4
100	16	$\pi/8$	>1000	42	45	60	-	52	56	197	-
20	32	$\pi/16$	203	35	37	42	744	44	49	125	.1
50	32	$\pi/16$	764	44	48	66	-	49	60	333	.1e-1
100	32	$\pi/16$	>1000	45	50	80	-	52	60	588	-

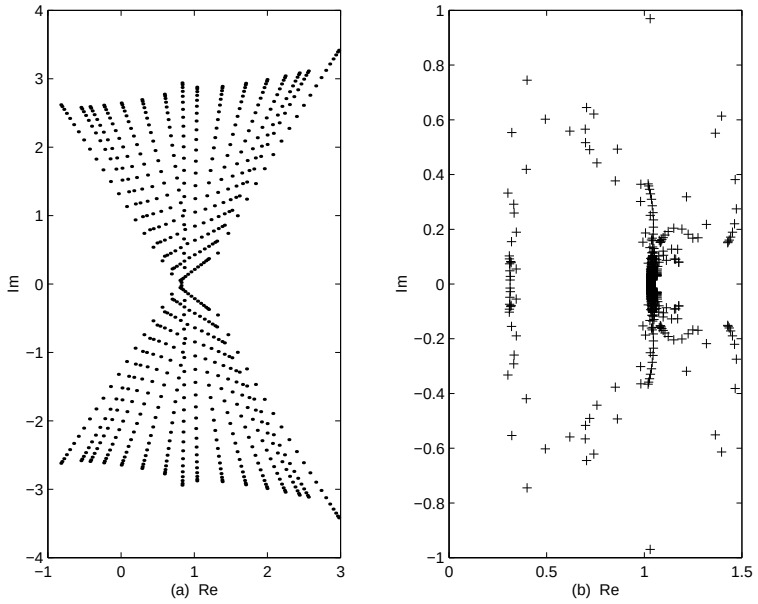

FIG. 5.1. *Wave equation* (5.1). *Spectrum of eigenvalues of the matrix* M (2.6) *before and after P-circulant preconditioning.* ($c = 1$, *GAM with* $k = 3$, $s = 16$, $N = 20$, $T = 2\pi$.)

The matrix H_{2N} is a Hamiltonian definite one:

$$H_{2N} = \begin{pmatrix} \mathbf{0} & I_N \\ T_N & \mathbf{0} \end{pmatrix}_{2N \times 2N}. \tag{5.3}$$

The matrix H_{2N} has the spectrum of eigenvalues on the imaginary axis. In Table 5.1, we can see the effect of the block preconditioner on the number of iterations needed to solve the IVP (5.2) using the fourth order GAM ($k = 3$), $T = 2\pi$, $g_1(x) = 0$, $g_2(x) = x$. In Figures 5.1 and 5.2 we can see the effect on the spectrum of eigenvalues of the matrix M.

Notice that the block preconditioner based on the Strang's circulants does not show the best preconditioned behavior, despite the fact that $M_\delta^{(1)}$, the small norm matrix in (4.7), is zero; see also Table 5.2. To this end, see also the remarks at the end of this section.

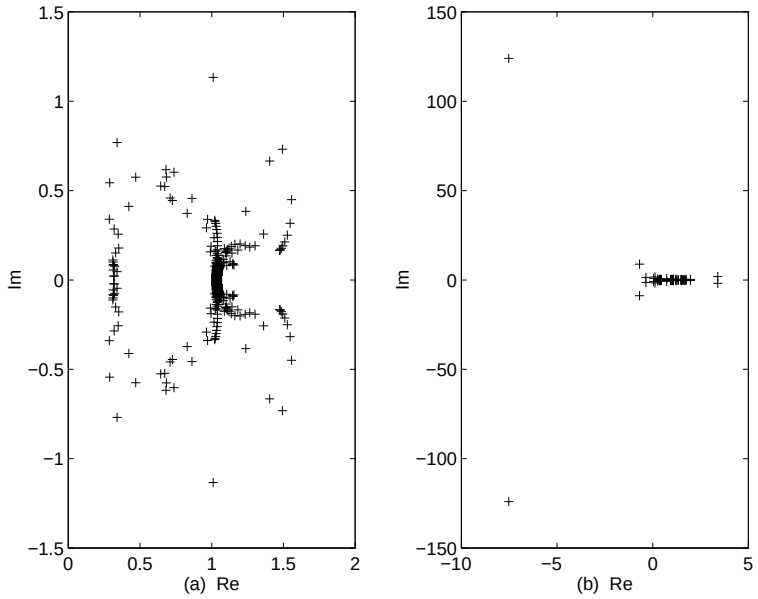

FIG. 5.2. *Wave equation* (5.1). *Spectrum of eigenvalues of the matrix* $P^{-1}M$ *as in Figure* 5.1(b) *but after block preconditioning using* (a) *Chan's and* (b) *Strang's circulants.*

TABLE 5.2
Wave equation (5.1)*, same parameters as Table* 5.1. *Comparison of the computational costs related to the block preconditioner using Strang's approximations.*

			GMRES			BiCGstab(2)		
N	s	size	P_s	C_s	S_s	P_s	C_s	S_s
20	8	320	1	1.04	1	.63	.7	1
50	8	800	.71	.73	1	.57	.66	1
100	8	1600	.4	.43	1	.45	.49	1
20	16	640	.87	.94	1	.43	.48	1
50	16	1600	.6	.68	1	.24	.28	1
100	16	3200	.27	.29	1	.63	.69	1
20	32	1280	.8	.85	1	.35	.39	1
50	32	3200	.15	.188	1	.59	.66	1
100	32	6400	.089	.1	1	.47	.57	1

Example 2. *Heat equation in a rectangle.*

Consider the two-dimensional heat equation defined in a rectangular domain:

$$\begin{cases} u_t - (u_{xx} + u_{yy}) = 0, & (x,y) \in \Omega = [0,\pi] \times [0,\pi], \\ u((x,y),0) = g(x,y), & (x,y) \in \Omega, \\ u((x,y),t) = 0, & (x,y) \in \partial\Omega,\ 0 \le t \le T. \end{cases} \tag{5.4}$$

Using centered differences to approximate the Laplacian operator in the rectangle Ω with a uniform grid

$$\Delta x = \Delta y = \Delta, \quad \Delta = \pi/(N+1)$$

gives the IVP

$$\begin{cases} y'(t) = \dfrac{1}{\Delta^2} \hat{L}_N y(t), & t \in [0,T], \\ y(0) = \eta, & \eta = (g(x_1,y_1)), \ldots, g(x_N,y_N)^T, \end{cases} \tag{5.5}$$

TABLE 5.3
Heat equation (5.4). *Number of matrix-vector products, GAM with* $k = 4$.

			GMRES				BiCGstab				
N	s	h	I_s	P_s	C_s	S_s	I_s	P_s	C_s	S_s	rfp
4	8	$\pi/4$	51	8	8	7	74	14	14	12	.3
8	8	$\pi/4$	121	8	8	7	216	14	14	12	.4
20	8	$\pi/4$	368	7	7	6	631	14	16	12	.5e-1
4	16	$\pi/8$	53	7	7	6	72	12	10	10	.3
8	16	$\pi/8$	126	7	7	6	175	12	10	10	.5
20	16	$\pi/8$	376	6	6	6	648	12	10	10	.5e-1
4	24	$\pi/12$	45	7	7	7	61	10	10	10	.6
8	24	$\pi/12$	107	7	7	7	154	12	12	10	.2
20	24	$\pi/12$	320	6	6	6	518	12	12	10	.7e-1

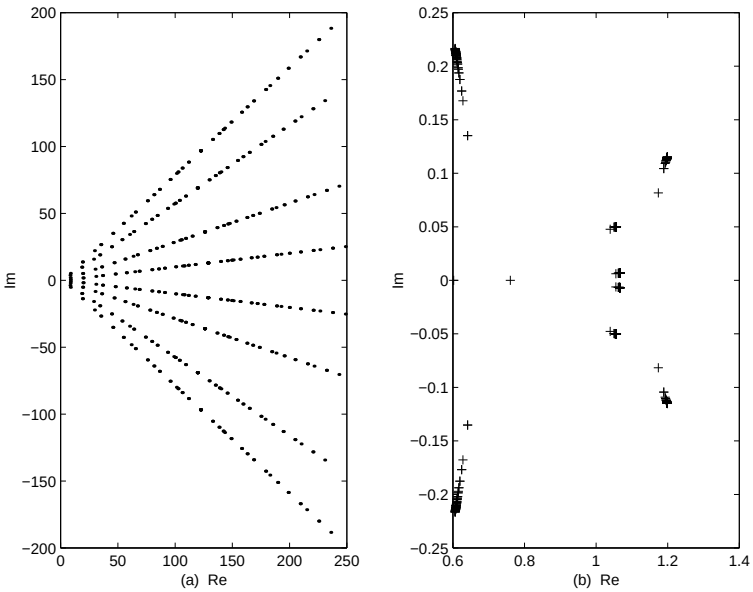

FIG. 5.3. *Heat equation* (5.4). *Spectrum of the eigenvalues of the matrix* M (2.6) *before and after P-circulant block preconditioning. (GAM,* $k = 4$, $s = 16$, $N = 20$.*)*

where $\hat{L}_N$ is a $N^2 \times N^2$ block tridiagonal matrix:

$$\hat{L}_N = \begin{pmatrix} \hat{T}_N & I_N & & \\ I_N & \ddots & \ddots & \\ & \ddots & \ddots & I_N \\ & & I_N & \hat{T}_N \end{pmatrix}, \qquad \hat{T}_N = \begin{pmatrix} -4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -4 \end{pmatrix} \tag{5.6}$$

and x_i, y_j, $i, j = 1, \ldots, N$, in (5.5) are defined accordingly. In Table 5.3, we can see the effect of the block preconditioner on the number of iterations needed to solve the IVP (5.5) using fifth order GAM ($k = 4$), $T = 2\pi$, $g(x, y) = x\,y$. In Figure 5.3 we can see the effect on the spectrum of eigenvalues of the matrix M.

As an example, notice that the built-in Matlab sparse direct solver for the linear system (2.6) used for Example 2 with $N = 20$, $s = 8$ ($N = 20$, $s = 16$) needs 6

(18) times more flops than the P-circulant block preconditioned GMRES. In general, we can save flops provided that M in (2.6) is large enough and the convergence of preconditioned iterations is achieved in a moderate number of iterations.

Example 3. *Wave equation of first order.*

$$\begin{cases} u_t - u_x = 0, \\ u(x,0) = g(x), & x \in [0,\pi], \\ u(\pi,t) = 0, & t \in [0,2\pi]. \end{cases} \tag{5.7}$$

We discretize the partial derivative $\partial/\partial x$ with the first order forward difference and stepsize $\Delta x = \pi/N$, $x_j = j\Delta x$ (upwind discretization). We obtain the system of N ODEs

$$\begin{cases} y'(t) = L_N y(t), & t \in [0,2\pi], \\ y(0) = \eta, & \eta = (g(x_0) \ \cdots \ g(x_{N-1}))^T, \end{cases} \tag{5.8}$$

$$L_N = \frac{1}{\Delta x} \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & -1 \end{pmatrix}_{N\times N}.$$

In Table 5.4, we can see the effect of the block preconditioner on the number of iterations needed to solve the IVP (5.7) using fifth order GAM ($k = 4$), $T = 2\pi$, $g(x) = \sqrt{x(\pi - x)}$. In Figure 5.4 we can see the effect on the spectrum of eigenvalues of the matrix M.

As expected, the block preconditioner (3.1) is effective for those classes of differential problems for which the preconditioned matrix $P^{-1}M$, M in (2.6), has a clustered spectrum. For Theorem 4.1, this is true if $M_\delta^{(1)}$ in (4.7) either is zero (as is the case of Strang's circulants) or is a suitable perturbation of the null matrix (e.g., has eigenvalues clustered around the origin of the complex plane, as is the case of P-circulants) *and* P is not ill-conditioned. Moreover, the small rank matrix $M_\delta^{(2)}$ should have few and possibly clustered outliers (i.e., the eigenvalues outside the multiple eigenvalue in the origin of $\mathbb{C}$).

Notice that the above properties may not hold if the preconditioner is ill-conditioned, e.g., because $\breve{A}$ or $\breve{B}$ are ill-conditioned or even singular. As an example, see Example 1 for the block preconditioner using Strang's or Chan's approximations. A similar behavior can be observed for Example 2 if a fast decaying diffusion coefficient is considered. The effect of the ill-conditioning on the convergence of preconditioned iterations can be reduced significantly if special initial conditions for the underlying partial differential equations are considered; see the numerical tests in [9]. More details and analysis of the rate of convergence of preconditioned iterations will be given in a forthcoming paper.

In this section, we have considered methods based on the formula (2.9). In general, the same behavior can be observed also for the other formulas of section 2.1, e.g., (2.10) for Example 1 and (2.8) for Examples 2 and 3.

6. Concluding remarks. In this work we have considered a block circulant preconditioner we introduced in [1] for the linear systems of certain codes for ordinary differential equations. Such preconditioners are used here implicitly with iterative Krylov subspace methods for nonsymmetric systems such as Bi-CGSTAB, Bi-CGSTAB(2) and GMRES.

TABLE 5.4
Wave equation of first order (5.7). *Number of matrix-vector products, GAM with* $k = 4$.

			GMRES				BiCGstab				
N	s	h	I_s	P_s	C_s	S_s	I_s	P_s	C_s	S_s	rfp
20	8	$\pi/4$	38	10	9	10	88	10	10	10	.2
50	8	$\pi/4$	100	12	11	12	-	12	12	13	.7e-1
100	8	$\pi/4$	239	13	12	13	-	18	16	18	.1e-1
20	16	$\pi/8$	28	9	8	8	49	10	10	10	.7
50	16	$\pi/8$	87	10	9	9	-	12	12	12	.7e-1
100	16	$\pi/8$	222	10	10	10	-	12	14	12	.1e-1
20	32	$\pi/16$	36	7	7	7	41	8	8	9	.3
50	32	$\pi/16$	70	8	8	8	179	9	9	9	.1
100	32	$\pi/16$	176	9	9	9	-	10	12	10	.2e-1

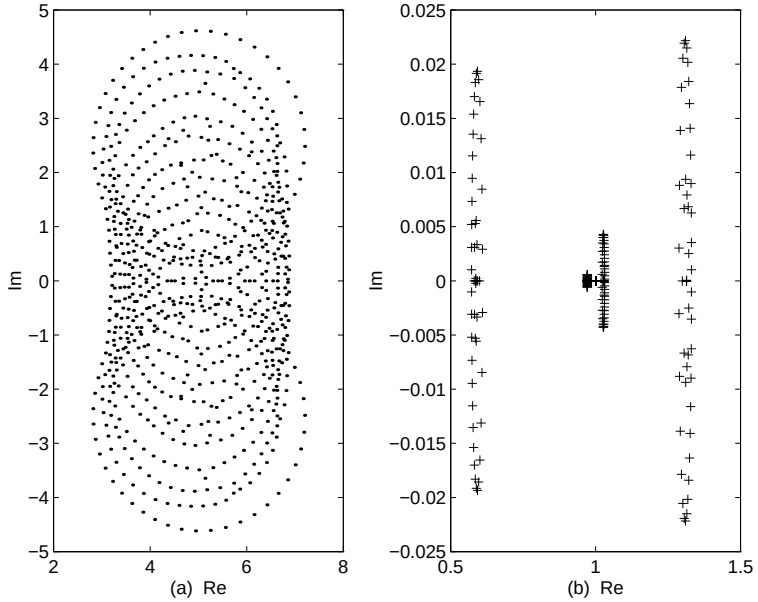

FIG. 5.4. *Wave equation of first order* (5.7). *Spectrum of the eigenvalues of the matrix* M (2.6) *before and after P-circulant preconditioning. (GAM,* $k = 3$, $s = 16$, $N = 20$.*)*

It has been observed that, for some classes of differential problems, the preconditioned iterations are almost independent from the discretization.

Moreover, a recently introduced circulant approximation, called P-circulant, has been found to be promising in comparison to optimal approximations, in the sense of the norm (see, e.g., [8, p. 432–434]), such as Strang's and Chan's. This has been confirmed by the analysis in [2].

Despite the fast convergence, the above block preconditioner has a moderate theoretical parallel complexity and an interesting serial computational cost for the classes of differential problems such that the spectrum of the block preconditioned matrix $P^{-1}M$ is clustered.

Acknowledgments. The author would like to thank both the referees and Michele Benzi of the Los Alamos National Laboratory for very helpful comments that have improved this presentation. The author is grateful to Luigi Brugnano and Donato Trigiante of the University of Firenze for their suggestions and constructive criticism.

REFERENCES

[1] D. Bertaccini, *P-circulant preconditioners and the systems of ODE codes*, in Iterative Methods in Scientific Computation IV, IMACS Ser. Comput. Appl. Math., D. R. Kincaid and A. C. Elster, eds, New Brunswick, NJ, 1999, pp. 179–193.

[2] D. Bertaccini, *An Analysis of Circulant Preconditioners for Linear Multistep Integrators*, Math. Comp., submitted.

[3] L. Brugnano and D. Trigiante, *Solving ODE by Linear Multistep Methods: Initial and Boundary Value Methods*, Gordon and Breach, Reading, UK, 1998.

[4] T. F. Chan and K. R. Jackson, *The use of iterative linear-equation solvers in codes for large systems of stiff IVPs for ODEs*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 378–417.

[5] T. F. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[6] R. H. Chan, *Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.

[7] R. H. Chan and M.-C. Yeung, *Circulant preconditioners for complex Toeplitz matrices*, SIAM J. Numer. Anal., 30 (1993), pp. 1193–1207.

[8] R. H. Chan and M. K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.

[9] R. H. Chan, M. K. Ng, and X. Jin, *Circulant preconditioners for solving ordinary differential equations*, in Structured Matrices, D. Bini, E. Tyrtyshnikov, and P. Yalamov, eds., Nova Science, to appear.

[10] P. J. Davis, *Circulant Matrices*, John Wiley, New York, 1979.

[11] F. Di Benedetto, G. Fiorentino, and S. Serra, *C.G. preconditioning for Toeplitz matrices*, Comput. Math. Appl., 25 (1993), pp. 33–45.

[12] C. W. Gear and Y. Saad, *Iterative solution of linear equations in ODE codes*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 583–601.

[13] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II. *Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin, Heidelberg, 1991.

[14] S. Holmgreen and K. Otto, *A framework for polynomial preconditioners based on fast transforms* I: *Theory*, BIT, 38 (1998), pp. 544–559.

[15] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994.

[16] T. Huckle, *Some aspects of circulant preconditioners*, SIAM J. Sci. Comput., 14 (1993), pp. 531–541.

[17] F. Iavernaro and F. Mazzia, *Solving ordinary differential equations by generalized Adams methods: Properties and implementations techniques*, Appl. Numer. Math. 28 (1198), pp. 107–126.

[18] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, John Wiley, Chichester, 1991.

[19] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Comput. Sci. Appl. Math., Academic Press, New York, 1985.

[20] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[21] S. Serra, *The rate of convergence of Toeplitz based PCG methods for second order nonlinear boundary value problems*, Numer. Math., 81 (1999), pp. 461–495.

[22] G. L. G. Sleijpen, H. A. van der Vorst, and D. R. Fokkema, *BiCGstab(l) and other hybrid Bi-CG methods*, Numer. Algor., 7 (1994), pp. 75–109.

[23] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74, (1986), pp. 171–176.

[24] V. V. Strela and E. E. Tyrtyshnikov, *Which circulant preconditioner is better?*, Math. Comp., 65 (1996), pp. 137–150.

[25] H. A. van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 10 (1992), pp. 631–644.