BIT
2001, Vol. 41, No. 3, pp. 433–450
0006-3835/01/4103-0433 $16.00

THE CONVERGENCE RATE OF BLOCK PRECONDITIONED SYSTEMS ARISING FROM LMF-BASED ODE CODES *

DANIELE BERTACCINI [1] and MICHAEL K. NG[2†]

[1] *Dipartimento di Matematica "G. Castelnuovo", Universitá di Roma "La Sapienza" P. le Aldo Moro 2, IT-00185 Roma, Italy. email: bertaccini@uniroma1.it*

[2] *Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. email: mng@maths.hku.hk*

Abstract.

The solution of ordinary and partial differential equations using implicit linear multistep formulas (LMF) is considered. More precisely, boundary value methods (BVMs), a class of methods based on implicit formulas will be taken into account in this paper. These methods require the solution of large and sparse linear systems $\hat{M}\mathbf{x} = \mathbf{b}$. Block-circulant preconditioners have been proposed to solve these linear systems. By investigating the spectral condition number of $\hat{M}$, we show that the conjugate gradient method, when applied to solving the normalized preconditioned system, converges in at most $O(\log s)$ steps, where the integration step size is $O(1/s)$. Numerical results are given to illustrate the effectiveness of the analysis.

AMS subject classification: 65F10, 65N22.

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

1 Introduction.

We consider the solution of ordinary differential equations (ODEs) by using implicit linear multistep formulas (LMF). By applying the above formulas, the solution to a given ODE is given by the solution of a linear system

$$\hat{M}\mathbf{y} = \mathbf{b}, \tag{1.1}$$

where $\hat{M}$ depends on the LMF used. In this paper, we concentrate on the linear initial value problem

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{y}(t)}{\mathrm{d}t} = J\mathbf{y}(t) + \mathbf{g}(t), \quad t \in (t_0, T], \\ \\ \mathbf{y}(t_0) = \mathbf{z}, \end{cases} \tag{1.2}$$

*Received April 2000. Communicated by Lothar Reichel.

†Research supported in part by Hong Kong Research Grants Council Grant Nos. HKU 7147/99P and 7132/00P and HKU CRCG Grant No. 10203408.

where $\mathbf{y}(t)$, $\mathbf{g}(t) : \mathbb{R} \to \mathbb{R}^m$, $\mathbf{z} \in \mathbb{R}^m$, and $J \in \mathbb{R}^{m\times m}$. By applying BVMs on this ODE, the matrix $\hat{M}$ in (1.1) can be written as

$$\hat{M}_s = \hat{A}_s \otimes I - h\hat{B}_s \otimes J, \tag{1.3}$$

where $\hat{A}_s$ and $\hat{B}_s$ are $(s+1)$-by-$(s+1)$ matrices with entries depending on the LMF used, I is the m-by-m identity matrix and $h = (T - t_0)/s$ is the integration step size. In this paper, we assume that m is fixed and independent of s. We note that both $\hat{A}_s$ and $\hat{B}_s$ can be reduced to a sum of a Toeplitz matrix plus a low rank matrix. (A matrix is said to be Toeplitz if its entries are constant along its diagonals.) The size of the matrix $\hat{M}_s$ is very large when h is small and/or m is large. If a direct method is used to solve the system, the operation count can be too expensive and slow for practical, large scale applications.

In [4, 5], the first author proposed to use Krylov subspace methods such as the GMRES method [22] to solve (1.1). In order to speed up the convergence rate of Krylov subspace methods, he proposed circulant matrices as preconditioners. The first preconditioner proposed in [4, 5] for the matrix $\hat{M}_s$ in (1.1) is the well-known T. Chan circulant preconditioner; see [13]. The second one proposed in [4, 5] is a new preconditioner called the P-circulant preconditioner P. Moreover, Bertaccini [4] and Chan et al. [15] proposed the Strang-type preconditioner for (1.1). They showed that both the P-circulant and Strang-type preconditioned systems converge very quickly. However, when the Jacobian matrix J has some small (or zero) singular values (see, e.g., [4, 7, 8]), or if some eigenvalues have imaginary part large in absolute value relative to the real part (see, e.g. [5]), the Strang-type preconditioner can be very ill-conditioned or even singular. Therefore, in this paper, we will focus on P-circulant preconditioners.

The main aim of this paper is to study the convergence rate of these normalized block P-circulant preconditioned systems when we use the conjugate gradient method. We will investigate the spectral condition number $\kappa(\hat{M}_s)$ of $\hat{M}_s$, and show that $\kappa(\hat{M}_s)$ is at most $O(s)$, where $1/s$ is integration step size. In [5] it has been shown that the spectra of the P-circulant preconditioned matrices $P^{-1}\hat{M}_s$ are clustered around one. From that we get a bound on the convergence rate of the conjugate gradient method that depends on the spectral condition number of $\hat{M}_s$. It follows that the number of iterations required for convergence is at most $O(\log s)$. By noting that the number of operations per iteration of the conjugate gradient method is of $O(s \log s)$, the total complexity of the algorithm is therefore of $O(s \log^2 s)$. In the case when $\hat{M}_s$ is well-conditioned, the method converges in $O(1)$ steps. Hence the complexity is reduced to $O(s \log s)$.

The paper is organized as follows. In Section 2, we recall the general scheme of BVMs. We study the condition numbers of matrices arising from BVMs in Section 3. In Section 4, we introduce some circulant preconditioners and analyze the convergence rate of the preconditioned matrices. Finally, numerical examples and concluding remarks are given in Section 5.

2 Boundary value methods and their matrix forms.

The boundary value methods (BVMs) for differential equations are a generalization as a boundary value technique of the implicit linear multistep formulas

(see, e.g., [2, 11] and the references therein). In order to briefly describe them, suppose for simplicity that we have the linear IVP (1.2). A BVM approximates the solution of (1.2) 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 $t_j = t_0 + j\,h$, $j = 0, \ldots, s$, $h = (T - t_0)/s$:

$$\sum_{i=-\nu}^{k-\nu} \alpha_{i+\nu} \mathbf{y}_{n+i} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu} \mathbf{f}_{n+i}, \quad n = \nu, \ldots, s-k+\nu. \tag{2.1}$$

As usual, $\mathbf{y}_n$ is the discrete approximation to $\mathbf{y}(t_n)$, $\mathbf{f}_n = J\,\mathbf{y}_n + \mathbf{g}_n$, $\mathbf{g}_n = \mathbf{g}(t_n)$, and the values

$$\mathbf{y}_0, \ldots, \mathbf{y}_{\nu-1}, \quad \mathbf{y}_{s-k+\nu+1}, \ldots, \mathbf{y}_s \tag{2.2}$$

are given. We observe that the IVP (1.2) provides only the initial value $\mathbf{y}_0$. It is possible to avoid to supply the other conditions in (2.2) by coupling the *main method* (2.1) with other difference schemes of order p_1, where usually $p_1 = k$, called *additional methods*, which provide the following set of equations:

$$\sum_{i=0}^{q_1} \alpha_i^{(j)} \mathbf{y}_i = h \sum_{i=0}^{q_1} \beta_i^{(j)} \mathbf{f}_i, \quad j = 1, \ldots, \nu-1, \tag{2.3}$$

$$\sum_{i=0}^{q_2} a_{k-i}^{(j)} \mathbf{y}_{s-i} = h \sum_{i=0}^{q_2} \beta_{k-i}^{(j)} \mathbf{f}_{s-i}, \quad j = s-k+\nu+1, \ldots, s, \tag{2.4}$$

independent of those in (2.1) and where $k \le q_1, q_2 \le s$ (e.g., $q_1 = q_2 = k$ in the methods considered in [4, 5]). The equations (2.1), (2.3), and (2.4) define the use of a BVM on problem (1.2).

Let us cast BVMs in matrix form. This is done by introducing the matrices $\hat{A}_s$, $\hat{B}_s \in \mathbb{R}^{(s+1)\times(s+1)}$. If $q_1 = k$, we have

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

and $\hat{B}_s$ similarly, but with β_js instead of α_js, and all zeros in its first row. The discrete problem generated by the application of the BVM (2.1)–(2.4) to problem

(1.2) is then given by

$$\begin{aligned} \hat{M}_s\,\mathbf{y} &= \mathbf{e}_1 \otimes \eta + h\,(\hat{B}_s \otimes I)\mathbf{g}, \quad \mathbf{e}_1 = (1,0,\ldots,0)^* \in \mathbb{R}^{s+1}, \\ \mathbf{y} &= (\mathbf{y}_0,\ldots,\mathbf{y}_s)^*, \quad \mathbf{g} = (\mathbf{g}_0,\ldots,\mathbf{g}_s)^*, \\ \hat{M}_s &= \hat{A}_s \otimes I - h\,\hat{B}_s \otimes J. \end{aligned} \tag{2.6}$$

The matrix $\hat{M}_s$ in (2.6) turns out to be large and sparse when $s \gg k$, q_1, q_2 and/or J is large and sparse.

2.1 Some families of BVMs.

Here we give the definitions of some families of BVMs (see [11] for details). All considered methods are consistent, i.e., they satisfy the conditions

$$\rho(1) = 0, \qquad \rho'(1) = \sigma(1), \tag{2.7}$$

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

$$\rho(z) = z^{\nu} \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu}\; z^j, \qquad \sigma(z) = z^{\nu} \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu}\; z^j. \tag{2.8}$$

The generalized Adams methods, or GAM, are a generalization of the Adams–Moulton methods (see [17]) as a boundary value technique. They can be written in the following form:

$$\mathbf{y}_{n+\nu} - \mathbf{y}_{n+\nu-1} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu}\, \mathbf{f}_{n+i}, \tag{2.9}$$

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 rule, or ETR, because they share the same stability properties of the trapezoidal rule. Such methods turns out to be well-suited either for approximating Hamiltonian problems or continuous BVPs. When k is even, GAM are well suited for stiff problems (see [11, Chapters 6 and 7] for details).

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

$$\sum_{i=-\nu}^{\nu-1} \alpha_{i+\nu}\, \mathbf{y}_{n+i} = \frac{h}{2}(\mathbf{f}_n + \mathbf{f}_{n-1}), \tag{2.10}$$

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 (2.9), (2.10) 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 both for approximating Hamiltonian problems and continuous BVPs (see [11, Chapter 7] for details).

3 Condition numbers of Toeplitz matrices.

As observed in the previous section, the matrix $\hat{M}_s$ as in (2.6) has a block quasi-Toeplitz pattern. The perturbation to the Toeplitz pattern is due to the presence of the additional methods (2.3) and (2.4) in $\hat{A}_s$ as shown as in (2.5), and similarly in $\hat{B}_s$. In our analysis, we will assume that the effect of the above will be negligible. More precisely, we will consider only those formulas (2.3) and (2.4) such that the condition numbers of the matrices $\hat{M}_s$ and M_s behave similarly with respect to their size $s+1$ (see, e.g., [11, Chapter 11]) where

$$M_s = A_s \otimes I_m - hB_s \otimes J \tag{3.1}$$

is the block Toeplitz matrix obtained from the main formula (2.1). The matrix A_s is a $(s+1)\times(s+1)$ banded Toeplitz matrix, i.e.,

$$A_s = \begin{pmatrix} \alpha_\nu & \cdots & \alpha_{k-1} & \alpha_k & 0 & \cdots & 0 \\ \vdots & \alpha_\nu & \ddots & \alpha_{k-1} & \alpha_k & \ddots & \vdots \\ \alpha_0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \alpha_{k-1} & \alpha_k \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \alpha_{k-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \alpha_0 & \cdots & \alpha_\nu \end{pmatrix}, \tag{3.2}$$

and B_s has the similar pattern. The assumption on the behavior of the condition number of $\hat{M}_s$ and M_s is usually reasonable. This is the case of the families of methods of Section 2.1 and those considered in [4, 5], at least up to a certain number of step k (see, e.g., Chapter 11 in [11]).

Let us consider the Dahlquist test problem

$$\begin{cases} y'(t) &= \lambda y(t), \quad t \in [0,T] \\ y(t_0) &= \eta, \end{cases} \tag{3.3}$$

as an example. Here we choose $J = \lambda$ in (1.2), where $\lambda \in \mathbb{C}$ is a scalar. It is customary to consider the above scalar problem in the theory of methods for ODEs (see, e.g., [20, 11]), where λ can be any one of the eigenvalues of the original Jacobian matrix J, supposed diagonalizable.

In the tables below, we give an example of the 2-norm condition numbers of A_s, B_s, M_s and of their small rank perturbations counterparts for the problem (3.3), $\lambda = -1$ and $\lambda = +1$, $h = 1/s$, the coefficients of A_s, B_s are given in the formulas (2.9) used with $k = 3$ (order four) and (2.10) used with $k = 5$ (order six). As can be observed, they are of the same order of magnitude.

A similar behavior can be observed also for the other schemes described in Section 2.1 and other values of λ, provided that $h > 0$ is chosen so that $\hat{M}_s$, M_s, are nonsingular. More examples can be found in Tables 3.3 and 3.4 for wave and heat equations, respectively, in the next subsection. However, notice that the

Table 3.1: Condition numbers for different sizes of the underlying Toeplitz matrices and their small rank perturbations counterparts, formula (2.9) with $k = 4$ for problem (3.3) with $\lambda = -1$ and $T = 1$.

s	$\kappa(A_s)$	$\kappa(\hat{A}_s)$	$\kappa(B_s)$	$\kappa(\hat{B}_s)$	$\kappa(M_s)$	$\kappa(\hat{M}_s)$
32	41.0	41.0	4.0	9.5	28.0	28.0
64	82.0	82.0	4.0	9.5	56.0	56.0
128	163.0	163.0	4.0	9.5	113.0	113.0
256	226.0	326.0	4.0	9.5	226.0	226.0

Table 3.2: Condition numbers for different sizes of the underlying Toeplitz matrices and their small rank perturbations counterparts. Formula (2.10), $k = 3$ for problem (3.3) with $\lambda = +1$ and $T = 1$.

s	$\kappa(A_s)$	$\kappa(\hat{A}_s)$	$\kappa(B_s)$	$\kappa(\hat{B}_s)$	$\kappa(M_s)$	$\kappa(\hat{M}_s)$
32	41.3	31.6	31.3	118.8	65.4	52.4
64	82.1	59.3	61.8	180.1	129.5	96.0
128	163.6	114.7	123.0	280.4	257.5	183.1
256	326.6	225.5	245.2	459.2	513.5	357.2

matrices of the model problem (3.3) are representative for a general system of ODEs (or PDEs by semidiscretization). Indeed, supposing J diagonalizable,

$$J = VDV^{-1}, \quad D = diag(\lambda_1, \ldots, \lambda_m),$$

$$\hat{M}_s = \hat{A}_s \otimes I_m - h\hat{B}_s \otimes \left(VDV^{-1}\right) = (I_s \otimes V)(\hat{A}_s \otimes I_m - h\hat{B}_s \otimes D)(I_s \otimes V^{-1}),$$

where I_s is the $(s+1) \times (s+1)$ identity matrix. Thus,

$$\kappa_2(M_s) \leq \kappa_2(V)\kappa_2\left(\hat{A}_s \otimes I_m - h\hat{B}_s \otimes D\right) = \kappa_2(V) \max_{1\leq j\leq m} \left\{\kappa_2(\hat{A}_s - h\lambda_j\hat{B}_s)\right\},$$

where $\kappa_2(V)$ does not depend on s.

The condition number of the matrix $\hat{M}_s$ as (2.6) is quite important for several reasons. As an example, under suitable conditions, a bound can be given for the global error of the methods based on LMF in term of $\kappa(M_s)$; see Section 4.9.1 in [11]. For our purposes, the condition number will serve to give an estimate of the rate of convergence of the preconditioned iterations of the conjugate gradient method; see Section 4.1 in the sequel. In the following discussion, we made the following assumption:

Assumption (A)*: The spectral condition number of $\hat{M}_s$ is of the same order as the spectral condition number of M_s, i.e., there exist two positive numbers c_1 and c_2 such that*

$$c_1\kappa(M_s) \leq \kappa(\hat{M}_s) \leq c_2\kappa(M_s).$$

We stress that properly chosen additional methods should obey the above assumption to avoid an ill-conditioned linear system [11, Section 5.3]. In Sections 4.9.1, 5.4.1, and 11.4 of [11], Brugnano and Trigiante have used this assumption to analyze the stability of boundary value methods. More examples can be found in Tables 3.3 and 3.4 for the wave and the heat equations, respectively, in the next subsection.

3.1 More examples.

It is interesting to observe that in the nonscalar case (i.e., when $J \in \mathbb{R}^{m\times m}$, $m > 1$), we have that $\kappa(M_s)$ ($\kappa(\hat{M}_s)$) can be independent of s or can grow at most as $O(s)$ (cf. Tables 3.1 and 3.2), at least if h is small enough, accordingly to Theorem 3.1 stated above.

Let us consider the wave equation of first order with periodic boundary conditions:

$$\begin{cases} u_t - u_x = 0, & \\ u(x,0) = x, & x \in [0,\pi] \\ u(\pi,t) = u(0,t), & t \in [0,2\pi] \end{cases} \tag{3.4}$$

We discretize the partial derivative $\partial/\partial x$ with the central differences and step size $\Delta x = \pi/m$, $x_j = j\Delta x$. We obtain a family of systems of ODEs

$$\begin{cases} y'(t) = L_m y(t), & t \in [0,2\pi] \\ y(0) = \eta, & \eta = (g(x_0) \ \dots \ g(x_{m-1}))^T, \end{cases} \tag{3.5}$$

$$L_m = \frac{1}{2\Delta x} \begin{pmatrix} 0 & -1 & & & -1 \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 1 & & & 1 & 0 \end{pmatrix}. \tag{3.6}$$

The matrices L_m have the spectrum of eigenvalues on the imaginary axis. In Table 3.3, we can see that the condition numbers of the matrices are related to the underlying method. In this example, we find that the condition numbers of M_s ($\hat{M}_s$) grow at most as $O(\log s)$ as s increases for each fixed m.

Let us consider the heat equation

$$\begin{cases} u_t - cu_{xx} = 0, & \\ u(0,t) = u(x_{max},t) = 0, & t \in (0,T] \\ u(x,0) = x(\pi - x), & x \in [0,x_{max}]. \end{cases} \tag{3.7}$$

If we discretize the operator $\partial^2/\partial x^2$ in (3.7) with centered differences and stepsize $\delta x = x_{max}/(m+1)$, we obtain a system of m ODEs

$$\begin{cases} y'(t) = T_m y(t), & t \in [0,T] \\ y(0) = \eta, & \eta = (g(x_1) \ \dots \ g(x_m))^T \end{cases} \tag{3.8}$$

Table 3.3: Wave equation. Condition numbers for different sizes of the underlying Toeplitz matrices and their small rank perturbations counterparts. Formula (2.9) with $k = 3$ is used here.

m	s	$\kappa(A_s)$	$\kappa(\hat{A}_s)$	$\kappa(B_s)$	$\kappa(\hat{B}_s)$	$\kappa(M_s)$	$\kappa(\hat{M}_s)$
10	8	10.7	10.7	3.2	9.3	18.3	19.3
10	16	21.0	21.0	16.0	20.0	26.4	26.6
10	32	41.3	41.0	31.3	35.4	44.8	45.0
20	8	10.7	10.7	3.2	9.3	29.7	31.3
20	16	21.0	21.0	16.0	20.0	36.2	36.6
20	32	41.3	41.0	31.3	35.4	52.1	52.3

and the Jacobian matrix T_m is symmetric tridiagonal:

$$T_m = \frac{c}{(\Delta x)^2} \begin{pmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & & \ddots & 1 \\ & & 1 & -2 \end{pmatrix}. \tag{3.9}$$

Table 3.4 shows the condition numbers of the matrices related to the underlying method. In this example, we find that the condition numbers of M_s ($\hat{M}_s$) are independent of s for each fixed m.

Table 3.4: Heat equation. Condition numbers for different sizes of the underlying Toeplitz matrices and their small rank perturbations counterparts. Formula (2.9) with $k = 4$ is used here.

m	s	$\kappa(A_s)$	$\kappa(\hat{A}_s)$	$\kappa(B_s)$	$\kappa(\hat{B}_s)$	$\kappa(M_s)$	$\kappa(\hat{M}_s)$
10	8	10.6	10.7	3.2	9.3	44.6	49.0
10	16	21.0	21.0	3.7	9.5	44.5	48.0
10	32	41.3	41.0	3.9	9.5	44.1	48.0
20	8	10.6	10.7	3.2	9.3	164.2	181.0
20	16	21.0	21.0	3.7	9.5	163.8	179.0
20	32	41.3	41.0	3.9	9.5	163.6	178.0

3.2 Condition numbers of M_s.

In our simplifying assumptions, we will consider the Toeplitz matrices M_s, A_s, B_s, where M_s is given by (3.1). Let us associate to the matrix M_s the two following functions: the *symbol*

$$g_{M_s}(z, q) = \sum_{j=-\nu}^{k-\nu} (\alpha_{j+\nu} - q\beta_{j+\nu})\, z^j,\ z \in \mathbb{C}, \tag{3.10}$$

and the *characteristic polynomial* of M_s, i.e.,

$$p(z,q) = z^{\nu} g_{M_s}(z,q),\ z \in \mathbb{C}, \tag{3.11}$$

where $q = h\lambda \in \mathbb{C}$. In 1963, Dahlquist introduced the definition of A-stability for an initial value method: a method is said to be *A-stable* if

$$D_{k,0} \supseteq \{q \in \mathbb{C}\ :\ \mathrm{Re}(q) < 0\},$$

where, for linear multistep formulas, $D_{k,0} \equiv D_k$ is the region of $\mathbb{C}$ where $p(z,q)$ has roots inside the unit circle (see, e.g., [20, pp. 70, 224]).

The methods (2.1) (i.e., the Toeplitz matrices M_s) we are interested in are $\mathrm{A}_{\nu,k-\nu}$-stable. A BVM is called *$A_{\nu,k-\nu}$-stable* (see [11]) if the left half complex plane is contained in the set

$$D_{\nu,k-\nu} = \{q \in \mathbb{C} : p(z,q) \text{ has } \nu,\ k-\nu \text{ roots in and out } |z| = 1,\ \text{respectively}\}.$$

The $\mathrm{A}_{\nu,k-\nu}$-stability is thus a generalization of the notion of the A-stability for the initial value linear multistep formulas, i.e. for $k > \nu$. Notice that, for us, $\mathrm{A}_{\nu,k-\nu}$ stability means also that the roots of the characteristic polynomial of M_s cannot move into, on or outside the unit circle $|z| = 1$ as $q = h\lambda$ varies in the complex left half plane. Conversely, we can have that at most one root of $p(z)$ can leave the interior of the unit circle for the exterior, in a certain region of the positive half plane; see [11, Chapter 4] for details.

We stress that the methods in Section 2.1 have the above properties and that the condition number of M_s (of $\hat{M}_s$) depends on q; see [3, 11].

In the style of the proofs in [3, 11], we can give a bound for the spectral condition number of M_s. To this end, we consider a sequence of initial value problems in (1.2) with s being increased or h being decreased. Correspondingly, we have two sequences of Toeplitz matrices $\{A_s\}_{s=1}^{\infty}$ and $\{B_s\}_{s=1}^{\infty}$, where the (j,l)-th entries of A_s and B_s are given by

$$[A_s]_{j,l} = \begin{cases} \alpha_{\nu+l-j}, & -\nu \le l-j \le k-\nu, \\ 0, & \text{otherwise.} \end{cases}$$

and

$$[B_s]_{j,l} = \begin{cases} \beta_{\nu+l-j}, & -\nu \le l-j \le k-\nu, \\ 0, & \text{otherwise,} \end{cases}$$

respectively. In this case, as we increase s, the influence of the matrix B_s (i.e., of the matrix $\hat{B}_s$) in M_s (in $\hat{M}_s$ as in (2.6)) decreases up to disappear, e.g., if J is scalar, we have

$$\lim_{s\to\infty} M_s = A_s, \quad \lim_{s\to\infty} \hat{M}_s = \hat{A}_s. \tag{3.12}$$

Below we will show that, for a given differential problem, the condition number of M_s is at most $O(s)$.

THEOREM 3.1. *Consider a family of nonsingular Toeplitz matrices $\{M_s\}$ defined as* (3.1)*, where $h = T/s$, $T > 0$, an A-stable ($A_{\nu,k-\nu}$-stable) formula* (2.1)

whose boundary locus is a regular Jordan curve. If the Jacobian matrix J has eigenvalues whose real part is non-positive, i.e., $\lambda(J) \in \mathbb{C}^-$, there exist a_1, a_2 and a_3 (independent of s) such that

$$\|M_s\|_2 \leq a_1, \quad \|M_s^{-1}\|_2 \leq a_2(s+1), \tag{3.13}$$

and, consequently,

$$\kappa(M_s) = \|M_s^{-1}\|_2 \|M_s\|_2 \leq a_3(s+1), \tag{3.14}$$

Proof. We note that

$$\|M_s\|_\infty \leq m\|A_s\|_\infty + hm\|J\|_\infty \|B_s\|_\infty \leq m \sum_{i=-\nu}^{k-\nu} |\alpha_{i+\nu}| + hm\|J\|_\infty \sum_{i=-\nu}^{k-\nu} |\beta_{i+\nu}|.$$

Since m and J are fixed and independent of s, $\|M_s\|_\infty$ is uniformly bounded from above by a positive constant independent of s. By using the similar argument, the result also holds for $\|M_s\|_1$. By using the inequality,

$$\|M_s\|_2 \leq \sqrt{\|M_s\|_\infty \|M_s\|_1} \leq a_1,$$

we obtain the result.

Without loss of generality, we assume now that $J = \lambda \in \mathbb{C}^-$. The underlying banded Toeplitz matrix $M_s = A_s - qB_s$, where $q = h\lambda$, has the characteristic polynomial $p(z, q)$ defined as (3.11). From the hypotheses, the formula (2.1) is A-stable ($A_{\nu,k-\nu}$-stable). Then, the number of the zeros of the polynomial in, on and outside $|z| = 1$, respectively, cannot change while varying $h > 0$ (or s). More precisely, if $\lambda \neq 0$, we have that $q = h\lambda \in D_{\nu,k-\nu}$ and then $p(z, q)$ has ν, $k - \nu$ roots in and out the unit circle, respectively. Otherwise, if λ is zero, $p(z, q)$ has $\nu - 1$, 1 and $k - \nu$ roots in, on and outside $|z| = 1$, respectively. Indeed, the boundary locus of the polynomial is a regular Jordan curve and one of the roots inside $|z| = 1$ reaches the unit circle as a simple root as q approaches zero for the consistency conditions (2.7).

Using Theorem 4 in [3], we have that the following bound holds true for the inverse of M_s:

$$|M_s^{-1}| \leq \alpha \left(I + C + \Delta^T\right), \tag{3.15}$$

where C and Δ are $(s+1) \times (s+1)$ matrices given by

$$C = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \vdots & \ddots & \\ 1 & \cdots & 1 & 0 \end{pmatrix}, \ \Delta = \begin{pmatrix} 0 & & & \\ \gamma & 0 & & \\ \vdots & \vdots & \ddots & \\ \gamma^s & \cdots & \gamma & 0 \end{pmatrix}, \tag{3.16}$$

respectively. Here $\alpha > 0$ and $0 < \gamma < 1$ are independent of s. The absolute value symbol $|\cdot|$ in the left hand of (3.15), i.e. used for a matrix, should be intended componentwise. At this point, it is straightforward to check that

$$\|M_s^{-1}\|_\infty \leq \alpha_1(s+1), \quad \|M_s^{-1}\|_1 \leq \alpha_2(s+1), \tag{3.17}$$

where α_1, α_2 are positive constants independent of s. Hence, if $a_2 = \max\{\alpha_1, \alpha_2\}$ and $a_3 = a_1\, a_2$, the result follows. □

The above claims are confirmed also by the results in Tables 3.1 and 3.2. For the first table, $q = h\lambda \in D_{\nu,k-\nu}$, while for the other, $q \in \mathbb{C} \setminus D_{\nu,k-\nu}$. Notice that, for each member of the family of formulas as (2.9) with k even, $D_{\nu,k-\nu}$ *includes* $\mathbb{C}^-$, while each member of the family of formulas as (2.9) and as (2.10) with k odd, $D_{\nu,k-\nu}$ *coincides* with $\mathbb{C}^-$.

Theorem 3.1 can be proved by using the more general approach of [9, 10]. The latest approach is based on the analysis of the generating function of M_s, i.e. (3.10), of the underlying Toeplitz matrix. To this end, the proof may use the results on the generating functions of the matrices A_s and B_s that can be found in [6]. However, the arguments used in the proof of Theorem 3.1 and in [3] can be useful to focus on the assumption (A), i.e. on the condition numbers of the matrices $\hat{A}_s$, $\hat{B}_s$, as claimed in the sequel.

4 Block circulant preconditioners.

Let $\hat{M}_s$ in (2.6) be the low rank perturbation of a block-Toeplitz matrix generated by a BVM as in Section 2. A way to solve such large and sparse linear systems is to use an iterative method; see for instance [16, 12]. To speed up the convergence of the iterative method, a preconditioner P_s should be chosen to approximate the matrix M_s while keeping the system $P_s\mathbf{x} = \mathbf{c}$ cheap to solve.

In order to obtain the preconditioner, let us consider the following approximation of the matrix $\hat{M}_s$:

$$P_s = \breve{A}_s \otimes I_m - h\breve{B}_s \otimes \breve{J}_m \tag{4.1}$$

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

$$\begin{aligned} &\breve{A}_s = \mathrm{circ}\,(\tilde{\alpha}_j), \quad \tilde{\alpha}_j = c_{j,1}(s)\alpha_{j+\nu} + c_{j,2}(s)\alpha_{j+\nu-(s+1)}, \\ &\breve{B}_s = \mathrm{circ}\,(\tilde{\beta}_j), \quad \tilde{\beta}_j = c_{j,3}(s)\beta_{j+\nu} + c_{j,4}(s)\beta_{j+\nu-(s+1)}, \quad j = 0, \ldots, s. \end{aligned} \tag{4.2}$$

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 (4.2). The coefficients $c_{i,j}(s)$ in (4.2) are chosen in such a way that $\breve{A}_s$, $\breve{B}_s$ are suitable approximations of A_s, B_s in (2.5), respectively.

The approximation of A, B with T. Chan's *optimal circulant* (see [13]) requires that

$$\begin{aligned} c_{j,1}(s) = c_{j,3}(s) &= 1 - \frac{j}{s+1}, \\ c_{j,2}(s) = c_{j,4}(s) &= \frac{j}{s+1}, \quad j = 0, \ldots, s, \end{aligned} \tag{4.3}$$

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

$$c_{j,1}(s) = c_{j,3}(s) = 1, \quad j = 0, \ldots, \left\lfloor \frac{s+1}{2} \right\rfloor,$$
$$c_{j,2}(s) = c_{j,4}(s) = 1, \quad j = \left\lfloor \frac{s+1}{2} \right\rfloor + 1, \ldots, s, \quad c_{r,j}(s) = 0 \text{ otherwise.}$$

Consider, instead of (4.3), the following definition of the coefficients $c_{j,i}(s)$:

$$c_{j,1}(s) = c_{j,3}(s) = 1 + \frac{j}{s+1}, \quad c_{j,2}(s) = c_{j,4}(s) = \frac{j}{s+1}, \quad j = 0, \ldots, s. \tag{4.4}$$

We will call *P-circulants* the circulant matrices defined in (4.2) with (4.4). We remark that the T. Chan preconditioners may be singular or very ill-conditioned for some BVMs; see [4]. Bertaccini [4] and Chan et al. [15] showed that both the P-circulant and generalized Strang preconditioned systems can converge very quickly. However, when the Jacobian matrix J has some small (or zero) eigenvalues (for instance in some ODEs; see, e.g. [4, 7, 8]), the Strang preconditioner can be severely ill-conditioned or even singular. Therefore, we will focus on P-circulant preconditioners in the following discussion.

4.1 Convergence rate of the preconditioned iterations.

In this subsection, we analyze the convergence rate of the conjugate gradient method when applied to solve the preconditioned normal system

$$(P_s^{-1}\hat{M}_s)^*(P_s^{-1}\hat{M}_s)\mathbf{x} = (P_s^{-1}\hat{M}_s)^* P_s^{-1}\mathbf{b}. \tag{4.5}$$

We show that the method converges in at most $O(\log s)$ steps where the spectral condition number $\hat{M}_s$ under Assumption **(A)** is $O(s)$. We begin by the following theorem (see [5] for the proof).

THEOREM 4.1. *Let $\hat{M}_s$ be the matrix of the linear system* (2.6) *for the schemes described in Section* 2.1 *using an $A_{\nu,k-\nu}$-stable formula, and P_s its block P-circulant preconditioner as* (4.1). *Then, P_s is invertible for all $s \geq k$ and there exists a positive number c' independent of s such that $\|P_s^{-1}\|_2 < c'$ for all s. Moreover, for any given $\delta > 0$, there exists a positive number r_0 (independent of s) and $s_0 \geq \mu$ such that, for all $s \geq s_0$, we have*

$$P_s^{-1}\hat{M}_s = I + L_{r_0} + V_\delta \tag{4.6}$$

where $\operatorname{rank}(L_{r_0}) \leq r_0$ *and* $\|V_\delta\|_2 \leq \delta$.

Next we derive a lower bound for the singular values of $P_s^{-1}\hat{M}_s$.

LEMMA 4.2. *Let Assumption* **(A)** *be satisfied. There exists a constant $\tilde{c} > 0$ independent of s such that*

$$\|\hat{M}_s\|_2 > \tilde{c}. \tag{4.7}$$

Hence we have

$$\|\hat{M}_s^{-1}P_s\|_2 \leq \frac{\|P_s\|_2}{\|\hat{M}_s\|_2}\kappa(\hat{M}_s) \leq cs \tag{4.8}$$

for some constant $c > 0$.

PROOF. We note from (1.3), (3.1) and (3.2) that

$$\hat{M}_s - M_s = L_k$$

where $\text{rank}(L_k) \le k$. It follows that

$$\hat{M}_s^* \hat{M}_s - M_s^* M_s = (\hat{M}_s - M_s)^*(\hat{M}_s - M_s) + M_s^*(\hat{M}_s - M_s) + (\hat{M}_s - M_s)^* M_s = L_{3k},$$

where $\text{rank}(L_{3k}) \le 3k$. By using Weyl's theorem [19, p. 184], we have

$$\lambda_{\max}(\hat{M}_s^* \hat{M}_s) \ge \lambda_{(s+1)m-3k}(M_s^* M_s).$$

Since the eigenvalues of $M_s^* M_s$ are equal to the square of the singular values of M_s, and the singular values of M_s are distributed as $|p(e^{\mathbf{i}\theta}, q)|$ for $\theta \in [0, 2\pi)$ (see, e.g., [24]), for sufficiently large s, we establish the inequality (4.7). Indeed, we can reduce the analysis to the scalar case (see Section 3) and thus M_s is a Toeplitz matrix whose related polynomial and symbol are $p(z, q)$ and $z^{-\nu} p(z, q)$, respectively. By Theorem 4.1, $\|P_s\|_2$ is uniformly bounded above by a constant independent of s. By Assumption (**A**) and using Theorem 3.1, we establish the inequality (4.8). □

As for the convergence rate of the preconditioned conjugate gradient method for the normal equations, we will show that the method will converge in at most $O(\log s)$ steps when the spectral condition number of $\hat{M}_s$ is of the order of $O(s)$. We begin by noting the following error estimate of the conjugate gradient method (see [1]).

THEOREM 4.3. *Let* $\mathbf{x}$ *be the solution of the system* (4.5) *and* $\mathbf{x}^{(j)}$ *be the* j*-th iterant of the ordinary conjugate gradient method applied to* (4.5). *If the eigenvalues* $\{\lambda_k\}$ *of* $(P_s^{-1} M_s)^*(P_s^{-1} M_s)$ *are such that*

$$0 < \lambda_1 \le \cdots \le \lambda_p \le b_1 \le \lambda_{p+1} \le \cdots \le \lambda_{(s+1)m-q} \le b_2 \le \lambda_{(s+1)m-q+1} \le \cdots \le \lambda_{(s+1)m},$$

then

$$\frac{||(P_s^{-1} M_s)(\mathbf{x} - \mathbf{x}^{(j)})||_2}{||(P_s^{-1} M_s)(\mathbf{x} - \mathbf{x}^{(0)})||_2} \le 2 \left(\frac{b-1}{b+1} \right)^{j-p-q} \cdot \max_{\lambda \in [b_1, b_2]} \left\{ \prod_{k=1}^{p} \left(\frac{\lambda - \lambda_k}{\lambda_k} \right) \right\}. \tag{4.9}$$

Here $b := (b_2/b_1)^{\frac{1}{2}} \ge 1$.

For the preconditioned system (4.5), the iteration matrix is given by

$$(P_s^{-1} M_s)^*(P_s^{-1} M_s).$$

Theorem 4.1 implies that we can choose $b_1 = 1 - \delta$ and $b_2 = 1 + \delta$, and therefore $(b-1)/(b+1) < \delta$. Then, p and q are constants that are independent of s. In order to use (4.9), we need a lower bound for λ_k, $1 \le k \le p$. It follows by Lemma 4.2 that

$$\lambda_k \ge \min_{\ell} \lambda_\ell = \frac{1}{\|(P_s^{-1} M_s)^*(P_s^{-1} M_s)\|_2^2} \ge \frac{1}{c^2 s^2},$$

for $1 \leq k \leq n$, where c is a positive constant. Thus, for $1 \leq k \leq p$ and $\lambda \in [1-\delta, 1+\delta]$, we have that

$$0 \leq \frac{\lambda - \lambda_k}{\lambda_k} \leq c^2 s^2.$$

Hence, (4.9) becomes

$$\frac{||(P_s^{-1}M_s)(\mathbf{x}-\mathbf{x}^{(j)})||_2}{||(P_s^{-1}M_s)(\mathbf{x}-\mathbf{x}^{(0)})||_2} < c^{2p}s^{2p}\delta^{j-p-q}.$$

Given arbitrary tolerance $\epsilon > 0$, an upper bound for the number of iterations required to make

$$\frac{||P_s^{-1}M_s(\mathbf{y}-\mathbf{y}_j)||_2}{||P_s^{-1}M_s(\mathbf{y}-\mathbf{y}_0)||_2} < \epsilon$$

is therefore given by

$$j_0 := p + q - \frac{2p\log c + 2p\log s - \log\epsilon}{\log\delta} = O(\log s).$$

Since each preconditioned conjugate gradient iteration requires $O(s\log s)$ operations, the total complexity of the preconditioned conjugate gradient method is at most $O(s\log^2 s)$. We note that when the condition number of $\hat{M}_s$ is of the order of $O(1)$, the number of iterations required for convergence will also be constant and hence the total number operations for solving (1.1) will be of the order of $O(s\log s)$.

5 Numerical results.

To show the effectiveness of our analysis, we will integrate some test problems already considered in Section 3. Further numerical examples using other formulas can be found in [4, 5, 7, 8, 15].

The initial guess for the iterative solvers is always zero. The stopping criterion is $\|\mathbf{r}_k\| < 10^{-6}\|\mathbf{b}\|$, where $\mathbf{r}_k$ is the residual $\mathbf{b} - \hat{M}_s\mathbf{x}$. All calculations are done in Matlab.

We compare three different iterative algorithms: the conjugate gradient method for the normal system (CGN) [18], the BiCGStab [25] and GMRES [22] methods. More precisely, we apply the BICGStab and (unrestarted) GMRES methods to solving the linear system

$$P_s^{-1}\hat{M}_s\mathbf{x} = P_s^{-1}\mathbf{b} \tag{5.1}$$

and conjugate gradient method to solving the system of normal equations

$$(P_s^{-1}\hat{M}_s)^*(P_s^{-1}\hat{M}_s)\mathbf{x} = (P_s^{-1}\hat{M}_s)^*P_s^{-1}\mathbf{b}. \tag{5.2}$$

Here P_s is the identity matrix when we consider unpreconditioned iterations (the columns "I" in the tables), otherwise P_s is the block circulant preconditioner (4.1). In the tests, we use P-circulant approximations considered in Section 4 (the

columns "P" in the tables) and also the skew circulant approximations introduced in [8] (the column "S" in the tables). The latter one has been found to be more competitive with respect to Strang's and T. Chan's circulant preconditioners, that can be severely ill-conditioned or singular even when the matrix $\hat{M}_s$ (2.6) is well-conditioned (see [5, 7, 6, 8]). A "$-$" means that convergence was not attained.

We stress that each iteration of the BiCGStab and conjugate gradient method requires two matrix vector products, while one is required for the GMRES method (but the latter has a computational cost that increases with the number of iterations until restarting; see [22] for details). The computational cost of each preconditioned iteration is of $O(m\,s\,\log s) + O(m\,s)$ flops, where $m = 1$ for the scalar test problems, while m is the size of the sparse Jacobian matrices (3.6), (3.9) for the others. For more details on the computational cost and the implementation; see [5].

Table 5.1: Number of iterations, formula (2.9) with $k = 4$ for problem (3.3) with $\lambda = -1$ and $T = 1$.

	CGN			BiCGStab		GMRES	
s	I	P	S	I	P	I	P
32	32	6	4	28	5	32	7
64	64	6	4	54	5	64	7
128	128	7	4	101	4	128	7
256	256	7	4	188	4	256	7
512	512	7	4	$-$	4	512	7

Table 5.2: Number of iterations, formula (2.10), $k = 3$ for problem (3.3) with $\lambda = +1$ and $T = 1$.

	CGN			BiCGStab		GMRES	
s	I	P	S	I	P	I	P
32	32	8	4	46	7	32	7
64	64	8	4	$-$	6	64	7
128	128	9	4	$-$	6	123	7
256	256	9	4	$-$	6	235	7
512	512	10	4	$-$	6	512	7

Notice that the Strang-type circulant preconditioner is singular for all odd m for the wave equation problem. We observe that the convergence of the conjugate gradient method depends on the eigenvalues of the preconditioned normal system (5.2). These eigenvalues are the singular values squared. More precisely, if the condition number $\kappa(M_s)$ ($\kappa(\hat{M}_s)$) does not depend on s, the same is true for the number of preconditioned iterations (cf. the heat equation example in Table 5.4). On the contrary, if $\kappa(M_s)$ ($\kappa(\hat{M}_s)$) is (at most) of $O(s)$, the number of the

Table 5.3: Number of iterations for the wave equation problem. Formula (2.9) with $k = 3$ is used here.

		CGN			BiCGStab		GMRES	
m	s	I	P	S	I	P	I	P
10	8	25	14	15	79	11	45	16
10	16	51	15	14	88	12	44	16
10	32	79	15	15	96	12	55	16
20	8	47	20	21	–	16	86	22
20	16	98	24	21	–	16	88	21
20	32	135	24	25	–	12	77	19

Table 5.4: Number of iterations for the heat equation problem. Formula (2.9) with $k = 4$ is used here.

		CGN			BiCGStab		GMRES	
m	s	I	P	S	I	P	I	P
10	8	128	24	26	51	7	63	9
10	16	171	23	23	63	7	79	8
10	32	186	20	20	70	8	89	8
20	8	439	28	29	130	7	130	8
20	16	613	31	37	162	6	173	8
20	32	679	29	31	184	6	207	8

preconditioned iterations may increase as $O(\log s)$ (cf. the examples in Tables 5.1, 5.2 and 5.3).

The convergence of the BiCGStab and GMRES methods depends on the (pseudo) eigenvalues of the (preconditioned) matrix of the system in (5.1); see for instance, [21]. Thus, the analysis of Section 4 cannot be applied to the BiCGStab and GMRES methods. Nonetheless, it can be observed that usually their rates of convergence behavior are similar to the conjugate gradient method. We can also see from the tables that the number of the GMRES and BiCGStab iterations required for convergence stays constant for increasing s when m is fixed. A detailed analysis of the convergence behavior of the GMRES iterations for the underlying class of block circulant preconditioners will be shown in a forthcoming paper.

In summary, we have found that Krylov subspace methods in combination with the block circulant preconditioners can be efficient methods for solving ODEs and PDEs with BVMs; see [4, 5, 8, 15]. By studying the spectral condition number of $\hat{M}_s$, we have shown in this paper that the conjugate gradient method, when applied to solving the normalized preconditioned system, converges in at most $O(\log s)$ steps.

Acknowledgments.

The first author would like to thank The University of Hong Kong for the support and hospitality offered in January–February 2000, when part of this work was completed. The authors wish to thank an anonymous referee who provide some useful comments.

REFERENCES

1. O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation*, Academic Press, Orlando, FL, 1984.
2. A. O. H. Axelsson and J. G. Verwer, *Boundary value techniques for initial value problems in ordinary differential equations*, Math. Comp., 45 (1985), pp. 153–171.
3. P. Amodio and L. Brugnano, *The conditioning of Toeplitz band matrices*, Math. Comput. Model., 23:10 (1996), pp. 29–42.
4. D. Bertaccini, *P-circulant preconditioners and the systems of ODE codes*, in Iterative Methods in Scientific Computation IV, IMACS Series in Computational and Applied Mathematics, D. R. Kincaid et al., eds., IMACS, NJ, 1999, pp. 179–193.
5. D. Bertaccini, *A Circulant Preconditioner for the systems of LMF-based ODE codes*, SIAM J. Sci. Comput., 22:3 (2000), pp. 767–786.
6. D. Bertaccini, *An analysis of circulant preconditioners for linear multistep integrators*, submitted, available at URL http://www.math.unifi.it/~bertaccini.
7. D. Bertaccini, *Reliable preconditioned iterative linear solvers for some numerical integrators*, Numer. Linear Algebra Appl., 8:2 (2000), pp. 111–125.
8. D. Bertaccini and M. K. Ng, *Skew-circulant preconditioners for systems of LMF-based ODE codes*, in Lecture Notes in Comp. Sci., L. Vulkov et al., eds., Springer-Verlag, New York, 2001, pp. 93–101.
9. A. Böttcher and S. Grudsky, *On the condition numbers of large semi-definite Toeplitz matrices*, Linear Algebra Appl., 279 (1998), pp. 285–301.
10. A. Böttcher and B. Silbermann, *Introduction to large truncated Toeplitz matrices*, Springer-Verlag, New York, 1998.
11. L. Brugnano and D. Trigiante, *Solving ODE by Linear Multistep Methods: Initial and Boundary Value Methods*, Gordon and Breach, Reading, UK, 1998.
12. 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. Stat. Comput., 7 (1986), pp. 378–417.
13. T. F. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 766–771.
14. R. H. Chan and M. K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Review, 38:3 (1996), pp. 427–482.
15. R. H. Chan, M. K. Ng, and X. Jin, *Circulant preconditioners for solving ordinary differential equations*, in Structured matrices, D. Bini et al., eds., Nova Science Pub., to appear.
16. C. W. Gear and Y. Saad, *Iterative solution of linear equations in ODE codes*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 583–601.
17. E. Hairer and G. Wanner, *Solving Ordinary Differentials Equations II. Stiff and Differential Algebraic Problems*, Springer-Verlag, Berlin, 1991.

18. M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand., 49 (1952), pp. 409–436.
19. R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
20. J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, Wiley, Chichester, 1991.
21. N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matr. Anal. Appl., 13 (1992), pp. 778–795.
22. Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.
23. G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
24. E. E. Tyrtyshnikov, *A unifying approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl., 232 (1996), 1–43.
25. H. A. van der Vorst, *BICGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 10 (1992), pp. 631–644.