SIAM J. NUMER. ANAL.
Vol. 40, No. 5, pp. 1798–1822
© 2002 Society for Industrial and Applied Mathematics

THE SPECTRUM OF CIRCULANT-LIKE PRECONDITIONERS FOR SOME GENERAL LINEAR MULTISTEP FORMULAS FOR LINEAR BOUNDARY VALUE PROBLEMS*

DANIELE BERTACCINI†

Abstract. The spectrum of the eigenvalues, the conditioning, and other related properties of circulant-like matrices used to build up block preconditioners for the nonsymmetric algebraic linear equations of time-step integrators for linear boundary value problems are analyzed. Moreover, results concerning the entries of a class of Toeplitz matrices related to the latter are proposed. Generalizations of implicit linear multistep formulas in boundary value form are considered in more detail.

It is proven that there exists a new class of approximations which are well conditioned and whose eigenvalues have positive and bounded real and bounded imaginary part. Moreover, it is observed that preconditioners based on other circulant-like approximations, which are well suited for Hermitian linear systems, can be severely ill conditioned even if the matrices of the nonpreconditioned system are well conditioned.

Key words. trigonometric preconditioners, nonsymmetric Toeplitz matrices, eigenvalues, linear systems of time-step integrators, general linear multistep formulas in boundary value form, boundary value problems

AMS subject classifications. 65F10, 65F15, 65N22, 15A18, 15A48

PII. S0036142901397447

1. Introduction.

In this paper we investigate the properties of some classes of circulant approximations and some generalizations used in the preconditioners for (small rank perturbations of) block nonsymmetric Toeplitz matrices introduced in [4]. These matrices arise in the numerical approximation of time-dependent partial differential equations by means of generalizations of implicit linear multistep formulas.

An $n \times n$ matrix $A_n = (a_{j,k})$ is said to be Toeplitz if $a_{j,k} = a_{j-k}$, $j, k = 1, \dots, n$, i.e., A_n is constant along its diagonals, quasi Toeplitz if it is a small rank perturbation of a Toeplitz matrix. An $n \times n$ matrix $\breve{A}_n$ is said to be circulant if it is Toeplitz and its diagonals satisfy $\breve{a}_{n-j} = \breve{a}_{-j}$, $j = 1, \dots, n-1$. The circulant matrices $\breve{A}_n$ are diagonalized by the Fourier matrix $F = (F_{j,k})$, $F_{j,k} = e^{2\pi \mathbf{i} jk/n}/\sqrt{n}$, $j, k = 0, \dots, n-1$, $\mathbf{i}$ is the imaginary unit; see [13]. From the previous arguments, it follows that such matrices are easily and efficiently invertible using the fast Fourier transform (FFT); see, e.g., [11]. Other circulant-like matrices will be mentioned in section 4.

The matrices of the underlying linear systems can be written as follows:

$$M = A \otimes I - h\, B \otimes J, \tag{1.1}$$

where A and B are $n \times n$ (small rank perturbations of) band Toeplitz matrices whose entries are given by the coefficients of the scheme involved, I is the identity, and J is an $m \times m$ matrix which can be large and sparse. More precisely, J is the Jacobian matrix of a system of ordinary or partial differential equations discretized in space by

*Received by the editors November 5, 2001; accepted for publication (in revised form) June 9, 2002; published electronically November 22, 2002. This work was partially supported by an INdAM-GNCS project and by a grant from the MURST project "Progetto giovani ricercatori anno 2000."
http://www.siam.org/journals/sinum/40-5/39744.html
†Università di Roma "La Sapienza," Dipartimento di Matematica, P.le A. Moro 2, 00185 Roma, Italy (bertaccini@mat.uniroma1.it).

finite differences; see [4] for details. It is worth noting that J can have a (multilevel) structure as well. For example, J can be block-banded, block-Toeplitz, etc.

Unfortunately, when m and/or n are (even moderately) large, iterative solvers for (1.1), used without preconditioners or with general purpose preconditioners such as those based on incomplete factorizations, often converge very slowly or do not converge at all; see [4, section 5]. Moreover, direct methods are not appropriate because they cannot exploit the block structure of (1.1). On the other hand, the preconditioners we consider here take into account the block structure in (1.1). More precisely, they are block-circulant and, in matrix form, can be written as

$$P = \breve{A} \otimes I - h\,\breve{B} \otimes \tilde{J}, \tag{1.2}$$

where $\breve{A}$ and $\breve{B}$ are circulant-like approximations for A, B, respectively, and $\tilde{J}$ is a suitable approximation for J.

In [4] we have observed that $P^{-1}M$, the preconditioned matrix, can be written as a perturbation of the identity matrix (see also section 5.3), which can result in fast convergence of Krylov subspace methods for nonsymmetric linear systems such as GMRES and BiCG-like methods such as BiCGstab. The computational cost for a possible implementation has been considered in detail in [4, section 4.1], showing that the cost per iteration is of the order of $O(mn \log n)$ if J is banded, say.

Here we will prove that there exists a class of circulant approximations, introduced in [4], which have a moderate 2-norm condition number increasing at most linearly with their size n. Moreover, the spectrum of the eigenvalues of several of the possible approximations for the nonsymmetric matrices A, B in (1.1) will be investigated as well, showing that it lies in the right half plane. It is worth noting that this holds true for the original matrices A, B in (1.1) considered here; see [6].

We stress that the condition number and the spectrum of the component matrices $\breve{A}$ and $\breve{B}$ of the preconditioner (1.2) are very important to have fast convergence. Indeed, as observed in [4, 5, 6], the matrix J in (1.1) can have very small (and/or very large) singular values in different subintervals of integration (see [5]), and this is difficult (if not impossible) to know in advance. Recall that J is the Jacobian matrix of the given continuous time-dependent problem; see [15, 19]. Thus, if $\breve{A}$ or $\breve{B}$ are ill conditioned, we can have an ill conditioned preconditioner even if the original matrix M is well conditioned; see, e.g., the end of section 5.2 and Figure 5.3. As observed in [4, 5], this can slow down the convergence process (see [14]), giving unacceptably slowly convergent (or even divergent!) preconditioned iterations.

We observe that, in the case of nonsymmetric linear systems, solved by Krylov accelerators, the condition number of the *preconditioned* matrix $P^{-1}M$ (say), assumed to be not too large, is much less important for the convergence than the clustering of the spectrum of its eigenvalues; see, e.g., [18, 14]. On the other hand, the condition number of the matrix M in (1.1) is crucial for the rate of convergence of conjugate gradients preconditioned iterations for the normal linear system; see [6].

Here we will consider certain general linear multistep methods (or GLMs, see [19]) used in boundary value form and called boundary value methods. These methods are used to solve continuous boundary value problems for differential equations (see [2, 9] and references therein). However, the asymptotic techniques considered here could be adapted, at least in principle, to other discretization schemes.

Notice that, in this paper, we will consider multistep formulas of arbitrarily high order merely to state bounds and the asymptotic behavior of the spectrum of the underlying circulant(-like) approximations involved in (1.2). In practice, the best

performance of the underlying preconditioners seems to be achieved for formulas (2.3) whose number of steps k is not too large (typically 3 to 9, say). On the other hand, we have observed in [4, 5] that the preconditioners (1.2) can be effective for any order of magnitude of n, either if it is small (4 to 8, say) or (very) large ($n > 1024$, say) as well.

In section 2 we summarize some information on numerical integrators based on linear multistep formulas in boundary value form. Section 3 contains some introductory lemmas. In section 4 we recall some circulant approximations. Section 5 is devoted to the investigation on the spectrum and the conditioning of the underlying matrices. Finally, some remarks on the convergence of preconditioned iterations and the use of different approximations in (1.2) are given in section 5.3.

2. Families of numerical integrators.

2.1. Formulas in boundary value form. The boundary value methods for differential equations are a generalization of implicit linear multistep formulas; see [2, 9] and references therein. They approximate the solution of a continuous differential boundary value problem by means of a discrete boundary value problem. For simplicity, let us consider the linear boundary value problem

$$\begin{cases} y'(t) = f(t, y(t)) := J\, y(t) + g(t), \quad t \in (t_0, T], \\ y(t_0) = \eta_1, \quad y(T) = \eta_2, \end{cases} \tag{2.1}$$

where $y(t), g(t) : \mathbb{R} \to \mathbb{R}^m$, $J \in \mathbb{R}^{m\times m}$, $\eta_j \in \mathbb{R}^m$, $j = 1, 2$. The continuous problem (2.1) can be reduced to a discrete boundary value problem by the following k-step linear multistep formula of order p used with $\nu > 0$ initial and $k-\nu > 0$ final conditions over a uniform mesh $t_j = t_0 + j\, h,\ j = 0, \ldots, s$:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = h \sum_{i=0}^{k} \beta_i f_{n+i}, \qquad n = 0, \ldots, s-k, \tag{2.2}$$

where y_n is the discrete approximation to $y(t_n)$, $f_n = f(t_n, y_n) \equiv J\, y_n + g_n$, $g_n = g(t_n)$, while the values $y_0, \ldots, y_{\nu-1}, y_{s-k+\nu+1}, \ldots, y_s$ of the approximation computed in the mesh points $t_0, \ldots, t_{\nu-1}, t_{s-k+\nu+1}, \ldots, t_s$, respectively, are assumed to be given. We observe that the boundary value problem (2.1) provides only the initial and final values y_0 and y_s, respectively. The missing values are supplied by coupling the method (2.2) with other difference schemes of order p, sometimes called additional methods, which provide an additional set of equations, independent of those in (2.2). For simplicity, we can assume that these formulas have the same number of steps as (2.2) but different coefficients $\alpha_j^{(r)}$, $\beta_j^{(r)}$, $r = 1, \ldots, \nu-1, s-k+\nu+1, \ldots, s-1$, $j = 0, \ldots, k > \nu$; see [4] for details.

In order to stress the dependence of the formula on the ν initial and $k-\nu$ final values, it is useful to rewrite (2.2) in the following shifted form:

$$\sum_{i=-\nu}^{k-\nu} \alpha_{i+\nu} y_{n+i} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu} f_{n+i}, \qquad n = \nu, \ldots, s-k+\nu. \tag{2.3}$$

To have order $p \geq 1$, the coefficients α_j, β_j in (2.3) should satisfy the order conditions (see, e.g., [19])

$$\sum_{j=0}^{k} \left(j^i \alpha_j - i j^{i-1} \beta_j \right) = 0, \qquad i = 0, \ldots, p, \tag{2.4}$$

where the first two equations of (2.4) ($i = 0, 1$) are usually called consistency conditions. By rewriting (2.4) in shifted form we have

$$\sum_{j=-\nu}^{k-\nu} \left(j^r \alpha_{j+\nu} - r j^{r-1} \beta_{j+\nu} \right) = 0, \qquad r = 0, \dots, p, \tag{2.5}$$

where we assume, as is customary, that $0 j^{-1} = 0$, $j^0 = 1$ for all j.

For practical implementation, we cast the above discrete problem in matrix form:

$$\begin{aligned} &MY = b,\ Y = \left(y_0^T, y_1^T, \dots, y_s^T \right)^T, \qquad M = A \otimes I_m - hB \otimes J, \\ &b = e_1 \otimes \eta_1 + e_{s+1} \otimes \eta_2 + h(B \otimes I_m) g, \qquad g = (g(t_0) \cdots g(t_s))^T, \end{aligned} \tag{2.6}$$

where $e_i \in \mathbb{R}^{s+1}$, $i = 1, \dots, s+1$, is the ith column of the identity matrix and A, $B \in \mathbb{R}^{(s+1)\times(s+1)}$ are quasi-Toeplitz matrices whose pattern is

$$A = \begin{pmatrix} 1 & \cdots & 0 & & & \\ \alpha_0^{(1)} & \cdots & \alpha_k^{(1)} & & & \\ \vdots & \vdots & \vdots & & & \\ \alpha_0^{(\nu-1)} & \cdots & \alpha_k^{(\nu-1)} & & & \\ \alpha_0 & \cdots & \alpha_k & & & \\ & \alpha_0 & \cdots & \alpha_k & & \\ & & \ddots & \ddots & \ddots & \\ & & & \alpha_0 & \cdots & \alpha_k \\ & & & \alpha_0^{(s-k+\nu+1)} & \cdots & \alpha_k^{(s-k+\nu+1)} \\ & & & \vdots & \vdots & \vdots \\ & & & \alpha_0^{(s-1)} & \cdots & \alpha_k^{(s-1)} \\ & & & 0 & \cdots & 1 \end{pmatrix}, \tag{2.7}$$

where $\alpha_j^{(r)}$, $j = 0, \dots, k$, are the coefficients of the additional formulas. The matrix B is similarly defined, but with β_j ($\beta_j^{(r)}$) instead of α_j ($\alpha_j^{(r)}$) and with the entries of the first and last rows set to zero; see [4] for details. We stress that the matrix M in (2.6) is usually nonsymmetric, nondiagonally dominant, and large and sparse if, e.g., n or m are large and J is sparse.

2.2. Some linear multistep schemes. Let us recall some families of formulas (2.3) that will be considered in subsequent sections.

A minimal requirement for a multistep formula (2.3) is consistency, see, e.g., [19], i.e., they must satisfy the conditions $\rho(1) = 0$, $\rho'(1) = \sigma(1)$, where $\rho(z)$ and $\sigma(z)$ denote the two characteristic polynomials associated with the given method, i.e.,

$$\rho(z) = \sum_{j=0}^{k} \alpha_j \ z^j, \quad \sigma(z) = \sum_{j=0}^{k} \beta_j \ z^j$$

or, in shifted form,

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

The backward differentiation formulas are a class of well-known initial value methods for the numerical integration of stiff problems (see, e.g., [15, 19]). A generalization of these as a boundary value scheme, called generalized backward differentiation formulas, has been proposed in [9] and can be written in the form

$$\sum_{i=-\nu}^{k-\nu} \alpha_{i+\nu}\, y_{n+i} = h\, f_n, \qquad n = \nu, \dots, s-k+\nu, \tag{2.9}$$

where $\nu = (k+2)/2$ if k is even, and $\nu = (k+1)/2$ if k is odd; see [9]. Notice that backward differentiation formulas have $\nu = k$ in (2.9). The coefficients $\{\alpha_i\}$ are determined by imposing maximum order for (2.9), i.e., order k, $k \geq 1$.

Another popular class of initial value methods is the Adams–Moulton formulas; see, e.g., [15, 19]. Let us consider their generalization in the boundary value form, proposed in [9], called generalized Adams–Moulton methods, that can be written in the following form:

$$y_n - y_{n-1} = h \sum_{i=-\nu}^{k-\nu} \beta_{i+\nu}\, f_{n+i}, \qquad n = \nu, \dots, s-k+\nu, \tag{2.10}$$

i.e., the only nonzero coefficients in the first characteristic polynomial are $\alpha_\nu = 1$ and $\alpha_{\nu-1} = -1$, $\nu = k/2$ if k is even, and $\nu = (k+1)/2$ if k is odd. The coefficients $\{\alpha_i\}$ are determined by imposing that the method has maximum order, i.e., $k+1$. Notice that the classical Adams–Moulton methods have $\nu = k$; see, e.g., [19]. When k is odd, the scheme shares the same stability properties of the trapezoidal rule. Such methods can be suitable for approximating Hamiltonian problems and continuous boundary value problems.

Another generalization of the trapezoidal rule proposed in [9] is given by the following formula:

$$\sum_{i=-\nu}^{\nu-1} \alpha_{i+\nu}\, y_{n+i} = \frac{h}{2}(f_n + f_{n-1}), \qquad n = \nu, \dots, s-k+\nu, \tag{2.11}$$

where $\nu = (k+1)/2$ if k is odd and $\nu = k/2$ if k is even. The coefficients $\{\alpha_i\}$ are determined by imposing that the above formula has maximum order, i.e., $k+1$. Such methods can be suitable for approximating Hamiltonian problems and continuous boundary value problems.

It will be useful in the following sections to have some of the order conditions (2.4) for the above mentioned schemes written in a different form. For the formulas (2.11), we consider (2.4) with $\beta_\nu = \beta_{\nu-1} = 1/2$. Therefore, we have

$$\sum_{j=-\nu}^{\nu-1} j^r \alpha_{j+\nu} = (-1)^{r+1}\frac{r}{2}, \quad r = 0, 2, \dots, k+1, \qquad \sum_{j=-\nu}^{\nu-1} j\alpha_{j+\nu} = 1. \tag{2.12}$$

Similarly, for (2.9), $\beta_\nu = 1$, $\beta_j = 0$ for $j \neq \nu$. Thus, the α_j, $j = 0 \dots, k$, satisfy consistency conditions and

$$\sum_{j=-\nu}^{k-\nu} j^r \alpha_{j+\nu} = 0, \qquad r = 0, 2, \dots, k. \tag{2.13}$$

Finally, for (2.10), $\alpha_\nu = -\alpha_{\nu-1} = 1$ while the coefficients β_j, $j = 0 \dots, k$, satisfy

$$\sum_{j=-\nu}^{k-\nu} j^r \beta_{j+\nu} = (-1)^r \frac{1}{r+1}, \qquad r = 0, 1, \dots, k. \tag{2.14}$$

3. The entries of a class of Toeplitz matrices. Let us consider the $n \times n$ band Toeplitz matrix $\hat{A}_n = (\alpha_j)$, $n > k$,

$$\hat{A}_n = \begin{pmatrix} \alpha_\nu & \dots & \alpha_{k-1} & \alpha_k & 0 & \dots & 0 \\ \vdots & \alpha_\nu & \ddots & \alpha_{k-1} & \alpha_k & \ddots & \vdots \\ \alpha_0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \alpha_{k-1} & \alpha_k \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \alpha_{k-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \alpha_0 & \dots & \alpha_\nu \end{pmatrix}, \tag{3.1}$$

and $\hat{B}_n = (\beta_j)$ having similar pattern, but with β_j instead of α_j, $j = 0, \dots, k$. If α_j, β_j, $j = 0, \dots, k$, are the coefficients of (2.3), $E_n^{(A)} = A_n - \hat{A}_n$, $E_n^{(B)} = B_n - \hat{B}_n$ are small rank matrices if $n \gg k$, where $A \equiv A_n$ is defined in (2.7) and similarly for $B \equiv B_n$.

It can be checked that all such matrices are, in general, nonsymmetric, nondiagonally dominant, with real entries of nonconstant sign. Moreover, let us associate to the matrices $\hat{A}_n$, $\hat{B}_n$ (A_n and B_n) as above the functions $g_A(z)$, $g_B(z)$, respectively. It is customary to call $g_A(z)$ the symbol of the matrix A_n, see, e.g., [8], where

$$g_A(z) = z^{-\nu}\rho(z) = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} z^j, \qquad z \in \mathbb{C}, \tag{3.2}$$

and $\rho(z)$ is the characteristic polynomial of $\hat{A}_n$ while $g_B(z)$ is defined similarly for $\hat{B}_n$ from $\sigma(z)$ in (2.8). The set $\{q \in \mathbb{C} : q = g_A(e^{\mathbf{i}\theta}), 0 \le \theta < 2\pi\}$ is called the boundary locus of the Toeplitz matrix $\hat{A}_n$. It is worth noting that $g_A(e^{\mathbf{i}j\theta})$, $g_B(e^{\mathbf{i}j\theta})$ are the generating functions of the band Toeplitz matrices $\hat{A}_n$, $\hat{B}_n$, respectively; see, e.g., [8, 11].

The Toeplitz matrices related to the linear multistep formulas considered in section 2.2 have the boundary locus and their spectrum of eigenvalues in the right half plane; see [9]. We recall that the families of matrices $\{\hat{A}_n\}$, $\{\hat{B}_n\}$ are such that their entries α_j, β_j, $j = 0, \dots, k$, satisfy the system of linear equations (2.4), where p in (2.4) is the largest integer such that those equations are independent. Notice that the choice of ν is strictly related to the condition number of the underlying Toeplitz matrices; see [6, 8, 9]. For example, with the choice suggested in section 2.2, the matrices $\{\hat{A}_n\}$, $\{\hat{B}_n\}$ related to the formulas (2.9), (2.10), (2.11) have a condition number which increases at most linearly with their size; see [6]. On the other hand, the boundary locus and the spectrum of the eigenvalues of the matrices related to linear multistep formulas are not necessarily contained in one half plane for all k. For example, one can consider the matrices associated with well-known families of formulas used as initial value methods for a sufficiently large value of k in (2.3). This is the

case of the backward differentiation formulas for $k > 2$ and of the Adams–Moulton methods for $k > 1$. However, if ν is chosen differently from the choice suggested in section 2.2, the eigenvalues of $\{\hat{A}_n\}$, $\{\hat{B}_n\}$ can have both positive and negative (or zero!) real part. Indeed, it is easy to check that this is the case of formulas (2.10) used with $k = 5$ but with $\nu = 4$ instead of $\nu = (k+1)/2 = 3$.

We will assume, as is the case in practice for the methods described in section 2.2, that the influence of the small rank perturbations $E_n^{(A)} = A_n - \hat{A}_n$, $E_n^{(B)} = B_n - \hat{B}_n$ on the spectral properties of $\hat{A}_n$, $\hat{B}_n$ is moderate. More precisely, here we refer to suitably chosen additional schemes such that their related matrices A_n, B_n have the spectrum of eigenvalues in the right half plane, and the condition number of these is still of the order of $O(n)$, where n is their size. These hypotheses are usually reasonable; see [6]. On the other hand, notice that, in general, the influence of low rank modifications in the non-Hermitian case can very much change the spectral properties of a given matrix; see [22]. However, in this paper we will focus mainly on the preconditioner and on the spectrum of the component matrices of the preconditioner (1.2), which are normal and defined by using the coefficients of (2.3), i.e., by the entries of $\hat{A}_n$, $\hat{B}_n$ only.

We will need an explicit expression of the coefficients of the formulas (2.9) and (2.10). To this end, there are at least two (equivalent) strategies. In the first one, the coefficients can be computed by writing the formula of the GLM in backward difference form; see, e.g., [19, chapter 3]. Thus, expanding the backward differences of y_{n+j} for the formulas (2.9) and of f_{n+j} for the formulas (2.10), equating the coefficients of y_{n+j} and of f_{n+j}, $j = 0, \dots, k$, to the corresponding expressions, and using an induction argument gives the coefficients α_j, β_j, $j = 0, \dots, k$.

PROPOSITION 3.1. *The coefficients of the formulas* (2.9) *are given by*

$$\alpha_i = (-1)^{k-i} \sum_{j=k-i}^{k} \binom{j}{k-i} \delta_j, \qquad i = 0, \dots, k, \tag{3.3}$$

where

$$\delta_i = \begin{cases} 0, & i = 0; \\ 1, & i = 1; \\ \dfrac{1}{i!} \displaystyle\sum_{s=0}^{i-1} \prod_{\substack{j=0, \\ j \neq s}}^{i-1} (-(k-\nu)+j), & i \leq k-\nu,\ i \geq 1; \\ \dfrac{1}{i!} \displaystyle\prod_{\substack{j=0, \\ j \neq k-\nu}}^{i-1} (-(k-\nu)+j), & i > (k-\nu) \geq 1. \end{cases}$$

The coefficients of the formulas (2.10) *are given by*

$$\beta_i = \frac{(-1)^{k-i}}{(k-i)!} \sum_{j=k-i}^{k} \frac{1}{(j-(k-i))!} \int_{-(k-\nu)-1}^{-(k-\nu)} \prod_{m=0}^{j-1} (r+m)\, dr, \quad i = 0, \dots, k. \tag{3.4}$$

Proof. The expression (3.3) is derived by writing (2.9) in backward difference form (see [19, chapter 3]), i.e.,

$$\sum_{j=0}^{k} \delta_i \nabla^i y_{n+k-\nu} = h\, f_n,$$

where

$$\delta_i = (-1)^i \frac{d}{dr}\binom{-r}{i},$$

and the above derivative is computed at $r = k - \nu$. Thus, by observing that

$$\binom{-r}{j} = \frac{(-r-j+1)\cdots(-r)}{j!} = \frac{(-1)^j}{j!}\prod_{m=0}^{j-1}(r+m), \tag{3.5}$$

we have (3.3). The other expression, i.e., (3.4), is derived by observing that (2.10) can be written as (see [19, chapter 3])

$$y_{n+1} - y_n = h\sum_{j=0}^{k}\gamma_i \nabla^i f_{n+k-\nu+1}, \tag{3.6}$$

where

$$\gamma_j = (-1)^j \int_{-(k-\nu)-1}^{-(k-\nu)} \binom{-r}{j}\, dr.$$

Thus, from (3.5), we have (3.4). ∎

The other strategy is based on the explicit solution of linear equations in the unknowns α_j, β_j, $j = 0, \ldots, k$, by writing (2.4) in matrix form. Thus, we have to solve a linear system whose matrix is a Vandermonde-like one, and several combinatorial identities can be used. The coefficients of (2.9) and of (2.10) were computed following this strategy in [3]. We stress that the derivation of a useful expression can be rather lengthy. Full details can be found in [3, pp. 46–50, 66–69].

PROPOSITION 3.2. *The coefficients of the formula* (2.9) *are given by*

$$\begin{aligned}
\alpha_i &= \frac{(-1)^{\nu-i}}{\nu-i}\frac{\nu!(k-\nu)!}{i!(k-i)!}, \qquad i \neq \nu,\ i = 0, \ldots, k,\\
&= \frac{1}{\nu} = \frac{2}{k+1}, \qquad i \equiv \nu,\ k\ \textit{odd},\\
&= \frac{2\nu-1}{\nu(\nu-1)} = \frac{4(k+1)}{k(k+2)}, \qquad i \equiv \nu,\ k\ \textit{even},\ k \geq 1.
\end{aligned} \tag{3.7}$$

The coefficients of the formula (2.10) *are given by*

$$\beta_i = \frac{(-1)^{k-i}}{i!\,(k-i)!}\int_{\nu}^{\nu+1}\prod_{\substack{m=0,\\ m\neq i}}^{k}(t-m)\, dt, \qquad i = 0, \ldots, k. \tag{3.8}$$

Proof. The proof follows after some manipulations of the results in Theorem 4.1.1 for (3.7) and in Remark 4.2.2 for (3.8) in [3] by recalling that $\nu = (k+1)/2$ if k is odd, while $\nu = (k+2)/2$ for (2.9), $\nu = k/2$ for (2.10) if k is even. ∎

We remark that there is a third approach to derive the coefficients of the formulas (2.9) and (2.10) that is simpler than the other two. It is based on generating functions and symbolic operators; see, e.g., [19, sections 3.9–3.12]. Using that approach, we have that the generating function for δ_i in (3.3) is given by

$$G_1(z) = -(1-z)^{k-\nu}\log(1-z) = \sum_{i=0}^{\infty}\delta_i z^i.$$

Therefore,

$$\delta_i = \sum_{s=0}^{i-1} (-1)^s \binom{k-\nu}{s} \frac{1}{i-s}.$$

Similarly, the generating function for γ_i in (3.6) is given by

$$G_2(z) = \frac{-z(1-z)^{k-\nu}}{\log(1-z)} = \sum_{i=0}^{\infty} \gamma_i z^i,$$

and an explicit expression for γ_i can be derived accordingly.

Obviously, suitably manipulating the expressions derived by one strategy (e.g., (3.3), (3.4)) gives the expressions derived by the others (see, e.g., (3.7), (3.8), respectively).

COROLLARY 3.3. *The coefficients of the formulas* (2.9) *are uniformly bounded by* 2 *for all* $k \geq 1$. *Moreover,* $|\alpha_{i+1}| < |\alpha_i|$ *for* $i = \nu+1, \ldots, k-1$; $|\alpha_{i+1}| > |\alpha_i|$ *for* $i = 0, \ldots, \nu-2$; $|\alpha_\nu| < |\alpha_{\nu+1}|$; $|\alpha_{\nu-1}| > |\alpha_\nu|$; *and* $\lim_{k\to\infty} \alpha_j = 0$, $j = 0, k, \nu$.

Proof. By considering the expression (3.7) we have that $|\alpha_{i+1}| < |\alpha_i|$ for $i = \nu+1, \ldots, k-1$ and $|\alpha_{i+1}| > |\alpha_i|$ for $i = 0, \ldots, \nu-2$. For k odd, $\nu = (k+1)/2$, and, by (3.7), we have $\alpha_{\nu-1} = -\frac{(k+1)/2}{k-(k+1)/2+1} = -1$ while, for k even, $\nu = (k+2)/2$ and $\alpha_{\nu-1} = \frac{k+2}{k} \leq 2$ for $k \geq 2$. Similarly, for k odd, we have $\alpha_{\nu+1} = (k-1)/(k+4) < 1$, otherwise $\alpha_{\nu+1} = (k-2)/(k+4) < 1$, $k \geq 2$, and the proof is complete by recalling (3.7) again for $i = 0, \nu, k$. ∎

On the other hand, we can observe that for many families of linear multistep formulas the above results do not hold. This is the case of popular initial value methods such as the backward differentiation formula and the Adams–Moulton methods or of the schemes in section 2.2 with some choices of ν different from those suggested there. More precisely, some of the coefficients α_j and β_j, $j = 0, \ldots, k$, for the methods above, can grow boundlessly very fast for $k \to \infty$.

4. Circulant approximations for general linear multistep formulas. Let us consider the block preconditioners in (1.2) for the linear systems in (2.6) based on circulant-like matrices introduced in [4, 5, 7]. The approximating operators $\breve{A}$, $\breve{B}$ in (1.2) are computed by taking into account the coefficients of the formula (2.3), i.e., they are defined for the Toeplitz matrices $\hat{A}_n$, $\hat{B}_n$.

In what follows, we will recall in brief the main trigonometric approximations for the nonsymmetric matrices $\hat{A}_n$, $\hat{B}_n$ (and for A, B in (1.1)) we have found effective for the preconditioner (1.2); see also section 5.3. To this end, let $T_n = (t_j)$ be an $n \times n$ Toeplitz matrix whose diagonal entries are t_j, $j = -(n-1), \ldots, n-1$.

Strang's $s(T_n)$ (see [21]), sometimes called simple circulant approximation, is such that if $s_0, \ldots, s_{n-1}$ are the entries of the first row of the corresponding $n \times n$ preconditioner for T_n, we have

$$s_j = \begin{cases} t_j, & 0 < j \leq \left\lfloor \frac{n}{2} \right\rfloor, \\ \\ t_{j-n}, & \left\lfloor \frac{n}{2} \right\rfloor < j < n, \end{cases} \qquad j = 0, \ldots, n-1. \tag{4.1}$$

The spectrum of the Hermitian Toeplitz matrices preconditioned using Strang's preconditioner was analyzed in [10]. Notice that $s(T_n)$ is singular for the Toeplitz matrices

T_n whose generating function $f(\theta)$ is zero in $\theta = 0$, as observed, e.g., in [5, 24]. Unfortunately, the generating function of the matrix $\hat{A}_n$ always has a zero of multiplicity one in $\theta = 0$ because of the consistency condition $0 = \rho(1) = \sum_{j=0}^{k} \alpha_j$. Thus, as observed in [5], the approximation (4.1) cannot be safely used in the preconditioner (1.2), e.g., when the Jacobian matrix J in (2.6) has some very small or zero eigenvalues; see [5] for more details.

T. Chan's circulant preconditioner for the Toeplitz matrix T_n, denoted by $c(T_n)$, is defined such that $\|c(T_n) - T_n\|_F$ is minimum, where $c(T_n)$ is chosen in the set of $n \times n$ circulant matrices and $\|\cdot\|_F$ is the Frobenius norm. If $c_0, \ldots, c_{n-1}$ are the entries of the first row of $c(T_n)$ and t_j, $j = -(n-1), \ldots, n-1$, are the elements on the diagonals of the Toeplitz matrix T_n, we have (see [12])

$$c_j = \frac{(n-j)t_j + jt_{j-n}}{n}, \qquad j = 0, \ldots, n-1. \tag{4.2}$$

If the Toeplitz matrix T_n is Hermitian and positive definite, then these properties hold true for $c(T_n)$ as well; see [23]. Unfortunately, if T_n is nonsymmetric, $s(T_n)$ and $c(T_n)$ can have eigenvalues in the right and left half plane or zero as well, even for those matrices T_n whose eigenvalues have strictly positive real part. For example, this holds true for the underlying linear systems based on the formulas in section 2. Moreover, there are families of formulas (2.3) such that the circulant approximation (4.2) can be ill conditioned or even singular (e.g., those based on the midpoint method in boundary value form; see at the end of section 5.2).

Let us consider the P-circulant approximation introduced in [4]. Again, if T_n is a Toeplitz matrix whose entries of the diagonals are $t_{-(n-1)}, \ldots, t_{n-1}$, we have that the entries $p_0, \ldots, p_{n-1}$ of the first row of the P-circulant preconditioner $p(T_n)$ for T_n are given by

$$p_j = \frac{(n+j)t_j + jt_{j-n}}{n}, \qquad j = 0, \ldots, n-1. \tag{4.3}$$

Notice that the P-circulant and simple and T. Chan's circulants are equivalent in the sense of the linear approximation processes; see [20]. In practice, P-circulant matrices come from using the Frobenius norm weight $(n-j)/n$ for the lower and the weight $(n+j)/n$ for the upper diagonals, respectively. The circulant matrices, whose entries are defined in (4.3), have been called P-circulant in [4] because, for some classes of Toeplitz matrices (and thus for formulas (2.3)), their eigenvalues have positive real part; see section 5.2. This property can speed up the convergence process with respect to the other basic approximations described here; see section 5.3. We stress that P-circulants neither preserve symmetry (but for our purpose this is not essential) nor minimize the "distance" with the original Toeplitz matrix. More precisely, $||p(T_n) - T_n||$ is not minimized with respect to the p-norms (e.g., $p = 1, 2, \infty$) nor the Frobenius norm.

The MS-circulant approximation for T_n is given by a rank-one perturbation of the simple circulant preconditioner whose zero eigenvalue in (5.1) is set to a suitable nonzero value c for those Toeplitz matrices T_n whose generating functions have a zero. We achieved interesting results in [5] by setting $c = 1/n$ and $c = \min_r\{\mathrm{Re}(\phi_r)\} > 0$, where ϕ_r, $r = 1, \ldots, n$, are the eigenvalues of $s(T_n)$.

Finally, the $\{\omega\}$-circulant approximation can be considered as another extension of the simple circulant approximation. Let T_n be a n_1-band Toeplitz matrix, $n_1 < \lceil n/2 \rceil$. The $\{\omega\}$-circulant matrix $\tilde{s}(t_n)$ differs from the simple circulant $s(T_n)$ because

the entries outside the diagonals $-n_1, \dots, n_1$ of $\tilde{s}(t_n)$ are given by those of $s(T_n)$ multiplied by $\omega = \exp(\mathbf{i}\theta)$, $0 < \theta \leq \pi$, and $\tilde{s}(T_n)$ is nonsingular even if the generating function of T_n has a zero for $\theta = 0$; see [7].

We observe that some combinations of the above approximations can give further useful preconditioners as well. For example, it is straightforward to define $\{\omega\}$–P-circulant preconditioners by using (4.3) and $\{\omega\}$-circulant matrices instead of circulant matrices. The arguments used in the following sections can be adapted for these preconditioners as well, in general, and we will focus only on the "basic" approximations above.

5. The spectrum of the circulant approximations. The Toeplitz matrices $\hat{A}_n$, $\hat{B}_n$ in (3.1) are positive stable for the linear multistep formulas (2.9), (2.10), (2.11); see [9, chapter 11]. We recall that a square matrix is said to be (semi)positive stable if its eigenvalues have positive (nonnegative) real part; see, e.g., [16]. It is straightforward to note that positive stable matrices are nonsingular.

Let us consider $n = s + 1$ and the $(s + 1) \times (s + 1)$ P-circulant matrix $p(A)$ defined in (4.3) for the Toeplitz matrix A in (2.7) (and then for $\hat{A}_{s+1}$ in (3.1)). The eigenvalues ϕ_j, $j = 0, \dots, s$, of $p(A)$ can be computed by a linear combination of the entries of the first row (see Davis [13]):

$$\phi_l = \sum_{j=0}^{s} p_j \epsilon^{jl}, \qquad l = 0, \dots, s, \quad \epsilon = e^{2\pi\,\mathbf{i}/(s+1)}. \tag{5.1}$$

From (4.3) we have

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

Therefore,

$$\phi_l = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \epsilon^{jl}, \qquad l = 0, \dots, s. \tag{5.2}$$

A similar expression holds for the eigenvalues $\psi_0, \dots, \psi_s$ of $p(B)$:

$$\psi_l = \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu} \left(1 + \frac{j}{s+1}\right) \epsilon^{jl}, \qquad l = 0, \dots, s. \tag{5.3}$$

Notice that (5.2) and (5.3) are trigonometric sums. Let us define

$$\hat{\Phi}_k(x) = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \cos(j\,x), \qquad x \in \mathbb{R}, \tag{5.4}$$

$$\hat{\Psi}_k(x) = \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu} \left(1 + \frac{j}{s+1}\right) \cos(j\,x), \qquad x \in \mathbb{R}. \tag{5.5}$$

We observe that (5.4) and (5.5) are analytic functions (for $k < \infty$). From (5.2), we have that

$$\hat{\Phi}_k\left(\frac{2\pi l}{s+1}\right) = \mathrm{Re}(\phi_l), \quad \hat{\Psi}_k\left(\frac{2\pi l}{s+1}\right) = \mathrm{Re}(\psi_l), \qquad l = 0, \dots, s.$$

Thus, it is straightforward to see that if $\hat{\Phi}_k(x)$, $\hat{\Psi}_k(x)$ are positive for real values of x, then $p(A)$ and $p(B)$ are positive stable.

By using similar arguments, we can derive the expression of the eigenvalues of $s(A_{s+1})$, $s(B_{s+1})$ and $c(A_{s+1})$, $c(B_{s+1})$, respectively:

$$\gamma_l = \sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu}\epsilon^{jl}, \qquad \delta_l = \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu}\epsilon^{jl}, \qquad l = 0, \ldots, s, \tag{5.6}$$

and

$$\sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu}\left(1 - \frac{|j|}{s+1}\right)\epsilon^{jl}, \qquad \sum_{j=-\nu}^{k-\nu} \beta_{j+\nu}\left(1 - \frac{|j|}{s+1}\right)\epsilon^{jl}, \qquad l = 0, \ldots, s. \tag{5.7}$$

Notice that the eigenvalues of $s(A_{s+1})$, $s(B_{s+1})$ lie on the boundary locus of A_{s+1}, B_{s+1}, respectively.

5.1. Preliminary results. First, let us give some properties of the trigonometric sums (5.4), (5.5).

We recall that a sequence $\{c_j\}$ is of bounded variation (see [25]) if the series $\sum_{j=0}^{\infty} |c_{j+1} - c_j|$ converges. If $\{c_j\}$ tends monotonically to zero, then $\{c_j\}$ is of bounded variation. It is useful to simplify the expressions (5.4) and (5.5) by observing that $\cos(x)$ is an even function.

LEMMA 5.1. *The function* $\hat{\Phi}_k(x)$ *in* (5.4) *can be expressed for* (2.9) *as*

$$\hat{\Phi}_k(x) = \frac{a_0}{2} + \sum_{n=1}^{k-\nu} (-1)^n a_n \cos(n\,x), \tag{5.8}$$

where $a_0 = 2\alpha_\nu$, $\nu = (k+1)/2$ *if* k *is odd,* $\nu = (k+2)/2$ *if* k *is even, and* $a_n = (-1)^n \tilde{a}_n$,

$$\tilde{a}_n = \alpha_{n+\nu}\left(1 + \frac{n}{s+1}\right) + \alpha_{-n+\nu}\left(1 - \frac{n}{s+1}\right), \qquad n = 1, \ldots, k-\nu. \tag{5.9}$$

It is intended that α_j *is zero if* $j < 0$ *or* $j > k$. *The sequence* $\{a_n\}$ *has the following properties:*

(1) $a_n \geq 0$, $n \geq 0$;
(2) a_n *tends to zero if* $n \to \infty$;
(3) a_n *is uniformly bounded (i.e.,* $0 \leq a_n < 2$, $n \geq 0$*);*
(4) $\{a_n\}$ *is monotonic decreasing;*
(5) $\{a_n\}$ *is of bounded variation.*

Proof. The expression (5.8) follows by observing that from (3.7), (5.4), and (5.9) we have

$$\tilde{a}_n = \frac{(-1)^n}{n}\left\{ -\frac{\binom{k}{\nu+n}}{\binom{k}{\nu}}\left(1 + \frac{n}{s+1}\right) + \frac{\binom{k}{\nu-n}}{\binom{k}{\nu}}\left(1 - \frac{n}{s+1}\right)\right\} = (-1)^n \cdot a_n.$$

(1) Let us check for first that $a_n > 0$ for $n \geq 1$, $n \leq k - \nu$ (recall that $a_0 = 2\alpha_\nu$ is positive; see (3.7)). From here on, it is intended that $a_n = 0$ if $n > k - \nu$. Again,

from the expression (3.7), we have

$$
\begin{aligned}
n\cdot a_n &= \frac{\binom{k}{\nu-n}}{\binom{k}{\nu}}\left(1-\frac{n}{s+1}\right)-\frac{\binom{k}{\nu+n}}{\binom{k}{\nu}}\left(1+\frac{n}{s+1}\right)\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\cdot\left(1-\frac{n}{s+1}\right)\right]-\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-n}{\nu+n}\cdot\left(1+\frac{n}{s+1}\right)\right]\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\left[\left(1-\frac{n}{s+1}\right)-\frac{\nu-n}{\nu+n}\left(1+\frac{n}{s+1}\right)\right]\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\frac{2n}{(\nu+n)(s+1)}\cdot(s+1-\nu)>0.
\end{aligned}
$$

Indeed, notice that the term in square brackets above can assume values in $(0,1)$, and $(s+1-\nu)$ is greater than zero because $s\geq k\geq\nu\geq 1$ by hypothesis; see section 2.2.

(2) Now, let us check that a_n converges to zero for $n\to\infty$. From the last expression, we have

$$
\begin{aligned}
a_n &= \left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\frac{2(s+1-\nu)}{(\nu+n)(s+1)}\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\frac{2}{\nu+n}\cdot\left(1-\frac{\nu}{s+1}\right)\\
&\leq\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\frac{2}{n}\leq\frac{2}{n} \qquad (5.10)
\end{aligned}
$$

because the term in square brackets above assumes values in $(0,1)$, $n\leq\nu$, and $0<1-\nu/(s+1)<1$ because $s\geq k\geq\nu\geq 1$.

(3) It is an immediate consequence of the bound in (5.10).

(4) $\{a_n\}$ is monotonic (decreasing). Indeed,

$$
\begin{aligned}
&a_{n+1}-a_n\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot\left[\frac{2\nu(\nu-n)}{(\nu+n+1)(s+1)(\nu+n)}-\frac{2\nu}{(\nu+n)(s+1)}\right]\\
&=\left[\frac{\nu-1}{\nu+1}\cdot\ldots\cdot\frac{\nu-(n-1)}{\nu+(n-1)}\right]\cdot-\frac{2\nu}{s+1}\cdot\frac{2n+1}{(\nu+n+1)(\nu+n)}<0
\end{aligned}
$$

by using similar arguments as in (1) and (2).

(5) Finally, for (1)–(4), $\{a_n\}$ is of bounded variation. □

We recall that a sequence of functions is said to converge locally uniformly on a set $\mathcal{S}$ if it converges uniformly on every compact subset of $\mathcal{S}$; see, e.g., [17, p. 160].

PROPOSITION 5.2. *The sequence of functions $\{\hat{\Phi}_k(x)\}$ for (2.9) converges locally uniformly with respect to k (and then with respect to $\nu=O(k)$, see section 2) in $(-\pi,\pi)$.*

Proof. It is a consequence of Lemma 5.1 and of [25, Theorem 2.7, p. 4]. □

COROLLARY 5.3. *Under the hypotheses of Proposition 5.2, the function*

$$\hat{\Phi}(x)=\lim_{k\to\infty}\hat{\Phi}_k(x) \qquad (5.11)$$

is an analytic function for $x\in(-\pi,\pi)$ and continuous for $x\in\mathbb{R}$.

Proof. It is a consequence of Proposition 5.2 and of [17, Corollary 3.4c, p. 161]. The continuity over the whole real axis derives from Abel's limit theorem applied in $x = \pm\pi$ and considering that, for n integer, we have

$$\hat{\Phi}_k(\pm x + 2n\pi) = \hat{\Phi}_k(x), \qquad \hat{\Phi}_k(2\pi - x)' = -\hat{\Phi}_k(x)', \tag{5.12}$$

$$\hat{\Psi}_k(\pm x + 2n\pi) = \hat{\Psi}_k(x), \qquad \hat{\Psi}_k(2\pi - x)' = -\hat{\Psi}_k(x)', \qquad x \in [0, 2\pi],\ k \geq 1. \tag{5.13}$$

Similar expressions hold true for $\hat{\Phi}(x)$ and $\hat{\Psi}(x)$. □

LEMMA 5.4. *Let k be an integer and $\nu = \lceil (k+1)/2 \rceil$ (k even or odd) or $\nu = (k+2)/2$ (k even). For $-(k-\nu) \leq n \leq (k-\nu)$, we have*

$$(\nu + n)!\,(k - \nu - n)! \geq (\nu - n)!\,(k - \nu + n)!, \tag{5.14}$$

$$\nu!\,(k-\nu)! \leq (\nu + n)!\,(k - \nu - n)!. \tag{5.15}$$

Proof. Let us consider (5.14). We have

$$\frac{(\nu-n)!(k-\nu+n)!}{(\nu+n)!(k-\nu-n)!} = \frac{(\nu-n)!(k-(\nu-n))!}{k!} \cdot \frac{k!}{(\nu+n)!(k-(\nu+n))!} = \frac{\binom{k}{\nu+n}}{\binom{k}{\nu-n}},$$

where the ratio above is equal to 1 if k is even and it is less than 1 otherwise. Indeed,

$$\binom{k}{\nu-n} = \binom{k}{k-(\nu-n)} = \begin{cases} \binom{k}{\nu+n} & \text{if } k-\nu = \nu \text{ (and } k \text{ is even)}, \\ \binom{k}{\nu+n-1} & \text{if } k-\nu = \nu-1 \text{ (and } k \text{ is odd)}\ . \end{cases}$$

Thus, for k odd, we have

$$\binom{k}{\nu+n-1} = \binom{k+1}{\nu+n} - \binom{k}{\nu+n} \Rightarrow \frac{\binom{k}{\nu+n}}{\binom{k}{\nu-n}} = \frac{1}{\frac{\nu+n-1}{\nu-n}} < 1, \quad 1 \leq n \leq (k-\nu).$$

Using similar arguments, we can see that (5.14) holds true for $\nu = (k+2)/2$ and k even as well. Now, let us consider (5.15) for $n \geq 1$ (for negative values of n a similar argument can be used). We have

$$\frac{\nu!\,(k-\nu)!}{(\nu+n)!\,(k-\nu-n)!} = \frac{k-\nu}{\nu+1} \cdot \frac{k-\nu-1}{\nu+2} \cdots \frac{k-\nu-(n-1)}{\nu+n},$$

where we have that the above expression is equal to

$$\begin{cases} \dfrac{\nu-1}{\nu+1} \cdots \dfrac{\nu-n}{\nu+n} < 1 & \text{if } k-\nu = \nu \text{ and } k \text{ is odd}, \\ \dfrac{\nu}{\nu+1} \cdots \dfrac{\nu-(n-1)}{\nu+n} < 1 & \text{if } k-\nu = \nu-1 \text{ and } k \text{ is even}\ . \end{cases}$$

Using similar arguments, we can see that (5.15) holds true for $\nu = (k+2)/2$ and k even as well. □

LEMMA 5.5. *The function* $\hat{\Psi}_k(x)$ *can be expressed for* (2.10) *as*

$$\hat{\Psi}_k(x) = \frac{b_0}{2} + \sum_{n=1}^{k-\nu} (-1)^{n+1} b_n \cos(n\,x), \tag{5.16}$$

where $b_0 = 2\beta_\nu$, $\nu = \lceil (k+1)/2 \rceil$, *and* $b_n = (-1)^{n+1}\tilde{b}_n$, $n \geq 1$,

$$\tilde{b}_n = \beta_{n+\nu}\left(1 + \frac{n}{s+1}\right) + \beta_{-n+\nu}\left(1 - \frac{n}{s+1}\right), \qquad n = 1, \ldots, k-\nu. \tag{5.17}$$

It is intended that β_j *is zero if* $j < 0$ *or* $j > k$. *The sequence* $\{b_n\}$ *has the following properties:*

(1) $b_n \geq 0$, $n \geq 0$;
(2) b_n *tends to zero if* $n \to \infty$;
(3) b_n *is uniformly bounded (i.e.,* $0 \leq b_n < 2$, $n \geq 0$*);*
(4) $\{b_n\}$ *is monotonic decreasing;*
(5) $\{b_n\}$ *is of bounded variation.*

Proof. (1) By expanding (5.17), we have

$$\begin{aligned}\tilde{b}_n = &\left[\frac{(-1)^{k-\nu-n}}{(\nu+n)!(k-\nu-n)!}\int_\nu^{\nu+1} \prod_{\substack{m=0,\\ m\neq\nu+n}}^{k} (t-m)\,dt\right]\left(1 + \frac{n}{s+1}\right)\\ &+ \left[\frac{(-1)^{k-\nu+n}}{(\nu-n)!(k-\nu+n)!}\int_\nu^{\nu+1} \prod_{\substack{m=0,\\ m\neq\nu-n}}^{k} (t-m)\,dt\right]\left(1 - \frac{n}{s+1}\right).\end{aligned} \tag{5.18}$$

Thus, for $n \geq 1$, we have

$$\begin{aligned}\tilde{b}_n = (-1)^{n+1}\int_\nu^{\nu+1} &\left[\frac{1+n/(s+1)}{(\nu+n)!(k-\nu-n)!}\frac{1}{|t-\nu-n|}\right.\\ &\left. - \frac{1-n/(s+1)}{(\nu-n)!(k-\nu+n)!}\frac{1}{(t-\nu+n)}\right]\prod_{m=0}^{k}|t-m|\,dt = (-1)^{n+1}\cdot b_n,\end{aligned} \tag{5.19}$$

while, by observing that $t-m$, $m = 0, \ldots, k$, do not change sign for $t \in (\nu, \nu+1)$,

$$\tilde{b}_0 \equiv b_0 = 2\beta_\nu = \frac{2(-1)^{k-\nu}}{\nu!(k-\nu)!}\int_\nu^{\nu+1}\prod_{\substack{m=0,\\ m\neq\nu}}^{k}(t-m)\,dt = \frac{2}{\nu!(k-\nu)!}\int_\nu^{\nu+1}\prod_{\substack{m=0,\\ m\neq\nu}}^{k}|t-m|\,dt, \tag{5.20}$$

therefore (5.20) is positive. To check that $b_n > 0$, $n \geq 1$, and $n \leq k-\nu$ (it is intended that $b_n = 0$ for $n > k-\nu$), it is enough to see that the part in square bracket in (5.19) is positive or zero. For brevity, let us consider k even. By Lemma 5.4, the part in square brackets in (5.19) can be rewritten as

$$\frac{1}{(\nu+n)!(k-\nu-n)!}\left(\frac{1+n/(s+1)}{n-(t-\nu)} - \frac{1-n/(s+1)}{n+(t-\nu)}\right)$$

$$= \frac{(s+1+n)\,(n+(t-\nu)) - (s+1-n)\,(n-(t-\nu))}{(\nu+n)!\,(k-\nu-n)!\,(n-(t-\nu))\,(n+(t-\nu))\,(s+1)}.$$

Then, we can observe that the ratio above is positive because the denominator of the related expression is positive, $s+1+n>0$, $n+(t-\nu)>0$, and

$$\frac{(s+1-n)(n-(t-\nu))}{(s+1+n)(n+(t-\nu))} < 1, \quad n \geq 1, \quad 0 \leq t \leq \nu+1.$$

For k odd a similar argument can be used, and (1) and the expression (5.16) are verified.

(2) Let us check first that b_0 is bounded. We observe that, for $t = t^* = \nu + \epsilon(\nu)$, $0 < \epsilon(\nu) \to 0$ for $k, \nu \to \infty$ (recall that $\nu = O(k)$), the following function

$$f(t) = \prod_{\substack{m=0,\\ m\neq\nu}}^{k} |t-m|, \qquad \nu \leq t \leq (\nu+1), \tag{5.21}$$

reaches its (unique) maximum in the segment $\nu \leq t \leq (\nu+1)$. This can be checked by considering the derivative df/dt of (5.21) in $(\nu, \nu+1)$ and applying an induction argument on k. Thus, $f(t^*) = c\cdot\nu!\,(k-\nu)!$, where $c = c(\nu)$ is a parameter of the order of 1 that converges fast to 1 as $\nu \to \infty$ and, by (5.20), b_0 is uniformly bounded above by 2. As a corollary of the above result, by (3.8), β_ν is uniformly bounded above by 1. Now, to check that b_n is bounded for $n \geq 1$, it is enough to observe that both the following factors in (5.19)

$$\frac{1}{(\nu+n)!(k-\nu-n)!}\int_{\nu}^{\nu+1}\prod_{m=0}^{k}|t-m|\,dt, \quad \frac{1}{(\nu-n)!(k-\nu+n)!}\int_{\nu}^{\nu+1}\prod_{m=0}^{k}|t-m|\,dt$$

are positive and bounded by a constant of the order of unity. To this end, notice that

$$G(t) = \prod_{m=0}^{k}(t-m) = \begin{cases} (-1)^{k+1}\cdot\dfrac{\Gamma(k+1-t)}{\Gamma(-t)}, & \nu < t < \nu+1, \\ 0, & t = \nu \text{ or } t = \nu+1, \end{cases} \tag{5.22}$$

where $\Gamma(z)$ is the Gamma function (see [1] for definitions and some properties). The equality (5.22) can be derived by using arguments in [1, pp. 12–13]. It is straightforward to observe that

$$\int_{\nu}^{\nu+1}\prod_{m=0}^{k}|t-m| \leq \sup_{\nu<t<\nu+1}\prod_{m=0}^{k}|t-m|.$$

Let us denote by t^* the maximum of the function $|G(t)|$ in $\nu \leq t \leq \nu+1$; $G(t)$ is defined in (5.22). By considering dG/dt in $(\nu, \nu+1)$ and using an induction argument on k, we have that $t^* = \nu + 1/2 + \epsilon(\nu)$, where $\epsilon(\nu) \to 0$ as $\nu \to \infty$ ($k \to \infty$). By using the definition of $\Gamma(z)$, we have $\Gamma(x+1) = x\,\Gamma(x) \Rightarrow \Gamma(x) = \Gamma(x+1)/x$, where x cannot be a negative integer or zero. Applying repeatedly $\Gamma(x) = \Gamma(x+1)/x$, we have

$$|\Gamma(-t)|^{-1} \leq \nu!\,(t^*-\nu)\,\frac{1}{|\Gamma(-t^*+\nu+1)|}. \tag{5.23}$$

By observing that

$$(k-\nu-1)! < \Gamma(k+1-t^*) < (k-\nu)!,\ 0 < c \equiv \Gamma(-t^*+\nu+1) < 1,\ 1 < 1/(t^*-\nu) < 2$$

and by recalling Lemma 5.4, we can write

$$|G(t)| \le c \cdot (t^*-\nu)^{-1} \cdot \nu! \cdot (k-\nu)!;$$

thus, b_n is bounded above by 2.

(3) To check that $\{b_n\}$ converges to zero as $n \to \infty$, arguments similar to those used to prove (1), (2) give that b_n in (5.19) can be written as $b_n = \frac{1}{n} \cdot h_n$, where $\{h_n\}$ is a uniformly bounded sequence.

(4) To check that $\{b_n\}$ is a monotonic nonnegative decreasing sequence, we observe that the expression (5.19) for $n > 1$ and for k even gives

$$\begin{aligned} b_n - b_{n+1} = \frac{1}{(\nu+n)!(k-\nu-n)!} \int_\nu^{\nu+1} \left\{ \left[\frac{1+n/(s+1)}{n-(t-\nu)} - \frac{1-n/(s+1)}{n+(t-\nu)} \right] \right. \\ \left. - \left[\frac{1+(n+1)/(s+1)}{(n+1)-(t-\nu)} - \frac{1-(n+1)/(s+1)}{(n+1)+(t-\nu)} \right] \right\} \prod_{m=0}^{k} |t-m|\, dt. \end{aligned} \tag{5.24}$$

Let us consider the expression in curly brackets in (5.24). We have the following lower bound:

$$\begin{aligned} \{\cdot\} &> \frac{1+n/(s+1)}{n-1} - \frac{1+(n+1)/(s+1)}{n+1} - \frac{1-n/(s+1)}{n-1} + \frac{1-(n+1)/(s+1)}{n+1} \\ &= \frac{2n}{(s+1)(n-1)} - \frac{2(n+1)}{(s+1)(n+1)} = \frac{2}{s+1}\left(\frac{n}{n-1} - 1\right) > 0, \end{aligned} \tag{5.25}$$

and thus, by (5.24), $b_n - b_{n+1} > 0$ and the sequence $\{b_n\}$ is monotonic decreasing.

(5) For (1)–(4), $\{b_n\}$ is of bounded variation. ∎

Finally, by using similar arguments such as in the Proposition 5.2 and Corollary 5.3, we have the following results.

PROPOSITION 5.6. *The sequence of functions* $\{\hat{\Psi}_k(x)\}$ *for* (2.10) *converges locally uniformly with respect to* k *(and thus with respect to* $\nu = O(k)$*) in* $(-\pi, \pi)$.

Proof. It is a consequence of Lemma 5.5 and of [25, Theorem 2.7, p. 4]. ∎

COROLLARY 5.7. *Under the hypotheses of Proposition* 5.6, *the function*

$$\hat{\Psi}(x) = \lim_{k\to\infty} \hat{\Psi}_k(x) \tag{5.26}$$

is analytic for $x \in (-\pi, \pi)$ *and continuous for* $x \in \mathbb{R}$. ∎

5.2. Main results. As a consequence of the results in the previous section, we can give bounds for the eigenvalues for some of the underlying approximations.

THEOREM 5.8. *The P-circulant matrices* $p(A_{s+1})$, $p(B_{s+1})$ *related to the formulas* (2.9), (2.10) *are positive stable and, if* ϕ_j, ψ_j, $j = 0, \dots, s$, *are the eigenvalues of* $p(A_{s+1})$, $p(B_{s+1})$, *respectively, we have*

$$\frac{1}{s+1} \le \mathrm{Re}(\phi_j) < 2, \qquad \mathrm{Re}(\psi_j) = 1 \text{ for } (2.9), \tag{5.27}$$

$$\frac{1}{s+1} \le \mathrm{Re}(\phi_j) < \frac{2s+1}{s+1} < 2, \qquad \frac{2}{\pi^2(s+1)} < \mathrm{Re}(\psi_j) < 1 \ \text{ for } (2.10). \tag{5.28}$$

Proof. Let us observe that, from (4.3) and (5.2), considering the scaling and consistency conditions $\rho(1) = 0$, $\sigma(1) = 1$, we have

$$\sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} = 0, \quad \sum_{j=-\nu}^{k-\nu} j\alpha_{j+\nu} = 1. \tag{5.29}$$

Thus, the expression (5.2) gives $\phi_0 = \hat{\Phi}(0) = 1/(s+1)$ for the formulas (2.3).

Let us check (5.27). To this end, we consider formulas (2.9) and expand $\cos(jx)$ in the right-hand side of (5.4) using power series in a neighborhood of the origin. If $\mathcal{P}(f)$ is the formal power series expansion of a function f, we can write

$$\mathcal{P}\left(\hat{\Phi}_k(x)\right) = \frac{1}{s+1} + \sum_{n=1}^{\infty} \left\{ (-1)^n \frac{x^{2n}}{(2n)!} \sum_{j=-\nu}^{k-\nu} j^{2n} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \right\}. \tag{5.30}$$

For brevity, we consider k odd ($\Rightarrow \nu = (k+1)/2$). From (2.13), it is worth noting that (5.30) is equivalent to the following expression:

$$\mathcal{P}\left(\hat{\Phi}_k(x)\right) = \frac{1}{s+1} + \sum_{n=\nu}^{\infty} \left\{ (-1)^n \frac{x^{2n}}{(2n)!} \sum_{j=-\nu}^{k-\nu} j^{2n} \alpha_{j+\nu} \left(1 + \frac{j}{s+1}\right) \right\}. \tag{5.31}$$

However, in Proposition 5.2, we observed that $\{\hat{\Phi}_k(x)\}$ converges locally uniformly in $\mathcal{S} = (-\pi, \pi)$ with respect to k (i.e., to ν because $k = 2\nu - 1$) for (2.9). Moreover, the functions $f_n(x) = (-1)^n a_n \cos(n\,x)$ in (5.8) are analytic in $\mathcal{S}$. Then, by [17, Corollary 3.4c, p. 161], we have that $\sum f_n(x)$, i.e., $\hat{\Phi}(x)$ in (5.11), is analytic in $\mathcal{S}$ and that the sequence $\{\hat{\Phi}_k(x)\}$ converges in $\mathcal{S}$ and the series related to (5.30) (and (5.31)) is the Taylor series of $\hat{\Phi}(x)$. However, by Abel's limit theorem for the power expansions, we have that the Taylor expansion in (5.30) (and (5.31)) converges for $x = \pm\pi$ as well,

$$\hat{\Phi}(x) := \lim_{k\to+\infty} \hat{\Phi}_k(x) = \lim_{k\to+\infty} \mathcal{P}(\hat{\Phi}_k(x)), \qquad x \in [-\pi, \pi], \tag{5.32}$$

and thus, by (5.12), for $x \in \mathbb{R}$. To conclude the first part of the proof, we observe that the quantity in the curly brackets in (5.31) is positive for $n = \nu$ (i.e., the first term of the sum), vanishes fast for $k \to \infty$ (recall (2.4) and (2.13)), and we can see that (see Figure 5.1, right)

$$\frac{1}{s+1} \le \Phi_k(x) \le \Phi_1(x) \le 2, \qquad k \ge 1, \qquad -\pi \le x \le \pi.$$

Let us now check (5.28). We can expand $\cos(jx)$ in the expression (5.5) in Taylor series in a neighborhood of the origin, and, for $k < \infty$, we have

$$\hat{\Psi}_k(x) = \sum_{n=0}^{\infty} \left\{ (-1)^n \frac{x^{2n}}{(2n)!} \sum_{j=-\nu}^{k-\nu} j^{2n} \beta_{j+\nu} \left(1 + \frac{j}{s+1}\right) \right\}. \tag{5.33}$$

By considering the order conditions (2.14), recalling the power series expansion of $\sin(x)$ and of $\cos(x)$ in a neighborhood of the origin and arguments similar to those

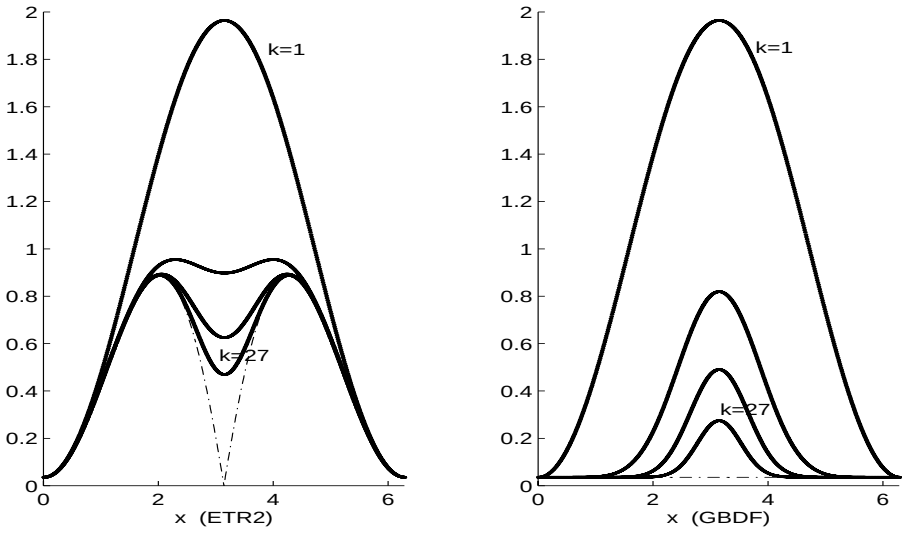

FIG. 5.1. $\hat{\Phi}_k(x)$, $k = 1, 7, 15, 27$, $s = 30$ *for formula* (2.11) *(left) and for formula* (2.9) *(right). The dashed curves give* $\hat{\Phi}(x)$.

used to prove (5.27), we have that, for $k \to \infty$ (i.e., $\nu \to \infty$ because $\nu = O(k)$),

$$\lim_{k\to\infty} \hat{\Psi}_k(x) = \hat{\Psi}(x) = \begin{cases} \dfrac{\sin(x)}{x} - \dfrac{1}{s+1}\left(\dfrac{\sin(x)}{x} - \dfrac{1}{2}\dfrac{\sin^2(x/2)}{(x/2)^2}\right), \\ \quad x \in (-\pi, 0) \cup (0, \pi), \\ \\ 1 - \dfrac{1}{2(s+1)}, \qquad x = 0. \end{cases}$$

Thus, by using similar arguments as before, the expressions (5.12), (5.13), and Abel's limit theorem, we see that the following inequalities hold true for (2.10):

$$\hat{\Psi}_k(x) \geq \hat{\Psi}(\pi) = \frac{2}{\pi^2(s+1)} > 0, \qquad x \in \mathbb{R},$$

$$\frac{1}{s+1} \leq \hat{\Phi}_k(x) \equiv \hat{\Phi}_1(x) \leq \frac{2s+1}{s+1} < 2, \qquad x \in \mathbb{R},\ k \geq 1.$$

The behavior of $\hat{\Psi}_k(x)$ for some values of k is displayed in Figure 5.2. □

We observe that the imaginary parts of the eigenvalues of the circulant approximations (4.1) and (4.3) for the matrices A and B in (2.6) for the formulas (2.9) and (2.10) are uniformly bounded by constants of the order of unity.

THEOREM 5.9. *If* ϕ_j, ψ_j, $j = 0, \ldots, s$, *are the eigenvalues of* $p(A_{s+1})$, $p(B_{s+1})$, *respectively, we have*

$$-\pi < \mathrm{Im}(\phi_j) < \pi, \qquad \mathrm{Im}(\psi_j) = 0 \qquad \textit{for } (2.9), \tag{5.34}$$

$$-\frac{s}{s+1} \leq \mathrm{Im}(\phi_j) < \frac{s}{s+1}, \qquad -c < \mathrm{Im}(\psi_j) < c \qquad \textit{for } (2.10),\ j = 0, \ldots, s, \tag{5.35}$$

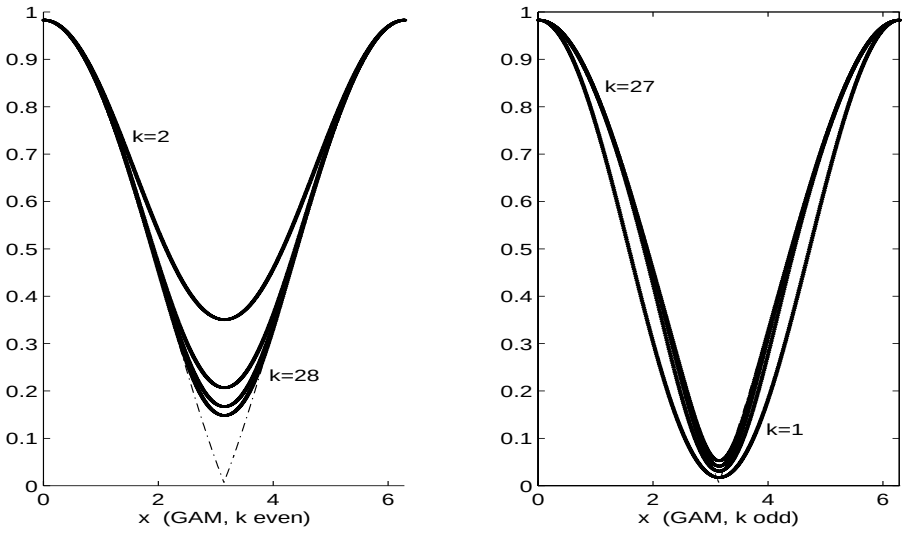

FIG. 5.2. *Left:* $\hat{\Psi}_k(x)$, $k = 2, 6, 14, 28$ *(k even); right:* $\hat{\Psi}_k(x)$, $k = 1, 7, 15, 27$ *(k odd) for formula* (2.10), $s = 28$. *The dashed curve gives* $\hat{\Psi}(x)$.

where

$$c = \max_{0<x<\pi} \left| \frac{\cos(x) - 1}{x} \right| \quad (\Rightarrow 0.7246 < c < 0.7247).$$

Proof. The proof uses arguments similar to those in the proof of Theorem 5.8. □

Again, as a corollary of Theorem 5.8, we have the following results.

THEOREM 5.10. *The preconditioners* $s(A_{s+1})$, $s(B_{s+1})$ *defined in* (4.1) *and related to the formulas* (2.9), (2.10) *are semipositive stable. If* γ_j, δ_j, $j = 0, \ldots, s$, *are the eigenvalues of* $s(A_{s+1})$, $s(B_{s+1})$, *respectively, we have*

$$0 \le \mathrm{Re}(\gamma_j) \le 2, \qquad \mathrm{Re}(\delta_j) = 1 \qquad \textit{for (2.9)}, \tag{5.36}$$

$$0 \le \mathrm{Re}(\gamma_j) \le 2, \qquad 0 \le \mathrm{Re}(\delta_j) \le 1 \qquad \textit{for (2.10)}, \ j = 0, \ldots, s. \tag{5.37}$$

THEOREM 5.11. *The* $\{\omega\}$*-circulant preconditioners* $\tilde{s}(A_{s+1})$, $\tilde{s}(B_{s+1})$ *for* $\omega = \exp(\mathbf{i}\theta)$, $0 < \theta \le \pi$, *and the MS-circulant preconditioners defined in section* 4, *related to the formulas* (2.9), (2.10), *are positive stable. If* $\tilde{\gamma}_j$, $\tilde{\delta}_j$, $j = 0, \ldots, s$, *are the eigenvalues of* $\tilde{s}(A_{s+1})$, $\tilde{s}(B_{s+1})$, *respectively, we have*

$$0 < \mathrm{Re}(\tilde{\gamma}_j) \le 2, \qquad \mathrm{Re}(\tilde{\delta}_j) = 1 \qquad \textit{for (2.9)}, \tag{5.38}$$

$$0 < \mathrm{Re}(\tilde{\gamma}_j) \le 2, \qquad 0 \le \mathrm{Re}(\tilde{\delta}_j) \le 1 \qquad \textit{for (2.10)}, \ j = 0, \ldots, s. \tag{5.39}$$

The bounds (5.38), (5.39) *hold true for the eigenvalues of the MS-circulant approximations as well.*

Similarly, it is straightforward to derive a result analogous to Theorem 5.9 for simple $\{\omega\}$-circulant and MS-circulant preconditioners by using the results in [7, 5].

It is worth noting that Theorems 5.8 and 5.10 can give results beyond linear algebra. The following corollary suggests a proof for the A-stability of formulas (2.9) using different tools, shorter than in [3, pp. 50–65].

COROLLARY 5.12. *The formulas* (2.9)*, used in boundary value form with* ν *initial and* $k-\nu$ *final conditions, are A-stable.*

Proof. As observed in [9], a linear multistep formula used in boundary value form is A-stable if its boundary locus is in the right half plane. In fact, the expression of the real part of the boundary locus of the formulas (2.9) is given by

$$\sum_{j=-\nu}^{k-\nu} \alpha_{j+\nu} \cos(j\,x), \qquad x \in \mathbb{R},\ k \geq 1;$$

see (5.6). Thus, by using the bound (5.36), we have that the boundary locus of formulas (2.9) is in the right half plane. ∎

Notice that the condition number of the underlying P-circulant approximations has a favorable behavior, e.g., for the methods based on formulas (2.9) and (2.10).

COROLLARY 5.13. *Consider the sequences* $\{K_2(p(A_{s+1}))\}$, $\{K_2(p(B_{s+1}))\}$. *We have that*

$$K_2(p(A_{s+1})) < (s+1)\sqrt{\pi^2+1}, \qquad K_2(p(B_{s+1})) = 1 \qquad \textit{for } (2.9),$$

and

$$K_2(p(A_{s+1})) < 2(s+1), \qquad K_2(p(B_{s+1})) < (s+1)\frac{\pi^2}{2} \qquad \textit{for } (2.10),$$

where $K_2(\cdot)$ *is the* 2*-norm condition number.*

Proof. The proof follows from Theorems 5.8 and 5.9 by considering that circulant matrices are normal (see [13]), and thus the singular values are given by the modulus of the eigenvalues. ∎

We observe that the bounds in the Corollary 5.13 could be not very tight for all values of k and s. However, for our purposes, it is enough to stress the linear dependence of the condition number from the size of the underlying matrices. Again, recall that $K_2(A_{s+1}) = O(s)$ and $K_2(B_{s+1}) = O(s)$ as well; see [6].

On the other hand, we cannot give an upper bound for $K_2(p(A_{s+1}))$ for the matrices related to the formulas in (2.11). Indeed, applying arguments similar to those used in the proofs of Theorem 5.8 and of Corollary 5.13, we would have $\hat{\Phi}(n\,\pi) = 0$, $n \neq 0$ integer (see Figure 5.1, left). Therefore, $K_2(p(A_{s+1}))$ cannot be bounded, and we have not considered in detail formulas (2.11). Moreover, we experienced that the methods based on formulas (2.9) and (2.10) can perform better than those based on (2.11) with the underlying preconditioners. For example, less preconditioned iterations are often required to solve the linear systems (2.6) for (2.9) and (2.10).

Notice that, for several families of non-Hermitian Toeplitz matrices, the real parts of the eigenvalues of their circulant approximations can be positive, negative, or zero even when the nonpreconditioned matrix is positive stable. For example, this is the case of backward differentiation formulas, Adams–Moulton methods, formulas in section 2.2 for choices of ν different from those suggested there. Moreover, for non-Hermitian matrices, the circulant approximation in (4.2) may give ill conditioned preconditioners as well. For example, the condition number of $c(A_{s+1})$ can grow fast with k, e.g., for the families of k-step formulas in section 2.2; see Figure 5.3. Moreover, the block preconditioners using the approximations (4.2) can be singular

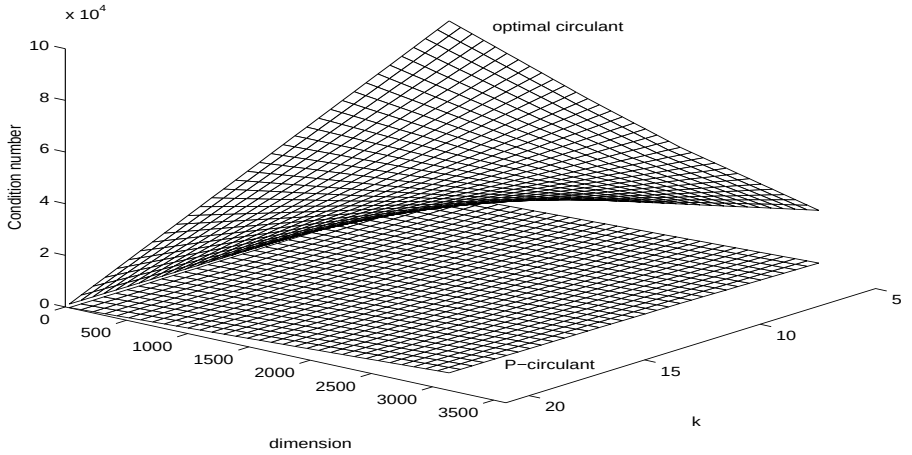

FIG. 5.3. *Condition number of the P-circulant and of the circulant approximation based on* (4.2) *for the matrices A as in* (2.7) *related to k-step formulas* (2.9).

for stable multistep formulas in boundary value form whose component matrices are nonsingular. This is the case of the midpoint method using with one initial and one final condition ($k = 2$, $\nu = 1$ in (2.3)) introduced in [2]. Indeed, by using (5.7), and by recalling that $\alpha_2 = 1 = -\alpha_0$ in (2.3), we have the expression of the eigenvalues of $c(A_{s+1})$ for the above-mentioned method:

$$\left(1 - \frac{1}{s+1}\right)(\epsilon^l - \epsilon^{-l}) = 2\mathbf{i}\left(1 - \frac{1}{s+1}\right)\sin\left(\frac{2\pi\mathbf{i}l}{s+1}\right), \qquad l = 0, \ldots, s,$$

which is zero for $l = 0$ for any $s \geq 2$ and, if s is odd, for $l = (s+1)/2$ as well. On the other hand, by (5.2), the eigenvalues of the corresponding P-circulant matrix $p(A_{s+1})$ are given by

$$\frac{1}{s+1}\cos\left(\frac{2\pi l}{s+1}\right) + \mathbf{i}\sin\left(\frac{2\pi l}{s+1}\right), \qquad l = 0, \ldots, s,$$

which cannot be zero. However, we have experienced that, for low order formulas in section 2.2, both the preconditioners based on P-circulants and on T. Chan's circulants (4.2) can be effective to solve (2.6) with Krylov subspace accelerators; see [4, 5].

5.3. How to choose the approximations and convergence of preconditioned iterations. In the previous sections, we have considered the spectrum of the component matrices of the block preconditioners in (1.2). Further information on the spectrum of the matrix M in (2.6) and its component matrices can be found in [2, 6, 9]. On the other hand, the convergence of preconditioned iterations using, e.g., GMRES, BiCGstab, and some other BiCG-like methods, is essentially decided by the distribution of the spectrum of the eigenvalues and by the eigenvectors of the preconditioned matrix; see, e.g., [18]. Notice that the analysis of the preconditioned linear system, in the nonsymmetric case, cannot be performed by using the arguments in the previous sections. In fact, one should explicitly manipulate the related characteristic polynomial. For example, if we consider the preconditioner (1.2) and the linear

systems (2.6), we need to derive an analytic expression for λ from

$$\det((1-\lambda)I_{m(s+1)} + P^{-1}E) = 0, \qquad E = (A - \breve{A}) \otimes I_m - h(B - \breve{B}) \otimes J, \tag{5.40}$$

where we recall that $I_{m(s+1)} + P^{-1}E = P^{-1}M$; see [4, Theorem 4.1] for details. Unfortunately, the above approach can fail to give complete and useful information on the convergence process. Indeed, the above analysis must consider a specific Jacobian matrix J and a formula (2.3) with k fixed. Moreover, the derivation of λ from (5.40) is usually rather lengthy even for low order schemes and sometimes it is difficult to handle in view of the behavior of the eigenvalues.

Therefore, in what follows, we will give some general suggestions in order to decide whether approximation could be more suitable to precondition the underlying linear system. These hints could be adapted for the solution of other problems based on nonsymmetric (block-)Toeplitz-like matrices.

Recall that, for Krylov subspace methods, we expect fast convergence of preconditioned iterations if the spectrum of the eigenvalues of the block preconditioned matrix is clustered around $(1, 0) \in \mathbb{C}$. Let T_n be a nonsymmetric band Toeplitz matrix, $p(T_n)$ the P-circulant approximation for T_n, and $l(T_n)$ a trigonometric approximation, e.g., one of those described in section 4. Defining $E_p = T_n - p(T_n)$, $E_l = T_n - l(T_n)$, and using similar arguments as in [4], we can write

$$p(T_n)^{-1}T_n = I + p(T_n)^{-1}E_p = I + p(T_n)^{-1}(E_p^{(1)} + E_p^{(2)}), \tag{5.41}$$

$$l(T_n)^{-1}T_n = I + l(T_n)^{-1}E_l = I + l(T_n)^{-1}(E_l^{(1)} + E_l^{(2)}), \tag{5.42}$$

where $E_p^{(2)}$, $E_l^{(2)}$ have small rank with respect to n and $E_p^{(1)}$, $E_l^{(1)}$ have small norm (with respect to T_n, say). From (5.41), (5.42), we expect that the P-circulant-based approximation will perform better than the other if, e.g.,

(C1) $||p(T_n)^{-1}||_2 < ||l(T_n)^{-1}||_2$;

(C2) $||p(T_n)^{-1}E_p^{(1)}||_2 < ||l(T_n)^{-1}E_l^{(1)}||_2$ (if the underlying approximation $l(T_n)$ is such that $||E_l^{(1)}|| \neq 0$; otherwise we require that $||p(T_n)^{-1}E_p^{(1)}||_2$ is moderate);

(C3) the outlying eigenvalues of $p(T_n)^{-1}T_n$ (i.e., the eigenvalues outside the cluster in $(1, 0) \in \mathbb{C}$) have positive real part whereas some of $l(T_n)^{-1}T_n$ have negative real part.

Notice that condition (C1) is equivalent to, say, that of $K_2(p(T_n)) < K_2(l(T_n))$ because $||p(T_n)||_2$, $||l(T_n)||_2$ are uniformly bounded with n. (We assume, as is customary, that the entries of T_n are uniformly bounded with respect to n.) By condition (C2) alone and (5.41), (5.42), it would appear that preconditioners based on simple circulant-like approximations such as Strang's, MS-circulant, and $\{\omega\}$-circulant will perform definitively better than a P-circulant based one (or, e.g., better than (4.2)) because they have $E_l^{(1)} = 0$ in (5.42), i.e., no small norm perturbation. Unfortunately, this is false in general. Finally, the third condition (C3) can be very important for the convergence of GMRES and BiCG-like Krylov methods. Indeed, as observed in [18], if the convex hull of the eigenvalues includes the origin of the complex plane, then the convergence can be slow.

Let us consider some examples in which P-circulant-like block preconditioners in (1.2) can outperform preconditioners based on other approximations for the linear systems (2.6). For simplicity, we assume $J = VDV^{-1}$ diagonalizable, $D =$

$diag(\mu_1, \dots, \mu_m)$, and $\mathrm{Re}(\mu_r) \leq 0$. By using the notation of the previous sections, we have the following decomposition for P as in (1.2):

$$P = (F^* \otimes V)\, \mathrm{diag}\,(\phi_0 - h\psi_0\mu_1, \dots, \phi_0 - h\psi_0\mu_m, \dots \\ \dots, \phi_s - h\psi_s\mu_1, \dots, \phi_s - h\psi_s\mu_m)\,(F \otimes V^{-1}). \tag{5.43}$$

Then, the eigenvalues of the block preconditioner are given by $\phi_j - h\psi_j\mu_r$, $j = 0, \dots, s$, $r = 1, \dots, m$, and

$$||P^{-1}||_2 \leq K_2(V) \min_{j,r}\{|\phi_j - h\psi_j\mu_r|\}^{-1},$$

where $K_2(V)$ does not depend on s. If we consider the matrices related to the schemes (2.9), we have $\psi_j \equiv 1$, $j = 0, \dots, s$, and using P-circulant approximations for $\breve{A}$, $\breve{B}$ in (1.2) gives

$$||P^{-1}||_2 \leq K_2(V) \frac{s+1}{1 + (T - t_0)\tilde{\mu}} = O(s), \qquad \tilde{\mu} = \min_r\{|\mu_r|\}.$$

On the other hand, similar bounds cannot be stated for non-P-circulant-like approximations because, in general, we have

$$||P^{-1}||_2 \leq K_2(V) \frac{s+1}{(T - t_0)\tilde{\mu}}, \qquad \tilde{\mu} = \min_r\{|\mu_r|\},$$

which can be unbounded if some eigenvalues of J are very small or zero in modulus. A similar effect can be observed for some classes of matrices J with purely imaginary eigenvalues and other matrices A, B in (2.6), T_n in (5.41), (5.42) as well.

Notice that, by using similar arguments as before, we can write $P^{-1}M = I + P^{-1}(E^{(1)} + E^{(2)})$; see, [4, Theorem 4.1]. Therefore, if we take the 2-norm of the perturbation of the identity in the right-hand side above, we get

$$||P^{-1}(E^{(1)} + E^{(2)})|| \leq ||P^{-1}|| \cdot (||E^{(1)}|| + ||E^{(2)}||).$$

By the above arguments, $||P^{-1}||$ can be larger for the preconditioner not based on P-circulant matrices. As a result, the amplification of the perturbations $E^{(1)} + E^{(2)}$ given by the multiplication by P^{-1} can give (C3); see (5.41), (5.42). Moreover, recall that the spectrum of the eigenvalues can be much more sensitive to perturbations with respect to the Hermitian case; see, e.g., [22].

On the other hand, if the eigenvalues μ_r, $r = 1, \dots, m$, are, e.g., negative and bounded from below by a constant $c < 0$, then preconditioners based on simple circulant-like approximations (i.e., based on Strang's, $\{\omega\}$-circulant, and MS-circulant matrices) may give better performances for large s as well. For numerical examples, see [4, 5, 7].

Acknowledgments. The author would like to thank three anonymous referees, Lionello Pasquini, and the editor for helpful comments and useful suggestions which have improved this presentation. This work is dedicated to my wife Vittoria.

REFERENCES

[1] E. Artin, *The Gamma Function*, Holt, Rinehart and Winston, New York, 1964.
[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] L. Aceto, *On the Stability Problem Arising in Numerical Methods for ODEs*, Ph.D thesis, Università di Genova, Italy, 1999.
[4] D. Bertaccini, *A circulant preconditioner for the systems of LMF-based ODE codes*, SIAM J. Sci. Comput., 22 (2000), pp. 767–786.
[5] D. Bertaccini, *Reliable preconditioned iterative linear solvers for some numerical integrators*, Numer. Linear Algebra Appl., 8 (2001), pp. 111–125.
[6] D. Bertaccini and M. K. Ng, *The convergence rate of block preconditioned systems arising from LMF-based ODE codes*, BIT, 41 (2001), pp. 433–450.
[7] D. Bertaccini and M. K. Ng, *Skew-circulant preconditioners for systems of LMF-based ODE codes,* Lecture Notes in Comp. Sci. 1988, Springer-Verlag, Berlin, 2001, pp. 93–101.
[8] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, New York, 1998.
[9] L. Brugnano and D. Trigiante, *Solving ODE by Linear Multistep Methods: Initial and Boundary Value Methods*, Gordon & Breach, Reading, UK, 1998.
[10] R. H. Chan and G. Strang, *Toeplitz equations by conjugate gradients with circulant preconditioner*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 104–119.
[11] R. H. Chan and M. K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.
[12] T. F. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 766–771.
[13] P. J. Davis, *Circulant Matrices*, John Wiley, New York, 1979.
[14] T. A. Driscoll, K. C. Toh L. N. Trefethen, *From potential theory to matrix iterations in six steps*, SIAM Rev., 40 (1998), pp. 547–578.
[15] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II. *Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin, 1991.
[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994.
[17] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley, New York, 1974.
[18] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Front. Appl. Math. 17, SIAM, Philadelphia, 1997.
[19] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, John Wiley, New York, 1991.
[20] S. Serra Capizzano, *Toeplitz preconditioners constructed from linear approximation processes*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 446–465.
[21] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
[22] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, CA, 1990.
[23] E. E. Tyrtyshnikov, *Optimal and superoptimal circulant preconditioners*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 459–473.
[24] E. E. Tyrtyshnikov, *Circulant preconditioners with unbounded inverses*, Linear Algebra Appl., 216 (1995), pp. 1–23.
[25] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge, UK, 1959.