

## Efficient Algorithms for the Evaluation of the Eigenvalues of (Block) Banded Toeplitz Matrices\*

By D. Bini and V. Pan

**Abstract\*\*.** Let  $A$  be an  $n \times n$  banded block Toeplitz matrix of bandwidth  $k$  with  $m \times m$  blocks having entries in a field  $\mathbf{F}$ . We present algorithms for computing  $p(\lambda) = \det(A - \lambda I)$  as well as the ratio  $p(\lambda)/p'(\lambda)$ , where  $p'(\lambda)$  is the first derivative of  $p(\lambda)$  with respect to  $\lambda$ , in roughly  $(3/2)k^2 \log n + O(k^3)$  block multiplications. If the field  $\mathbf{F}$  supports FFT, then the cost is reduced to  $O((m^2 k \log k + m^3 k) \log n + k^3 m^3)$  scalar multiplications. The algorithms generalize an algorithm given by W. Trench for computing  $p(\lambda)$  in the case  $m = 1$  in roughly  $k \log n + O(k^3)$  multiplications and rely on powering a companion matrix associated with the linear recurrence relation representing the original problem.

**1. Introduction.** Toeplitz matrices, either with scalar entries or with blocks, arise in a variety of problems of Applied Mathematics; a wide literature deals with the problem of their inversion. Recently, some interest has been focused on the eigenvalue problem for banded Toeplitz matrices. In [1] and [2], spectral properties are expressed in terms of a suitable polynomial associated with a banded Toeplitz matrix  $A$ , and parallel algorithms have been devised for the computation of the characteristic polynomial  $p(\lambda) = \det(A - \lambda I)$  at a point  $\lambda$  as well as of the ratio  $p(\lambda)/p'(\lambda)$ , where  $p'(\lambda)$  is the first derivative of  $p(\lambda)$ .

More recently, an algorithm for the computation of  $p(\lambda)$  has been proposed by Trench [8]; its cost is roughly  $k \log_2 n + O(k^3)$  multiplications, where  $n$  is the dimension of the matrix and  $k$  is its bandwidth. That algorithm works over any algebraically closed field (even though it was originally presented over the complex field), and it is assumed that the zeros of a suitable  $k$ th degree polynomial, associated with the banded Toeplitz matrix, have been precomputed together with their multiplicities.

In this paper we consider the eigenvalue problem for banded block Toeplitz matrices and present two new algorithms for the computation of  $p(\lambda)$ . The first algorithm, which is obtained dealing with the problem in terms of a linear difference ma-

---

Received October 17, 1986; revised August 28, 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65F15.

*Key words and phrases.* Banded Toeplitz matrices, block matrices, eigenvalues, computational complexity, matrix difference equation, cyclic reduction.

\*The first author was supported by NSF Grant DCR 8507573 and by M.P.I. 40% funds, and the second author was supported by NSF Grant DCR 8507573.

\*\*The abstract appeared in *Abstracts of papers presented to the Amer. Math. Soc.*, v. 8, no. 1, 1987, p. 145 (87T-68-14), and in *Abstracts of ICIAM'87*, INRIA, France 1987. The main results of this paper have been presented at the Oberwolfach conference on Complexity, November 1986.





Moreover,  $\lambda$  is an eigenvalue of  $A$  if and only if not all of the  $x_i$  are zero, or, equivalently, if and only if not all of the  $\alpha_j$  are zero, since the row vector formed by  $x_i$  equals the product of the row vector  $\alpha_j$  times a Vandermonde matrix. Imposing the boundary conditions (2.3) yields

$$\begin{bmatrix} 1 & \dots & 1 \\ z_1 & \dots & z_k \\ \vdots & & \vdots \\ z_1^{r-1} & \dots & z_k^{r-1} \\ z_1^{n+r} & \dots & z_k^{n+r} \\ \vdots & & \vdots \\ z_1^{n+k-1} & \dots & z_k^{n+k-1} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \mathbf{0},$$

that is,  $\lambda$  is an eigenvalue of  $A$  if and only if  $\det V_n = 0$ , where  $V_n$  is the  $k \times k$  matrix of this last system. Trench shows that

$$\det(A - \lambda I) = (-1)^{ns} a_s^n \det V_n / \det V_0,$$

where  $V_0 = (z_j^{i-1})$  is the  $k \times k$  Vandermonde matrix associated with the zeros  $z_j$ ; he also extends this last equation to the case where the condition  $z_i \neq z_j$ ,  $i \neq j$ , is not satisfied. Observe that if the zeros  $z_j$  are known, or if they can be easily approximated, then it is possible to compute  $p(\lambda)$  in roughly  $k \log_2 n + O(k^3)$  multiplications. In this way, to approximate the eigenvalues of  $A$  as the zeros of  $p(\lambda)$ , it is possible to apply the secant method or any rootfinding method to the polynomial equation  $p(\lambda) = 0$ , which uses only the values of  $p(\lambda)$  at  $\lambda$ .

Now we present a different approach, which yields an algorithm for the evaluation of  $p(\lambda)$  over an arbitrary field of constants  $\mathbf{F}$ . That algorithm does not require to precompute approximations to the zeros of an auxiliary polynomial; furthermore, it enables us to extend the result to the case of block Toeplitz matrices and to compute the ratio  $p(\lambda)/p'(\lambda)$ , which in turn allows us to apply Newton's method to the polynomial equation  $p(\lambda) = 0$  in order to evaluate the eigenvalues of the matrix  $A$ .

First we observe that the matrix vector equation (2.1) is equivalent to the linear recurrence (2.2) with the boundary condition (2.3) and can be rewritten as follows:

$$(2.5) \quad \begin{bmatrix} x_i \\ \vdots \\ x_{i+k-1} \end{bmatrix} = F^{r+i} \begin{bmatrix} x_{-r} \\ \vdots \\ x_{s-1} \end{bmatrix}, \quad i = -r, -r+1, \dots, n-r.$$

Here,  $F$  is the  $k \times k$  companion (Frobenius) matrix,

$$(2.6) \quad F = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ -\gamma_0 & -\gamma_1 & \cdots & & -\gamma_{k-1} \end{bmatrix},$$

$$\gamma_j = a_s^{-1} a_{j-r}, \quad j = 0, \dots, k-1, \quad j \neq r, \quad \gamma_r = a_s^{-1} (a_0 - \lambda).$$

Setting  $\mathbf{y} = (x_{-r}, x_{-r+1}, \dots, x_{s-1})^T \in \mathbf{F}^k$ , we can rewrite (2.3) as follows,

$$\begin{aligned} [I_r \mid 0] \mathbf{y} &= \mathbf{0} && \text{(initial conditions),} \\ [0 \mid I_s] F^n \mathbf{y} &= \mathbf{0} && \text{(terminal conditions),} \end{aligned}$$

where  $F$  is defined by (2.6). Here and hereafter,  $I_m$  denotes the  $m \times m$  identity matrix,  $\mathbf{v}^T$  denotes the transpose of  $\mathbf{v}$ . The second equation follows from (2.3) and (2.5) where  $i = n - r$ . Setting

$$(2.7) \quad [V \mid U] = [0 \mid I_s] F^n,$$

we deduce that (2.1) is equivalent to

$$\begin{bmatrix} I_r & 0 \\ V & U \end{bmatrix} \mathbf{y} = \mathbf{0}, \quad \mathbf{y} \neq \mathbf{0},$$

that is, to the equation  $\det U = 0$ , so the following relation holds:

$$p(\lambda) = 0 \quad \text{if and only if} \quad \det U = 0.$$

From this, we deduce the following result.

**PROPOSITION 2.1.** *Let  $U$  be the right lower  $s \times s$  submatrix of  $F^n$ ; then for the characteristic polynomial  $p(\lambda)$  of  $A$  we have*

$$(2.8) \quad p(\lambda) = (-1)^{ns} a_s^n \det U.$$

**PROOF.** First observe that  $(-1)^n p(\lambda)$  is a monic polynomial of degree  $n$ ; we show that  $(-1)^{n(s+1)} a_s^n \det U$  is also a monic polynomial of degree  $n$ , that is, the term of  $\det U$  having the highest degree in  $\lambda$  is  $(-1)^{n(s+1)} \lambda^n a_s^{-n}$ . For this purpose, let  $U_q$ ,  $q = 1, 2, \dots, n$ , denote the  $s \times s$  right lower submatrix of  $F^q$  in which the entries not contributing to the term of the highest degree in  $\lambda$  are set to zero. Then, if  $g$  and  $e$  are natural numbers such that  $q = gs + e$ ,  $e < s$ , we have that

$$U_q = \begin{bmatrix} 0 & (\lambda a_s^{-1})^g I_{s-e} \\ (\lambda a_s^{-1})^{g+1} I_e & 0 \end{bmatrix}.$$

The last equation is trivial for  $q = 1$  and is easily verified by induction on  $q$ , using the relation

$$F^{q+1} = FF^q = HF^q + \mathbf{i}_k \mathbf{g}^T F^q,$$

where  $H = (h_{i,j})$ ,  $h_{i,i+1} = 1$ ,  $h_{i,j} = 0$  otherwise,  $\mathbf{i}_k$  is the last column of  $I_k$  and  $-\mathbf{g}^T = (\gamma_0, \dots, \gamma_{k-1})$ . (Indeed,  $HF^q$  is obtained by shifting upwards the rows of  $F^q$ , and the highest-degree entry of the matrix  $\mathbf{i}_k \mathbf{g}^T F^q$  is  $(\lambda a_s^{-1})^{g+1}$ , in the last

row and in the  $(r + e + 1)$ st column. Therefore, the term of highest degree in  $\det U$  is given by

$$\det U_n = (-1)^{n(s+1)} \lambda^n a_s^{-n}.$$

Now, (2.8) follows if  $p(\lambda)$  has pairwise distinct zeros, since

$$(-1)^n p(\lambda) \quad \text{and} \quad (-1)^{n(s+1)} a_s^n \det U$$

are both monic polynomials and have  $n$  distinct common zeros. The result is immediately extended via a continuity argument to the case where  $p(\lambda)$  has multiple zeros.  $\square$

We will also need the following

**PROPOSITION 2.2.** *If  $\psi(z) = z^n \bmod \varphi(z)$ , then  $\psi(F) = F^n$ .*

Proposition 2.2 immediately follows from the well-known fact [6, p. 325] that the polynomial (2.4) is the minimal polynomial of the Frobenius matrix  $F$ , that is,  $\varphi(F) = 0$ .

Now we are ready to present the following algorithm for the evaluation of  $p(\lambda)$ , which, for simplicity, will be described under the assumption that  $n = 2^h$ ,  $h$  integer.

*Stage 1.* Compute the coefficients of the polynomial

$$\psi(z) = \sum_{i=0}^{k-1} \alpha_i z^i = z^n \bmod \varphi(z),$$

by using the repeated squaring technique:

$$(2.9) \quad \psi_0 = z, \quad \psi_{i+1} = \psi_i^2 \bmod \varphi(z), \quad i = 0, \dots, h-1, \quad h = \log_2 n, \quad \psi_h = \psi.$$

*Stage 2.* Compute the matrix

$$U = [0 \mid I_s] F^n \begin{bmatrix} 0 \\ I_s \end{bmatrix} = [0 \mid I_s] \psi(F) \begin{bmatrix} 0 \\ I_s \end{bmatrix}$$

by means of Horner's rule,

$$U = \alpha_0 L + (\alpha_1 L + (\dots + (\alpha_{k-2} L + \alpha_{k-1} L F) F) \dots) F \begin{bmatrix} 0 \\ I_s \end{bmatrix},$$

where  $L = [0 \mid I_s]$  is a  $s \times k$  matrix, according to the following steps: successively compute

$$(2.10) \quad \begin{aligned} Y_1 &= \alpha_{k-1} L, \\ Y_i &= (\alpha_{k-i} L + Y_{i-1} F), \quad i = 2, \dots, k, \\ U &= Y_k \begin{bmatrix} 0 \\ I_s \end{bmatrix}. \end{aligned}$$

*Stage 3.* Compute  $(-1)^{ns} a_s^n \det U$ .

If we compute over the field of complex numbers or over any other field that supports FFT, we may use fast polynomial arithmetic in order to multiply and divide a pair of  $k$ th degree polynomials in  $O(k \log k)$  arithmetic operations (see [3]); in that case, the cost of the algorithm is  $O(k \log k \log(n/k) + k^3)$  multiplications. On the other hand, we may apply customary polynomial arithmetic with the cost of roughly  $(3/2)k^2 \log(n/k) + O(k^3)$  multiplications.

For comparison, the method given in [8] evaluates  $p(\lambda)$  as follows.

*Stage 1.* Compute all the zeros  $z_1, \dots, z_k$  of  $\varphi(z)$  together with their multiplicities.

*Stage 2.* Compute a  $k \times k$  Vandermonde-like matrix  $V_n$  depending on the multiplicities of the zeros. If  $z_i \neq z_j$  for  $i \neq j$ , this matrix has elements  $z_j^i$  for  $i = 0, \dots, r - 1$  and  $z_j^{n+i}$  for  $i = r, \dots, r + s - 1$ .

*Stage 3.* Compute  $(-1)^{ns} a_s^n \det V_n / \det V_0$ .

The cost of this method is given by  $k \log_2 n + O(k^3)$  multiplications, not counting the cost of Stage 1, which can be rather large, particularly in the case of clustered zeros (estimated in [5]). This is less than the cost of our algorithm by a factor of  $\log k$ , if we compute over fields supporting FFT, and by a factor of  $k$  otherwise.

The new approach yields a “rational” algorithm, that is, an algorithm which does not compute any approximations to polynomial zeros and thus, in exact arithmetic, delivers the exact result in a finite number of arithmetical operations. Moreover, this method can be slightly modified in order to deal with block matrices and to compute the ratio  $p(\lambda)/p'(\lambda)$ , where  $p'(\lambda)$  is the first derivative of  $p(\lambda)$  with respect to  $\lambda$ . These two generalizations are the subject of the next three sections.

**3. Evaluation of the Ratio  $p(\lambda)/p'(\lambda)$ .** In order to evaluate the eigenvalues of  $A$ , applying Newton’s method to the polynomial equation  $p(\lambda) = 0$ , we must compute the quotient  $p(\lambda)/p'(\lambda)$ . From (2.5) and (2.8) we deduce the equation  $p'(\lambda)/p(\lambda) = (\det U)' / (\det U)$ . Recall that

$$(\det U)' = \sum_{i,j} \frac{\partial \det U}{\partial u_{ij}} \frac{du_{ij}}{d\lambda} = \text{trace}((\text{Adj } U)U'),$$

where  $U' = (u'_{ij})$ ,  $u'_{ij} = du_{ij}/d\lambda$ , and  $\text{Adj } U = U^{-1} \det U$  is the adjugate matrix of  $U$ . Consequently,

$$(3.1) \quad (\det U) / (\det U)' = 1 / \text{trace}(U^{-1}U').$$

Thus, we essentially reduced the problem to computing  $U^{-1}$  and  $U'$ , the  $s \times s$  lower right submatrices of  $F^n$  and  $(F^n)'$ . Since  $F^n = \psi(F)$ , where

$$\psi(z) = \sum_{i=0}^{k-1} \alpha_i z^i = z^n \text{ mod } \varphi(z),$$

and  $\varphi(z) = \varphi(z, \lambda)$  is defined by (2.4), we have the relation

$$(3.2) \quad (F^n)' = \sum_{i=0}^{k-1} \alpha'_i F^i + \sum_{i=0}^{k-1} \alpha_i (F^i)',$$

so that

$$(3.3) \quad U' = G + H, \quad G = L \sum_{i=0}^{k-1} \alpha'_i F^i \begin{bmatrix} 0 \\ I_s \end{bmatrix}, \quad H = L \sum_{i=0}^{k-1} \alpha_i (F^i)' \begin{bmatrix} 0 \\ I_s \end{bmatrix},$$

where  $L = [0 \mid I_s]$ .

Next we will show how to compute simultaneously  $H$  and  $U$ . Observe that

$$\begin{aligned} \sum_{i=0}^{k-1} \alpha_i (F^i)' &= \sum_{i=1}^{k-1} \alpha_i (F^{i-1} F)' = \sum_{i=1}^{k-1} \alpha_i (F^{i-1} F' + (F^{i-1})' F) \\ &= \sum_{i=1}^{k-1} \alpha_i F^{i-1} F' + \sum_{i=1}^{k-1} \alpha_i (F^{i-1})' F = \sum_{i=1}^{k-1} \alpha_i F^{i-1} F' + \sum_{i=0}^{k-2} \alpha_{i+1} (F^i)' F. \end{aligned}$$

Setting  $Y_i = L \sum_{j=1}^i \alpha_{k-i+j-1} F^{j-1}$  and  $Z_i = L \sum_{j=0}^{i-1} \alpha_{k-i+j} (F^j)'$ , we get the relations

$$(3.4) \quad \begin{aligned} Y_1 &= L\alpha_{k-1}, & Z_1 &= L\alpha_{k-1}, \\ Y_i &= Y_{i-1}F + L\alpha_{k-i}, & Z_i &= Y_{i-1}F' + Z_{i-1}F, & i &= 2, 3, \dots, k, \\ H &= Z_k \begin{bmatrix} 0 \\ I_s \end{bmatrix}, & U &= Y_k \begin{bmatrix} 0 \\ I_s \end{bmatrix}. \end{aligned}$$

It remains to compute the first term in the right-hand side of (3.2). Rewrite (2.9) as follows,

$$\begin{aligned} \psi_0 &= z, \\ \psi_i^2 &= Q_i \varphi + \psi_{i+1}, & i &= 0, \dots, h-1, & h &= \log_2 n, \\ \psi_h &= \psi(z), \end{aligned}$$

where  $\varphi = \varphi(z)$ . Taking the derivatives with respect to  $\lambda$ , we obtain that

$$\begin{aligned} \psi'_0 &= 0, \\ 2\psi_i \psi'_i &= Q'_i \varphi + Q_i \varphi' + \psi'_{i+1}, & i &= 0, \dots, h-1, \\ \psi'_h &= \psi'. \end{aligned}$$

Since  $\varphi' = -z^r$ , we have that  $2\psi_i \psi'_i + Q_i z^r = Q'_i \varphi + \psi'_{i+1}$  (that is,  $\psi'_{i+1} = 2\psi_i \psi'_i + Q_i z^r \pmod{\varphi}$ ). Therefore, the derivative with respect to  $\lambda$  of the polynomial  $\psi(z)$  can be evaluated together with  $\psi(z)$  using the following equations:

$$(3.5) \quad \left. \begin{aligned} \psi_0 &= z, & \psi'_0 &= 0, \\ \psi_i^2 &= Q_i \varphi + \psi_{i+1} \\ 2\psi_i \psi'_i + z^r Q_i &= Q'_i \varphi + \psi'_{i+1} \end{aligned} \right\}, \quad i = 0, 2, \dots, h-1.$$

The entire algorithm consists of the following stages.

*Stage 1.* Compute the coefficients of the polynomial  $\psi(z)$  and of its derivative with respect to  $\lambda$  by means of (3.5).

*Stage 2.* Compute the matrices  $U$  and  $H$  by means of (3.4).

*Stage 3.* Compute the matrix  $G$ , the  $s \times s$  right lower submatrix of  $\psi'(F)$ , by using Horner's rule.

*Stage 4.* Compute the  $s \times s$  right lower submatrix  $U'$  of  $(F^n)'$  by means of (3.3).

*Stage 5.* Compute  $\text{trace}(U^{-1}U')$ .

In view of the structures of the matrices involved, the computational cost of the above algorithm for the evaluation of  $p(\lambda)/p'(\lambda)$  is roughly twice the cost of the algorithm for the evaluation of  $p(\lambda)$ , shown in Section 2.



$Q(z)$  is said to be the *right quotient* and  $R(z)$  is said to be the *right remainder* of the division of  $P(z)$  by  $D(z)$ .

We observe that the evaluation of the right quotient and of the right remainder of the division of two matrix polynomials can be performed using known standard algorithms for conventional polynomial division (see [3]), replacing the coefficients with block matrices.

We will use the following two propositions:

**PROPOSITION 4.1.** *Let  $Q(z)$  and  $R(z)$  be the right quotient and remainder of the division of the  $m \times m$  matrix polynomials  $P(z)$  by  $D(z)$  and let  $F$  be a  $(pm) \times (pm)$  matrix such that the right value  $D(F)$  is the null matrix. Then  $P(F) = R(F)$ .*

*Proof.* We cannot simply substitute  $z = F$  into the matrix equation  $P(z) = Q(z)D(z) + R(z)$ , for the products of the coefficients of  $P(z)$ ,  $Q(z)$ ,  $D(z)$ ,  $R(z)$  with  $F$  do not commute. With a bit more caution, however, from the relation

$$P(z) = Q(z)D(z) + R(z) = \sum_{i=0}^{p-d} Q_i D(z) z^i + R(z),$$

where  $Q(z) = \sum_{i=0}^{p-d} Q_i z^i$ , we deduce that  $P(F) = \sum_{i=0}^{p-d} Q_i D(F) F^i + R(F)$ , whence  $P(F) = R(F)$ .  $\square$

**PROPOSITION 4.2.** *Let  $Q(z)$  and  $R(z)$  be the right quotient and remainder of the division of the matrix polynomials  $Iz^p$  by  $D(z)$ , respectively, that is,*

$$(4.1) \quad Iz^p = Q(z)D(z) + R(z),$$

*and let  $S(z)$ ,  $T(z)$  be the right quotient and remainder of the division of  $(R(z))^2$  by  $D(z)$ , that is,*

$$(4.2) \quad (R(z))^2 = S(z)D(z) + T(z).$$

*Then  $T(z)$  is the right remainder of the division of  $Iz^{2p}$  by  $D(z)$ , i.e., there exists a matrix polynomial  $U(z)$  such that  $Iz^{2p} = U(z)D(z) + T(z)$  (here and hereafter,  $I$  denotes the identity matrix of appropriate size).*

*Proof.* From (4.1) we obtain that

$$(R(z))^2 = (Iz^p - Q(z)D(z))^2 = Iz^{2p} + Q(z)D(z)Q(z)D(z) - 2Q(z)D(z)z^p.$$

Therefore, from (4.2) we get

$$Iz^{2p} = 2Q(z)D(z)z^p - Q(z)D(z)Q(z)D(z) + S(z)D(z) + T(z),$$

whence,

$$Iz^{2p} = (2Q(z)z^p - Q(z)D(z)Q(z) + S(z))D(z) + T(z). \quad \square$$

Similarly to the scalar case, the assumption that  $(\mathbf{x}, \lambda)$  is an eigenpair of  $A$ , that is,  $(A - \lambda I)\mathbf{x} = \mathbf{0}$ ,  $\mathbf{x} \neq \mathbf{0}$ , where  $\mathbf{x} = (\mathbf{x}_i)$ ,  $\mathbf{x}_i \in \mathbf{F}^m$ ,  $i = 1, 2, \dots, n$ , is a block vector, can be rewritten as the following matrix difference equation of order  $k = r + s$  with constant coefficients (see [4, p. 512]),

$$A_s \mathbf{x}_{i+k} + A_{s-1} \mathbf{x}_{i+k-1} + \dots + (A_0 - \lambda I) \mathbf{x}_r + \dots + A_{-r} \mathbf{x}_i = \mathbf{0},$$

$$i = -r, -r + 1, \dots, n - r - 1.$$

This equation is complemented with the following “boundary conditions”

$$(4.3) \quad \mathbf{x}_{-r}, \dots, \mathbf{x}_{-1} = \mathbf{0} \quad (\text{initial conditions}),$$

$$(4.4) \quad \mathbf{x}_n, \dots, \mathbf{x}_{n+s-1} = \mathbf{0} \quad (\text{terminal conditions}).$$

If  $A_s$  is nonsingular (the singular case is dealt with in Section 5), the above matrix difference equation is equivalent to

$$(4.5) \quad G_s \mathbf{x}_{j+k} + G_{s-1} \mathbf{x}_{j+k-1} + \dots + G_{-r} \mathbf{x}_j = \mathbf{0},$$

where  $G_i = A_s^{-1} A_{i-r}$ ,  $i = 0, \dots, k$ ,  $i \neq r$ ,  $G_r = A_s^{-1} (A_0 - \lambda I)$ , so that  $G_k = I$ . Now the matrix difference equation (4.5) with constant coefficients can be rewritten equivalently as follows,

$$\begin{bmatrix} \mathbf{x}_i \\ \vdots \\ \mathbf{x}_{i+k-1} \end{bmatrix} = F^{r+i} \begin{bmatrix} \mathbf{x}_{-r} \\ \vdots \\ \mathbf{x}_{s-1} \end{bmatrix}, \quad i = -r, -r+1, \dots$$

Here,  $F$  denotes the  $k \times k$  block companion matrix, compare (2.6), defined by

$$(4.6) \quad F = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ -G_0 & -G_1 & \dots & & -G_{k-1} \end{bmatrix}.$$

Moreover, from (4.3) and (4.4), we conclude that the equation  $\det(A - \lambda I) = 0$  is equivalent to  $\det U = 0$ , where  $[V \mid U] = [0 \mid I_{sm}] F^n$ . The same argument as in the scalar case enables us to prove that  $\det U$  is a polynomial in  $\lambda$  of degree  $nm$  whose term of the highest degree in  $\lambda$  is given by

$$\det \begin{bmatrix} 0 & I_{s-e} \otimes (\lambda A_s^{-1})^g \\ I_e \otimes (\lambda A_s^{-1})^{g+1} & 0 \end{bmatrix}.$$

Here,  $g$  and  $e$  are natural numbers such that  $n = gs + e$ ,  $e < s$ . Therefore, we get the following block version of Proposition 2.1, which has also been proved by Tismenetsky [7], using the spectral theory of matrix polynomials.

**PROPOSITION 4.3.** *Let  $U$  be the right lower  $s \times s$  block submatrix of the block Frobenius matrix  $F^n$ . Then for the characteristic polynomial  $p(\lambda)$  of  $A$  we have*

$$p(\lambda) = (-1)^{ns} (\det A_s)^n \det U.$$

Next consider the  $m \times m$  matrix polynomial  $\Phi(z) = G_0 + G_1 z + \dots + G_k z^k$  associated with the matrix linear difference equation (4.5). Similarly to the scalar case, we deduce that its right value taken on at  $F$  is the null matrix, that is,

$$\Phi(F) = \sum_{i=0}^k (I_m \otimes G_i) F^i = 0.$$

Therefore, by virtue of Propositions 4.1 and 4.2 it is possible to compute  $F^n$  by means of the repeated squaring technique,

$$\begin{aligned} \Psi_0 &= z, \\ \Psi_{i+1} &= \Psi_i \text{ mod } \Phi(z), \quad i = 0, 1, \dots, h-1, \quad h = \log n, \\ F^n &= \Psi_h(F). \end{aligned}$$



and for simplicity, assume  $\lambda = 0$ . Then the bandwidth is  $k = 2$ . Setting  $B = A_1^{-1}A_{-1}$ ,  $C = A_1^{-1}A_0$ , we have

$$\begin{aligned} \Phi(x) &= Ix^2 + Cx + B, \\ \Psi_1 &= x^2 \pmod{\Phi(x)} = -Cx - B, \\ \Psi_{i+1} &= \Psi_i^2 \pmod{\Phi(x)} = C_{i+1}x + B_{i+1}, \quad i = 1, \dots, h - 1, \quad h = \log_2 n. \end{aligned}$$

Since

$$\begin{aligned} (C_i x + B_i)^2 \pmod{\Phi(x)} &= C_i^2 x^2 + (C_i B_i + B_i C_i)x + B_i^2 \pmod{\Phi(x)} \\ &= (C_i B_i + B_i C_i - C_i^2 C)x + (B_i^2 - C_i^2 B), \end{aligned}$$

we have the relations

$$\begin{aligned} C_{i+1} &= -C_i^2 C + C_i B_i + B_i C_i, & i = 1, \dots, h - 1, \quad h = \log_2 n. \\ (4.7) \quad B_{i+1} &= B_i^2 - C_i^2 B, \\ C_1 &= -C, \quad B_1 = -B. \end{aligned}$$

For each  $i$  we can compute  $C_{i+1}$ ,  $B_{i+1}$  (given  $C$ ,  $B$ ,  $C_i$ ,  $B_i$ ) in five matrix multiplications, using (4.7) and the equation

$$C_i B_i + B_i C_i = (C_i + B_i)^2 - B_i^2 - C_i^2.$$

Now, recall that  $F^n = (I_m \otimes C_h)F + (I_m \otimes B_h)I$ , so that

$$U = -C_h C + B_h,$$

see (4.6).

This means that the overall cost of computing  $U$  is not greater than the cost of  $5 \log_2(n/2) + 1$  matrix multiplications or  $m^3(5 \log_2(n/2) + 1)$  scalar multiplications.

**5. The Degenerate Case.** When we evaluated the characteristic polynomial of the block Toeplitz banded matrix in Section 4, we have assumed that the block  $A_s$  is nonsingular. In that case, it is possible to reduce the original matrix difference equation to the form (4.5) and to solve that equation via powering the companion matrix (4.6). In this section we consider the case where  $\det A_s = 0$ .

Suppose we are given the matrix difference equation

$$(5.1) \quad A_s \mathbf{x}_{j+k} + A_{s-1} \mathbf{x}_{j+k-1} + \dots + A_{-r} \mathbf{x}_j = \mathbf{0}, \quad j = -r, \dots, n - r - 1,$$

with boundary conditions

$$(5.2) \quad \mathbf{x}_{-r}, \dots, \mathbf{x}_{-1} = \mathbf{0},$$

$$(5.3) \quad \mathbf{x}_n, \dots, \mathbf{x}_{n+s-1} = \mathbf{0},$$

where the matrix  $A_s$  is singular. We may assume, without loss of generality, that

$$A_s = \begin{bmatrix} B_s \\ 0 \end{bmatrix},$$

where  $B_s$  is a  $q \times m$  full-rank matrix,  $q < m$ ; otherwise, consider a nonsingular  $m \times m$  matrix  $M$  such that

$$MA_s = \begin{bmatrix} B_s \\ 0 \end{bmatrix}$$

and replace (5.1) with  $MA_s\mathbf{x}_{j+k} + MA_{s-1}\mathbf{x}_{j+k-1} + \dots + MA_{-r}\mathbf{x}_j = \mathbf{0}$ ,  $j = -r, \dots, n - r - 1$ .

Now partition the matrix  $A_i$  such that

$$A_i = \begin{bmatrix} B_i \\ C_i \end{bmatrix},$$

where  $B_i$  is a  $q \times m$  matrix, so that (5.1) can be split into two equations,

$$\begin{aligned} B_s\mathbf{x}_{j+k} + B_{s-1}\mathbf{x}_{j+k-1} + \dots + B_{-r}\mathbf{x}_j &= \mathbf{0}, & j &= -r, \dots, n - r - 1, \\ C_{s-1}\mathbf{x}_{j+k-1} + \dots + C_{-r}\mathbf{x}_j &= \mathbf{0}, & j &= -r, \dots, n - r - 1. \end{aligned}$$

These equations are equivalent to the following system,

$$\begin{aligned} E_s\mathbf{x}_{j+k} + E_{s-1}\mathbf{x}_{j+k-1} + \dots + E_{-r}\mathbf{x}_j &= \mathbf{0}, & j &= -r, \dots, n - r - 2, \\ C_{s-1}\mathbf{x}_{s-1} + \dots + C_{-r}\mathbf{x}_{-r} &= \mathbf{0}, \\ B_s\mathbf{x}_{n+s-1} + \dots + B_{-r}\mathbf{x}_{n-r-1} &= \mathbf{0}, \end{aligned}$$

where

$$E_i = \begin{bmatrix} B_i \\ C_{i-1} \end{bmatrix}, \quad i = -r, \dots, s, \quad C_{-r-1} = \mathbf{0}.$$

Assuming  $E_s$  nonsingular (if  $E_s$  were singular, we would apply this technique again to the above matrix difference equation), we can express  $\mathbf{x}_i$  in terms of the block companion matrix

$$F = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ -G_0 & -G_1 & \dots & & -G_{k-1} \end{bmatrix}, \quad G_i = -E_s^{-1}E_{i-r}, \quad i = 0, \dots, k - 1,$$

by means of the relations

$$(5.4) \quad \begin{bmatrix} \mathbf{x}_{j+1} \\ \vdots \\ \mathbf{x}_{j+k} \end{bmatrix} = F^{r+j+1} \begin{bmatrix} \mathbf{x}_{-r} \\ \vdots \\ \mathbf{x}_{s-1} \end{bmatrix}, \quad j = -r, -r + 1, \dots, n - r - 2,$$

$$\begin{aligned} (5.5) \quad & C_{s-1}\mathbf{x}_{s-1} + \dots + C_{-r}\mathbf{x}_{-r} = \mathbf{0}, \\ (5.6) \quad & B_s\mathbf{x}_{n+s-1} + \dots + B_{-r}\mathbf{x}_{n-r-1} = \mathbf{0}. \end{aligned}$$

(5.4) implies that

$$\begin{bmatrix} \mathbf{x}_{n-r-1} \\ \vdots \\ \mathbf{x}_{n+s-r} \end{bmatrix} = F^{n-1} \begin{bmatrix} \mathbf{x}_{-r} \\ \vdots \\ \mathbf{x}_{s-1} \end{bmatrix}.$$



In this case it is possible to apply a cyclic reduction method to compute  $\det(A - \lambda I)$ . Observe that permuting block rows and block columns of the  $n \times n$  block matrix  $A - \lambda I$  according to the permutation  $(1, 3, 5, \dots, 2^h - 1, 2, 4, 6, \dots, 2^h - 2)$ , yields the matrix

$$(6.1) \quad P(A - \lambda I)P^T = A_n^* = \left[ \begin{array}{c|c} \begin{array}{ccc} L & & \\ & L & \\ & & \ddots \\ & & & L \end{array} & \begin{array}{ccc} G & & \\ H & G & \\ & \ddots & \\ & & HG & H \\ & & & H \end{array} \\ \hline \begin{array}{ccc} HG & & \\ & HG & \\ & & \ddots \\ & & & HG \end{array} & \begin{array}{ccc} L & & \\ & L & \\ & & \ddots \\ & & & L \end{array} \end{array} \right], \quad L = E - \lambda I,$$

where  $P$  is the permutation matrix defining the above permutation.

Moreover, if  $\det L \neq 0$ , we may perform one step of block Gaussian elimination and write

$$A_n^* = \left[ \begin{array}{c|c} \begin{array}{ccc} I & & \\ & I & \\ & & \ddots \\ & & & I \end{array} & \begin{array}{c} 0 \\ \\ \\ \end{array} \\ \hline \begin{array}{ccc} H^* G^* & & \\ & H^* G^* & \\ & & \ddots \\ & & & H^* G^* \end{array} & \begin{array}{ccc} I & & \\ & I & \\ & & \ddots \\ & & & I \end{array} \end{array} \right] \left[ \begin{array}{c|c} \begin{array}{ccc} L & & \\ & L & \\ & & \ddots \\ & & & L \end{array} & \begin{array}{ccc} G & & \\ H & G & \\ & \ddots & \\ & & HG & H \\ & & & H \end{array} \\ \hline \begin{array}{c} 0 \\ \\ \\ \end{array} & \begin{array}{ccc} L_1 G_1 & & \\ H_1 L_1 G_1 & & \\ & \ddots & \\ & & H_1 L_1 \end{array} \end{array} \right]$$

where  $H^* = HL^{-1}$ ,  $G^* = GL^{-1}$ , and

$$\begin{aligned} L_1 &= L - (HL^{-1}G + GL^{-1}H), \\ G_1 &= -GL^{-1}G, \\ H_1 &= -HL^{-1}H. \end{aligned}$$

Therefore,  $\det(A_n - \lambda I) = \det A_n^* = (\det L)^{(n+1)/2} \det(A_{(n+1)/2-1} - \lambda I)$ ,

$$A_{(n+1)/2-1} = \begin{bmatrix} L_1 & G_1 & & \\ H_1 & L_1 & G_1 & 0 \\ & & & \\ & 0 & & \\ & & & H_1 L_1 \end{bmatrix}.$$

Now, if  $L_1$  is nonsingular, the same process can be applied to  $A_{(n+1)/2-1}$  and so on. We arrive at the following relations,

$$(6.2) \quad \det A_n^* = (\det L_0)^{2^{h-1}} (\det L_1)^{2^{h-2}} \dots (\det L_{h-2})^2 \det L_{h-1},$$

$$(6.3) \quad \left. \begin{aligned} L_{i+1} &= L_i - (H_i L_i^{-1} G_i + G_i L_i^{-1} H_i) \\ G_{i+1} &= -G_i L_i^{-1} G_i \\ H_{i+1} &= -H_i L_i^{-1} H_i \end{aligned} \right\}, \quad i = 0, \dots, h-2,$$

$$L_0 = L, \quad H_0 = H, \quad G_0 = G,$$

provided that all matrices  $L_i$  are nonsingular.

The cost of this process is one matrix inversion and six matrix multiplications in each step of the evaluation of  $L_i$ ,  $G_i$ ,  $H_i$ ; moreover, the cost of evaluating  $h$  determinants should be added. The overall cost is roughly

$$(22/3)m^3 \log_2(n+1)$$

scalar multiplications.

Observe that the cost of this algorithm is higher than the cost of the algorithm described in Section 5, which required roughly  $5m^3 \log_2(n/2)$  multiplications. Note, however, that the approach of this section can be applied even when  $G$  and  $H$  are singular, provided that all  $L_i$  are nonsingular.

In order to compute the ratio  $p(\lambda)/p'(\lambda)$  for performing one step of Newton's method, we deduce from (6.1) and (6.2) that

$$\begin{aligned} p'(\lambda)/p(\lambda) &= 2^{h-1}(\det L_0)' / (\det L_0) + 2^{h-2}(\det L_1)' / (\det L_1) \\ &+ \cdots + (\det L_{h-1})' / (\det L_{h-1}). \end{aligned}$$

Since  $(\det B)' / (\det B) = \text{trace}(B^{-1}B')$ , for any nonsingular matrix  $B$ , it follows that

$$p'(\lambda)/p(\lambda) = 2^{h-1} \text{trace}(L_0^{-1}L'_0) + 2^{h-2} \text{trace}(L_1^{-1}L'_1) + \cdots + \text{trace}(L_{h-1}^{-1}L'_{h-1}).$$

The evaluation of the matrices  $L'_i$  can be carried out together with the evaluation of  $L_i$  by means of the following formulae obtained by taking derivatives of (6.3) with respect to  $\lambda$ :

$$\left. \begin{aligned} L'_{i+1} &= L'_i - \left. \begin{aligned} &(H'_i L_i^{-1} G_i + H_i L_i^{-1} G'_i + H_i (L_i^{-1})' G_i \\ &+ G'_i L_i^{-1} H_i + G_i L_i^{-1} H'_i + G_i (L_i^{-1})' H_i) \end{aligned} \right\}, \quad i = 0, \dots, h-2, \\ G'_{i+1} &= -(G'_i L_i^{-1} G_i + G_i L_i^{-1} G'_i + G_i (L_i^{-1})' G_i) \\ H'_{i+1} &= -(H'_i L_i^{-1} H_i + H_i L_i^{-1} H'_i + H_i (L_i^{-1})' H_i) \end{aligned}$$

$$L'_0 = I, \quad G'_0 = 0, \quad H'_0 = 0.$$

Here  $(L_i^{-1})'$  can be computed using the formula

$$(L_i^{-1})' = -L_i^{-1} (L_i)' L_i^{-1}.$$

Dipartimento di Matematica  
II Università di Roma "Tor Vergata"  
Rome, Italy

Department of Computer Science  
SUNY at Albany  
Albany, New York 12222

1. D. BINI & M. CAPOVANI, "Fast parallel and sequential computations and spectral properties concerning band Toeplitz matrices," *Calcolo*, v. 22, 1983, pp. 176-189.

2. D. BINI & M. CAPOVANI, "Spectral and computational properties of band symmetric Toeplitz matrices," *Linear Algebra Appl.*, v. 52/53, 1983, pp. 99-126.

3. D. BINI & V. PAN, "Polynomial division and its computational complexity," *J. Complexity*, v. 2, 1986, pp. 179-203.

4. P. LANCASTER & M. TISMENETSKY, *The Theory of Matrices*, Academic Press, New York, 1985.

5. V. PAN, *On the Sequential and Parallel Approximate Evaluation of Polynomial Zeros*, Tech. Rep. 87-14, Computer Science Dept., SUNYA, Albany, N. Y., 1987, to appear in *Comput. Math. Appl.*
6. J. STOER & R. BULIRSCH, *Introduction to Numerical Analysis*, Springer-Verlag, Berlin, 1980.
7. M. TISMENETSKY, "Determinant of block-Toeplitz band matrices," *Linear Algebra Appl.*, v. 85, 1987, pp. 165-187.
8. W. F. TRENCH, "On the eigenvalue problem for Toeplitz band Matrices," *Linear Algebra Appl.*, v. 64, 1985, pp. 199-214.