SIAM J. ALG. DISC. METH.
Vol. 7, No. 2, April 1986

ON THE EIGENVALUE PROBLEM FOR A CLASS OF BAND MATRICES INCLUDING THOSE WITH TOEPLITZ INVERSES*

WILLIAM F. TRENCH†

Abstract. We study the eigenvalue problem for a class $\mathscr{H}$ of band matrices which includes as a proper subclass all band matrices with Toeplitz inverses. Toeplitz matrices of this kind occur, for example, as autocorrelation matrices of purely autoregressive stationary time series. A formula is given for the characteristic polynomial $p_n(\lambda)$ of an nth order matrix H_n in $\mathscr{H}$, with bandwidth $k+1 \leqq n$, as the ratio of $k \times k$ determinants whose entries are polynomials in the zeros of a certain kth degree polynomial which is independent of n and has one coefficient which depends upon λ. The formula permits the evaluation of $p_n(\lambda)$ by means of a computation with complexity independent of n. Also given is a formula for the eigenvectors in terms of these zeros and k coefficients which can be obtained by solving a $k \times k$ homogeneous system.

AMS(MOS) subject classification. 15A18

1. Introduction. We consider the eigenvalue problem for the class $\mathscr{H}$ of matrices

$$H_n = (h_{ijn})_{i,j=0}^{n-1}, \tag{1}$$

defined as follows. Let

$$A(z) = \sum_{\nu=0}^{r} a_\nu z^\nu, \qquad B(z) = \sum_{\mu=0}^{s} b_\mu z^\mu,$$

where

$$a_0 b_0 \neq 0 \quad \text{and} \quad r+s=k<n, \tag{2}$$

and $\{h_{ijn}\}$ are defined by the generating functions

$$H_{in}(z) = \sum_{j=0}^{n-1} h_{ijn} z^j = \begin{cases} z^i A(z) \sum_{\mu=0}^{i} b_\mu z^{-\mu}, & 0 \leqq i \leqq s-1, \\ z^i A(z) B(1/z), & s \leqq i \leqq n-r-1, \\ z^i B(1/z) \sum_{\nu=0}^{n-i-1} a_\nu z^\nu, & n-r \leqq i \leqq n-1. \end{cases}$$

Explicitly,

$$h_{ijn} = c_{j-i} - \sum_{\nu=i+1}^{s} a_{j-i+\nu} b_\nu - \sum_{\mu=n-i}^{r} b_{i-j+\mu} a_\mu, \qquad 0 \leqq i,j \leqq n-1, \tag{3}$$

if we define

$$a_l = 0 \text{ if } l>r \text{ or } l<0, \quad b_l = 0 \text{ if } l>s \text{ or } l<0, \quad \sum_{q}^{p} = 0 \text{ if } q>p, \tag{4}$$

$$c_\nu = 0 \quad \text{if } \nu > r \text{ or } \nu < -s, \tag{5}$$

and

$$C(z) = A(z)B(1/z) = \sum_{\nu=-s}^{r} c_\nu z^\nu. \tag{6}$$

* Received by the editors March 1, 1984, and in revised form December 5, 1984.
† Department of Mathematics and Computer Science, Drexel University, Philadelphia, Pennsylvania 19104.

The class $\mathscr{H}$ is connected with Toeplitz matrices; i.e., matrices of the form

$$T_n = (\phi_{j-i})_{i,j=0}^{n-1}.$$

From (3) and (4),

$$h_{ijn} = c_{j-i} \quad \text{if} \begin{cases} 0 \leqq i \leqq s-1 \text{ and } r \leqq j \leqq n-1, \\ \text{or } s \leqq i \leqq n-r-1, \\ \text{or } n-r \leqq i \leqq n-1 \text{ and } 0 \leqq j \leqq n-s-1; \end{cases} \tag{7}$$

thus, H_n is *quasi-Toeplitz* (a term used in [7]) in that h_{ijn} is a function of $j-i$ alone except in the $s \times r$ submatrix in the upper left corner of H_n and the $r \times s$ submatrix in the lower right corner. Moreover, H_n is *banded*; i.e.,

$$h_{ijn} = 0 \quad \text{if } j-i > r \text{ or } i-j > s, \tag{8}$$

from (3), (4), and (5).

Matrices in the class $\mathscr{H}$ have been encountered by the author [10] in connection with prediction of stationary time series, and by Greville [4], [5], [6], in connection with a smoothing problem. Greville and the author studied them in [7], and obtained results which can be summarized as follows.

THEOREM 1 (Greville–Trench). *The matrices H_n $(n > k)$ are invertible if and only if $A(z)$ and $z^s B(1/z)$ are relatively prime, in which case their inverses are the Toeplitz matrices*

$$H_n^{-1} = T_n = (\phi_{j-i})_{i,j=0}^{n-1}, \qquad n > k,$$

where $\{\phi_j\}$ is determined as follows: Obtain $[\phi_{s-1}, \phi_{s-2}, \cdots, \phi_{-r}]$ as the (unique) solution of the $k \times k$ system

$$\text{(9)} \quad \begin{aligned} &\text{(a)} \quad \sum_{\nu=0}^{r} a_\nu \phi_{j-\nu} = b_0^{-1} \delta_{j0}, \qquad 0 \leqq j \leqq s-1, \\ &\text{(b)} \quad \sum_{\mu=0}^{s} b_\mu \phi_{-j+\mu} = 0, \qquad 1 \leqq j \leqq r, \end{aligned}$$

and then compute

$$\phi_j = -a_0^{-1} \sum_{\nu=1}^{r} a_\nu \phi_{j-\nu}, \qquad j \geqq s, \tag{10}$$

and

$$\phi_{-j} = -b_0^{-1} \sum_{\mu=1}^{s} b_\mu \phi_{-j+\mu}, \qquad j > r. \tag{11}$$

Moreover, if H_n $(n > k)$ is a matrix of the form (1) *such that* (8) *holds and H_n^{-1} is a Toeplitz matrix, then $H_n \in \mathscr{H}$.*

Greville continued the investigation of these matrices in [2] and [3].

The main result of this paper reduces the evaluation of the characteristic polynomial $p_n(\lambda)$ of H_n to finding the zeros of the polynomial

$$P(z; \lambda) = \sum_{\mu=-s}^{r} c_\mu z^{\mu+s} - \lambda z^s \tag{12}$$

(which are obviously independent of n) and evaluating a kth order determinant whose entries are polynomials in these zeros. The complexity of this representation of $p_n(\lambda)$

depends only on k (cf. (2)), and is independent of n. Moreover, we give an explicit formula for the eigenvectors of H_n corresponding to a given eigenvalue, which depends on k coefficients that can be obtained by solving a kth order homogeneous system with complexity independent of n. The results are analogous to those obtained in [11] for Toeplitz band matrices

$$T_n = (c_{j-i})_{i,j=0}^{n-1}, \tag{13}$$

where $\{c_\nu\}$ satisfies (5) and $r+s=k<n$. However, the arguments needed here are considerably more complicated than those in [11].

Our results here are not restricted to the case where $A(z)$ and $z^sB(1/z)$ are relatively prime, so that H_n is invertible; however, this case is especially important, since Theorem 1 implies that the eigenvalue problems for invertible matrices in $\mathcal{H}$ and for Toeplitz matrices with banded inverses are equivalent. Although there is a large body of literature on inverting Toeplitz matrices and solving systems with Toeplitz matrices, little has been published on approaches to the eigenvalue problem for these matrices which take full advantage of their simple structure. (For examples, see Grunbaum [8], [9]; Dini and Capovani [1]; and Trench [11].)

2. Preliminary definitions and lemmas. We take the underlying field to be the complex numbers.

It can be seen from (7) and (8) that if r or s is zero, then H_n is a triangular Toeplitz matrix. Since the eigenvalue problem for such matrices is trivial, we assume henceforth that (2) holds, and also that

$$rsa_rb_s \neq 0. \tag{14}$$

Then $rsc_rc_{-s} \neq 0$, so $P(0;\lambda)\neq 0$.

It was shown in [11] that there are at most k values of λ for which $P(z;\lambda)$ has fewer than k distinct zeros. We call such points *critical points* of $P(z;\lambda)$. All other values of λ are *ordinary points.* For completeness, we phrase all definitions so as to include the case where λ is a critical point; however, for notational convenience we illustrate the definitions only for ordinary points.

DEFINITION 1. For a fixed λ, let $z_1,\cdots,z_q$ be the distinct zeros of (12) with multiplicities $m_1,\cdots,m_q$; thus,

$$q\leqq k,\quad m_j\geqq 1\ (1\leqq j\leqq k),\quad m_1+\cdots m_q=k.$$

If $Q_1(z),\cdots,Q_k(z)$ are given polynomials, define the k-vector function

$$w(z)=\operatorname{col}[Q_1(z),Q_2(z),\cdots,Q_k(z)],$$

and let Ω be the $k\times k$ matrix defined as follows: its first m_1 columns are $w^{(l)}(z_1)$ $(0\leqq l\leqq m_1-1)$; its next m_2 columns are $w^{(l)}(z_2)$ $(0\leqq l\leqq m_2-1)$; and so forth. Let V be the matrix resulting from this construction in the special case where $Q_i(z)=z^{i-1}$, $1\leqq i\leqq k$, and define

$$\Delta(\lambda)=\frac{\det\Omega}{\det V}. \tag{15}$$

Thus, if λ is an ordinary point of $P(z;\lambda)$, then $q=k$, $m_1=\cdots=m_k=1$,

$$\Omega=(Q_i(z_j))_{i,j=1}^{k}$$

and V is the Vandermonde matrix

$$V=(z_j^{i-1})_{i,j=1}^{k}.$$

It can be shown in general that

$$\det V = K_{ij} \prod_{1 \leqq i < j \leqq q} (z_j - z_i)^{r_{ij}} \qquad (K_{ij} = \text{constant} \neq 0),$$

where the r_{ij}'s are positive integers (all ones if $q = k$); hence, $\det V \neq 0$.

We refrain from using the functional notation $z_i(\lambda)$ for the root z_i, since this would necessitate an irrelevant appeal to the theory of multiple-valued algebraic functions. Note that there is an ambiguity in the definitions of Ω and V, since the numbering of the roots is arbitrary; however, since renumbering $z_1, \cdots, z_q$ would simply permute the columns of both matrices in the same way, the ratio of the determinants in (15) is uniquely defined for each λ.

To avoid cumbersome two-dimensional displays in proofs which follow, we also denote the function $\Delta(\lambda)$ defined by (15) in the form

$$\Delta(\lambda) = |Q_1(z), \cdots, Q_k(z)|(\lambda); \tag{16}$$

thus, if λ is an ordinary point of $P(z; \lambda)$, then

$$|Q_1(z), \cdots, Q_k(z)|(\lambda) = \frac{\det (Q_i(z_j))_{i,j=1}^k}{\det (z_j^{i-1})_{i,j=1}^k}. \tag{17}$$

There is an abuse of notation here, since (17) is not a function of z as the symbol on the left appears to indicate; however, the convenience of the notation outweighs this drawback.

In the following we adopt the convention that the polynomial $p = 0$ has degree $-\infty$.

LEMMA 1. *Let $n_1, \cdots, n_k$ be nonnegative integers, and $m = \max\{n_1, \cdots, n_k\}$. Then the function*

$$|z^{n_1}, \cdots, z^{n_k}|(\lambda) \tag{18}$$

is a polynomial of degree $\leqq m - k + 1$.

This lemma was proved in [11], where the function in (18) was denoted by $q(\lambda; n_1, \cdots, n_k)$. The main result in [11] is that the characteristic polynomial of the Toeplitz band matrix T_n in (13) is

$$\det [\lambda I_n - T_n] = (-1)^{(r-1)n} c_r^n |1, z, \cdots, z^{s-1}, z^{n+s}, \cdots, z^{n+k-1}|(\lambda).$$

We will obtain an analogous result here for H_n.

LEMMA 2. *Let $Q_1(z), \cdots, Q_k(z)$ and $\Delta(\lambda)$ be as in Definition 1, and let $m = \max_{1 \leqq i \leqq k} \{\deg Q_i(z)\}$. Then $\Delta(\lambda)$ is a polynomial of degree $\leqq m - k + 1$. Moreover, $\Delta(\lambda) = 0$ for a given complex number λ if and only if there are constants $c_1, \cdots, c_k$ (not all zero) such that the polynomial*

$$Q(z) = c_1 Q_1(z) + \cdots + c_k Q_k(z) \tag{19}$$

is divisible by $P(z; \lambda)$. In particular, if $m \leqq k - 1$, so that $\Delta(\lambda) = C$ (constant), then $C = 0$ if and only if $Q_1(z), \cdots, Q_k(z)$ are linearly dependent.

Proof. If

$$Q_i(z) = \sum_{j=0}^{m} a_{ij} z^j, \qquad 1 \leqq i \leqq k,$$

then

$$\Delta(\lambda) = \sum a_{1j_1} \cdots a_{kj_k} |z^{j_1}, \cdots, z^{j_k}|(\lambda),$$

where the sum is over all $j_1, \cdots, j_k$ such that $0 \leqq j_1, \cdots, j_k \leqq m$, so $\Delta(\lambda)$ is a polynomial of degree $\leqq m-k+1$, by Lemma 1. From (15), we see that $\Delta(\lambda)=0$ if and only if the system

$$\Omega' Y = 0 \quad ('=\text{transpose}) \tag{20}$$

has a nontrivial solution $Y=[c_1, \cdots, c_k]$. From the definition of Ω, it can be seen that (20) is equivalent to

$$Q^{(l)}(z_j)=0, \qquad 0 \leqq l \leqq m_j-1, \quad 1 \leqq j \leqq q, \tag{21}$$

with $Q(z)$ as in (19). But (21) holds if and only if $P(z;\lambda)$ divides $Q(z)$. This implies the stated conclusions.

LEMMA 3. *With the assumptions of Lemma 2, suppose that*

$$Q_i(z)=\alpha z^m+\cdots$$

for some $m \geqq k$ and i in $\{1, \cdots . k\}$, and that

$$\deg Q_j(z) < m \quad \text{if } j \neq i. \tag{22}$$

Then $\Delta(\lambda)$ as defined in (16) can be written as

$$\Delta(\lambda)=\alpha\lambda c_r^{-1}|Q_1(z), \cdots, Q_{i-1}(z), z^{m-r}, Q_{i+1}(z), \cdots, Q_k(z)|(\lambda)+O(\lambda^{m-k}), \tag{23}$$

where $O(\lambda^{m-k})$ denotes a polynomial of degree $\leqq m-k$.

Proof. From elementary properties of determinants,

$$\Delta(\lambda)=|\cdots, \alpha z^m, \cdots|(\lambda)+|\cdots, Q_i(z)-\alpha z^m, \cdots|(\lambda), \tag{24}$$

where the first "$\cdots$" denotes "$Q_i(z), \cdots, Q_{i-1}(z)$" and the second "$\cdots$" denotes "$Q_{i+1}(z), \cdots, Q_k(z)$" throughout this proof. From (22) and Lemma 2, the second term on the right of (24) is $O(\lambda^{m-k})$; hence,

$$\Delta(\lambda)=\alpha|\cdots, z^m, \cdots|(\lambda)+O(\lambda^{m-k}). \tag{25}$$

Now we use the identity

$$z^m=c_r^{-1}\left[\lambda z^{m-r}-\sum_{\mu=-s}^{r-1} c_\mu z^{\mu+m-r}+z^{m-k}P(z,\lambda)\right]$$

(see (12)) to write

$$\begin{aligned} |\cdots, z^m, \cdots|(\lambda) &= \lambda c_r^{-1}|\cdots, z^{m-r}, \cdots|(\lambda) \\ &\quad -\sum_{\mu=-s}^{r-1} c_\mu c_r^{-1}|\cdots, z^{\mu+m-r}, \cdots|(\lambda) \\ &\quad + c_r^{-1}|\cdots, z^{m-k}P(z,\lambda), \cdots|(\lambda). \end{aligned} \tag{26}$$

From Definition 1, the ith row of the determinant in the numerator of

$$|\cdots, z^{m-k}P(z,\lambda), \cdots|(\lambda)$$

consists entirely of zeros, so the last term on the right of (26) vanishes. Lemma 2 and (22) imply that the sum on the right of (26) is $O(\lambda^{m-k})$. Therefore,

$$|\cdots, z^m, \cdots|(\lambda)=\lambda c_r^{-1}|\cdots, z^{m-r}, \cdots|(\lambda)+O(\lambda^{m-k}).$$

This and (25) imply (23).

We now establish a useful connection between the eigenvalue problem for H_n and a boundary value problem for a related difference equation. Let

$$U = \mathrm{col}\,[u_0, \cdots, u_{n-1}]$$

and

$$V = H_n U = \mathrm{col}\,[v_0, \cdots, v_{n-1}]. \tag{27}$$

Then,

$$v_i = \sum_{j=0}^{n-1} h_{ijn} u_j = \sum_{j=-i}^{n-i-1} h_{i,j+i,n} u_{j+i}, \qquad 0 \leqq i \leqq n-1.$$

Therefore, from (3),

$$v_i = \sum_{j=-i}^{n-i-1} \left[c_j - \sum_{\nu=i+1}^{s} a_{j+\nu} b_\nu - \sum_{\mu=n-i}^{r} b_{-j+\mu} a_\mu \right] u_{j+i}, \qquad 0 \leqq i \leqq n-1. \tag{28}$$

If $s \leqq i \leqq n-r-1$, then the sums with respect to μ and ν both vanish, and (28) reduces to

$$v_i = \sum_{j=-s}^{r} c_j u_{j+i}. \tag{29}$$

For our purposes it is convenient to have this equation hold for $0 \leqq i \leqq n-1$, but this is impossible as things stand, since (29) would then involve the undefined quantities $u_{-s}, \cdots, u_{-1}$ and $u_n, \cdots, u_{n+r-1}$. This defect can be remedied by defining *extrapolated components* for the vector U, as in the following lemma.

LEMMA 4. *The components of V in* (27) *are given by* (29) *for* $0 \leqq i \leqq n-1$ *if and only if the extrapolated components* $u_{-s}, \cdots, u_{-1}$ *and* $u_n, \cdots, u_{n+r-1}$ *satisfy the equations*

$$\sum_{l=0}^{r} a_l u_{l-p} = 0, \qquad 1 \leqq p \leqq s, \tag{30}$$

and

$$\sum_{l=0}^{s} b_l u_{n+p-l-1} = 0, \qquad 1 \leqq p \leqq r. \tag{31}$$

Proof. The extrapolated components are uniquely defined by $u_0, \cdots, u_{n-1}$ and (30) and (31). They can be computed recursively from the equations

$$u_{-p} = -a_0^{-1} \sum_{l=1}^{r} a_l u_{l-p}, \qquad 1 \leqq p \leqq s, \tag{32}$$

and

$$u_{n+p-1} = -b_0^{-1} \sum_{l=1}^{s} b_l u_{n+p-l-1}, \qquad 1 \leqq p \leqq r. \tag{33}$$

We have already verified (29) for $s \leqq i \leqq n-r-1$. If $0 \leqq i \leqq s-1$, then (28) reduces to

$$v_i = \sum_{j=-i}^{r} \left(c_j - \sum_{\nu=i+1}^{s} a_{j+\nu} b_\nu \right) u_{j+i} \tag{34}$$

because of (2), (4), and (5). From (6),

$$c_j = \sum_{\nu=i+1}^{s} a_{j+\nu} b_\nu, \qquad -s \leqq j \leqq -i-1,$$

(since $a_{j+\nu}=0$ if $j+\nu<0$); hence, we can change the lower limit of summation in (34) to $j=-s$ for any choice of $u_{-s},\cdots,u_{-1}$. Therefore,

$$v_i=\sum_{j=-s}^{r} c_j u_{j+i}-\sum_{j=-s}^{r}\left(\sum_{\nu=i+1}^{s} a_{j+\nu}b_\nu\right)u_{j+i},\qquad 0\leqq i\leqq s-1. \tag{35}$$

The double sum in (35) can be rewritten as

$$\sum_{\nu=i+1}^{s} b_\nu \sum_{j=-s}^{r} a_{j+\nu}u_{j+i}=\sum_{\nu=i+1}^{s} b_\nu \sum_{l=\nu-s}^{\nu+r} a_l u_{l+i-\nu}$$

$$=\sum_{\nu=i+1}^{s} b_\nu \sum_{l=0}^{r} a_l u_{l-(\nu-i)},\qquad 0\leqq i\leqq s-1$$

(since $a_l=0$ if $l<0$ or $l>r$). From this it can be seen that (29) holds for $0\leqq i\leqq s-1$ if and only if (30) holds.

Now suppose $n-r\leqq i\leqq n-1$. Then (28) reduces to

$$v_i=\sum_{j=-s}^{n-i-1}\left(c_j-\sum_{\mu=n-i}^{r} b_{-j+\mu}a_\mu\right)u_{j+i}. \tag{36}$$

Since (6) implies that

$$c_j=\sum_{\mu=n-i}^{r} b_{-j+\mu}a_\mu,\qquad n-i\leqq j\leqq r$$

(recall that $b_{-j+\mu}=0$ if $-j+\mu<0$), we can change the upper limit of summation in (36) to $j=r$ for any choice of $u_n,\cdots,u_{n+r-1}$. Therefore,

$$v_i=\sum_{j=-s}^{r} c_j u_{j+i}-\sum_{j=-s}^{r}\left(\sum_{\mu=n-i}^{r} b_{-j+\mu}a_\mu\right)u_{j+i},\qquad n-r\leqq i\leqq n-1.$$

The double sum on the right can be rewritten as

$$\sum_{\mu=n-i}^{r} a_\mu \sum_{j=-s}^{r} b_{-j+\mu}u_{j+i}=\sum_{\mu=n-i}^{r} a_\mu \sum_{l=\mu-r}^{\mu+s} b_l u_{\mu-l+i}$$

$$=\sum_{\mu=n-i}^{r} a_\mu \sum_{l=0}^{s} b_l u_{\mu-l+i},\qquad n-r\leqq i\leqq n-1$$

(since $b_l=0$ if $l<0$ or $l>s$). From this it can be seen that (29) holds for $n-r\leqq i\leqq n-1$ if and only if (31) holds.

Lemma 4 obviously implies the following lemma.

LEMMA 5. *A complex number λ is an eigenvalue of H_n if and only if there are complex numbers*

$$u_{-s},\cdots,u_{n+r-1}, \tag{37}$$

not all zero, which satisfy the difference equation

$$\sum_{j=-s}^{r} c_j u_{j+i}=\lambda u_i,\qquad 0\leqq i\leqq n-1, \tag{38}$$

and the boundary conditions (30) *and* (31). *In this case the vector*

$$U=\operatorname{col}[u_0,\cdots,u_{n-1}] \tag{39}$$

is an eigenvector of H_n corresponding to λ.

It is important to observe that if the sequence (37) satisfies these hypotheses, then U in (39) is nonzero, since if $u_0=\cdots=u_{n-1}=0$, then (32) and (33) imply that the remaining elements in (37) vanish.

3. The main results. In the following $(x)^{(l)}$ is the factorial polynomial:

$$(x)^{(0)}=1, \qquad (x)^{(l)}=x(x-1)\cdots(x-l+1), \quad l\geqq 1.$$

THEOREM 2. *Let* λ *satisfy the assumptions of Definition* 1, *and let* Ω_n *be the* $k\times k$ *matrix which results from the construction specified in Definition* 1 *when*

$$Q_i(z)=\begin{cases} z^{i-1}A(z), & 1\leqq i\leqq s,\\ z^{n+i-1}B(1/z), & s+1\leqq i\leqq k;\end{cases}$$

thus

$$\Omega_n=\begin{bmatrix} A(z_1) & A(z_2) & \cdots & A(z_k)\\ \vdots & \vdots & & \vdots\\ z_1^{s-1}A(z_1) & z_2^{s-1}A(z_2) & \cdots & z_k^{s-1}A(z_k)\\ z_1^{n+s}B(1/z_1) & z_2^{n+s}B(1/z_2) & \cdots & z_k^{n+s}B(1/z_k)\\ \vdots & \vdots & & \vdots\\ z_1^{n+k-1}B(1/z_1) & z_2^{n+k-1}B(1/z_2) & \cdots & z_k^{n+k-1}B(1/z_k)\end{bmatrix} \tag{40}$$

if λ *is an ordinary point of* $P(z;\lambda)$. *Then* λ *is an eigenvalue of* H_n *if and only if* Ω_n *is singular, in which case the components of the eigenvector* (39) *are given by*

$$u_i=\sum_{j=1}^{q}\sum_{\nu=0}^{m_j-1}\alpha_{\nu j}(s+i)^{(\nu)}z_j^{s+i-\nu} \tag{41}$$

for $0\leqq i\leqq n-1$, *where the vector*

$$X=\operatorname{col}[\alpha_{01},\cdots,\alpha_{m_1-1,1},\alpha_{02},\cdots,\alpha_{m_2-1,2},\cdots,\alpha_{0q},\cdots,\alpha_{m_q-1,q}] \tag{42}$$

is a nontrivial solution of the $k\times k$ *system*

$$\Omega_n X=0. \tag{43}$$

(*Note that* (41) *and* (42) *can be written more simply as*

$$u_i=\sum_{j=1}^{k}\alpha_j z_j^{s+i}, \qquad 0\leqq i\leqq n-1,$$

and

$$X=\operatorname{col}[\alpha_1,\alpha_2,\cdots,\alpha_k]$$

if λ *is an ordinary point of* $P(z;\lambda)$.)

Proof. We use Lemma 5. The general solution of the difference equation (38) is of the form (41) for $-s\leqq i\leqq n+r-1$. (See the proof of Theorem 1 in [11].) Substituting (41) into (30) and summing first on l yields

$$\sum_{j=1}^{q}\sum_{\nu=0}^{m_j-1}\alpha_{\nu j}\sum_{l=0}^{r}a_l(s+l-p)^{(\nu)}z_j^{s+l-p-\nu}=0, \qquad 1\leqq p\leqq s.$$

This is equivalent to

$$\sum_{j=1}^{q}\sum_{\nu=0}^{m_j-1}\alpha_{\nu j}[(z^{s-p}A(z))^{(\nu)}|_{z=z_j}]=0, \qquad 1\leqq p\leqq s. \tag{44}$$

By a similar argument, substituting (41) into (31) yields

$$\sum_{j=1}^{q}\sum_{\nu=0}^{m_j-1}\alpha_{\nu j}[(z^{s+n+p-1}B(1/z))^{(\nu)}|_{z=z_j}]=0, \qquad 1\leqq p\leqq r. \tag{45}$$

Since (44) and (45) together are equivalent to (43), the conclusion follows.

If we let $\Omega=\Omega_n$ in (15), then $\Delta(\lambda)$ in (16) becomes

$$\Delta_n(\lambda)=|A(z),\cdots,z^{s-1}A(z),z^{n+s}B(1/z),\cdots,z^{n+k-1}B(1/z)|(\lambda), \tag{46}$$

which is a polynomial of degree $\leqq n$, by Lemma 2. Since Ω_n is singular if and only if $\Delta_n(\lambda)=0$, Theorem 2 clearly implies that there is a connection between $\Delta_n(\lambda)$ and the characteristic polynomial

$$p_n(\lambda)=\det[\lambda I_n-H_n]. \tag{47}$$

The following results make this connection precise.

THEOREM 3. *If $A(z)$ and $z^sB(1/z)$ are relatively prime, then*

$$p_n(\lambda)=(-1)^{(r-1)n}R^{-1}c_r^n\Delta_n(\lambda), \qquad n>k, \tag{48}$$

where R is the (nonzero) value of the $k\times k$ determinant with rows $i=1,\cdots,k$ as follows:

(a) *For $1\leqq i\leqq s$, there are $i-1$ zeros, then $a_0,\cdots,a_r$, then $s-i$ zeros.*

(b) *For $s+1\leqq i\leqq k$, there are $i-s-1$ zeros, then $b_s,\cdots,b_0$, then $k-i$ zeros.*

Proof. Although $p_n(\lambda)$ has meaning only if $n>k$, $\Delta_n(\lambda)$ is defined for all $n\geqq 0$. We first prove by induction that

$$\Delta_n(\lambda)=(-1)^{(r-1)n}Rc_r^{-n}\lambda^n+g_n(\lambda), \qquad n\geqq 0, \tag{49}$$

where $\deg g_n(\lambda)<n$. It suffices to consider only the case where λ is an ordinary point of $P(z;\lambda)$, since there are at most k critical values of λ, and we already know that $\Delta_n(\lambda)$ is a polynomial of degree $\leqq n$.

From (40), $\Omega_0=WV$, where V is the Vandermonde matrix of Definition 1 and $\det W=R$. This implies (49) for $n=0$, with $g_0=0$. To see that $R\neq 0$, suppose $R=0$. Then (40) with $n=0$ and the last sentence of Lemma 2 imply that

$$A(z),\cdots,z^{s-1}A(z),z^sB(1/z),\cdots,z^{k-1}B(1/z)$$

are linearly dependent. Therefore, there are polynomials $f(z)$ and $g(z)$, not identically zero, such that $\deg f(z)<s$, $\deg g(z)<r$, and $f(z)A(z)=g(z)z^sB(1/z)$. By an argument in [12, § 27], this implies that $A(z)$ and $z^sB(1/z)$ have a nonconstant common factor, which contradicts our assumption.

We now complete the proof of (49) by showing that

$$\Delta_{n+1}(\lambda)=(-1)^{r-1}c_r^{-1}\lambda\Delta_n(\lambda)+O(\lambda^n), \qquad n\geqq 0, \tag{50}$$

where $O(\lambda^n)$ denotes a polynomial of degree $\leqq n$. From (46) with n replaced by $n+1$,

$$\Delta_{n+1}(\lambda)=|\cdots,z^{n+s+1}B(1/z),\cdots,z^{n+k}B(1/z)|(\lambda),$$

where the first "$\cdots$" denotes "$A(z),\cdots,z^{s-1}A(z)$" throughout this proof. The polynomial of highest degree appearing in the definition of $\Delta_{n+1}(\lambda)$ is $z^{n+k}B(1/z)$; hence, Lemma 3 implies that

$$\Delta_{n+1}(\lambda)=b_0\lambda c_r^{-1}|\cdots,z^{n+s+1}B(1/z),\cdots,z^{n+k-1}B(1/z),z^{n+s}|(\lambda)+O(\lambda^n), \tag{51}$$

where $z^{n+s+1}B(1/z), \cdots, z^{n+k-1}B(1/z)$ is absent if $r=1$. We rewrite (51) as

$$\Delta_{n+1}(\lambda)=(-1)^{r-1}\lambda c_r^{-1}|\cdots, b_0 z^{n+s}, z^{n+s+1}B(1/z), \cdots, z^{n+k-1}B(1/z)|(\lambda)+O(\lambda^n)$$

$$=(-1)^{r-1}\lambda c_r^{-1}[\Delta_n(\lambda)+\Gamma_n(\lambda)]+O(\lambda^n), \tag{52}$$

where

$$\Gamma_n(\lambda)=|\cdots, z^{n+s}(b_0-B(1/z)), z^{n+s+1}B(1/z), \cdots, z^{n+k-1}B(1/z)|(\lambda). \tag{53}$$

(See (46).)

We will now show that

$$\Gamma_n(\lambda)=O(\lambda^{n-1}). \tag{54}$$

If $r=1$, then $k=s+1$ and (53) reduces to

$$\Gamma_n(\lambda)=|\cdots, z^{n+s}(b_0-B(1/z))|(\lambda):$$

so Lemma 2 implies (54). If $r>1$, then successively applying Lemma 3 to (53) $r-1$ times yields

$$\Gamma_n(\lambda)=(b_0\lambda c_r^{-1})^{r-1}|\cdots, z^{n+s}(b_0-B(1/z)), z^{n+s-r+1}, \cdots, z^{n+s-1}|(\lambda)$$

$$=(b_0\lambda c_r^{-1})^{r-1}\sum_{\mu=1}^{s} b_\mu|\cdots, z^{n+s-\mu}, z^{n+s-r+1}, \cdots, z^{n+s-1}|(\lambda).$$

The terms in this sum are identically zero for $1\leqq\mu\leqq\max(s, r-1)$ (since they are essentially determinants with two identical rows), and $O(\lambda^{n-r})$ for $r\leqq\mu\leqq s$ (by Lemma 2). This implies (54). Since (52) and (54) imply (50), this completes the proof of (49).

Now (49) implies that the polynomial

$$\tilde{p}_n(\lambda)=(-1)^{(r-1)n}R^{-1}c_r^n\Delta_n(\lambda)$$

is monic and of exact degree n, as is the characteristic polynomial $p_n(\lambda)$ in (47). From Theorem 2, $\tilde{p}_n(\lambda)$ and $p_n(\lambda)$ have the same zeros; therefore certainly $\tilde{p}_n(\lambda)=p_n(\lambda)$ if H_n has n distinct eigenvalues. There remains the possibility that H_n has only m ($<n$) distinct eigenvalues and

$$\tilde{p}_n(\lambda)=(\lambda-\lambda_1)^{r_1}\cdots(\lambda-\lambda_m)^{r_m}, p_n(\lambda)=(\lambda-\lambda_1)^{s_1}\cdots(\lambda-\lambda_m)^{s_m}$$

with $r_i\neq s_i$ for some i; however, this possibility can be excluded by a continuity argument of the kind given in [11].

Theorems 1 and 3 yield the following result, which makes explicit the connection between our results and the eigenvalue problem for Toeplitz matrices with band inverses.

THEOREM 4. *Suppose $A(z)$ and $B(z)$ satisfy* (2) *and* (14), *and $A(z)$ and $z^sB(1/z)$ are relatively prime. Let T_n and $\{\phi_r\}$ be as in Theorem* 1. *Then the characteristic polynomial of T_n is given by*

$$\det[\lambda I_n-T_n]=[\Delta_n(0)]^{-1}\lambda^n\Delta_n(1/\lambda).$$

Moreover, if λ is an eigenvalue of T_n, then the corresponding eigenvectors (39) *can be obtained as in Theorem* 2.

Our results have specific applications to statistics in the case where

$$B(z)=A^*(z)=\sum_{\nu=0}^{r}\bar{a}_\nu z^\nu,$$

so that the matrices H_n ($n>2r$) are Hermitian. Greville [3] has shown that in this case

H_n is positive definite for all $n > 2r$ if and only if the zeros of $A(z)$ are all outside the unit circle, or positive semidefinite if and only if none are inside the unit circle. He also obtained results on the spectral radii of the matrices $\{H_n\}$.

If the roots of $A(z)$ are all outside the unit circle, then $A(z)$ and $z^r A^*(1/z)$ are relatively prime. It can be shown in this case that the sequence $\{\phi_r\}$ defined by (9), (10), and (11) (with $b_\nu = \bar{a}_\nu$) is proportional to the autocorrelation sequence of the purely autoregressive weakly stationary time series $\{y_m\}$ defined by the stochastic difference equation

$$a_0 y_m + a_1 y_{m-1} + \cdots + a_r y_{m-r} = x_m, \qquad -\infty < m < \infty,$$

where $\{x_m\}$ is uncorrelated and weakly stationary.

The formula (48) is clearly invalid if $A(z)$ and $z^s B(1/z)$ have a nonconstant common factor, since R is the resultant of $z^r A(1/z)$ and $B(z)$, which would also have a nonconstant common factor, and therefore $R = 0$ [12, § 27]. In this case we have the following result.

THEOREM 5. *Suppose $A(z)$ and $z^s B(1/z)$ have greatest common divisor*

$$D(z) = (z - \zeta_1) \cdots (z - \zeta_m) \qquad (m \geqq 1), \tag{55}$$

and let

$$\frac{A(z)}{D(z)} = A_1(z) = \alpha_0 + \cdots + \alpha_{r-m} z^{r-m}, \tag{56}$$

$$\frac{z^s B(1/z)}{D(z)} = z^{s-m} B_1(1/z) = \beta_{s-m} + \cdots \beta_0 z^{s-m}. \tag{57}$$

Then the characteristic polynomial $p_n(\lambda)$ in (47) is given by

$$p_n(\lambda) = \frac{(-1)^{m(k+1)+(r-1)(n-m)} c_r^n}{R_1[\zeta_1 \cdots \zeta_m]^s} \lambda^m \tilde{\Delta}_n(\lambda), \qquad n > k, \tag{58}$$

where

$$\tilde{\Delta}_n(\lambda) = |A_1(z), \cdots, z^{s-1} A_1(z), z^{n+s-m} B_1(1/z), \cdots, z^{n+k-m-1} B_1(1/z)|(\lambda), \tag{59}$$

and R_1 is the (nonzero) value of the $k \times k$ determinant with rows $i = 1, \cdots, k$ as follows:

(a) *For $1 \leqq i \leqq s$ there are $i-1$ zeros; then $\alpha_0, \cdots, \alpha_{r-m}$, then $s - i + m$ zeros.*

(b) *For $s+1 \leqq i \leqq k$ there are $m + i - s - 1$ zeros; then $\beta_{s-m}, \cdots, \beta_0$, then $k - i$ zeros.*

Proof. Again we consider only ordinary points λ of $P(z; \lambda)$. For $1 \leqq j \leqq k$, $D(z_j)$ is a common factor of the jth column of the determinant in the numerator of $\Delta_n(\lambda)$. (See (46) and recall (40).) Removing these common factors shows that

$$\Delta_n(\lambda) = D(z_1) \cdots D(z_k) \tilde{\Delta}_n(\lambda), \tag{60}$$

with $\tilde{\Delta}_n(\lambda)$ as in (59), because of (56) and (57). From (55),

$$D(z_1) \cdots D(z_k) = \prod_{l=1}^{m} (z_1 - \zeta_l) \cdots (z_k - \zeta_l). \tag{61}$$

Since $z_1, \cdots, z_k$ are the zeros of $P(z; \lambda)$, (12) implies that

$$(z_1 - \zeta_l) \cdots (z_k - \zeta_l) = (-1)^k c_r^{-1} P(\zeta_l; \lambda), \qquad 1 \leqq l \leqq m.$$

But $A(\zeta_l) = 0$, so (6) and (12) imply that $P(\zeta_l; \lambda) = -\lambda \zeta_l^s$. This and (61) imply that

$$D(z_1) \cdots D(z_k) = (-1)^{m(k+1)} c_r^{-m} (\zeta_1 \cdots \zeta_m)^s \lambda^m.$$

This and (60) yield

$$\Delta_n(\lambda)=(-1)^{m(k+1)}c_r^{-m}(\zeta_1\cdots\zeta_m)^s\lambda^m\tilde{\Delta}_n(\lambda). \tag{62}$$

An induction argument like the one used to prove (49) shows that

$$\tilde{\Delta}_n(\lambda)=(-1)^{(r-1)(n-m)}R_1c_r^{-n+m}\lambda^{n-m}+\tilde{g}_n(\lambda),\qquad n\geqq m, \tag{63}$$

where $\deg \tilde{g}_n(\lambda)<n-m$. To see that $R_1\neq 0$, we observe that since $A_1(z)$ and $z^{s-m}B_1(1/z)$ are relatively prime and $A_1(0)\neq 0$, it follows that $A_1(z)$ and $z^sB_1(1/z)$ are relatively prime. Therefore, an argument like the one used in Theorem 2 to prove that $R\neq 0$ applies here.

From (62) and (63), the polynomial on the right of (58) is monic and of exact degree n. An argument similar to that used in the proof of Theorem 2 now establishes (58). This completes the proof of Theorem 4.

Laplace's development provides a convenient method for expanding the determinants in (46) and (59); see [11, § 5].

Now let $E_n(\lambda)$ be the solution space of the system

$$H_nX=\lambda X.$$

The following lemma is analogous to a lemma obtained in [11] for Toeplitz band matrices.

LEMMA 6. *Let λ and $z_1,\cdots,z_q$ be as in Definition 1. Then λ is an eigenvalue of H_n if and only if there are polynomials*

$$f(z)=C_0+\cdots+C_{s-1}z^{s-1},\qquad g(z)=D_0+\cdots+D_{r-1}z^{r-1}, \tag{64}$$

such that the polynomial

$$h(z)=f(z)A(z)+z^{n+s}g(z)B(1/z) \tag{65}$$

is not identically zero and has zeros at $z_1,\cdots,z_q$ with multiplicities at least $m_1,\cdots,m_q$; i.e.,

$$h^{(l)}(z_j)=0,\quad 0\leqq l\leqq m_j-1,\quad 1\leqq j\leqq q. \tag{66}$$

Moreover, if $S_n(\lambda)$ is the vector space of polynomials h of the form (64) *and* (65) *which satisfy* (66), *then*

$$\dim S_n(\lambda)=\dim E_n(\lambda).$$

Proof. A polynomial h of the stated form satisfies (66) if and only if the vector

$$Y=\mathrm{col}\,[C_0,\cdots,C_{s-1},D_0,\cdots,D_{r-1}]$$

satisfies the system

$$\Omega_n^t Y=0.$$

Therefore, $\dim S_n(\lambda)=$ nullity of $\Omega_n^t=$ nullity of $\Omega_n=\dim E_n(\lambda)$. (See the proof of Theorem 2.)

Lemma 6 implies the next two theorems. Since the proofs of these theorems are the same as those of Theorems 3 and 4 of [11], we omit them.

THEOREM 6. *If λ is an eigenvalue of H_n then*

$$\dim E_n(\lambda)\leqq \min\,(r,s).$$

THEOREM 7. *Suppose* λ *is an eigenvalue of* H_n *and* $\dim E_n(\lambda) \geqq 2$. *Then* λ *is also an eigenvalue of* H_{n-1} (*if* $n > k+1$) *and* H_{n+1}; *moreover,*

$$\dim E_{n-1}(\lambda) \geqq -1 + \dim E_n(\lambda)$$

and

$$\dim E_{n+1}(\lambda) \geqq -1 + \dim E_n(\lambda).$$

Acknowledgments. The author thanks T. N. E. Greville for reviving his interest in the matrices discussed here several years after the publication of [10]. The present paper was strongly influenced by the joint paper [7], and by Greville's subsequent work [2]–[6] on these matrices and on their applications to data smoothing. In particular, the crucial step of introducing the extrapolated components of U in Lemmas 4 and 5 was motivated by a similar device used by Greville in [4], [5], and [6] to extend symmetric smoothing formulas to the extremities of data.

REFERENCES

[1] D. BINI AND M. CAPOVANI, *Spectral and computational properties of band symmetric Toeplitz matrices*, Lin. Alg. Appl., 52/53 (1983), pp. 99–126.

[2] T. N. E. GREVILLE, *On a problem concerning band matrices with Toeplitz inverses*, Proc. 8th Manitoba Conference on Numerical Methods of Computation, Utilitas, Winnipeg, 1978, pp. 275–283.

[3] ———, *Bounds for the eigenvalues of Hermitian Trench matrices*, Proc. 9th Manitoba Conference Numerical Methods of Computation, Utilitas, Winnipeg, 1980, pp. 241–256.

[4] ———, *Moving-weighted-average smoothing extended to the extremities of the data*. I. *Theory*, Scand. Actuar. J. (1981), pp. 39–55.

[5] ———, *Moving-weighted-average smoothing extended to the extremities of the data*. II. *Methods*, Scand. Actuar. J. (1981), pp. 65–81.

[6] ———, *Moving-weighted-average smoothing extended to the extremities of the data*. III. *Stability and optimal properties*, J. Approx. Th., 33 (1981), pp. 43–58.

[7] T. N. E. GREVILLE AND W. F. TRENCH, *Band matrices with Toeplitz inverses*, Lin. Alg. Appl., 27 (1979), pp. 199–209.

[8] F. A. GRUNBAUM, *Toeplitz matrices commuting with tridiagonal matrices*, Lin. Alg. Appl., 40 (1981), pp. 25–36.

[9] ———, *Eigenvectors of a Toeplitz matrix: discrete version of prolate spheroidal wave functions*, this Journal, 2 (1981), pp. 136–141.

[10] W. F. TRENCH, *Weighting coefficients for the prediction of stationary time series from the finite past*, SIAM J. Appl. Math., 15 (1967), pp. 1502–1510.

[11] ———, *On the eigenvalue problem for Toeplitz band matrices*, Lin. Alg. Appl., 64 (1985), pp. 199–214.

[12] B. L. VAN DER WAERDEN, *Modern Algebra*, Frederick Ungar, New York, 1949.