1985

On the eigenvalue problem for Toeplitz matrices generated by rational functions

William F. Trench, *Trinity University*

Linear and Multilinear Algebra, 1985, Vol. 17, pp. 337–353
0308-1087/85/1704-0337 $20.00

On the Eigenvalue Problem for Toeplitz Matrices Generated by Rational Functions

WILLIAM F. TRENCH
Department of Mathematics and Computer Science, Drexel University, Philadelphia, PA 19104

(Received August 15, 1984; in final form November 3, 1984)

Formulas are given for the characteristic polynomials $\{p_n(\lambda)\}$ and the eigenvectors of the family $\{T_n\}$ of Toeplitz matrices generated by a formal Laurent series of a rational function $R(z)$. The formulas are in terms of the zeros of a certain fixed polynomial with coefficients which are simple functions of λ and the coefficients of $R(z)$. The complexity of the formulas is independent of n.

1. INTRODUCTION

We consider the eigenvalue problem for Toeplitz matrices

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

generated by rational functions; thus,

$$\sum_{j=-\infty}^{\infty} t_j z^j \sim \frac{C(z)}{A(z)B(1/z)} = R(z), \tag{1}$$

where "$\sim$" is explained below,

$$A(z) = \sum_{\mu=0}^{r} a_\mu z^\mu, \tag{2}$$

$$B(z) = \sum_{\nu=0}^{s} b_\nu z^\nu, \tag{3}$$

and

$$C(z) = \sum_{j=-q}^{p} c_j z^j. \tag{4}$$

We take the underlying field to be the complex numbers, and assume throughout that no two of the polynomials $A(z)$, $z^s B(1/z)$, and $z^q C(z)$ have a common factor. We also assume that

$$a_0 a_r b_0 b_s c_p c_{-q} \neq 0 \qquad \text{and} \qquad p + q > 0;$$

however, we do not assume that p and q have the same sign. To describe our results, it is convenient to denote

$$M = \max(p, r), \qquad N = \max(q, s), \tag{5}$$

and

$$k = M + N.$$

We will assume that $M \geqslant 1$ and $N \geqslant 1$, which is not a serious restriction, since T_n is triangular if $MN = 0$, and the eigenvalue problem for triangular Toeplitz matrices is essentially trivial. Now let

$$P(z; \lambda) = z^N (C(z) - \lambda A(z) B(1/z)), \tag{6}$$

which is a polynomial of exact degree k in z for all except at most one value of λ. Our main result is an expression for the characteristic polynomial

$$p_n(\lambda) = \det[\lambda I_n - T_n] \tag{7}$$

in the form

$$p_n(\lambda) = K_n (c_M - \lambda a_M b_0)^n \det \Omega_n / \det V, \tag{8}$$

where $c_M = 0$ if $M = r > p$, $a_M = 0$ if $M = p > r$, K_n is a specifically defined constant, V is a $k \times k$ Vandermonde matrix involving the zeros $\{z_j\}$ of $P(z; \lambda)$, and Ω_n is a $k \times k$ matrix with elements of the form $z_j^{i-1} A(z_j)$ in rows $i = 1, \ldots, N$ and $z_j^{n+i-1} B(1/z_j)$ in rows $i = N + 1, \ldots, k$. (If λ is such that $P(z; \lambda)$ has repeated roots, then V becomes a generalized Vandermonde matrix and the entries in some columns of Ω_n involve derivatives of these functions, as defined precisely below.) Therefore, (8) provides a representation of $p_n(\lambda)$ with complexity essentially independent of n. We also give an explicit formula for the eigenvectors of T_n corresponding to a given eigenvalue λ. The formula depends upon k coefficients which can be obtained by solving a

homogeneous system of k equations in k unknowns; again, the complexity of this computation is essentially independent of n.

The results of this paper extend those obtained in [10] for the case where

$$A(z) = B(z) = 1, \tag{9}$$

so that T_n is a Toeplitz band matrix. They are also related to results obtained in [11].

There is an extensive literature on Toeplitz matrices generated by rational functions, devoted mainly to studying the asymptotic distribution of the eigenvalues of T_n as $n \to \infty$. For example, Widom [12] considered the case where (9) holds and

$$C(z) = \sum_{j=-p}^{p} c_j z^j,$$

with $c_j = \bar{c}_{-j}$, so that T_n is Hermitian. Schmidt and Spitzer [9] extended his results by assuming (9), but allowing $C(z)$ to be of the general form (4), so that T_n need not be Hermitian. Formulas obtained by the author in [10] for the characteristic polynomials of Toeplitz band matrices can be viewed as generalizations of formulas obtained by quite different methods in [9] and [12].

Day [3] extended some of the results of Widom and Schmidt and Spitzer, by eliminating the restriction (9). He represented $p_n(\lambda)$ as a determinant of order and form depending only on the zeros of $A(z)$, $B(z)$, and $C(z)$, and not on n. His results are quite different from those presented here, and his assumptions more stringent. Recently, Bottcher and Silbermann [2, Theorem 6.28] have obtained a formula which generalizes Day's, but does not seem to be directly related to ours. They also mention the existence of a formula for the determinant of T_n in an apparently unpublished manuscript of Gorodeckii [7], which, from their description, may be related to our formula (49) (below) with $\lambda = 0$. However, they also indicate that Gorodeckii imposed conditions on the locations of the zeros of $A(z)$ and $B(z)$ which imply that the series in (1) and (11) converge in an annular region containing $|z| = 1$, while we make no such assumption.

Dickinson [4] has devised an efficient method for solving the system $T_n X = Y$, where T_n is a Toeplitz matrix generated by a rational function.

2. PRELIMINARY CONSIDERATIONS

The rational function in (1) is written in the form shown so that we may conveniently draw on results from [6], and because it is the natural form for applications to stationary time series. We use "$\sim$" rather than "$=$" in (1) because we do not assume that the series on the left converges for any z, but only that it is a certain formal Laurent series generated by $R(z)$. The coefficients $\{t_j\}$ in any such series are of the form

$$t_j = \sum_{l=-q}^{p} c_l \phi_{j-l}, \tag{10}$$

where

$$\sum_{j=-\infty}^{\infty} \phi_j z^j \sim [A(z)B(1/z)]^{-1} = R_0(z); \tag{11}$$

therefore, to define $\{t_j\}$ we must specify which formal expansion of $R_0(z)$ is being considered. To this end, and because of its usefulness below, we present the following lemma, which is a convenient restatement of part of the results obtained in [6]. Here and throughout the paper, we write

$$A(z)B(1/z) = \sum_{j=-s}^{r} \theta_j z^j, \tag{12}$$

and define

$$a_j = 0 \quad \text{if } j < 0 \text{ or } j > r; \qquad b_j = 0 \quad \text{if } j < 0 \text{ or } j > s; \tag{13}$$

$$c_j = 0 \quad \text{if } j < -q \text{ or } j > p; \qquad \theta_j = 0 \quad \text{if } j < -s \text{ or } j > r; \tag{14}$$

and

$$\sum_{i}^{j} = 0 \qquad \text{if} \quad i > j. \tag{15}$$

With these definitions, we can write, for example,

$$\theta_j = \sum_{\nu=0}^{s} a_{j+\nu} b_\nu = \sum_{\mu=0}^{r} a_\mu b_{\mu-j}, \qquad -\infty < j < \infty. \tag{16}$$

Lemma 1 (Greville–Trench) *Let $A(z)$ and $B(z)$ be as in* (2) *and* (3), *with $a_0 b_0 \neq 0$ and $A(z)$ and $z^s B(1/z)$ relatively prime. Then there is a unique sequence $\{\phi_j\}_{-\infty}^{\infty}$ such that*

$$\sum_{\nu=0}^{r} a_\nu \phi_{j-\nu} = b_0^{-1} \delta_{j0}, \qquad j \geqslant 0,$$

and

$$\sum_{\mu=0}^{s} b_\mu \phi_{-j+\mu} = a_0^{-1}\delta_{j0}.$$

Moreover, if $m > r + s$, then the Toeplitz matrix

$$\Phi_m = (\phi_{j-i})_{i,j=0}^{m-1}$$

is invertible, with inverse

$$\Phi_m^{-1} = (h_{ijm})_{i,j=0}^{m-1}, \tag{17}$$

where

$$h_{ijm} = \theta_{j-i} - \sum_{\nu=i+1}^{s} a_{j-i+\nu} b_\nu - \sum_{\mu=m-i}^{r} b_{i-j+\mu} a_\mu, \tag{18}$$

$$0 \leqslant i, j \leqslant m-1.$$

Greville [5] has given a particularly clear presentation of the precise manner in which the sequence $\{\phi_j\}$ is generated by $R_0(z)$, as follows. Since $A(z)$ and $z^s B(1/z)$ are relatively prime, there are unique polynomials $f(z)$ and $g(z)$ of degrees less than s and r, respectively, such that

$$1 = f(z)A(z) + g(z)z^s B(1/z).$$

Dividing this identity by $A(z)B(1/z)$ yields

$$R_0(z) = f(z)[B(1/z)]^{-1} + g(z)z^s[A(z)]^{-1}. \tag{19}$$

Substituting the convergent expansions

$$[A(z)]^{-1} = \sum_{\nu=0}^{\infty} \sigma_\nu z^\nu, \qquad |z| < R_1$$

and

$$[B(1/z)]^{-1} = \sum_{\nu=0}^{\infty} \eta_\nu z^{-\nu}, \qquad |z| > R_2$$

into (19) yields the series in (11). Whether the series converges in some annulus depends upon the relative locations of the zeros of $A(z)$ and $B(z)$; however, convergence is irrelevant to the problem that we are considering. In any case, formal manipulation yields

$$A(z)B(1/z) \sum_{j=-\infty}^{\infty} \phi_j z^j = 1,$$

so the series is a formal reciprocal of $A(z)B(1/z)$.

From (4), (5), and (12), the polynomial defined in (6) can be written as

$$P(z;\lambda) = \sum_{j=-N}^{M} (c_j - \lambda\theta_j) z^{j+N}. \tag{20}$$

Since $C(z)$ and $A(z)B(1/z)$ have no common zeros, $P(z;\lambda)$ is irreducible. Therefore, the equation

$$P(z;\lambda) = 0 \tag{21}$$

defines z as a multiple-valued analytic function of the complex variable λ. (See, e.g. [1], pp. 300–306.) Since our results can be established without function-theoretic arguments, we make little use of this theory here. (It is to be hoped that function-theoretic arguments will be useful in devising computational procedures from the theoretical results given here, but we cannot justify any such claim as yet.) We make only the following observations, which are used below:

(i) As a function of z, $P(z;\lambda)$ is a polynomial of exact degree k, unless

$$c_M - \lambda\theta_M = 0, \tag{22}$$

which obviously occurs for at most one value of λ, and then only if $M = r$.

(ii) For all values of λ which do not satisfy (22), (21) has k roots, counting multiplicities, and they are all nonzero unless

$$c_{-N} - \lambda\theta_{-N} = 0, \tag{23}$$

which occurs for at most one value of λ, and then only if $N = s$.

(iii) If λ does not satisfy (22), then the k roots of (21) are distinct unless the resultant of $P(z;\lambda)$ and $P_z(z;\lambda)$ vanishes ([1], p. 301). Since this resultant is a polynomial in λ, this occurs for at most finitely many values of λ.

The following definition has been used previously in [10] and [11]. The notation of this definition applies throughout the paper.

Definition 1 For a fixed λ which does not satisfy (22), let $z_1, \dots, z_L$ be the distinct roots of (21), with multiplicities $m_1, \dots, m_L$; thus,

$$L \leqslant k, \qquad m_j \geqslant 1 \quad (1 \leqslant j \leqslant L); \qquad m_1 + \cdots + m_L = k.$$

If $Q_1(z), \dots, Q_k(z)$ are given polynomials, let

$$w(z) = \mathrm{col}[Q_1(z), \dots, Q_k(z)],$$

and let Ω be the $k \times k$ matrix constructed as follows: its first m_1 columns are $w^{(l)}(z_1)$ $(0 \leqslant l \leqslant m_1 - 1)$; its next m_2 columns are $w^{(l)}(z_2)$ $(0 \leqslant l \leqslant m_2 - 1)$; and so forth. Let V be the generalized Vandermonde matrix resulting from this construction when

$$Q_i(z) = z^{i-1}, \qquad 1 \leqslant i \leqslant k,$$

and define

$$\Delta(\lambda) = \frac{\det \Omega}{\det V}. \tag{24}$$

If (21) has k distinct roots $z_1, \dots, z_k$, then V is the ordinary Vandermonde matrix, and (24) becomes

$$\Delta(\lambda) = \left[\prod_{1 \leqslant i < j \leqslant k} (z_j - z_i) \right]^{-1} \det[Q_i(z_j)]_{i,j=1}^{k}.$$

The function $\Delta(\lambda)$ is well defined in any case, since it can be shown in general that V is nonsingular [8]; moreover, the numbering of the roots $z_1, \dots, z_L$ does not affect the value of $\Delta(\lambda)$, since renumbering them would affect only the signs of the numerator and denominator in (24), and in the same way.

We will sometimes find it convenient to write

$$\Delta(\lambda) = |Q_1(z), \dots, Q_k(z)|(\lambda) \tag{25}$$

when we wish to make the choice of $Q_1(z), \dots, Q_k(z)$ explicit. (The reader should not be misled by this notation, which appears to suggest that $\Delta(\lambda)$ is also a function of z; it is not.)

3. THE MAIN RESULTS

In the special case where

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

we will denote the matrix Ω and the function $\Delta(\lambda)$ of Definition 1 by Ω_n and $\Delta_n(\lambda)$. Thus, if the roots of (21) are distinct, then

$$\Omega_n = \begin{bmatrix} A(z_1) & A(z_2 & \cdots & A(z_k) \\ \vdots & \vdots & & \vdots \\ z_1^{N-1}A(z_1) & z_2^{N-1}A(z_2) & \cdots & z_k^{N-1}A(z_k) \\ z_1^{n+N}B(1/z_1) & z_2^{n+N}B(1/z_2) & \cdots & z_k^{n+N}B(1/z_k) \\ \vdots & \vdots & & \vdots \\ z_1^{n+k-1}B(1/z_1) & z_2^{n+k-1}B(1/z_2) & \cdots & z_k^{n+k-1}B(1/z_k) \end{bmatrix}. \quad (26)$$

Whether the roots are distinct or not,

$$\begin{aligned} \Delta_n(\lambda) &= \frac{\det \Omega_n}{\det V} \\ &= |A(z), \ldots, z^{N-1}A(z), z^{n+N}B(1/z), \ldots, z^{n+k-1}B(1/z)|(\lambda). \end{aligned} \quad (27)$$

Clearly, Ω_n is singular if and only if $\Delta_n(\lambda) = 0$.

In the following,

$$(x)^{(0)} = 1; \qquad (x)^{(\nu)} = x(x-1)\cdots(x-\nu+1), \qquad \nu \geqslant 1;$$

and $E_n(\lambda)$ denotes the eigenspace of T_n corresponding to a given eigenvalue λ.

THEOREM 1 *Suppose that* $\boxed{\text{See the addendum}}$

$$n > \max(r + s - k, 0) \quad (28)$$

and λ does not satisfy (22) *or* (23). *Then λ is an eigenvalue of T_n if and only if Ω_n is singular. In this case, $E_n(\lambda)$ consists of the vectors*

$$U = \mathrm{col}[u_0, \ldots, u_{n-1}] \quad (29)$$

of the form

$$u_i = \sum_{j=1}^{L} \sum_{r=0}^{m_j-1} \alpha_{rj}[z^{N+i}A(z)B(1/z)]^{(r)}|_{z=z_j}, \quad (30)$$

$$0 \leqslant i \leqslant n-1,$$

where the vector

$$X = \mathrm{col}[\alpha_{01}, \ldots, \alpha_{m_1-1,1}, \ldots, \alpha_{0q}, \ldots, \alpha_{m_L-1,L}] \quad (31)$$

satisfies the equation

$$\Omega_n X = 0. \tag{32}$$

Proof A vector U as in (29) is in $E_n(\lambda)$ if and only if

$$\sum_{j=0}^{n-1} t_{j-i} u_j = \lambda u_i, \qquad 0 \leqslant i \leqslant n-1. \tag{33}$$

From (10), this can be rewritten as

$$\sum_{l=-q}^{p} c_l v_{l+i} = \lambda u_i, \qquad 0 \leqslant i \leqslant n-1, \tag{34}$$

where

$$v_i = \sum_{j=0}^{n-1} \phi_{j-i} u_j, \qquad -N \leqslant i \leqslant n+M-1. \tag{35}$$

Actually, (34) is equivalent to (33) if v_i is defined by (35) only for $-q \leqslant i \leqslant n+p-1$; however, it is convenient to extend the definition (35) for $-N \leqslant i \leqslant n+M-1$. We can rewrite (35) in the form of an $(n+k) \times (n+k)$ system by simply defining

$$u_i = 0 \qquad \text{if} \quad -N \leqslant i \leqslant -1 \quad \text{or} \quad n \leqslant i \leqslant n+M-1. \tag{36}$$

Doing this, letting $i' = N + i$ and $j' = N + j$, and then dropping the primes yields

$$v_{-N+i} = \sum_{j=0}^{n+k} \phi_{j-i} u_{-N+j}, \qquad 0 \leqslant i \leqslant n+k-1.$$

The matrix of this system is Φ_{n+k}; therefore, (17) and (18) with $m = n + k$ imply that if $n + k > r + s$, then

$$u_{-N+i} = \sum_{j=0}^{n+k} \left[\theta_{j-i} - \sum_{\nu=i+1}^{s} a_{j-i+\nu} b_\nu - \sum_{\mu=n+k-i}^{r} b_{i-j+\mu} a_\mu \right] v_{-N+j},$$
$$0 \leqslant i \leqslant n+k-1.$$

Letting $i' = -N + i$ and $j' = j - N - i'$ and then dropping the primes yields

$$u_i = \sum_{j=-N-i}^{n+M-i} \left[\theta_j - \sum_{\nu=N+i+1}^{s} a_{j+\nu} b_\nu - \sum_{\mu=n+M-i}^{r} b_{-j+\mu} a_\mu \right] v_{j+i}, \tag{37}$$
$$-N \leqslant i \leqslant n+M-1.$$

(Henceforth we use (13), (14), and (15) without explicit citations.) Now (37) implies that

$$u_i = \sum_{l=-s}^{r} \theta_l v_{l+i}, \qquad 0 \leqslant i \leqslant n-1. \tag{38}$$

From (16), (36), and (37),

$$\sum_{j=-N-i}^{r} \left(\sum_{\nu=0}^{N+i} a_{j+\nu} b_\nu \right) v_{j+i} = 0, \qquad -N \leqslant i \leqslant -1, \tag{39}$$

and

$$\sum_{j=-s}^{n+M-i} \left(\sum_{\mu=0}^{n+M-i-1} b_{-j+\mu} a_\mu \right) v_{j+i} = 0, \qquad n \leqslant i \leqslant n+M-1. \tag{40}$$

Changing the order of summation in (39) and performing the first summation of the result on $l = j + \nu$ yields

$$\sum_{\nu=0}^{N+i} b_\nu \sum_{l=0}^{r} a_l v_{l-\nu+i} = 0, \qquad -N \leqslant i \leqslant -1. \tag{41}$$

Changing the order of summation in (40) and performing the first summation of the result on $l = -j + \mu$ yields

$$\sum_{\mu=0}^{n+M-i-1} a_\mu \sum_{l=0}^{s} b_l v_{i-l+\mu} = 0, \qquad n \leqslant i \leqslant n+M-1. \tag{42}$$

Since $a_0 b_0 \neq 0$, (41) is equivalent to

$$\sum_{l=0}^{r} a_l v_{l-i} = 0, \qquad 1 \leqslant i \leqslant N, \tag{43}$$

and (42) is equivalent to

$$\sum_{l=0}^{s} b_l v_{n+i-l-1} = 0, \qquad 1 \leqslant i \leqslant M. \tag{44}$$

After substituting (38) into (34), we can summarize the results thus far as follows: an arbitrary complex number λ is an eigenvalue of T_n if and only if there is a nonzero vector

$$\hat{V} = [v_{-N}, \ldots, v_{n+M-1}]$$

such that

$$\sum_{j=-N}^{M} (c_j - \lambda \theta_j) v_{j+i} = 0, \qquad 0 \leqslant i \leqslant n-1, \tag{45}$$

and (43) and (44) hold. In this case the vector U defined by (29) and (38) is in $E_n(\lambda)$. To put it another way, we have reduced the eigenvalue problem for T_n to finding values of λ for which the difference equation (45) has nontrivial solutions which satisfy the boundary conditions (43) and (44).

Thus far we have not used our assumption that λ does not satisfy (22) or (23). Now let us invoke this assumption. Then the general solution of (45) can be written as

$$v_i = \sum_{j=1}^{L} \sum_{\nu=0}^{m_j-1} \alpha_{\nu j}(N+i)^{(\nu)} z_j^{N+i-\nu}, \qquad -N \leqslant i \leqslant n+M-1. \tag{46}$$

Substituting (46) into (43) and then summing first on l yields

$$\sum_{j=1}^{L} \sum_{\nu=0}^{m_j-1} \alpha_{\nu j} \sum_{l=0}^{r} a_l (N+l-i)^{(\nu)} z_j^{N+l-i-\nu} = 0, \qquad 1 \leqslant i \leqslant N,$$

which can be rewritten as

$$\sum_{j=1}^{L} \sum_{\nu=0}^{m_j-1} \alpha_{\nu j}[z^{N-i}A(z)]^{(\nu)}|_{z=z_j} = 0, \qquad 1 \leqslant i \leqslant N. \tag{47}$$

By a similar argument, substituting (46) into (44) yields

$$\sum_{j=1}^{L} \sum_{\nu=0}^{m_j-1} \alpha_{\nu j}[z^{N+n+i-1}B(1/z)]^{(\nu)}|_{z=z_j} = 0, \qquad 1 \leqslant i \leqslant M. \tag{48}$$

Since (31), (47), and (48) are equivalent to (32), λ is an eigenvalue of T_n if and only if (32) has nontrivial solutions; i.e. if and only if Ω_n is singular.

Substituting (46) into (38), summing first on l, and then recalling (12) yields (30), which completes the proof.

Since (21) can have repeated roots for at most finitely many values of λ, it seems worthwhile to state the following corollary of Theorem 1.

COROLLARY 1 *In addition to the assumptions of Theorem* 1, *suppose* λ *is such that* (21) *has distinct roots* $z_1, \ldots, z_k$. *Then* λ *is an eigenvalue of* T_n *if and only if there are constants* $\alpha_1, \ldots, \alpha_k$, *not all zero, which satisfy the* $k \times k$ *system*

$$\sum_{j=1}^{k} \alpha_j z_j^{N-i} A(z_j) = 0, \qquad 1 \leqslant i \leqslant N,$$

$$\sum_{j=1}^{k} \alpha_j z_j^{N+n+i-1} B(1/z_j) = 0, \qquad 1 \leqslant i \leqslant M.$$

In this case the eigenvector (29) is given by

$$u_i = \sum_{j=1}^{k} \alpha_j z_j^{N+i} A(z_j) B(1/z_j), \qquad 0 \leqslant i \leqslant n-1.$$

From (26) and (27), Theorem 1 implies that there is a connection between $\Delta_n(\lambda)$ and the characteristic polynomial $p_n(\lambda)$ of T_n. The following theorem makes this connection precise. The proof of this theorem is presented in the form of a sequence of lemmas, some of which extend results from [10] and [11], and may be of interest in their own right. See the addendum

THEOREM 2 *If n satisfies (28), then the characteristic polynomial (7) is given by*

$$p_n(\lambda) = (-1)^{(M-1)n} R^{-1} (a_0 b_0)^{-n} (c_M - \lambda \theta_M)^n \Delta_n(\lambda), \tag{49}$$

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

(a) *For $1 \leqslant i \leqslant N$, there are $i-1$ zeros, then $a_0, \ldots, a_M$, then $N - i$ zeros.*

(b) *For $N + 1 \leqslant i \leqslant k$, there are $i - N - 1$ zeros, then $b_N, \ldots, b_0$, then $k - i$ zeros.*

4. PROOF OF THEOREM 2

In this section the assumptions of Definition 1 apply, the polynomial $p = 0$ has degree $-\infty$, and $0(\lambda^\nu)$ denotes a polynomial of degree $\leqslant \nu$.

LEMMA 2 *If $Q_1(z), \ldots, Q_{i-1}(z)$, $Q_{i+1}(z), \ldots, Q_k(z)$ are polynomials and μ is a nonnegative integer, then*

$$|Q_1(z), \ldots, Q_{i-1}(z), z^\mu P(z;\lambda), Q_{i+1}(z), \ldots, Q_k(z)|(\lambda) = 0 \tag{50}$$

for all $\lambda \neq \theta_M / c_M$.

Proof Under the assumptions of Definition 1,

$$[z^\mu P(z;\lambda)]^{(l)}|_{z=z_j} = 0, \qquad 0 \leqslant l \leqslant m_j - 1,\ 1 \leqslant j \leqslant L.$$

Hence the ith row of the matrix Ω in Definition 1 (with $Q_i(z) = z^\mu P(z;\lambda)$) consists entirely of zeros. Therefore, (24) implies (50).

LEMMA 3 *If*

$$\deg Q_i(z) \leqslant m, \qquad 1 \leqslant i \leqslant k,$$

then

$$(c_M - \lambda\theta_M)^{m-k+1}|Q_1(z), \ldots, Q_k(z)|(\lambda) = O(\lambda^{m-k+1}). \tag{51}$$

Proof First suppose that

$$Q_i(z) = z^{n_i}, \qquad 1 \leqslant i \leqslant k, \tag{52}$$

where $n_1, \ldots, n_k$ are nonnegative integers. We will show that if

$$\max\{n_1, \ldots, n_k\} \leqslant m, \tag{53}$$

then

$$(c_M - \lambda\theta_M)^{m-k+1}|z^{n_1}, \ldots, z^{n_k}|(\lambda) = O(\lambda^{m-k+1}). \tag{54}$$

If $0 \leqslant m \leqslant k-1$ and (52) holds, then Ω (cf. Definition 1) has two identical rows or its rows are a permutation of the rows of V; hence, (24) and (25) imply that

$$|z^{n_1}, \ldots, z^{n_k}|(\lambda) = 0,\ 1,\ \text{or}\ -1.$$

This proves (54) if $0 \leqslant m \leqslant k-1$. Now suppose we have established (54) for any set of nonnegative integers $n_1, \ldots, n_k$ satisfying (53), and let

$$\max\{n_1, \ldots, n_k\} = m + 1.$$

Without loss of generality, we may assume that

$$n_k = m + 1 \qquad \text{and} \qquad n_j \leqslant m, \qquad 1 \leqslant j \leqslant k-1.$$

We can complete the induction by showing that

$$(c_M - \lambda\theta_M)^{m-k+2}|\ldots, z^{m+1}|(\lambda) = O(\lambda^{m-k+2}),$$

where "$\ldots$" denotes "$z^{n_1}, \ldots, z^{n_{k-1}}$" for the rest of the proof. From Lemma 2,

$$|\ldots, z^{m-k+1}P(z;\lambda)|(\lambda) = 0.$$

This and (20) imply that

$$(c_M - \lambda\theta_M)|\ldots, z^{m+1}|(\lambda) = -\sum_{j=-N}^{M-1} (c_j - \lambda\theta_j)|\ldots, z^{j+m-M+1}|(\lambda). \tag{55}$$

Our induction assumption implies that

$$(c_j - \lambda\theta_j)^{m-k+1}|\ldots, z^{j+m-M+1}|(\lambda) = O(\lambda^{m-k+1}),$$
$$-N \leqslant j \leqslant M-1,$$

which, with (55), yields

$$(c_M - \lambda\theta_M)^{m-k+2}|\ldots, z^{m+1}|(\lambda) = O(\lambda^{m-k+2}).$$

This completes the induction, proving (54). To obtain (51) from this, we observe that if

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

then

$$|Q_1(z), \ldots, Q_k(z)|(\lambda) = \sum a_{1j_1} \cdots a_{kj_k} |z^{j_1}, \ldots, z^{j_k}|(\lambda), \tag{56}$$

where the sum is over all $j_1, \ldots, j_k$ such that $0 \leqslant j_1, \ldots, j_k \leqslant m$. Multiplying (56) by $(c_M - \lambda\theta_M)^{m-k+1}$ and invoking (54) yields (51).

LEMMA 4 *For a given complex number* λ,

$$|Q_1(z), \ldots, Q_k(z)|(\lambda) = 0 \tag{57}$$

if and only if there are constants $\gamma_1, \ldots, \gamma_k$, *not all zero, such that the polynomial*

$$Q(z) = \gamma_1 Q_1(z) + \cdots + \gamma_k Q_k(z) \tag{58}$$

is divisible by $P(z; \lambda)$. *In particular, if* $m \leqslant k-1$, *so that* $\Delta(\lambda) = C$ *(constant, from Lemma 3), then* $C = 0$ *if and only if* $Q_1(z), \ldots, Q_k(z)$ *are linearly dependent.*

Proof From (24) and (25), (57) is equivalent to the existence of a nontrivial solution $Y = \mathrm{col}[\gamma_1, \ldots, \gamma_k]$ of the system

$$\Omega^t Y = 0 \qquad ({}^t = \text{transpose}). \tag{59}$$

From the definition of Ω, (59) is equivalent to

$$Q^{(l)}(z_j) = 0, \qquad 0 \leqslant l \leqslant m_j - 1,\ 1 \leqslant j \leqslant L, \tag{60}$$

with $Q(z)$ as in (58). But (60) holds if and only if $P(z; \lambda)$ divides $Q(z)$. This completes the proof.

In particular, Lemma 4 implies that the function

$$q_n(\lambda) = (c_M - \lambda \theta_M)^n \Delta_n(\lambda) \qquad \text{(cf. (27))} \tag{61}$$

is a polynomial of degree $\leqslant n$.

LEMMA 5 *The polynomial $q_n(\lambda)$ in (61) is of the form*

$$q_n(\lambda) = (-1)^{(M-1)n}(a_0 b_0)^n R\lambda^n + O(\lambda^{n-1}), \tag{62}$$

where R is the constant introduced in the statement of Theorem 2. Moreover,

$$R \neq 0. \tag{63}$$

Proof To evaluate the constant $q_0(\lambda)$, we may without loss of generality choose λ so that (21) has distinct roots $z_1, \ldots, z_k$. Then it can be seen that $\Omega_0 = WV$, where

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

and W is a constant matrix with rows as described in the statement of Theorem 2; therefore, det $W = R$. This and (61) imply (62) with $n = 0$. If $R = 0$, then $\Delta_0(\lambda) = 0$, and the last sentence in Lemma 4 implies that

$$A(z), \quad \ldots, \quad z^{N-1}A(z), \quad z^N B(1/z), \quad \ldots, \quad z^{k-1}B(1/z)$$

are linearly dependent. (See (27) with $n = 0$.) Therefore, there are polynomials $f(z)$ and $g(z)$, not identically zero, such that deg $f(z) < N$, deg $g(z) < M$, and

$$f(z)A(z) = g(z)z^N B(1/z).$$

This implies that $A(z)$ and $z^s B(1/z)$ have a nonconstant common factor, contrary to our assumption. This proves (63).

We now complete the proof of (60) inductively by showing that

$$q_{n+1}(\lambda) = (-1)^{M-1} a_0 b_0 \lambda q_n(\lambda) + O(\lambda^n), \qquad n \geqslant 0. \tag{64}$$

By adding appropriate multiples of the last $M - 1$ rows of Ω_n to its $(N + 1)$st row, we see from (26) that

$$\begin{aligned} \Delta_n(\lambda) = a_0^{-1}|\ldots, z^{n+N}(A(z) - a_M z^M)B(1/z), \\ z^{n+N+1}B(1/z), \ldots, z^{n+k-1}B(1/z)|(\lambda), \end{aligned} \tag{65}$$

where the first "..." denotes "$A(z), \ldots, z^{N-1}A(z)$" throughout the rest of

the proof, and "$z^{n+N+1}B(1/z),\ldots,z^{n+k-1}B(1/z)$" is to be omitted if $M=1$. From (6),

$$\begin{aligned}
&z^{n+N}(A(z)-a_M z^M)B(1/z)\\
&\quad=\lambda^{-1}(z^{n+N}C(z)-z^nP(z;\lambda))-a_M z^{n+k}B(1/z)\\
&\quad=(c_M-\lambda a_M b_0)(\lambda b_0)^{-1}z^{n+k}B(1/z)+\lambda^{-1}(g_n(z)-z^nP(z;\lambda)),
\end{aligned}\tag{66}$$

where

$$g_n(z)=z^{n+N}C(z)-c_M b_0^{-1}z^{n+k}B(1/z),$$

and therefore

$$\deg g_n(z)\leqslant n+k-1.\tag{67}$$

Recalling that $a_M b_0=\theta_M$ (see (16)), substituting the last member of (66) into (65), and multiplying the resulting equation by $\lambda a_0 b_0$ yields

$$\begin{aligned}
&\lambda a_0 b_0\Delta_n(\lambda)\\
&\quad=(-1)^{M-1}(c_M-\lambda\theta_M)\Delta_{n+1}(\lambda)\\
&\qquad+b_0|\ldots,z^nP(z;\lambda),z^{n+N+1}B(1/z),\ldots,z^{n+k+1}B(1/z)|(\lambda)\\
&\qquad+b_0|\ldots,g_n(z),z^{n+N+1}B(1/z),\ldots,z^{n+k-1}B(1/z)|(\lambda).
\end{aligned}\tag{68}$$

Lemma 2 implies that the second term on the right vanishes identically, while (67) and Lemma 3 imply that the last term is of the form $(c_M-\lambda\theta_M)^{-n}O(\lambda^n)$. Therefore, multiplying (68) by $(c_M-\lambda\theta_M)^n$ and recalling (61) yields (64). This completes the proof.

Proof of Theorem 2 Let $Q_n(\lambda)$ denote the right side of (49). By Lemma 5, $Q_n(\lambda)$ is a monic polynomial of degree n, as is the characteristic polynomial $p_n(\lambda)$. Therefore, Theorem 2 implies that

$$p_n(\lambda)=Q_n(\lambda)\tag{69}$$

if $p_n(\lambda)$ has n distinct zeros $\lambda_1,\ldots,\lambda_n$ such that

$$\lambda_i\neq c_M/\theta_M\qquad\text{and}\qquad\lambda_i\neq c_{-N}/\theta_{-N},\qquad 1\leqslant i\leqslant n.\tag{70}$$

Now suppose $p_n(\lambda)$ does not have this property, and let $\varepsilon>0$. Then we can choose

$$\hat{A}(z)=\sum_{\mu=0}^{r}\hat{a}_\mu z^\mu,\qquad\hat{B}(z)=\sum_{\nu=0}^{s}\hat{b}_\nu z^\nu,$$

Lane [illegible]

and

$$\hat{C}(z) = \sum_{j=-q}^{p} \hat{c}_j z^j,$$

with

$$\sum_{\mu=0}^{r} |\hat{a}_\mu - a_\mu|^2 + \sum_{\nu=0}^{s'} |\hat{b}_\nu - b_\nu|^2 + \sum_{j=-q}^{p} |\hat{c}_j - c_j|^2 < \varepsilon^2, \tag{71}$$

so that $\hat{A}(z)$, $\hat{B}(z)$, and $\hat{C}(z)$ satisfy our hypotheses and the characteristic polynomial $\hat{p}_n(\lambda)$ of the corresponding Toeplitz matrix $\hat{T}_n$ has n distinct roots which satisfy the counterpart of (70). Then, if $\hat{Q}_n(\lambda)$ is related to $\hat{A}(z)$, $\hat{B}(z)$, and $\hat{C}(z)$ as $Q_n(\lambda)$ is related to $A(z)$, $B(z)$, and $C(z)$, it follows that $\hat{p}_n(\lambda) = \hat{Q}_n(\lambda)$. Since the coefficients of $\hat{p}_n(\lambda)$ and $\hat{Q}_n(\lambda)$ are continuous functions of the coefficients of $\hat{A}(z)$, $\hat{B}(z)$, and $\hat{C}(z)$, we can now let $\varepsilon \to 0$ in (70) to conclude that (69) holds in general. This completes the proof.

Acknowledgement

The author is grateful to the referee for calling references [2] and [7] to his attention.

References

[1] L. V. Ahlfors, *Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1979.

[2] A. Böttcher and B. Silbermann, *Invertibility and Asymptotics of Toeplitz Matrices*, *Mathematical Research 17*, Akademie-Verlag, Berlin, 1983.

[3] K. M. Day, Toeplitz matrices generated by the Laurent series expansion of an arbitrary rational function, *Trans. Amer. Math. Soc.* **206** (1975), 224–245.

[4] B. W. Dickinson, Solution of linear equations with rational Toeplitz matrices, *Math. Comp.* **34** (1980), 227–233.

[5] T. N. E. Greville, On a problem concerning band matrices with Toeplitz inverses, *Proc. 8th Manitoba Conf. Numer. Math. Comput.* (1978), 275–283 (Utilitas Mathematica Publishing Company).

[6] T. N. E. Greville and W. F. Trench, Band matrices with Toeplitz inverses, *Lin. Alg. Appl.* **27** (1979), 199–209.

[7] M. B. Gorodeckii, Toeplitz determinants generated by rational functions (Russian; manuscript deposited in "VINITI," Dec. 1981, NO. 5451-81 Dep.).

[8] D. Kalman, The generalized Vandermonde matrix, *Math. Mag.* **57** (1984), 15–21.

[9] P. Schmidt and F. Spitzer, The Toeplitz matrices of an arbitrary Laurent polynomial, *Math. Scand.* **8** (1960), 15–38.

[10] W. F. Trench, On the eigenvalue problem for Toeplitz band matrices, *Lin. Alg. Appl.* **64** (1985), 199–214.

[11] W. F. Trench, On the eigenvalue problem for a class of band matrices including those with Toeplitz inverses (to appear, *SIAM J. Alg. Disc. Meth.*).

[12] H. Widom, On the eigenvalues of certain Hermitean operators, *Trans. Amer. Math. Soc.* **88** (1958), 491–522.

LAMA 669 ①

Linear and Multilinear Algebra, 1986, Vol. 20, pp. 000–000
0000–0000/86/0000–0000 $10.00/0

Addendum

On the Eigenvalue Problem for Toeplitz Matrices Generated by Rational Functions

WILLIAM F. TRENCH

(*Received January 24, 1986*)

The requirement that

$$n > \max(r + s - k, 0) \tag{28}$$

in Theorem 1 of [1] is superfluous, since (5) and the immediately following definition of k obviously imply that the quantity on the right is always zero; therefore, Theorem 1 holds for all $n \geqslant 1$. The same is true of Theorem 2.

Reference

[1] W. F. Trench, On the eigenvalue problem for Toeplitz matrices generated by rational functions, *Lin. Multilin. Algebra* 19 (1985), 169–185.