Eigenvalues of tridiagonal pseudo-Toeplitz matrices

Devadatta Kulkarni, Darrell Schmidt, Sze-Kai Tsui

Eigenvalues of tridiagonal pseudo-Toeplitz matrices

Devadatta Kulkarni, Darrell Schmidt, Sze-Kai Tsui

Department of Mathematical Sciences, Oakland University, Rochester, MI 48309 4485, USA

In this article we determine the eigenvalues of sequences of tridiagonal matrices that contain a Toeplitz matrix in the upper left block.

Keywords: Eigenvalues; Toeplitz matrices; Tridiagonal matrices; Chebyshev polynomials; Graph ical method for location of roots

1. Introduction

Although Hermitian matrices are known to have real eigenvalues only, the evaluation of these eigenvalues remains as misty as ever. For tridiagonal matrices there are several known methods describing their eigenvalues such as Gershgorin's theorem [5], Sturm sequences for Hermitian tridiagonal matrices [1,4], etc. The eigenvalues of a tridiagonal Toeplitz matrix can be completely determined [11]. Attempts have been made to resolve the eigenvalue problem for matrices which are like tridiagonal Toeplitz matrices but not entirely Toeplitz (see [2,3,12,13]). This paper falls in the same general direction of investigation.

We study tridiagonal matrices which contain a Toeplitz matrix in the upper left block. We call them pseudo-Toeplitz to make a distinction from all the matrices studied before in [2,3,12], etc. The major feature of our treatment is the connection between the characteristic polynomials of these tridiagonal pseudo-Toeplitz matrices and the Chebyshev polynomials of the second kind, whereby we can locate the eigenvalues that fall in the intervals determined by the roots of some Chebyshev polynomials of the second kind. In other words, we use these intervals derived from roots of some Chebyshev polynomial as a reference to determine the eigenvalues of the original pseudo-Toplitz matrix. In fact, we are able to determine the location of all eigenvalues of some tridiagonal pseudo-Toeplitz matrices which have either all entries real with a nonnegative product from each off-diagonal pair or the entries on the main diagonal purely imaginary with a negative product from each off-diagonal pair (see Corollary 3.4). In Section 2, we lay down a basic tool for finding the eigenvalues of tridiagonal Toeplitz matrices, which is markedly different from the traditional approach used in [2,11,12]. In Section 3, we give a detailed account of the number of eigenvalues in each such interval whose end points are consecutive roots of a pair of Chebyshev polynomials related to the given tridiagonal pseudo-Toeplitz matrix. We also show that, for a sequence of (real) tridiagonal matrices with a positive product from each pair of off-diagonal entries, the eigenvalues of two consecutive matrices in the sequence interlace (see Proposition 3.1). Furthermore, we discuss a lower bound for the number of real eigenvalues for tridiagonal pseudo-Toeplitz matrices of a fixed dimension (see Theorem 3.6). In Section 4, we demonstrate examples of tridiagonal pseudo-Toeplitz matrices for which we can completely determine their real eigenvalues graphically.

These techniques have been also applied in infinite dimensional programming [10] and in numerical solutions of heat equations [3]. Standard references for the Chebyshev polynomials are [6,8,9].

2. Eigenvalues of tridiagonal Toeplitz matrices

It is known that the eigenvalues of tridiagonal Toeplitz matrices can be determined analytically. The method employs the boundary value difference equation [11]. In this section, we provide a different approach to the solution which will be extended to determine eigenvalues of several more general matrices in the later sections.

Let $T_n(a,b,c)$ be an $n \times n$ tridiagonal matrix defined by

$$T_n(a,b,c) = \begin{bmatrix} a & c & & \mathbf{0} \\ b & \ddots & \ddots & \\ & \ddots & \ddots & c \\ \mathbf{0} & & b & a \end{bmatrix}.$$

When $b = c = 1$ we denote $T_n(a, b, c)$ by $T_n(a)$. $T_n(a, b, c)$ is the same matrix denoted by $T_n^0(a, b, c)$ in the later sections. We denote the characteristic polynomial of $T_n(a)$ by $\phi_n(a)(\lambda)$, and it is related to the nth degree Chebyshev polynomial of the second kind. Indeed, expanding $\det(T_n(a) - \lambda I)$ by the last row, we have

$$\phi_n(a)(\lambda) = (a - \lambda)\phi_{n-1}(a)(\lambda) - \phi_{n-2}(a)(\lambda) \tag{1}$$

for $n \geqslant 2$, with $\phi_0(a)(\lambda) = 1, \phi_1(a)(\lambda) = a - \lambda$. Substituting $a - \lambda = 2x$, (1) becomes

$$\phi_n(a)(x) = 2x\phi_{n-1}(a)(x) - \phi_{n-2}(a)(x) \tag{2}$$

for $n \geqslant 2$, with $\phi_0(a)(x) = 1$, $\phi_1(a)(x) = 2x$. Thus, $\phi_n(a)(x)$ is the nth degree Chebyshev polynomial of the second kind, denoted by U_n. It is well-known that

$$U_n(x) = \frac{\sin((n+1)\cos^{-1}(x))}{\sin(\cos^{-1} x)} \quad \text{for } |x| \leqslant 1,$$

and the roots of $U_n(x)$ are $\cos(k\pi/(n+1))(k = 1, 2, \ldots, n)$. Thus, we have the following proposition.

Proposition 2.1. *The eigenvalues of $T_n(a)$ are*

$$a - 2\cos(k\pi/(n+1)) \quad \textit{for } k = 1, 2, \ldots, n.$$

Next, we relate $T_n(a, b, c)$ to $T_n(a)$. Note that

$$\frac{1}{\sqrt{bc}} T_n(a, b, c) = T_n\left(\frac{a}{\sqrt{bc}}, \frac{b}{\sqrt{bc}}, \frac{c}{\sqrt{bc}}\right),$$

and the eigenvalues of $\alpha T_n(a, b, c)$ are just α times the eigenvalues of $T_n(a, b, c)$. Thus, it suffices to consider the characteristic polynomial $\phi_n(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc})$ of $T_n(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc})$. As above, we can see that $\phi_n(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc})$ satisfies the same recurrence relation and initial conditions as $\phi_n(a/\sqrt{bc})$:

$$\phi_n(\lambda) = \left(\frac{a}{\sqrt{bc}} - \lambda\right)\phi_{n-1}(\lambda) - \phi_{n-2}(\lambda)(n \leqslant 2),$$

where $\phi_0(\lambda) = 1$ and $\phi_1(\lambda) = a/\sqrt{bc} - \lambda$. Thus,

$$\phi_n\left(\frac{a}{\sqrt{bc}},\frac{b}{\sqrt{bc}},\frac{c}{\sqrt{bc}}\right)=\phi_n\left(\frac{a}{\sqrt{bc}}\right)=U_n.$$

From Proposition 2.1 we know that the eigenvalues for $T_n\left(a/\sqrt{bc}\right)$ are $a/\sqrt{bc}-2\cos(k\pi/(n+1))$ $(k=1,2,\ldots,n)$. Hence, we have the following result.

Theorem 2.2. *The eigenvalues of* $T_n(a,b,c)$ *are*

$$a-2\sqrt{bc}\cos(k\pi/(n+1))\quad \textit{for } k=1,2,\ldots,n.$$

3. Main results

In this section we study the eigenvalues of those tridiagonal matrices the upper left block of which are Toeplitz matrices. That is, we consider

$$T_n^k(a,b,c)=\left[\begin{array}{c|cccc} \overbrace{T_n(a,b,c)}^{n} & & \overbrace{\mathbf{0}}^{k} & & \\ & c_k & & & \\ \hline b_k & a_k & \ddots & & \\ \mathbf{0} & \ddots & \ddots & \ddots & c_1 \\ & & & b_1 & a_1 \end{array}\right]\begin{array}{l}\}\,n\\ \\ \}\,k\end{array}=\left[\begin{array}{c|c} T_n(a,b,c) & \\ & c_k \\ \hline b_k & \\ & B_k \end{array}\right]$$

where

$$T_n(a,b,c)=\begin{bmatrix} a & c & & \mathbf{0}\\ b & \ddots & \ddots & \\ & \ddots & \ddots & c\\ \mathbf{0} & & b & a\end{bmatrix},\quad B_k=\begin{bmatrix} a_k & c_{k-1} & & \mathbf{0}\\ b_{k-1} & \ddots & \ddots & \\ & \ddots & \ddots & c_1\\ \mathbf{0} & & b_1 & a_1\end{bmatrix}.$$

In the previous section we have determined the eigenvalues of $T_n(a,b,c)$ completely. Now we consider $T_n^k(a/\sqrt{bc},b/\sqrt{bc},c/\sqrt{bc})$. We denote the characteristic polynomials of $T_n^k(a/\sqrt{bc},b/\sqrt{bc},c/\sqrt{bc})$ and $(1/\sqrt{bc})B_k$ by $\phi_n^k(\lambda)$ and $\psi_k(\lambda), k\geqslant 1$, respectively, where $\phi_0^0(\lambda)=1$, $\psi_0(\lambda)=1$. Expanding $\det\left(T_n^k(a/\sqrt{bc},b/\sqrt{bc},c/\sqrt{bc})-\lambda I\right)$ by the last k rows using the Laplace development, we have for $n\geqslant 1$ and $k\geqslant 1$ that

$$\phi_n^k(\lambda)=\phi_n^0(\lambda)\psi_k(\lambda)-\frac{b_kc_k}{bc}\phi_{n-1}^0(\lambda)\psi_{k-1}(\lambda). \tag{3}$$

It follows from Eq. (2) in Section 2 that $\phi_n^0(x)$ is the nth degree Chebyshev polynomial $U_n(x)$ of the second kind with

$$\frac{a}{\sqrt{bc}} - \lambda = 2x. \tag{4}$$

If λ is a root of (3) but not a common root of U_{n-1} and ψ_k, then

$$\frac{U_n(\lambda)}{U_{n-1}(\lambda)} = \frac{(b_k c_k/bc)\psi_{k-1}(\lambda)}{\psi_k(\lambda)}.$$

Let $U_n(x)/U_{n-1}(x) = p_n(x)$, $n \geqslant 1$ and $p_0(x) = 1$. Then

$$p_n(x) = \frac{U_n(x)}{U_{n-1}(x)} = \frac{2xU_{n-1} - U_{n-2}(x)}{U_{n-1}} = 2x - \frac{1}{p_{n-1}(x)}, \quad n \geqslant 1,$$

and

$$p_n'(x) = 2 + \frac{p_{n-1}'(x)}{p_{n-1}^2(x)} = 2 + \sum_{k\ 1}^{n-1} \frac{2}{p_{n-1}^2 p_{n-2}^2 \cdots p_{n-k}^2}, \quad n \geqslant 1,$$

$$= 2 + \frac{2}{U_{n-1}^2} \sum_{k\ 1}^{n-1} U_{n-1-k}^2. \tag{5}$$

Thus, $p_n'(x) > 0$ for all $n \geqslant 1$ and for all x in the domain of p_n. Next we denote for $1 \leqslant j \leqslant k$,

$$g_j(\lambda) = \frac{\psi_{j-1}(\lambda) b_j c_j / bc}{\psi_j(\lambda)}.$$

We compare their graphs.

Let $\eta_1, \ldots, \eta_{n-1}$, be the zeros of U_{n-1} and $\xi_1, \ldots, \xi_n$ be the zeros of U_n. It is known that $-1 < \xi_1 < \eta_1 < \xi_2 < \eta_2 < \cdots < \xi_{n-1} < \eta_{n-1} < \xi_n < 1$. Also denote $\eta_0 = \xi_0 = -\infty$ and $\eta_n = \xi_{n+1} = \infty$. It follows from (5) that p_n is strictly increasing in each interval (η_{j-1}, η_j), $1 \leqslant j \leqslant n$. The graph of $p_n(x)$ is shown in Fig. 1.

In order to describe the behavior of $g_j, 1 \leqslant j \leqslant k$, we impose the following conditions for the ensuing paragraphs through Corollary 3.4

$$\frac{a_j}{\sqrt{bc}}, \frac{a}{\sqrt{bc}} \text{ are real} \quad \text{and} \quad \frac{b_j c_j}{bc} \geqslant 0, \quad 1 \leqslant j \leqslant k. \tag{6}$$

For $2 \leqslant j \leqslant k$, expanding the determinant that generates $\psi_j(\lambda)$, by the first row, we have

$$\psi_j(\lambda) = \left(\frac{a_j}{\sqrt{bc}} - \lambda\right)\psi_{j-1}(\lambda) - \frac{b_{j-1}c_{j-1}}{bc}\psi_{j-2}(\lambda).$$

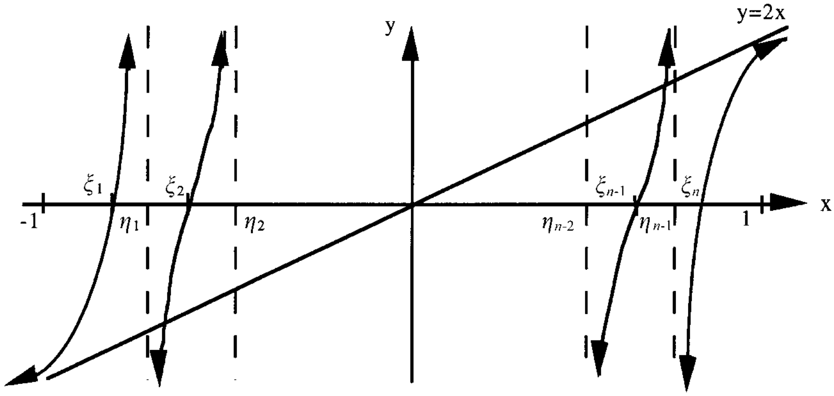

Fig. 1. $y \quad p_n(x)$.

It follows that

$$\begin{aligned} g_j(\lambda) &= \frac{(b_jc_j/bc)\psi_{j-1}(\lambda)}{\left(a_j/\sqrt{bc}-\lambda\right)\psi_{j-1}(\lambda) - (b_{j-1}c_{j-1}/bc)\psi_{j-2}(\lambda)} \\ &= \frac{b_jc_j/bc}{\left(a_j/\sqrt{bc}-\lambda\right) - g_{j-1}(\lambda)} \end{aligned}$$

and

$$\begin{aligned} g_j'(\lambda) &= \frac{(b_jc_j/bc)(1+g_{j-1}'(\lambda))}{((a_j/bc-\lambda)-g_{j-1}(\lambda))^2}, \quad 2 \leqslant j \leqslant k \\ g_1'(\lambda) &= \frac{b_1c_1/bc}{(a/\sqrt{bc}-\lambda)^2}. \end{aligned} \tag{7}$$

By induction, $g_k'(\lambda)$ is nonnegative, and hence $g_k'(x) \leqslant 0$ in view of (4). Due to (6) the tridiagonal matrices $(1/\sqrt{bc})B_k$ are similar to symmetric matrices and hence they have exactly k real eigenvalues, counting multiplicities (see [7, p. 174]).

Next, we look into the situation where the eigenvalues of $(1/\sqrt{bc})B_k$ and the eigenvalues of $(1/\sqrt{bc})B_{k-1}$ are interlacing. For this we prefer to denote

$$A_k = \begin{bmatrix} a_1 & c_1 & & \mathbf{0} \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{k-1} \\ \mathbf{0} & & b_{k-1} & a_k \end{bmatrix}$$

as a sequence of tridiagonal matrices satisfying $b_jc_j > 0, 1 \leqslant j \leqslant k-1$, and a_j, $1 \leqslant j \leqslant n$, are real. In this notation we have the following proposition.

Proposition 3.1. *The eigenvalues of A_k are distinct and interlace strictly with eigenvalues of A_{k-1} for $k \geqslant 2$.*

Proof. The proof is by induction on k. We denote the characteristic polynomial of A_k by $\varphi_k(\lambda)$. The root of $\varphi_1(\lambda)$ is a_1 and

$$\varphi_2(\lambda) = (\lambda - a_1)(\lambda - a_2) - b_1 c_1.$$

It follows from $b_1c_1 > 0$ that $\varphi_2(\lambda)$ has one root in $(\max(a_1, a_2), \infty)$ and one root in $(-\infty, \min(a_1, a_2))$. Thus, the assertion holds for $k = 2$. Assume that the assertion holds for order $k-1$. Let $\rho_1 < \rho_2 < \cdots < \rho_{k-1}$ be the eigenvalues of A_{k-1}, and $\zeta_1, < \cdots <, \zeta_{k-2}$ be the eigenvalues of A_{k-2}, where $\rho_1 < \zeta_1 < \rho_2 \cdots < \zeta_{n-2} < \rho_{n-1}$ by hypothesis. Now $\varphi_{k-2}(\lambda) = (-1)^{k-2}\lambda^{k-2} +$ lower order terms so that $\varphi_{k-2}(\lambda) = \prod_{j\ 1}^{k-2}(\zeta_j - \lambda)$. It follows that $(-1)^{k-2+j}\varphi_{k-2} > 0$ on $(\zeta_{k-2-j}, \zeta_{k-1-j}), 0 \leqslant j \leqslant k-2$, where $\zeta_0 = -\infty, \zeta_{k-1} = \infty$. Since $\rho_{k-1-j} \in (\zeta_{k-2-j}, \zeta_{k-1-j})$, $(-1)^{k+j}\varphi_{k-2}(\rho_{k-1-j}) > 0$, $0 \leqslant j \leqslant k-2$.

Expanding the determinant that generates $\varphi_k(\lambda)$ by the last row yields

$$\varphi_k(\lambda) = (a_k - \lambda)\varphi_{k-1}(\lambda) - b_{k-1}c_{k-1}\varphi_{k-2}(\lambda), \quad k \geqslant 2. \tag{8}$$

Then it follows from (8) and $b_{k-1}c_{k-1} > 0$ that $(-1)^{k+j}\varphi_k(\rho_{k-1-j}) < 0$, $0 \leqslant j \leqslant k-2$. So φ_k has a zero in (ρ_{j-1}, ρ_j), $2 \leqslant j \leqslant k-1$. It remains to show that φ_k has a zero in each of $(-\infty, \rho_1)$ and (ρ_{k-1}, ∞). Observe that $(-1)^k\varphi_k(\lambda) = \lambda^k +$ lower terms, so $(-1)^k\varphi_k(\rho_{k-1}) < 0 < (-1)^k\varphi_k(\lambda)$ for λ sufficiently larger than ρ_{n-1}. Thus, φ_k has a zero in (ρ_{k-1}, ∞). Also $\varphi_k(\rho_1) < 0 < \varphi_k(\lambda)$ for λ sufficiently smaller than ρ_1. Thus, φ_k has a zero in $(-\infty, \rho_1)$. □

Now, let $\zeta_1 \leqslant \cdots \leqslant \zeta_k$ be the roots of $\psi_k(x)$ and $\rho_1 \leqslant \cdots \leqslant \rho_{k-1}$ be the roots of $\psi_{k-1}(x)$. Then, it follows from Proposition 3.1 that $\zeta_1 < \rho_1 < \zeta_2 < \rho_2 < \cdots < \rho_{k-1} < \zeta_k$ if $b_jc_j/bc > 0$ for $1 \leqslant j \leqslant k$. However, if $b_jc_j = 0$ for some $1 \leqslant j \leqslant k$, then ψ_k and ψ_{k-1} have common root(s). Let l be the largest index, j, such that $b_jc_j = 0$. Then, expanding the determinant that yields $\phi_n^k(\lambda)$ by the last l rows according to the Laplace development, we have

$$\phi_n^k(\lambda) = \tilde{\phi}_n^{k-l}(\lambda)\psi_l(\lambda),$$

where $\tilde{\phi}_n^{k-l}(\lambda)$ is the characteristic polynomial of $\tilde{T}_n^k$ which is the $n + (k-l)$ order square matrix in the upper left corner of T_n^k. $\tilde{T}_n^k$ is of the form T_n^k if we reindex the entries in the lower right corner of $\tilde{T}_n^k$. We also note that $g_k(x)$, in its reduced form, has exactly $k-l$ poles and $k-1-l$ zeros. The l real roots of $\psi_l(x)$ are roots of $\phi_n^k(\lambda)$. In this decoupled case, we may focus our attention on determining roots of $\tilde{\phi}_n^{k-l}(\lambda)$. We also have an equation analogous to (3)

$$\tilde{\phi}_n^{k-l}(\lambda) = \phi_n^0(\lambda)\tilde{\psi}_{k-l}(\lambda) - \frac{b_kc_k}{bc}\phi_{n-1}^0(\lambda)\tilde{\psi}_{k-1-l}(\lambda),$$

where for $0 \leqslant j \leqslant k-l-1$, $\tilde{\psi}_{k-j-l}(\lambda)$ is the characteristic polynomial of the matrix $(1/\sqrt{bc})\tilde{B}_{k-j-l}$, where

$$\tilde{B}_{k-j-l} = \begin{bmatrix} a_{k-j} & c_{k-j-1} & & \mathbf{0} \\ b_{k-j-1} & \ddots & \ddots & \\ & \ddots & \ddots & c_{l+1} \\ \mathbf{0} & & b_{l+1} & a_{l+1} \end{bmatrix}.$$

Thus g_k, in its reduced form, is $\tilde{g}_{k-l} = (b_k c_k/bc)\tilde{\psi}_{k-1-l}/\tilde{\psi}_{k-l}$.

It follows from the intermediate value theorem that $p_n(x)$ and $g_k(x)$, in its reduced form, must agree with each other at least once in every interval (η_{j-1}, η_j), $1 \leqslant j \leqslant n$, for $p_n(x)$ is strictly increasing and $g_k'(x) < 0$ wherever $g_k(x)$ is defined in (η_{j-1}, η_j), $1 \leqslant j \leqslant n$ from (7). If a pole of $g_k(x), \zeta_i, 1 \leqslant i \leqslant k-l$ is not a pole of $p_n(x)$, then ζ_i must fall in an interval (η_{j-1}, η_j), for some j, $1 \leqslant j \leqslant n$. If $\zeta_i, \zeta_{i+1}, \ldots, \zeta_{i+r}$ are the poles of $g_k(x)$ that lie in (η_{j-1}, η_j) for some j, then by the intermediate value theorem, $p_n(x)$ and $g_k(x)$ must agree exactly once in each of the following intervals, $(\eta_{j-i}, \zeta_i), (\zeta_i, \zeta_{i+1}), \ldots, (\zeta_{i+r}, \eta_j)$, giving rise to $r+1$ roots of $\phi_n^k(\lambda)$ in Eq. (3). In this notation we have the following theorem.

Theorem 3.2. *Suppose that $g_k(x)$ is in the reduced form. Then*

(i) *for $1 \leqslant j \leqslant n$, (η_{j-1}, η_j) contains one more root of $\phi_n^k(x)$ than poles of $g_k(x)$;*

(ii) *$\phi_n^k(x)$ has $n+k$ real roots. Furthermore, these roots are distinct, if $b_j c_j \neq 0, 1 \leqslant j \leqslant k$.*

Proof. (i) Follows immediately from the discussion before the theorem. For (ii) it suffices to show that $\tilde{\phi}_n^{k-l}$ has $n+(k-l)$ real roots. We note that each common pole of $g_k(x)$ and $p_n(x)$ gives rise to a root of $\phi_n^k(x)$ in Eq. (3). We may now assume that $\tilde{g}_{k-l}(x)$ and $p_n(x)$ have no common poles. Thus, it follows from part (i) that the $\tilde{\phi}_n^{k-l}(x)$ must have $n+(k-l)$ real roots. If $b_j c_j \neq 0$ for all $1 \leqslant j \leqslant k$, then it follows from Proposition 3.1 that $\phi_n^k(x)$ has $n+k$ distinct real roots. □

A similar analysis of the location of roots of $\phi_n^k(x)$ can be done with regards to intervals $(\xi_{j-1}, \xi_j), 1 \leqslant j \leqslant n+1$, which is in the following theorem.

Theorem 3.3. *Suppose $b_j c_j \neq 0$, $1 \leqslant j \leqslant k$. Each (ξ_{j-1}, ξ_j) for $1 \leqslant j \leqslant n+1$ contains one more root of $\phi_n^k(x)$ than zeros of $g_k(x)$.*

Proof. The graph of $p_n(x)$ in $[\xi_{j-1}, \xi_j]$ is depicted in Fig. 2.

If $g_k(x)$ has no zeros in (ξ_{j-1}, ξ_j), then it follows from Proposition 3.1 that $g_k(x)$ can have at most one pole, ρ_l, in $[\xi_{j-1}\xi_j]$. In addition, if $g_k(x)$ has no pole in $[\xi_{j-1}\xi_j]$, then the graph of $g_k(x)$ is strictly decreasing on this interval, and $g_k(x) > 0$ or $g_k(x) < 0$ for all x in $[\xi_{j-1}, \xi_j]$, and hence $g_k(x)$ and $p_n(x)$ agree

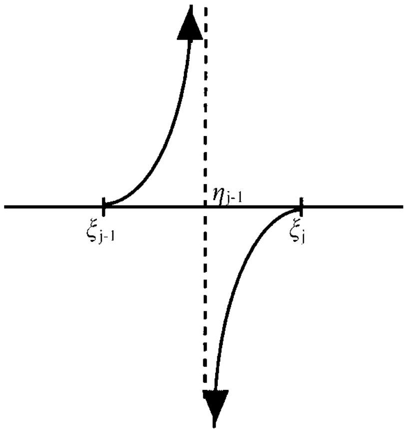

Fig. 2. $y \quad p_n(x)$.

exactly once in (ξ_{j-1}, ξ_j). If there is exactly one pole, ζ_i, of $g_k(x)$ such that $\xi_{j-1} \leqslant \zeta_i \leqslant \xi_j$, then $g_k(x)$ is strictly decreasing and $g_k(x) < 0$ in (ξ_{j-1}, ζ_i), and is strictly decreasing and $g_k(x) > 0$ in (ζ_i, ξ_j). If $\xi_{j-1} \leqslant \zeta_i < \eta_{j-1}$, $g_k(x)$ and $p_n(x)$ agree exactly once in (ζ_i, ξ_j). If $\eta_{j-1} < \zeta_i \leqslant \xi_j$, then $g_k(x)$ and $p_n(x)$ agree exactly once in (ξ_{j-1}, ζ_i). If $\zeta_i = \eta_{j-1}$, then $g_k(x)$ and $p_n(x)$ are not equal for all x in (ξ_{j-1}, ξ_j). But, $\zeta_i = \eta_{j-1}$ is a root of $\phi_n^k(x)$.

In general, suppose $g_k(x)$ has r zeros, $\rho_i < \rho_{i+1} \cdots < \rho_{i+(r-1)}$, in (ξ_{j-1}, ξ_j). Then $p_n(x)$ has no zeros in each of $(\xi_j, \rho_i), (\rho_i, \rho_{i+1}), \ldots, (\rho_{i+(r-1)}, \xi_j)$. Repeating the argument in the previous paragraph with the roles of $p_n(x)$ and $g_k(x)$ reversed, we conclude that there is exactly one root of $\phi_n^k(x)$ in each of the intervals $(\xi_j, \rho_i), (\rho_i, \rho_{i+1}), \ldots, (\rho_{i+(r-1)}, \xi_j)$, a total of $r+1$ roots. □

Corollary 3.4. (i) *If* $b_jc_j \geqslant 0,\ 1 \leqslant j \leqslant k, bc > 0$ *and* a, a_j *are real*, $1 \leqslant j \leqslant k$, *then* $T_n^k(a, b, c)$ *has* $n + k$ *real eigenvalues. Furthermore, these eigenvalues are distinct if* $b_jc_j > 0, 1 \leqslant j \leqslant k$.

(ii) *If* $bc < 0,\ b_jc_j < 0,$ *and* a, a_j *are purely imaginary complex numbers for* $1 \leqslant j \leqslant k$, *then* $T_n^k(a, b, c)$ *has* $n + k$ *distinct eigenvalues.*

Proof. If x_0 is a root of $\phi_n^k(x)$, then $a/\sqrt{bc} - 2x_0 (\equiv \lambda_0)$ is a root of $\phi_n^k(\lambda)$, and thus $a - 2\sqrt{bc}x_0$ is an eigenvalue of $T_n^k(a, b, c)$. If $b_jc_j \geqslant 0, 1 \leqslant j \leqslant k$, $bc > 0$ and a, a_j are real $1 \leqslant j \leqslant k$, then condition (6) is satisfied. If $bc < 0$ and $b_jc_j < 0$, and a, a_j are purely imaginary complex numbers for $1 \leqslant j \leqslant k$, then condition (6) is also satisfied. The result follows from Theorem 3.2. The eigenvalues of $T_n^k(a, b, c)$ are of the form $a - 2\sqrt{bc}x_i$, $1 \leqslant i \leqslant n \leqslant k$, where the x_i's are distinct real roots of $\phi_n^k(x)$. □

In the next section we demonstrate examples of $T_n^k(a,b,c)$ in which the eigenvalues can be determined completely by graphing in the case of $k=1$ or 2.

Next, we will determine a lower bound for the number of real eigenvalues of $T_n^k(a,b,c)$. For the rest of this section, we assume that $bc>0, n \geqslant k$, and $a, a_j, b_j, c_j,\ 1 \leqslant j \leqslant k$, are real. With these assumptions, all entries of $T_n^k(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc})$ and $T_n^k(a,b,c)$ are real and $bc=1$. In order to facilitate an induction argument on k, we rename the entries in $T_n^k(a,b,c)$ by reversing $\{a_1,\ldots,a_k\}$, $\{b_1,\ldots,b_k\}$, $\{c_1,\ldots,c_k\}$ *as* $\{a_k,\ldots,a_1\}$, $\{b_k,\ldots,b_1\}$, $\{c_k,\ldots,c_1\}$, respectively, and thereby write

$$T_n^k(a,b,c) = \left[\begin{array}{c|cccc} T_n(a,b,c) & & & \mathbf{0} & \\ & c_1 & & & \\ \hline b_1 & a_1 & \ddots & & \\ & \ddots & \ddots & \ddots & c_k \\ \mathbf{0} & & & b_k & a_k \end{array}\right]$$

Expanding the determinant for $\phi_n^k(\lambda)$ by the last row we have

$$\phi_n^k(\lambda) = (a_k - \lambda)\phi_n^{k-1}(\lambda) - b_k c_k \phi_n^{k-2}(\lambda), \quad k \geqslant 2. \tag{9}$$

In case $k=1$, we have

$$\phi_n^1(x) = (a_1 - a + 2x)\phi_n^0(x) - b_1 c_1 \phi_{n-1}^0(x)$$

with $a - \lambda = 2x$. Thus, by Eq. (2), we have

$$\begin{aligned} \phi_n^1(x) &= (a_1 - a)U_n(x) + 2xU_n(x) - b_1c_1U_{n-1}(x), \\ &= (a_1 - a)U_n(x) + U_{n+1}(x) + (1 - b_1c_1)U_{n-1}(x). \end{aligned}$$

Thus, $\phi_n^1(x)$ is a linear combination of $U_{n+1}(x)$, $U_n(x)$ and $U_{n-1}(x)$. In general, we show that $\phi_n^k(x)$ is a linear combination of $U_{n+k}(x)$, $U_{n+k-1}(x),\ldots,U_n(x)$, $U_{n-1}(x),\ldots,U_{n-k}(x)$ by induction in k. Suppose that $\phi_n^1(x),\ldots,\phi_n^{k-1}(x)$ satisfy the above assertion. From (9) we have

$$\phi_n^k(x) = (a_k - a + 2x)\phi_n^{k-1}(x) - b_k c_k \phi_n^{k-2}(x),$$

and so

$$\phi_n^k(x) = (a_k - a)\phi_n^{k-1}(x) - b_k c_k \phi_n^{k-2}(x) + 2x\phi_n^{k-1}(x). \tag{10}$$

By the induction hypothesis, the first two terms on the right side of (10) are linear combinations of $U_{n+k-1}(x),\ldots,U_{n-k+1}(x)$. The last term on the right side of (10) is of the form

$$\sum_{n-k+1 \leqslant j \leqslant n+k-1} \alpha_j 2xU_j(x).$$

By (2), $2xU_j(x) = U_{j+1}(x) + U_{j-1}(x)$ and hence

$$\sum_j \alpha_j 2xU_j(x) = \sum_j \alpha_j[U_{j+1}(x) + U_{j-1}(x)] = \sum_{n-k \leqslant j \leqslant n+k} \beta_j U_j(x)$$

with β_j real. Thus, $\phi_n^k(x)$ is a linear combination of $U_{n+k}(x), \ldots, U_{n-k}(x)$. We summarize this result in a proposition.

Proposition 3.5. *$\phi_n^k(x)$ is a linear combination of $U_{n+k}(x)$, $U_{n+k-1}(x), \ldots, U_{n-k}(x)$ with real coefficients.*

Theorem 3.6. *$\phi_n^k(\lambda)$ has at least $n-k$ real roots located in $(a-2, a+2)$.*

Proof. Suppose that $\phi_n^k(x)$ has fewer than $n-k$ real roots in $(-1,1)$, say $\zeta_1 \leqslant \zeta_2 \leqslant \cdots \leqslant \zeta_m$, $m < n-k$. Consider a polynomial $f(x) = \prod_{j\ \ 1}^{m}(x - \zeta_j)$ of degree m on $(-1,1)$. $f(x)$ can be written as a linear combination of $U_0(x), U_1(x), \ldots, U_m(x)$, i.e., $f(x) = \sum_{j\ \ 0}^{m} \alpha_j U_j(x)$. Next consider the weighted inner product

$$\langle \phi_n^k(x), f(x) \rangle \equiv \int_{-1}^{1} (1-x^2)^{-1/2} \phi_n^k(x) f(x)\, \mathrm{d}x,$$

which is nonzero since $\phi_n^k(x)$ and $f(x)$ are of either the same sign or the opposite sign over each of the following intervals $(-1, \zeta_1), (\zeta_1, \zeta_2), \ldots, (\zeta_m, 1)$. On the other hand, it follows from Proposition 3.5 that

$$\phi_n^k(x) = \sum_{n-k \leqslant j \leqslant n+k} \beta_j U_j(x)$$

and hence

$$\langle \phi_n^k, f \rangle = \left\langle \sum_{n-k \leqslant j \leqslant n+k} \beta_j U_j(x), \sum_{j\ \ 0}^{m} \alpha_j U_j(x) \right\rangle = 0, \quad m < n-k,$$

a contradiction since the polynomials $u_j(x)$ are orthogonal with respect to this inner product. Hence, $\phi_n^k(x)$ has at least $n-k$ real roots in $(-1,1)$, and therefore, $\phi_n^k(\lambda)$ has at least $n-k$ real roots in $(a-2, a+2)$. □

4. Examples

For the sake of simplicity, we require $bc > 0$ in this section.

We study the eigenvalues of a matrix

$$T_n^1(a,b,c) = \left[\begin{array}{c|c} T_n(a,b,c) & \begin{matrix}\mathbf{0} \\ b_1\end{matrix} \\ \hline \mathbf{0} \quad c_1 & a_1 \end{array}\right]$$

which corresponds to the case of $k = 1$.

By examining the roots of the characteristic polynomial of $(1/\sqrt{bc})T_n^1(a,b,c)$ and using the substitution $a/\sqrt{bc} - \lambda = 2x$, we get from (3) that, for $n \geqslant 1$, the roots x of $\phi_n^k(a)(x)$ satisfy the equation

$$bc(e_1 + 2x)U_n(x) - b_1c_1U_{n-1}(x) = 0, \tag{11}$$

and if x is not a common root of $U_{n-1}(x)$ and $e_1 + 2x$, then

$$\frac{U_n(x)}{U_{n-1}(x)} = \frac{b_1c_1}{bc(e_1 + 2x)}, \tag{12}$$

where $e_1 = (a_1 - a)/\sqrt{bc}$ and $U_n(x)$ denotes the nth degree Chebyshev polynomial of the second kind. We have seen in Corollary 3.4 that if $b_1c_1 > 0$ and $bc > 0$, $T_n^1(a,b,c)$ has $n + 1$ real distinct eigenvalues, obtained by studying the intersection of the graphs

$$g_1(x) = b_1c_1/bc(e_1 + 2x) \text{ with } p_n(x) = U_n(x)/U_{n-1}(x)$$

in the xy-plane. By looking at the graph of $y = g_1(x)$, we can determine the location of eigenvalues of $T_n^1(a,b,c)$ precisely.

Let $\eta_0 < \xi_1 < \eta_1 < \xi_2 < \cdots < \eta_{i-1} < \xi_i < \eta_i < \cdots < \eta_{n-1} < \xi_n < \eta_n$ where $\eta_0 = -\infty$, $\eta_n = \infty$ and $\xi_1, \xi_2, \ldots, \xi_n$ are the roots of $U_n(x)$ and $\eta_1, \eta_2, \ldots, \eta_{n-1}$ are the roots of $U_{n-1}(x)$. If $bc > 0$ and $(a - a_1)/2\sqrt{bc}$ coincides with one of the η_i's, it is a root of (11). Otherwise, we call the interval (η_{i-1}, η_i) the distinguished interval if $\eta_{i-1} < (a - a_1)/2\sqrt{bc} < \eta_i$. With this notation we have the following result.

Theorem 4.1. *If $bc > 0$ and $\eta_{i-1} < (a - a_1)/2\sqrt{bc} < \eta_i$ for some i, there is exactly one root of* (12) *in each of the $n - 1$ intervals (η_{j-1}, η_j) where $j \neq i, 1 \leqslant j \leqslant n$. If $b_1c_1 > 0$, then there are precisely two additional roots of* (12), *exactly one lying in each of the intervals*

$$\left(\eta_{i-1}, \frac{a - a_1}{2\sqrt{bc}}\right) \quad \textit{and} \quad \left(\frac{a - a_1}{2\sqrt{bc}}, \eta_i\right).$$

If $b_1c_1 < 0$, then there may be zero, one or two additional roots of (12) *in the interval (η_{i-1}, η_i).*

Proof. Let δ_1 and δ_2 be the parts of the graph of $g_1(x)$ for $x < -e_1/2$ and for $x > e_1/2$, respectively. We observe that if $\eta_{i-1} < -e_1/2 < \eta_i$, from Fig. 1, we see that δ_1 meets each component of the graph $y = U_n(x)/U_{n-1}(x)$ once in the $i-2$ intervals on the left of (η_{i-1}, η_1), and δ_2 meets each component in $n-i+1$ intervals once on the right of (η_{i-1}, η_1), producing $n-1$ roots of (12). This holds, if $b_1c_1 > 0$, then $y = g_1(x) = b_1c_1/\sqrt{bc}(e_1 + 2x)$ is decreasing on each interval $(-\infty, -e_1/2)$ and $(-e_1/2, \infty)$ as depicted in Fig. 3; or if $b_1c_1 \leqslant 0$. Now if $b_1c_1 > 0$, the component of the graph of $y = U_n(x)/U_{n-1}(x)$ in the distinguished interval, (η_{i-1}, η_i), meets both δ_1 and δ_2, and we get two additional roots of (12) (see Fig. 1 along with Fig. 3). If $b_1c_1 < 0$, the graph of $y = g_1(x)$ is increasing on $(-\infty, -e_1/2)$ and on $(-e_1/2, \infty)$. With bc, a and a_1 fixed, b_1 and c_1 can be chosen so that $b_1c_1 < 0$ and each of the three illustrations in Fig. 4 occurs. $\square$

Now we study the eigenvalues of a matrix

$$T_n^2(a,b,c) = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{T_n(a,b,c)} & \multicolumn{2}{c}{\mathbf{0}} \\ & & b_2 & 0 \\ \hline \mathbf{0} & c_2 & a_2 & b_1 \\ & 0 & c_1 & a_1 \end{array}\right]$$

which is the case $k = 2$.

By examining the roots of the characteristic polynomial of $(1/\sqrt{bc})T_n^2(a,b,c)$ and using the substitution $a/\sqrt{bc} - \lambda = 2x$, we get from (3), for $n \geqslant 1$,

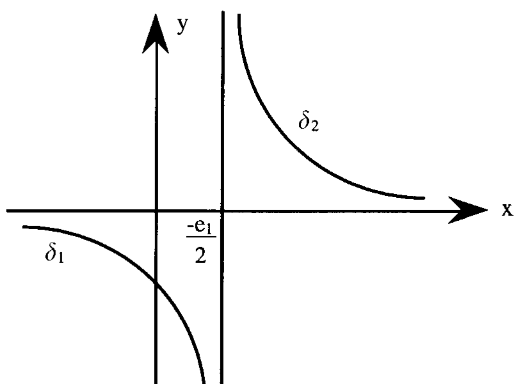

Fig. 3. y $g_1(x)$.

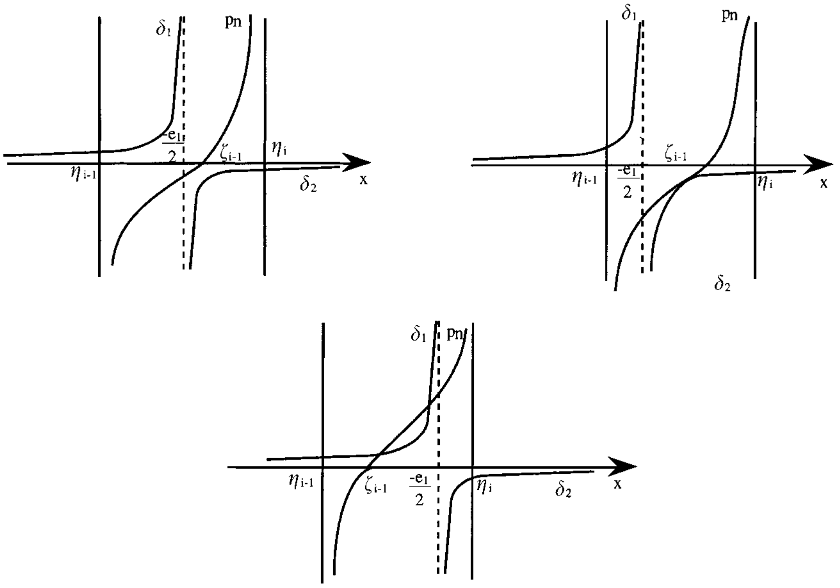

Fig. 4. Intersections when $b_1c_1 < 0$.

$$bc(4x^2 + (e_1 + e_2)2x + e_1e_2 - d)U_n(x) - b_2c_2(e_1 + 2x)U_{n-1}(x) = 0. \qquad (13)$$

If x is not a common root of factors in two summands,

$$\frac{U_n(x)}{U_{n-1}(x)} = \frac{b_2c_2(e_1 + 2x)}{bc(4x^2 + (e_1 + e_2)2x + e_1e_2 - d)}, \qquad (14)$$

where

$$\frac{a_1 - a}{\sqrt{bc}} = e_1, \quad \frac{a_2 - a}{\sqrt{bc}} = e_2, \quad \frac{b_1c_1}{bc} = d$$

and $U_n(x)$ denotes the nth degree Chebyshev polynomial of the second kind. We have seen in Corollary 3.4, that if $b_2c_2 > 0$, $b_1c_1 > 0$ and $bc > 0$, $T_n^2(a, b, c)$ has $n + 2$ real distinct eigenvalues.

With the help of the graph of

$$y = g_2(x) = \frac{b_2c_2(e_1 + 2x)}{bc(4x^2 + (e_1 + e_2)2x + e_1e_2 - d)},$$

we can determine the locations of roots of (14). Let $\eta_0 < \xi_1 < \eta_1 < \cdots < \eta_{i-1} < \xi_i < \eta_i < \cdots < \eta_{n-1} < \xi_n < \eta_n$, where $\eta_0 = -\infty$, $\eta_n = \infty$ and $\xi_1, \xi_2, \ldots, \xi_n$ are the roots of $U_n(\lambda)$ and $\eta_1, \eta_2, \ldots, \eta_{n-1}$ are the roots of $U_{n-1}(\lambda)$. We set

$$\theta_1 = \frac{-(e_1 + e_2) - \sqrt{(e_1 - e_2)^2 + 4d}}{4},$$

$$\theta_2 = \frac{-(e_1 + e_2) + \sqrt{(e_1 - e_2)^2 + 4d}}{4}$$

and

$$\Delta = (e_1 - e_2)^2 + 4d.$$

Note that $x = \theta_1$ and $x = \theta_2$ are vertical asymptotes of $y = g_2(x)$ if $\Delta \geqslant 0$. If $\Delta \geqslant 0$ and θ_1 or θ_2 is a root of $U_{n-1}(x)$, it is a root of (13); otherwise, let us denote by J_1 and J_2 the intervals to which θ_1 and θ_2 belong respectively, amongst intervals (η_{i-1}, η_i) for $i = 1, 2, \ldots, n$. In this notation, we have the following result.

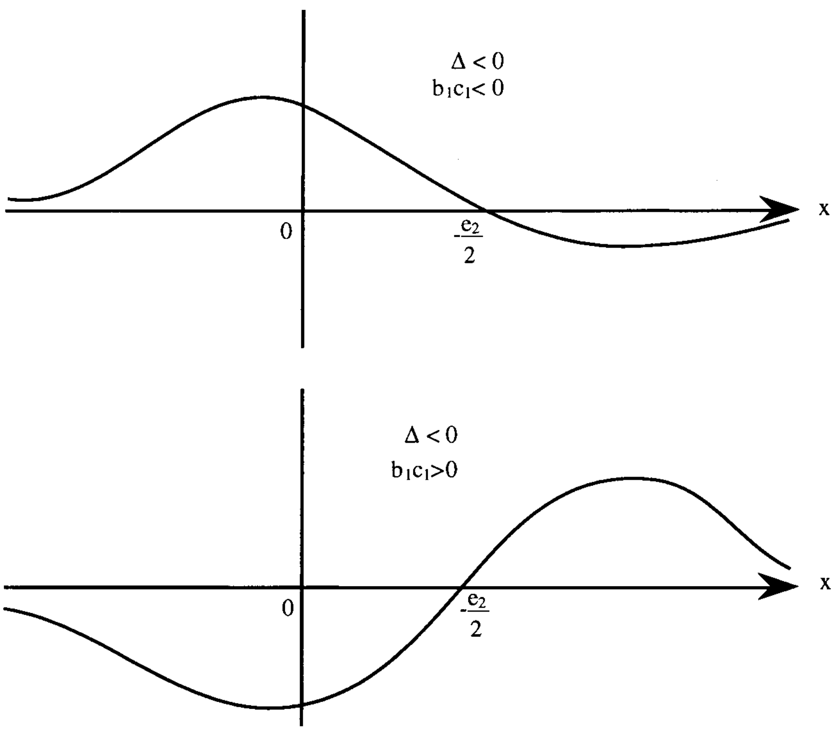

Fig. 5. $g_2(x) \quad \frac{b_2 c_2 (e_1 + 2x)}{bc(4x^2 + (e_1 + e_2)2x + e_1 e_2 \quad d)}$.

Theorem 4.2. *If $\Delta < 0$, the Eq.* (14) *has n real roots, at least one lying in each interval* (η_{j-1}, η_j) *for* $j = 1, 2, \dots, n$.

Proof. The result follows from looking at the graphs given in Fig. 5 comparing them with the graph of $y = U_n(x)/U_{n-1}(x)$ in Fig. 1. □

Theorem 4.3. *Suppose $\Delta > 0$, and θ_1, θ_2 are not roots of $U_{n-1}(x)$.*

(i) *Then Eq.* (15) *has at least one real root in each interval* (η_{j-1}, η_j), *for* $1 \leqslant j \leqslant n$, *which is not distinguished, accounting for at least* $n-2$ *or* $n-1$ *roots of* (14) *depending on whether* $J_1 \neq J_2$ *or* $J_1 = J_2$.

(ii) *If* $b_2c_2 > 0$, $b_1c_1 > 0$, *then there are exactly* $n+2$ *distinct real roots of* (14) *and each nondistinguished interval* (η_{i-1}, η_j), $1 \leqslant i \leqslant n$, *contains exactly one root of* (14). *If* $b_2c_2 > 0$, $b_1c_1 < 0$, *then there are at least n real roots of* (14).

(iii) *If* $b_2c_2 < 0$ *and* $J_1 \neq J_2$, *then there are at least n real roots of* (14) *when* $b_1c_1 < 0$, *and at least* $n-2$ *real roots of* (14), *when* $b_1c_1 > 0$.

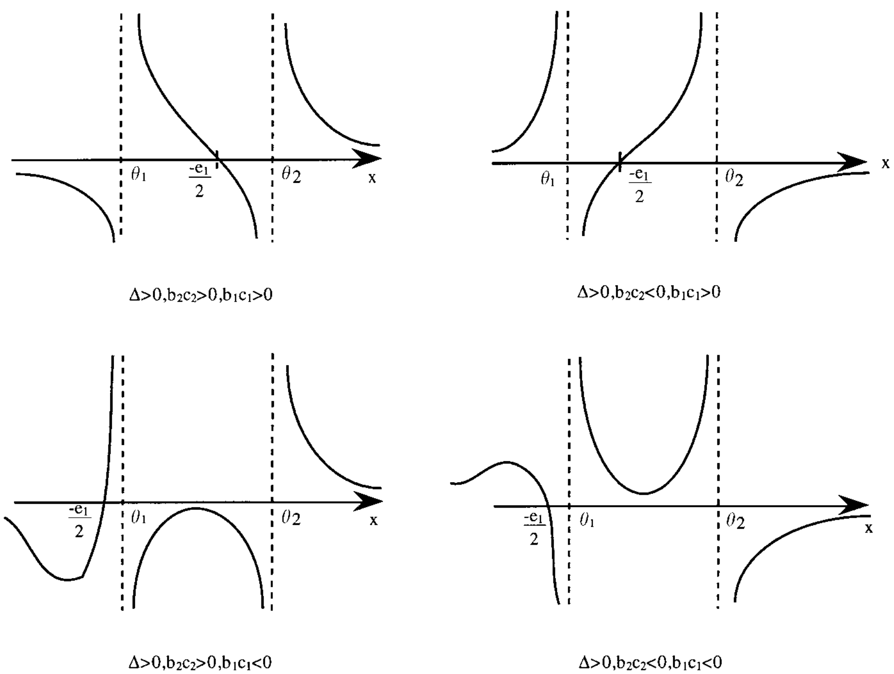

Fig. 6. $y \quad g_2(x), \Delta > 0$.

Proof. (i) Fig. 6 depicts the graph of $y = g_2(x)$ for four cases. From Figs. 1 and 6, it can be seen that $g_2(x)$ and $p_n(x)$ have at least one intersection in any given nondistinguished interval (η_{j-1}, η_j), for some $1 \leqslant j \leqslant n$, for $x = \eta_{j-1}$, and $x = \eta_j$ are vertical asymptotes of $p_n(x)$, and the x-axis is a horizontal asymptote of $g_2(x)$.

(ii) Suppose $b_2c_2 > 0$ and $b_1c_1 > 0$. Then $g_2'(x) < 0$ for all $x \neq \theta_1, \theta_2$. If $J_1 \neq J_2$, then each J_j, j contains two roots accounting for all $2 + 2 + (n-2) = n+2$ roots using (i). If $\sigma_1 = J_2$, then J_1 contains three roots accounting for all $3 + (n-1) = n+2$ roots using (i). In either case, all roots of (14) are accounted for, and thus all nondistinguished intervals contain exactly one root of (14) again. Suppose $b_2c_2 > 0$ and $b_1c_1 > 0$ and $b_1c_1 < 0$. The distinguished interval J_1 or J_2 that contains θ_2 contain two roots (Fig. 6). If $J_1 \neq J_2$, this interval contains at least two roots of (14) accounting for $2 + (n-2) = n$ roots, using (i). If $J_1 = J_2$, J_1 contains at least one root of (14) in $J_1 \cap (\theta_1, \infty)$ accounting for $1 + (n-1) = n$ roots, by (i) again.

(iii) The number of real roots is at least $n-2$ by Theorem 3.6. In addition, if $b_2c_2 < 0$, $b_1c_1 < 0$ and $J_1 \neq J_2$, then it can be seen from Fig. 6 that the distinguished interval that contains θ_1 necessarily contains two roots of (14) accounting for $2 + (n-2) = n$ roots. Finally, if $J_1 = J_2$, then one root of (14) is guaranteed in $J_1 \cap (-\infty, \theta_1)$ of (14), accounting for $1 + (n-1) = n$ roots. □

Theorem 4.4. *If $\Delta = 0$, then there are at least n distinct real roots of* (13).

Proof. The graph of $g_2(x)$ is given in Fig. 7, where $x = \theta_1$ is the vertical asymptote of $g_2(x)$.

If $\Delta = 0$, then $c_1b_1 < 0$. If θ_1 does not lie in the interval $[\eta_{j-1}, \eta_j], 1 \leqslant j \leqslant n$, then Eq. (14) has at least one root in $[\eta_{j-1}, \eta_j]$. If $\theta_1 \in (\eta_{j-1}, \eta_j)$, it can be seen easily from Fig. 7 that only one root is guaranteed in any open interval (η_{j-1}, η_j) that contains θ_1. Note that if θ_1 coincides with η_j, then one of the intervals $[\eta_{j-1}, \eta_j]$ and (η_j, η_{j+1}) necessarily contains a root of (14) while the other might not contain a root. In this case, θ_1 is a root of (13). □

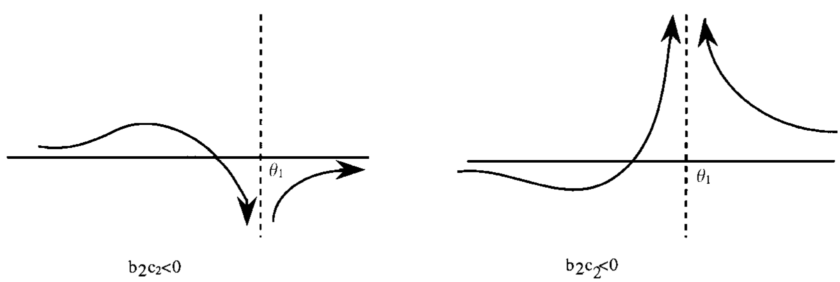

Fig. 7. y $g_2(x), \Delta$ 0.

Remark 4.5. The positions of the real roots discussed in Theorems (4.2) (4.4) can be determined completely by the graphs of $p_n(x)$ *and* $g_2(x)$.

References

[1] K.E. Atkinson, An Introduction to Numerical Analysis, Wiley, New York, 1978.
[2] R.M. Beam, R.F. Warming, The asymptotic spectra of banded Toeplitz and quasi Toeplitz matrices, SIAM J. Sci. Comput. 14 (4) (1993) 971 1006.
[3] B. Cahlon, D.M. Kulkarni, P. Shi, Stepwise stability for the heat equation with a nonlocal constraint, SIAM J. Numer. Anal. 32 (2) (1995) 571 593.
[4] J.M. Franklin, Matrix Theory, Prentice Hall, New Jersey, 1968.
[5] G.H. Golub, C.F. van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1989.
[6] U.W. Hochstrasser, Orthogonal polynomials, in: M. Abramowitz, I.A. Stegun, (Eds.), Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, 1966, pp. 771 802.
[7] R.A. Horn, C.A. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[8] A. Korsak, C. Schubert, A determinant expression of Chebychev polynomials, Canad. Math. Bull. 3 (1960) 243 246.
[9] T.J. Rivlin, Chebychev Polynomials, Wiley, New York, 1974.
[10] I.E. Schochetman, R.L. Smith, S. K. Tsui, in: A.V. Fiacco (Ed.), Solutions existence for infinite quadratic programming, the positive semi definite case, to appear in Mathematical Programming with Data Perturbations, Marcel Dekker, 1997.
[11] G.D. Smith, Numerical Solution of Partial Differential Equations: Finite Difference Methods, 3rd ed., Clarendon Press, Oxford, 1985.
[12] W.F. Trench, On the eigenvalue problem for Toeplitz band matrices, Linear Algebra Appl. 64 (1985) 199 214.
[13] S. K. Tsui, Matrices related to orthogonal polynomials, Preprint 1995.