JOURNAL OF APPROXIMATION THEORY **51**, 254–266 (1987)

Representations and Bounds for Zeros of Orthogonal Polynomials and Eigenvalues of Sign-Symmetric Tri-diagonal Matrices

ERIK A. VAN DOORN

Faculty of Applied Mathematics, University of Twente, Enschede, The Netherlands

Communicated by R. Bojanic

Received July 2, 1985

We consider a sequence $\{P_n\}$ of orthogonal polynomials defined by a three-term recurrence formula. Representations and bounds are derived for the endpoints of the smallest interval containing the (real and distinct) zeros of P_n in terms of the parameters in the recurrence relation. These results are brought to light by viewing $(-1)^n P_n$ as the characteristic polynomial of a sign-symmetric tri-diagonal matrix of order n. Our findings are subsequently used to obtain new proofs for a number of bounds on the endpoints of the true and limit intervals of orthogonality for the sequence $\{P_n\}$.

1. INTRODUCTION

We are concerned with the zeros of polynomials Q_n satisfying a recurrence relation with real coefficients

$$Q_n(x) = (\alpha_n x - \beta_n)\, Q_{n-1}(x) - \gamma_n Q_{n-2}(x), \qquad n = 1, 2, \ldots \tag{1.1}$$

where $Q_{-1}(x) = 0$, $Q_0(x) = \alpha_0 \neq 0$ (α_0 real) and $\alpha_{n-1}\alpha_n\gamma_n > 0$ ($n > 1$). Letting $P_n(x) \equiv (\alpha_0\alpha_1 \cdots \alpha_n)^{-1} Q_n(x)$, $c_n \equiv \alpha_n^{-1}\beta_n$ and $\lambda_n \equiv (\alpha_{n-1}\alpha_n)^{-1}\gamma_n$, it is seen that

$$P_n(x) = (x - c_n)\, P_{n-1}(x) - \lambda_n P_{n-2}(x), \qquad n = 1, 2, \ldots \tag{1.2}$$

where $P_{-1}(x) = 0$, $P_0(x) = 1$. So without loss of generality we can take the simpler recurrence formula (1.2), where c_n is real and $\lambda_{n+1} > 0$ ($n > 0$), as a starting point for our analysis. Note that the value of λ_1 is irrelevant; it will be convenient, however, to assume $\lambda_1 = 1$.

It is well known (see [3]) that $P_n(x)$ has n real, distinct zeros $x_{n1} < x_{n2} < \cdots < x_{nn}$. Moreover, the zeros of $P_n(x)$ and $P_{n+1}(x)$ interlace, that is,

$$x_{n+1,i} < x_{ni} < x_{n+1,i+1}, \qquad i = 1, 2, \ldots, n. \tag{1.3}$$

0021-9045/87 $3.00

Hence the limits

$$\xi_i \equiv \lim_{n \to \infty} x_{ni} \qquad \text{and} \qquad \eta_j \equiv \lim_{n \to \infty} x_{n,n-j+1} \tag{1.4}$$

exist, where we allow of $-\infty$ and $+\infty$, respectively. The quantities ξ_1 and η_1 are of particular interest, since they are the endpoints of the *true interval of orthogonality* for $\{P_n\}$: the smallest interval containing the support of a mass distribution with respect to which the polynomials P_n are orthogonal.

It is evident from (1.3) and (1.4) that

$$\xi_i \leqslant \xi_{i+1} < \eta_{j+1} \leqslant \eta_j, \tag{1.5}$$

so that

$$\sigma \equiv \lim_{i \to \infty} \xi_i \qquad \text{and} \qquad \tau \equiv \lim_{j \to \infty} \eta_j \tag{1.6}$$

exist, again allowing of $\pm\infty$. The quantities σ and τ are also of interest, since they are, if not $\sigma = \tau = \pm\infty$, the endpoints of the *limit interval of orthogonality* for $\{P_n\}$: the smallest interval containing the limit points of the support of a mass distribution with respect to which the polynomials P_n are orthogonal. We will have use for other representations for σ and τ. Namely, let $\xi_1^{(k)}$ and $\eta_1^{(k)}$ denote the endpoints of the true interval of orthogonality for the polynomials $P_n^{(k)}$ which are determined through the recurrence formula (1.2) by the sequences $\{c_n^{(k)} \equiv c_{n+k}\}_{n=1}^{\infty}$ and $\{\lambda_n^{(k)} \equiv \lambda_{n+k}\}_{n=2}^{\infty}$. Then [3, Theorem III.4.2]

$$\xi_1^{(k)} \leqslant \xi_1^{(k+1)} < \eta_1^{(k+1)} \leqslant \eta_1^{(k)}, \qquad k = 0, 1, \ldots \tag{1.7}$$

and [5]

$$\lim_{k \to \infty} \xi_1^{(k)} = \sigma, \qquad \lim_{k \to \infty} \eta_1^{(k)} = \tau. \tag{1.8}$$

Our foremost aim is to obtain information on the interval $[x_{n1}, x_{nn}]$, the smallest interval containing the zeros of P_n, in terms of the parameters defining P_n. That is, we will look for representations and bounds for x_{n1} and x_{nn} in terms of c_i and λ_i $(i = 1, 2, \ldots, n)$. Actually, without loss of generality we can confine attention to the point x_{n1}. Namely, if $x_{n1} < x_{n2} < \cdots < x_{nn}$ are the zeros of $P_n(x)$, then $-x_{nn} < -x_{n,n-1} < \cdots < -x_{n1}$ are the zeros of $P_n(-x)$. Furthermore, it is readily seen that the polynomials $\bar{P}_n(x) = (-1)^n P_n(-x)$ satisfy a recurrence relation of the type (1.2) with parameters $\bar{c}_i = -c_i$ and $\bar{\lambda}_i = \lambda_i$. It follows that a lower (upper) bound for x_{n1} yields an upper (lower) bound for x_{nn}, and vice versa, simply by reversing the sign of the c_i's in the pertinent bound and the sign of the bound itself.

Once we have representations and bounds for x_{n1} at our disposal it is of course easy to derive similar results for ξ_1 and σ (and, via the procedure outlined above, for η_1 and τ) by virtue of (1.4) and (1.8). The second objective of this paper is to show that many bounds for ξ_1 and σ that were derived in the past by various techniques (see [3, 5] and the references mentioned there) can be obtained in this way.

Our approach to generate representations and bounds for x_{n1} is based upon the observation in Section 2 that $(-1)^n P_n(x)$ can be interpreted as the characteristic polynomial of a sign-symmetric tri-diagonal matrix, so that, actually, x_{n1} is the smallest eigenvalue of such a matrix. In Sections 3–5 various ways to exploit this observation are elaborated.

2. Orthogonal Polynomials and Tri-Diagonal Matrices

Suppose we are given sequences of real numbers $\{a_i\}_{i=2}^{\infty}$, $\{b_i\}_{i=2}^{\infty}$ and $\{c_i\}_{i=1}^{\infty}$ with the property $\operatorname{sign}(a_i)=\operatorname{sign}(b_i)$. With these numbers we form the tri-diagonal matrices

$$T_n \equiv \begin{pmatrix} c_1 & b_2 & & & & \\ a_2 & c_2 & b_3 & & 0 & \\ & a_3 & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & 0 & & \cdot & \cdot & b_n \\ & & & & a_n & c_n \end{pmatrix}, \qquad n=1, 2, ..., \tag{2.1}$$

and we ask for the eigenvalues of T_n. If $a_i=b_i=0$ for some $i\leqslant n$, then determination of the eigenvalues of T_n reduces to determination of the eigenvalues of two sign-symmetric tri-diagonal matrices of lower order, so there is no loss of generality in confining attention to the case $a_i b_i>0$ for all i. Now writing $\lambda_i \equiv a_i b_i$ and expanding $\det(T_n - xI_n)$ by its last row, it is readily verified that

$$\det(T_n - xI_n) = (-1)^n P_n(x), \qquad n=1, 2, ...,$$

where the P_n are the polynomials of (1.2). Here I_n denotes the $n\times n$ identity matrix. (Incidentally, this is yet another way to prove that the eigenvalues of a sign-symmetric tri-diagonal matrix are real and distinct; cf. [7] and the Solution to Problem 80–4, *SIAM Rev.* **23** (1981), p. 112.) Thus our original problem of finding representations and bounds for x_{n1}, the smallest zero of the polynomial $P_n(x)$ defined by (1.2), can be phrased as the problem of finding representations and bounds for the smallest eigenvalue of the matrix T_n of (2.1), where a_i and b_i are any real numbers such that $a_i b_i = \lambda_i > 0$ ($i = 2, 3, \ldots$).

The connection between orthogonal polynomials and *symmetric* tri-diagonal matrices is well known (see, e.g., [11, Sects. 7–8]). It would seem that our generalization to sign-symmetric tri-diagonal matrices is of no consequence, since, in determining the eigenvalues of T_n, one might as well choose $a_i = b_i = \pm\lambda_i^{1/2}$, which makes T_n symmetric. We shall see, however, that certain bounds for the eigenvalues of T_n do depend on the particular values chosen for a_i and b_i, so that it is advantageous not to specify these values for the present.

Three approaches to obtain information on the smallest eigenvalue x_{n1} of T_n, and hence on the quantities ξ_1 and σ, will be discussed in the following three sections. Throughout n will be a fixed but otherwise arbitrary natural number.

3. Gerschgorin Discs

We will assume that the numbers a_i and b_i of (2.1) satisfy

$$a_i = -\chi_i^{-1}\lambda_i, \quad b_i = -\chi_i, \qquad i = 2, 3, \ldots, n, \tag{3.1}$$

where $\chi_2, \chi_3, \ldots, \chi_n$ are positive numbers. From Gerschgorin's Theorem [9, Sect. 2.2.1] it then follows immediately that

$$x_{n1} \geq \min_{1 \leq i \leq n} \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}, \tag{3.2}$$

where $\chi_1 = \infty$, $\chi_{n+1} = 0$. Moreover, if we choose

$$\chi_i = -P_{i-1}(x_{n1})/P_{i-2}(x_{n1}) \tag{3.3}$$

($i = 2, 3, \ldots, n$), which is positive by virtue of (1.2) and (1.3), then equality is readily seen to prevail in (3.2). Summarizing, we can state the following theorem, which was obtained independently by Gilewicz and Leopold [12] (in terms of x_{nn}) using quite a different argument.

THEOREM 1. *With $\chi \equiv \{\chi_1, \chi_2, \ldots, \chi_{n+1}\}$ ranging over all sequences such that $\chi_1 = \infty$, $\chi_{n+1} = 0$ and $\chi_i > 0$ $(1 < i \leq n)$ one has*

$$x_{n1} = \max_{\chi} \{ \min_{1 \leq i \leq n} \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}\}. \tag{3.4}$$

The next corollary of Theorem 1 is originally due to Leopold [8].

COROLLARY 1.1. *For any real number* ϕ *such that* $\phi \leqslant c_1$ *and* $\phi < c_i$ $(i=2, 3, ..., n)$, *one has*

$$x_{n1} \geqslant \min_{1<i\leqslant n} \{\phi - (c_i - \phi)^{-1} \lambda_i\}. \tag{3.5}$$

Proof. Defining $\chi_1 = \infty$, $\chi_{n+1} = 0$ and $\chi_i = (c_i - \phi)^{-1} \lambda_i$ $(i=2, 3, ..., n)$, we have

$$c_i - \chi_i^{-1}\lambda_i - \chi_{i+1} = \begin{cases} c_1 - (c_2 - \phi)^{-1} \lambda_2, & i=1 \\ \phi - (c_{i+1} - \phi)^{-1} \lambda_{i+1}, & i = 2, 3, ..., n-1 \\ \phi, & i = n, \end{cases}$$

so that the result is implied by Theorem 1. ∎

By choosing $\chi_i = \lambda_i^{1/2}$ $(i=2, 3, ..., n)$ we obtain a lower bound for x_{n1} which is well known [1]. (Here and elsewhere δ_{ij} denotes Kronecker's delta.)

COROLLARY 1.2. *One has*

$$x_{n1} \geqslant \min_{1\leqslant i\leqslant n} \{c_i - (1-\delta_{i1})\lambda_i^{1/2} - (1-\delta_{in})\lambda_{i+1}^{1/2}\}. \tag{3.6}$$

As for ξ_1 and σ Theorem 1 leads to the following representation theorems, which were obtained earlier in [5] by studying oscillatory behaviour of solutions of certain second order linear difference equations.

THEOREM 2. *With* $\boldsymbol{\chi} \equiv \{\chi_1, \chi_2, ...\}$ *ranging over all infinite sequences such that* $\chi_1 = \infty$ *and* $\chi_i > 0$ $(i>1)$ *one has*

$$\xi_1 = \max_{\boldsymbol{\chi}} \{\inf_{i\geqslant 1} \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}\}. \tag{3.7}$$

Proof. From (1.4) and (3.4) it follows that for any sequence $\boldsymbol{\chi}$

$$\xi_1 \geqslant \inf_{i\geqslant 1} \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}. \tag{3.8}$$

If $\xi_1 = -\infty$ we obviously have equality in (3.8). To show that equality may be obtained if $\xi_1 > -\infty$ one should take

$$\chi_i = -P_{i-1}(\xi_1)/P_{i-2}(\xi_1) \tag{3.9}$$

$(i=2, 3, ...)$, which is positive in view of (1.2)–(1.4), and use the recurrence relation (1.2). ∎

THEOREM 3. *With $\chi \equiv \{\chi_1, \chi_2, \ldots\}$ ranging over all infinite sequences such that $\chi_1 = \infty$ and $\chi_i > 0$ $(i > 1)$ one has*

$$\sigma = \sup_{\chi} \{ \lim_{i \to \infty} \inf \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}\}. \tag{3.10}$$

Proof. Suppose χ is such that

$$\sigma < \lim_{i \to \infty} \inf \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}.$$

Then, for k sufficiently large,

$$\sigma < \inf_{i > k} \{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\} = \inf_{i \geq 1} \{c_i^{(k)} - \chi_{k+i}^{-1}\lambda_i^{(k)} - \chi_{k+i+1}\}.$$

This, however, contradicts (3.7) interpreted for $\xi_1^{(k)}$, since $\xi_1^{(k)} \leq \sigma$. Consequently,

$$\sigma \geq \lim_{i \to \infty} \inf\{c_i - \chi_i^{-1}\lambda_i - \chi_{i+1}\}. \tag{3.11}$$

If $\sigma = -\infty$ we clearly have equality in (3.11). If $\sigma > -\infty$, and hence $\xi_1^{(k)} > -\infty$ (see (1.7) and [3, Theorem II.4.6]), then the right-hand side of (3.11) can be made arbitrarily close to σ. Namely, consider the sequences $\chi^{(k)} \equiv \{\chi_1^{(k)}, \chi_2^{(k)}, \ldots\}$ $(k = 1, 2, \ldots)$, where $\chi_1^{(k)} = \infty$,

$$\chi_i^{(k)} = -P_{i-k-1}^{(k)}(\xi_1^{(k)})/P_{i-k-2}^{(k)}(\xi_1^{(k)}) \tag{3.12}$$

for $i > k+1$ and $\chi_i^{(k)}$ is positive but otherwise arbitrary for $2 \leq i \leq k+1$. Substitution of $\chi^{(k)}$ in the right-hand side of (3.11) makes it equal to $\xi_1^{(k)}$. The statement follows since $\xi_1^{(k)} \uparrow \sigma$ as $k \to \infty$. ∎

We refer to [5] for a survey of lower bounds for ξ_1 and σ that have appeared in the literature and may be conceived as corollaries to Theorems 2 and 3.

4. OVALS OF CASSINI

A theorem of Brauer's [9, Sect. 2.4.2] states that for any $n \times n$ matrix $A \equiv (a_{ij})$ each eigenvalue lies in at least one of the $n(n-1)/2$ *ovals of Cassini*,

$$|z - a_{ii}|\,|z - a_{jj}| \leq P_i P_j, \qquad i, j = 1, 2, \ldots, n;\ i \neq j, \tag{4.1}$$

where $P_i \equiv \sum_{j \neq i} |a_{ij}|$. If A is a tri-diagonal matrix a stronger result is available, however. First note the following.

THEOREM 4. *Let $A \equiv (a_{ij})$ be an $n \times n$ tri-diagonal matrix. If*

$$|a_{ii}|\,|a_{i+1,i+1}| > P_i P_{i+1}, \qquad i = 1, 2, ..., n-1, \tag{4.2}$$

then $\det(A) \neq 0$.

Proof. Suppose that $\det(A) = 0$. Then there is a vector $\mathbf{x} \equiv (x_1, x_2, ..., x_n) \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$. For convenience we define $x_0 = x_{n+1} = 0$. Now let r be such that

$$|x_r| = \max_{1 \leqslant i \leqslant n} \{|x_i|\},$$

and let $s \in \{r-1, r+1\}$ be such that

$$|x_s| = \max\{|x_{r-1}|, |x_{r+1}|\}.$$

Clearly, $|x_r| > 0$ since $\mathbf{x} \neq \mathbf{0}$. Moreover, $|x_s| > 0$ because the opposite conclusion together with $A\mathbf{x} = \mathbf{0}$ would imply $x_r = 0$. We then have

$$|a_{rr} x_r| = \left| -\sum_{j \neq r} a_{rj} x_j \right| \leqslant |x_s|\, P_r$$

$$|a_{ss} x_s| = \left| -\sum_{j \neq s} a_{sj} x_j \right| \leqslant |x_r|\, P_s,$$

whence $|a_{rr}|\,|a_{ss}| \leqslant P_r P_s$. Since $|r - s| = 1$, it follows that (4.2) does not hold. ∎

As a simple corollary we have the result that we were referring to.

COROLLARY 4.1. *An eigenvalue of the $n \times n$ tri-diagonal matrix $A \equiv (a_{ij})$ lies in at least one of the $n-1$ ovals*

$$|z - a_{ii}|\,|z - a_{i+1,i+1}| \leqslant P_i P_{i+1}, \qquad i = 1, 2, ..., n-1. \tag{4.3}$$

Back to our original context we let a_i and b_i be as in (3.1) with χ_i being a positive number. From Corollary 4.1 we then get after some algebra

$$x_{n1} \geqslant \min_{1 \leqslant i < n} \tfrac{1}{2}\{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + 4t_i)^{1/2}\}, \tag{4.4}$$

where

$$t_i = (\chi_{i+1} + \chi_i^{-1}\lambda_i)(\chi_{i+2} + \chi_{i+1}^{-1}\lambda_{i+1}), \tag{4.5}$$

and $\chi_1 = \infty$, $\chi_{n+1} = 0$. Moreover, substitution of (3.9) in (4.5) is readily seen to yield equality in (4.4), so that we have a new representation theorem for x_{n1}. It will be convenient to reformulate this result. Namely, we observe that

$$t_i = \lambda_{i+1}\left(\left(1 - \frac{\lambda_i}{\lambda_i + \chi_i\chi_{i+1}}\right)\left(\frac{\lambda_{i+1}}{\lambda_{i+1} + \chi_{i+1}\chi_{i+2}}\right)\right)^{-1},$$

so that t_i is dependent only on the products $\chi_i\chi_{i+1}$ and $\chi_{i+1}\chi_{i+2}$. Now defining

$$h_i = \lambda_i(\lambda_i + \chi_i\chi_{i+1})^{-1}, \qquad i = 1, 2, ..., n, \tag{4.6}$$

it follows that

$$t_i = \lambda_{i+1}((1 - h_i)\, h_{i+1})^{-1}, \tag{4.7}$$

where $h_1 = 0$, $h_n = 1$ and $0 < h_i < 1$ $(1 < i < n)$. On the other hand, any such sequence $\{h_1, h_2, ..., h_n\}$ uniquely determines the products $\chi_i\chi_{i+1}$ $(1 \leqslant i \leqslant n)$ via (4.6), so our representation theorem may be stated as follows.

THEOREM 5. *With* $\mathbf{h} \equiv \{h_1, h_2, ..., h_n\}$ *ranging over all sequences such that* $h_1 = 0$, $h_n = 1$ *and* $0 < h_i < 1$ $(1 < i < n)$, *one has for* $n > 1$

$$x_{n1} = \max_{\mathbf{h}} \left\{ \min_{1 \leqslant i < n} \frac{1}{2} \left\{ c_i + c_{i+1} - \left((c_{i+1} - c_i)^2 + \frac{4\lambda_{i+1}}{(1 - h_i)\, h_{i+1}} \right)^{1/2} \right\} \right\}. \tag{4.8}$$

If $n = 2$ we can write down the exact value of $x_{n1} = x_{21}$ directly from Theorem 5. If $n > 2$ a simple lower bound for x_{n1} is obtained by taking $h_i = \frac{1}{2}$ $(1 < i < n)$.

COROLLARY 5.1. *For* $n > 2$ *one has*

$$x_{n1} \geqslant \min_{1 \leqslant i < n} \tfrac{1}{2} \{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + e_i\lambda_{i+1})^{1/2}\}, \tag{4.9}$$

where $e_i = 8$ *if* $i = 1, n - 1$ *and* $e_i = 16$ *otherwise.*

The representation (4.8) (as compared with (4.4)) shows to full advantage when we use the result to obtain representations and bounds for ξ_1 and σ. Before doing so we must introduce the following concept. An infinite sequence $\boldsymbol{\beta} \equiv \{\beta_1, \beta_2, ...\}$ is called a *chain sequence* if each β_i admits a representation of the form

$$\beta_i = (1 - g_i)\, g_{i+1}, \tag{4.10}$$

where $0 \leqslant g_1 < 1$ and $0 < g_i < 1$ $(i > 1)$. The sequence $\mathbf{g} \equiv \{g_1, g_2, ...\}$ is then called a *parameter sequence* for $\boldsymbol{\beta}$. Of interest to us is the fact that if $\boldsymbol{\beta}$ is a chain sequence, then there exists a parameter sequence $\mathbf{h} \equiv \{h_1, h_2, ...\}$ for $\boldsymbol{\beta}$

(the *minimal parameter sequence*) with the property $h_1 = 0$ (see [3, Sect. III.5] for a proof and other useful properties of chain sequences).

We are now ready to state the representation theorems for ξ_1 and σ that are implied by Theorem 5. These results are largely due to Chihara [3].

THEOREM 6. *With $\boldsymbol{\beta} \equiv \{\beta_1, \beta_2, ...\}$ ranging over all chain sequences, one has*

$$\xi_1 = \max_{\boldsymbol{\beta}} \{\inf_{i \geq 1} \tfrac{1}{2}\{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + 4\lambda_{i+1}/\beta_i)^{1/2}\}\}. \tag{4.11}$$

Proof. The inequality that is implied by (4.11) for any fixed $\boldsymbol{\beta}$ follows directly from Theorem 5. If $\xi_1 > -\infty$ one should choose

$$\beta_i = \lambda_{i+1}(c_{i+1} - \xi_1)^{-1}(c_i - \xi_1)^{-1}, \tag{4.12}$$

which yields a chain sequence by [3, Theorem IV.2.1], to obtain equality. ∎

A simple corollary to Theorem 6 is the analogue to Corollary 5.1, obtained by letting $\beta_1 = \frac{1}{2}$, $\beta_i = \frac{1}{4}$ $(i > 1)$.

COROLLARY 6.1. *One has*

$$\xi_1 \geq \inf_{i \geq 1} \tfrac{1}{2}\{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + e_i\lambda_{i+1})^{1/2}\}, \tag{4.13}$$

where $e_i = 8$ if $i = 1$ and $e_i = 16$ otherwise.

THEOREM 7. *With $\boldsymbol{\beta} \equiv \{\beta_1, \beta_2, ...\}$ ranging over all chain sequences, one has*

$$\sigma = \sup_{\boldsymbol{\beta}} \{\lim_{i \to \infty} \inf \tfrac{1}{2}\{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + 4\lambda_{i+1}/\beta_i)^{1/2}\}\}. \tag{4.14}$$

Proof. Considering that $\{\beta_{k+i}\}_{i=1}^{\infty}$ is a chain sequence if $\{\beta_i\}_{i=1}^{\infty}$ is a chain sequence, it follows readily from (1.7), (1.8) and (4.11) that for any chain sequence $\boldsymbol{\beta}$

$$\sigma \geq \lim_{i \to \infty} \inf \tfrac{1}{2}\{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + 4\lambda_{i+1}/\beta_i)^{1/2}\} \tag{4.15}$$

(cf. the proof of Theorem 3). If $\sigma = -\infty$ we obviously have equality in (4.15), so let us assume $\sigma > -\infty$. Now consider a sequence of points $y^{(k)}$, $k = 1, 2, ...$, such that $y^{(k)} \uparrow \sigma$ as $k \to \infty$ and $y^{(k)} < \sigma$ for all k (if $\xi_1^{(k)} < \sigma$ for all k, then these numbers will do). Subsequently, consider a sequence of chain sequences $\boldsymbol{\beta}^{(k)}$, $k = 1, 2, ...$, which are such that

$$\beta_i^{(k)} = \lambda_{i+1}(c_{i+1} - y^{(k)})^{-1}(c_i - y^{(k)})^{-1}$$

for i sufficiently large. Using [3, Theorem IV.3.2 and the Corollary to Theorem III.5.5] it is not difficult to show that such chain sequences exist. Substitution of $\boldsymbol{\beta}^{(k)}$ in the right-hand side of (4.15) makes it equal to $y^{(k)}$. It follows that the right-hand side of (4.15) can be made arbitrarily close to σ by choosing $\boldsymbol{\beta} = \boldsymbol{\beta}^{(k)}$ with k sufficiently large. ▮

We refer to [3, 4] for bounds on σ that can be obtained from Theorem 7.

5. The Field of Values

In this final section we shall assume

$$a_i = b_i = -\lambda_i^{1/2}, \qquad i = 2, 3, \dots, \tag{5.1}$$

which renders the matrix T_n of (2.1) symmetric. It then follows from [9, Sects. 5.2.2 and 5.2.6] (see also [2, Theorem 7.2]) that the interval $[x_{n1}, x_{nn}]$ is precisely the *field of values* of T_n, i.e., the set of numbers of the form $\mathbf{y}T_n\mathbf{y}'$, where $\mathbf{y} \equiv (y_1, y_2, \dots, y_n)$ and $\mathbf{y}\mathbf{y}' = \sum y_i^2 = 1$. We note that

$$\mathbf{y}T_n\mathbf{y}' = \sum_{i=1}^{n} (y_i^2 c_i - 2y_{i-1} y_i \lambda_i^{1/2}), \tag{5.2}$$

where $y_0 = 0$. Moreover, in determining the minimum of (5.2) over all n-dimensional unit vectors $\mathbf{y}$, we can evidently restrict ourselves to $\mathbf{y}$ in the non-negative orthant. So, generalizing a result of Freud's [13] (who phrased it in terms of x_{nn}), we can conclude the following.

Theorem 8. *With $\boldsymbol{\theta} \equiv \{\theta_0, \theta_1, \dots, \theta_n\}$ ranging over all sequences of non-negative real numbers such that $\theta_0 = 0$ and $\sum \theta_i = 1$, one has*

$$x_{n1} = \min_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{n} (\theta_i c_i - 2(\theta_{i-1}\theta_i\lambda_i)^{1/2}) \right\}. \tag{5.3}$$

It is known (see, e.g., [10]), but also easy to observe from (5.2), that the unit eigenvectors of T_n yield the stationary points of $\mathbf{y}T_n\mathbf{y}'$. Now if $\mathbf{y}$ is a unit eigenvector, then $\mathbf{y}T_n\mathbf{y}'$ equals the corresponding eigenvalue, so the eigenvalues of T_n are precisely the (local) extrema of $\mathbf{y}T_n\mathbf{y}'$, while the global minimum x_{n1} of $\mathbf{y}T_n\mathbf{y}'$ is attained by the unit eigenvectors corresponding to the smallest eigenvalue x_{n1}. In fact, since the eigenvalues of T_n are simple, the space of eigenvectors corresponding to a particular eigenvalue is one dimensional. Hence, having observed already that the minimum over all unit vectors of $\mathbf{y}T_n\mathbf{y}'$ is attained by some vector in the non-negative orthant, we can conclude that for any sequence

$\boldsymbol{\theta} \equiv \{\theta_0, \theta_1, ..., \theta_n\}$ of non-negative numbers such that $\theta_0 = 0$ and $\sum \theta_i = 1$, one has

$$x_{n1} \leqslant \sum_{i=1}^{n} (\theta_i c_i - 2(\theta_{i-1}\theta_i\lambda_i)^{1/2}) \tag{5.4}$$

with equality prevailing for precisely one sequence $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} \equiv \{\hat{\theta}_0, \hat{\theta}_1, ..., \hat{\theta}_n\}$. It is not difficult to show that, actually,

$$\hat{\theta}_i = p_{i-1}^2(x_{n1}) \left(\sum_{j=0}^{n-1} p_j^2(x_{n1}) \right)^{-1} \tag{5.5}$$

$(1 \leqslant i \leqslant n)$, where

$$p_i(x) = (\lambda_1 \lambda_2 \cdots \lambda_{i+1})^{-1/2} P_i(x), \qquad i = 0, 1, ... \tag{5.6}$$

By virtue of (1.3), $\hat{\theta}_i$ is positive for $i = 1, 2, ..., n$. It follows in particular that the inequality in (5.4) is strict if some of the θ_i $(i > 0)$ are zero. Below we give two corollaries to Theorem 8, of which the first one is known (see, e.g., [6]).

Corollary 8.1. *For $n \geqslant 1$ one has*

$$x_{n1} \leqslant \min_{1 \leqslant i \leqslant n} \{c_i\} \tag{5.7}$$

and the inequality is strict if and only if $n > 1$.

Proof. Letting $\theta_j = \delta_{ij}$, $j = 0, 1, ..., n$, we obtain from (5.4) that $x_{n1} \leqslant c_i$ for all $i = 1, 2, ..., n$. Clearly, $x_{1,1} = c_1$, so the last assertion follows from the assertion below (5.6) and the fact that $\theta_j = 0$ for some $j > 0$ if $n > 1$. ∎

Corollary 8.2. *For $n \geqslant 2$ one has*

$$x_{n1} \leqslant \min_{1 \leqslant i < n} \tfrac{1}{2} \{c_i + c_{i+1} - ((c_{i+1} - c_i)^2 + 4\lambda_{i+1})^{1/2}\}, \tag{5.8}$$

and the inequality is strict if and only if $n > 2$.

Proof. Choosing $\boldsymbol{\theta}$ in (5.4) such that

$$\theta_i = \tfrac{1}{2} - \tfrac{1}{2}(c_i - c_{i+1})((c_{i+1} - c_i)^2 + 4\lambda_{i+1})^{-1/2}, \qquad \theta_{i+1} = 1 - \theta_i \tag{5.9}$$

and $\theta_j = 0$ for $j \neq i$, $i+1$ $(i = 1, 2, ..., n-1)$ leads readily to (5.8). Since $x_{2,1} = \frac{1}{2}(c_1 + c_2 - ((c_2 - c_1)^2 + 4\lambda_2)^{1/2})$ and $\theta_j = 0$ for some $j > 0$ if $n > 2$, equality previals if and only if $n = 2$. ∎

We note that the particular choice for $\boldsymbol{\theta}$ in (5.9) results from minimizing the right-hand side of (5.4) on the condition that $\theta_j = 0$ for $j \neq i$, $i+1$.

From (1.4) we have $x_{n1} \to \xi_1$ $(n \to \infty)$, so Theorem 8 immediately leads to a representation formula for ξ_1. As before we will reformulate the result in such a way that bounds for ξ_1 can be easily obtained from it.

THEOREM 9. *With* $\boldsymbol{\theta} \equiv \{\theta_0, \theta_1, ...\}$ *ranging over all infinite sequences of non-negative numbers such that* $\theta_0 = 0$ *and* $\sum \theta_i = 1$, *one has*

$$\xi_1 = \inf_{\boldsymbol{\theta}} \left\{ \lim_{n \to \infty} \inf \left\{ \sum_{i=1}^{n} (\theta_i c_i - 2(\theta_{i-1} \theta_i \lambda_i)^{1/2}) \right\} \right\}. \tag{5.10}$$

Proof. Let $\boldsymbol{\theta}$ be a fixed but otherwise arbitrary sequence. We let

$$s_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \theta_i$$

and we denote by $f_n(\boldsymbol{\theta})$ the expression between the inmost braces in (5.10). Finally, $f(\boldsymbol{\theta}) = \lim \inf_{n \to \infty} f_n(\boldsymbol{\theta})$. We shall prove that $\xi_1 \leqslant f(\boldsymbol{\theta})$. If $\xi_1 = -\infty$, then the statement trivially holds, so let us assume $\xi_1 > -\infty$. Now suppose $\xi_1 - f(\boldsymbol{\theta}) > 2\varepsilon > 0$. Then there exists an integer $N \equiv N(\varepsilon)$ such that both $\xi_1 - f_N(\boldsymbol{\theta}) > \varepsilon$ and $s_N(\boldsymbol{\theta})\, \xi_1 > \xi_1 - \varepsilon$ (the latter condition being relevant only if $\xi_1 > 0$.) Also, let $\boldsymbol{\theta}^* \equiv \{\theta_0^*, \theta_1^*, ...\}$ be such that $\theta_i^* = s_N^{-1}(\boldsymbol{\theta})\, \theta_i$ for $i = 0, 1, ..., N$ and $\theta_i^* = 0$ otherwise. We then have, by (5.4), (1.3) and (1.4),

$$\xi_1 - \varepsilon > f_N(\boldsymbol{\theta}) = s_N(\boldsymbol{\theta})\, f_N(\boldsymbol{\theta}^*) \geqslant s_N(\boldsymbol{\theta})\, x_{N1} > s_N(\boldsymbol{\theta})\, \xi_1 > \xi_1 - \varepsilon,$$

which is a contradiction. Consequently, $\xi_1 \leqslant f(\boldsymbol{\theta})$.

A proof for the fact that, actually, $\xi_1 = \inf_{\boldsymbol{\theta}} f(\boldsymbol{\theta})$ is established by considering the sequences $\boldsymbol{\theta}^{(n)} \equiv \{\theta_0^{(n)}, \theta_1^{(n)}, ...\}$, where $\theta_i^{(n)}$ equals the right-hand side of (5.5) if $1 \leqslant i \leqslant n$ and $\theta_i^{(n)} = 0$ otherwise. One then has $\xi_1 \leqslant f(\boldsymbol{\theta}^{(n)}) = x_{n1}$, while $x_{n1} \downarrow \xi_1$ as $n \to \infty$. ∎

It is easily seen that Theorem 9 implies [5, Theorem 4]. We refer to [5] for a survey of known upper bounds for ξ_1 that are corollaries to Theorem 9.

Regarding σ we have the following result.

THEOREM 10. *For any sequence* $\boldsymbol{\theta} \equiv \{\theta_0, \theta_1, ...\}$ *such that* $\theta_0 = 0$ *and* $\sum \theta_i = 1$, *one has*

$$\sigma \leqslant \lim_{k \to \infty} \inf \left\{ \lim_{n \to \infty} \inf \left\{ \sum_{i=1}^{n} (\theta_i c_{k+i} - 2(\theta_{i-1} \theta_i \lambda_{k+i})^{1/2}) \right\} \right\}. \tag{5.11}$$

Proof. Suppose the opposite of (5.11) holds true. Then, for $\varepsilon > 0$ sufficiently small, there are infinitely many numbers k such that

$$\sigma > \varepsilon + \lim_{n \to \infty} \inf \left\{ \sum_{i=1}^{n} (\theta_i c_{k+i} - 2(\theta_{i-1} \theta_i \lambda_{k+1})^{1/2}) \right\}. \tag{5.12}$$

However, by Theorem 9, the right-hand side of (5.12) is not less than $\varepsilon + \xi_1^{(k)}$, which contradicts (1.8). ∎

Theorem 10 is very similar, but not quite identical to [5, Theorem 5]. However, the latter theorem and all other upper bounds for σ that were derived in [5] may be obtained from Theorem 8. Since we can add nothing new to these results, we shall not pursue a detailed derivation.

REFERENCES

1. W. BARTH, R. S. MARTIN, AND J. H. WILKINSON, Calculation of the eigenvalues of a symmetric tridiagonal matrix by the method of bisection, *Numer. Math.* **9** (1967), 386–393.
2. R. BELLMAN, "Introduction to Matrix Analysis," 2nd ed., McGraw–Hill, New York, 1970.
3. T. S. CHIHARA, "An Introduction to Orthogonal Polynomials," Gordon & Breach, New York, 1978.
4. T. S. CHIHARA, Spectral properties of orthogonal polynomials on unbounded sets, *Trans. Amer. Math. Soc.* **270** (1982), 623–639.
5. E. A. VAN DOORN, On oscillation properties and the interval of orthogonality of orthogonal polynomials, *SIAM J. Math. Anal.* **15** (1984), 1031–1042.
6. G. H. GOLUB, Bounds for eigenvalues of tridiagonal symmetric matrices computed by the *LR* method, *Math. Comp.* **16** (1962), 438–445.
7. J. Z. HEARON, Roots of an irreducible continuant, *Linear Algebra Appl.* **3** (1970), 125–127.
8. E. LEOPOLD, Location of the zeros of polynomials satisfying three-term recurrence relations. III. Positive coefficients case, *J. Approx. Theory* **43** (1985), 15–24.
9. M. MARCUS AND H. MINC, "A Survey of Matrix Theory and Matrix Inequalities," Allyn & Bacon, Boston, 1964.
10. B. N. PARLETT, The Rayleigh quotient iteration and some generalizations for nonnormal matrices, *Math. Comp.* **28** (1974), 679–693.
11. B. N. PARLETT, "The Symmetric Eigenvalue Problem," Prentice–Hall, Englewood Cliffs, NJ, 1981.
12. J. GILEWICZ AND E. LEOPOLD, On the sharpness of results in the theory of location of zeros of polynomials defined by three term recurrence relations, *in* "Polynômes Orthogonaux et Applications," Proceedings of the Laguerre Symposium, Bar-le-Duc, 1984, pp. 259–266, Lecture Notes in Mathematics 1171, Springer-Verlag, Berlin, 1985.
13. G. FREUD, On the greatest zero of an orthogonal polynomial, *J. Approx. Theory* **46** (1986), 16–24.