SENSITIVITY ANALYSIS FOR SZEGŐ POLYNOMIALS

SUN-MI KIM* AND LOTHAR REICHEL†

Abstract. Szegő polynomials are orthogonal with respect to an inner product on the unit circle. Numerical methods for weighted least-squares approximation by trigonometric polynomials conveniently can be derived and expressed with the aid of Szegő polynomials. This paper discusses the conditioning of several mappings involving Szegő polynomials and, thereby, sheds light on the sensitivity of some approximation problems involving trigonometric polynomials.

1. Introduction.

Let $\mu(t)$ be a distribution function with infinitely many points of increase in the interval $[-\pi, \pi]$ and let

$$\mu_j := \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ijt} d\mu(t), \qquad j = 0, \pm 1, \pm 2, \dots ,$$

be the moments associated with $\mu(t)$, where $i := \sqrt{-1}$. For notational convenience, we assume throughout this paper that $\mu(t)$ is scaled so that $\mu_0 = 1$. We note that the moment matrices $M_n := [\mu_{j-k}]_{j,k=0}^{n-1}$, $n = 1, 2, 3, \dots$, are Hermitian positive definite and of Toeplitz form. There is an infinite sequence of monic polynomials $\{\psi_j\}_{j=0}^{\infty}$, known as *Szegő polynomials*, that are orthogonal with respect to the inner product on the unit circle

$$(f, g) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f(e^{it})} g(e^{it}) d\mu(t), \tag{1.1}$$

where the bar denotes complex conjugation. Thus, $(\psi_j, \psi_k) = 0$ for $j \neq k$, and $(\psi_j, \psi_j) > 0$ for all $j \geq 0$.

Properties of Szegő polynomials are discussed in, e.g., [15, 21, 26, 29]. Of particular importance is the fact that the monic Szegő polynomials satisfy recurrence relations of the form

$$\begin{array}{rcl} \psi_j(z) &=& z\psi_{j-1}(z) + \gamma_j \widetilde{\psi}_{j-1}(z), \\ \widetilde{\psi}_j(z) &=& \bar{\gamma}_j z\psi_{j-1}(z) + \widetilde{\psi}_{j-1}(z), \end{array} \qquad j = 1, 2, 3, \dots , \tag{1.2}$$

with $\psi_0(z) := 1$ and $\widetilde{\psi}_0(z) := 1$. The functions $\widetilde{\psi}_j$ often are referred to as *reverse polynomials*, because they satisfy $\widetilde{\psi}_j(z) = z^j \bar{\psi}_j(z^{-1})$; thus, the coefficients of $\widetilde{\psi}_j(z)$, when represented in terms of powers of z, are the conjugated coefficients of $\psi_j(z)$ in reverse order. The recursion coefficients $\gamma_j \in \mathbb{C}$, also known as *Schur parameters*, and the complementary parameters $\sigma_j \in \mathbb{R}^+$ are determined by

$$\begin{array}{rcl} \gamma_j &=& -(1, z\psi_{j-1})/\sigma_{j-1}, \\ \sigma_j &=& \sigma_{j-1}(1 - |\gamma_j|^2), \end{array} \qquad j = 1, 2, 3, \dots ,$$

with $\sigma_0 = 1$; see, e.g., [15, 21, 26, 29]. It is well known that $|\gamma_j| < 1$ for $j \geq 1$, and that all the zeros of Szegő polynomials lie in the open unit disk in $\mathbb{C}$.

---

[*] Department of Mathematical Sciences, Kent State University, Kent, OH 44242. E-mail: `sukim@math.kent.edu`. Research supported in part by NSF grant DMS-0107858.

[†] Department of Mathematical Sciences, Kent State University, Kent, OH 44242. E-mail: `reichel@math.kent.edu`. Research supported in part by NSF grant DMS-0107858.

Consider the integral

$$I(f) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})d\mu(t).$$

An n-point *Szegő quadrature rule* for this integral is of the form

$$S_\mu^{(n)}(f) := \sum_{j=1}^{n} \omega_j^2 f(\lambda_j), \qquad \omega_j^2 > 0,\ \lambda_j \in \Gamma, \tag{1.3}$$

where Γ denotes the unit circle in $\mathbb{C}$. The defining property of $S_\mu^{(n)}$ is that

$$S_\mu^{(n)}(p) = I(p), \qquad \forall p \in \Lambda_{-(n-1),n-1}, \tag{1.4}$$

where $\Lambda_{-(n-1),n-1}$ is the linear space of Laurent polynomials

$$L(z) := \sum_{k=-(n-1)}^{n-1} c_k z^k, \qquad c_k \in \mathbb{C},$$

of degree at most $n-1$. Equivalently, $S_\mu^{(n)}$ is exact for all trigonometric polynomials of degree less than n. In particular,

$$\sum_{j=1}^{n} \omega_j^2 = \mu_0 = 1. \tag{1.5}$$

Differently from the situation for Gauss quadrature rules, the nodes of Szegő rules are not uniquely determined. One node can be chosen arbitrarily on Γ; see, e.g., [12, 15, 19, 21] for discussions on Szegő quadrature rules.

Following Gragg [12], we express the recursion relations (1.2) for $1 \le j \le n$, with $\gamma_n = \tau$, in matrix form. Define the upper Hessenberg matrix

$$\widehat{H}_\tau := \begin{bmatrix} -\bar{\gamma}_0\gamma_1 & -\bar{\gamma}_0\gamma_2 & \cdots & -\bar{\gamma}_0\gamma_{n-1} & -\bar{\gamma}_0\tau \\ 1-|\gamma_1|^2 & -\bar{\gamma}_1\gamma_2 & \cdots & -\bar{\gamma}_1\gamma_{n-1} & -\bar{\gamma}_1\tau \\ & 1-|\gamma_2|^2 & \cdots & -\bar{\gamma}_2\gamma_{n-1} & -\bar{\gamma}_2\tau \\ & & \ddots & \vdots & \vdots \\ 0 & & & 1-|\gamma_{n-1}|^2 & -\bar{\gamma}_{n-1}\tau \end{bmatrix}$$

with $\gamma_0 := 1$. Then, for $1 \le j \le n$, the recursions (1.2) can be written as

$$[\psi_0(z), \psi_1(z), \dots, \psi_{n-1}(z)]\widehat{H}_\tau = z[\psi_0(z), \psi_1(z), \dots, \psi_{n-1}(z)] - [0, \dots, 0, \Psi_{n,\tau}(z)],$$

where $\Psi_{n,\tau}(z)$ is a polynomial of degree n in z. Its zeros are the eigenvalues of $\widehat{H}_\tau$. If $\tau = \gamma_n$, then $\Psi_{n,\tau} = \psi_n$, the nth Szegő polynomial, and the eigenvalues of $\widehat{H}_\tau$ are the zeros of this polynomial. Note that $\widehat{H}_\tau$ may have multiple eigenvalues and be defective. This is the case when $\mu(t) := t$ and $\tau := 0$. Then $\psi_j(z) = z^j$ for $j = 0, 1, 2, \dots$.

We will let $\tau \in \Gamma$. Gragg [12] showed that for such τ, the matrix

$$H_\tau := D_n^{-1/2}\widehat{H}_\tau D_n^{1/2} \tag{1.6}$$

with $D_n := \mathrm{diag}[\sigma_0, \sigma_1, \ldots, \sigma_{n-1}]$ is unitary, and that its eigenvalues are the nodes, λ_j, and the squared magnitude of the first component of associated normalized eigenvectors are the weights ω_j^2 of an n-point Szegő rule (1.3). The rule satisfies (1.4) for all $\tau \in \Gamma$. We will denote the matrix H_τ simply by H, since the τ-dependence is not important for our purpose. Efficient algorithms for the computation of Szegő quadrature rules are presented in [13, 14, 16, 30, 31].

Szegő polynomials arise in many applications, such as in signal processing, time series analysis, and operator theory; see, e.g., [1, 4, 15, 20, 22, 23, 24]. In particular, least-squares approximation by trigonometric polynomials gives rise to the following *inverse eigenvalue problem* for the unitary upper Hessenberg matrix (1.6): Given n distinct eigenvalues $\{\lambda_j\}_{j=1}^n$ on Γ and the first components $\{\omega_j\}_{j=1}^n$ of the corresponding eigenvectors normalized to have unit length with $\omega_j > 0$, determine the unique unitary upper Hessenberg matrix H with positive subdiagonal elements, such that

$$HW = W\Lambda,$$

where $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$ and W is a unitary eigenvector matrix with

$$e_1^t W e_j = \omega_j, \qquad 1 \le j \le n. \tag{1.7}$$

Throughout this paper the superscript t denotes transposition, and the superscript $*$ transposition and complex conjugation. The uniqueness of the matrix H follows from the Implicit Q Theorem [11, Theorem 7.4.2]. Existence and uniqueness of the solution of this inverse eigenvalue problem are discussed in [2, 25], where also algorithms for the computation of the solution are presented.

We are interested in investigating the sensitivity of the mapping

$$\mathcal{F}_n : [\omega, \lambda] \mapsto H, \qquad [\omega, \lambda] := [\omega_1, \omega_2, \ldots, \omega_n, \lambda_1, \lambda_2, \ldots, \lambda_n] \tag{1.8}$$

to perturbations in the data and, in particular, how the size of the ω_j and the location of the λ_j affects the conditioning of this map.

A fundamental problem in Scientific Computation is the determination of orthogonal polynomials from given moments. Several algorithms for computing Szegő polynomials from the moments have been proposed; see, e.g., [12, 17, 24, 25]. It is therefore of interest to study the conditioning of the mapping

$$\mathcal{K}_n : \mu \mapsto H, \qquad \mu := [\mu_0, \mu_1, \ldots, \mu_{n-1}].$$

It is convenient to consider $\mathcal{K}_n$ a composition of two maps $\mathcal{K}_n = \widetilde{\mathcal{F}}_n \circ \mathcal{G}_n$, where

$$\widetilde{\mathcal{F}}_n : [\omega^2, \lambda] \mapsto H, \qquad [\omega^2, \lambda] := [\omega_1^2, \omega_2^2, \ldots, \omega_n^2, \lambda_1, \lambda_2, \ldots, \lambda_n], \tag{1.9}$$

maps the quadrature data to the matrix (1.6) and

$$\mathcal{G}_n : \mu \mapsto [\omega^2, \lambda]$$

maps the moments to the quadrature data. This approach to studying the conditioning of $\mathcal{K}_n$ has been applied by Gautschi [8, 9] in his investigations of the sensitivity of polynomials orthogonal with respect to an inner product on an interval on the real axis as a function of the moments or modified moments.

Having determined a bound for the condition number of the map $\mathcal{F}_n$, we easily can compute a bound for the condition number of $\widetilde{\mathcal{F}}_n$. This is carried out in Section

2. The conditioning of the mapping $\mathcal{G}_n$ is investigated in [18]. Define the condition number

$$\kappa(\mathcal{G}_n)(\mu) := \frac{||\mu||_2}{||[\omega^2, \lambda]||_2} ||J_{\mathcal{G}_n}||_2, \tag{1.10}$$

where $J_{\mathcal{G}_n}$ is the Jacobian of $\mathcal{G}_n$. Then, it is shown in [18] that

$$\begin{aligned} &\kappa(\mathcal{G}_n)(\mu) < \frac{1}{\sqrt{2n-1}} \cdot \\ &\left\{ \sum_{k=1}^{2n-1} |P_1(e^{i\theta_k})|^2 + \sum_{j=2}^{n} \sum_{k=1}^{2n-1} \left(|P_j(e^{i\theta_k})|^2 + \frac{|Q_j(e^{i\theta_k})|^2}{\omega_j^2} \right) \right\}^{1/2}, \end{aligned} \tag{1.11}$$

where $\theta_k = 2\pi k/(2n-1)$ and P_j, Q_j are the Laurent polynomials

$$\begin{aligned} P_1(z) &= l_1(z)\bar{l}_1(z^{-1}), \\ P_j(z) &= \hat{l}_j(z)\bar{l}_j(z^{-1}) \left\{ \frac{z-\lambda_1}{\lambda_j - \lambda_1} - \left(l_j'(\lambda_j) - \lambda_j^{-2}\bar{l}_j'(\lambda_j^{-1}) \right) (z - \lambda_j) \right\}, \\ Q_j(z) &= \hat{l}_j(z)\bar{l}_j(z^{-1})(z - \lambda_j). \end{aligned}$$

Here $l_j(z)$ and $\hat{l}_j(z)$ are the Lagrange polynomials defined by

$$l_j(z) := \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{z - \lambda_k}{\lambda_j - \lambda_k}, \qquad \hat{l}_j(z) := \prod_{\substack{k=2 \\ k \neq j}}^{n} \frac{z - \lambda_k}{\lambda_j - \lambda_k}. \tag{1.12}$$

The bound (1.11) is modest unless there is a "tiny" weight ω_j^2, there are pairs of close nodes λ_j, or most nodes are clustered on a small subarc of Γ. We will see that, except for in these situations, the mapping $\widetilde{\mathcal{F}}_n$ also is fairly well-conditioned. We therefore may conclude that the map $\mathcal{K}_n$ is fairly well-conditioned when there are no tiny weights and the nodes are fairly evenly spread out over Γ with no pair of adjacent nodes very close.

It is well known that the map from the moments to the recursion coefficients for polynomials orthogonal with respect to an inner product on an interval on the real axis generally is severely ill-conditioned; the condition number typically grows rapidly and exponentially with n. Therefore, Gautschi studied the use of so-called *modified moments* instead of the ordinary moments, and showed that the determination of recursion coefficients from modified moments can be less sensitive to perturbations in the modified moments; see [7, 8, 9]. The modified moments are moments with respect to a family of polynomials other than the power basis. It is usually advantageous to choose the family to be a sequence of orthogonal polynomials with respect to an inner product on $\mathbb{R}$, such as a sequence of Chebyshev or Laguerre polynomials; see Beckermann and Bourreau [3], as well as Fischer [5] and Gautschi [7, 8, 9] for discussions. Related results for Sobolev inner products on a real interval are discussed by Zhang [32]. Gautschi and Zhang [10] describe numerical methods using modified moments in this context.

The monomials z^j are orthogonal with respect to the inner product (1.1) when $\mu(t) = t$. Therefore the moments μ_j are analogous to modified moments. This suggests that the determination of recursion coefficients for Szegő polynomials from moments generally should be less sensitive to perturbations than the computation of recursion coefficients for polynomials orthogonal with respect to an inner product on the real axis from the moments. Combining results of this paper with those in [18] supports this observation. In particular, the numerical examples of Section 3 show the condition numbers of $\mathcal{F}_n$ and $\widetilde{\mathcal{F}}_n$ to grow slower than exponentially.

The usefulness of Szegő polynomials in signal processing stems from their connection with symmetric positive definite Toeplitz matrices. In signal processing applications the recursion coefficients γ_j often are computed by factoring a Toeplitz autocorrelation matrix associated with a wide sense stationary signal by the Schur or Levinson algorithms; see, e.g., [1, 12, 15, 20, 22, 23]. Block generalizations of the latter algorithms are available and their numerical properties are discussed in, e.g., [6, 28]. These block algorithms compute matrix-valued recursion coefficients, which determine matrix-valued Szegő polynomials; see, e.g., [4] for discussions on these polynomials, which are associated with matrix-valued measures on the unit circle. We are presently extending the approach of the present paper to investigate the conditioning of mappings analogous to (1.8) and (1.9) for matrix-valued Szegő polynomials in order to better understand the numerical properties of these polynomials.

We finally remark that the matrix (1.6) can be expressed as a product of n elementary unitary matrices,

$$H = G_1(\gamma_1)G_2(\gamma_2)\cdots G_{n-1}(\gamma_{n-1})\hat{G}_n(\tau),$$

where the $G_j(\gamma_j)$, $1 \le j < n$, are $n \times n$ Givens matrices

$$G_j(\gamma_j) := \begin{bmatrix} I_{j-1} & & & \\ & -\gamma_j & \eta_j & \\ & \eta_j & \overline{\gamma}_j & \\ & & & I_{n-j-1} \end{bmatrix}, \qquad \eta_j \in \mathbb{R}^+, \qquad |\gamma_j|^2 + \eta_j^2 = 1,$$

and

$$\hat{G}_n(\tau) := \mathrm{diag}[1, \ldots, 1, -\tau].$$

Therefore small perturbations in H correspond to small perturbations in the γ_j and τ. It follows that the mapping from $[\omega, \lambda]$ to the coefficients $[\gamma_1, \gamma_2, \ldots, \gamma_{n-1}, \tau]$ is well conditioned when the mapping $\mathcal{F}_n$ is well conditioned.

2. The conditioning of $\mathcal{F}_n$ and $\widetilde{\mathcal{F}}_n$.

The aim of this section is to bound the condition number of the nonlinear mappings (1.8) and (1.9). Our investigation follows the approach used by Beckermann and Bourreau [3] in their study of the choice of modified moments with respect to an inner product on an interval on the real axis.

In order to define the condition number of a matrix-valued map $\mathcal{A} : \mathbb{C}^m \ni x \longmapsto A \in \mathbb{C}^{n\times n}$, where $x = [x_1, x_2, \ldots, x_m]^t$, we introduce

$$\frac{\partial A}{\partial x_j} := \lim_{\Delta x_j \to 0} \frac{\Delta A}{\Delta x_j}.$$

Here ΔA is the perturbation of A caused by the error Δx_k in the data x_k. Then we could define the condition number of $\mathcal{A}$ as

$$\kappa(\mathcal{A})(x) := \frac{||x||_2}{||A||_2} \left(\sum_{j=1}^{m} \left\| \frac{\partial A}{\partial x_j} \right\|_F^2 \right)^{1/2}. \tag{2.1}$$

Here and below $\|\cdot\|_2$ denotes the Euclidean vector norm and $\|\cdot\|_F$ the Frobenius norm. The condition number (2.1) has previously been used by Zhang [32, p. 573] in his analysis of the sensitivity of orthogonal polynomials of Sobolev-type to perturbations in the moments and modified moments. However, in order to obtain more meaningful results when the quantities involved are of different orders of magnitude, we allow nonnegative scaling factors δ_j, similarly as Beckermann and Bourreau in their investigation [3]. Thus, introduce the scaled condition number

$$\kappa^{(\delta)}(\mathcal{A})(x) := \frac{1}{||A||_2}\left(\sum_{j=1}^{m}\left|\frac{x_j}{\delta_j}\right|^2\right)^{1/2}\left(\sum_{j=1}^{m}\delta_j^2\left\|\frac{\partial A}{\partial x_j}\right\|_F^2\right)^{1/2},$$

which we minimize over all scaling factors $\delta_j > 0$ to obtain

$$\hat{\kappa}(\mathcal{A})(x) := \inf_{\delta_j>0}\kappa^{(\delta)}(\mathcal{A})(x) = \frac{1}{||A||_2}\sum_{j=1}^{m}|x_j|\left\|\frac{\partial A}{\partial x_j}\right\|_F.$$

It follows from the Cauchy inequality that $\hat{\kappa}(\mathcal{A})(x) \le \kappa^{(\delta)}(\mathcal{A})(x)$. The corresponding condition number for the mapping $\mathcal{F}_n$ is given by

$$\hat{\kappa}_{[\omega,\lambda]}(\mathcal{F}_n)([\omega,\lambda]) := \sum_{j=1}^{n}\left(\omega_j\left\|\frac{\partial H}{\partial \omega_j}\right\|_F + \left\|\frac{\partial H}{\partial \lambda_j}\right\|_F\right), \tag{2.2}$$

where we have used the fact that $\|H\|_2 = 1$. The subscript $[\omega,\lambda]$ of $\hat{\kappa}_{[\omega,\lambda]}$ indicates that we consider perturbations in all ω_j and λ_j. We also are interested in the sensitivity of $\mathcal{F}_n$ to perturbation in the ω_j, only, and denote the the associated condition number by

$$\hat{\kappa}_{[\omega]}(\mathcal{F}_n)([\omega,\lambda]) := \sum_{j=1}^{n}\omega_j\left\|\frac{\partial H}{\partial \omega_j}\right\|_F. \tag{2.3}$$

Beckermann and Bourreau [3] studied the relation between Gauss quadrature data and the related symmetric tridiagonal Jacobi matrix. Using their techniques, we obtain upper bounds for the partial derivatives in (2.2). This yields the following bound for the condition number of $\mathcal{F}_n$.

Theorem 2.1. *The condition number (2.2) satisfies the bound*

$$\hat{\kappa}_{[\omega,\lambda]}(\mathcal{F}_n)([\omega,\lambda]) \le 7n + 6\sum_{j=1}^{n}\omega_j\left(\sum_{k=1}^{n}\frac{|l_k'(\lambda_j)|^2}{\omega_k^2}\right)^{1/2}, \tag{2.4}$$

where l_k' denotes the derivative of the Lagrange polynomial l_k defined by (1.12).

The proof of this bound requires a few auxiliary results, which we formulate as lemmas. The first of these recalls some decompositions of matrices related to the recursion coefficient matrix H derived by Gragg [12].

Lemma 2.2. *Let*

$$V_n := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix},$$

$$\Lambda_n := \operatorname{diag}[\lambda_1, \lambda_2, \dots, \lambda_n], \qquad W_n := \operatorname{diag}[\omega_1, \omega_2, \dots, \omega_n],$$

and $S := V_n W_n$. *Then* H^t *has the spectral decomposition* $H^t = Q_n \Lambda_n Q_n^*$ *with a unitary matrix* Q_n. *The columns are scaled to have positive first component. Thus, the* j*th column of* Q_n *has first component* $\omega_j > 0$*; cf. (1.7). Moreover,* S^* *has the QR-factorization with unitary factor* Q_n^* *and an upper triangular factor with positive diagonal entries. Further,* $SS^* = \overline{M}_n$*, where* $M_n = [\mu_{j-k}]_{j,k=0}^{n-1}$ *is the moment matrix.*

A variant of Lemma 2.3 below is shown by Beckermann and Bourreau [3, Lemma 3] for the case when $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$. The proof in [3] is based on results by Stewart [27, Theorem 3.1] on the perturbation of the QR-factorization of a real matrix. We modify [3, Lemma 3] in two ways: We allow $A \in \mathbb{C}^{n\times n}$ and $x \in \mathbb{C}^n$, and only show a first order inequality in $\|x\|_2$. In our application of this result, $\|x\|_2 = \mathcal{O}(\epsilon)$ with ϵ "tiny." In the following lemma and below, the symbols $\doteq$ and $\dot{\leq}$ denote first order equalities and inequalities in ϵ, respectively.

Lemma 2.3. *Consider* $A \in \mathbb{C}^{n\times n}$ *and its perturbed counterpart* $\tilde{A} = (I_n + e_j x^t)A$*, with* $x \in \mathbb{C}^n$ *and* e_j *the* j*th axis vector. Assume that* $||x||_2 = \mathcal{O}(\epsilon)$*. Then, for the QR-factorizations* $A = QR$*,* $\tilde{A} = \tilde{Q}\tilde{R}$*, with* R *and* $\tilde{R}$ *having real diagonal entries, we have*

$$||\tilde{Q} - Q||_F \dot{\leq} 3||x||_2. \tag{2.5}$$

Proof. The result can be shown by following the development of Stewart [27, pp. 516–517] for the special case when $A = Q = R = I$. This situation allows some simplifications, which yield (2.5). Stewart [27] shows his bounds for $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$, but points out that his results can be extended to the complex case under the condition that the diagonal elements of the upper triangular factors in the QR-factorizations are real. □

We are in a position to show Theorem 2.1.

Proof. (i) For each fixed $j \in \{1, 2, \dots, n\}$, we consider the perturbed matrices $\tilde{H}, \tilde{Q}_n, \tilde{\Lambda}_n$, and $\tilde{S}$ obtained by replacing ω_j by $\tilde{\omega}_j = \omega_j + \epsilon$ for some small $|\epsilon|$, $\tilde{\omega}_k = \omega_k$ for $k \neq j$, and $\tilde{\lambda}_k = \lambda_k$ for $k = 1, 2, \dots, n$. Our aim is to estimate

$$\left\| \frac{\partial H}{\partial \omega_j} \right\|_F = \lim_{\epsilon\to 0} \frac{||\tilde{H} - H||_F}{|\epsilon|}.$$

Let $A_{j\cdot}$ denote the jth row of the matrix A. Then

$$\tilde{S}^* = \tilde{W}_n V_n^* = \begin{bmatrix} \omega_1 (V_n^*)_{1\cdot} \\ \vdots \\ (1 + \frac{\epsilon}{\omega_j})\omega_j (V_n^*)_{j\cdot} \\ \vdots \\ \omega_n (V_n^*)_{n\cdot} \end{bmatrix} = (I_n + e_j x^t) S^*, \qquad x := (\frac{\epsilon}{\omega_j}) e_j.$$

Application of Lemma 2.3 shows that

$$||\tilde{Q}_n^* - Q_n^*||_F \dot{\leq} \frac{3|\epsilon|}{\omega_j}$$

for $|\epsilon|$ sufficiently small. Thus, it follows from Lemma 2.2 that

$$\begin{aligned}
||\tilde{H} - H||_F &= ||\tilde{H}^t - H^t||_F = ||\tilde{Q}_n \Lambda_n \tilde{Q}_n^* - Q_n \Lambda_n Q_n^*||_F \\
&= ||(\tilde{Q}_n - Q_n)\Lambda_n \tilde{Q}_n^* + Q_n \Lambda_n (\tilde{Q}_n^* - Q_n^*)||_F \\
&\leq 2||\tilde{Q}_n - Q_n||_F \\
&\dot{\leq} \frac{6|\epsilon|}{\omega_j}.
\end{aligned}$$

Therefore,

$$\left\| \frac{\partial H}{\partial \omega_j} \right\|_F \leq \frac{6}{\omega_j}.$$

(ii) Similarly, for each fixed $j \in \{1, 2, \ldots, n\}$, let $\tilde{H}, \tilde{Q}_n, \tilde{\Lambda}_n$, and $\tilde{S}$ be the perturbed matrices obtained by replacing λ_j by $\tilde{\lambda}_j = \lambda_j + \epsilon$ for some small $|\epsilon|$, $\tilde{\lambda}_k = \lambda_k$ for $k \neq j$, and $\tilde{\omega}_k = \omega_k$ for $k = 1, 2, \ldots, n$. We would like to estimate

$$\left\| \frac{\partial H}{\partial \lambda_j} \right\|_F = \lim_{\epsilon \to 0} \frac{||\tilde{H} - H||_F}{|\epsilon|}.$$

It follows from

$$[0, 1, 2\lambda_j, \ldots, (n-1)\lambda_j^{n-2}]^t = V_n^t [l_1^{'}(\lambda_j), l_2^{'}(\lambda_j), \ldots, l_n^{'}(\lambda_j)]^t$$

that

$$\begin{array}{rcll}
\tilde{S} & = & \tilde{V}_n W_n \doteq V_n(I_n + \tilde{x}e_j^t)W_n, & \text{where } \tilde{x} = \epsilon[l_1^{'}(\lambda_j), \ldots, l_n^{'}(\lambda_j)]^t, \\
& = & V_n W_n (I_n + x e_j^t), & \text{where } x = \epsilon[\frac{\omega_j}{\omega_1} l_1^{'}(\lambda_j), \ldots, \frac{\omega_j}{\omega_n} l_n^{'}(\lambda_j)]^t.
\end{array}$$

Application of Lemma 2.3 to $\tilde{S}^* = (I_n + e_j x^*) W_n V_n^*$ with $|\epsilon|$ sufficiently small yields

$$||\tilde{Q}_n^* - Q_n^*||_F \,\dot{\leq}\, 3|\epsilon| \left(\omega_j^2 \sum_{k=1}^n \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2}.$$

Therefore,

$$\begin{aligned}
||\tilde{H} - H||_F &= ||\tilde{Q}_n \tilde{\Lambda}_n \tilde{Q}_n^* - Q_n \Lambda_n Q_n^*||_F \\
&= ||(\tilde{Q}_n - Q_n)\tilde{\Lambda}_n \tilde{Q}_n^* + Q_n(\tilde{\Lambda}_n - \Lambda_n)\tilde{Q}_n^* + Q_n \Lambda_n (\tilde{Q}_n^* - Q_n^*)||_F \\
&\dot{\leq} 2||\tilde{Q}_n - Q_n||_F + ||\tilde{\Lambda}_n - \Lambda_n||_F \\
&\dot{\leq} 6|\epsilon| \left(\omega_j^2 \sum_{k=1}^n \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2} + |\epsilon|
\end{aligned}$$

and, consequently,

$$\left\| \frac{\partial H}{\partial \lambda_j} \right\|_F \leq 6 \left(\omega_j^2 \sum_{k=1}^n \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2} + 1.$$

(iii) Substituting the bounds from (i) and (ii) into (2.2) completes the proof. □

The bound (2.4) indicates that the condition number (2.2) may be large if there are tiny ω_j or if the nodes λ_j are clustered on a subarc of the unit circle. However, the above proof shows that $\mathcal{F}_n$ is quite well-conditioned with respect to perturbations in the ω_j only, even when some $\omega_j > 0$ are tiny.

Corollary 2.4. *The condition number (2.3) satisfies the bound*

$$\hat{\kappa}_{[\omega]}(\mathcal{F}_n)([\omega,\lambda]) \leq 6n.$$

The following corollary simplifies the bound (2.4).

Corollary 2.5. *Let l_k be the Lagrange polynomials defined by (1.12) and introduce*

$$\alpha_k := \max_{|z|=1} |l_k(z)|, \qquad 1 \leq k \leq n. \tag{2.6}$$

Then

$$\hat{\kappa}_{[\omega,\lambda]}(\mathcal{F}_n)([\omega,\lambda)] \leq 7n + 6n^2 \max_{1\leq k\leq n} \frac{\alpha_k}{\omega_k}. \tag{2.7}$$

Proof. The Cauchy inequality yields

$$\sum_{j=1}^{n} \omega_j \left(\sum_{k=1}^{n} \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2} \leq \left(\sum_{j,k=1}^{n} \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2}. \tag{2.8}$$

Since the Lagrange polynomials l_k are of degree $n-1$, Bernstein's inequality gives

$$\max_{|z|=1} |l_k^{'}(z)| \ \leq \ (n-1) \max_{|z|=1} |l_k(z)|, \qquad 1 \leq k \leq n.$$

Therefore,

$$\sum_{j,k=1}^{n} \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \leq n(n-1)^2 \sum_{k=1}^{n} \max_{|z|=1} \frac{|l_k(z)|^2}{\omega_k^2} \leq n^3 \sum_{k=1}^{n} \frac{\alpha_k^2}{\omega_k^2} \leq n^4 \max_{1\leq k\leq n} \frac{\alpha_k^2}{\omega_k^2}.$$

Substitution into the right-hand side of (2.8) and using this bound in (2.4) shows (2.7). □

We turn to a bound for the condition number

$$\hat{\kappa}_{[\omega^2,\lambda]}(\widetilde{\mathcal{F}}_n)([\omega^2,\lambda]) := \sum_{j=1}^{n} \left(\omega_j^2 \left\| \frac{\partial H}{\partial \omega_j^2} \right\|_F + \left\| \frac{\partial H}{\partial \lambda_j} \right\|_F \right) \tag{2.9}$$

for the mapping (1.9).

Theorem 2.6. *The condition number (2.9) satisfies*

$$\hat{\kappa}_{[\omega^2,\lambda]}(\widetilde{\mathcal{F}}_n)([\omega^2,\lambda]) \leq 4n + 6 \sum_{j=1}^{n} \omega_j \left(\sum_{k=1}^{n} \frac{|l_k^{'}(\lambda_j)|^2}{\omega_k^2} \right)^{1/2}, \tag{2.10}$$

where the l_k are given by (1.12).

Proof. The inequality is shown similarly as the bound (2.4). We use the same notation as in the proof of Theorem 2.1. Some details are omitted. Consider the perturbed matrices $\tilde{H}, \tilde{Q}_n, \tilde{\Lambda}_n$, and $\tilde{S}$ obtained by replacing ω_j^2 by $\tilde{\omega}_j^2 = \omega_j^2 + \epsilon$ for some small $|\epsilon|$. Let $\tilde{\omega}_k = \omega_k$ for $k \neq j$, and $\tilde{\lambda}_k = \lambda_k$ for $k = 1, 2, \ldots, n$. It follows from $\tilde{\omega}_j \doteq \omega_j(1 + \frac{\epsilon}{2\omega_j^2})$ that

$$\tilde{S}^* \doteq (I_n + e_j x^*)S^*, \qquad x := (\frac{\epsilon}{2\omega_j^2})e_j.$$

Lemma 2.3 yields, for $|\epsilon|$ sufficiently small,

$$||\tilde{Q}_n^* - Q_n^*||_F \overset{.}{\leq} \frac{3|\epsilon|}{2\omega_j^2}.$$

Therefore, by Lemma 2.2,

$$||\tilde{H} - H||_F \overset{.}{\leq} \frac{3|\epsilon|}{\omega_j^2}$$

and it follows that

$$\left\|\frac{\partial H}{\partial \omega_j}\right\|_F \leq \frac{3}{\omega_j^2}.$$

The bounds for the partial derivatives $\left\|\frac{\partial H}{\partial \lambda_j}\right\|_F$ are the same as in the proof of Theorem 2.1. This shows the bound (2.10). ▯

3. Numerical examples. This section presents computed examples which illustrate the bounds of the previous section. The computations were carried in MATLAB with about 16 significant digits.

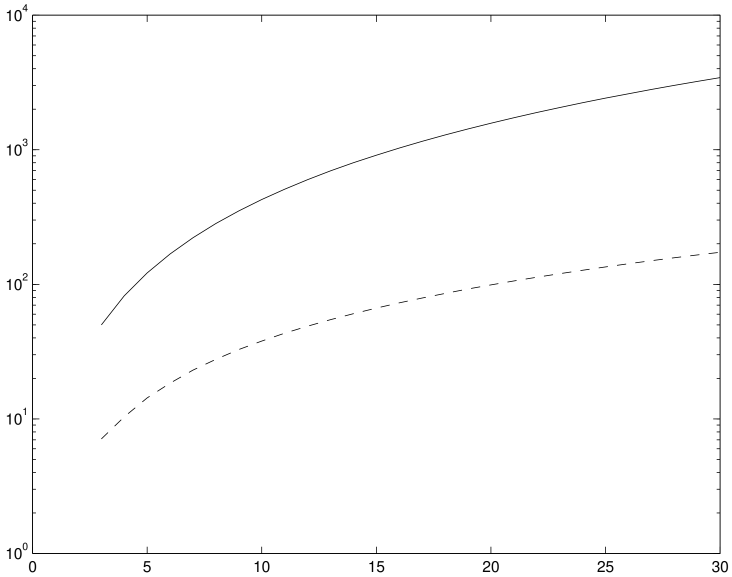

FIG. 3.1. *Example 3.1: The condition number of the map $\mathcal{F}_n$ (dashed graph) and the bound of Theorem 2.1 (continuous graph).*

Example 3.1: Let the weights ω_j and nodes λ_j be given by

$$\omega_1 := 1 - \frac{n-1}{n^3}, \qquad \omega_j := \frac{1}{n^3}, \qquad 2 \leq j \leq n,$$
$$\lambda_j := \exp(2\pi i(\frac{j-1}{n})), \qquad 1 \leq j \leq n.$$

Figure 3.1 displays the condition number of the map $\mathcal{F}_n$ and the upper bound of Theorem 2.1 for $1 \leq n \leq 50$. The quality of the bound can be seen to deteriorate as n increases. Nevertheless, the bound, and therefore also the condition number of $\mathcal{F}_n$, grow slower than exponentially with n.

Evaluation of the condition numbers of $\mathcal{F}_n$ and $\widetilde{\mathcal{F}}_n$ shows the latter to be somewhat smaller than the former, in agreement with the fact that the bound of Theorem 2.6 is smaller than the bound of Theorem 2.1. □

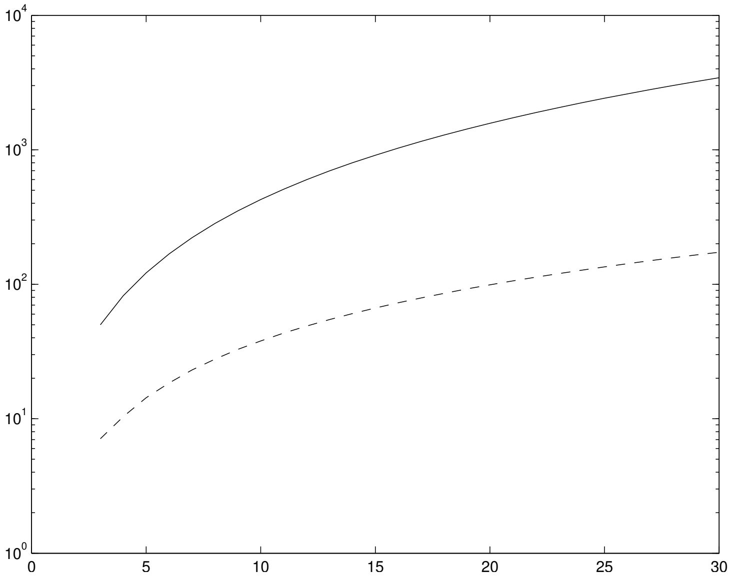

Fig. 3.2. *Example 3.2: The condition number of the map $\mathcal{F}_n$ (dashed graph) and the bound of Theorem 2.1 (continuous graph).*

Table 3.1
Example 3.2: Decrease of the smallest weights as a function of n.

n	5	10	15	20	25	30
ρ_n	$3.8 \cdot 10^{-3}$	$6.9 \cdot 10^{-5}$	$3.4 \cdot 10^{-6}$	$2.7 \cdot 10^{-7}$	$2.9 \cdot 10^{-8}$	$4.0 \cdot 10^{-9}$

Example 3.2: Let the recursion coefficients be given by

$$\gamma_j := \frac{1}{\sqrt{j+1}}, \qquad 1 \leq j < n. \tag{3.1}$$

Let $\{\omega_j\}_{j=1}^n$ denote the weights of the n-point Szegő quadrature rule determined by the recursion coefficients (3.1) and $\tau := -1$. Define $\rho_n := \min_{1 \leq j \leq n} \omega_j$. Table 3.1 shows ρ_n to decrease rapidly as n increases. Therefore the condition number of the map $\mathcal{F}_n$ grows quite rapidly with n; see Figure 3.2, which shows the condition number of $\mathcal{F}_n$ as well as the upper bound of Theorem 2.1 for $1 \leq n \leq 30$. Note that the sum $\sum_{j=1}^{n-1} \gamma_j^2$ grows unboundedly with n. □

Acknowledgment. We would like to thank Paco Marcellan for discussions and the referees for suggestions.

REFERENCES

[1] G. S. Ammar, W. B. Gragg, and L. Reichel, Determination of Pisarenko frequency estimates as eigenvalues of an orthogonal matrix, in: F. T. Luk, Ed., Advanced Algorithms and Architectures for Signal Processing II, SPIE 826, Internat. Soc. Optical Engrg., Bellingham, 1987, pp. 143–145.

[2] G. S. Ammar, W. B. Gragg, and L. Reichel, Constructing a unitary Hessenberg matrix from spectral data, in: G. H. Golub and P. Van Dooren, Eds., Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms, Springer, Berlin, 1991, pp. 385–395.

[3] B. Beckermann and E. Bourreau, How to choose modified moments?, J. Comput. Appl. Math., 98 (1998), pp. 81–98.

[4] R. L. Ellis and I. Gohberg, Orthogonal Systems and Convolution Operators, Birkhäuser, Boston, 2002.

[5] H.-J. Fischer, On the condition of orthogonal polynomials via modified moments, Z. Anal. Anwendungen, 15 (1996), pp. 1–18.

[6] K. A. Gallivan, S. Thirumalai, P. Van Dooren, and V. Vermaut, High performance algorithms for Toeplitz and block Toeplitz matrices, Linear Algebra Appl., 241-243 (1996), pp. 343–388.

[7] W. Gautschi, On the construction of Gaussian quadrature rules from modified moments, Math. Comp., 24 (1970), pp. 245–260.

[8] W. Gautschi, On generating orthogonal polynomials, SIAM J. Sci. Statist. Comput., 3 (1982), pp. 289–317.

[9] W. Gautschi, On the sensitivity of orthogonal polynomials to perturbations in the moments, Numer. Math., 48 (1986), pp. 369–382.

[10] W. Gautschi and M. Zhang, Computing orthogonal polynomials in Sobolev spaces, Numer. Math., 71 (1995), pp. 159–183.

[11] G. H. Golub and C. F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[12] W. B. Gragg, Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle, J. Comput. Appl. Math., 46 (1993), pp. 183–198. This is a slightly edited version of a paper published in Russian: in E. S. Nikolaev, Ed., Numerical Methods in Linear Algebra, Moscow University Press, Moscow, 1982, pp. 16–32.

[13] W. B. Gragg, The QR algorithm for unitary Hessenberg matrices, J. Comput. Appl. Math., 16 (1986), pp. 1–8.

[14] W. B. Gragg and L. Reichel, A divide and conquer method for the unitary and orthogonal eigenproblems, Numer. Math., 57 (1990), pp. 695–718.

[15] U. Grenander and G. Szegő, Toeplitz Forms and Their Applications, Chelsea, New York, 1984.

[16] M. Gu, R. Guzzo, X.-B. Chi, and X.-Q. Cao, A stable divide and conquer algorithm for the unitary eigenproblem, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 385–404.

[17] C. Jagels and L. Reichel, The isometric Arnoldi process and an application to iterative solutions of large linear systems, in: R. Beauwens and P. de Groen, Eds., Iterative Methods in Linear Algebra, Elsevier, Amsterdam, 1992, pp. 361–369.

[18] C. Jagels and L. Reichel, On the construction of Szegő polynomials, J. Comput. Appl. Math., 46 (1993), pp. 241–254.

[19] C. Jagels and L. Reichel, Szegő-Lobatto quadrature rules, J. Comput. Appl. Math., 200 (2007), pp. 116–126.

[20] W. B. Jones, O. Njåstad, and E. B. Saff, Szegő polynomials associated Wiener-Levinson filters, J. Comput. Appl. Math., 32 (1990), pp. 387–406.

[21] W. B. Jones, O. Njåstad, and W. J. Thron, Moment theory, orthogonal polynomials, quadrature and continued fractions associated with the unit circle, Bull. London Math. Soc., 21 (1989), pp. 113–152.

[22] W. B. Jones, W. J. Thron, O. Njåstad, and H. Waadeland, Szegő polynomials applied to frequency analysis, J. Comput. Appl. Math., 46 (1993), pp. 217–228.

[23] T. Kailath, Linear estimation for stationary and near-stationary processes, in: T. Kailath, Ed., Modern Signal Processing, Hemisphere, Washington, 1985, pp. 59–128.

[24] L. Reichel and G. S. Ammar, Fast approximation of dominant harmonics by solving an orthogonal eigenvalue problem, in: J. G. McWhirter, Ed., Mathematics in Signal Processing II, Clarendon Press, Oxford, 1990, pp. 575–591.

[25] L. Reichel, G. S. Ammar, and W. B. Gragg, Discrete least squares approximation by trigonometric polynomials, Math. Comp., 57 (1991), pp. 273–289.

[26] B. Simon, Orthogonal Polynomials on the Unit Circle, Amer. Math. Soc., Providence, 2004.

[27] G. W. Stewart, Perturbation bounds for the QR factorization of a matrix, SIAM J. Numer. Anal., 14 (1977), pp. 509–517.

[28] M. Stewart and P. Van Dooren, Stability issues in the factorization of structured matrices, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 104–118.

[29] G. Szegő, Orthogonal Polynomials, 4th ed., Amer. Math. Soc., Providence, 1975.

[30] T.-L. Wang and W. B. Gragg, Convergence of the shifted QR algorithm for unitary Hessenberg matrices, Math. Comp., 71 (2002), pp. 1473–1496.

[31] T.-L. Wang and W. B. Gragg, Convergence of the unitary QR algorithm with a unimodular Wilkinson shift, Math. Comp., 72 (2003), pp. 375–385.

[32] M. Zhang, Sensitivity analysis for computing orthogonal polynomials of Sobolev type, in: R. V. M. Zahar, Ed., Approximation and Computation, Birkhäuser, Boston, 1994, pp. 563–576.