JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 98 (1998) 81–98

# How to choose modified moments?

Bernhard Beckermann*, Emmanuel Bourreau

*Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA - M3, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France*

Received 3 April 1997; received in revised form 23 September 1997

**Abstract**

Modified moments of a nonnegative measure determine uniquely the recurrence coefficients of the polynomials orthogonal with respect to the given nonnegative measure. The sensitivity of the underlying nonlinear maps with regard to perturbations in the modified moments has been studied by Gautschi and Fischer, and is further analyzed in this paper. In particular, we discuss the question of how (not) to choose the support of the underlying two measures. © 1998 Elsevier Science B.V. All rights reserved.

*AMS classification:* 65D20; 33C45

*Keywords:* Orthogonal polynomials; Modified moments; Modified Chebyshev algorithm

## 1. Introduction

Given some nonnegative measure $\sigma$ supported on the real line, in many applications one is concerned with the problem of computing monic orthogonal polynomials $\hat{\pi}_n$, $n \geqslant 0$, satisfying $\hat{\pi}_n = x^n +$ lower powers and $\int \hat{\pi}_n(x)\hat{\pi}_k(x)\,\mathrm{d}\sigma(x) = 0$ for $n \neq k$. It is well known that these polynomials are connected via a three-term recurrence

$$\hat{\pi}_{n+1}(x) = (x - \alpha_n)\hat{\pi}_n(x) - \beta_n^2\hat{\pi}_{n-1}(x), \quad n \geqslant 0,$$

where $\hat{\pi}_0(x) = 1$, $\hat{\pi}_{-1}(x) = 0$, and $\alpha_n, \beta_n, n \geqslant 0$ are some real numbers, with $\beta_n > 0$, $\beta_0 := [\int \mathrm{d}\sigma(x)]^{1/2}$. However, for applications like evaluating the orthogonal polynomial $\hat{\pi}_n$, determining its zeros, or computing the data of the corresponding $n$th Gaussian quadrature rule, one does not need explicitly the coefficients of this polynomial. From a numerical standpoint [8, p. 290], it is much more interesting to compute the recurrence coefficients $(\alpha, \beta) := (\alpha_0, \ldots, \alpha_{n-1}, \beta_0, \ldots, \beta_{n-1})^{\mathrm{t}}$ and

* E-mail: bbecker@ano.univ-lille1.fr.


PII: S0377-0427(98)00116-2

then to apply classical methods such as the Clenshaw algorithm [14, p. 181], or to analyze the corresponding Jacobi matrix

$$\mathcal{J} = \begin{pmatrix} \alpha_0 & \beta_1 & 0 & \cdots & \cdots & \cdots & 0 \\ \beta_1 & \alpha_1 & \beta_2 & 0 & \cdots & \cdots & 0 \\ 0 & \beta_2 & \alpha_2 & \beta_3 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & \beta_{n-3} & \alpha_{n-3} & \beta_{n-2} & 0 \\ 0 & \cdots & \cdots & 0 & \beta_{n-2} & \alpha_{n-2} & \beta_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \beta_{n-1} & \alpha_{n-1} \end{pmatrix} \tag{1}$$

In a number of papers [6–10], Gautschi studied the question of how to compute the recurrence coefficients of $\sigma$ in a numerical stable way. A number of classical methods such as the Stieltjes procedure or the (classical) Chebyshev algorithm suffer from numerical instabilities (see [8, Section 2]). On the other hand, Gautschi mentioned that the use of the so-called *modified moments* instead of the ordinary moments $\sigma_k := \int x^k \, d\sigma(x)$, $k = 0, \ldots, 2n-1$, may lead to a significant stabilization of the Chebyshev algorithm (see, for instance, the numerical results reported in [13]). The resulting *modified Chebyshev algorithm* has been proposed in [15, 19], it takes as initial data the vector of modified moments

$$m = (m_0, \ldots, m_{2n-1})^{\mathrm{t}}, \qquad m_k = \int \hat{p}_k(x) \, d\sigma(x),$$

where usually $(\hat{p}_k)_{k \geqslant 0}$ is chosen to be a known family of monic polynomials orthogonal with respect to a nonnegative measure $s$, again supported on the real line. [1]

As a preliminary step for establishing stability of the modified Chebyshev algorithm, we need to consider the numerical condition of the problem, i.e., of the nonlinear map $K_n : m \mapsto (\alpha, \beta)$ mapping the modified moments to the recurrence coefficients. Obviously, for fixed $\sigma$, the condition number will depend very much on the choice of the measure $s$. It is the aim of this paper to show that the two supports should essentially coincide, since otherwise the condition number of $K_n$ will increase exponentially as a function of $n$. This result confirms some numerical experiments reported in [8, Examples 4.4, 4.5, 4.8, 4.9, and especially Example 4.7].

Following Gautschi, we will consider the decomposition $K_n = H_n \circ G_n$, with $G_n : m \mapsto (\lambda, \tau)$ mapping the $2n$-modified moments to the data $(\lambda, \tau) := (\lambda_1, \ldots, \lambda_n, \tau_1, \ldots, \tau_n)^{\mathrm{t}}$ of the Gaussian quadrature formula

$$Q_n(f) = \sum_{\nu=1}^{n} \lambda_\nu \cdot f(\tau_\nu). \tag{2}$$

[1] Following Gautschi, we adopt the convention that quantities connected with $\sigma$ are denoted by Greek letters, whereas Latin letters are reserved for the measure s.

We recall that $\tau_1, \ldots, \tau_n$ are the distinct real zeros of $\hat{\pi}_n$ (or the eigenvalues of the Jacobi matrix (1)), and that the Christoffel numbers $\lambda_1, \ldots, \lambda_n$ can be obtained from the first component of the corresponding eigenvectors of $\mathscr{J}$. The map $H_n : (\lambda, \tau) \mapsto (\alpha, \beta)$ does only depend on the measure $\sigma$.

As condition number of a (smooth) nonlinear map $M : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ one usually adopts the quantity

$$\limsup_{\|\Delta x\| \to 0} \frac{\|M(x + \Delta x) - M(x)\|}{\|M(x)\|} \cdot \frac{\|x\|}{\|\Delta x\|} = \frac{\|x\| \cdot \left\|\frac{\partial y}{\partial x}(x)\right\|}{\|M(x)\|},$$

where $\partial y/\partial x = (\partial y_j/\partial x_k)_{j,k}$ is the Jacobian matrix of $M$, and $\|.\|$ denotes suitable vector norms and subordinate matrix norms. Thus, this quantity gives a measure for the maximal magnification of (infinitesimal) relative errors. However, this definition is less meaningful if the components of the argument or value greatly vary in magnitude. Therefore, in the sequel of this paper we will allow for an additional scaling factor

$$\mathrm{cond}_D(M)(x) := \frac{\left\|\frac{\partial y}{\partial x}(x) D\right\|_{\mathrm{F}}}{\|y\|_2} \|D^{-1}x\|_2, \tag{3}$$

with $D$ being a nonsingular diagonal matrix, $\|.\|_2$ denoting the 2-Hölder norm, and $\|.\|_{\mathrm{F}}$ the Froebenius matrix norm. [2] The reader may easily verify with aid of the Cauchy-Schwarz inequality that

$$\min_D \mathrm{cond}_D(M)(x) = \frac{\sum_{j=1}^{2n} |x_j| \left\|\frac{\partial y}{\partial x_j}(x)\right\|_2}{\|M(x)\|_2}, \tag{4}$$

with the minimum being attained for some optimal scaling matrix $D_{\mathrm{opt}}$. In what follows we will also consider the trivial scaling matrix $D = I$, as well as $D = D_{\mathrm{nor}}$ being defined by $D_{\mathrm{nor}}^2 = \mathrm{diag}(\int \hat{p}_k(t)^2 \, \mathrm{d}s(t))_{k=0,\ldots,2n-1}$ (i.e., $D_{\mathrm{nor}}^{-1} \cdot m$ is a vector of modified moments built up with orthonormal polynomials).

For the map $G_n$, Gautschi showed [8, Theorem 3.1] that

$$(\mathrm{cond}_{D_{\mathrm{nor}}} G_n)(m) = \frac{\|D_{\mathrm{nor}}^{-1} \cdot m\|_2}{\|(\lambda, \tau)\|_2} \left( \int_{\mathbb{R}} g_n(t) \, \mathrm{d}s(t) \right)^{1/2}$$

with some explicit polynomial $g_n$ of degree $4n - 2$. Later, Fischer [4, Theorems 1, 6] gave a similar formula [3] for the map $K_n$,

$$(\mathrm{cond}_{D_{\mathrm{nor}}} K_n)(m) = \frac{\|D_{\mathrm{nor}}^{-1} \cdot m\|_2}{\|(\alpha, \beta)\|_2} \left( \int_{\mathbb{R}} \omega_n(t) \, \mathrm{d}s(t) \right)^{1/2}, \tag{5}$$

$$\omega_n(x) = \sum_{j=0}^{n-1} \left\{ \frac{\beta_j^2}{4} (\pi_j^2(x) - \pi_{j-1}^2(x))^2 + (\beta_{j+1}\pi_{j+1}(x) - \beta_j \pi_{j-1}(x))^2 \pi_j^2(x) \right\}. \tag{6}$$

[2] The use of the compatible Froebenius matrix norm $\|.\|_{\mathrm{F}}$ instead of the subordinate spectral matrix norm $\|.\|_2$ is for convenience only. The reader may easily switch to an alternate definition using the fact that $\|B\|_2 \leqslant \|B\|_{\mathrm{F}} \leqslant \sqrt{n}\|B\|_2$ for any $B \in \mathbb{R}^{n \times n}$.

[3] The factor 1/4 in formula (6) does not occur in [4, Theorem 6] according to the different choice of $\beta_j$.

As mentioned already by Gautschi (without providing explicit formulas), for estimating the quantity $\mathrm{cond}_I(H_n^{-1})$ it is sufficient to apply classical results on the perturbation of eigendata while perturbing the Jacobi matrix: In the appendix we show that

$$\mathrm{cond}_I(H_n^{-1})(\alpha,\beta) \leqslant \frac{\|(\alpha,\beta)\|_2}{\|(\tau,\lambda)\|_2}\sqrt{4\sigma_0 + 2n^2 + 8\sigma_0^2 n \sum_{\substack{v,k=1\\ v\neq k}}^{n} \frac{1}{(\tau_k-\tau_v)^2}}. \tag{7}$$

As far as we know, explicit estimates for the condition number of the map $H_n$ in terms of, e.g., the weights and the nodes of the Gaussian quadrature rule (2) have been obtained only very recently[4] by Fischer [5, Section 4], we will comment on his findings in Remark 2 below. Here the situation is more involved, since a perturbation of the vector $(\lambda,\tau)$ of quadrature data does not immediately imply that we know the perturbation of the other eigendata, namely the changes of the orthonormal polynomials

$$\pi_k(x) = \hat{\pi}_k(x) \Big/ \sqrt{\int \hat{\pi}_k(t)^2\,\mathrm{d}\sigma(t)}, \quad k \geqslant 0, \tag{8}$$

satisfying the three-term recurrence

$$\beta_{n+1}\pi_{n+1}(x) = (x-\alpha_n)\pi_n(x) - \beta_n\pi_{n-1}(x), \quad n \geqslant 0, \ \pi_0(x) = \frac{1}{\beta_0}, \ \pi_{-1}(x) = 0.$$

We show in Section 2 that their behavior can be obtained by discussing the QR-decomposition of a matrix connected to the Hankel matrix of $\sigma$. However, here, we have to apply carefully Stewart's perturbation theory of QR-decomposition [17] since the underlying Hankel matrix is ill-conditioned.[5] In the second part of Section 2 we express our findings for $H_n$ and $H_n^{-1}$ in terms of the Szegö kernel of the measure $\sigma$ which enables us to determine in Corollary 8 the asymptotic behavior of their condition numbers for fairly general classes of measures $\sigma$. In Section 3, we provide lower and upper bounds for the polynomial introduced in (6). As a consequence, we show in Theorem 11 that the supports of the measures $s$ and $\sigma$ must essentially coincide in order to have a well-conditioned map $K_n$. Finally, for $\sigma$ being a continuous modification of the Chebyshev measure $s$ we establish the relation $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n) = \mathcal{O}(n^{3/2})$.

## 2. The condition of $H_n$

The aim of this part is to estimate the condition number of the nonlinear map $H_n : (\lambda,\tau) \mapsto (\alpha,\beta)$ mapping the quadrature data $(\lambda_1,\ldots,\lambda_n,\tau_1,\ldots,\tau_n)^t$ to the recurrence coefficients $(\alpha_0,\ldots,\alpha_{n-1},\beta_0,\ldots,\beta_{n-1})^t$. In what follows, we denote by $\ell_1,\ldots,\ell_n$ the Lagrange polynomials

$$\ell_j(t) = \prod_{\substack{k=1\\ k\neq j}}^{n} \frac{t-\tau_k}{\tau_j-\tau_k}. \tag{9}$$

[4] The formulas given in [8, Section 3.1] are erroneous, see [10, p. 371, footnote].

[5] We know from [3, Theorem 3.6] that the 2-condition number of a positive-definite Hankel matrix of order $n \geqslant 3$ is bounded below by $(16n)^{-1} \cdot 3.21^{n-1}$.

**Theorem 1.** *We have, using* (4),

$$\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)(\lambda,\tau) \leqslant 6\sqrt{2n}\left[n + \left(\sigma_2 \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\ell_k'(\tau_j)^2}{\lambda_k}\right)^{1/2}\right].$$

In order to prove the theorem, we will make use of the two lemmas below. Lemma 2 restates some well-known factorisations of the Jacobi matrix and the Hankel matrix in terms of the quadrature data $(\lambda,\tau)$ and the corresponding orthonormal polynomials (8). The perturbation of these polynomials will be estimated by using Lemma 3, where we discuss a special case of QR-perturbation theory.

**Lemma 2.** *We define*

$$\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n), \qquad X = \mathrm{diag}(\tau_1,\ldots,\tau_n), \qquad V = \begin{pmatrix} 1 & \tau_1 & \cdots & \tau_1^{n-1} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 1 & \tau_n & \cdots & \tau_n^{n-1} \end{pmatrix},$$

*then the Jacobi matrix has the Jordan normal form* $\mathscr{J} = Q^{\mathrm{t}} \cdot X \cdot Q$, *with*

$$Q = \Lambda^{1/2} \begin{pmatrix} \pi_0(\tau_1) & \cdots & \pi_{n-1}(\tau_1) \\ \vdots & & \vdots \\ \pi_0(\tau_n) & \cdots & \pi_{n-1}(\tau_n) \end{pmatrix},$$

*and* $Q$ *being the orthogonal left-hand factor in the QR-decomposition of the scaled Vandermonde matrix* $S = \Lambda^{1/2}V$. *Furthermore,* $S^{\mathrm{t}}S$ *is the positive-definite Hankel matrix of moments* $(\int_{\mathbb{R}} x^{j+k}\,\mathrm{d}\sigma(x))_{j,k=0,\ldots,n-1}$, *and* $\|\mathscr{J}\|_{\mathrm{F}} = \|X\|_{\mathrm{F}} = \|(\tau_1,\ldots,\tau_n)\|_2 \leqslant \|(\lambda,\tau)\|_2$.

**Proof.** The representations of $\mathscr{J}$ and $S^{\mathrm{t}} \cdot S$ as well as the orthogonality of Q may be easily verified with help of the Gaussian quadrature rule (2). Denote by $C$ the matrix containing in its $j$th column, $j = 1,\ldots,n$, the coefficients of the ascending powers of $\pi_{j-1}$. Then $C$ is upper triangular with positive elements on the diagonal, and so is $R := C^{-1}$. This yields the unique QR-decomposition $\Lambda^{1/2}V = QR$. □

In terms of Lemma 2, a modification of $(\lambda,\tau)$ implies a modification of the QR-decomposition of the matrix $S = \Lambda^{1/2}V$. In this way, we may get perturbations bounds for the Jacobi matrix and thus for $(\alpha,\beta) = H_n(\lambda,\tau)$. However, classical bounds on the perturbation of QR-factorizations involve the condition number of $S$, the square root of the condition number of the Hankel matrix $S^{\mathrm{t}} \cdot S$, which is exponentially increasing in $n$. Thus, we first have to specify the QR-bounds, by considering only rank one modifications.

**Lemma 3.** *Consider $A \in \mathbb{R}^{n\times n}$ and its perturbed counterpart $\tilde{A} = (I + e_j \cdot x^t) \cdot A$, with $x \in \mathbb{R}^n$, $\|x\|_2 < 1/2\sqrt{n}$. Then for the orthogonal factors in the two QR-factorizations $A = Q \cdot R$, $\tilde{A} = \tilde{Q} \cdot \tilde{R}$ we have the estimate*

$$\|\tilde{Q} - Q\|_F \leqslant \frac{3\|x\|_2\sqrt{n}}{1 - 2\|x\|_2\sqrt{n}}.$$

**Proof.** From $\tilde{A} = \tilde{Q} \cdot \tilde{R} = (I + e_j \cdot x^t) \cdot Q \cdot R$ we get the factorization $Q^t \cdot \tilde{Q} \cdot \tilde{R} = [I + Q^t \cdot e_j \cdot x^t \cdot Q] \cdot R$, and thus $Q_1 = Q^t \cdot \tilde{Q}$ is the orthogonal factor in the QR-decomposition of $I + E$, $E := Q^t \cdot e_j \cdot x^t \cdot Q$. Since $\|Q - \tilde{Q}\|_F = \|I - Q_1\|_F$, we are left with the problem of estimating the distance $W = I - Q_1$ of the orthogonal factors in the QR-decomposition of $I$, and of $I + E$, respectively. Our assumption on $\|E\|_F = \|x\|_2 < 1/2\sqrt{n}$ allows us to apply [17, Theorem 3.1], leading to

$$\|W\|_F \leqslant \frac{3\|E\|_F\|I\|_F}{1 - 2\|E\|_F\|I\|_F},$$

as claimed in the assertion of the Lemma. □

**Remark 4.** Following a remark on [17, p. 511], it is possible to choose different norms in [17, Theorem 3.1]. As a consequence, in the assertion of Lemma 3, the factors $\sqrt{n}$ may be dropped.

**Proof of Theorem 1.** We will apply Lemma 3 to $A^{1/2}V$

(i) In a first step, let us describe the dependence of $\beta_0$ on the quadrature data $(\lambda, \tau)$. From $\beta_0^2 = \sigma_0 = \sum_{k=1}^{n} \lambda_k$ we get

$$\frac{\partial \beta_0}{\partial \lambda_j} = \frac{1}{2\beta_0}, \qquad \frac{\partial \beta_0}{\partial \tau_j} = 0.$$

(ii) For some fixed $j \in \{1, \dots, n\}$, we now consider the perturbed quantities $\tilde{\mathscr{J}}$, $\tilde{Q}$, $\tilde{\Lambda}$, obtained by replacing $\lambda_j$ by $\tilde{\lambda}_j = \lambda_j + \varepsilon$ with some small $\varepsilon$ and $\tilde{\lambda}_k = \lambda_k$ for $k \neq j$, $\tilde{\tau}_k = \tau_k$ for $k \in \{1, \dots, n\}$. Our aim is to estimate $\|\tilde{\mathscr{J}} - \mathscr{J}\|_F$ since by (i)

$$\left\| \frac{\partial(\alpha_0, \dots, \alpha_{n-1}, \beta_0, \dots, \beta_{n-1})}{\partial \lambda_j} \right\|_2 \leqslant \lim_{\varepsilon \to 0} \frac{\|\tilde{\mathscr{J}} - \mathscr{J}\|_F}{|\varepsilon|} + 1.$$

In the sequel of the proof, we use $\doteq$ and $\dot{\leqslant}$ for first-order (in)equalities. Writing $V_k$ for the $k$th row of the Vandermonde matrix $V$ of Lemma 2, we have

$$\tilde{A} := \tilde{\Lambda}^{1/2} V \doteq \begin{pmatrix} \sqrt{\lambda_1} V_1 \\ \vdots \\ (\sqrt{\lambda_j} + \frac{\varepsilon}{2\sqrt{\lambda_j}}) V_j \\ \vdots \\ \sqrt{\lambda_n} V_n \end{pmatrix} = (I + e_j \cdot x^{\mathrm{t}}) \cdot A$$

with $x = (\varepsilon/2\lambda_j) \cdot e_j$, $A = \Lambda^{1/2} \cdot V$. An application of Lemma 3 gives

$$\|Q - \tilde{Q}\|_{\mathrm{F}} \dot{\leqslant} 3 \frac{|\varepsilon|\sqrt{n}}{2\lambda_j},$$

and thus by Lemma 2

$$\begin{aligned} \|\tilde{\mathscr{J}} - \mathscr{J}\|_{\mathrm{F}} &= \|\tilde{Q}^{\mathrm{t}} X \tilde{Q} - Q^{\mathrm{t}} X Q\|_{\mathrm{F}} = \|(\tilde{Q} - Q)^{\mathrm{t}} \cdot X \cdot \tilde{Q} + Q^{\mathrm{t}} \cdot X \cdot (\tilde{Q} - Q)\|_{\mathrm{F}} \\ &\leqslant 2\|\tilde{Q} - Q\|_{\mathrm{F}} \|X\|_{\mathrm{F}} \dot{\leqslant} 3\sqrt{n} \frac{|\varepsilon|}{\lambda_j} \|\mathscr{J}\|_{\mathrm{F}}. \end{aligned}$$

According to $\|\mathscr{J}\|_{\mathrm{F}} \leqslant \sqrt{2} \cdot \|(\alpha, \beta)\|_2$, $\beta_0 \leqslant \|(\alpha, \beta)\|_2$, we obtain for $|\varepsilon|$ tending to zero

$$\frac{\lambda_j}{\|(\alpha, \beta)\|_2} \left\| \frac{\partial(\alpha, \beta)}{\partial \lambda_j} \right\|_2 \leqslant 3\sqrt{2n} + \frac{\lambda_j}{2\beta_0^2},$$

and consequently

$$\frac{1}{\|(\alpha, \beta)\|_2} \sum_{j=1}^{n} \lambda_j \left\| \frac{\partial(\alpha, \beta)}{\partial \lambda_j} \right\|_2 \leqslant 3n\sqrt{2n} + \frac{1}{2} \leqslant 6n\sqrt{2n} - \sqrt{2n}.$$

(iii) Similarly to (ii), let $\tilde{\mathscr{J}}$, $\tilde{Q}$, $\tilde{V}$ be obtained by the data $\tilde{\tau}_j = \tau_j + \varepsilon$, $\tilde{\tau}_k = \tau_k$ for $k \neq j$, and $\tilde{\lambda}_k = \lambda_k$ for $k = 1, \ldots, n$. By (i), we have

$$\left\| \frac{\partial(\alpha_0, \ldots, \alpha_{n-1}, \beta_0, \ldots, \beta_{n-1})}{\partial \tau_j} \right\|_2 \leqslant \lim_{\varepsilon \to 0} \frac{\|\tilde{\mathscr{J}} - \mathscr{J}\|_{\mathrm{F}}}{|\varepsilon|}.$$

Note that

$$(0, 1, 2\tau_j, \ldots, (n-1)\tau_j^{n-2}) = (\ell_1'(\tau_j), \ldots, \ell_n'(\tau_j)) \cdot V,$$

and therefore $\Lambda^{1/2} \cdot \tilde{V} \doteq (I + e_j \cdot x^{\mathrm{t}}) \cdot \Lambda^{1/2} \cdot V$ with

$$x = \varepsilon \cdot \left( \frac{\sqrt{\lambda_j}}{\sqrt{\lambda_1}} \cdot \ell_1'(\tau_j), \ldots, \frac{\sqrt{\lambda_j}}{\sqrt{\lambda_n}} \cdot \ell_n'(\tau_j) \right)^{\mathrm{t}}.$$

Applying Lemma 3 leads to

$$\|\tilde{Q} - Q\|_F \dot{\leqslant} 3|\varepsilon|\sqrt{n}\sqrt{\lambda_j \cdot \sum_{k=1}^{n} \frac{\ell_k'(\tau_j)^2}{\lambda_k}},$$

and therefore

$$\begin{aligned}\|\tilde{\mathscr{J}} - \mathscr{J}\|_F &= \|(\tilde{Q} - Q)^t \tilde{X}\tilde{Q} + Q^t(\tilde{X} - X)\tilde{Q} + Q^t X(\tilde{Q} - Q)\|_F \\ &\dot{\leqslant} 2 \cdot \|\tilde{Q} - Q\|_F \cdot \|J\|_F + \|\tilde{X} - X\|_F \\ &\dot{\leqslant} 6\sqrt{n}|\varepsilon|\|\mathscr{J}\|_F \sqrt{\lambda_j \sum_{k=1}^{n} \frac{\ell_k'(\tau_j)^2}{\lambda_k}} + |\varepsilon|.\end{aligned}$$

Since $\|\tau\|_2 = \|\mathscr{J}\|_F$, we may conclude that

$$\frac{|\tau_j|}{\|(\alpha,\beta)\|_2} \left\| \frac{\partial(\alpha,\beta)}{\partial \tau_j} \right\|_2 \leqslant 6\sqrt{2n}\sqrt{\lambda_j \tau_j^2}\sqrt{\sum_{k=1}^{n} \frac{l_k'(\tau_j)^2}{\lambda_k}} + \frac{\sqrt{2}|\tau_j|}{\|\tau\|_2},$$

and, by using the Cauchy–Schwarz inequality,

$$\frac{1}{\|(\alpha,\beta)\|_2} \sum_{j=1}^{n} |\tau_j| \left\| \frac{\partial(\alpha,\beta)}{\partial \tau_j} \right\|_2 \leqslant 6\sqrt{2n}\sqrt{\sigma_2} \left( \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\ell_k'(\tau_j)^2}{\lambda_k} \right)^{1/2} + \sqrt{2n}.$$

(iv) The assertion of the theorem now follows by combining (4) with the final results of parts (ii) and (iii). □

**Remark 5.** While writing this paper, we became aware of a recent result of Fischer [5] who computed explicitly the Jacobian of the map $H_n$. Using the (completely different) tools provided in [5, Section 4, Lemma 13] one may establish two further estimates for $\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)$ as summarized below. First, in the proof of [5, Corollary 1] Fischer shows that the quantity discussed in part (ii) of the proof of Theorem 1 is uniformly bounded (if one takes the 1–Hölder norm)

$$\sum_{j=1}^{n} \frac{\lambda_j}{\|(\alpha,\beta)\|_1} \left\| \frac{\partial(\alpha,\beta)}{\partial \lambda_j} \right\|_1 \leqslant 4.$$

Hence (as already seen in our proof), the coefficients $\alpha_j, \beta_j$ are quite insensitive to perturbations in $\lambda$. For the analogue of the quantity of part (iii), Fischer provides implicitly two different bounds, the first

$$\sum_{j=1}^{n} \frac{|\tau_j|}{\|(\alpha,\beta)\|_1} \left\| \frac{\partial(\alpha,\beta)}{\partial \tau_j} \right\|_1 \leqslant \frac{\|(\tau)\|_1}{\|(\alpha,\beta)\|_1} + 6 \max_{1\leqslant j\leqslant n} |\tau_j| \max_{0\leqslant k\leqslant n-1} \left[ \sum_{v=1}^{n} \lambda_v \cdot \pi_k'(\tau_v)^2 \right]^{1/2}$$

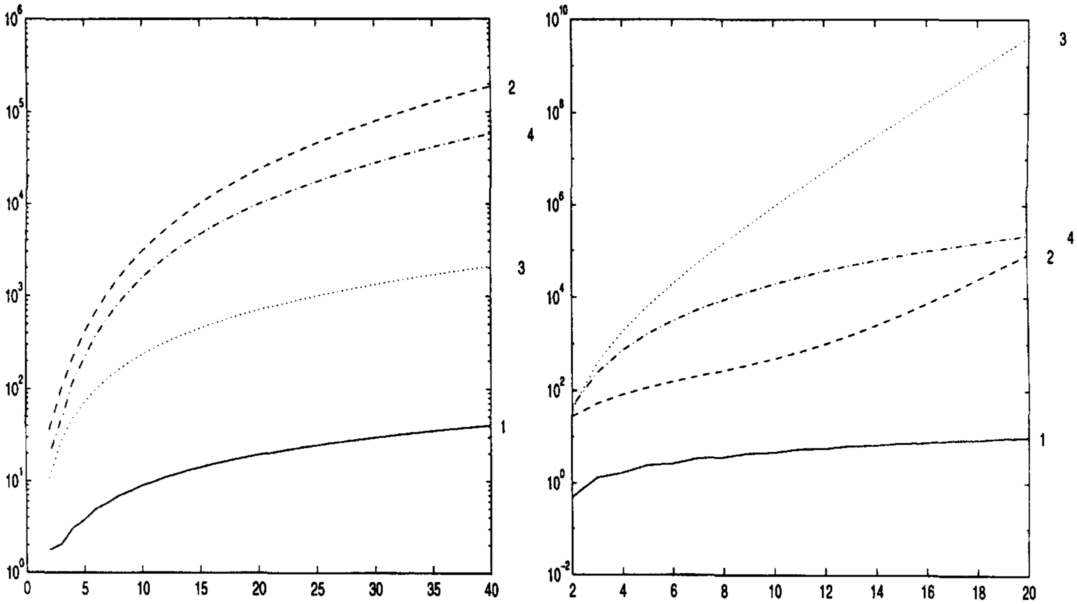


Fig. 1. The value $\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)(\lambda,\tau)$ (1: solid line), our estimate from Theorem 1 (2: dashed), Fischer's estimate in terms of norms of derivatives (3: dotted), and Fischer's estimate in terms of quadrature data (4: dashdotted). On the left-hand side we have chosen the Chebyshev quadrature data $\lambda_j = 1/n$, $\tau_j = \cos(\pi(2j-1)/(2n))$, $j=1,\dots,n$ for $2 \leqslant n \leqslant 40$. On the right-hand side we display an example where our estimate is tighter. Here we take equidistant nodes $\tau_1,\dots,\tau_n$ on $[0,1]$, $\lambda_1 = 1-(n-1)/n^5$, and $\lambda_2 = \cdots = \lambda_n = 1/n^5$, $2 \leqslant n \leqslant 20$.

in terms of the norm of $\pi_k'$, following from an argument given in the proof of [5, Corollary 1], and the second

$$\sum_{j=1}^{n} \frac{|\tau_j|}{\|(\alpha,\beta)\|_1} \left\| \frac{\partial(\alpha,\beta)}{\partial \tau_j} \right\|_1 \leqslant 2n \frac{\|(\tau)\|_1}{\|(\alpha,\beta)\|_1} + 6 \sum_{\substack{v,j=1\\ v\neq j}}^{n} (1+\sqrt{\lambda_v/\lambda_j}) \frac{|\tau_j|}{|\tau_j-\tau_v|}$$

only in terms of the quadrature data, being a consequence of [5, Lemma 13 and Proof of Corollary 2].

We have computed the exact value of $\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)(\lambda,\tau)$ for several quadrature data explicitly, and compared it with our estimate and the estimates given by Fischer. Our findings for two particular classes are displayed in Fig. 1. It seems to be impossible to know in advance which of these estimates is the tightest. Hence, according to their different nature, Fischer's estimates may be considered as complementary to Theorem 1.

In order to determine the asymptotic behavior of the condition numbers of $H_n$ and $H_n^{-1}$ for a larger class of measures, in the sequel of this section we will give estimates in terms of the Szegö kernel (see [18, (3.1.9), Theorem 14.2.2, (14.2.3)])

$$K_{n-1}^{\sigma}(t) := \sum_{j=0}^{n-1} \pi_j(t)^2 = \sum_{j=1}^{n} \frac{\ell_j(t)^2}{\lambda_j}, \qquad \kappa_n^{\sigma} = \max_{x\in \mathrm{supp}(\sigma)} K_{n-1}^{\sigma}(x), \tag{10}$$

where $\pi_j$, $\ell_j$, and $\lambda_j$, have been defined in (8), (9), and (2), respectively. In view of Theorem 1 and Eq. (7), the following lemma will be helpful:

**Lemma 6.**

$$\sum_{j=1}^{n}\left(\sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k}\right) \leqslant 9n^5\kappa_n^\sigma, \qquad \sum_{\substack{v,k=1\\ v\neq k}}^{n}\frac{1}{(\tau_k-\tau_v)^2} \leqslant 9\sigma_0 n^4\kappa_n^\sigma.$$

**Proof of Lemma 6.** The reader may easily verify, using (9) and (10), that

$$\frac{K_{n-1}^{\sigma\prime\prime}(\tau_j)}{2} = \sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k} + \frac{\ell_j''(\tau_j)}{\lambda_j}, \qquad \ell_j'(\tau_j) = \sum_{\substack{k=1\\ k\neq j}}^{n}\frac{1}{\tau_j-\tau_k},$$

$$\ell_j''(\tau_j) = \ell_j'(\tau_j)^2 - \sum_{\substack{k=1\\ k\neq j}}^{n}\frac{1}{(\tau_j-\tau_k)^2}, \qquad \lambda_j = \frac{1}{K_{n-1}^\sigma(\tau_j)}.$$

We will suppose in what follows that $[\alpha,\beta]=[-1,1]$ in order to simplify consideration. However, the case of an arbitrary compact interval is included by applying a linear transformation. By the Bernstein's inequality [12, p. 67, (5.7.4)], for any polynomial $P$ of degree $\leqslant k$ there holds

$$\max_{x\in[-1,1]}|P'(x)| \leqslant k^2 \max_{x\in[-1,1]}|P(x)|.$$

Applying this inequality twice we get $|K_{n-1}^{\sigma\prime\prime}(\tau_j)| \leqslant (2n-3)^2(2n-2)^2\kappa_n^\sigma$.

Notice that $\tau_j\in[-1,1]$, since zeros of orthogonal polynomials lie in the convex hull of the support of the corresponding measure. Furthermore

$$\left|\frac{\ell_j''(\tau_j)}{\lambda_j}\right| \leqslant (n-1)^2n^2\max_{t\in[-1,1]}\left|\frac{\ell_j(t)}{\lambda_j}\right| \leqslant (n-1)^2n^2\max_{t\in[-1,1]}\frac{1}{\sqrt{\lambda_j}}\sqrt{\sum_{k=1}^{n}\frac{\ell_k(t)^2}{\lambda_k}},$$

and thus

$$\left|\frac{\ell_j''(\tau_j)}{\lambda_j}\right| \leqslant (n-1)^2n^2\kappa_n^\sigma.$$

We are now prepared to prove the inequalities of the assertion. In fact,

$$\begin{aligned}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k}\right) &\leqslant \frac{1}{2}(2n-3)^2(2n-2)^2\kappa_n^\sigma\sum_{j=1}^{n}1 + \sum_{j=1}^{n}\frac{|\ell_j''(\tau_j)|}{\lambda_j}\\ &\leqslant \frac{1}{2}n(2n-3)^2(2n-2)^2\kappa_n^\sigma + n\max_{k=1,\dots,n}\frac{|\ell_k''(\tau_k)|}{\lambda_k}\\ &\leqslant \frac{1}{2}n(2n-3)^2(2n-2)^2\kappa_n^\sigma + n(n-1)^2n^2\kappa_n^\sigma \leqslant 9n^5\kappa_n^\sigma\end{aligned}$$

Table 1
Chebyshev polynomials of first kind (numbers between parentheses indicate decimal exponents)

| $n$ | $\sum_{j=1}^{n}\left(\sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k}\right)$ | $\sum_{\substack{v,k=1\\v\neq k}}^{n}\frac{1}{(\tau_k-\tau_v)^2}$ | $9n^5\kappa_n^\sigma$ | $9\sigma_0 n^4\kappa_n^\sigma$ |
|---|---|---|---|---|
| 5 | 3.18 (2) | 5.00 (1) | 8.95 (4) | 5.63 (4) |
| 10 | 1.29 (4) | 8.25 (2) | 5.73 (6) | 1.80 (6) |
| 20 | 4.60 (5) | 1.33 (4) | 3.67 (8) | 5.76 (7) |
| 40 | 1.55 (7) | 2.13 (5) | 2.35 (10) | 1.84 (9) |
| 80 | 5.09 (8) | 3.41 (6) | 1.50 (12) | 5.90 (10) |

Table 2
Chebyshev polynomials of second kind

| $n$ | $\sum_{j=1}^{n}\left(\sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k}\right)$ | $\sum_{\substack{v,k=1\\v\neq k}}^{n}\frac{1}{(\tau_k-\tau_v)^2}$ | $9n^5\kappa_n^\sigma$ | $9\sigma_0 n^4\kappa_n^\sigma$ |
|---|---|---|---|---|
| 5 | 9.29 (2) | 5.6 (1) | 9.85 (5) | 3.09 (5) |
| 10 | 6.97 (4) | 7.02 (2) | 2.21 (8) | 3.46 (7) |
| 20 | 6.78 (6) | 9.61 (3) | 5.26 (10) | 4.13 (9) |
| 40 | 7.57 (8) | 1.41 (5) | 1.30 (13) | 5.10 (11) |
| 80 | 9.06 (10) | 2.15 (6) | 3.26 (15) | 6.41 (13) |

and

$$\sum_{\substack{v,k=1\\v\neq k}}^{n}\frac{1}{(\tau_k-\tau_v)^2}=\sum_{j=1}^{n}(\ell_j'(\tau_j)^2-\ell_j''(\tau_j))\leqslant\sum_{j=1}^{n}\left[\lambda_j\frac{\ell_j'(\tau_j)^2}{\lambda_j}+|\ell_j''(\tau_j)|\right]$$

$$\leqslant\sum_{j=1}^{n}\left[\lambda_j\left(\sum_{k=1}^{n}\frac{\ell_k'(\tau_j)^2}{\lambda_k}\right)+|\ell_j''(\tau_j)|\right]$$

$$\leqslant\left(\sum_{j=1}^{n}\lambda_j\right)\left(\frac{(2n-3)^2(2n-2)^2}{2}+(n-1)^2n^2\right)\kappa_n^\sigma\leqslant 9\sigma_0 n^4\kappa_n^\sigma. \qquad \square$$

**Example 7.** The Chebyshev polynomials of first kind and second kind may serve to illustrate more precisely the estimates of Lemma 6, since here the quadrature data are explicitly known, and one can compute the maximum $\kappa_n^\sigma$ of the corresponding Szegö kernel on the interval $[-1,1]$. As a consequence, we may determine explicitly the two quantities studied in Lemma 6, together with their upper bounds, as displayed for these two families of orthogonal polynomials in Tables 1, and 2, respectively.

Taking into account that $1 = \sigma_0 \cdot \pi_0(x)^2 \leqslant \sigma_0 \cdot K_{n-1}^{\sigma}(x)$, by combining Lemma 6, Theorem 1, and estimate (7), we get

**Corollary 8.** *Provided that* $\mathrm{supp}(\sigma) = [-1,1]$, *there holds*

$$\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)(\lambda,\tau) \leqslant 24\sqrt{2\sigma_0}n^3\sqrt{\kappa_n^{\sigma}},$$

$$\mathrm{cond}_I(H_n^{-1})(\alpha,\beta) \leqslant \sqrt{4\sigma_0 + 6n^2 + (2 + 72\sigma_0^2(n+\sigma_0))n^5\kappa_n^{\sigma}} = \mathcal{O}(n^3\sqrt{\kappa_n^{\sigma}}).$$

**Remark 9.** It is not difficult to show that both condition numbers discussed in Corollary 8 are bounded below by $C \cdot n^{-1/2}$ with suitable constants $C$. Consequently, if in addition $\sigma \in \mathbf{Reg}$ (see Section 3 below) then the $n$th root of both condition numbers tends to 1.

**Example 10.** In order to illustrate Corollary 8, we will discuss more generally Jacobi polynomials associated with the measure $w(x) = (x-1)^{\alpha}(x+1)^{\beta}$, $\mathrm{d}\sigma(x) = w(x)\,\mathrm{d}x$ on $[-1,1]$, with $\gamma := \max(\alpha,\beta) \geqslant -\frac{1}{2}$. According to [18, Eqs. (4.5.8), (7.32.2)],

$$K_{n-1}^{w}(1) = 2^{-\alpha-\beta-1}\frac{\Gamma(n+\alpha+\beta+1)\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\Gamma(\alpha+2)\Gamma(n)\Gamma(n+\beta)}, \qquad \kappa_n^{w} = \max\{K_{n-1}^{w}(-1), K_{n-1}^{w}(1)\},$$

showing that $\kappa_n^{w} = \mathcal{O}(n^{2\gamma+2})$. Thus by Corollary 8 we obtain

$$\mathrm{cond}_{D_{\mathrm{opt}}}(H_n)(\lambda,\tau) = \mathcal{O}(n^{\gamma+4}), \quad \text{and} \quad \mathrm{cond}_I(H_n^{-1})(\alpha,\beta) = \mathcal{O}(n^{\gamma+4}).$$

## 3. The condition of $K_n$

The aim of this section is to provide lower and upper bounds for the quantity $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$ for measures $\sigma$, $s$ with compact supports $\mathrm{supp}(\sigma)$, and $\mathrm{supp}(s)$, by exploiting the explicit formulas (5) and (6) given by Fischer. As shown in Theorem 11 below, both supports have essentially to coincide in order to insure non-exponential growth of $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$ as a function of $n$. In what follows, we will express growth rates in terms of the Green function $g_S$ of a compact set, being positive outside of $S$, e.g.,

$$g_{[a,b]}(z) = \log\left|\frac{2z-a-b}{b-a} + \sqrt{\left(\frac{2z-a-b}{b-a}\right)^2 - 1}\right|$$

is the Green function of the interval $[a,b]$. Also, we will consider in particular measures of the set $\mathbf{Reg}$ being introduced in [16], several characterization may be found in [16, Sections 3,4]. For instance, if $\sigma \in \mathbf{Reg}$ then $(\pi_n^{1/n})_{n\geqslant 0}$ converges locally uniformly outside the convex hull of $\mathrm{supp}(\sigma)$ to $\exp(g_{\mathrm{supp}(\sigma)}(z))$. Let us recall from [16, Theorem 4.1.1] that, e.g., measures $\mathrm{d}\sigma(x) = w(x)\,\mathrm{d}x$ supported on an interval with $w > 0$ satisfy $\sigma \in \mathbf{Reg}$, and the same remains true if we add a finite number of mass points to $\sigma$.

**Theorem 11.** (a) *If* $\operatorname{supp}(s) \not\subset \operatorname{supp}(\sigma)$ *then* $\operatorname{cond}_{D_{\rm nor}}(K_n)$ *grows at least with the geometric rate*

$$\liminf_{n\to\infty} \operatorname{cond}_{D_{\rm nor}}(K_n)^{1/n} \geqslant \max_{x\in\operatorname{supp}(s)} e^{2g_{\operatorname{supp}(\sigma)}(x)} > 1.$$

(b) *If* $s \in \mathbf{Reg}$ *is supported on an interval and if* $\operatorname{supp}(\sigma) \not\subset \operatorname{supp}(s)$ *then* $\operatorname{cond}_{D_{\rm nor}}(K_n)$ *grows at least with the geometric rate*

$$\liminf_{n\to\infty} \operatorname{cond}_{D_{\rm nor}}(K_n)^{1/n} \geqslant \max_{x\in\operatorname{supp}(\sigma)} e^{g_{\operatorname{supp}(s)}(x)} > 1.$$

(c) *If* $S = \operatorname{supp}(s) = \operatorname{supp}(\sigma)$ *then with* $M := \max_{x\in S}|x|$,

$$\frac{\sqrt{\sigma_0}/2}{\sqrt{\sigma_0}+\sqrt{n}\cdot M} \leqslant \operatorname{cond}_{D_{\rm nor}}(K_n) \leqslant \sqrt{8}M\left[\int K_n^\sigma(t)K_{n-1}^\sigma(t)\,\mathrm{d}s(t)\int K_{2n-1}^s(x)\,\mathrm{d}\sigma(x)\right]^{1/2}.$$

*holds. If in addition* $s,\sigma\in\mathbf{Reg}$ *and if* $S$ *is a set being regular with respect to the Dirichlet problem (e.g., a finite union of intervals) then* $\lim_{n\to\infty}\operatorname{cond}_{D_{\rm nor}}(K_n)^{1/n} = 1$.

As our starting point in the investigation of the numerical condition of the nonlinear map $K_n$, we conjectured that there is a link to the Euclidean condition number of the upper triangular matrix of transmission coefficients

$$T_n(\sigma,s) = \left(\int \pi_j(x)p_k(x)\,\mathrm{d}\sigma(x)\right)_{j,k=0,\dots,n},$$

which allows to express $p_k$ in terms of $\pi_0,\dots,\pi_k$ for $k=0,1,\dots,n$. In fact, the elements of this matrix occur (up to a scaling) as intermediate quantities in the modified Chebyshev algorithm. Thus, we suspect that for the computation of the recurrence coefficients via the modified Chebyshev algorithm in a numerically stable way it is necessary that the elements of $T_{2n-1}(\sigma,s)$ (and probably also of its inverse $T_{2n-1}(\sigma,s)^{-1} = T_{2n-1}(s,\sigma)$) are not "too large".

Estimates for the condition number of such matrices and of Gram matrices $T_n(\sigma,s)^{\rm T}T_n(\sigma,s) = (\int p_j(x)p_k(x)\,\mathrm{d}\sigma(x))$ have been given in [1, Sections 3 and 5], it is quite instructive to compare them with the findings of Theorem 11: it is shown for instance in [1, Corollary 3.6] that the condition number will increase exponentially if the supports of $s$ and $\sigma$ do not coincide. Also, there is an analogue of the upper bound of Theorem 11(c): From [1, Eq. (3.8)] we get

$$\|T_n(\sigma,s)\|_{\rm F}\|T_n(\sigma,s)^{-1}\|_{\rm F} = \left[\int K_n^\sigma(t)\,\mathrm{d}s(t)\int K_n^s(x)\,\mathrm{d}\sigma(x)\right]^{1/2}.$$

In order to prove Theorem 11, we first need to give lower and upper bounds for the polynomial $\omega_n$ introduced in (6) in terms of the Szegö kernel for which we have information on its asymptotic behavior.

**Lemma 12.** *We have for* $n\geqslant 1$

$$\frac{\operatorname{dist}(x,\operatorname{supp}(\sigma))^2}{2}\sin^2\left(\frac{\pi}{4n+2}\right)K_{n-1}^\sigma(x)K_{n-2}^\sigma(x) \leqslant \omega_n(x) \leqslant 2(|x|+\|\mathscr{J}\|_2)^2K_n^\sigma(x)K_{n-1}^\sigma(x).$$

**Proof.** By examining the singular values of the associated matrix one verifies that for any real numbers $a_0 = 0, a_1, a_2, \dots, a_n$

$$2\sin^2\left(\frac{\pi}{2n+4}\right)\sum_{j=1}^{n} a_j^2 \leqslant \sum_{j=1}^{n}(a_j - a_{j-1})^2 \leqslant 2\sum_{j=1}^{n} a_j^2.$$

holds. Consequently,

$$\begin{aligned}
\omega_n(x) &\geqslant \sum_{j=0}^{n-1}\left\{\frac{\beta_j^2}{4}(\pi_j(x)^2 - \pi_{j-1}(x)^2)^2\right\} + 2\sin^2\left(\frac{\pi}{4n+2}\right)\sum_{j=0}^{n-1}(\beta_{j+1}^2\pi_{j+1}(x)^2\pi_j(x)^2)\\
&\geqslant 2\sin^2\left(\frac{\pi}{4n+2}\right)\left[\frac{\beta_0^2}{4}(\pi_0(x)^2)^2 + \beta_n^2\pi_n(x)^2\pi_{n-1}(x)^2 + \sum_{j=1}^{n-1}\frac{\beta_j^2}{4}(\pi_j(x)^2 + \pi_{j-1}(x)^2)^2\right]\\
&\geqslant \frac{1}{2}\sin^2\left(\frac{\pi}{4n+2}\right)\left[\sum_{j=0}^{n-1}\beta_j^2\,(\pi_j(x)^2 + \pi_{j-1}(x)^2)(K_j^\sigma(x) - K_{j-2}^\sigma(x))\right],
\end{aligned}$$

and similarly

$$\omega_n(x) \leqslant 2\left[\beta_n^2\pi_n^2(x)(K_n^\sigma(x) - K_{n-2}^\sigma(x)) + \sum_{j=0}^{n-1}(K_j^\sigma(x) - K_{j-2}^\sigma(x))\beta_j^2(\pi_j^2(x) + \pi_{j-1}^2(x))\right].$$

The remaining polynomials may now be estimated in terms of the Jacobi matrix: observing that

$$(xI - \mathcal{J})(\pi_0(x), \dots, \pi_{j-1}(x), 0, \dots, 0)^t = (\underbrace{0, \dots, 0}_{j-1}, \beta_j\pi_j(x), -\beta_j\pi_{j-1}, (x)0, \dots, 0)^t$$

for $j = 0, 1, \dots, n-1$ (and also for $j = n$ if we drop the last term) we obtain $\beta_j^2(\pi_j(x)^2 + \pi_{j-1}(x)^2) \leqslant \|xI - \mathcal{J}\|_2^2 K_{j-1}^\sigma(x) \leqslant (|x| + \|\mathcal{J}\|_2)^2 K_{j-1}^\sigma(x)$, as required for the upper bound for $\omega_n$ in the assertion. For establishing the lower bound, we just need to recall from [2, Proof of Theorem 2.1] that $\|R(x)\|^2 \cdot \beta_j^2(\pi_j(x)^2 + \pi_{j-1}(x)^2) \geqslant K_{j-1}^\sigma(x)$ for all $j \geqslant 0$ and for $x \notin \operatorname{supp}(\sigma)$, where $R$ denotes the resolvent operator of the (self-adjoint) Jacobi operator associated to $\sigma$, and thus $\|R(x)\| = 1/\operatorname{dist}(x, \operatorname{supp}(\sigma))$. $\square$

**Remark 13.** We see from the proof that the term $\operatorname{dist}(x, \operatorname{supp}(\sigma))$ in Lemma 12 may be replaced by the distance of $x$ to the set of zeros of $\pi_N$ for any $N \geqslant n$. Consequently, the quantity $\int K_n^\sigma(x)^2\,\mathrm{d}s(x)$ is a quite good estimate for the magnitude of $\int \omega_n(x)\,\mathrm{d}s(x)$ in (5). Also, for $\sigma$ being the (unit) measure of Chebyshev polynomials of the first kind we obtain from Lemma 12 the rough estimate $\omega_n(x) \leqslant 3 \cdot (n+2)^2$, for $x \in \operatorname{supp}(\sigma)$ whereas a more careful analysis [4, Theorem 8] gives $\frac{1}{4} \leqslant \omega_n(x) \leqslant 2n$.

In what follows, the following statement will be helpful: If $(f_n)_n$ is a sequence of continuous functions which converge to a function $f$ uniformly on compact subsets of the complex plane then

$$\lim_{n\to\infty}\left(\int_{\mathbb{R}}|f_n(t)|^n\,\mathrm{d}\sigma(t)\right)^{1/n} = \max_{t\in\operatorname{supp}(\sigma)}|f(t)|. \tag{11}$$

Relation (11) is well-known in the case $f_n = f$, $n \geqslant 0$; a proof for the general case is left to the reader.

**Proof of Theorem 11(a).** We first notice that

$$\frac{\|D_{\text{nor}}^{-1}m\|_2}{\|(\alpha,\beta)\|_2} \geqslant \frac{|\int p_0(x)\,d\sigma(x)|}{\beta_0 + \|\mathscr{J}\|_F} \geqslant \frac{\sigma_0/\sqrt{s_0}}{\sqrt{\sigma_0} + \sqrt{n}\max_{x\in\text{supp}(\sigma)}|x|}. \tag{12}$$

By assumption there exist a closed set $V \subset \text{supp}(s)\backslash\text{supp}(\sigma)$. Using (5) and Lemma 12 we obtain

$$\text{cond}_{D_{\text{nor}}}(K_n) \geqslant \frac{\sigma_0 \text{dist}(V, \text{supp}(\sigma)) \sin(\frac{\pi}{4n+2})}{2\sqrt{s_0}(\sqrt{\sigma_0} + \sqrt{n}\max_{x\in\text{supp}(\sigma)}|x|)} \left[\int_V K_{n-1}^\sigma(x) K_{n-2}^\sigma(x)\,ds(t)\right]^{1/2},$$

and thus it remains to discuss $\int_V K_n^\sigma(z)^2\,ds(x)$. In what follows, we denote the $L_p(\sigma)$-norms, $p \in [1,\infty]$ by

$$\|f\|_{\sigma,p} := \left[\int f(t)^p\,d\sigma(t)\right]^{1/p}, \qquad \|f\|_{\sigma,\infty} := \max_{t\in\text{supp}(\sigma)} |f(t)|.$$

Since $\|f\|_{\sigma,p_1} \leqslant \sigma_0^{1/p_1 - 1/p_2}\|f\|_{\sigma,p_2}$ for $p_1 \leqslant p_2$, we have

$$\sqrt{\sigma_0}\sqrt{K_n^\sigma(x)} = \sqrt{\sigma_0}\max_{\deg P\leqslant n}\frac{|P(x)|}{\|P\|_{2,\sigma}} \geqslant \max_{\deg P\leqslant n}\frac{|P(x)|}{\|P\|_{\infty,\sigma}} =: \Delta_n^\sigma(x).$$

It is well-known that $(\Delta_n^\sigma(x)^{1/n})_{n\geqslant 0}$ converges locally uniformly outside of $\text{supp}(\sigma)$ to $\exp(g_{\text{supp}(\sigma)}(x))$. Taking into account (11), we obtain

$$\liminf_{n\to\infty}\left[\int_V K_n^\sigma(z)^2\,ds(x)\right]^{1/2n} \geqslant \lim_{n\to\infty}\left[\int_V \Delta_n^\sigma(z)^4\,ds(x)\right]^{1/2n} = \max_{x\in V}\exp(g_{\text{supp}(\sigma)}(x))^2,$$

as claimed in the assertion. □

**Proof of Theorem 11(b).** Since $\omega_n(x) \geqslant \omega_1(x) \geqslant 1/(4\sigma_0)$ and $\|D_{\text{nor}}^{-1}m\|_2 \geqslant |\int p_{2n-2}(x)\,d\sigma(x)|$, the assertion follows by showing that the latter quantity has the specified exponential growth. The orthonormal polynomial $p_{2n-2}$ has all its zeros in the compact interval $\text{supp}(s) =: S$, and thus is positive outside of $S$. Therefore,

$$\left|\int p_{2n-2}(x)\,d\sigma(x)\right| \geqslant \int_{\text{supp}(\sigma)\backslash S} |p_{2n-2}(x)|\,d\sigma(x) - \int_S |p_{2n-2}(x)|\,d\sigma(x).$$

Since $S$ is a convex set being regular with respect to the Dirichlet problem in $\mathbb{C}\backslash S$, we know from [16, Theorems 3.1.1(ii) and 3.2.3(iv)] that $\limsup_{n\to\infty}|p_n(x)|^{1/n} \leqslant 1$ uniformly in $S$, and that $(|p_n(x)|^{1/n})_{n\geqslant 0}$ tends to $\exp(g_S(x))$ locally uniformly outside of $S$. From (11) it follows that

$$\lim_{n\to\infty}\left|\int p_{2n-2}(x)\,d\sigma(x)\right|^{1/n} = \max_{x\in\text{supp}(\sigma)}\exp(2g_S(x)),$$

again as claimed in the assertion. □

**Proof of Theorem 11(c).** The lower bound for $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$ follows from (12) together with the inequality $\omega_n(x) \geqslant 1/(4\sigma_0)$. For the upper bound we first notice that $\|(\alpha,\beta)\|_2 \geqslant \sqrt{\sigma_0}$, and $|x| + \|\mathscr{J}\|_2 \leqslant 2M$ for $x \in S$. Also,

$$\frac{\|D_{\mathrm{nor}}^{-1}m\|_2}{\|(\alpha,\beta)\|_2} \leqslant \frac{\|(\|p_k\|_{\sigma,1})_{k=0,\ldots,2n-1}\|_2}{\sqrt{\sigma_0}} \leqslant \|(\|p_k\|_{\sigma,2})_{k=0,\ldots,2n-1}\|_2 = \left[\int K_{2n-1}^s(x)\,\mathrm{d}\sigma(x)\right]^{1/2}.$$

Combining these estimates with Lemma 12 yields the claimed upper bound for $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$. In order to show the last part we apply again (11), since both $(K_n^\sigma(x)^{1/n})_{n\geqslant 0}$ and $(K_n^s(x)^{1/n})_{n\geqslant 0}$ converge to the constant 1 uniformly in $S$ by [16, Theorem 3.2.3(iii)]. □

**Remark 14.** The last sentence of Theorem 11 remains valid in the case where we have a partition $S = S_1 \cup S_2$, with $S_1$ being regular with respect to the Dirichlet problem, and $S_2$ containing a finite number of additional mass points. In fact, it is known that both $(K_n^\sigma(x))_{n\geqslant 0}$ and $(K_n^s(x))_{n\geqslant 0}$ converge for isolated mass points $x \in S_2$. Moreover, by a slight extension of the arguments used in [16, Theorem 3.2.3] we also have that both $(K_n^\sigma(x)^{1/n})_{n\geqslant 0}$ and $(K_n^s(x)^{1/n})_{n\geqslant 0}$ converge to the constant 1 uniformly in $S_1$.

Theorem 11(c) may also be used to derive (quite sharp) estimates for $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$ in the case where $\sigma$ is just a (bounded) modification of $s$ – a very typical situation for employing successfully modified moments (see, e.g., [8, 13]).

**Example 15.** Consider the case $\mathrm{d}\sigma(x) = \rho(x)\cdot \mathrm{d}s(x)$, $x \in \mathrm{supp}(s) =: S$, where we suppose the existence of constants $\rho_{\min}, \rho_{\max}$ with $0 < \rho_{\min} \leqslant \rho(x) \leqslant \rho_{\max}$ for all $x \in S$. Then $\sqrt{\rho_{\min}}\|P\|_{2,s} \leqslant \|P\|_{2,\sigma} \leqslant \sqrt{\rho_{\max}}\|P\|_{2,s}$ for any polynomial, and consequently

$$\rho_{\min} K_n^\sigma(x) \leqslant K_n^s(x) \leqslant \rho_{\max} K_n^\sigma(x), \quad x \in \mathbb{R},\ n \geqslant 0.$$

In particular, we have $\int K_{2n-1}^s(x)\,\mathrm{d}\sigma(x) \leqslant 2n\rho_{\max}$, and from Theorem 11(c) we may conclude that

$$\mathrm{cond}_{D_{\mathrm{nor}}}(K_n) \leqslant \frac{4\sqrt{n\rho_{\max}}}{\rho_{\min}} \max_{x\in S}|x| \left[\int K_n^s(t)K_{n-1}^s(t)\,\mathrm{d}s(t)\right]^{1/2}. \tag{13}$$

As mentioned before in Remark 13, we believe that this estimate is sharp: the quantity $\mathrm{cond}_{D_{\mathrm{nor}}}(K_n)$ is probably also bounded below[6] by a constant times $n^{-1}[\int K_{n-1}^s(t)K_{n-2}^s(t)\,\mathrm{d}s(t)]^{1/2}$. We are not able to give for general $s$ an upper bound for this integral, which apparently may become quite large for increasing $n$.

However, according to their simplicity, for the auxiliary polynomials $p_n$ one chooses quite often (shifted) Chebyshev polynomials: for the (unit) measure $\mathrm{d}s(x) = 1/(\pi\sqrt{1-x^2})$ on $[-1,1]$ we have $p_0 = 1$ and $p_n = \sqrt{2}T_n$ for $n \geqslant 1$, and thus $\int K_n^s(t)^2\,\mathrm{d}s(t) \leqslant (n+1)(n+2)$, showing that in this case

$$\mathrm{cond}_{D_{\mathrm{nor}}}(K_n) \leqslant \frac{4(n+2)\sqrt{n\rho_{\max}}}{\rho_{\min}} \max_{x\in S}|x|. \tag{14}$$

[6] As one may easily see from the proof of Lemma 12, this is for instance true if the $\beta_j$ are bounded away from zero.

Estimate (14) confirms an observation made by Gautschi [9, Example 4.1]: one may compute quite accurately the recurrence coefficients of the measure $d\sigma(z) = ds(z)/\sqrt{1 - k^2x^2}$ $(k > 1)$ via modified moments based on Chebyshev polynomials. In fact, here one also easily shows [1, Section 3.2] that the Euclidean condition number of the underlying Gram matrix (or matrix of transmission coefficients) is bounded independently of $n$.

## Appendix. The condition of $H_n^{-1}$

For the sake of completeness, we add a proof for the estimate (7) of $\mathrm{cond}_I(H_n^{-1})$. Recall from Lemma 2 that $(\lambda, \tau)$ are obtained as eigendata from the Jacobi matrix (1), more precisely, $\mathscr{J} x_k = \tau_k x_k$, $k = 1, \dots, n$, with $\|x_k\|_2 = 1$, $\lambda_k = \beta_0^2 (e_1^{\mathrm{t}} \cdot x_k)^2$. Thus, in order to prove (7), we are left with the classical problem of estimating the changes of eigendata if $\mathscr{J}$ is perturbed by $\varepsilon \cdot F$ for some small $\varepsilon \in \mathbb{R}$ (see, e.g., [11, Sections 7.2.2, 7.2.4, pp. 344–346]). In particular, it is natural that an upper bound of $\mathrm{cond}_I(H_n^{-1})$ involves the minimal mutual distance of the eigenvalues of $\mathscr{J}$.

Writing more explicitly $(\mathscr{J} + \varepsilon F) x_k(\varepsilon) = \tau_k(\varepsilon) x_k(\varepsilon)$, it is known that

$$\dot{\tau}_k(0) = x_k^{\mathrm{t}} F x_k \quad \text{and} \quad \dot{x}_k(0) = \sum_{\substack{v=1 \\ v \neq k}}^{n} \frac{x_v^{\mathrm{t}} F x_k}{\tau_k - \tau_v} x_v.$$

In particular, if $\|F\|_2 \leqslant 1$, using the Cauchy–Schwarz inequality we obtain the simple estimates

$$|\dot{\tau}_k(0)| \leqslant 1 \quad \text{and} \quad |e_1^{\mathrm{t}} \dot{x}_k(0)| \leqslant \sqrt{\sum_{\substack{v=1 \\ v \neq k}}^{n} \frac{1}{(\tau_k - \tau_v)^2}}. \tag{15}$$

If now the quantity $\alpha_j$ in $\mathscr{J}$ is replaced by $\tilde{\alpha}_j = \alpha_j + \varepsilon$ for a $j \in \{0, \dots, n-1\}$, we perturb $\mathscr{J}$ by $\varepsilon F$, $F = e_{j+1} \cdot e_{j+1}^{\mathrm{t}}$, with $\|F\|_2 = 1$. Thus, by (15),

$$\left\| \frac{\partial \tau}{\partial \alpha_j} \right\|_2 \leqslant \sqrt{n}, \qquad \left\| \frac{\partial \lambda}{\partial \alpha_j} \right\|_2 \leqslant 2\sigma_0 \sqrt{\sum_{\substack{v,k=1 \\ v \neq k}}^{n} \frac{1}{(\tau_k - \tau_v)^2}}.$$

The same bounds are valid if $\beta_j$ is replaced by $\tilde{\beta}_j = \beta_j + \varepsilon$ for a $j \in \{1, \dots, n-1\}$, i.e., $\mathscr{J}$ is perturbed by $\varepsilon \cdot F$, $F = e_j \cdot e_{j+1}^{\mathrm{t}} + e_{j+1} \cdot e_j^{\mathrm{t}}$, with $\|F\|_2 = 1$. In addition, since $\mathscr{J}$ does not depend on $\beta_0$, we immediately get

$$\frac{\partial \tau}{\partial \beta_0} = 0 \quad \text{and} \quad \left\| \frac{\partial \lambda}{\partial \beta_0} \right\|_2 \leqslant 2\beta_0.$$

Consequently,

$$\left\| \frac{\partial(\lambda, \tau)}{\partial(\alpha, \beta)} \right\|_{\mathrm{F}}^2 \leqslant 4\sigma_0 + 2n^2 + 8\sigma_0^2 n \sum_{\substack{v,k=1 \\ v \neq k}}^{n} \frac{1}{(\tau_k - \tau_v)^2}$$

as claimed in assertion (7).

## References


[1] B. Beckermann, On the numerical condition of polynomial bases: estimates for the condition number of Vandermonde, Krylov and Hankel matrices, Habilitation Thesis, Universität Hannover, 1995.

[2] B. Beckermann, On the Classification of the Spectrum of Second Order Difference Operators, Publication ANO 379, Université de Lille, 1997.

[3] B. Beckermann, The Condition Number of real Vandermonde, Krylov and Positive Definite Hankel Matrices, Publication ANO 380, Université de Lille, 1997.

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

[5] H.-J. Fischer, On generating orthogonal polynomials for discrete measures, preprint 1997.

[6] W. Gautschi, Construction of Gauss–Christoffel quadrature formulas, Math. Comput., 22 (1968) 251–270.

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

[8] W. Gautschi, On generating orthogonal polynomials, Siam J. Sci. Statist. Comput. 3(3) (1982) 298–317.

[9] W. Gautschi, Orthogonal polynomials – Constructive theory and applications, J. Comput. Appl. Math. 12&13 (1985) 61–76.

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

[11] G.H. Golub, C.F. Van Loan, Matrix Computations, 2nd ed., Johns Hopkins University Press, Baltimore/London, 1993.

[12] G. Meinardus, Approximation of Functions, Theory and Numerical Methods, Springer, Berlin, 1967.

[13] M.M. Cecchi, M.R. Zaglia, Computing the coefficients of a recurrence formula for numerical integration by moments and modified moments, J. Comput. Appl. Math. 49 (1–3) (1993) 207–216.

[14] W.H. Press, B.P. Flannery, S. Teukolsky, W.T. Vetterling, Numerical Recipes, The art of Scientific Computing, Cambridge University Press, Cambridge, 1986.

[15] R.A. Sack, A.F. Donovan, An algorithm for Gaussian quadrature given modified moments, Numer. Math. 18 (1971/72) 465–478.

[16] H. Stahl, V. Totik, General Orthogonal polynomials, Cambridge University Press, New York, 1992.

[17] G.W. Stewart, Perturbation bounds for the QR factorization of a matrix, SIAM J. Numer. Anal. 14(3) (1977) 509–517.

[18] G. Szegö, Orthogonal Polynomials, 4th ed., AMS Colloquium Publications, vol. 23, American Mathematical Society, Providence, RI, 1975.

[19] J.C. Wheeler, Modified moments and Gaussian quadrature, Rocky Mountain J. Math. 4 (1974) 287–296.