Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 396 (2005) 103–125
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On the sensitivity of Lanczos recursions to the spectrum

Vladimir Druskin [a], Liliana Borcea [b], Leonid Knizhnerman [c,*]

[a] *Schlumberger-Doll Research Center, Old Quarry Road, Ridgefield, CT 06877-4108, United States*
[b] *Computational and Applied Mathematics, MS 134, Rice University, 6100 Main Street, Houston, TX 77005-1892, United States*
[c] *Central Geophysical Expedition, House 38, Bldg. 3, Narodnogo Opolcheniya Street, Moscow 123298, Russia*

Received 21 May 2004; accepted 31 August 2004

Submitted by R.A. Brualdi

Abstract

We obtain novel, explicit formulas for the sensitivity of Jacobi matrices to small perturbations of their spectra. Our derivation is based on the connection between Lanczos's algorithm and the discrete Gel'fand–Levitan inverse spectral method. We prove uniform stability of Lanczos recursions in discrete primitive norms, for perturbations of the eigenvalues relative to their separations. A stronger, l^1 norm stability bound is also derived, under additional assumptions of rate of decay of the perturbations of the spectrum, which arise naturally for Sturm–Liouville operators.

AMS classification: 65F10; 65F15; 65F18; 34B24

Keywords: Lanczos recurrence; Jacobi inverse eigenvalue problem; Discrete Gel'fand–Levitan method; Orthogonal polynomials

* Corresponding author.
E-mail addresses: vdruskin@ridgefield.oilfield.slb.com (V. Druskin), borcea@caam.rice.edu (L. Borcea), mmd@cge.ru (L. Knizhnerman).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.08.031

1. Introduction

Let $\mathscr{T}$ be a real, symmetric, tridiagonal $n \times n$ (Jacobi) matrix, with entries α_i and $\beta_j > 0$ on the diagonal and subdiagonal, respectively, for $i = 1, \ldots, n$ and $j = 1, \ldots, n-1$. Let also $\lambda_i \in \mathbb{R}$ and $\mathbf{s}_i \in \mathbb{R}^n$, $i = 1, \ldots, n$, be the eigenvalues and orthonormal eigenvectors of $\mathscr{T}$, such that the matrices $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $S = (\mathbf{s}_1, \ldots, \mathbf{s}_n)$ satisfy

$$\mathscr{T} S = S\Lambda.$$

The inverse problem of finding $\mathscr{T}$ from spectral data Λ and $\mathbf{e}_1^{\mathrm{T}} S$ (the first row of S) is classic one, and it is solved uniquely [1–7] and efficiently, with a Lanczos recursion for matrix Λ, and initial vector $\mathbf{e}_1^{\mathrm{T}} S$ [8]. Equivalently, $\mathscr{T}$ is determined by the coefficients in the three term recursion of polynomials orthogonal with respect to the Stieltjes measure

$$\sum_{p=1}^{n} S_{1p}^2 H(t - \lambda_p), \tag{1}$$

where H is the Heaviside (step) function and S_{1p} are the entries in $\mathbf{e}_1^{\mathrm{T}} S$ [3,9–11]. In this paper, we study the sensitivity of the entries of $\mathscr{T}$, with respect to perturbations of Λ and $\mathbf{e}_1^{\mathrm{T}} S$. In particular, we seek stability bounds which do not depend on the dimension n of the problem.

We shall use the following notation convention: Matrices are denoted by capital letters and the entries of, say matrix S, are S_{ij}. All vectors are denoted with small, boldfaced letters. We let $\mathbf{e}_j$, for $j = 1, \ldots, n$, be the canonical basis vectors of $\mathbb{R}^n$ (columns of the identity matrix) and we use the super index T for transpose.

Let us define perturbed spectral data

$$\widetilde{\Lambda} = \Lambda + \mathrm{d}\Lambda, \quad \mathbf{e}_1^{\mathrm{T}} \widetilde{S} = \mathbf{e}_1^{\mathrm{T}} S + \mathbf{e}_1^{\mathrm{T}}\, \mathrm{d}S,$$

corresponding (in the linearized sense) to a symmetric, tridiagonal matrix

$$\widetilde{\mathscr{T}} = \mathscr{T} + \mathrm{d}\mathscr{T},$$

where the ‘d’ symbol denotes a differential. Standard perturbation theory gives [12]

$$\mathrm{d}\lambda_i = \mathbf{s}_i^{\mathrm{T}}\, \mathrm{d}\mathscr{T}\, \mathbf{s}_i, \quad \mathrm{d}S_{1i} = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\mathbf{s}_j^{\mathrm{T}}\, \mathrm{d}\mathscr{T}\, \mathbf{s}_i}{\lambda_i - \lambda_j} S_{1j}, \quad i = 1, \ldots, n,$$

but explicit compact formulas for $\mathrm{d}\mathscr{T}$, in terms of $\mathrm{d}\Lambda$ and $\mathbf{e}_1^{\mathrm{T}}\, \mathrm{d}S$ have not been known. Instead, various stability bounds on $\mathrm{d}\mathscr{T}$ have been derived [10,13–16], but they exhibit at least polynomial growth in the space dimension n.

In this paper, we derive explicit formulas for perturbation $\mathrm{d}\mathscr{T}$, using a Gel’fand–Levitan approach [17–20], in discrete form [21,22], as introduced by Natterer in

[22]. A number of new stability estimates for Lanczos recursions follow directly from the formulas derived here. In particular, we obtain uniform stability bounds, in a discrete primitive norm, for perturbations of the eigenvalues relative to their separations. Moreover, under additional assumptions on rates of decay of the perturbations of the spectral data, which arise naturally in inverse Sturm–Liouville problems [18], we obtain a stronger uniform stability estimate, in the l^1 norm.

This paper is organized as follows: in Section 2, we review the discrete Gel'fand–Levitan formulation due to Natterer [22], which we then use in our derivation of explicit formulas for $\mathrm{d}\mathscr{T}$, in Section 3. We give the stability estimates of Lanczos recursions in Section 4. For a special class of Jacobi matrices, that arise in the discretization of Sturm–Liouville equations, we show that the stability estimates are independent of the dimension n of the problem. Concluding remarks are given in Section 6. In Appendix A, we show an alternative proof of the sensitivity formulae, using the theory of orthogonal polynomials. Finally, in Appendix B, we explain how to extend the stability analysis to the classic version of Jacobi inverse eigenvalue problems, where, instead of eigenvector components $\mathbf{e}_1^{\mathrm{T}} S$, we specify the eigenvalues of the right lower $(n-1) \times (n-1)$ block of $\mathscr{T}$.

2. Discrete Gel'fand–Levitan method

In this section, we review the discrete Gel'fand–Levitan algorithm derived by Natterer [22], in a slightly modified form, tailored to our objectives, and discuss its connection with the continuum setting. We should point out that the perturbations do not have to be small in this section.

2.1. Perturbations and transmutation matrices

Let us introduce a so-called transmutation matrix $G \in \mathbb{R}^{n \times n}$, satisfying

$$EG\widetilde{\mathscr{T}} - E\mathscr{T}G = 0, \quad \mathbf{e}_1^{\mathrm{T}} G = \mathbf{e}_1^{\mathrm{T}}, \tag{2}$$

where $E = I - \mathbf{e}_n \mathbf{e}_n^{\mathrm{T}}$ is the projection on the orthogonal complement of span $\{\mathbf{e}_n\}$.

Theorem 1 (Natterer [22]). *G is uniquely defined by Eqs.* (2) *and it is lower triangular*.

Proof. Eqs. (2) can be solved by a Lanczos type iteration, as follows: Let $\mathbf{g}_i = \mathbf{e}_i^{\mathrm{T}} G$ be the rows of G, and rewrite (2) as

$$\begin{aligned}
&\mathbf{g}_1 = \mathbf{e}_1^{\mathrm{T}}, \\
&\alpha_1 \mathbf{g}_1 + \beta_1 \mathbf{g}_2 = \mathbf{g}_1 \widetilde{\mathscr{T}}, \\
&\beta_{i-1}\mathbf{g}_{i-1} + \alpha_i \mathbf{g}_i + \beta_i \mathbf{g}_{i+1} = \mathbf{g}_i \widetilde{\mathscr{T}}, \quad i = 2, \ldots, n-1.
\end{aligned} \tag{3}$$

By assumption, $\beta_j \neq 0$, so we can determine uniquely, from (3), all the rows of G. That G is lower triangular follows easily from (3) and the tridiagonal structure of $\widetilde{\mathscr{T}}$. □

Let now $\widetilde{S}$ and $\widetilde{\Lambda}$ be the matrices of orthonormal eigenvectors and eigenvalues of $\widetilde{\mathscr{T}}$, respectively, and define $S_0 \in \mathbb{R}^{n\times n}$, such that

$$E\mathscr{T} S_0 = ES_0\widetilde{\Lambda}, \quad \mathbf{e}_1^{\mathrm{T}}(S_0 - \widetilde{S}) = 0. \tag{4}$$

Theorem 2 (Natterer [22]). *S_0 is determined uniquely by Eqs. (4) and it satisfies*

$$S_0 = G\widetilde{S}. \tag{5}$$

Proof. The uniqueness of S_0 is obtained easily from (4), which we write row by row as

$$\begin{aligned}
&\mathbf{e}_1^{\mathrm{T}} S_0 = \mathbf{e}_1^{\mathrm{T}}\widetilde{S},\\
&\alpha_1\mathbf{e}_1^{\mathrm{T}} S_0 + \beta_1\mathbf{e}_2^{\mathrm{T}} S_0 = \tilde{\mathbf{e}}_1^{\mathrm{T}} S_0\widetilde{\Lambda},\\
&\beta_{i-1}\mathbf{e}_{i-1}^{\mathrm{T}} S_0 + \alpha_i\mathbf{e}_i^{\mathrm{T}} S_0 + \beta_i\mathbf{e}_{i+1}^{\mathrm{T}} S_0 = \tilde{\mathbf{e}}_i^{\mathrm{T}} S_0\widetilde{\Lambda}, \quad i = 2,\dots,n-1.
\end{aligned} \tag{6}$$

To prove (5), we check that $G\widetilde{S}$ satisfies Eqs. (4). Using (2), we have

$$E\mathscr{T}(G\widetilde{S}) = EG\widetilde{\mathscr{T}}\widetilde{S} = E(G\widetilde{S})\widetilde{\Lambda}, \quad \mathbf{e}_1^{\mathrm{T}}\widetilde{S} = \mathbf{e}_1^{\mathrm{T}}(G\widetilde{S})$$

and (5) follows from the uniqueness of solution of (4), established above. □

Algorithm 1 (*The inversion algorithm*). Assuming that we know a matrix $\mathscr{T}$ and spectral data $\widetilde{\Lambda}$, $\mathbf{e}_1^{\mathrm{T}}\widetilde{S}$, calculate the matrix S_0, using recursion (6). The orthogonality of $\widetilde{S}$ and (5) imply

$$GG^{\mathrm{T}} = S_0 S_0^{\mathrm{T}}, \tag{7}$$

so G can be computed from the Cholesky factorization of $S_0S_0^{\mathrm{T}}$. Finally, $\widetilde{\mathscr{T}}$ is calculated in terms of $\mathscr{T}$ and G, using recursion (3).

The key point of the inversion algorithm is: To find $\widetilde{T}$, we seek first $\widetilde{S}$, which is related to the matrix of "generalized" eigenvectors S_0 of the given $\mathscr{T}$, through the transmutation matrix (kernel) G. The matrix S_0 solves the "initial value problem" (4), and it differs from S, the matrix of eigenvectors of $\mathscr{T}$, when $\widetilde{\Lambda} \neq \Lambda$. Finally, the orthogonality of $\widetilde{S}$ requires that kernel G satisfy Eq. (7).

2.2. Continuum interpretation

Consider Schrödinger's equation in a unit interval,

$$\frac{\mathrm{d}^2 y(z)}{\mathrm{d}z^2} - q(z)y(z) = \lambda y(z), \quad 0 < z < 1,$$

$$\frac{dy(0)}{dz} = 0, \qquad (8)$$
$$y(1) = 0,$$

with some real scattering potential function $q(z)$. Suppose that we discretize (8) with a finite difference scheme, on a staggered grid

$$0 = z_1 = \hat{z}_0 < \hat{z}_1 < z_1 < \hat{z}_2 < \cdots < \hat{z}_n < z_{n+1},$$

with spacing $h_j = z_{j+1} - z_j$ and $\hat{h}_j = \hat{z}_j - \hat{z}_{j-1}$, for $j = 1, \ldots, n$, between the primary and dual points, respectively. The discretized equations are

$$\frac{1}{\hat{h}_i}\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right) - q_i y_i = \lambda y_i, \quad i = 1, \ldots, n,$$
$$\frac{y_1 - y_0}{h_0} = 0, \qquad (9)$$
$$y_{n+1} = 0,$$

where y_0 is assigned to some dummy node z_0 that we can take at arbitrarily small distance h_0 to the left of z_1. Using a diagonal scaling matrix

$$\widehat{H}^{\frac{1}{2}} = \mathrm{diag}\left(\hat{h}_1^{\frac{1}{2}}, \ldots, \hat{h}_n^{\frac{1}{2}}\right),$$

we rewrite the system of equations (9) with excluded y_0, in compact form,

$$\mathscr{T}\mathbf{s} - \lambda\mathbf{s} = 0,$$

where

$$\mathbf{s} = \widehat{H}^{\frac{1}{2}}\mathbf{y}, \quad \mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{T}},$$

and $\mathscr{T}$ is a Jacobi matrix, with entries

$$\beta_i = \frac{1}{h_i\sqrt{\hat{h}_i\hat{h}_{i+1}}}, \quad i = 1, \ldots, n-1,$$
$$\alpha_1 = -\frac{1}{h_1\hat{h}_1} - q_1, \qquad (10)$$
$$\alpha_i = -\frac{1}{h_{i-1}\hat{h}_i} - \frac{1}{h_i\hat{h}_i} - q_i, \quad i = 2, \ldots, n.$$

Now, let us similarly consider another potential $\tilde{q}(z)$ and denote by $\widetilde{\mathscr{T}}$ the corresponding finite difference Schrödinger operator (Jacobi matrix). We associate to the column and row indices of matrix G discrete spatial and temporal coordinates z_j and t_i, for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant n$. Then, Eq. (2) can be interpreted as the finite-difference approximation of wave problem

$$\frac{\partial^2 g(t,z)}{\partial z^2} - \frac{\partial^2 g(t,z)}{\partial t^2} = [\widetilde{q}(z) - q(t)]g(t,z),$$

$$\left.\frac{\partial g}{\partial z}\right|_{z=0} = 0, \quad g|_{z=1} = 0, \quad \left.\frac{\partial g}{\partial t}\right|_{t=0} = 0, \quad g|_{t=0} = \delta(z)$$

on $[0, 1]^2$, and $G_{ij}/\hat{h}_1$ becomes the finite difference approximation of the wave Green's function $g(t, z)$. At $t_1 = 0$, we have a unit impulse at location $z_1 = 0$, as given by initial condition $\mathbf{e}_1^{\mathrm{T}} G = \mathbf{e}_1^{\mathrm{T}}$. This impulse propagates at later times, as described by finite difference equations (3). When $\widetilde{\mathscr{T}} \neq \mathscr{T}$, by causality, the scattered impulse advances one grid cell at each time step, so G is lower triangular. In this case, in the continuum setting, we have a nonzero perturbation of Schrödinger potential $\widetilde{q} - q$, so we obtain nonzero g for $1 \geqslant t > z$. However, if $\widetilde{\mathscr{T}} = \mathscr{T}$, there is no scattering and the impulse travels undisturbed, which makes G an identity. This case corresponds to $\widetilde{q} - q = 0$ and the continuous wave solution $g = \delta(z - t)$.

Finally, kernel G determines the eigenvectors of $\widetilde{T}$ (see (5)) and, since they are orthonormal, we obtain (7), the discrete counterpart of the Gel'fand–Levitan integral equation for $g(t, z)$ [17–20,23].

Remark 1. Here our objective is to show the connection of the discrete and continuous Gel'fand–Levitan settings, and we do not intend to use Algorithm 1 for numerical calculations. In the above continuum interpretation we did not specify the choice of the grid, which is essential for solving numerically the inverse spectral problem for (8). Take for example a grid size n, and suppose that the grid points are arbitrarily distributed in [0,1]. The finite-difference inverse problem usually gives answers that are far from the true $q(z)$ and the results do not improve as we increase n. This is because arbitrary grids lead to discrete Sturm–Liouville operators (matrices) with eigenvalues that have different asymptotes than those in the continuum [24]. However, we experimentally showed in [25] that there exists a class of so called "optimal grids" with "correct" spectral asymptotes, which lead to the true solution $q(z)$ for large n. The perturbation analysis developed in this work has been used by us for the convergence proof of the discrete inversion approach on the optimal grids in [26]. We discuss one of such grids in more details in Section 5.

3. Sensitivity analysis

To calculate the sensitivity of $\widetilde{\mathscr{T}}$ to perturbations of the spectrum, we linearize the discrete Gel'fand–Levitan equations of Section 2 around $\mathscr{T}$, Λ and S. Substituting (1) in (2), we have

$$E\,\mathrm{d}G\,\mathscr{T} - E\mathscr{T}\,\mathrm{d}G = -E\,\mathrm{d}\mathscr{T}, \quad \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}G = \mathbf{0}^{\mathrm{T}}, \tag{11}$$

where $G = I + \mathrm{d}G + \mathrm{o}(\|\mathrm{d}G\|)$. These equations are similar to (2) and they can be solved with a three-term recursion for rows $\mathrm{d}\mathbf{g}_i = \mathbf{e}_i^{\mathrm{T}}\,\mathrm{d}G$,

$$
\begin{aligned}
&\mathrm{d}\mathbf{g}_1 = \mathbf{0},\\
&\mathrm{d}\mathbf{g}_2 = \frac{\mathrm{d}\beta_1}{\beta_1}\mathbf{e}_2^{\mathrm{T}} + \frac{\mathrm{d}\alpha_1}{\beta_1}\mathbf{e}_1^{\mathrm{T}}, \qquad (12)\\
&\beta_{i-1}\,\mathrm{d}\mathbf{g}_{i-1} + \alpha_i\,\mathrm{d}\mathbf{g}_i + \beta_i\,\mathrm{d}\mathbf{g}_{i+1} - \mathrm{d}\mathbf{g}_i\,\mathscr{T} = \mathrm{d}\beta_{i-1}\,\mathbf{e}_{i-1}^{\mathrm{T}} + \mathrm{d}\alpha_i\,\mathbf{e}_i^{\mathrm{T}} + \mathrm{d}\beta_i\,\mathbf{e}_{i+1}^{\mathrm{T}},\\
&\quad i = 2, \ldots, n-1.
\end{aligned}
$$

For example, let us calculate the diagonal of $\mathrm{d}G$. Taking the $(i+1)$st component of the last equation in (12) and using the lower triangular structure of $\mathrm{d}G$, we have

$$\mathrm{d}G_{i+1,i+1} - \mathrm{d}G_{i,i} = \frac{\mathrm{d}\beta_i}{\beta_i}, \quad 1 \leqslant i < n, \qquad \mathrm{d}G_{1,1} = 0,$$

or, equivalently,

$$\mathrm{d}G_{i+1,i+1} = \sum_{j=1}^{i} \frac{\mathrm{d}\beta_j}{\beta_j}, \quad 1 \leqslant i < n. \qquad (13)$$

Similarly, we get the subdiagonal of $\mathrm{d}G$, from the ith component of the last equation in (12),

$$\mathrm{d}G_{i+1,i} = \frac{1}{\beta_i}\sum_{j=1}^{i} \mathrm{d}\alpha_j. \qquad (14)$$

Now, we wish to determine $\mathrm{d}G$, in terms of the perturbations of the spectral data. To achieve this, linearize (7),

$$\mathrm{d}G + \mathrm{d}G^{\mathrm{T}} = \mathrm{d}S_0\, S^{\mathrm{T}} + S\,\mathrm{d}S_0^{\mathrm{T}},$$

where $S_0 = S + \mathrm{d}S_0 + \mathrm{o}(\|\mathrm{d}S_0\|)$, and define the matrix (matrix differential form) $\mathrm{d}V = \mathrm{d}S_0\, S^{\mathrm{T}}$, so that (13) and (14) become

$$\sum_{j=1}^{i} \frac{\mathrm{d}\beta_j}{\beta_j} = \mathrm{d}V_{i+1,i+1} \quad \text{and} \quad \frac{1}{\beta_i}\sum_{j=1}^{i} \mathrm{d}\alpha_j = \mathrm{d}V_{i+1,i} + \mathrm{d}V_{i,i+1}, \quad 1 \leqslant i < n. \qquad (15)$$

The sensitivity analysis reduces to calculating $\mathrm{d}V$, which solves a matrix equation derived by linearizing (4),

$$E\mathscr{T}\,\mathrm{d}S_0 - E\,\mathrm{d}S_0\,\Lambda = ES\,\mathrm{d}\Lambda, \quad \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}S_0 = \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}S$$

and multiplying by S^{T}, from the right,

$$E\mathscr{T}\,\mathrm{d}V - E\,\mathrm{d}V\,\mathscr{T} = ES\,\mathrm{d}\Lambda\, S^{\mathrm{T}}, \quad \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}V = \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}S\, S^{\mathrm{T}}. \qquad (16)$$

Next, we take advantage of the linearity of Eqs. (16) and we write

$$\mathrm{d}V = \mathrm{d}V^{\lambda} + \mathrm{d}V^{s},$$

where $\mathrm{d}V^\lambda$ and $\mathrm{d}V^s$ are the contributions of $\mathrm{d}\Lambda$ and $\mathbf{e}_1^\mathrm{T}\,\mathrm{d}S$, to the solution of equation (16). We calculate $\mathrm{d}V^\lambda$ and $\mathrm{d}V^s$, separately, in Sections 3.1 and 3.2.

3.1. Sensitivity to perturbations of the eigenvalues

To calculate the sensitivity to perturbations of the eigenvalues, we solve

$$E\mathscr{T}\,\mathrm{d}V^\lambda - E\,\mathrm{d}V^\lambda\,\mathscr{T} = ES\,\mathrm{d}\Lambda\,S^\mathrm{T}, \quad \mathbf{e}_1^\mathrm{T}\,\mathrm{d}V^\lambda = \mathbf{0}^\mathrm{T}.$$

This equation differs from (11) just through the right-hand side, so it has a unique solution that can be found with a three-term recursion. Since we are interested in just a few components of $\mathrm{d}V^\lambda$, appearing in Eqs. (15), we decompose the solution as

$$\mathrm{d}V^\lambda = \mathrm{d}W^{(i)} + \mathrm{d}P^{(i)} + \mathrm{d}Q^{(i)}, \tag{17}$$

where

$$\mathscr{T}\,\mathrm{d}W^{(i)} - \mathrm{d}W^{(i)}\,\mathscr{T} = S\,\mathrm{d}\Lambda\,S^\mathrm{T} + \mathbf{e}_i\,\mathrm{d}\mathbf{r}^\mathrm{T}, \quad \mathbf{s}_j^\mathrm{T}\,\mathrm{d}W^{(i)}\,\mathbf{s}_j = 0, \quad j = 1, \ldots, n, \tag{18}$$

for some vector $\mathrm{d}\mathbf{r}$ to be found,

$$E\mathscr{T}\,\mathrm{d}P^{(i)} - E\,\mathrm{d}P^{(i)}\,\mathscr{T} = 0, \quad \mathbf{e}_1^\mathrm{T}\,\mathrm{d}P^{(i)} = -\mathbf{e}_1^\mathrm{T}\,\mathrm{d}W^{(i)} \tag{19}$$

and

$$E\mathscr{T}\,\mathrm{d}Q^{(i)} - E\,\mathrm{d}Q^{(i)}\mathscr{T} = -E\mathbf{e}_i\,\mathrm{d}\mathbf{r}^\mathrm{T}, \quad \mathbf{e}_1^\mathrm{T}\,\mathrm{d}Q^{(i)} = 0.$$

Note that, by construction, the first i rows of $\mathrm{d}Q^{(i)}$ are identically zero and

$$\mathbf{e}_{i+1}^\mathrm{T}\,\mathrm{d}Q^{(i)} = -\frac{\mathrm{d}\mathbf{r}^\mathrm{T}}{\beta_i}. \tag{20}$$

Thus, it suffices to get $\mathrm{d}W^{(i)}$ and $\mathrm{d}P^{(i)}$, for $i = 1, \ldots, n$, and substitute them in (15).

Let us solve (18). We write the unknown $\mathrm{d}\mathbf{r}$ in the basis of eigenvectors of $\mathscr{T}$,

$$\mathrm{d}\mathbf{r} = \sum_{j=1}^{n} \mathrm{d}C_j\,\mathbf{s}_j,$$

and we multiply (19) from the left and right by S^T and S, respectively,

$$\Lambda S^\mathrm{T}\,\mathrm{d}W^{(i)}\,S - S^\mathrm{T}\,\mathrm{d}W^{(i)}\,S\Lambda = \mathrm{d}\Lambda + \sum_{j=1}^{n} \mathrm{d}C_j\,S^\mathrm{T}\mathbf{e}_i\mathbf{e}_j^\mathrm{T}. \tag{21}$$

Taking the diagonal part of (21) and using (18), we have for the jth entry,

$$\mathrm{d}\lambda_j + \mathrm{d}C_j\,S_{ij} = 0,$$

so the desired vector $\mathrm{d}\mathbf{r}$ is

$$\mathrm{d}\mathbf{r} = -\sum_{j=1}^{n} \frac{\mathrm{d}\lambda_j}{S_{ij}}\mathbf{s}_j. \tag{22}$$

Then, substituting (22) in (21) gives

$$\mathbf{s}_p^{\mathrm{T}}\,\mathrm{d}W^{(i)}\,\mathbf{s}_q = \frac{\mathrm{d}\lambda_q\, S_{ip}}{(\lambda_q-\lambda_p)S_{iq}} \quad \text{for } p \neq q, \quad p,q=1,\ldots,n,$$

or, equivalently,

$$\mathrm{d}W^{(i)} = \sum_{q=1}^{n}\sum_{\substack{p=1\\p\neq q}}^{n} \frac{\mathrm{d}\lambda_q\, S_{ip}}{(\lambda_q-\lambda_p)S_{iq}}\mathbf{s}_p\mathbf{s}_q^{\mathrm{T}}. \tag{23}$$

Next, we seek the solution of (19) in the form

$$\mathrm{d}P^{(i)} = \sum_{q=1}^{n}\mathrm{d}D_q\,\mathbf{s}_q\mathbf{s}_q^{\mathrm{T}}.$$

Such a matrix satisfies the first equation in (19), and scalar differentials $\mathrm{d}D_q$ are determined by the “initial” condition on $\mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}P^{(i)}$. The result is

$$\mathrm{d}P^{(i)} = -\sum_{q=1}^{n}\sum_{\substack{p=1\\p\neq q}}^{n} \frac{\mathrm{d}\lambda_q\, S_{ip}S_{1p}}{(\lambda_q-\lambda_p)S_{iq}S_{1q}}\mathbf{s}_q\mathbf{s}_q^{\mathrm{T}}, \tag{24}$$

the unique solution of Eq. (19).

Theorem 3. *Assuming that* $\mathbf{e}_1^{\mathrm{T}}\mathrm{d}S=\mathbf{0}$, *we have, for* $i=2,\ldots,n$,

$$\sum_{j=1}^{i-1}\frac{\mathrm{d}\beta_j}{\beta_j} = \sum_{q=1}^{n}\mathrm{d}\lambda_q\sum_{\substack{p=1\\p\neq q}}^{n}\frac{1}{\lambda_q-\lambda_p}\left(S_{ip}^2-\frac{S_{1p}}{S_{1q}}S_{ip}S_{iq}\right), \tag{25}$$

whereas, for $i=1,\ldots,n-1$,

$$\sum_{j=1}^{i}\mathrm{d}\alpha_j = \sum_{q=1}^{n}d\lambda_q + \beta_i\sum_{q=1}^{n}d\lambda_q\sum_{\substack{p=1\\p\neq q}}^{n}\frac{1}{\lambda_q-\lambda_p} \times\left[2S_{ip}S_{i+1p}-\frac{S_{1p}}{S_{1q}}(S_{ip}S_{i+1q}+S_{i+1p}S_{iq})\right]. \tag{26}$$

Finally,

$$\mathrm{d}\alpha_n = -\beta_{n-1}\sum_{q=1}^{n}\mathrm{d}\lambda_q\sum_{\substack{p=1\\p\neq q}}^{n}\frac{1}{\lambda_q-\lambda_p} \times\left[2S_{n-1p}S_{np}-\frac{S_{1p}}{S_{1q}}\left(S_{n-1p}S_{nq}+S_{np}S_{n-1q}\right)\right]. \tag{27}$$

Proof. The proof is simply the substitution of (23), (24) and (20) in (17) and, subsequently, in (15). Equality (27) follows from (26) and identity $\sum_{j=1}^{n} \mathrm{d}\alpha_j = \sum_{j=1}^{n} \mathrm{d}\lambda_j$. □

3.2. Sensitivity to perturbations of the weights

Here, we suppose that the eigenvalues are unperturbed, so (16) becomes

$$E\mathscr{T}\,\mathrm{d}V^s - E\,\mathrm{d}V^s\,\mathscr{T} = 0, \quad \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}V^s = \mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}S\,S^{\mathrm{T}}$$

and

$$\mathrm{d}V^s = \sum_{q=1}^{n} \frac{\mathrm{d}S_{1q}}{S_{1q}} \mathbf{s}_q \mathbf{s}_q^{\mathrm{T}} \tag{28}$$

is the unique solution. The sensitivity result follows from (28) and (15):

Theorem 4. *If* $\mathrm{d}\Lambda = 0$, *we have*

$$\sum_{j=1}^{i-1} \frac{\mathrm{d}\beta_j}{\beta_j} = \sum_{q=1}^{n} \frac{\mathrm{d}S_{1q}}{S_{1q}} S_{iq}^2, \quad i = 2, \ldots, n, \tag{29}$$

and

$$\frac{1}{\beta_i} \sum_{j=1}^{i} \mathrm{d}\alpha_j = 2 \sum_{q=1}^{n} \frac{\mathrm{d}S_{1q}}{S_{1q}} S_{i+1q} S_{iq}, \quad i = 1, \ldots, n-1. \tag{30}$$

Moreover,

$$\mathrm{d}\alpha_n = -\sum_{j=1}^{n-1} \mathrm{d}\alpha_j = -2\beta_{n-1} \sum_{q=1}^{n} \frac{\mathrm{d}S_{1q}}{S_{1q}} S_{nq} S_{n-1q}. \tag{31}$$

4. Stability estimates for Lanczos recursions

4.1. Estimates with respect to the eigenvalues

Recalling that Jacobi matrices have distinct eigenvalues, let us take a decreasing ordering of λ_j, for $j = 1, \ldots, n$, and let

$$\Delta_j = \min(\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}) > 0, \quad j = 2, \ldots, n-1,$$
$$\Delta_1 = \lambda_1 - \lambda_2, \quad \Delta_n = \lambda_{n-1} - \lambda_n$$

be the eigenvalue separation. The perturbation of λ_j, relative to the separation, is denoted by

$$\mathrm{d}\theta_j = \frac{\mathrm{d}\lambda_j}{\Delta_j}$$

and (25), (26) become

$$\sum_{j=1}^{i-1} \frac{\mathrm{d}\beta_j}{\beta_j} = \sum_{q=1}^{n} \mathrm{d}\theta_q\, \rho_q^{(i)}, \quad i = 2, \ldots, n, \tag{32}$$

$$\frac{1}{\beta_i} \sum_{j=1}^{i} \mathrm{d}\alpha_j = \sum_{q=1}^{n} \mathrm{d}\theta_q\, \zeta_q^{(i)}, \quad i = 1, \ldots, n-1, \tag{33}$$

where

$$\rho_q^{(i)} = \sum_{\substack{p=1\\p\neq q}}^{n} \frac{\Delta_q}{\lambda_q - \lambda_p} \left(S_{ip}^2 - \frac{S_{1p}}{S_{1q}} S_{ip} S_{iq} \right) \tag{34}$$

and

$$\zeta_q^{(i)} = \frac{\Delta_q}{\beta_i} + \sum_{\substack{p=1\\p\neq q}}^{n} \frac{\Delta_q}{\lambda_q - \lambda_p} \left[2 S_{ip} S_{i+1p} - \frac{S_{1p}}{S_{1q}} (S_{ip} S_{i+1q} + S_{i+1p} S_{iq}) \right]. \tag{35}$$

We have the following, weak (discrete primitive) stability estimates:

Theorem 5. *Assume that* $\mathbf{e}_1^{\mathrm{T}}\, \mathrm{d}S = 0$ *and let* $\xi = \max_{1 \leqslant p,q \leqslant n} \left| \frac{S_{1p}}{S_{1q}} \right|$. *We have, for all* $i = 2, \ldots, n$,

$$\frac{1}{n} \left| \sum_{j=1}^{i-1} \frac{\mathrm{d}\beta_j}{\beta_j} \right| \leqslant (1 + \xi) \max_{1 \leqslant q \leqslant n} |\mathrm{d}\theta_q| \tag{36}$$

and

$$\frac{1}{\beta_{i-1} n} \left| \sum_{j=1}^{i-1} \mathrm{d}\alpha_j \right| \leqslant 2(1 + \xi) \max_{1 \leqslant q \leqslant n} |\mathrm{d}\theta_q| + \frac{1}{\beta_{i-1} n} \sum_{q=1}^{n} \Delta_q |\mathrm{d}\theta_q|. \tag{37}$$

Proof. From (34), we obtain

$$|\rho_q^{(i)}| \leqslant \sum_{\substack{p=1\\p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} S_{ip}^2 + \xi \sum_{\substack{p=1\\p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} |S_{ip} S_{iq}|,$$

where, by definition, $\Delta_q / |\lambda_q - \lambda_p| \leqslant 1$. Cauchy–Schwarz's inequality gives

$$|\rho_q^{(i)}| \leqslant \sum_{\substack{p=1\\p\neq q}}^{n} \left(S_{ip}^2 + \xi |S_{iq}||S_{ip}| \right) \leqslant \sum_{p=1}^{n} S_{ip}^2 + \xi |S_{iq}| \sqrt{n} \left(\sum_{p=1}^{n} S_{ip}^2 \right)^{\frac{1}{2}}$$
$$= 1 + \xi \sqrt{n} |S_{iq}|$$

and, from (32), we have

$$\frac{1}{n}\left|\sum_{j=1}^{i-1}\frac{\mathrm{d}\beta_j}{\beta_j}\right| \leqslant \frac{1}{n}\sum_{q=1}^{n}|\mathrm{d}\theta_q|[1+\xi\sqrt{n}|S_{iq}|] \leqslant (1+\xi)\max_{1\leqslant q\leqslant n}\left|\mathrm{d}\theta_q\right|,$$

where Cauchy–Schwarz's inequality was used, once more. Similarly, (35) satisfies

$$\begin{aligned}\left|\zeta_q^{(i-1)}\right| &\leqslant \frac{\Delta_q}{\beta_{i-1}} + \sum_{\substack{p=1\\p\neq q}}^{n}\left[2|S_{i-1\,p}S_{ip}| + \xi\left(|S_{i-1\,p}S_{iq}| + |S_{ip}S_{i-1q}|\right)\right]\\ &\leqslant \frac{\Delta_q}{\beta_{i-1}} + 2 + \xi\sqrt{n}(|S_{i-1q}| + |S_{iq}|)\end{aligned}$$

and (37) follows as above, from (33) and Cauchy–Schwarz's inequality. □

We also have the strong, l^1 stability estimates:

Theorem 6. *Assume that* $\mathbf{e}_1^{\mathrm{T}}\,\mathrm{d}S = 0$ *and let* ξ *be defined as above, in Theorem 5. Then*,

$$\sum_{i=1}^{n-1}\frac{|\mathrm{d}\beta_i|}{\beta_i} \leqslant 2(1+\xi)\sum_{q=1}^{n}|\mathrm{d}\theta_q|\sum_{\substack{p=1\\p\neq q}}^{n}\frac{\Delta_q}{|\lambda_q-\lambda_p|} \tag{38}$$

and

$$\begin{aligned}&\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1}\frac{|\mathrm{d}\alpha_i|}{\max(\beta_i,\beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}}\\ &\quad\leqslant \sum_{q=1}^{n}|\mathrm{d}\theta_q|\frac{\Delta_q}{\beta_1} + 4(1+\xi)\sum_{q=1}^{n}|\mathrm{d}\theta_q|\sum_{\substack{p=1\\p\neq q}}^{n}\frac{\Delta_q}{|\lambda_q-\lambda_p|}.\end{aligned} \tag{39}$$

Proof. From (32), we have

$$\frac{|\mathrm{d}\beta_i|}{\beta_i} \leqslant \sum_{q=1}^{n}|\mathrm{d}\theta_q|\sum_{\substack{p=1\\p\neq q}}^{n}\frac{\Delta_q}{|\lambda_q-\lambda_p|}\left(S_{i+1p}^2 + S_{ip}^2 + \xi|S_{i+1p}S_{i+1q}| + \xi|S_{ip}S_{iq}|\right).$$

Bound (38) follows by summation over i and by using Cauchy–Schwarz's inequality,

$$\sum_{i=1}^{n}\left|S_{ip}S_{iq}\right| \leqslant \left(\sum_{i=1}^{n}S_{ip}^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}S_{iq}^2\right)^{\frac{1}{2}} = 1.$$

Next, we obtain from (33), for all $i = 2, \ldots, n-1$,

$$\frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} \leqslant \sum_{q=1}^{n} |\mathrm{d}\theta_q| \sum_{\substack{p=1 \\ p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} \big[2|S_{ip}S_{i+1p}| + 2|S_{ip}S_{i-1p}| + \xi\big(|S_{ip}S_{i+1q}| + |S_{i+1p}S_{iq}| + |S_{ip}S_{i-1q}| + |S_{i-1p}S_{iq}|\big)\big],$$

whereas, for $i = 1$,

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} \leqslant \sum_{q=1}^{n} |\mathrm{d}\theta_q| \frac{\Delta_q}{\beta_1} + \sum_{q=1}^{n} |\mathrm{d}\theta_q| \sum_{\substack{p=1 \\ p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} \times\big[2|S_{1p}S_{2p}| + \xi\big(|S_{1p}S_{2q}| + |S_{2p}S_{1q}|\big)\big],$$

and, for $i = n$, in view of (27)

$$\frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \leqslant \sum_{q=1}^{n} |\mathrm{d}\theta_q| \sum_{\substack{p=1 \\ p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} \big[2|S_{n-1p}S_{np}| + \xi\big(|S_{n-1p}S_{nq}| + |S_{np}S_{n-1q}|\big)\big].$$

Summing over i, we get

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1} \frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}}$$
$$\leqslant \sum_{q=1}^{n} |\mathrm{d}\theta_q| \frac{\Delta_q}{\beta_1} + \sum_{q=1}^{n} |\mathrm{d}\theta_q| \sum_{\substack{p=1 \\ p\neq q}}^{n} \frac{\Delta_q}{|\lambda_q - \lambda_p|} \times \left[4\sum_{i=1}^{n-1} |S_{ip}S_{i+1p}| + 2\xi \sum_{i=1}^{n-1} \big(|S_{ip}S_{i+1q}| + |S_{i+1p}S_{iq}|\big)\right],$$

and (39) follows from Cauchy–Schwarz's inequality. □

4.2. *Estimates with respect to the weights*

If only the weights S_{1i}^2, $i = 1, \ldots, n$, are perturbed, we have the following stability bounds:

Theorem 7. *Assume that* $\mathrm{d}\Lambda = 0$. *Then, for all* $i = 2, \ldots, n$, *we have the weak (discrete primitive) estimates*

$$\frac{1}{n} \left| \sum_{j=1}^{i-1} \frac{\mathrm{d}\beta_j}{\beta_j} \right| \leqslant \frac{1}{n} \max_{1\leqslant q\leqslant n} \left| \frac{\mathrm{d}S_{1q}}{S_{1q}} \right| \tag{40}$$

and

$$\frac{1}{n\beta_{i-1}}\left|\sum_{j=1}^{i-1} \mathrm{d}\alpha_j\right| \leqslant \frac{2}{n} \max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|. \tag{41}$$

Proof. Eq. (29) gives

$$\frac{1}{n}\left|\sum_{j=1}^{i-1}\frac{\mathrm{d}\beta_j}{\beta_j}\right| \leqslant \frac{1}{n}\sum_{q=1}^{n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right| S_{iq}^2 \leqslant \frac{1}{n}\max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|\sum_{q=1}^{n} S_{iq}^2$$
$$= \frac{1}{n}\max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|,$$

so (40) is proved. Similarly, from (30), we have

$$\frac{1}{n\beta_{i-1}}\left|\sum_{j=1}^{i-1}\mathrm{d}\alpha_j\right| \leqslant \frac{2}{n}\sum_{q=1}^{n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right| |S_{iq}S_{i+1q}| \leqslant \frac{2}{n}\max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|\sum_{q=1}^{n}|S_{iq}S_{i+1q}|$$
$$\leqslant \frac{2}{n}\max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|\left(\sum_{q=1}^{n}S_{iq}^2\right)^{\frac{1}{2}}\left(\sum_{q=1}^{n}S_{i+1q}^2\right)^{\frac{1}{2}}$$
$$= \frac{2}{n}\max_{1\leqslant q\leqslant n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|. \qquad \square$$

We also have the strong, l^1 stability bounds:

Theorem 8. *If* $\mathrm{d}\Lambda = 0$,

$$\sum_{i=1}^{n-1}\frac{|\mathrm{d}\beta_i|}{\beta_i} \leqslant 2\sum_{q=1}^{n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|$$

and

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1}\frac{|\mathrm{d}\alpha_i|}{\max(\beta_i,\beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \leqslant 4\sum_{q=1}^{n}\left|\frac{\mathrm{d}S_{1q}}{S_{1q}}\right|.$$

Proof. From (29), we have

$$\sum_{i=1}^{n-1}\frac{|\mathrm{d}\beta_i|}{\beta_i} \leqslant \sum_{q=1}^{n}\frac{|\mathrm{d}S_{1q}|}{S_{1q}}\sum_{i=1}^{n-1}\left(S_{i+1q}^2 + S_{iq}^2\right) \leqslant 2\sum_{q=1}^{n}\frac{|\mathrm{d}S_{1q}|}{S_{1q}}.$$

Further, (30) gives, for $i = 2, \ldots, n-1$,

$$\frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} \leqslant 2 \sum_{q=1}^{n} \frac{|\mathrm{d}S_{1q}|}{S_{1q}} \left(|S_{iq} S_{i+1q}| + |S_{i-1q} S_{iq}|\right)$$

whereas, for $i = 1$,

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} \leqslant 2 \sum_{q=1}^{n} \frac{|\mathrm{d}S_{1q}|}{S_{1q}} |S_{1q} S_{2q}|$$

and, for $i = n$, by virtue of (31)

$$\frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \leqslant 2 \sum_{q=1}^{n} \frac{|\mathrm{d}S_{1q}|}{S_{1q}} |S_{n-1q} S_{nq}|.$$

Now, summing over i, we have

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1} \frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \leqslant 4 \sum_{q=1}^{n} \frac{|\mathrm{d}S_{1q}|}{S_{1q}} \sum_{i=1}^{n-1} |S_{iq} S_{i+1q}|$$

$$\leqslant 4 \sum_{q=1}^{n} \frac{|\mathrm{d}S_{1q}|}{S_{1q}}. \qquad \square$$

5. Discussion of the stability estimates. Connection with discrete Sturm–Liouville problems

In Section 4, we derived stability estimates of two kinds. The first kind, given by Theorems 5 and 7, considers discrete primitive norms of $\mathrm{d}\beta_j/\beta_j$ and $\mathrm{d}\alpha_j$, respectively. We call them discrete primitive, because, if we associated with each j a point z_j, in a uniform mesh, of spacing $1/n$, the left-hand sides in (36), (37), (40) and (41) would correspond to discretizations of integrals, from z_1 to z_{i-1}, of expressions $\mathrm{d}\beta(z)/\beta(z)$ and $\mathrm{d}\alpha(z)$, some interpolations of nodal values $\mathrm{d}\beta_j/\beta_j$ and $\mathrm{d}\alpha_j$, on the grid. This discrete primitive norm is obviously weaker than the l^1 norm that we consider in Theorems 6 and 8 but the tradeoff is that, in general, the upper bounds on the l^1 norm can be very large and growing to infinity as $n \to \infty$ because of the factor

$$\sum_{\substack{p=1 \\ p \neq r}}^{n} \frac{\Delta_r}{|\lambda_r - \lambda_p|}$$

appearing in the right-hand sides of (38) and (39). There are however important cases of Jacobi matrices, for which this factor grows slow enough (uniformly with n) and perturbations $\mathrm{d}\theta_r$ decay at a fast enough rate as r increases, in order to achieve l^1

estimates that are bounded independently of the dimension n of the problem. Such matrices arise, for example, from the discretization of Sturm–Liouville equations, with sufficiently smooth coefficients. We illustrate next this fact, through an example motivated by our recent study of continuum limits of solutions of discrete inverse Sturm–Liouville problems, on so called optimal finite difference grids [26].

Now, let us revisit the finite-difference interpretation given in Section 2.2. The Schrödinger operator in (8) has infinitely many distinct, negative eigenvalues λ_j and orthonormal eigenfunctions $y_j(z)$, for $j \geqslant 1$. The eigenvalues of $\mathscr{T}$ depend on the grid that we choose. Among all possible grids, we distinguish the "optimal" one, which ensures that the eigenvalues of $\mathscr{T}$ are the n largest λ_j, and that the first components of the eigenvectors S_{1j} are given by $\hat{h}_1^{\frac{1}{2}} y_j(0)$, for $j = 1, \ldots, n$. The existence and uniqueness of solution of the inverse spectral problem for Jacobi matrix $\mathscr{T}$ [2–7] guarantees that such an optimal grid exists and it is unique. For simplicity of the explanation, let us take the case $q(z) = \overline{q}$ = constant, where

$$\lambda_j = -\left(j - \frac{1}{2}\right)^2 \pi^2 - \overline{q}, \quad S_{1j} = \sqrt{2\hat{h}_1} = \frac{1}{\sqrt{n}}, \quad j = 1, \ldots, n,$$

and where we can write $\mathscr{T}$ and therefore the optimal grid, explicitly (see [26])

$$\begin{aligned} h_j &= \frac{2 + \mathrm{O}((n-j)^{-1} + j^{-2})}{\pi\sqrt{n^2 - j^2}}, \quad 1 \leqslant j \leqslant n-1, \quad h_n = \frac{\sqrt{2} + \mathrm{O}(n^{-1})}{\sqrt{\pi n}}, \\ \hat{h}_j &= \frac{2 + \mathrm{O}((n+1-j)^{-1} + j^{-2})}{\pi\sqrt{n^2 - (j-1/2)^2}}, \quad 1 \leqslant j \leqslant n. \end{aligned} \tag{42}$$

Next, consider the discretization of another Schrödinger equation, with potential

$$\tilde{q}(z) = \bar{q} + \varepsilon \Delta q(z), \quad \varepsilon \to +0,$$

where $\Delta q(z)$ is a mean zero, sufficiently smooth function, say in $\mathscr{H}^1(0, 1)$. The discretized equation is

$$\widetilde{\mathscr{T}}\widetilde{\mathbf{s}} - \widetilde{\lambda}\widetilde{\mathbf{s}} = 0,$$

and $\widetilde{\mathscr{T}}$ is a Jacobi matrix, similar to $\mathscr{T}$, with entries $\widetilde{\beta}_i$ and $\widetilde{\alpha}_i$. As explained above, the discretization is done on the optimal grid obtained by solving the inverse spectral problem for $\widetilde{\mathscr{T}}$, with the n largest eigenvalues $\widetilde{\lambda}_j$ of the Schrödinger operator and the weights $\tilde{S}_{1j} = \widetilde{\hat{h}_1^{\frac{1}{2}}} \widetilde{y}_j(0)$, for $j = 1, \ldots, n$. Because the perturbation of the potential is $\mathrm{O}(\varepsilon)$, the spectral data are of the form

$$\widetilde{\lambda}_j = \lambda_j + \mathrm{d}\lambda_j + \mathrm{o}(\varepsilon), \quad \tilde{S}_{1j} = S_{1j} + \mathrm{d}S_{1j} + \mathrm{o}(\varepsilon), \quad j = 1, \ldots, n$$

and the stability bounds derived in Sections 4.1 and 4.2 give us estimates of the entries in $\widetilde{\mathscr{T}}$.

From now on, symbol $\lesssim$ has the same meaning as 'O'. Moreover, multiplicative constants, hidden under the $\lesssim$ and 'O' symbols, are independent of n.

Under our assumptions on $\Delta q(z)$, the eigenvalue and weight perturbations obey [23]

$$\mathrm{d}\lambda_j \lesssim \frac{\varepsilon}{j},$$

$$\frac{\mathrm{d}S_{1j}}{S_{1j}} = \frac{\mathrm{d}\hat{h}_1}{2\hat{h}_1} + \frac{\mathrm{d}y_j(0)}{y_j(0)} = \frac{\mathrm{d}\hat{h}_1}{2\hat{h}_1} + \mathrm{O}\left(\frac{1}{j^2}\right)\varepsilon \lesssim \left(\frac{1}{n} + \frac{1}{j^2}\right)\varepsilon,$$

$$j = 1, \ldots, n,$$

and, since the eigenvalue separation is

$$\Delta_j = 2\pi^2[\max(j, 2) - 1], \quad j = 1, \ldots, n,$$

we have that

$$\mathrm{d}\theta_j \lesssim \frac{\varepsilon}{j^2}, \quad j = 1, \ldots, n. \tag{43}$$

Moreover,

$$\xi = \max_{1 \leqslant p, r \leqslant n} \left|\frac{S_{1p}}{S_{1r}}\right| = 1.$$

Then, the l^1 bounds in Theorems 6 and 8 give

$$\begin{aligned}\sum_{i=1}^{n-1} \frac{|\mathrm{d}\beta_i|}{\beta_i} &\lesssim \sum_{r=1}^{n} |\mathrm{d}\theta_r| \sum_{\substack{p=1\\p\neq r}}^{n} \frac{\Delta_r}{|\lambda_r - \lambda_p|} + \sum_{r=1}^{n} \left|\frac{\mathrm{d}S_{1r}}{S_{1r}}\right| \\ &\lesssim |\mathrm{d}\theta_1| + \sum_{r=2}^{n} |\mathrm{d}\theta_r| \sum_{\substack{p=1\\p\neq r}}^{n} \frac{2(r-1)}{|(p+r-1)(p-r)|} + \sum_{r=1}^{n} \frac{|\mathrm{d}S_{1r}|}{S_{1r}}\end{aligned} \tag{44}$$

and

$$\begin{aligned}&\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1} \frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \\ &\quad\lesssim \sum_{r=1}^{n} |\mathrm{d}\theta_r| \frac{\Delta_r}{\beta_1} + \sum_{r=1}^{n} |\mathrm{d}\theta_r| \sum_{\substack{p=1\\p\neq r}}^{n} \frac{\Delta_r}{|\lambda_r - \lambda_p|} + \sum_{r=1}^{n} \frac{|\mathrm{d}S_{1r}|}{S_{1r}} \\ &\quad\lesssim \frac{1}{\beta_1} \sum_{r=1}^{n} |\mathrm{d}\theta_r|\, r + |\mathrm{d}\theta_1| \\ &\qquad + \sum_{r=2}^{n} |\mathrm{d}\theta_r| \sum_{\substack{p=1\\p\neq r}}^{n} \frac{2(r-1)}{|(p+r-1)(p-r)|} + \sum_{r=1}^{n} \frac{|\mathrm{d}S_{1r}|}{S_{1r}}.\end{aligned} \tag{45}$$

Now,

$$\sum_{r=1}^{n} \frac{|\mathrm{d}S_{1r}|}{S_{1r}} \lesssim \sum_{r=1}^{n} \left(\frac{1}{n} + \frac{1}{r^2}\right) \varepsilon \lesssim \varepsilon.$$

Furthermore,

$$\sum_{\substack{p=1\\ p\neq r}}^{n} \frac{2(r-1)}{|(p+r-1)(p-r)|} \leqslant \sum_{p=1}^{r-1} \left(\frac{1}{r-p} + \frac{1}{p+r-1}\right) + \sum_{p=r+1}^{n} \left(\frac{1}{p-r} - \frac{1}{p+r-1}\right) \lesssim \log r,$$

because [27]

$$\sum_{m=1}^{j} \frac{1}{m} = \log j + \mathrm{O}(1),$$

so, using (43), we obtain

$$|\mathrm{d}\theta_1| + \sum_{r=2}^{n} |\mathrm{d}\theta_r| \sum_{\substack{p=1\\ p\neq r}}^{n} \frac{2(r-1)}{|(p+r-1)(p-r)|} \lesssim \varepsilon \sum_{r=1}^{n} \frac{\log r}{r^2} \lesssim \varepsilon.$$

Thus, (44) becomes

$$\sum_{i=1}^{n-1} \frac{|\mathrm{d}\beta_i|}{\beta_i} \lesssim \varepsilon.$$

Finally, (10) and (42) give

$$\frac{1}{\beta_1} = h_1\sqrt{\hat{h}_1\hat{h}_2} = \mathrm{O}(n^{-2})$$

and, in view of (45) and the inequality

$$\sum_{r=1}^{n} |\mathrm{d}\theta_r|\, r \lesssim \varepsilon \sum_{r=1}^{n} \frac{1}{r} \lesssim \varepsilon \log n,$$

we obtain

$$\frac{|\mathrm{d}\alpha_1|}{\beta_1} + \sum_{i=2}^{n-1} \frac{|\mathrm{d}\alpha_i|}{\max(\beta_i, \beta_{i-1})} + \frac{|\mathrm{d}\alpha_n|}{\beta_{n-1}} \lesssim \varepsilon.$$

So we have proved here that the perturbations of $\log \beta_i$ are uniformly bounded. Comparing that with at least $\mathrm{O}(n)$ growth of the entries of the unperturbed matrix which follows from (42) and (10), we can see that the offdiagonal components of the Jacobi matrix do not depend on q in the asymptotic limit $n \to \infty$. A similar statement can be made for the diagonal elements. In other words, we have arrived at an interesting conclusion that the coefficients of Lanczos recursions, using truncated spectral measures of Schrödinger differential operators, are asymptotically independent of smooth perturbations of the potential. This result is related to the convergence result of [26] for the discrete inverse problems.

6. Concluding remarks

Remark 9. The stability estimates of Sections 4 and 5 apply for infinitesimal perturbations but they can be easily extended, by integration, to finite perturbations.

Remark 10. It is well known [28,29] that Lanzcos recursions are unstable to computer roundoff. According to [30], roundoff influence (in particular, the loss of orthogonality) on computer arithmetic Lanczos recursions is equivalent to the appearance, in the vicinity of the true spectrum, of clusters of spurious eigenvalues of an extended tridiagonal operator (with the computed recursion as the left upper block). The total weight of a cluster approximates the weight of the corresponding true eigenvalue, so the spectral measure of Greenbaum's operator approximates the measure of the basic matrix. A natural question arises, if this fact is in contradiction with our stability results. The answer is "no contradiction", because our differential expressions deal with an arbitrary, but fixed dimension n, while Greenbaum's operator has larger dimensions than the basic matrix.

Note also that elements of Greenbaum's cluster have separations of the same order as the cluster's width. Imagine then that we work with the extended tridiagonal matrix, and let its eigenvalues choose their location more or less chaotically inside the cluster. Generally, in this case spectral perturbations become of the same order as the separations within the cluster, so according to our perturbation formulas, we can obtain large perturbations of the recursion coefficients, including its initial part of length n.

Remark 11. Finally, we note that the sensitivity formulas (not the estimates) can be extended by means of analytical continuation to non-Hermitian Lanczos algorithms using pseudo inner products, for example, to the modification intended for complex symmetric matrices [31].

Appendix A. Relation to orthogonal polynomials

Natterer [22–Section 6] noticed that the discrete Gel'fand–Levitan theory can be formulated in the language of orthogonal polynomials. His remark can also be applied to derivation of formulae (32)–(35).

Indeed, consider measure (1) with nodes λ_j and weights $y_j = S_{1j}^2$. Introduce the corresponding orthonormal polynomials Q_i and second kind polynomials P_i.

Perturbed measure parameters $\widetilde{\lambda}_j$ and $\widetilde{y}_j$ induce the perturbed scalar product $\langle f, g\rangle_{\text{pert}} = \sum_{j=1}^{n} \widetilde{y}_j f(\widetilde{\lambda}_j)g(\widetilde{\lambda}_j)$. Put $e_{ij} = \langle Q_i, Q_j\rangle_{\text{pert}} - \delta_{ij}$. A standard small perturbation analysis in junction with the Lanczos recurrence

$$\lambda Q_k(\lambda) = \beta_{k+1} Q_{k+1}(\lambda) + \alpha_{k+1} Q_k(\lambda) + \beta_k Q_{k-1}(\lambda), \quad k \geqslant 0,\ \beta_0 \equiv 0,\ Q_{-1} \equiv 0,$$

leads to the differential expressions

$$\mathrm{d}\alpha_i = \beta_i\, \mathrm{d}e_{i,i-1} - \beta_{i-1}\, \mathrm{d}e_{i-1,i-2}, \quad \mathrm{d}\beta_i = \frac{\beta_i}{2}(\mathrm{d}e_{ii} - \mathrm{d}e_{i-1,i-1}),$$

whence

$$\frac{1}{\beta_i}\sum_{j=1}^{i} \mathrm{d}\alpha_j = \mathrm{d}e_{i,i-1}, \quad \sum_{j=1}^{i} \frac{\mathrm{d}\beta_i}{\beta_i} = \frac{\mathrm{d}e_{ii}}{2}$$

(cf. (15)).

We shall use the following formulae concerning orthonormal polynomials:

$$\frac{Q_m(\lambda) - Q_m(\lambda_0)}{\lambda - \lambda_0} = \sum_{k=0}^{m-1} Q_k(\lambda)\big[Q_k(\lambda_0)P_m(\lambda_0) - Q_m(\lambda_0)P_k(\lambda_0)\big], \tag{A.1}$$

$$\beta_i\big[P_i(\lambda)Q_{i-1}(\lambda) - P_{i-1}(\lambda)Q_i(\lambda)\big] = 1, \tag{A.2}$$

$$y_j = \left[\sum_{i=0}^{n-1} Q_i(\lambda_j)^2\right]^{-1} \tag{A.3}$$

[9–Chapter 2, the proof of Lemma 7.4, formulae (5.30), (5.24)],

$$S_j = \sqrt{y_j}\,\big(Q_0(\lambda_j), \ldots, Q_{n-1}(\lambda_j)\big)^{\mathrm{T}} \tag{A.4}$$

[29–formula (7.23) with proper normalization].

Consider perturbation of an individual node λ_k. Using the orthonormality and taking the limit value for the indefiniteness $\frac{0}{0}$), we derive

$$\frac{1}{2}\frac{\partial e_{ii}}{\partial \lambda_k} = y_k Q_{i-1}(\lambda_k) Q'_{i-1}(\lambda_k) \overset{(\text{A.1})}{=} \sum_{\substack{j=1\\ j\neq k}}^{n} \frac{y_j}{\lambda_k - \lambda_j} Q_{i-1}(\lambda_j)[Q_{i-1}(\lambda_j) - Q_{i-1}(\lambda_k)] \overset{(\text{A.3,A.4})}{=} \frac{\rho_k^{(i)}}{\Delta_k}.$$

Further,

$$\frac{\partial e_{i,i-1}}{\partial \lambda_k} = y_k (Q_{i-1}Q_i)'(\lambda_k) \overset{(\text{A.1})}{=} \sum_{j=1}^{n} \frac{y_j}{\lambda_k - \lambda_j} Q_{i-1}(\lambda_j)[Q_i(\lambda_j) - Q_i(\lambda_k)] + y_k Q_{i-1}(\lambda_k) Q'_i(\lambda_k)$$

$$
\begin{aligned}
&+\sum_{j=1}^{n}\frac{y_j}{\lambda_k-\lambda_j}Q_i(\lambda_j)[Q_{i-1}(\lambda_j)-Q_{i-1}(\lambda_k)]+y_kQ_i(\lambda_k)Q'_{i-1}(\lambda_k)\\
&+[Q_{i-1}(\lambda_k)P_i(\lambda_k)-Q_i(\lambda_k)P_{i-1}(\lambda_k)]\\
\overset{(A.2)}{=}\;&\frac{1}{\beta_i}+\sum_{\substack{j=1\\ j\neq k}}^{n}\frac{y_j}{\lambda_k-\lambda_j}Q_{i-1}(\lambda_j)[Q_i(\lambda_j)-Q_i(\lambda_k)]\\
&+\sum_{\substack{j=1\\ j\neq k}}^{n}\frac{y_j}{\lambda_k-\lambda_j}Q_i(\lambda_j)[Q_{i-1}(\lambda_j)-Q_{i-1}(\lambda_k)]=\frac{1}{\beta_i}\\
&+2\sum_{\substack{j=1\\ j\neq k}}^{n}\frac{y_j}{\lambda_k-\lambda_j}\big[2Q_{i-1}(\lambda_j)Q_i(\lambda_j)\\
&-Q_{i-1}(\lambda_j)Q_i(\lambda_k)-Q_i(\lambda_j)Q_{i-1}(\lambda_k)\big]\\
\overset{(A.3,A.4)}{=}\;&\frac{\zeta_k^{(i)}}{\Delta_k}.
\end{aligned}
$$

Looking at formulae (34) and (35), one may think that small weights can make the right-hand sides huge. The orthogonal polynomial approach shows that this is not so. Imagine that y_k is tiny. Define the orthonormal polynomials R_i and the polynomials of second kins S_i corresponding to the scalar product $\langle f,g\rangle_k\equiv\sum_{\substack{j=1\\ j\neq k}}^{n}y_jf(\lambda_j)g(\lambda_j)$. We have $R_i(\lambda)\approx Q_i(\lambda)$ for $0\leqslant i\leqslant n-2$. This implies

$$
\frac{\left|\rho_k^{(i)}\right|}{\Delta_k}\approx\left|\sum_{\substack{j=1\\ j\neq k}}^{n}\frac{y_j}{\lambda_k-\lambda_j}R_{i-1}(\lambda_j)[R_{i-1}(\lambda_j)-R_{i-1}(\lambda_k)]\right|\overset{(A.1)}{=}0
$$

and, analogously,

$$
\frac{\zeta_k^{(i)}}{\Delta_k}\approx\frac{1}{\beta_i}.
$$

In fact, paper [14] gives estimates for arbitrary positive weights, but at the expense of a polynomial in n growth of bounds (even for equal weights).

Appendix B. Classic Jacobi inverse eigenvalue problem

The classic Jacobi inverse eigenvalue problem is stated in terms of the mixed spectrum

$$
\lambda_1<\mu_1<\lambda_2<\cdots<\mu_{n-1}<\lambda_n, \tag{B.1}
$$

where λ_i are the same as earlier and μ_i are the eigenvalues of the principal lower $(n-1) \times (n-1)$ submatrix of $\mathscr{T}$. In this formulation, eigenvalue perturbations are to be compared to the separations in the mixed spectrum (B.1). The weights y_i are expressed in terms of the elements of (B.1) by means of the simple formula

$$y_i = \prod_{j=1}^{n-1}(\lambda_i - \mu_j) \Bigg/ \prod_{1\leqslant j\leqslant n,\ j\neq i} (\lambda_i - \lambda_j),$$

which allows one to easily estimate relative perturbations of y_i when the mixed spectrum is perturbed. Therefore, we can utilize the chain rule to obtain estimates in terms of $d\lambda_j$ and $d\mu_i$.

Acknowledgments

We thank F. Natterer for the publication of manuscript [22] on the Internet. The work of L. Borcea was partially supported by the Office of Naval Research, under grant N00014-02-1-0088 and by the National Science Foundation, under grant DMS-0305056.

References

[1] D. Boley, G.H. Golub, A survey of matrix inverse eigenvalue problems, Inverse Problems 3 (1987) 595–622.
[2] M.T. Chu, G.H. Golub, Structured inverse eigenvalue problems, Acta Numerica (2001) 1–70.
[3] C. de Boor, G.H. Golub, The numerically stable reconstruction of a Jacobi matrix from spectral data, Linear Algebra Appl. 21 (1978) 245–260.
[4] H. Hochstadt, On some inverse problems in matrix theory, Arch. Math. 18 (1967) 201–207.
[5] H. Hochstadt, On the construction of a Jacobi matrix from spectral data, Linear Algebra Appl. 8 (1974) 435–446.
[6] L.J. Gray, D.G. Wilson, Construction of a Jacobi matrix from spectral data, Linear Algebra Appl. 14 (1976) 131–134.
[7] O.H. Hald, Inverse eigenvalue problems for Jacobi matrices, Linear Algebra Appl. 14 (1976) 63–85.
[8] B.N. Parlett, D.S. Scott, The Lanczos algorithm with selective orthogonalization, Math. Comput. 33 (1979) 217–238.
[9] E.M. Nikishin, V.N. Sorokin, Rational Approximations and Orthogonality, in: Translations of Mathematical Monographs, vol. 92, American Mathematical Society, Providence, RI, 1991.
[10] W. Gautschi, On generating orthogonal polynomials, SIAM J. Sci. Statist. Comput. 3 (3) (1982) 289–317.
[11] H.J. Fischer, On generating orthogonal polynomials for discrete measures, J. Anal. Appl. (Z. Anal. Anwendungen) 17 (1998) 183–206.
[12] G.H. Golub, C.F. van Loan, Matrix Computations, second ed., The John Hopkins University Press, Baltimore and London, 1989.
[13] W. Gautschi, Computational aspects of orthogonal polynomials, in: P. Nevai (Ed.), Orthogonal Polynomials, Kluwer Academic Publishers, 1990, pp. 181–216.
[14] L. Knizhnerman, Stability estimates on the Jacobi and unitary Hessenberg inverse eigenvalue problems, SIAM J. Matrix Anal. Appl. 26 (1) (2004) 154–169.

[15] S.F. Xu, A stability analysis of the Jacobi inverse eigenvalue problem, BIT 33 (4) (1993) 695–702.
[16] M.T. Chu, Inverse eigenvalue problems, SIAM Rev. 40 (1998) 1–39.
[17] I.M. Gel'fand, B.M. Levitan, On the determination of a differential equation from its spectral function, AMS Translations (1956) 253–304.
[18] B.M. Levitan, Inverse Sturm–Liouville Problems, VNU Science Press, Utrecht, 1987.
[19] R. Burridge, The Gel'fand–Levitan, the Marchenko and the Gopinath–Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse–response problems, Wave Motion 2 (1980) 305–323.
[20] T. Habashy, A generalized Gel'fand–Levitan–Marchenko integral equation, Inverse Problems 7 (1991) 703–711.
[21] G.M.L. Gladwell, N.B. Willms, A discrete Gel'fand–Levitan method for band-matrix inverse eigenvalue problems, Inverse Problems 5 (2) (1989) 165–179.
[22] F. Natterer, Discrete Gel'fand–Levitan theory, electronic, from the author's web page.
[23] K. Chadan, D. Colton, L. Päivärinta, W. Rundell, An Introduction to Inverse Scattering and Inverse Spectral Problems, SIAM, Philadelphia, PA, 1997.
[24] J.W. Paine, F.R. de Hoog, R.S. Andersen, On the correction of finite difference eigenvalue approximations for Sturm–Liouville problems, Computing 26 (1981) 123–139.
[25] L. Borcea, V. Druskin, Optimal finite difference grids for direct and inverse Sturm–Liouville problems, Inverse Problems 18 (2002) 979–1001.
[26] L. Borcea, V. Druskin, L. Knizhnerman, On the continuum limit of a discrete inverse spectral problem on optimal finite difference grids, Comm. Pure Appl. Math., accepted.
[27] A.A. Bukhshtab, Number Theory, Prosveschchenie, Moscow, 1966.
[28] C.C. Paige, Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix, J. Inst. Math. Appl. 18 (1976) 341–349.
[29] B.N. Parlett, The Symmetric Eigenvalue Problem, SIAM, Prentice-Hall, 1998.
[30] A. Greenbaum, Behavior of slightly perturbed Lanczos and conjugate-gradient recurrences, Linear Algebra Appl. 113 (1989) 7–63.
[31] J.K. Cullum, R.H. Willoughby, Lanczos algorithms for large symmetric eigenvalue computations, SIAM, Philadelphia, PA, 2002.