Math. Nachr. **?** (1997),

Dependence of Eigenvalues on the Problem

By Q. Kong, H. Wu and A. Zettl

(Received)

Abstract. The eigenvalues of linear, regular, two point boundary value problems depend continuously on the problem. In the important self-adjoint case studied by Naimark and Weidmann this dependence is differentiable and the derivatives of the eigenvalues with respect to a given parameter: an endpoint, a boundary condition, a coefficient, or the weight function, are found. Monotone properties of the eigenvalues with respect to the coefficients and the weight function are established without using the variational (min-max) characterization.

1. Introduction

It is part of the folklore of Mathematics that the eigenvalues of linear, ordinary, self-adjoint, regular boundary value problems (BVP's) depend continuously on the problem. The state of the art software package SLEDGE, developed by Fulton and Pruess [8], for computing eigenvalues of separated regular, and some singular, Sturm-Liouville problems is based on approximating the coefficients of the differential equation by piece-wise constant functions and then computing the eigenvalues of the approximate problem. The code SLEUTH, developed by Greenberg and Marletta [9], computes eigenvalues of fourth order problems and is also based on piece-wise constant approx-

1991 *Mathematics Subject Classification.* Primary 34B24, 34L15; Secondary: 34L05
Keywords and phrases. eigenvalues, linear boundary value problems, continuous dependence on parameters

imations of the coefficients. The code SLEIGN, developed by Bailey, Gordon and Shampine [2], as well as the new code SLEIGN2, developed by Bailey, Everitt and Zettl [1], for computing eigenvalues of regular and singular Sturm-Liouville problems use interval truncations, at least in some cases.

When computing an eigenvalue of a BVP on the interval $[0, \pi]$ with these or other codes, it is common practice to use a truncated interval, say $[0, 3.14159265359]$, and expect to get a good approximation. Indeed, many authors do this without comment. This expectation is, perhaps, not unreasonable if the eigenvalues depends continuously on the right end-point of the interval.

In this paper we show that any isolated eigenvalue of a regular self-adjoint or non-self-adjoint ordinary linear n-th order BVP depends continuously on the problem: on the end-points, the boundary conditions, the coefficients of the differential equation, and the weight function. The continuous dependence on the coefficients and on the weight function is with respect to the L^1 norm, which we believe is the "natural" norm for the study of regular BVP. We also show that the eigenfunctions depend continuously on the problem.

For an important special class of regular BVP - the class studied by Naimark in his celebrated book [15], and by Weidmann [17] - we show that the eigenvalues depend differentiably on the problem data and find their derivatives.

The monotone properties of the eigenvalues with respect to the coefficients and the weight function are established as a consequence of the differentiability results and their proofs - without using the variational min-max characterization. This appears to be a new, and very general, method for proving eigenvalue comparisons.

With all coefficients and the weight function fixed with one exception, say q, we study the dependence of a given eigenvalue $\lambda(q)$ on q, and show that the set of BVP's for which $\lambda(q)$ is simple is an open set in L^1; on the other hand, the set of q such that the BVP for which $\lambda(q)$ is not simple is generically a closed and nowhere dense set in L^1.

Our continuity results extend theorems in Kong and Zettl [11] from the second order self-adjoint case to the n-th order non-self-adjoint case; the differentiability results here also extend theorems in [11] from $n = 2$ to the general even order case. The continuity proof in [11] depends on norm resolvent convergence of self-adjoint operators in different Hilbert spaces. Our proof here is based on a basic result in complex variable theory. The proof of the differentiability of the eigenvalues depends on their continuity and on continuity properties of eigenfunctions.

2. Notation and Quasi-Differential Expressions

In this section we introduce our notation and discuss some basic facts about quasi-differential equations needed below. For a more comprehensive discussion of quasi-differential equations the reader is referred to [19] and to [6] for the scalar coefficient case and to [7] for the general case with matrix coefficients.

Throughout this paper, $\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, and $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$; $J = (a', b')$ denotes an arbitrary open interval of the real line, bounded or unbounded:

$$J = (a' , b') \quad -\infty \le a' < b' \le \infty. \tag{2.1}$$

Let $n, m \in \mathbb{N}$ with $n > 1$. For a given set S, $M_{n,m}(S)$ denotes the set of $n \times m$ matrices with entries from S. If $n = m$, we write also $M_n(S)$, and if $m = 1$, we may also write S^n. For any $A \in M_{n,m}(\mathbb{C})$, A^T and A^* denote the transpose and the complex conjugate transpose of A, respectively. We denote by $L(J), L_{loc}(J), AC_{loc}(J)$ the sets of complex valued functions which are integrable on J, integrable on every compact subinterval of J, and absolutely continuous on every compact subinterval of J, respectively. We also denote by $L^R(J), L^R_{loc}(J)$ the sets of real valued functions with similar meanings to $L(J), L_{loc}(J)$. For $a' < a < b < b'$ and $w \in L^R_{loc}(a', b')$, $w > 0$ *a.e.* on (a, b), $L^2_w(a, b)$ is used for the Hilbert space of (equivalence classes of) functions u such that $|u|^2 w \in L(a, b)$ with the norm $\|u\| = (\int_a^b |u|^2 w)^{1/2}$.

Following Everitt and Race [5] we let

$$Z_n(J) = \left\{ P = (p_{rs})_{r,s=1}^n \in L_{\text{loc}}(J) : \begin{array}{l} p_{r,r+1} \neq 0 \text{ a.e. for } 1 \le r \le n-1 \\ p_{r,s} = 0 \text{ a.e. for } 2 \le r+1 < s \le n \end{array} \right\}. \tag{2.2}$$

Let $P \in Z_n(J)$, $\mathcal{D}_\prime := \{\dagger : \mathcal{J} \to \mathbb{C},\ \dagger \text{ measurable}\}$, and $y^{[0]} := y$ for $y \in \mathcal{D}_\prime$. Inductively, for $r = 1, \cdots, n$, we define

$$\mathcal{D}_\nabla = \{y \in \mathcal{D}_{\nabla-\infty} : \dagger^{[\nabla-\infty]} \in \mathcal{AC}_{\text{loc}}(\mathcal{J})\}, \tag{2.3}$$

$$y^{[r]} = p_{r,r+1}^{-1} \left((y^{[r-1]})' - \sum_{s=1}^{r} p_{r,s} y^{[s-1]} \right) \quad \text{for } y \in \mathcal{D}_\nabla, \tag{2.4}$$

where $p_{n,n+1} := 1$. Finally we set

$$\mathcal{M}_\mathcal{P} y := i^n y^{[n]} \quad \text{for } y \in \mathcal{D}_\backslash. \tag{2.5}$$

The expression $\mathcal{M} = \mathcal{M}_\mathcal{P}$ is called the quasi-differential expression associated with P. For $\mathcal{D}_\backslash$ we also use the notation $\mathcal{D}(\mathcal{P})$. The function $y^{[r]}$ $(0 \le r \le n)$ is called the r-th quasi-derivative of y; since the quasi-derivative depends on P, we write $y_P^{[r]}$ instead of $y^{[r]}$ when we wish to emphasize its dependence on P. For any $y \in \mathcal{D}(\mathcal{P})$, let

$$Y = \begin{pmatrix} y^{[0]} \\ y^{[1]} \\ \vdots \\ y^{[n-1]} \end{pmatrix} = \begin{pmatrix} y_P^{[0]} \\ y_P^{[1]} \\ \vdots \\ y_P^{[n-1]} \end{pmatrix} \tag{2.6}$$

be the vector function associated with y and P.

Remark 2.1. *The "expression domain" $\mathcal{D}(\mathcal{P})$ is a vector space over $\mathbb{C}$ and the operator $\mathcal{M} : \mathcal{D}(\mathcal{P}) \to (\mathcal{AC}_{\mathrm{loc}}(\mathcal{J}))^{\setminus}$ defined by (2.6) is linear. We think of $y^{[n]}$ and of $\mathcal{M}\dagger$ as a differential expression with expression domain $\mathcal{D}(\mathcal{P})$ and of the intermediate quasi-derivatives $y^{[r]}$, $r = 0, 1, \cdots n-1$, as generalizations of the classical derivatives $y^{(r)}$.*

Although the factor i^n in (2.5) plays no significant role in this paper, we include it here in order to be consistent with the widely accepted notation in the literature.

Theorem 2.2. *Let $u \in J$, $C = (c_j)_{j=1}^n \in M_{n,1}(\mathbb{C})$. Assume that $P \in Z_n(J)$, $f \in L_{loc}(J)$. Then there exists a unique $y \in \mathcal{D}_\setminus$ such that*

$$y_P^{[n]} = f \, on \, J \quad and \quad y_P^{[r]}(u) = c_r, \; r = 0, 1, \cdots, n-1. \tag{2.7}$$

Furthermore, this solution $y = y(\cdot, u, C, P, f)$ and its quasi-derivatives $y^{[r]} = y^{[r]}(\cdot, u, C, P, f)$, $r = 0, 1, 2, \ldots, n-1$, are continuous functions of the problem; in particular, they depend continuously on u, on $C = (c_j)_{j=0}^{n-1}$, on P, and on f, uniformly on compact subintervals of J. More precisely, given a compact subinterval $[a, b]$ of J and given an $\epsilon > 0$ there exists a $\delta > 0$ such that if $u, v \in [a, b]$, $C, D \in M_n(\mathbb{C})$, $P, Q \in Z_n(J)$, $f, g \in L_{loc}(J)$ satisfy

$$|u - v| + ||C - D|| + \int_a^b |P - Q| + \int_a^b |f - g| < \delta,$$

then

$$|y^{[r]}(t, u, C, P, f) - y^{[r]}(t, v, D, Q, g)| < \epsilon \quad for \; all \; t \in [a, b], \quad r = 0, 1, \ldots, n-1.$$

Here $|| \cdot ||$ denotes any matrix norm and

$$\int_a^b |P - Q| = \sum_{s \le r} \int_a^b |p_{rs} - q_{rs}| + \sum_{r=1}^{n-1} \int_a^b \left| \frac{1}{p_{r,r+1}} - \frac{1}{q_{r,r+1}} \right|. \tag{2.8}$$

Proof. See [12]. □

3. Regular Boundary Value Problems

In this section we establish the characterization of the eigenvalues as zeros of an entire function and prove the continuity of the eigenvalues and eigenfunctions for two point boundary value problems, self-adjoint or not.

Let $P \in Z_n(J), w \in L^R_{loc}(J)$, and for each $u \in J$, let $\Phi(\cdot, u, P, w, \lambda)$ be the matrix solution of the initial value problem (IVP)

$$Y' = (P + (-i)^n \lambda W)Y \quad \text{on } J, \tag{3.1}$$

with $Y(u) = I$, where I is the identity matrix and $W = (w_{ij})$ is the $n \times n$ matrix such that $w_{n1} = w$, and $w_{ij} = 0$ otherwise.

We study BVP defined on compact subintervals of J. Let $a' < a < b < b'$, $A, B \in M_n(\mathbb{C})$, and consider the two point, in general non-self-adjoint, BVP:

$$\mathcal{M}_{\mathcal{P}}\ \dagger = \rangle^{\backslash}\dagger^{[\backslash]} = \lambda \sqsupseteq \dagger \text{ on } [\dashv, \lfloor], \quad \lambda \in \mathbb{C}, \tag{3.2}$$

$$AY(a) + BY(b) = 0 \tag{3.3}$$

where Y is the vector function associated with y and P according to (2.6). Note that this is a regular BVP since the components of P and w are in $L(a, b)$; consequently, the solution y and its quasi-derivatives can be continuously extended to the end-points a and b, see [12].

The characteristic function Δ of problem (3.2), (3.3) is defined as follows:

$$\Delta(\lambda) = \det[A + B\,\Phi(b, a, P, w, \lambda)], \ \lambda \in \mathbb{C}. \tag{3.4}$$

Theorem 3.3. *Let $P \in Z_n(J)$, $w \in L^R_{loc}(J)$, $A, B \in M_n(\mathbb{C})$, and let $\Delta(\lambda)$ be defined by (3.4) for $\lambda \in \mathbb{C}$. Then*

1. *Δ is an entire function of λ and consequently, either $\Delta(\lambda)$ is identically zero or each of its zeros, if there are any, is an isolated point in $\mathbb{C}$. Thus there are only the following four possibilities:*

   (i) Δ is identically zero on $\mathbb{C}$,

   (ii) Δ has no zero in $\mathbb{C}$,

   (iii) Δ has a finite number of zeros in $\mathbb{C}$,

   (iv) Δ has an infinite but countable number of zeros in $\mathbb{C}$.

2. *A complex number λ is an eigenvalue of the boundary value problem (3.2), (3.3) if and only if $\Delta(\lambda) = 0$.*

Proof. The fact that Δ is an entire function of λ follows from its definition and from the fact that solutions of IVP depend analytically on λ - see [12]. For the proof of part (2) we notice, see [14], that the scalar equation (3.2) is equivalent to the first order system (3.1). Noting that

$$Y(t) = \Phi(t, a, P, w, \lambda)\, Y(a)$$

and using the boundary conditions (3.3) we get

$$[A + B\Phi(b, a, P, w, \lambda)]\, Y(a) = 0.$$

This linear algebraic equation has a non-trivial solution for $Y(a)$, and hence for $Y(t)$ and for $y(t)$, if and only if $\Delta(\lambda) = 0$. □

We want to show that a small change of the problem results in only a small change in the eigenvalues and eigenfunctions. How does one measure the closeness of two BVP to each other? To answer this question we introduce the " boundary value problem space" $\mathbf{\Omega}$.

Define

$$\text{(3.5)} \quad \mathbf{\Omega} = \left\{ \boldsymbol{\omega} = (a, b, A, B, P, w) : \begin{array}{l} a' < a < b < b',\ A, B \in M_n(\mathbb{C}),\ P \in Z_n(J), \\ w \in L^R_{loc}(J),\ w > 0 \ a.e. \text{ on } J \end{array} \right\}.$$

Note that for any $a' < a < b < b'$, we have $P \in M_n(L(a, b)),\ w \in L(a, b)$ so that the BVP (3.2), (3.3) is a well defined regular problem on the interval $[a, b]$. Each $\boldsymbol{\omega} \in \mathbf{\Omega}$ determines a unique BVP: a and b the interval, A and B the boundary condition, and the restriction of P and of w to the interval $[a, b]$ the equation. Observe that the values of P and w outside $[a, b]$ do not affect the spectrum of the problem determined by $\boldsymbol{\omega} \in \mathbf{\Omega}$. To account for this and to facilitate comparing problems which may be defined on different intervals we introduce

$$\text{(3.6)} \qquad \tilde{\mathbf{\Omega}} = \{\tilde{\boldsymbol{\omega}} = (a, b, A, B, \tilde{P}, \tilde{w})\}$$

where

$$\text{(3.7)} \qquad \tilde{P} = \begin{cases} P & \text{on } [a, b] \\ 0 & \text{otherwise} \end{cases}$$

and $\tilde{w}$ is defined similarly. We introduce the space

$$\text{(3.8)} \qquad X = (a', b') \times (a', b') \times M_n(\mathbb{C}) \times M_n(\mathbb{C}) \times M_n(L(a', b')) \times L(a', b')$$

with its product topology. The topology on $M_n(\mathbb{C})$ is determined by any fixed matrix norm, on $L(a', b')$ by the usual norm, on $M_n(L(a', b')$ we use the topology determined by

$$\int_{a'}^{b'} |\tilde{P}| = \int_a^b |P| = \sum_{r \geq s} \int_a^b |p_{rs}| + \sum_{r=1}^{n-1} \int_a^b \frac{1}{|p_{r,r+1}|}. \tag{3.9}$$

The space X is the "natural" setting for the study of regular BVP. Note that, since p_{rs} for $r \geq s$, $1/p_{r,r+1}$, $r = 1, 2, \ldots, k$ and w are only assumed to be in $L_{loc}(a', b')$, $\mathbf{\Omega}$ is not a subset of X but $\tilde{\mathbf{\Omega}}$ is since $\tilde{P}, \tilde{w}$ are in $M_n(L(a', b'))$, $L(a', b')$, respectively. Now we identify $\mathbf{\Omega}$ with $\tilde{\mathbf{\Omega}}$ as a subset of X. Then $\mathbf{\Omega}$ inherits the norm from X, and the convergence in $\mathbf{\Omega}$ is determined by this norm. Note that every point in $\mathbf{\Omega}$ is an accumulation point of $\mathbf{\Omega}$ with respect to this norm from X and hence it makes sense to discuss convergence of BVP's with respect to this norm.

The isolated eigenvalues of regular - not necessarily self-adjoint - BVP's depend continuously on the problem. More precisely we have

Theorem 3.4. *Let $\mathbf{\Omega}$ be defined as above. Let $\boldsymbol{\omega}_0 = (a, b, A, B, P, w) \in \mathbf{\Omega}$. Assume that a complex number $\mu = \lambda(\boldsymbol{\omega}_0)$ is an isolated eigenvalue of the BVP (3.2), (3.3) determined by $\boldsymbol{\omega}_0$, i.e. μ is an isolated eigenvalue of the BVP*

$$\mathcal{M}_{\mathcal{P}} \dagger = \rangle^{\backslash} \dagger^{[\backslash]}_{\mathcal{P}} = \lambda \sqsupseteq \dagger \text{ on } [\dashv, \lfloor], \quad \mathcal{AY}(\dashv) + \mathcal{BY}(\lfloor) = \prime. \tag{3.10}$$

Then, given any $\epsilon > 0$, there exists a $\delta > 0$ such that for any $\boldsymbol{\omega} = (c, d, C, D, Q, v) \in \mathbf{\Omega}$ satisfying the inequality

$$||\boldsymbol{\omega} - \boldsymbol{\omega}_0|| = |a - c| + |b - d| + ||A - C|| + ||B - D|| + \int_{a'}^{b'} \left(|\tilde{P} - \tilde{Q}| + |\tilde{w} - \tilde{v}| \right) < \delta, \tag{3.11}$$

the BVP

$$\mathcal{M}_{\mathcal{Q}} \ddagger = \rangle^{\backslash} \ddagger^{[\backslash]}_{\mathcal{Q}} = \lambda \sqsubseteq \ddagger \text{ on } [\rfloor, \lceil], \quad \mathcal{CZ}(\rfloor) + \mathcal{DZ}(\lceil) = \prime, \tag{3.12}$$

has an isolated eigenvalue $\lambda(\boldsymbol{\omega})$ satisfying the inequality

$$|\lambda(\boldsymbol{\omega}) - \lambda(\boldsymbol{\omega}_0)| \; < \; \epsilon. \tag{3.13}$$

Here Y, Z are the vectors associated with y, P and with z, Q, respectively, according to (2.6).

Proof. The proof is based on the characterization of the eigenvalues as zeros of the function Δ given by (3.4). Let $\Delta = \Delta(\boldsymbol{\omega}, \lambda)$ for $\boldsymbol{\omega} = (a, b, A, B, P, w) \in \mathbf{\Omega}$, $\lambda \in \mathbb{C}$. Then for any $\boldsymbol{\omega} \in \mathbf{\Omega}$, $\Delta(\boldsymbol{\omega}, \lambda)$ is an entire function of λ; furthermore, it is continuous in $\boldsymbol{\omega}$ - see Theorems 2.7, 2.8 of [12], and $\Delta(\boldsymbol{\omega}_0, \mu) = 0$. Thus $\Delta(\boldsymbol{\omega}_0, \lambda)$ is not constant in λ since μ is an isolated eigenvalue. Hence there exists $\rho > 0$ such that $\Delta(\boldsymbol{\omega}_0, \lambda) \neq 0$ for $\lambda \in S_\rho := \{\lambda \in \mathbb{C} : |\lambda - \mu| = \rho\}$. By the well known theorem on continuity of the roots of an equation as a function of parameters, see [3], 9.17.4, the statement of Theorem 3.2 follows. □

Remark 3.5. *Below, when we speak of an eigenvalue of $\boldsymbol{\omega} = (a, b, A, B, P, w)$ we mean an eigenvalue of the BVP (3.2), (3.3) determined by $\boldsymbol{\omega}$.*

Note that, given the existence of an isolated eigenvalue $\lambda(\boldsymbol{\omega}_0)$ for a BVP determined by $\boldsymbol{\omega}_0$, the proof of Theorem 3.2 establishes the existence of isolated eigenvalues for all BVP associated with $\boldsymbol{\omega}$ if $\boldsymbol{\omega}$ is sufficiently close to $\boldsymbol{\omega}_0$, and shows that these eigenvalues $\lambda(\boldsymbol{\omega})$ are close to the given one $\lambda(\boldsymbol{\omega}_0)$. If the problem associated with $\boldsymbol{\omega}_0$ is self-adjoint, then it has countably many eigenvalues λ_r. Hence, it follows from Theorem 3.2 that, given any positive integer m, if $\boldsymbol{\omega}$ is sufficiently close to $\boldsymbol{\omega}_0$, then the BVP determined by $\boldsymbol{\omega}$ - whether it is self-adjoint or not - must have at least m eigenvalues, and they are close to the corresponding eigenvalues of $\boldsymbol{\omega}_0$. Multiple eigenvalues are counted according to their algebraic multiplicity.

Theorem 3.2 shows that for any fixed isolated eigenvalue μ associated with $\boldsymbol{\omega}_0$ there exists a continuous eigenvalue branch $\lambda(\boldsymbol{\omega})$ satisfying $\lambda(\boldsymbol{\omega}_0) = \mu$. However, even for a self-adjoint BVP, this does not mean that for a fixed n, the n-th eigenvalue $\lambda_n(\boldsymbol{\omega})$ is always continuous in $\boldsymbol{\omega}$, see [11, remark 3.1].

Below we consider only the continuous eigenvalue branches as $\boldsymbol{\omega}$ varies in $\boldsymbol{\Omega}$.

Theorem 3.6. *Let the hypotheses and notation of Theorem 3.2 hold. Assume that $\lambda = \lambda(\boldsymbol{\omega}_0)$ is an isolated eigenvalue of $\boldsymbol{\omega}_0$. If λ has geometric multiplicity l at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, for $l = 1, 2, \ldots, n$, then there exists a neighborhood $\mathcal{N}$ of $\boldsymbol{\omega}_0$ in $\boldsymbol{\Omega}$ such that $\lambda(\boldsymbol{\omega})$ is an isolated eigenvalue of $\boldsymbol{\omega}$ of geometric multiplicity at most l for each $\boldsymbol{\omega} \in \mathcal{N}$.*

Proof. For $\boldsymbol{\omega} \in \mathcal{N}$ let

$$D(\boldsymbol{\omega}) = A + B\,\Phi(b, a, P, w, \lambda(\boldsymbol{\omega})).$$

Note that D depends continuously on $\boldsymbol{\omega}$. The solutions of BVP (3.2), (3.3) are the first components of the vector functions $\Phi(\cdot, a, P, w, \lambda(\boldsymbol{\omega}))d$ with $d \in \mathbb{C}^n$, $D(\boldsymbol{\omega})d = 0$, and these functions are not identically zero if $d \neq 0$. Hence the geometric multiplicity of $\lambda(\boldsymbol{\omega})$ equals the dimension of the kernel $N(D(\boldsymbol{\omega}))$ of $D(\boldsymbol{\omega})$. In our case, this means that rank$D(\boldsymbol{\omega}_0) = n - l$. Since D depends continuously on $\boldsymbol{\omega}$, $D(\boldsymbol{\omega})$ must have rank at least $n - l$ in a neighborhood of $\boldsymbol{\omega}_0$. This complete the proof. □

By a normalized eigenfunction u of the BVP (3.2), (3.3), we mean an eigenfunction u that satisfies

$$\int_a^b |u|^2 w = 1. \tag{3.14}$$

Theorem 3.7. *Let the hypotheses and notation of Theorem 3.2 hold. Assume that $\lambda = \lambda(\boldsymbol{\omega})$ is an isolated eigenvalue of $\boldsymbol{\omega}$ of geometric multiplicity l for each $\boldsymbol{\omega}$ in some neighborhood $\mathcal{N}$ of $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$. Then there exist l linearly independent normalized eigenfunctions $u_k(\cdot, \boldsymbol{\omega})$ such that*

$$u_k^{[j]}(\cdot, \boldsymbol{\omega}) \to u_k^{[j]}(\cdot, \boldsymbol{\omega}_0), \quad \text{as } \boldsymbol{\omega} \to \boldsymbol{\omega}_0 \in \boldsymbol{\Omega}, \quad k = 1, 2, \cdots, l, \quad j = 0, 1, \cdots, n-1, \tag{3.15}$$

all uniformly on any compact subinterval of J.

In particular, if $\lambda(\boldsymbol{\omega}_0)$ is an isolated simple eigenvalue, then $\lambda(\boldsymbol{\omega})$ is an isolated simple eigenvalue for each $\boldsymbol{\omega}$ in some neighborhood $\mathcal{N}$ of $\boldsymbol{\omega}_0$, and there exist normalized eigenfunctions $u_1 = u_1(\cdot, \boldsymbol{\omega})$ such that (3.15) holds with $k = 1$.

Proof. Let $E_2 = N(D(\boldsymbol{\omega}_0))$, E_1 a complementary space of E_2 in $\mathbb{C}^n$, $F_1 = R(D(\boldsymbol{\omega}_0))$, F_2 a complementary space of F_1 in $\mathbb{C}^n$. Then

$$D(\boldsymbol{\omega}) = \begin{pmatrix} D_{11}(\boldsymbol{\omega}) & D_{12}(\boldsymbol{\omega}) \\ D_{21}(\boldsymbol{\omega}) & D_{22}(\boldsymbol{\omega}) \end{pmatrix} : E_1 \oplus E_2 \to F_1 \oplus F_2$$

with $D_{11}(\boldsymbol{\omega}_0)$ invertible. Since D_{11} depends continuously on $\boldsymbol{\omega}$, also $D_{11}(\boldsymbol{\omega})$ is invertible in a neighborhood $\mathcal{N}' \subset \mathcal{N}$ of $\boldsymbol{\omega}_0$, and the Schur factorization

$$D(\boldsymbol{\omega}) = C_1(\boldsymbol{\omega}) \begin{pmatrix} D_{11}(\boldsymbol{\omega}) & 0 \\ 0 & S(\boldsymbol{\omega}) \end{pmatrix} C_2(\boldsymbol{\omega})$$

holds, where

$$C_1(\boldsymbol{\omega}) = \begin{pmatrix} id_{F_1} & 0 \\ D_{21}(\boldsymbol{\omega}) D_{11}(\boldsymbol{\omega})^{-1} & id_{F_2} \end{pmatrix}, \quad C_2(\boldsymbol{\omega}) = \begin{pmatrix} id_{E_1} & D_{11}(\boldsymbol{\omega})^{-1} D_{12}(\boldsymbol{\omega}) \\ 0 & id_{E_2} \end{pmatrix},$$

$$S(\boldsymbol{\omega}) = D_{22}(\boldsymbol{\omega}) - D_{21}(\boldsymbol{\omega}) D_{11}(\boldsymbol{\omega})^{-1} D_{12}(\boldsymbol{\omega}).$$

By assumption, $D(\boldsymbol{\omega})$ has rank $n - l$. Therefore, also $C_1(\boldsymbol{\omega})^{-1} D(\boldsymbol{\omega}) C_2(\boldsymbol{\omega})^{-1}$ has rank $n - l$, which implies that $S(\boldsymbol{\omega}) = 0$. Then

$$D(\boldsymbol{\omega}) C_2(\boldsymbol{\omega})^{-1} = C_1(\boldsymbol{\omega}) \begin{pmatrix} D_{11}(\boldsymbol{\omega}) & 0 \\ 0 & 0 \end{pmatrix},$$

and hence

$$D(\boldsymbol{\omega})C_2(\boldsymbol{\omega})^{-1}\begin{pmatrix}0\\ id_{E_2}\end{pmatrix} = 0.$$

Thus

$$C_2(\boldsymbol{\omega})^{-1}\begin{pmatrix}0\\ id_{E_2}\end{pmatrix} = \begin{pmatrix}-D_{11}(\boldsymbol{\omega})^{-1}D_{12}(\boldsymbol{\omega})\\ id_{E_2}\end{pmatrix}$$

maps F_2 one-to-one into the null space of $D(\boldsymbol{\omega})$. From the consideration at the beginning of the proof of Theorem 3.3, the statement of Theorem 3.4 follows if we observe that normalization is continuous. □

Remark 3.8. *As a special case of Theorem 3.3, we can fix all the components of* $\boldsymbol{\omega}$ *except one, say* a*, and let this one vary in its appropriate space. Thus if an isolated eigenvalue of the BVP for* $\boldsymbol{\omega} = (a, b, A, B, P, w)$ *has geometric multiplicity* l *for a particular value of* a*, then it has geometric multiplicity at most* l *for all left endpoints near* a*, similarly for the right end-point* b*, the boundary matrices* A*,* B*, the components* $(p_{rs},\ r \geq s;\ 1/p_{r,r+1})$*, of* P *and the weight function* w*.*

Moreover, as in Theorem 3.4, we can show that if the geometric multiplicity of any isolated eigenvalue of the BVP (3.2), (3.3) is constant in some neighborhood of any one of the components of $\boldsymbol{\omega}$ *in its appropriate space, then (3.15) holds for an appropriate sequence of normalized eigenfunctions.*

4. Differentiability of Eigenvalues with respect to the Data

In the last section we showed that the isolated eigenvalues depend continuously on all the data; here we show that this dependence is in fact differentiable, at least for the important class of self-adjoint BVP studied by Naimark [15] and Weidmann [17].

The next theorem is a basic result in the spectral theory of differential operators; it is stated here for the convenience of the reader and because we know of no single reference containing this result in the generality given here.

Theorem 4.9. *Let* $\boldsymbol{\omega} = \boldsymbol{\omega}(a, b, A, B, P, w) \in \boldsymbol{\Omega}$*. In addition, assume that the coefficient matrix* P *satisfies the symmetry condition*

$$P = P^{+} = -E^{-1}P^{*}E; \quad E = ((-1)^{i}\delta_{i,n+1-j}) \tag{4.1}$$

and the boundary condition matrices A, B *satisfy the self-adjointness condition*

$$AE^{-1}A^{*} = BE^{-1}B^{*}, \quad \operatorname{rank}(A : B) = n \tag{4.2}$$

where $(A : B)$ *denotes the* $n \times 2n$ *matrix obtained by placing* B *to the right of* A*. Then*

(i) The BVP (3.2), (3.3) has a discrete spectrum σ which consists of an infinite but countable number of all real eigenvalues, each has a multiplicity between 1 and n. The multiplicity may be different for different eigenvalues.

(ii) The spectrum σ is not bounded. It may be bounded above or below but not both. The corresponding set of eigenfunctions $\{u_r\}$ can be orthonormalized in the Hilbert space $L^2_w(a, b)$.

(iii) The orthonormalized sequence of eigenfunctions is a bases of $L^2_w(a, b)$. Thus each $f \in L^2_w(a, b)$ has a Fourier series expansion:

$$f = \sum (f, u_r)u_r\, ; \tag{4.3}$$

where the convergence is in the norm of $L^2_w(a, b)$ and the sum extends over all orthonormalized eigenfunctions u_r.

(iv) If the order n is odd, $n > 1$, then the spectrum is unbounded both above and below. Thus we have

$$\sigma = \{\lambda_j : \ j \in \mathbb{Z}\} \tag{4.4}$$

with corresponding eigenfunctions u_j. The eigenvalues can be ordered to satisfy

$$\ldots \le \lambda_{-2} \le \lambda_{-1} \le \lambda_0 \le \lambda_1 \le \lambda_2 \le \ldots \tag{4.5}$$

and we adopt the convention that λ_0 denotes the smallest non-negative eigenvalue.

(v) If $n = 2k$ is even and the leading coefficient $p_{k,k+1}$ is positive *a.e.*, then the spectrum is bounded below but not above. The eigenvalues can be ordered such that

$$\sigma = \{\lambda_j : \ j \in \mathbb{N}_0\} \tag{4.6}$$

and

$$-\infty < \lambda_0 \le \lambda_1 \le \lambda_2 \le \ldots \tag{4.7}$$

(vi) If the order $n = 2k$ is even and the leading coefficient $p_{k,k+1}$ changes its sign in the interval $[a, b]$, i.e., both the sets $\{t \in [a, b], p_{k,k+1}(t) > 0\}$ and $\{t \in [a, b], p_{k,k+1}(t) < 0\}$ have positive Lebesgue measures, then the spectrum is unbounded above and below. In this case, the ordering of the eigenvalues is as in the odd order case described above.

Proof. Parts (i), (ii) and (iii) follow from the standard theory of regular, self-adjoint differential operators in Hilbert space, see [15], [17]; for (iv), (v) and (vi) see [13]. □

Remark 4.10. *We comment here on conditions (4.1) and (4.2). Given $P \in Z_n(J)$ the symmetry condition (4.1) guarantees that the differential expression $\mathcal{M}_\mathcal{P}$ generated by P is a symmetric (formally self-adjoint) differential expression. Condition (4.2) on the boundary matrices A, B guarantees that the boundary condition determined by A, B is self-adjoint. These two conditions, together with $w \in L^R_{loc}(J), w > 0$ a.e. on $J = (a', b')$, imply that the BVP (3.2), (3.3) is self-adjoint and can be identified with a self-adjoint operator in the Hilbert space $L^2_w(a, b)$ for $a, b \in J, a < b$, see [6], [15], [17], [18], [19].*

In the following we assume $w > 0$ *a.e.* on J, and consider matrices P which are of even order $n = 2k$ and have the following form:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & \cdots & 0 & 0 & 1/p_k & 0 & \cdots & 0 \\ 0 & \cdots & 0 & p_{k-1} & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & & \\ 0 & p_1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ p_0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.8}$$

To simplify the notation we write $P = \langle p_0, p_1, \ldots, p_k \rangle$. Note that $1/p_k$ is in the k-th row and $k + 1$ column, p_{k-1} is in row $k + 1$ and column k, ..., p_0 is in the n-th row and first column; the order of P is $n = 2k$, and $P \in Z_n(J)$ if

$$1/p_k,\ p_j \in L_{loc}(J), \quad j = 1, \ldots, k - 1, \quad p_k \neq 0 \text{ a.e.} \tag{4.9}$$

Also, P satisfies the symmetry condition (4.1) if and only if all the p_j are real valued on J. We denote by NF_k the class of real valued matrices P of the form (4.8) satisfying the conditions (4.9). These are the matrices P which generate the symmetric quasi-differential expressions $\mathcal{M}_\mathcal{P}$ studied by Naimark [15] and by Weidmann [17], see also [4]. We define

$$\mathbf{\Omega}_1 = \{\boldsymbol{\omega} = (a, b, A, B, P, w) \in \mathbf{\Omega} :\ A, B \text{ satisfy (4.2)},\ P = \langle p_0, p_1, \ldots, p_k \rangle \in NF_k\}.$$

Let $P = \langle p_0, p_1, \ldots, p_k \rangle$, and let $\mathcal{M} = \mathcal{M}_\mathcal{P}$ be the differential expression associated with P according to section 2. The quasi-derivatives of y are given by

$$\begin{aligned} y^{[j]} &= y^{(j)},\ j = 0, 1, \ldots, k-1, \\ y^{[k]} &= p_k y^{(k)}, \\ y^{[j]} &= (y^{[j-1]})' - p_{n-j} y^{[n-j]},\ j = k+1, \ldots, n. \end{aligned} \tag{4.10}$$

The cases $n = 2$, $n = 4$, $n = 6$ are of special interest:

$$\mathcal{M}\dagger = -(\surd_{\infty}\dagger')' + \surd_{\prime}\dagger, \tag{4.11}$$

$$\mathcal{M}\dagger = [(\surd_{\in}\dagger'')' - \surd_{\infty}\dagger']' - \surd_{\prime}\dagger, \tag{4.12}$$

$$\mathcal{M}\dagger = -\left([(\surd_{\ni}\dagger''')' + \surd_{\in}\dagger'']' + \surd_{\infty}\dagger'\right)' + \surd_{\prime}\dagger. \tag{4.13}$$

In general, for $n = 2k$

$$\mathcal{M}\dagger = (-\infty)^{\|}\left\{\left\{\cdots\left\{[(\surd_{\|}\dagger^{(\|)})' - \surd_{\|-\infty}\dagger^{(\|-\infty)}]' - \surd_{\|-\in}\dagger^{(\|-\in)}\right\}' - \cdots - \surd_{\infty}\dagger'\right\}' - \surd_{\prime}\dagger\right\}. \tag{4.14}$$

The equation corresponding to (3.2) is

$$(-1)^k\left\{\left\{\cdots\left\{[(p_k y^{(k)})' - p_{k-1}y^{(k-1)}]' - p_{k-2}y^{(k-2)}\right\}' - \cdots - p_1 y'\right\}' - p_0 y\right\} = \lambda w y. \tag{4.15}$$

In the following the definition of the Frechet derivative in Banach spaces will be used.

Definition. A map T from a Banach space X into a Banach space Y is Frechet differentiable at a point $x \in X$ if there exists a bounded linear operator dT_x which maps X into Y such that for $h \in X$

$$|T(x+h) - T(x) - dT_x(h)| = o(h),\ \text{as } h \to 0. \tag{4.16}$$

In the following we first state theorems on the Frechet differentiability of the eigenvalue λ of the BVP consisting of (4.15) and (3.3), and leave the proofs to the latter part of this section.

Theorem 4.11. *Let $n = 2k$ and $\boldsymbol{\omega} = \boldsymbol{\omega}(a, b, A, B, P, w) \in \mathbf{\Omega}_1$ with $P = \langle p_0, p_1, \ldots, p_k\rangle$. Let $\lambda = \lambda(\boldsymbol{\omega})$ be an eigenvalue of the BVP consisting of (4.15) and (3.3) associated with $\boldsymbol{\omega}$, and let $u = u(\boldsymbol{\omega})$ be a normalized eigenfunction for this λ.*

1. *Fix all components of* $\boldsymbol{\omega}$ *except* p_j *for some* $j \in \{0, 1, \ldots, k-1\}$, *and consider* $\lambda = \lambda(p_j)$ *as a function of* $p_j \in L^R(a,b)$. *Thus* $\lambda(p_j)$ *is a (nonlinear) functional on the Banach space* $L^R(a,b)$. *Assume that* $\lambda(p_j)$ *has constant multiplicity in a neighborhood* $\mathcal{N}$ *of* p_j *in* $L^R(a,b)$. *Then* λ *is Frechet differentiable at* p_j *in* $L^R(a,b)$ *and its Frechet derivative is given by*

$$d\lambda_{p_j}(h) = (-1)^{k+j-1} \int_a^b |u^{(j)}|^2\, h, \quad h \in L^R(a,b), \quad j = 0, 1, \ldots, k-1. \tag{4.17}$$

2. *Fix all components of* $\boldsymbol{\omega}$ *except* p_k *and consider* $\lambda = \lambda(1/p_k)$ *as a function of* $1/p_k$ *in* $L^R(a,b)$. *Thus* $\lambda(1/p_k)$ *is a (nonlinear) functional on* $L^R(a,b)$. *Assume that* $\lambda(1/p_k)$ *has constant multiplicity in a neighborhood* $\mathcal{N}$ *of* $1/p_k$ *in* $L^R(a,b)$. *Then* λ *is Frechet differentiable at* $1/p_k$ *in* $L^R(a,b)$ *and its Frechet derivative is given by*

$$d\lambda_{1/p_k}(h) = -\int_a^b |p_k u^{(k)}|^2\, h, \quad h \in L^R(a,b). \tag{4.18}$$

3. *Fix all components of* $\boldsymbol{\omega}$ *except* w *and consider* $\lambda = \lambda(w)$ *as a function of* $w \in L^R(a,b)$, $w > 0$ *a.e.. Thus* $\lambda(w)$ *is a (nonlinear) functional on the positive cone* V *in* $L^R(a,b)$ *of such* w. *Assume that* $\lambda(w)$ *has constant multiplicity in a neighborhood* $\mathcal{N}$ *of* w *in* V. *Then* λ *is Frechet differentiable at* w *in* V *and its Frechet derivative is given by*

$$d\lambda_w(h) = -\lambda(w) \int_a^b |u|^2\, h, \quad h \in L^R(a,b). \tag{4.19}$$

Theorem 4.12. *Let the assumptions and the notation of Theorem 4.2 hold.*

1. *Fix all components of* $\boldsymbol{\omega}$ *except* b *and consider* $\lambda = \lambda(b)$ *as a function of* b *for* $b \in (a,b')$. *Thus* $\lambda(b)$ *is a function on the interval* (a,b'). *Assume that* $\lambda(b)$ *has constant multiplicity in an interval* $\mathcal{N} \subset (\dashv, \lfloor')$. *Then* λ *is differentiable a.e. in* $\mathcal{N}$ *and*

$$\lambda'(b) = \sum_{j=0}^{n-1} (-1)^{k+j-1} u^{[j]}(b) \left(\overline{u}^{[n-j-1]}\right)'(b), \quad \text{a.e. for } b \in \mathcal{N}. \tag{4.20}$$

In particular, if $n = 2\,(k=1)$, *(4.20) reduces to*

$$\lambda'(b) = -\frac{1}{p_1(b)} |(p_1 u')(b)|^2 + |u(b)|^2 [p_0(b) - \lambda(b) w(b)], \quad \text{a.e. in } \mathcal{N}; \tag{4.21}$$

if $n = 4\,(k=2)$, *(4.20) reduces to*

$$\begin{aligned} \lambda'(b) &= -\frac{1}{p_2(b)} |(p_2 u'')(b)|^2 - |u(b)|^2 [p_0(b) + \lambda(b) w(b)] \\ &\quad -p_1(b)|u'(b)|^2 + 2Re\,(u'(b)(p_2 \bar{u}'')'(b)) \quad \text{a.e. in } \mathcal{N}. \end{aligned} \tag{4.22}$$

2. *Fix all components of $\boldsymbol{\omega}$ except a and consider $\lambda = \lambda(a)$ as a function of a for $a \in (a', b)$. Thus $\lambda(a)$ is a function on (a', b). Assume that $\lambda(a)$ has constant multiplicity in an interval $\mathcal{N} \subset (\dashv', \lfloor)$. Then λ is differentiable a.e. in $\mathcal{N}$ and*

$$\lambda'(a) = -\sum_{j=0}^{n-1} (-1)^{k+j-1} u^{[j]}(a) \left(\overline{u}^{[n-j-1]}\right)'(a) \quad \text{a.e. for } a \in \mathcal{N}. \tag{4.23}$$

In particular, if $n = 2\,(k = 1)$, (4.23) reduces to

$$\lambda'(a) = \frac{1}{p_1(a)} |(p_1 u')(a)|^2 - |u(a)|^2 [p_0(a) - \lambda(a) w(a)], \quad \text{a.e. in } \mathcal{N}; \tag{4.24}$$

if $n = 4\,(k = 2)$, (4.23) reduces to

$$\begin{aligned} \lambda'(a) \;=\; & \frac{1}{p_2(a)} |(p_2 u'')(a)|^2 + |u(a)|^2 [p_0(a) + \lambda(a) w(a)] \\ & + p_1(a) |u'(a)|^2 - 2Re\,(u'(a)(p\overline{u}'')'(a)) \quad \text{a.e. in } \mathcal{N}. \end{aligned} \tag{4.25}$$

Furthermore, all the equalities (4.20)-(4.22) hold at a point $b \in \mathcal{N}$ if $p_1, \ldots, p_k$ are continuous at b and $p_k(b) \neq 0$. Similarly, the equalities (4.23) - (4.25) hold at a point $a \in \mathcal{N}$ if $p_1, \ldots, p_k$ are continuous at a and $p_k(a) \neq 0$.

Theorem 4.13. *Let the assumptions and the notation of Theorem 4.2 hold, and let A and B be nonsingular matrices in $M_n(\mathbb{C})$.*

1. *Fix all components of $\boldsymbol{\omega}$ except A and consider $\lambda = \lambda(A)$ as a function of A. Assume that $\lambda(A)$ has constant multiplicity in a neighborhood $\mathcal{N}$ of A.*

   Then λ is differentiable at A and

$$d\lambda_A(H) = (-1)^{k-1} U^*(a) E A^{-1} H U(a), \;\; H \in M_n(\mathbb{C}). \tag{4.26}$$

2. *Fix all components of $\boldsymbol{\omega}$ except B and consider $\lambda = \lambda(B)$ as a function of B. Assume that $\lambda(B)$ has constant multiplicity in a neighborhood $\mathcal{N}$ of B.*

   Then λ is differentiable at B and

$$d\lambda_B(H) = (-1)^{k} U^*(b) E B^{-1} H U(b), \;\; H \in M_n(\mathbb{C}). \tag{4.27}$$

Remark 4.14. By Theorem 3.3, if λ is simple at $\boldsymbol{\omega}_0$, then λ is simple in a neighborhood $\mathcal{N}$ of $\boldsymbol{\omega}_0$ in $\boldsymbol{\Omega}_1$. This implies that the constant multiplicity assumption in Theorems 4.2-4.4 is automatically satisfied if λ is simple at $\boldsymbol{\omega}_0$.

The following lemmas are needed to prove the above theorems.

Lemma 4.15. *Let $n = 2k$. Assume that $P = \langle p_0, p_1, \ldots, p_k\rangle$, $Q = \langle q_0, q_1, \ldots, q_k\rangle \in NF_k$. Then for any $y \in \mathcal{D}(\mathcal{P})$ and $z \in \mathcal{D}(\mathcal{Q})$ we have*

$$(4.28) \int_a^b (y\mathcal{M}_{\mathcal{Q}}\not\ddagger - \not\ddagger\mathcal{M}_{\mathcal{P}}\dagger) = (-1)^k[y,z]_a^b + \int_a^b \sum_{j=0}^{k-1}(-1)^{k+j-1}(q_j - p_j)y^{(j)}\bar{z}^{(j)} + \int_a^b (q_k - p_k)y^{(k)}\bar{z}^{(k)}$$

where

$$(4.29) \qquad [y,z] = Z^*EY = \sum_{j=0}^{n-1}(-1)^j y^{[j]}\bar{z}^{[n-j-1]}$$

for Y, Z defined by (2.6) according to y and P, z and Q, respectively, and

$$(4.30) \qquad [y,z]_a^b = [y,z](b) - [y,z](a).$$

In particular, if $P = Q$, then $\mathcal{M}_{\checkmark} = \mathcal{M}_{\mathcal{Q}} = \mathcal{M}$, and

$$(4.31) \qquad \int_a^b (y\mathcal{M}\not\ddagger - \not\ddagger\mathcal{M}\dagger) = (-\infty)^{\|}[\dagger,\ddagger]_{\dashv}^{\llcorner}, \quad \dagger,\ddagger \in \mathcal{D}(\mathcal{P}).$$

Note that equality (4.31) is Green's identity. Thus (4.28)- (4.30) can be viewed as an extension of Green's identity.

P r o o f. The proof follows from repeated integration by parts. □

Lemma 4.16. *Assume that $P = \langle p_0, p_1, \ldots, p_k\rangle \in NF_k$. Let $u, v \in \mathcal{D}(\mathcal{P})$ and both satisfy the self-adjoint BC (3.3), (4.2). Then $[u,v]_a^b = 0$.*

P r o o f. This is the well known self-adjointness characterization, see [15], Theorem 1, p.73. □

Lemma 4.17. *Assume $P = \langle p_0, p_1, \ldots, p_k\rangle \in NF_k$. Then for any $\alpha_j, \beta_j \in \mathbb{C}$, $j = 0, \ldots, n-1$, there exists $u \in \mathcal{D}(\mathcal{P})$, such that*

$$u^{[j]}(a) = \alpha_j, \quad u^{[j]}(b) = \beta_j, \quad j = 0, \ldots, n-1.$$

P r o o f. This is the "Naimark Lemma" [15], p.63, Lemma 2. □

Corollary 4.18. *Assume that $P = \langle p_0, p_1, \ldots, p_k\rangle$ and $Q = \langle q_0, q_1, \ldots, q_k\rangle \in NF_k$. Let $u \in \mathcal{D}(\mathcal{P})$ and $v \in \mathcal{D}(\mathcal{Q})$, and both satisfy BC (3.3), (4.2). Then $[u,v]_a^b = 0$.*

Proof. Let U, V be defined by (2.6) according to u and P, v and Q, respectively. Then

$$[u,v]_a^b = [V^*EU]_a^b = (V^*EU)(b) - (V^*EU)(a).$$

By Lemma 4.3 there exists $u_1 \in \mathcal{D}(\mathcal{P})$ such that

$$u_1^{[j]}(a) = v^{[j]}(a), \quad u_1^{[j]}(b) = v^{[j]}(b), \quad j = 0, \dots, n-1.$$

Hence with $U_1 = (u_1^{[0]}, \dots, u_1^{[n-1]})^T$ it follows

$$U_1(a) = V(a), \quad U_1(b) = V(b),$$

and so u_1 satisfies (3.3), (4.2). By Lemma 4.2, $[u, u_1]_a^b = 0$. This implies that $[u,v]_a^b = 0$. □

Corollary 4.19. *Assume that $P = \langle p_0, p_1, \dots, p_k\rangle \in NF_k$. Let $u, v \in \mathcal{D}(\mathcal{P})$ such that (3.3), (4.2) hold for u on $[a,b] \subset J$ and for v on $[a,d] \subset J$, respectively. Then*

$$[u,v]_a^b = -[V^*(d) - V^*(b)]EU(b). \tag{4.32}$$

Proof.

$$[u,v]_a^b = (V^*EU)(b) - (V^*EU)(a)$$
$$= V^*(d)EU(b) - (V^*EU)(a) - (V^*(d) - V^*(b))EU(b). \tag{4.33}$$

By Lemma 4.3 there exists $u_1 \in \mathcal{D}(\mathcal{P})$ such that

$$u_1^{[j]}(a) = v^{[j]}(a), \quad u_1^{[j]}(b) = v^{[j]}(d), \quad j = 0, \dots, n-1.$$

Hence with $U_1 = (u_1^{[0]}, \dots, u_1^{[n-1]})^T$ it follows

$$U_1(a) = V(a), \quad U_1(b) = V(d),$$

and so u_1 satisfies (3.3), (4.2) on $[a,b]$. By Lemma 4.2, $[u, u_1]_a^b = 0$. This implies that

$$V^*(d)EU(b) - (V^*EU)(a) = 0,$$

and (4.32) then follows from (4.33). □

Lemma 4.20. *Assume $g \in L_{loc}(J)$. Then*

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} g = g(t) \quad \text{a.e. in } J. \tag{4.34}$$

In particular, (4.34) holds when g is continuous at t.

Proof. See Lemma 4.3 in [10], also see the Lebesgue Lemma. □

Proof of Theorem 4.2:

1. Let $u = u(\cdot, p_j)$ and $v = u(\cdot, p_j + h)$ be chosen such that $u^j(\cdot, p_j + h) \to u^j(\cdot, p_j)$ uniformly on $[a, b]$ as $h \to 0$ in $L^R(a, b)$. From (4.15) and Lemma 4.1

$$[\lambda(p_j + h) - \lambda(p_j)] \int_a^b u\bar{v}w = (-1)^k [u, v]_a^b + (-1)^{k+j-1} \int_a^b u^{(j)} \bar{v}^{(j)} h.$$

By Corollary 4.1 and Theorems 3.2, 3.4

$$[\lambda(p_j + h) - \lambda(p_j)](1 + o(1)) = (-1)^{k+j-1} \int_a^b |u^{(j)}|^2 h + o(h),$$

and consequently

$$\begin{aligned} \lambda(p_j + h) - \lambda(p_j) &= [(-1)^{k+j-1} \textstyle\int_a^b |u^{(j)}|^2 h + o(h)](1 + o(1))^{-1} \\ &= (-1)^{k+j-1} \textstyle\int_a^b |u^{(j)}|^2 h + o(h) \end{aligned}$$

as $h \to 0$ in $L^R(a, b)$. This completes the proof of (4.17).

2. Let $u = u(\cdot, 1/p_k)$ and $v = u(\cdot, 1/q_k)$ where $1/q_k = 1/p_k + h$, be chosen such that $u(\cdot, 1/q_k) \to u(\cdot, 1/p_k)$ as $h \to 0$ in $L^R(a, b)$. Then $1/p_k \in L^R(a, b)$ implies that $1/q_k \in L^R(a, b)$ and

$$p_k - q_k = p_k q_k h.$$

By (4.15) and Lemma 4.1

$$[\lambda(1/q_k) - \lambda(1/p_k)] \int_a^b u\bar{v}w = (-1)^k [u, v]_a^b + \int_a^b (q_k - p_k) u^{(k)} \bar{v}^{(k)} h.$$

By Corollary 4.1

$$[\lambda(1/q_k) - \lambda(1/p_k)] \int_a^b u\bar{v}w = -\int_a^b (p_k u^{(k)})(q_k \bar{v}^{(k)}) h.$$

The rest of the proof is similar to part 1 and hence is omitted.

3. The proof of (4.19) follows immediately from (4.15) and Lemmas 4.1 and 4.2. ∎

Proof of Theorem 4.3:

1. Let $u = u(\cdot, b)$ and $v = u(\cdot, b + h)$ be chosen such that $u(\cdot, b + h) \to u(\cdot, b)$ uniformly on $[a, b]$ as $h \to 0$ in $L^R(a, b)$. From (4.15) and Corollary 4.2, Lemma 4.1

$$\begin{aligned} &[\lambda(b + h) - \lambda(b)] \int_a^b u\bar{v}w = (-1)^k [u, v]_a^b \\ = \; &(-1)^{k-1} [U^*(b + h, b + h) - U^*(b, b + h)] \, E \, U(b, b) \\ = \; &(-1)^{k-1} [\int_b^{b+h} (U')^*(s, b + h) \, ds] \, E \, U(b, b). \end{aligned} \tag{4.35}$$

We claim that

$$\int_b^{b+h} (U')^*(s, b+h)\, ds = \int_b^{b+h} (U')^*(s, b)\, ds + o(h) \quad a.e. \tag{4.36}$$

In fact, from (3.1) - the vector form of (4.15) - we have that for $h > 0$

$$\begin{aligned} & \left| \frac{1}{h} \int_b^{b+h} [(U')^*(s, b+h) - (U')^*(s, b)]\, ds \right| \\ = & \left| \frac{1}{h} \int_b^{b+h} [U^*(s, b+h) - U^*(s, b)] P^*(s)\, ds \right. \\ & \left. + \frac{i^n}{h} \int_b^{b+h} [\lambda(b+h) U^*(s, b+h) - \lambda(b) U^*(s, b)] W^*(s)\, ds \right| \\ \le & \left(\frac{1}{h} \int_b^{b+h} |P^*(s)|\, ds \right) \sup_{s \in (b, b+h)} |U^*(s, b+h) - U^*(s, b)| \\ & + \left(\frac{1}{h} \int_b^{b+h} |W^*(s)|\, ds \right) \sup_{s \in (b, b+h)} |\lambda(b+h) U^*(s, b+h) - \lambda(b) U^*(s, b)| \\ \to & \ 0 \quad a.e. \text{ as } h \to 0 \end{aligned}$$

by Lemma 4.4 and the continuity of λ and U. The same is true for $h < 0$. This shows (4.36). Combining (4.35) and (4.36) we see that

$$\begin{aligned} \lambda(b+h) - \lambda(b) &= (-1)^{k-1} [\int_b^{b+h} (U')^*(s, b)\, ds + o(h)]\, E\, U(b, b) \\ &= (-1)^{k-1} h (U')^*(b, b)\, E\, U(b, b) + o(h) \quad a.e. \text{ in } (a, b'). \end{aligned} \tag{4.37}$$

Dividing both sides of (4.37) by h, taking limit as $h \to 0$, and using (4.29) we obtain that

$$\lambda'(b) = (-1)^{k-1} (U')^*(b, b)\, E\, U(b, b) = (-1)^{k-1} [u, u'](b), \quad a.e..$$

Then (4.20) follows from (4.29). (4.21) and (4.22) are the interpretations of (4.20) for the cases $n = 2$ and $n = 4$ with the expression (4.10) for quasi-derivatives.

2. The proof is similar and hence is omitted. ∎

Proof of Theorem 4.4: We only give a proof for Part 1. Let $u = u(\cdot, A)$ and $v = u(\cdot, A+H)$ be chosen such that $u(\cdot, A+H) \to u(\cdot, A)$ uniformly on $[a, b]$ as $H \to 0$ in $M_n(\mathbb{C})$. By Lemma 4.1, (4.15) and (3.3)

$$\begin{aligned} [\lambda(A+H) - \lambda(A)] \textstyle\int_a^b u \bar{v} w &= (-1)^k [u, v]_a^b = (-1)^k [V^* E U]_a^b \\ &= (-1)^k [(V^* E U)(b) - (V^* E U)(a)] \\ &= (-1)^k [V^*(a)(A+H)^* B^{*-1} E B^{-1} A U(a) - V^*(a) E U(a)]. \end{aligned} \tag{4.38}$$

By (4.2) we see that

$$B^{*-1}EB^{-1} = A^{*-1}EA^{-1}$$

and then

$$A^*B^{*-1}EB^{-1}A = E. \tag{4.39}$$

Combining (4.38) and (4.39), and noting that $\int_a^b u\bar{v}w \to 1$ as $H \to 0$ in $M_n(\mathbb{C})$, we obtain

$$\lambda(A+H) - \lambda(A) = (-1)^k U^*(a)H^*A^{*-1}EU(a) + o(H).$$

Since the eigenvalues of selfadjoint problems are real, this is equivalent to say that

$$\lambda(A+H) - \lambda(A) = (-1)^{k-1}U^*(a)EA^{-1}HU(a) + o(H)$$

which completes the proof of (4.26). ▮

5. Multiplicity of Eigenvalues

In this section we study the multiplicity of the eigenvalues as a function of the coefficients and the weight function of the problem.

Theorem 5.21. *Let $n = 2k$ and $\boldsymbol{\omega} = (a, b, A, B, P, w) \in \boldsymbol{\Omega}_1$ with $P = \langle p_0, p_1, \ldots, p_k\rangle$. Let q be p_0 or p_1. Fix all components of $\boldsymbol{\omega}$ except q, and consider $\lambda = \lambda(q)$ as a function of $q \in L^R(a, b)$. Define subsets Q_m of $L^R(a, b)$ by*

$$Q_m = \{q \in L^R(a, b) : \ \lambda(q) \text{ has multiplicity } m\}, \quad m = 1, 2, \ldots, n. \tag{5.1}$$

Then

(i) $\bigcup_{l=1}^m Q_l$ is an open set in the space $L^R(a, b)$ for $m = 1, 2, \ldots, n$.

(ii) $\bigcup_{l=2}^n Q_l$ is a closed and nowhere dense set in the space $L^R(a, b)$.

Proof. (i) By Theorem 3.3 we see that $\bigcup_{l=1}^m Q_l$ is open in $L^R(a, b)$ for $m = 1, 2, \ldots, n$.

(ii) We now show that for any $m \in \{2, \ldots, n\}$, Q_m does not contain any non-empty open sets in $L^R(a, b)$. If the claim is not true, there exist $m \in \{2, \ldots, n\}$, $q_0 \in Q_m$, and a neighborhood $\mathcal{N}$ of q_0 in $L^R(a, b)$ such that $\mathcal{N} \subset \mathcal{Q}_{\Updownarrow}$. First, we discuss the case where $q = p_0$. Since $\lambda(q)$ has multiplicity m for all $q \in \mathcal{N}$, there exist m linearly independent normalized eigenfunctions $u_l(\cdot, q)$, $l = 1, \ldots, m$, corresponding to $\lambda(q)$.

The eigenfunctions can be chosen in such a way that at least two of them, say $u_1(\cdot, q)$ and $u_2(\cdot, q)$, do not satisfy that

$$|u_1(t,q)| = |u_2(t,q)| \quad \text{for all } t \in (a,b). \tag{5.2}$$

In fact, if (5.2) holds, then $u_1 = re^{i\alpha}$ and $u_2 = re^{i\beta}$ with $r \geq 0$ and $\alpha, \beta \in AC^R_{loc}(a,b)$. Since u_1 and u_2 are linearly independent, $\alpha - \beta$ is not identically constant. Let $v = u_1 - u_2$. Then $v \neq 0$, and v is an eigenfunction of $\lambda(q)$. Also, $|v| = r|e^{i(\alpha-\beta)} - 1|$, which is not a constant multiple of r. Hence, for the normalized eigenfunction u corresponding to v, we have that $|u(t,q)| \not\equiv |u_1(t,q)|$ for $t \in (a,b)$. By Theorem 4.2, (4.17) with $j = 0$ holds for $u = u_l(\cdot, q)$, $l = 1, \ldots, m$. This contradicts the uniqueness of the Frechet derivative of a Frechet differentiable function. Therefore, the claim is true when $q = p_0$. Next, we treat the case where $q = p_1$. If there is a constant eigenfunction corresponding to $\lambda(q)$, then by (4.17) with $j = 1$,

$$d\lambda_q(h) = 0 \quad \text{for all } h \in L^R(a,b),$$

and hence every eigenfunction corresponding to $\lambda(q)$ is a constant. This is impossible since $m \geq 2$. Thus, there are no constant eigenfunctions corresponding to $\lambda(q)$ and hence there exist m normalized eigenfunctions $u_l(\cdot, q)$, $l = 1, \ldots, m$, corresponding to $\lambda(q)$ such that $u_l'(\cdot, q)$, $l = 1, \ldots, m$, are linearly independent. Then, as in the previous case, we can assume that there does not hold

$$|u_1'(t,q)| = |u_2'(t,q)| \quad \text{for all } t \in (a,b),$$

and obtain a contradiction from (4.17) with $j = 1$. Therefore, the claim is also true when $q = p_1$.

For each $m \in \{1, \ldots, n-1\}$, $\bigcup_{l=m+1}^n Q_l$ is closed since $\bigcup_{l=1}^n Q_l = L^R(a,b)$ and $\bigcup_{l=1}^m Q_l$ is open. Now we show by induction that for each $m \in \{1, \ldots, n-1\}$, $\bigcup_{l=m+1}^n Q_l$ does not contain any non-empty open sets in $L^R(a,b)$. From the last paragraph, Q_n does not. Fix $m \in \{2, \ldots, n-1\}$ and assume that $\bigcup_{l=m+1}^n Q_l$ does not contain any non-empty open sets in $L^R(a,b)$. Let $\mathcal{N} \subset \bigcup_{\updownarrow=\Updownarrow}^{\backslash} \mathcal{Q}_{\updownarrow}$ be an open set of $L^R(a,b)$. Then

$$\mathcal{N} \bigcap \mathcal{Q}_{\Updownarrow} = \mathcal{N} \bigcap \left(\mathcal{L}^{\mathcal{R}}(\dashv, \lfloor) \backslash \bigcup_{\updownarrow=\Updownarrow+\infty}^{\backslash} \mathcal{Q}_{\updownarrow} \right)$$

is open and hence empty since Q_m does not contain any non-empty open sets of $L^R(a,b)$. So, $\mathcal{N}$ must be in $\bigcup_{l=m+1}^n Q_l$ and hence empty. Thus, $\bigcup_{l=m}^n Q_l$ does not contain any non-empty open sets. Therefore, in particular, $\bigcup_{l=2}^{n} Q_l$ is closed and nowhere dense. □

Let $j \geq 2$. Then it is easy to see that when we change p_j, the space

$$\{\text{polynomials } u \text{ in the eigenspace corresponding to } \lambda : \text{ degree of } u \leq j-1\}$$

stays the same. Let n_j be the larger of this dimension and the integer 1. Note that $n_j \leq j$ and that the multiplicity of λ is at least n_j for each p_j.

Theorem 5.22. *Let $n = 2k$ with $k \geq 2$, $\boldsymbol{\omega} = (a, b, A, B, P, w) \in \mathbf{\Omega}_1$ with $P = \langle p_0, p_1, \ldots, p_k \rangle$, and $j \in \{2, \cdots, k\}$. Fix all components of $\boldsymbol{\omega}$ except p_j, and consider $\lambda = \lambda(p_j)$ as a function of $p_j \in L^R(a, b)$. Define subsets Q_m of $L^R(a, b)$ by*

$$Q_m = \{p_j \in L^R(a, b) : \ \lambda(p_j) \text{ has multiplicity } m\}, \quad m = n_j, \ldots, n. \tag{5.3}$$

Then

(i) $\bigcup_{l=n_j}^{m} Q_l$ is an open set in the space $L^R(a, b)$ for $m = n_j, \ldots, n$.

(ii) $\bigcup_{l=n_j+1}^{n} Q_l$ is a closed and nowhere dense set in the space $L^R(a, b)$.

The proof of Theorem 5.2 is similar to that of Theorem 5.1, so we omit it. However, we give an example to indicate that it is necessary to separate Theorems 5.1 and 5.2.

Example. Let

$$A = \begin{pmatrix} 1 & b-a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & b-a \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and consider the BVP consisting of (4.15) with $k = 2$ and the BC

$$Y(b) = AY(a).$$

It is easy to see that A and $-I$ satisfy (4.2), the self-adjointness condition. For each fixed $w \in L^R(a, b)$ and any $\lambda_0 \in \mathbb{R}$, when

$$p_0 = -\lambda_0 w \qquad \text{and} \qquad p_1' = 0,$$

the BVP always has an eigenvalue λ_0 with multiplicity at least 2: $y(t) = 1$ and $y(t) = t$ are always two linearly independent eigenfunctions. So, $n_2 = 2$ in this case.

Theorem 5.23. *Let* $n = 2k$ *and* $\boldsymbol{\omega} = (a, b, A, B, P, w) \in \boldsymbol{\Omega}_1$ *with* $P = \langle p_0, p_1, \ldots, p_k \rangle$. *Fix all components of* $\boldsymbol{\omega}$ *except* w, *and consider* $\lambda = \lambda(w)$ *as a function of* $w \in L^R(a,b)$. *Assume that the zero set of* $\lambda(w)$ *is nonwhere dense in* $L^R(a,b)$. *Define subsets* Q_m *of* $L^R(a,b)$ *by*

$$Q_m = \{w \in L^R(a,b) : \ \lambda(w) \text{ has multiplicity } m\}, \quad m = 1, 2, \ldots, n. \tag{5.4}$$

Then

(i) $\bigcup_{l=1}^{m} Q_l$ is an open set in the space $L^R(a,b)$ for $m = 1, 2, \ldots, n$.

(ii) $\bigcup_{l=2}^{n} Q_l$ is a closed and nowhere dense set in the space $L^R(a,b)$.

Remark 5.24. It is interesting to note that a result analogous to Theorems 5.1 - 5.3 for the endpoints a, b does not hold. It is easy to construct examples, e.g., for the Fourier equation $-y'' = \lambda y$ such that $\lambda(b)$ has multiplicity 2 for all b, and (4.20), (4.23) hold for any normalized eigenfunctions u.

6. Eigenvalue Comparisons

The prevailing method for obtaining inequalities among eigenvalues of related problems is the variational method based on the min-max characterization of the eigenvalues. Here we obtain such inequalities based on methods similar to those used to establish the differentiability results in section 4. Thus our approach is more direct and more elementary than those based on the variational characterization and yet yields much more general results.

Consider the equations

$$(-1)^k \left\{ \left\{ \cdots \left\{ [(p_k y^{(k)})' - p_{k-1} y^{(k-1)}]' - p_{k-2} y^{(k-2)} \right\}' - \cdots - p_1 y' \right\}' - p_0 y \right\} = \lambda w y, \tag{6.1}$$

$$(-1)^k \left\{ \left\{ \cdots \left\{ [(q_k y^{(k)})' - q_{k-1} y^{(k-1)}]' - q_{k-2} y^{(k-2)} \right\}' - \cdots - q_1 y' \right\}' - q_0 y \right\} = \mu v y. \tag{6.2}$$

We have the following

Theorem 6.25. *Let* $n = 2k$. *Let* $P = \langle p_0, p_1, \ldots, p_k \rangle$ *and* $Q = \langle q_0, q_1, \ldots, q_k \rangle$ *be in* $NF_k(J)$, *let* $\boldsymbol{\omega}_1 = \boldsymbol{\omega}(a, b, A, B, P, w)$ *and* $\boldsymbol{\omega}_2 = \boldsymbol{\omega}(a, b, A, B, Q, v)$ *be in* $\boldsymbol{\Omega}_1$. *Denote by* $\lambda(\boldsymbol{\omega})$ *a fixed continuous eigenvalue branch which is defined at* $\boldsymbol{\omega}_1$ *and* $\boldsymbol{\omega}_2$. *Let*

$\lambda = \lambda(\boldsymbol{\omega}_1)$ and $\mu = \lambda(\boldsymbol{\omega}_2)$ be the corresponding eigenvalues of the BVP (6.1), (3.3) associated with $\boldsymbol{\omega}_1$ and the BVP (6.2), (3.3) associated with $\boldsymbol{\omega}_2$, respectively. Further assume that the following inequalities hold a.e. on (a, b) :

(i) $0 < 1/q_k \le 1/p_k$, *or* $1/q_k \le 1/p_k < 0$;

(ii) $(-1)^{k+j-1} p_j \le (-1)^{k+j-1} q_j,\ j = 0, 1, \dots, k-1$;

(iii) $w \ge v$ *in case that* $\lambda \ge 0$ *and* $\mu \ge 0$, $w \le v$ *in case that* $\lambda \le 0$ *and* $\mu \le 0$.

Then $\lambda \le \mu$. Furthermore, if strict inequality holds on a subset of the interval (a, b) with positive Lebesgue measure in (ii) for some $j = 0, 1, \dots, k-1$ or in (i) or in (iii), then $\lambda < \mu$.

Proof. We give the proof for the case $p_j = q_j,\ j = 1, \dots, k,\ w = v$, and $(-1)^{k-1} p_0 \le (-1)^{k-1} q_0$ only, since the proofs of the other cases are similar. To simplify the notation, we omit the subscripts and let $p = p_0,\ q = q_0$. It suffices to give a proof for the case that $(-1)^{k-1} p < (-1)^{k-1} q$ on (a, b), since the general case $(-1)^{k-1} p \le (-1)^{k-1} q$ then follows from a limit argument.

Assume $(-1)^{k-1} p(t) < (-1)^{k-1} q(t),\ t \in (a, b)$, define

$$p_s(t) = (1 - s)p(t) + sq(t),\ t \in (a, b),\ s \in [0, 1],$$

and denote by $\lambda(s) = \lambda(p_s)$, the eigenvalue of the BVP (4.15), (3.3) with p_0 replaced by p_s. Then $\lambda(0) = \lambda(p)$ and $\lambda(1) = \lambda(q)$. By assumption, $\lambda(s)$ is continuous on $[0, 1]$.

We will show that the Dini-derivative $D^+\lambda(s) > 0$ for all $s \in (0, 1)$. Then by a well-known theorem, see [16], Proposition 2, p.99, it follows that $\lambda(s)$ is increasing and hence $\lambda(0) < \lambda(1)$, i.e., $\lambda < \mu$.

For any $s \in (0, 1)$, let $S(s)$ be the l_s-dimensional unit sphere in $\mathbb{C}^n$ with $l_s \le n$ on which every point is an initial point of a vector-eigenfunction $U(\cdot, s)$ at a associated with $\lambda(s)$. Let $s_0 \in (0, 1)$. Then it is easy to see that there exists a sequence $\{s_k\} \subset (0, 1)$ and $s_k \to s_0$, such that $S(s_k),\ k = 1, 2, \dots$, all have the same dimension, say l, and

$$S(s_k) \to \tilde{S}$$

where $\tilde{S}$ is an l-dimensional unit sphere in $\mathbb{C}^n$ and $\tilde{S} \subset S(s_0)$. This implies that there exists at least one (actually l linearly independent) normalized eigenfunction(s) $u(\cdot, s)$ of $\lambda(s)$ such that the associated vector functions satisfy $U(\cdot, s_k) \to U(\cdot, s_0)$. With a similar argument to the proof of Theorem 4.2, part (1) (and noting that we do not need to use the hypothesis of constant multiplicity used in the proof of Theorem 4.2) we have

$$\lambda(s_k) - \lambda(s_0) = (-1)^{k-1}(s_k - s_0) \int_a^b |u(\sigma, s_0)|^2 (q(\sigma) - p(\sigma))\, d\sigma + o(s_k - s_0)$$

and hence

$$\lim_{k\to\infty} \frac{\lambda(s_k) - \lambda(s_0)}{s_k - s_0} = (-1)^{k-1} \int_a^b |u(\sigma, s_0)|^2 (q(\sigma) - p(\sigma))\, d\sigma \;>\; 0. \tag{6.3}$$

Therefore, $D^+\lambda(s_0) > 0$. This completes the proof. □

Remark 6.26. It is interesting that, in the case when the multiplicity of $\lambda(s_0)$ is greater than 1, the limit (6.3) is independent of the choice of the normalized eigenfunction u. Note however that, although the limit (6.3) exists for $h = q(\sigma) - p(\sigma)$, it may not exist for general $h \in L^R(a, b)$ since λ may not be differentiable at s_0; recall that we have not assumed, as in Theorem 4.2, that the multiplicity is constant in a neighborhood. This is why we need the Dini derivative in the proof of Theorem 6.1.

7. Comment with an Example

We give an example to show that the hypothesis that $\lambda(\boldsymbol{\omega})$ has constant multiplicity cannot be omitted in Theorems 4.2-4.4.

Consider the equation

$$-y'' = \lambda y \text{ on } [0, 1] \tag{7.1}$$

together with the BC

$$Y(1) = A\, Y(0), \tag{7.2}$$

where $A \in M_2(\mathbb{R})$ satisfies $\det A = 1$. It is easy to see that A and $-I$ satisfy (4.2), the self-adjointness condition.

Fact 1. *The BVP consisting of (7.1) and (7.2) has a positive eigenvalue of multiplicity 2 if and only if*

$$A = \alpha_r(s) := \begin{pmatrix} \cos s & \frac{1}{2r\pi + s} \sin s \\ -(2r\pi + s)\sin s & \cos s \end{pmatrix}$$

for some positive integer r and number $s \in (-2\pi, 0]$, with $(2r\pi + s)^2$ being the eigenvalue.

P r o o f. When $\lambda > 0$, the general solution of (7.1) is

$$y(t) = C\cos(\sqrt{\lambda}\,t) + D\sin(\sqrt{\lambda}\,t),$$

and then (7.2) can be rewritten as

$$\begin{cases} (\cos\sqrt{\lambda} - a)C + (\sin\sqrt{\lambda} - b\sqrt{\lambda}\,)D = 0, \\ -(\sqrt{\lambda}\,\sin\sqrt{\lambda} + c)C + \sqrt{\lambda}\,(\cos\sqrt{\lambda} - d)D = 0, \end{cases}$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

So, the BVP has a positive eigenvalue λ of multiplicity 2 if and only if

$$a = \cos\sqrt{\lambda}, \quad b = \frac{1}{\sqrt{\lambda}}\sin\sqrt{\lambda}, \quad c = -\sqrt{\lambda}\,\sin\sqrt{\lambda}, \quad d = \cos\sqrt{\lambda},$$

which imply our claims. □

Fix a positive integer r. By Theorem 3.2, there is a neighborhood $\mathcal{N}$ of the identity matrix I in the 3-dimensional space (actually a classical Lie group)

$$\{A \in M_2(\mathbb{R}) : \quad \det A = 1\}$$

such that for each $A \in \mathcal{N}$, the BVP has only either one eigenvalue of multiplicity 2 or two simple eigenvalues in $\big((2r\pi - \pi)^2,\ (2r\pi + \pi)^2\big)$. Let $\alpha_r(-\pi, \pi) = \{\alpha_r(s) : s \in (-\pi, \pi)\}$. Since $\mathcal{N} \setminus \alpha_\nabla(-\pi,\, \pi)$ is connected and there the BVP has only two simple eigenvalues in $\big((2r\pi - \pi)^2,\ (2r\pi + \pi)^2\big)$, we have exactly two continuous eigenvalue functions $\lambda_{r,u}$ and $\lambda_{r,l}$ on $\mathcal{N}$ defined by

$$\lambda_{r,u}(A) = \begin{cases} \text{the bigger eigenvalue in } \big((2r\pi - \pi)^2,\ (2r\pi + \pi)^2\big) & \text{if } A \in \mathcal{N} \setminus \alpha_\nabla(-\pi,\, \pi), \\ (2r\pi + s)^2 & \text{if } A = \alpha_r(s) \in \mathcal{N}, \end{cases}$$

$$\lambda_{r,l}(A) = \begin{cases} \text{the smaller eigenvalue in } \big((2r\pi - \pi)^2,\ (2r\pi + \pi)^2\big) & \text{if } A \in \mathcal{N} \setminus \alpha_\nabla(-\pi,\, \pi), \\ (2r\pi + s)^2 & \text{if } A = \alpha_r(s) \in \mathcal{N}. \end{cases}$$

For real s, set

$$\gamma(s) = \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix}.$$

Fact 2. *For $|s|$ sufficiently small we have*

$$\text{(7.3)} \quad \sqrt{\lambda_{r,u}(\gamma(s))} = 2r\pi + \arcsin\tanh|s|, \qquad \sqrt{\lambda_{r,l}(\gamma(s))} = 2r\pi - \arcsin\tanh|s|.$$

Proof. When $A = \gamma(s)$, the positive eigenvalues λ are characterized by the equation

$$(\cos\sqrt{\lambda} - e^s)\cdot\sqrt{\lambda}\,(\cos\sqrt{\lambda} - e^{-s}) + \sin\sqrt{\lambda}\cdot\sqrt{\lambda}\,\sin\sqrt{\lambda} = 0,$$

and hence also by

$$2 - (e^s + e^{-s})\cos\sqrt{\lambda} = 0.$$

The eigenvalues in $\big((2r\pi-\pi)^2,\ (2r\pi+\pi)^2\big)$ are

$$\left(2r\pi \pm \arccos\frac{2}{e^s + e^{-s}}\right)^2,$$

which implies (7.3). □

Fact 3. *$\lambda_{r,u}$ and $\lambda_{r,l}$ are not Frechet differentiable at $A = I$.*

Proof. By (7.3), direct calculations yield the one-sided derivatives

$$\big(\lambda_{r,u}\circ\gamma\big)'(0^+) = 4r\pi, \quad \big(\lambda_{r,u}\circ\gamma\big)'(0^-) = -4r\pi,$$

$$\big(\lambda_{r,l}\circ\gamma\big)'(0^+) = -4r\pi, \quad \big(\lambda_{r,l}\circ\gamma\big)'(0^-) = 4r\pi,$$

which imply our claim. □

Acknowledgment. The authors are indebted to the referee for a number of suggestions which have significantly improved the content and the exposition of the paper. In particular, the proof of Theorem 3.2 has been shortened, and the present proofs of Theorems 3.3 and 3.4 are due to the referee which are shorter and more direct than the original ones. We also thank the referee for finding a gap in the proof of the original Theorem 5.1, which has been divided into three theorems.

Department of Mathematics
Northern Illinois University
DeKalb, IL 60115
U.S.A.

Department of Mathematics
Northern Illinois University
DeKalb, IL 60115
U.S.A.

Department of Mathematics
Northern Illinois University
DeKalb, IL 60115
U.S.A.

References

[1] P. B. Bailey, W. N. Everitt, and A. Zettl. Sleign2: An eigenfunction- eigenvalue code for singular Sturm-Liouville problems. *pre-print.*

[2] P. B. Bailey, M. K. Gordon, and L. F. Shampine. Automatic solution ot the Sturm-Liouville problem. *ACM TOMS 4(1978), 193-208.*

[3] J. Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York/London, 1969.

[4] N. Dunford and J. T. Schwartz. *Linear operators , vol.II.* Interscience, New York, 1963.

[5] W. N. Everitt and D. Race. Some remarks on linear ordinary quasi-differential expressions. *Proc. London Math. Soc. v.3, no.54(1987), 300-320.*

[6] W. N. Everitt and A. Zettl. Generalized symmetric ordinary differential expressions, I. The basic theory. *Nieuw Archief. voor Wiskunde, (3), XXVII(1979), 363-397*, 1979.

[7] H. Frentzen. Equivalents, adjoints and symmetry of quasi-differential expressions with matrix valued coefficients and polynomials in them. *Proc. Roy. Soc. Edinburgh A 92 (1982), 123-146.*

[8] C.T. Fulton and S. Pruess. Mathematical software for Sturm-Liouville problems. *ACM. Trans. Math. Software 19 (1993), 360-376.*

[9] L. Greenberg and M. Marletta. The code SLEUTH for solving fourth order Sturm-Liouville problems. *pre-print.*

[10] Q. Kong and A. Zettl. Dependence of eigenvalues of Sturm-Liouville problems on the boundary. *J. Differential Equations, 126(1996), 389-407.*

[11] Q. Kong and A. Zettl. Eigenvalues of regular Sturm-Liouville problems. *J. Differential Equations, 131(1996), 1-19.*

[12] Q. Kong and A. Zettl. Linear ordinary differential equations. *WSSIAA v.3, (1994), (Special volume dedicated to W.Walter), Inequalities and Applications, edited by R. P. Agarwal, 381-397.*

[13] M. Moeller and A. Zettl. Semi-boundedness of ordinary differential operators. *J. Differential Equations, v.115, no.1, (1995), 24-49.*

[14] M. Moeller and A. Zettl. Weighted norm inequalities for quasi-derivatives. *Results in Mathematics, 24(1993), 153-160.*

[15] M. A. Naimark. *Linear Differential Operators.* Ungar, New York, 1968.

[16] H. L. Royden. *Real Analysis.* New York, Macmillan, London, 3rd edition, 1988.

[17] J. Weidmann. *Spectral Theory of Ordinary Differential Operators*, volume 1258 of *Lecture Notes in Mathematics 1258.* Springer Verlag, Berlin, 1987.

[18] A. Zettl. Adjoint and self-adjoint problems with interface conditions. *SIAM J. Applied Math. 16, (1968), 851-859.*

[19] A. Zettl. Formally self-adjoint quasi-differential operators. *Rocky Mountain J. Math. 5(1975), 453-474.*