

## ROOT VECTORS FOR GEOMETRICALLY SIMPLE TWO-PARAMETER EIGENVALUES

PAUL BINDING AND TOMAŽ KOŠIR

ABSTRACT. A class of two-parameter eigenvalue problems involving generally nonselfadjoint and unbounded operators is studied. A basis for the root subspace at a geometrically simple eigenvalue of Fredholm type is computed in terms of the underlying two-parameter system. Comparison with Faierman's work on two-parameter boundary value problems of Sturm-Liouville type is given as an application.

### 1. INTRODUCTION

We consider a pair of two-parameter equations of the form

$$(1.1) \quad W_i(\boldsymbol{\lambda})x_i = 0 \neq x_i,$$

where

$$(1.2) \quad W_i(\boldsymbol{\lambda}) = A_{i0}\lambda_0 + A_{i1}\lambda_1 + A_{i2},$$

$A_{ij}$  ( $j = 1, 2$ ) are bounded linear operators acting on Hilbert spaces  $H_i$  ( $i = 1, 2$ ) over the complex numbers,  $A_{i0}$  are closed densely defined with domain  $\mathcal{D}(A_{i0}) \subseteq H_i$ , and  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1) \in \mathbb{C}^2$ . Evidently,  $W_i(\boldsymbol{\lambda})$  is a linear operator with domain  $\mathcal{D}(A_{i0})$  for all such parameters  $\boldsymbol{\lambda}$ .

The spectral theory of such systems has been developed from various viewpoints; see [1, 8, 22, 24, 26] for books on the subject. In most of the literature, the  $A_{ij}$  are selfadjoint and satisfy definiteness conditions guaranteeing that the eigenvalues are simple in certain senses. A standard selfadjoint example is when (1.1) are Sturm-Liouville equations. One application of these ideas is to completeness and expansion theory of elliptic boundary value problems via separation of variables. For work on nonselfadjoint problems we refer to [5, 9, 11, 14, 15].

To introduce our analysis, we make the following constructions. The operators  $A_{ij}$ ,  $j = 1, 2$ , induce operators  $A_{ij}^\dagger$  on the Hilbert space tensor product  $H = H_1 \otimes H_2$  by means of

$$A_{1j}^\dagger(x_1 \otimes x_2) = A_{1j}x_1 \otimes x_2 \text{ and } A_{2j}^\dagger(x_1 \otimes x_2) = x_1 \otimes A_{2j}x_2$$

on decomposable tensors, extended by linearity and continuity to the whole of  $H$ . Similarly,  $A_{i0}$  induces an operator  $A_{i0}^\dagger$  with domain  $\mathcal{D}(A_{i0}^\dagger) \subset H$  (see [24, §2.3]).

---

Received by the editors June 15, 2001.

2000 *Mathematics Subject Classification*. Primary 35P10, 47A13; Secondary 35J55.

The first author's research was supported by NSERC of Canada.

The second author's research was supported by the Ministry of Science and Technology of Slovenia.

We denote by  $\mathcal{D}$  the intersection  $\mathcal{D}(A_{10}^\dagger) \cap \mathcal{D}(A_{20}^\dagger)$ , which is a dense subspace of  $H$ . We define the operator

$$(1.3) \quad \Delta_0 = \begin{vmatrix} A_{11}^\dagger & A_{12}^\dagger \\ A_{21}^\dagger & A_{22}^\dagger \end{vmatrix} = A_{11} \otimes A_{22} - A_{12} \otimes A_{21}$$

on  $H$  and operators

$$(1.4) \quad \Delta_1 = -A_{22}^\dagger A_{10}^\dagger + A_{12}^\dagger A_{20}^\dagger \text{ and } \Delta_2 = -A_{11}^\dagger A_{20}^\dagger + A_{21}^\dagger A_{10}^\dagger$$

on  $\mathcal{D}$ . In §2 we shall make assumptions guaranteeing the existence and commutativity of the operators  $\Gamma_j = \Delta_2^{-1} \Delta_j$ ,  $j = 0, 1$ , on  $\mathcal{D}$ .

The nullspace of an operator  $A$  is denoted by  $\mathcal{N}(A)$ . A pair  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1) \in \mathbb{C}^2$  is called an *eigenvalue of the pair of commuting linear transformations*  $\{\Gamma_0, \Gamma_1\}$  if

$$(1.5) \quad \mathcal{R}_1 = \mathcal{R}_1(\boldsymbol{\lambda}) = \mathcal{N}(\Gamma_0 - \lambda_0 I) \cap \mathcal{N}(\Gamma_1 - \lambda_1 I) \neq \{0\}.$$

The subspace  $\mathcal{R}_1$  is called the *geometric eigenspace* corresponding to  $\boldsymbol{\lambda}$ . The set of all the eigenvalues is called the *point spectrum* of  $\{\Gamma_0, \Gamma_1\}$ . We say that  $\boldsymbol{\lambda}$  is a *geometrically simple eigenvalue* of  $\{\Gamma_0, \Gamma_1\}$  if  $\dim \mathcal{R}_1 = 1$ . Under assumptions to be specified in §2, the eigenvalues  $\boldsymbol{\lambda}$  of (1.2) and that of its associated linear transformations  $\Gamma_0, \Gamma_1$  coincide and we have

$$(1.6) \quad \mathcal{R}_1(\boldsymbol{\lambda}) = \mathcal{N}(W_1(\boldsymbol{\lambda})) \otimes \mathcal{N}(W_2(\boldsymbol{\lambda})).$$

(See [5, §4].) We shall simply refer to such  $\boldsymbol{\lambda}$  as eigenvalues from now on.

We define the *m-th root subspace*

$$(1.7) \quad \mathcal{R}_m = \mathcal{R}_m(\boldsymbol{\lambda}) = \bigcap_{k=0}^m \mathcal{N}[(\Gamma_0 - \lambda_0 I)^k (\Gamma_1 - \lambda_1 I)^{m-k}],$$

for  $m \geq 2$ . The least integer  $M$ , if it exists, such that  $\mathcal{R}_M = \mathcal{R}_{M+1}$  is called the *ascent of*  $\boldsymbol{\lambda}$ . We call the subspace

$$\mathcal{R} = \mathcal{R}(\boldsymbol{\lambda}) = \bigcup_{m=1}^{\infty} \mathcal{R}_m$$

the *root subspace* of  $\mathbf{W} = (W_1, W_2)$  at  $\boldsymbol{\lambda}$ . Note that each subspace  $\mathcal{R}_m$  is an invariant subspace for both  $\Gamma_0$  and  $\Gamma_1$ . If the ascent of  $\boldsymbol{\lambda}$  is equal to 1, then  $\boldsymbol{\lambda}$  is called *semisimple*. If  $\boldsymbol{\lambda}$  is a geometrically simple eigenvalue and its ascent is equal to 1, then  $\mathcal{R} = \mathcal{R}_1$  is one-dimensional and  $\boldsymbol{\lambda}$  is called *algebraically simple*.

The elements of  $\mathcal{R}$  are called *root vectors* and it is our purpose to describe bases of root vectors for  $\mathcal{R}$  in terms of the  $A_{ij}$ . This problem was raised by Atkinson [1] and has been studied in certain definite selfadjoint cases by many authors. See [26] for a review of eigenvector completeness under various conditions including positive definiteness of  $\Delta_2$ . In these cases the eigenvalues are real (i.e.  $\boldsymbol{\lambda} \in \mathbb{R}^2$ ) and always  $\dim \mathcal{R} = \dim \mathcal{R}_1$ . Thus if an eigenvalue is geometrically simple, then it is also algebraically simple. Moreover, (1.6) shows that a basis for  $\mathcal{R} = \mathcal{R}_1$  can be constructed from the eigenvectors  $x_1 \otimes x_2$  for the  $\Gamma_j$ ,  $x_i$  as in (1.1).

Without definiteness, however, this may fail and various authors have tackled Atkinson's question. We cite [3] on real eigenvalues for selfadjoint problems and [8] on general eigenvalues for Sturm-Liouville (and therefore geometrically simple) problems. Faierman gives bases of  $\mathcal{R}_l$  for  $l = 2$  and 3 and gives a conjecture for general  $l$ . In an earlier paper [5] we gave a basis for  $\mathcal{R}_2$  for general nonselfadjoint

(including geometrically nonsimple) problems. A coalgebraic approach to this problem in [11] describes root subspaces  $\mathcal{R}_l$  for all  $l$  in terms of  $A_{ij}$ , but it is difficult to use it to construct bases explicitly. Here we shall give a general construction for a basis of  $\mathcal{R}$  for geometrically simple eigenvalues of problems that are not necessarily selfadjoint. We shall relate our construction in the Sturm-Liouville case to Faierman's work in §5.

Our plan is as follows. Our main assumptions are discussed in §2, followed by the construction of a basis for the root subspace in the so-called nonderogatory case. In §3, we lay foundations for the more difficult derogatory case. The main result (Theorem 4.1) is proved by induction on  $m$  (see (1.7)) in §4. §5 contains the specialization to the Sturm-Liouville case.

## 2. PRELIMINARIES

We begin with our assumptions on the  $A_{ij}$ . We make two regularity assumptions and a solvability assumption.

**Assumption I.** There exists  $\alpha \in \mathbb{C}$  such that the operator  $\Delta'_2 = \Delta_2 + \alpha\Delta_0 : \mathcal{D} \rightarrow H$  has a bounded inverse.

This assumption holds in many interesting cases. We refer to [5, p. 231] for a discussion.

To simplify the notation we assume, unless stated otherwise, that the operator  $\Delta_2$  has a bounded inverse. This form of Assumption I can be obtained by a shift in parameters.

Let  $\mathcal{A}_{ij}^\dagger$  denote the restriction of  $A_{ij}^\dagger$  to  $\mathcal{D}$ . The array

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{10}^\dagger & \mathcal{A}_{11}^\dagger & \mathcal{A}_{12}^\dagger \\ \mathcal{A}_{20}^\dagger & \mathcal{A}_{21}^\dagger & \mathcal{A}_{22}^\dagger \end{bmatrix}$$

then defines a linear map  $\mathcal{A} : \mathcal{D}^3 \rightarrow H^2$ . Here  $H^k$  is the direct sum of  $k$  copies of  $H$ . Omitting the  $j$ -th column we get a transformation  $\mathcal{A}_j$  acting on the (algebraic) direct sum  $\mathcal{D}^2$  for  $j = 0, 1, 2$ . Note that  $\Delta_j = (-1)^j \det \mathcal{A}_j$ . We denote by  $\mathcal{C}_j$  the  $j$ -th column of  $\mathcal{A}$ . Now we state the solvability assumption.

**Assumption II.** The equation  $\mathcal{A}_2 \mathbf{y} = \mathcal{C}_2 x$  has a solution  $\mathbf{y} \in \mathcal{D}^2$  for all  $x \in \mathcal{D}$ .

Again see [5, p. 236] for a discussion of this assumption.

The linear transformations  $\Gamma_j = \Delta_2^{-1} \Delta_j$ , ( $j = 0, 1$ ) act on  $\mathcal{D}$ . These are called the *associated linear transformations* of the two-parameter system (1.2). Under Assumption II it follows from [5, Thm. 3.2] that the linear transformations  $\Gamma_0$  and  $\Gamma_1$  commute on  $\mathcal{D}$ , i.e.,  $\Gamma_0 \Gamma_1 x = \Gamma_1 \Gamma_0 x$  for all  $x \in \mathcal{D}$ , and that

$$(2.1) \quad A_{i0}^\dagger \Gamma_0 x + A_{i1}^\dagger \Gamma_1 x + A_{i2}^\dagger x = 0$$

for  $i = 1, 2$  and all  $x \in \mathcal{D}$ . Relation (2.1) is a consequence of Assumption II that will be often used in the proofs.

Next we define the notions of eigenvalues and point spectra for a two-parameter system. A point  $\boldsymbol{\lambda} \in \mathbb{C}^2$  is called an *eigenvalue of the two-parameter system* (1.2) if both nullspaces  $\mathcal{N}(W_i(\boldsymbol{\lambda}))$ ,  $i = 1, 2$ , are nontrivial. The set of all eigenvalues is called the *point spectrum* of (1.2), and it is denoted by  $\sigma(\mathbf{W})$ . An eigenvalue  $\boldsymbol{\lambda} \in \sigma(\mathbf{W})$  is called *geometrically simple* if  $\dim \mathcal{N}(W_i(\boldsymbol{\lambda})) = 1$  for both  $i$ . Thus  $\boldsymbol{\lambda}$  is a geometrically simple eigenvalue of (1.2) if and only if it is a geometrically

simple eigenvalue for the pair of associated linear transformations  $(\Gamma_0, \Gamma_1)$  although perhaps not for either of  $\Gamma_j$  separately.

**Assumption III.**  $\lambda = (\lambda_0, \lambda_1)$  is a *geometrically simple eigenvalue of finite ascent*  $M$  and  $W_i(\lambda)$ ,  $i = 1, 2$ , are Fredholm of index 0 [25].

This assumption is satisfied, for example, in several cases arising from boundary value problems, e.g., of Sturm-Liouville type (see §5). Assumption III is in terms of the maps  $A_{ij}$  except for the assumption of finite ascent. This is defined in terms of the associated linear transformations. It follows from our results in §4 that an equivalent definition of the ascent is possible in terms of the  $A_{ij}$  – it is the least  $m$  such that  $\mathcal{N}(\mathcal{S}_m) = 0$ .

In view of Assumption III we have  $\dim \mathcal{N}(W_i(\lambda)) = \dim \mathcal{N}(W_i(\lambda)^*) = 1$  for both  $i$ ,  $\mathcal{R}_1 \subsetneq \mathcal{R}_2 \subsetneq \cdots \subsetneq \mathcal{R}_M$  and  $\mathcal{R}_{M+k} = \mathcal{R}_M$  for all  $k \geq 1$ .

Finally, we remark that Assumptions I–III of our present setup are stronger than Assumptions I–III of [5] or [21] in the case  $n = 2$ . Assumption I here and in [5, 21] is the same. Here and in [21] we use the same solvability assumption while in [5] a weaker solvability assumption is used. In [21] it is assumed further that  $\lambda$  is geometrically simple. In both [5, 21] it is assumed that the  $W_i(\lambda)$  are Fredholm operators not necessarily of index 0. The assumption of Fredholm index 0 here is made simply for ease of notation.

Next we state two-parameter versions of some results of [5, 18, 19, 20] on root subspaces at geometrically simple eigenvalues of commuting linear maps and of multiparameter systems. These versions are used in what follows.

We write

$$D_m = \dim \mathcal{R}_m \quad \text{and} \quad d_m = D_m - D_{m-1}$$

for  $m \geq 1$ . Here  $D_0 = 0$ .

**Lemma 2.1.** *If  $\lambda$  is a geometrically simple eigenvalue, then for each  $m \geq 2$  the subspace  $\mathcal{R}_m$  is finite-dimensional.*

*Proof.* The argument is similar to the one used in the proof of [5, Lemma 5.2]. It is by induction on  $m$ . Assume  $m \geq 2$ . The subspace  $\mathcal{R}_m$  is invariant for both  $\Gamma_j - \lambda_j I$ ,  $j = 0, 1$ . The range  $\mathcal{R}((\Gamma_j - \lambda_j I)|_{\mathcal{R}_m})$  is contained in  $\mathcal{R}_{m-1}$  and thus is finite-dimensional because of the inductive assumption. Each kernel  $\mathcal{N}((\Gamma_j - \lambda_j I)|_{\mathcal{R}_m})$  has finite codimension in  $\mathcal{R}_m$ , i.e., the orthogonal complement  $\mathcal{Q}_j$  of  $\mathcal{N}((\Gamma_j - \lambda_j I)|_{\mathcal{R}_m})$  in  $\mathcal{R}_m$  is finite-dimensional. Hence the linear span  $\mathcal{Q}$  of the  $\mathcal{Q}_j$ ,  $j = 0, 1$ , is finite-dimensional. Then  $\mathcal{R}_m = \mathcal{R}_{m-1} + \mathcal{Q}$  also has finite dimension.  $\square$

Since  $\lambda$  is geometrically simple, (1.6) implies that  $d_1 = 1$ . The restricted maps of  $\Gamma_0$  and  $\Gamma_1$  to  $\mathcal{R}_m$  commute, and so it follows that  $d_i \leq i$  for all  $i \geq 1$ . Suppose that

$$(2.2) \quad x_{i1} \in \mathcal{N}(W_i(\lambda)), \quad i = 1, 2,$$

are nonzero vectors. Then the vector

$$(2.3) \quad z_1 = x_{11} \otimes x_{21}$$

spans  $\mathcal{R}_1$ . Further, we choose nonzero vectors  $y_{i1} \in \mathcal{N}(W_i(\lambda)^*)$  ( $i = 1, 2$ ) and we write

$$(2.4) \quad \mathcal{S}_2 = \begin{bmatrix} \langle A_{10}x_{11}, y_{11} \rangle & \langle A_{11}x_{11}, y_{11} \rangle \\ \langle A_{20}x_{21}, y_{21} \rangle & \langle A_{21}x_{21}, y_{21} \rangle \end{bmatrix}.$$

Here we use  $\langle x, y \rangle$  to denote the scalar product of vectors  $x, y$ . It is clear from the context which Hilbert space  $H_i$  or  $H$  is meant. Now [5, Cor. 6.4] states that  $d_2 = \dim \mathcal{N}(\mathcal{S}_2)$ . If  $d_2 \leq 1$ , then we call  $\boldsymbol{\lambda}$  *nonderogatory*. Otherwise,  $d_2 = 2$ , and we call  $\boldsymbol{\lambda}$  *derogatory*. Bases for the root subspaces  $\mathcal{R}_m$  corresponding to a nonderogatory eigenvalue of an  $n$ -parameter system (for  $n \geq 2$ ) are constructed in [19] in the finite-dimensional setup. For completeness we will give an infinite-dimensional two-parameter version of the main result of [19] (see Theorem 2.2 below).

For  $\mathbf{c} = [c_1, c_2]^T \in \mathbb{C}^2$  and  $i = 1, 2$ , we write

$$U_i(\mathbf{c}) = A_{i0}c_1 + A_{i1}c_2.$$

If  $M = 1$ , then  $z_1$  spans  $\mathcal{R}_M$ . Assume from now on that  $M \geq 2$ . The following is a two-parameter analogue of the main result of [19]. It gives a basis for the root subspace at a nonderogatory eigenvalue.

**Theorem 2.2.** *Suppose that  $\boldsymbol{\lambda} \in \mathbb{C}^2$  is a nonderogatory eigenvalue for a two-parameter system (1.2) and suppose that  $2 \leq m \leq M$ . Then there exist  $\mathbf{c}_2, \mathbf{c}_3, \dots, \mathbf{c}_m \in \mathbb{C}^2$ ,  $\mathbf{c}_2 \neq 0$ , and  $x_{i2}, x_{i3}, \dots, x_{im} \in \mathcal{D}(A_{i0})$ ,  $i = 1, 2$ , such that*

$$(2.5) \quad \sum_{j=1}^{k-1} U_i(\mathbf{c}_{k+1-j}) x_{ij} + W_i(\boldsymbol{\lambda}) x_{ik} = 0 \quad \text{for } k = 2, 3, \dots, m, \quad i = 1, 2.$$

Moreover, for  $k = 1, 2, \dots, m$ , the vectors

$$(2.6) \quad z_k = \sum_{j=1}^k x_{1j} \otimes x_{2,k+1-j}$$

are such that

$$(2.7) \quad (\Gamma_{i-1} - \lambda_{i-1}I) z_k = \sum_{j=1}^{k-1} c_{k+1-j,i} z_j, \quad i = 1, 2,$$

where  $\mathbf{c}_k = [c_{k1} \quad c_{k2}]^T$  and  $\{z_1, z_2, \dots, z_m\}$  is a basis for the  $m$ -th root subspace  $\mathcal{R}_m$ .

*Proof.* The theorem is proved by induction on  $m$ . For each  $m$  we choose a vector  $z_m \in \mathcal{R}_m \setminus \mathcal{R}_{m-1}$  and use Assumptions II and III to prove the existence of vectors  $x_{ij} \in \mathcal{D}(A_{i0})$  such that (2.6) and (2.7) hold. Here the arguments follow closely those in the proof of [5, Thm. 6.2]. For the converse, (2.7) follows from (2.5) and (2.6) by a direct calculation (see [19, Lemma 13]).  $\square$

### 3. DEROGATORY EIGENVALUES

We assume now that  $\boldsymbol{\lambda}$  is a geometrically simple eigenvalue which is derogatory, i.e.,  $\mathcal{S}_2 = 0$  or equivalently  $d_2 = 2$ . First we recall some results on the structure of nilpotent commutative matrices with one-dimensional joint kernel and  $d_2 = 2$  (see [18, 20]).

**3.1. Structure of nilpotent commutative matrices.** We denote by  $\mathcal{R} = \mathcal{R}_M$  the root subspace corresponding to  $\lambda$  and we write  $N = \dim \mathcal{R}$ . The restricted linear maps  $C_i = (\Gamma_i - \lambda_i I)|_{\mathcal{R}}$ ,  $i = 0, 1$ , are commuting and nilpotent. There exists a basis

$$(3.1) \quad \mathcal{B} = \{z_1, z_{21}, z_{22}, z_{31}, \dots, z_{3d_3}, \dots, z_{M1}, z_{M2}, \dots, z_{Md_M}\}$$

for  $\mathcal{R}$  such that for  $m = 1, 2, \dots, M$ , the set

$$\mathcal{B}_m = \{z_1, z_{21}, z_{22}, z_{31}, \dots, z_{3d_3}, \dots, z_{m1}, z_{m2}, \dots, z_{md_m}\}$$

is a basis for  $\mathcal{R}_m$ . We call such a basis  $\mathcal{B}$  a *filtered basis (for  $\mathcal{R}$ )*. In a filtered basis the matrices for  $\mathbf{C} = \{C_0, C_1\}$  are simultaneously reduced to a special upper-triangular form and we view them as a cubic array of dimensions  $N \times N \times 2$

$$(3.2) \quad \mathbf{C} = \begin{bmatrix} \mathbf{0} & \mathbf{C}^{12} & \mathbf{C}^{13} & \dots & \mathbf{C}^{1,M} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{23} & \dots & \mathbf{C}^{2,M} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}^{M-1,M} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix},$$

where

$$\mathbf{C}^{km} = \begin{bmatrix} \mathbf{c}_{11}^{km} & \mathbf{c}_{12}^{km} & \dots & \mathbf{c}_{1,d_m}^{km} \\ \mathbf{c}_{21}^{km} & \mathbf{c}_{22}^{km} & \dots & \mathbf{c}_{2,d_m}^{km} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{c}_{d_k,1}^{km} & \mathbf{c}_{d_k,2}^{km} & \dots & \mathbf{c}_{d_k,d_m}^{km} \end{bmatrix}$$

is a cubic array of dimensions  $d_k \times d_m \times 2$  and  $\mathbf{c}_{gf}^{km} = [c_{1gf}^{km} \ c_{2gf}^{km}]^T \in \mathbb{C}^2$ . The *column cross-sections*

$$(3.3) \quad \mathcal{C}_f^{km} = [\mathbf{c}_{1f}^{km} \ \mathbf{c}_{2f}^{km} \ \dots \ \mathbf{c}_{d_k,f}^{km}], \quad f = 1, 2, \dots, d_m,$$

of  $\mathbf{C}^{km}$  are  $2 \times d_k$  matrices. The constants  $c_{igf}^{km}$  are determined by the relations

$$C_{i-1} z_{mf} = \sum_{k=1}^{m-1} \sum_{g=1}^{d_k} c_{igf}^{km} z_{kg}, \quad i = 1, 2,$$

where we write  $z_{11} = z_1$ . (For further details see [18, 20].)

For  $m \geq 4$  we denote by  $\Phi_m$  the set of multi-indices

$$\left\{ (l_1, l_2, l_3); \ l_j \geq 2, \ \sum_{j=1}^3 l_j \leq m+2 \right\},$$

and for  $\mathbf{l} = (l_1, l_2, l_3) \in \Phi_m$  we denote by  $\chi_{\mathbf{l}}$  the set of multi-indices

$$\{(h_1, h_2, h_3); \ 1 \leq h_j \leq d_{l_j} \text{ for } j = 1, 2, 3\}.$$

We call a matrix  $C$  *symmetric* if  $C = C^T$ .

For use in our proofs we state a version of some of results of [18, 20] for pairs of commutative matrices. Since  $\mathbf{c}_{11}^{12}$  and  $\mathbf{c}_{12}^{12}$  are linearly independent we can choose a basis for  $\mathcal{R}_2$  such that  $C_{i-1} z_{2j} = \delta_{ij} z_1$ , where  $\delta_{ij}$  is the Kronecker symbol. By adding appropriate linear combinations of  $z_{21}$  and  $z_{22}$  to other vectors  $z_{mf}$ ,  $m \geq 3$ , we find a basis  $\mathcal{B}$  such that the following result holds.

**Theorem 3.1.** *There exists a filtered basis  $\mathcal{B}$  for  $\mathcal{R}$  and associated with it a set of symmetric matrices*

$$(3.4) \quad C_f^m = [C_f^{klm}]_{k,l=2}^{m-1}, \quad m = 3, 4, \dots, M, \quad f = 1, 2, \dots, d_m,$$

where  $C_f^{klm} = [c_{ghf}^{klm}]_{g=1, h=1}^{d_k \quad d_l}$  are  $d_k \times d_l$  matrices, such that for the corresponding commutative array (3.2) we have:

- (a)  $C_1^{12} = [1 \quad 0]^T$ ,  $C_2^{12} = [0 \quad 1]^T$  and  $C_f^{1m} = [0 \quad 0]^T$  for  $m = 3, 4, \dots, M$  and  $f = 1, 2, \dots, d_m$ ,
- (b)  $C_f^{2km} = C_f^{km}$  are the column cross-sections (3.3),
- (c) for  $m = 3, 4, \dots, M$  the matrices  $C_f^{2, m-1, m}$ ,  $f = 1, 2, \dots, d_m$ , are linearly independent,
- (d)  $C_f^{klm} = 0$  if  $k + l > m + 1$ ,
- (e) for  $m = 4, 5, \dots, M$  and  $f = 1, 2, \dots, d_m$  the relations

$$(3.5) \quad \sum_{k=l_1+l_2-1}^{m-l_3+1} \sum_{g=1}^{d_k} c_{h_1 h_2 g}^{l_1 l_2 k} c_{h_3 g f}^{l_3 k m} = \sum_{k=l_2+l_3-1}^{m-l_1+1} \sum_{g=1}^{d_k} c_{h_1 h_3 g}^{l_1 l_3 k} c_{h_2 g f}^{l_2 k m}$$

hold for all  $\mathbf{l} = (l_1, l_2, l_3) \in \Phi_m$  and  $(h_1, h_2, h_3) \in \chi_1$ .

We remark that property (e) in the above theorem together with the symmetry of matrices  $C_f^m$  is equivalent to the commutativity of the matrices  $C_0$  and  $C_1$  and that  $c_{ghf}^{klm}$  are the structure constants for multiplication in the commutative algebra generated by  $C_0, C_1$  and the identity matrix (see [20]).

In the rest of the section we consider, unless stated otherwise, only filtered bases  $\mathcal{B}$  such that the associated matrices  $C_f^m$  satisfy the properties (a)–(e) of Theorem 3.1.

**3.2. Bases for  $\mathcal{R}_m$  for  $m = 2, 3$ .** We denote by  $\mathcal{D}'_i$  the (vector space) intersection  $\mathcal{D}(A_{i0}) \cap (\mathcal{N}(W_i(\boldsymbol{\lambda})))^\perp$ , where  $(\mathcal{N}(W_i(\boldsymbol{\lambda})))^\perp$  is the orthogonal complement of  $\mathcal{N}(W_i(\boldsymbol{\lambda}))$  in  $H_i$ .

**Lemma 3.2.** *For  $i = 1, 2$ , the subspace  $\mathcal{D}'_i$  is an infinite-dimensional vector subspace of  $\mathcal{D}_i = \mathcal{D}_i(A_{i0})$ .*

*Proof.* Since the lattice of vector subspaces is modular [7, p. 13] and since  $\mathcal{N}(W_i(\boldsymbol{\lambda})) \subset \mathcal{D}_i$ , it follows that

$$\mathcal{D}_i = (\mathcal{N}(W_i(\boldsymbol{\lambda})) \oplus (\mathcal{N}(W_i(\boldsymbol{\lambda})))^\perp) \cap \mathcal{D}_i = \mathcal{N}(W_i(\boldsymbol{\lambda})) + (\mathcal{N}(W_i(\boldsymbol{\lambda}))^\perp \cap \mathcal{D}_i).$$

Since  $\mathcal{N}(W_i(\boldsymbol{\lambda}))$  is finite-dimensional and  $\mathcal{D}_i$  infinite-dimensional,  $\mathcal{D}'_i$  is also an infinite-dimensional vector space.  $\square$

The following is a technical observation, which enables us to shorten proofs but has no significance otherwise. In actual calculations of basis vectors we can choose vectors  $x_{ik}^g \in \mathcal{D}(A_{i0})$  (i.e. not necessarily  $x_{ik}^g \in \mathcal{D}'_i$ ) that satisfy all other required conditions.

**Lemma 3.3.** *Suppose that a vector  $x_i \in (\mathcal{N}(W_i(\boldsymbol{\lambda})^*))^\perp$ . Then there exists a vector  $v_i \in \mathcal{D}'_i$  such that  $x_i = W_i(\boldsymbol{\lambda}) v_i$ .*

*Proof.* Suppose that  $x_i \in (\mathcal{N}(W_i(\boldsymbol{\lambda}))^\perp)^\perp$ . Since  $W_i(\boldsymbol{\lambda})$  is a Fredholm operator, its range is a closed subspace of  $H_i$  [25] and therefore  $(\mathcal{N}(W_i(\boldsymbol{\lambda}))^\perp)^\perp = \mathcal{R}(W_i(\boldsymbol{\lambda}))$ . Thus there exists a vector  $u_i \in \mathcal{D}(A_{i0})$  such that  $x_i = W_i(\boldsymbol{\lambda}) u_i$ . Since  $\mathcal{D}(W_i(\boldsymbol{\lambda})) = \mathcal{D}(A_{i0})$  we can find vectors  $v_i \in \mathcal{D}'_i$  and  $w_i \in \mathcal{N}(W_i(\boldsymbol{\lambda}))$  such that  $v_i + w_i = u_i$ . Then it follows that  $x_i = W_i(\boldsymbol{\lambda}) v_i$ .  $\square$

The following result is a special case of [5, Thm. 6.3]. Note that the additional condition that  $x_{i2}^k \in \mathcal{D}'_i$  is a consequence of Lemma 3.3.

**Theorem 3.4.** *Assume that  $\boldsymbol{\lambda} \in \sigma(\mathbf{W})$  is a derogatory eigenvalue, i.e.,  $\mathcal{S}_2 = 0$ , and that  $\mathbf{c}_{11}^{12} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$  and  $\mathbf{c}_{12}^{12} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ . Then there exist vectors  $x_{i2}^k \in \mathcal{D}'_i$  such that*

$$(3.6) \quad U_i(\mathbf{c}_{1k}^{12}) x_{i1} + W_i(\boldsymbol{\lambda}) x_{i2}^k = 0$$

for  $i, k = 1, 2$ . Furthermore, the vectors

$$(3.7) \quad z_{2k} = x_{11} \otimes x_{22}^k + x_{12}^k \otimes x_{21}$$

are such that

$$(\Gamma_{i-1} - \lambda_{i-1} I) z_{2k} = \delta_{ik} z_1$$

for  $i, k = 1, 2$ , and  $\{z_1, z_{21}, z_{22}\}$  is a basis for  $\mathcal{R}_2$ . Here  $\delta_{ik}$  is the Kronecker symbol.

Next we define the matrix  $\mathcal{S}_3$  by

$$\mathcal{S}_3 = \begin{bmatrix} \langle A_{10} x_{12}^1, y_{11} \rangle & \langle A_{10} x_{12}^2, y_{11} \rangle + \langle A_{11} x_{12}^1, y_{11} \rangle & \langle A_{11} x_{12}^2, y_{11} \rangle \\ \langle A_{20} x_{22}^1, y_{21} \rangle & \langle A_{20} x_{22}^2, y_{21} \rangle + \langle A_{21} x_{22}^1, y_{21} \rangle & \langle A_{21} x_{22}^2, y_{21} \rangle \end{bmatrix}.$$

We identify the vector space  $\Theta$  of all symmetric  $2 \times 2$  matrices with the vector space  $\mathbb{C}^3$  via the isomorphism  $\psi : \Theta \rightarrow \mathbb{C}^3$  defined by

$$(3.8) \quad \psi(C) = \begin{bmatrix} c_{11} & c_{12} & c_{22} \end{bmatrix}^T$$

for

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix} \in \Theta.$$

The following theorem gives the structure of basis vectors of  $\mathcal{R}_3$ .

**Theorem 3.5.** *Suppose that  $\{\mathbf{c}_l : 1 \leq l \leq d\}$  is a basis for  $\mathcal{N}(\mathcal{S}_3)$  and that  $C_l^3 = \left[ c_{jkl}^{223} \right]_{j,k=1}^2 = \psi^{-1}(\mathbf{c}_l) \in \Theta$  for  $1 \leq l \leq d$ . Then we have  $d = d_3$  and there exist vectors  $x_{i3}^l \in \mathcal{D}'_i$ ,  $i = 1, 2$ ,  $l = 1, 2, \dots, d$ , such that*

$$(3.9) \quad \sum_{k=1}^2 U_i(\mathbf{c}_{kl}^{23}) x_{i2}^k + W_i(\boldsymbol{\lambda}) x_{i3}^l = 0,$$

where  $\mathbf{c}_{kl}^{23} = \begin{bmatrix} c_{1kl}^{223} & c_{2kl}^{223} \end{bmatrix}^T$ . The vector

$$(3.10) \quad z_{31} = x_{11} \otimes x_{23}^l + x_{13}^l \otimes x_{21} + \sum_{j,k=1}^2 c_{jkl}^{223} x_{12}^j \otimes x_{22}^k$$

is in  $\mathcal{R}_3 \setminus \mathcal{R}_2$ , and

$$(3.11) \quad (\Gamma_{i-1} - \lambda_{i-1} I) z_{3l} = \sum_{k=1}^2 c_{ikl}^{223} z_{2k}$$



for  $i = 1, 2$  and  $1 \leq l \leq d$ . Moreover,  $\{z_1, z_{21}, z_{22}, z_{31}, z_{32}, \dots, z_{3d}\}$  is a basis for  $\mathcal{R}_3$ .

Conversely, choose vectors  $z_{31}, z_{32}, \dots, z_{3d_2} \in \mathcal{R}_3 \setminus \mathcal{R}_2$  such that  $\{z_1, z_{21}, z_{22}, z_{31}, z_{32}, \dots, z_{3d_3}\}$  is a basis for  $\mathcal{R}_3$  and such that (3.11) holds for both  $i$  and all  $l$ . Then

$$C_l^3 = \begin{bmatrix} c_{11l}^{223} & c_{12l}^{223} \\ c_{21l}^{223} & c_{22l}^{223} \end{bmatrix}$$

are symmetric,  $\{\psi(C_l^3) : 1 \leq l \leq d_3\}$  is a basis for  $\mathcal{N}(\mathcal{S}_3)$ , and there exist vectors  $x_{i3}^l \in \mathcal{D}'_i$ ,  $i = 1, 2$ , such that (3.9) and (3.10) hold.

*Proof.* Because  $\mathbf{c}_l \in \mathcal{N}(\mathcal{S}_3)$  and  $C_l^3 = \psi^{-1}(\mathbf{c}_l)$  it follows that

$$\sum_{j=1}^2 \sum_{k=1}^2 c_{jkl}^{223} \langle A_{i,j-1} x_{i2}^l, y_{i1} \rangle = 0$$

for  $i = 1, 2$ . Thus  $\sum_{k=1}^2 U_i(\mathbf{c}_{kl}^{223}) x_{i2}^k \in (\mathcal{N}(W_i(\boldsymbol{\lambda}^*))^\perp)$  and by Lemma 3.3 there exist vectors  $x_{i3} \in \mathcal{D}'_i$  such that (3.9) hold for  $i = 1, 2$ . Next we form the vector  $z_{3l}$  as in (3.10). In the following calculation we assume that  $i = 0$ . For  $i = 1$  the calculation is similar and we omit it. We have that

$$\begin{aligned} & (\Delta_0 - \lambda_0 \Delta_2) z_{3l} \\ &= (A_{11} \otimes W_2(\boldsymbol{\lambda}) - W_1(\boldsymbol{\lambda}) \otimes A_{21}) \left( x_{11} \otimes x_{23}^l + x_{13}^l \otimes x_{21} + \sum_{j,k=1}^2 c_{jkl}^{223} x_{12}^j \otimes x_{22}^k \right) \\ &= \sum_{j,k=1}^2 c_{jkl}^{223} \left( A_{1,k-1} x_{12}^j \otimes A_{21} x_{21} - A_{11} x_{11} \otimes A_{2,j-1} x_{22}^k \right) \\ &+ \sum_{j,k=1}^2 c_{jkl}^{223} \left( -A_{11} x_{12}^j \otimes A_{2,k-1} x_{21} + A_{1,j-1} x_{11} \otimes A_{21} x_{22}^k \right) \\ &= \sum_{j,k=1}^2 c_{jkl}^{223} (-A_{11} \otimes A_{2,j-1} + A_{1,j-1} \otimes A_{21}) (x_{12}^k + x_{11} \otimes x_{22}^k) \\ &= \Delta_2 \left( \sum_{k=1}^2 c_{1kl}^{223} z_{2k} \right). \end{aligned}$$

Conversely, suppose that  $z_3 \in \mathcal{R}_3 \setminus \mathcal{R}_2$ . Then there exist a symmetric matrix

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix} \in \Theta$$

and a vector  $\mathbf{a} = [a_0 \ a_1]^T \in \mathbb{C}^2$  such that for  $i = 0, 1$ , we have

$$(\Gamma_i - \lambda_i I) z_3 = \sum_{k=1}^2 c_{i+1,k} z_{2k} + a_i z_{11}.$$

We may and will assume that  $\mathbf{a} = 0$ . This is achieved by substituting  $z_3 - a_0 z_{21} - a_1 z_{22}$  for  $z_3$ . Thus we have

$$(3.12) \quad (\Gamma_i - \lambda_i I) z_3 = \sum_{k=1}^2 c_{i+1,k} z_{2k}.$$

Next it follows from (2.1) that

$$\sum_{j=0}^1 A_{ij}^\dagger (\Gamma_j - \lambda_j I) z_3 + W_i(\boldsymbol{\lambda})^\dagger z_3 = 0$$

and from (3.12) that

$$(3.13) \quad \sum_{j=0}^1 A_{ij}^\dagger \left( \sum_{k=1}^2 c_{j+1,k} z_{2k} \right) + W_i(\boldsymbol{\lambda})^\dagger z_3 = 0.$$

For  $i = 1, 2$ , we choose vectors  $v_i \in H_i$  so that  $\langle x_{i1}, v_i \rangle = 1$  and  $\langle x_{i2}^k, v_i \rangle = 0$  for  $k = 1, 2$ . This is possible because  $\text{Span}\{x_{i1}\} \cap \mathcal{D}'_i = \{0\}$ . For  $i = 1$  and  $i = 2$ , respectively, we form the scalar product of (3.13) by  $y_{11} \otimes v_2$  and  $v_1 \otimes y_{21}$ , respectively, to get

$$(3.14) \quad \sum_{j=0}^1 \left\langle A_{ij} \sum_{k=1}^2 c_{j+1,k} x_{i2}^k, y_{i1} \right\rangle = 0.$$

Hence, by Lemma 3.3, there exist vectors  $x_{i3} \in \mathcal{D}'_i$  such that

$$(3.15) \quad \sum_{k=1}^2 U_i(\mathbf{c}_k) x_{i2}^k + W_i(\boldsymbol{\lambda}) x_{i3} = 0,$$

where  $\mathbf{c}_k = [c_{1k} \ c_{2k}]^T$ . Now we form the vector

$$z_{31} = x_{11} \otimes x_{23} + x_{13} \otimes x_{21} + \sum_{j,k=1}^2 c_{jk} x_{12}^j \otimes x_{22}^k.$$

The same calculation as in the first part of the proof shows that

$$(\Gamma_i - \lambda_i I) z_{31} = \sum_{k=1}^2 c_{i+1,k} z_{2k}$$

for  $i = 0, 1$ . It follows that  $z_{31} - z_3 \in \mathcal{R}_1$  and so there exists a number  $\delta \in \mathbb{C}$  such that  $z_3 = z_{31} + \delta z_1$ . Without loss, we can use the vector  $x_{13} + \delta x_{11}$  in place of  $x_{13}$ . Then it follows that

$$z_3 = x_{11} \otimes x_{23} + x_{13} \otimes x_{21} + \sum_{j,k=1}^2 c_{jk} x_{12}^j \otimes x_{22}^k.$$

The final step is to verify that  $\psi(T) \in \mathcal{N}(\mathcal{S}_3)$ , which follows immediately from the equalities (3.14).  $\square$

Note that Theorem 3.1 implies the existence of vectors  $z_{3l} \in \mathcal{R}_3 \setminus \mathcal{R}_2$  ( $1 \leq l \leq d_3$ ) such that the following hold:

- $\{z_1, z_{21}, z_{22}, z_{31}, z_{32}, \dots, z_{3d_3}\}$  is a basis for  $\mathcal{R}_3$ ,
- equality (3.11) holds for both  $i$  and all  $l$ ,
- $\{C_l^3 : 1 \leq l \leq d_3\}$  are linearly independent.

Since  $\Gamma_0$  and  $\Gamma_1$  commute it follows that  $c_{12l}^{223} = c_{21l}^{223}$  ( $1 \leq l \leq d_3$ ) and therefore the matrices  $C_l^3$  are symmetric.

4. A BASIS FOR ROOT SUBSPACE  $\mathcal{R}_m$ ,  $m \geq 3$ , AT A DEROGATORY EIGENVALUE

Our goal in this section is to prove, by induction on  $m$ , the following theorem which is our main result.

**Theorem 4.1.** *Suppose that  $\{\mathbf{c}_{1f}^m, f = 1, 2, \dots, d\}$  is a basis for the kernel of  $\mathcal{S}_m$ . Then  $d = d_m$  and there exist vectors  $x_{im}^f \in \mathcal{D}'_i$  and matrices  $C_f^m$  such that conditions (i) to (iii) below hold when  $m-1$  is replaced by  $m$ . In particular, the union  $\mathcal{B}_m = \mathcal{B}_{m-1} \cup \{z_{m1}, z_{m2}, \dots, z_{md}\}$ , where*

$$(4.1) \quad z_{mf} = x_{1m}^f \otimes x_{21} + x_{11} \otimes x_{2m}^f + \sum_{k=2}^{\left[\frac{m}{2}\right]} \sum_{l=k}^{m+1-k} \sum_{g=1}^{d_k} \sum_{h=1}^{d_l} c_{ghf}^{klm} (x_{1k}^g \otimes x_{2l}^h + x_{1l}^h \otimes x_{2k}^g)$$

for  $f = 1, 2, \dots, d$ , is a basis for the root subspace  $\mathcal{R}_m$ .

Here  $\left[\frac{m}{2}\right]$  denotes the integer part of  $\frac{m}{2}$ .

We will prove Theorem 4.1 at the end of the section. Let us first introduce the inductive assumptions. For  $m = 1, 2, 3, \dots$  we suppose that we have vectors  $x_{i1} \in \mathcal{D}(A_{i0})$ ,  $x_{il}^h \in \mathcal{D}'_i$ ,  $i = 1, 2$ ,  $l = 2, 3, \dots, m-1$ ,  $h = 1, 2, \dots, d_l$ , and symmetric (structure) matrices (see (3.4))

$$(4.2) \quad C_h^l = \begin{bmatrix} C_h^{22l} & C_h^{23l} & \dots & C_h^{2,l-2,l-1} & C_h^{2,l-1,l} \\ (C_h^{23l})^T & C_h^{33l} & \dots & C_h^{3,l-2,l} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ (C_h^{2,l-2,l})^T & (C_h^{3,l-2,l})^T & \dots & 0 & 0 \\ (C_h^{2,l-1,l})^T & 0 & \dots & 0 & 0 \end{bmatrix}$$

such that

- (i) the matrices  $C_h^{2,l-1,l}$ ,  $h = 1, 2, \dots, d_l$ , are linearly independent for  $l = 3, 4, \dots, m-1$ , the column cross-sections (see (3.3))  $C_h^{jl}$  are equal to  $C_h^{2jl}$ , and the entries of matrices  $C_h^l$  satisfy the relations

$$(4.3) \quad \sum_{k=l_1+l_2-1}^{l-l_3+1} \sum_{g=1}^{d_k} c_{h_1h_2g}^{l_1l_2k} c_{h_3gh}^{l_3kl} = \sum_{k=l_2+l_3-1}^{l-l_1+1} \sum_{g=1}^{d_k} c_{h_1h_3g}^{l_1l_3k} c_{h_2gh}^{l_2kl}$$

for  $l = 4, 5, \dots, m-1$ ,  $\mathbf{l} \in \Phi_l$  and  $\mathbf{h} \in \chi_1$ ,

- (ii) the relations (2.2), (3.6) and

$$(4.4) \quad \sum_{k=2}^{l-1} \sum_{g=1}^{d_k} U_i(\mathbf{c}_{gh}^{kl}) x_{ik}^g + W_i(\boldsymbol{\lambda}) x_{il}^h = 0$$

for  $i = 1, 2$ ,  $l = 3, 4, \dots, m-1$  and  $h = 1, 2, \dots, d_l$  hold (here  $\mathbf{c}_{gh}^{kl} = \begin{bmatrix} c_{1gh}^{2kl} & c_{2gh}^{2kl} \end{bmatrix}^T$ ),

- (iii) the vectors  $z_1, z_{21}, z_{22}$ , of (2.3) and (3.7) together with vectors  $z_{31}, z_{32}, \dots, z_{3d_3}, \dots, z_{m-1,1}, z_{m-1,2}, \dots, z_{m-1,d_{m-1}}$ , that are defined by

$$z_{lh} = x_{1l}^h \otimes x_{21} + x_{11} \otimes x_{2l}^h + \sum_{k=2}^{\left[\frac{l}{2}\right]} \sum_{r=k}^{l+1-k} \sum_{g=1}^{d_k} \sum_{e=1}^{d_r} c_{geh}^{krl} (x_{1k}^g \otimes x_{2r}^e + x_{1r}^e \otimes x_{2k}^g)$$

for  $l = 3, 4, \dots, m-1$  and  $h = 1, 2, \dots, d_l$ , form a filtered basis for the root subspace  $\mathcal{R}_{m-1}$ . We also write

$$\mathcal{B}_{m-1} = \{z_1, z_{21}, z_{22}, \dots, z_{m-1,1}, z_{m-1,2}, \dots, z_{m-1,d_{m-1}}\}.$$

By Theorem 3.5 it follows that there exist vectors  $x_{i1}, x_{i2}^h, x_{i3}^h$  and matrices  $C_2^h$  such that conditions (i) to (iii) are satisfied for  $m-1=3$ . Now we assume that the above conditions are satisfied for  $m-1$  (with  $3 \leq m-1 \leq M-1$ ) and we prove that we can find vectors  $x_{im}^f \in \mathcal{D}'_i$ ,  $i = 1, 2$ ,  $f = 1, 2, \dots, d_m$ , and matrices  $C_f^m$ ,  $f = 1, 2, \dots, d_m$ , such that (i) to (iii) hold with  $m-1$  replaced by  $m$ .

Our next step is to introduce matrices  $\mathcal{S}_m$ . We do this in two stages. First we use matrices  $C_g^k$ ,  $k < m$ , and second, certain orthogonality relations.

For the purpose of calculation we write  $T^m = [T^{klm}]_{k,l=2}^{m-1}$  for an unknown symmetric matrix in the form (4.2). The entries of the  $d_k \times d_l$  matrix  $T^{klm}$  are denoted by  $t_{gh}^{klm}$ . They must satisfy the conditions

$$(4.5) \quad \sum_{k=l_1+l_2-1}^{m-l_3+1} \sum_{g=1}^{d_k} C_{h_1 h_2 g}^{l_1 l_2 k} t_{h_3 g}^{l_3 k m} - \sum_{k=l_1+l_3-1}^{m-l_2+1} \sum_{g=1}^{d_k} C_{h_1 h_3 g}^{l_1 l_3 k} t_{h_2 g}^{l_2 k m} = 0$$

for  $\mathbf{l} \in \Phi_m$  and  $\mathbf{h} \in \chi_1$ . We write the  $2 \times d_{m-1}$  matrix  $T^{1,m-1,m}$  also as a column

$$(4.6) \quad \mathbf{t}_1^m = \begin{bmatrix} t_{11}^{2,m-1,m} & t_{21}^{2,m-1,m} & t_{12}^{2,m-1,m} & t_{22}^{2,m-1,m} & \dots & t_{1d_{m-1}}^{2,m-1,m} & t_{2d_{m-1}}^{2,m-1,m} \end{bmatrix}^T.$$

We denote the sum  $\sum_{i=1}^l d_i$  by  $\widehat{d}_l$ . For every column of  $T^m$  we define a column vector  $\mathbf{t}_{kg}^m$ ,  $k = 2, 3, \dots, m-1$ ,  $g = 1, 2, \dots, d_k$ , of size  $\nu = \min \{ \widehat{d}_{k-1} + g, \widehat{d}_{m-k} \}$  by taking the first  $\nu$  entries in the  $(\widehat{d}_{k-1} + g)$ -th column of  $T^m$ . Note that  $\mathbf{t}_{kg}^m$  are defined so that they consist of all the entries above and including the main diagonal of the matrix  $T^m$  omitting those labelled 0 in (4.2). We define a column vector  $\mathbf{t}_2^m$  by

$$(4.7) \quad \mathbf{t}_2^m = [ (\mathbf{t}_{11}^m)^T \quad (\mathbf{t}_{12}^m)^T \quad (\mathbf{t}_{21}^m)^T \quad \dots \quad (\mathbf{t}_{2d_2}^m)^T \quad \dots \quad (\mathbf{t}_{m-2,1}^m)^T \quad \dots \quad (\mathbf{t}_{m-2,d_{m-2}}^m)^T ]^T.$$

We split the entries of the matrix  $T^m$  into two column vectors  $\mathbf{t}_1^m$  and  $\mathbf{t}_2^m$ . The mapping  $\psi_m : T^m \mapsto (\mathbf{t}_1^m, \mathbf{t}_2^m)$ , defined on all symmetric matrices of the form (4.2), is a generalization of the transformation  $\psi$  defined by (3.8). It is bijective and therefore it has an inverse, which maps two vectors  $\mathbf{t}_1^m$  and  $\mathbf{t}_2^m$  into a matrix  $T^m$  of the form (4.2). We use this inverse mapping in Lemma 4.3. We write the system of equations (4.5) in matrix form as

$$(4.8) \quad S_m^{21} \mathbf{t}_1^m + S_m^{22} \mathbf{t}_2^m = 0$$

where the entries of the matrices  $S_m^{21}$  and  $S_m^{22}$  are determined by the system (4.5). These entries are given because we assumed that the matrices  $C_g^k$  were known for  $k \leq m-1$ .

We also want the entries of the matrices  $T^{2km}$  to satisfy the scalar relations

$$(4.9) \quad \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} \langle (t_{1g}^{2km} A_{i0} + t_{2g}^{2km} A_{i1}) x_{ik}^g, y_{i1} \rangle = 0$$

for both  $i$ . These can be written equivalently in matrix form

$$(4.10) \quad S_m^{11} \mathbf{t}_1^m + S_m^{12} \mathbf{t}_2^m = 0.$$

Again the entries of the matrices  $S_m^{11}$  and  $S_m^{12}$  are determined by the equations (4.9).

Now we are prepared to define the matrices  $\mathcal{S}_m$ . First we choose a matrix  $B_m$  so that its columns form a basis for the kernel of

$$\begin{bmatrix} (S_m^{12})^* & (S_m^{22})^* \end{bmatrix}.$$

Then we define a matrix

$$\mathcal{S}_m = B_m^* \begin{bmatrix} S_m^{11} \\ S_m^{21} \end{bmatrix},$$

so equations (4.8) and (4.10) yield  $\mathcal{S}_m \mathbf{t}_1^m = 0$ .

We now choose vectors  $z_{mf} \in \mathcal{D}$ ,  $f = 1, 2, \dots, d_m$ , so that  $\mathcal{B}_{m-1} \cup \{z_{m1}, z_{m2}, \dots, z_{md_m}\}$ , is a basis for the space  $\mathcal{R}_m$ . By Theorem 3.1 there exist matrices  $C_f^m$  for  $f = 1, 2, \dots, d_m$  such that (i) holds and that we have

$$(4.11) \quad (\Gamma_{i-1} - \lambda_{i-1}I) z_{mf} = \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} c_{igf}^{2km} z_{kg}$$

for  $i = 1, 2$  and  $f = 1, 2, \dots, d_m$ . Next we prove three auxiliary results.

**Lemma 4.2.** *In the above setting it follows that  $\dim \mathcal{N}(\mathcal{S}_m) \geq d_m$ .*

*Proof.* Let  $\psi_m(C_f^m) = (\mathbf{c}_{1f}^m, \mathbf{c}_{2f}^m)$ . By Theorem 3.1(e) and the definition of  $\psi_m$  it follows that  $S_m^{21} \mathbf{c}_{1f}^m + S_m^{22} \mathbf{c}_{2f}^m = 0$ . Relation (2.1) implies

$$\sum_{j=0}^1 A_{ij}^\dagger (\Gamma_j - \lambda_j I) z_{mf} + W_i(\boldsymbol{\lambda})^\dagger z_{mf} = 0$$

for both  $i$ . By relations (4.11) we get that

$$(4.12) \quad \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} U_i(\mathbf{c}_{gf}^{km})^\dagger z_{kg} + W_i(\boldsymbol{\lambda})^\dagger z_{mf} = 0.$$

Here  $\mathbf{c}_{gf}^{km} = [c_{1gf}^{km} \ c_{2gf}^{km}]^T$ . Because we assume that  $x_{ik}^g \in \mathcal{D}'_i$  and  $\mathcal{D}'_i \cap \text{Span}\{x_{i1}\} = \{0\}$  it follows that for each  $i$  there exists a vector  $v_i \in H_i$  such that  $\langle x_{i1}, v_i \rangle = 1$  and  $\langle x_{ik}^g, v_i \rangle = 0$  for  $k = 2, 3, \dots, m-1$  and  $g = 1, 2, \dots, d_k$ . Then the scalar products of (4.12) by  $y_{11} \otimes v_2$  and  $v_1 \otimes y_{21}$ , for  $i = 1$  and  $i = 2$  respectively, are equal to 0. It follows, using the structure of the vectors  $z_{kg}$  in condition (iii), that

$$(4.13) \quad \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} \langle U_i(\mathbf{c}_{gf}^{km}) x_{ik}^g, y_{i1} \rangle = 0$$

for all  $f$ , and so

$$\sum_{j=1}^2 \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} c_{jgf}^{2km} \langle A_{i,j-1} x_{ik}^g, y_{i1} \rangle = 0.$$

Therefore  $\mathbf{c}_{1f}^m$  and  $\mathbf{c}_{2f}^m$  solve equation (4.10) and  $\mathbf{c}_{1f}^m$ ,  $f = 1, 2, \dots, d_m$ , are elements of the kernel of the matrix  $\mathcal{S}_m$ . Since these vectors are linearly independent, we have  $d_m \leq \dim \mathcal{N}(\mathcal{S}_m)$ .  $\square$

**Lemma 4.3.** *Suppose that  $\mathbf{c}_{11}^m$  is an element of the kernel  $\mathcal{N}(\mathcal{S}_m)$ . Then there exists a vector  $\mathbf{c}_{21}^m$  such that (4.8) and (4.10) hold. Furthermore, there exist vectors  $x_{im}^1 \in \mathcal{D}'_i$ ,  $i = 1, 2$ , such that*

$$(4.14) \quad \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} U_i(\mathbf{c}_{g1}^{km}) x_{ik}^g + W_i(\boldsymbol{\lambda}) x_{im}^1 = 0.$$

Here  $\mathbf{c}_{g1}^{km}$  are determined by the first block column of  $C_1^m = \psi_m^{-1}(\mathbf{c}_1^m, \mathbf{c}_2^m)$ .

*Proof.* From the structure of the matrix  $\mathcal{S}_m$  it follows that for an element  $\mathbf{c}_{11}^m \in \mathcal{N}(\mathcal{S}_m)$  there exists a vector  $\mathbf{c}_{21}^m$  such that relations (4.8) and (4.10) hold. Using the inverse of the isomorphism  $\psi_m^{-1}$  we associate with the pair of vectors  $\mathbf{c}_{11}^m$  and  $\mathbf{c}_{21}^m$  a matrix  $C_1^m$ . The relations (4.10) can be written equivalently in the form (4.13). Then it follows for both  $i$  that

$$\sum_{k=2}^{m-1} \sum_{g=1}^{d_k} U_i(\mathbf{c}_{g1}^{km}) x_{ik}^g \in (\mathcal{N}(W_i(\boldsymbol{\lambda})^*))^\perp = \mathcal{R}(W_i(\boldsymbol{\lambda})).$$

The latter equality holds since  $W_i(\boldsymbol{\lambda})$  is a Fredholm operator. By Lemma 3.3 there exists a vector  $x_{im}^1 \in \mathcal{D}'_i$  such that (4.14) holds.  $\square$

**Lemma 4.4.** *Suppose that we have the same setting as in Lemma 4.3. We construct a vector*

$$z_{m1} = x_{1m}^1 \otimes x_{21} + x_{11} \otimes x_{2m}^1 + \sum_{k=1}^{\lfloor \frac{m}{2} \rfloor} \sum_{l=k}^{m+1-k} \sum_{g=1}^{d_k} \sum_{h=1}^{d_l} c_{gh1}^{klm} (x_{1k}^g \otimes x_{2l}^h + x_{1l}^h \otimes x_{2k}^g).$$

Then it follows that

$$(4.15) \quad (\Gamma_{j-1} - \lambda_{j-1} I) z_{m1} = \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} c_{jg1}^{2km} z_{kg}$$

for  $j = 1, 2$ .

*Proof.* We use a direct calculation to show (4.15). We consider only the case  $j = 1$ . The case  $j = 2$  is proved in the same way. By the basic properties of operator determinants we have

$$(4.16) \quad (\Delta_0 - \lambda_0 \Delta_2) z_{m1} - \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} c_{1g1}^{2km} \Delta_2 z_{kg}$$

$$(4.17) \quad = \begin{vmatrix} -W_1(\boldsymbol{\lambda})^\dagger & A_{11}^\dagger \\ -W_2(\boldsymbol{\lambda})^\dagger & A_{21}^\dagger \end{vmatrix} (x_{1m}^1 \otimes x_{21} + x_{11} \otimes x_{2m}^1)$$

$$(4.18) \quad - \sum_{k=2}^{m-1} \sum_{g=1}^{d_k} \begin{vmatrix} U_1(\mathbf{c}_{g1}^{km})^\dagger & A_{11}^\dagger \\ U_2(\mathbf{c}_{g1}^{km})^\dagger & A_{21}^\dagger \end{vmatrix} (x_{1k}^g \otimes x_{21} + x_{11} \otimes x_{2k}^g)$$

$$(4.19) \quad + \begin{vmatrix} -W_1(\boldsymbol{\lambda})^\dagger & A_{11}^\dagger \\ -W_2(\boldsymbol{\lambda})^\dagger & A_{21}^\dagger \end{vmatrix} \sum_{k=2}^{\lfloor \frac{m}{2} \rfloor} \sum_{l=k}^{m+1-k} \sum_{g=1}^{d_k} \sum_{h=1}^{d_l} c_{gh1}^{klm} (x_{1k}^g \otimes x_{2l}^h + x_{1l}^h \otimes x_{2k}^g)$$

$$(4.20) \quad - \sum_{k=3}^{m-1} \sum_{g=1}^{d_k} c_{1g1}^{2km} \sum_{l=2}^{\lfloor \frac{k}{2} \rfloor} \sum_{r=l}^{k+1-l} \sum_{h=1}^{d_l} \sum_{e=1}^{d_r} c_{ehg}^{rlk} \Delta_2 (x_{1r}^e \otimes x_{2l}^h + x_{1l}^h \otimes x_{2r}^e).$$

Relation (4.14) implies that the sum of (4.17) and (4.18) is equal to 0. Since  $c_{ehg}^{rlk} = c_{heg}^{lrk}$  it follows that (4.20) is equal to

$$(4.21) \quad \begin{aligned} & - \sum_{k=3}^{m-1} \sum_{l=1}^{k-1} \sum_{r=l}^{k+1-l} \sum_{g=1}^{d_k} \sum_{h=1}^{d_l} \sum_{e=1}^{d_r} c_{1g1}^{2km} c_{ehg}^{rlk} \Delta_2 x_{1r}^e \otimes x_{2l}^h \\ & = - \sum_{r=2}^{m-2} \sum_{l=2}^{r-1} \sum_{h=1}^{d_l} \sum_{e=1}^{d_r} \left[ \sum_{k=l+r-1}^{m-1} \sum_{g=1}^{d_k} c_{1g1}^{2km} c_{ehg}^{rlk} \right] \Delta_2 x_{1r}^e \otimes x_{2l}^h. \end{aligned}$$

We use relations (3.5) and the definition of  $\Delta_2$ , and we rearrange the order of summation to show that (4.21) is equal to

$$(4.22) \quad - \sum_{k=2}^{\left[\frac{m}{2}\right]} \sum_{l=k}^{m+1-k} \sum_{g=1}^{d_k} \sum_{h=1}^{d_l} \sum_{r=k}^{l-1} \sum_{e=1}^{d_r} c_{gh1}^{klm} \begin{vmatrix} U_1 (\mathbf{c}_{eg}^{rk})^\dagger & A_{11}^\dagger \\ U_2 (\mathbf{c}_{eg}^{rk})^\dagger & A_{21}^\dagger \end{vmatrix} (x_{1r}^e \otimes x_{2l}^h + x_{1l}^h \otimes x_{2r}^e).$$

If we use (4.4) and (4.14) we see that the sum of (4.19) and (4.22) is 0. Therefore (4.16) is equal to 0, and (4.15) follows.  $\square$

Now we are ready to prove our main result.

*Proof of Theorem 4.1.* Suppose that we are given a basis  $\{\mathbf{c}_{1f}^m, f = 1, 2, \dots, d\}$  for  $\mathcal{N}(\mathcal{S}_m)$ . Then by Lemma 4.3 it follows that we can find vectors  $x_{im}^f \in \mathcal{D}'_i$  and matrices  $C_f^m$  of the form (4.2), where  $C_f^m = \psi_m^{-1} (\mathbf{c}_{1f}^m, \mathbf{c}_{2f}^m)$  for some  $\mathbf{c}_{2f}^m$ , such that (i) and (ii) hold when  $m-1$  is replaced by  $m$ . We apply Lemma 4.4 to show that for  $f = 1, 2, \dots, d$  the vectors  $z_{mf}$  are in  $\mathcal{R}_m$  but not in  $\mathcal{R}_{m-1}$ . They are linearly independent because  $\mathbf{c}_{1f}^m, f = 1, 2, \dots, d$ , are linearly independent. It follows that  $d \leq d_m$  and, because  $d \geq d_m$  by Lemma 4.2, we have  $d = d_m$ . Thus (iii) holds also.  $\square$

## 5. THE STURM-LIOUVILLE CASE

Faierman [8] considers a two-parameter eigenvalue problem involving a class of coupled Sturm-Liouville boundary value problems

$$(5.1) \quad \frac{d}{dt_i} \left( p_i(t_i) \frac{dy_i}{dt_i} \right) + ((-1)^{i+1} \mu_1 a_{i1}(t_i) + (-1)^i \mu_2 a_{i2}(t_i) - q_i(t_i)) y_i = 0, \quad i = 1, 2,$$

$$(5.2) \quad y_i(0) \cos \alpha_i - p_i(0) \frac{dy_i}{dt_i}(0) \sin \alpha_i = 0, \quad 0 \leq \alpha_i < \pi, \quad i = 1, 2,$$

and

$$(5.3) \quad y_i(1) \cos \beta_i - p_i(1) \frac{dy_i}{dt_i}(1) \sin \beta_i = 0, \quad 0 < \beta_i \leq \pi, \quad i = 1, 2,$$

where  $\boldsymbol{\mu} = (\mu_1, \mu_2)$  are parameters,  $\mathcal{I} = [0, 1]$  and  $t_i \in \mathcal{I}$ . As in [8, pp. 2 and 10] we assume that

- (i) for  $i = 1, 2$  the functions  $p_i, q_i, a_{ij}$ ,  $j = 1, 2$ , are real-valued,  $p_i, a_{ij}$  are Lipschitz continuous,  $p_i$  is positive and  $q_i$  is essentially bounded,
- (ii) the function  $\omega(t_1, t_2) = a_{11}(t_1)a_{22}(t_2) - a_{12}(t_1)a_{21}(t_2)$  on  $\mathcal{I}^2$  is not identically 0,
- (iii)  $a_{i1}(t_i) > 0$  for  $t_i \in \mathcal{I}$  and  $i = 1, 2$ .

Weaker assumptions would be possible (cf. [4, 5, 26]) but we adhere to Faierman's assumptions for the purpose of comparison.

Two-parameter problems which satisfy condition (iii) are called *elliptic*. In the literature also a formally stronger condition

(iii') each  $a_{ij}(t_i)$  is positively bounded below on  $\mathcal{I}$

is considered (e.g. [3]). Two-parameter problems which satisfy condition (iii') are called *uniformly elliptic*. We show in Lemma 5.1 that we can transform an elliptic problem to a uniformly elliptic one by a linear substitution in parameters.

To introduce the operators  $A_{ij}$  we take  $H_i = L_2(\mathcal{I})$ . We have

$$(5.4) \quad A_{ij}y_i(t_i) = a_{ij}(t_i)y_i(t_i), \quad i, j = 1, 2,$$

for  $y_i \in H_i$  and

$$(5.5) \quad A_{i0}y_i(t_i) = \frac{d}{dt_i} \left( p_i(t_i) \frac{dy_i}{dt_i} \right) - q_i(t_i)y_i, \quad i = 1, 2,$$

where the domain  $\mathcal{D}(A_{i0})$  consists of all functions  $y_i \in H_i$  such that  $y_i$  and  $\frac{dy_i}{dt_i}$  are absolutely continuous,  $\frac{d^2 y_i}{dt_i^2} \in H_i$  and the boundary conditions (5.2) and (5.3) hold.

**Lemma 5.1.** *Assume that the problem (5.1)–(5.3) is such that (i)–(iii) hold. Then there exists an invertible linear substitution of parameters  $\boldsymbol{\mu}$  such that the problem is uniformly elliptic after the substitution.*

*Proof.* The numerical range

$$\kappa_i = \{ (c_{i1}, c_{i2}) \in \mathbb{R}^2 : \text{for } j = 1, 2 \ c_{ij} = (A_{ij}x_i, x_i) \text{ for some } x_i \in H_i, \|x_i\| = 1 \}$$

is a bounded convex set for  $i = 1, 2$ . By conditions (i) and (iii) it follows that for some  $\epsilon > 0$  we have  $c_{11} > \epsilon$  for all elements of  $\kappa_1$  and  $c_{21} < -\epsilon$  for all elements of  $\kappa_2$ . Then there exist two lines through the origin in  $\mathbb{R}^2$  that separate  $\kappa_1$  and  $\kappa_2$  and we can find an invertible linear map  $\tau$  on  $\mathbb{R}^2$  that maps  $\kappa_1$  and  $\kappa_2$  into the first and the third quadrant, respectively. If we apply  $\tau$  to the parameters  $\boldsymbol{\mu}$  it follows that the transformed problem is uniformly elliptic.  $\square$

We hereafter assume, without loss of generality, that the parameters  $\boldsymbol{\mu}$  are chosen so that condition (iii') holds, i.e., that the two-parameter problem is uniformly elliptic. In order to apply the results of the previous sections we need to verify Assumptions I–III. By [8, Thm. 2.4] it follows that  $\Delta_2$  is a selfadjoint operator with dense domain  $\mathcal{D}$  of those functions in the Sobolev space  $W_2^2(\mathcal{I}^2)$  that satisfy the boundary condition on  $\mathcal{I}^2$  associated with (5.2) and (5.3). For details see [8, pp. 10–30]. Next we have:

**Lemma 5.2.**  $\mathcal{N}(\Delta_0) \cap \mathcal{N}(\Delta_2) = \{0\}$ .

*Proof.* Assume that  $\Delta_0 u = \Delta_2 u = 0$  for some  $u \in \mathcal{D}$ . By (ii) it follows that  $\omega$  is nonzero on some open set  $\Omega \subset \mathcal{I}^2$ . Hence  $u|_\Omega \equiv 0$ . By the unique continuation property for  $\Delta_2$  [8, Prop. 3.1] it follows that  $u \equiv 0$  on  $\mathcal{I}^2$ .  $\square$

Now it follows from [5, Prop. 2.1] that Assumption I holds for an  $\alpha \in \mathbb{R}$ . As in §2 we shall hereafter assume, without loss of generality, that  $\alpha = 0$ , i.e., that  $\Delta_2$  has a bounded inverse. This is achieved by a shift in parameters. Assumption II is a consequence of the following lemma. This is given in a special case in [5, §9.3]. Since the argument there appears to be incomplete we provide a more detailed proof here.



**Lemma 5.3.** *The system*

$$(5.6) \quad \begin{bmatrix} A_{10}^\dagger & A_{11}^\dagger \\ A_{20}^\dagger & A_{21}^\dagger \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} A_{12}^\dagger z \\ A_{22}^\dagger z \end{bmatrix}$$

has a solution  $u = (u_1, u_2) \in \mathcal{D}^2$  for all  $z \in \mathcal{D}$ .

*Proof.* By eliminating  $u_2$  in (5.6) we have

$$(5.7) \quad -\Delta_2 u_1 = \Delta_0 z.$$

By [8, Thm. 2.2] there exists  $u_1 \in \mathcal{D}$  which solves (5.7). Similarly, the elimination of  $u_2$  leads to  $-\Delta_2 u_2 = \Delta_1 z$ , which again is soluble for  $u_2 \in \mathcal{D}$  by [8, Thm. 2.2]. Thus it remains to show that  $u_1$  and  $u_2$  satisfy (5.6).

We shall do this first under the assumption that each  $A_{i0}$  is positive definite. For  $N \in \mathbb{N}$  we define bounded selfadjoint operators  $A_{i0}^{(N)}$  sharing the same eigenvectors and the lowest  $N$  eigenvalues with  $A_{i0}$ . The remaining eigenvalues of  $A_{i0}^{(N)}$  are defined as 1. The definition of  $A_{i0}^{(N)}$  is completed by linearity and continuity to  $H_i$ . We define  $\Delta_j^{(N)}$ ,  $j = 1, 2$ , analogously to  $\Delta_j$  and we note that  $A_{i0}^{(N)}$  and  $\Delta_j^{(N)}$  are uniformly bounded with inverses bounded uniformly in  $N$  for  $i, j = 1, 2$ .

Now consider the system of equations

$$(5.8) \quad A_{i0}^{(N)\dagger} z_0 = A_{i1}^\dagger z_1 + A_{i2}^\dagger z_2, \quad i = 1, 2,$$

where  $z_2 = \left(\Delta_1^{(N)}\right)^{-1} w$  for a fixed  $w \in H$ . Proceeding as in the proof of [2, Thms. 4.2 and 6.1] we obtain

$$z_0 = -\left(\Delta_2^{(N)}\right)^{-1} \Delta_0 \left(\Delta_1^{(N)}\right)^{-1} w \quad \text{and} \quad z_1 = -\left(\Delta_2^{(N)}\right)^{-1} w,$$

so (5.8) yields

$$(5.9) \quad A_{i0}^{(N)\dagger} x_N = -A_{i1}^\dagger \left(\Delta_2^{(N)}\right)^{-1} w - A_{i2}^\dagger \left(\Delta_1^{(N)}\right)^{-1} w, \quad i = 1, 2,$$

where  $x_N = \left(\Delta_2^{(N)}\right)^{-1} \Delta_0 \left(\Delta_1^{(N)}\right)^{-1} w$ . Now [2, Lemma 4.1] shows that the right-hand side in (5.9) has limit  $-A_{i1}^\dagger \Delta_2^{-1} w - A_{i2}^\dagger \Delta_1^{-1} w = y_i$ , say, as  $N \rightarrow \infty$ . Thus  $A_{i0}^{(N)\dagger} x_N$  converges to  $y_i$ . Since the inverses of  $A_{i0}^{(N)}$  are bounded uniformly in  $N$  it follows that  $x_N - \left(\left(A_{i0}^{(N)}\right)^{-1}\right)^\dagger y_i \rightarrow 0$  as  $N \rightarrow \infty$ . Similarly, since the inverses of  $\left(\Delta_2^{(N)}\right)^{-1}$  are bounded uniformly in  $N$  it follows that  $x_N \rightarrow \Delta_2^{-1} \Delta_0 \Delta_1^{-1} w$  as  $N \rightarrow \infty$  (cf. [2, Lemma 4.1]). Thus

$$\Delta_2^{-1} \Delta_0 \Delta_1^{-1} w = \left(A_{i0}^{-1}\right)^\dagger y_i,$$

whence

$$A_{i0}^\dagger \Delta_2^{-1} \Delta_0 \Delta_1^{-1} w = -A_{i1}^\dagger \Delta_2^{-1} w - A_{i2}^\dagger \Delta_1^{-1} w.$$

Now put  $w = \Delta_1 z$  to obtain (5.6) for  $u_1 = -\Delta_2^{-1} \Delta_0 z$  and  $u_2 = -\Delta_2^{-1} \Delta_1 z$ . To remove the restriction that  $A_{i0}$  are uniformly positive we use the technique of [6, Thms. 3.1 and 3.2].  $\square$

In the rest of this section we discuss the eigenvalues and associated root subspaces of (5.1)–(5.3). If there exist nonzero functions  $y_i \in \mathcal{D}(A_{i0})$ ,  $i = 1, 2$ , which solve (5.1)–(5.3) for a pair  $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{C}^2$ , then  $\boldsymbol{\mu}$  is called an *eigenvalue for the problem* (5.1)–(5.3). The function  $w(t_1, t_2) = y_1(t_1)y_2(t_2)$  is such that  $(\Delta_i - \mu_i \Delta_0)w = 0$  for  $i = 1, 2$ . Because we assume that  $\Delta_2$  has a bounded inverse, and in particular, is one-to-one, we have that  $\mu_2 \neq 0$ . Then it follows for each eigenvalue  $\boldsymbol{\mu}$  of (5.1)–(5.3) that  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1) = \left(\frac{1}{\mu_2}, \frac{\mu_1}{\mu_2}\right)$  is an eigenvalue of the associated two-parameter system

$$(5.10) \quad W_i(\boldsymbol{\lambda}) = A_{i0}\lambda_0 + A_{i1}\lambda_1 + A_{i2}, \quad i = 1, 2.$$

Here  $A_{ij}$  are given by (5.4) and (5.5). We shall hereafter keep the notation  $\boldsymbol{\mu}$  and  $\boldsymbol{\lambda}$  as above, so if  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1)$  is an eigenvalue of (5.10), then  $\boldsymbol{\mu} = (\mu_1, \mu_2) = \left(\frac{\lambda_1}{\lambda_0}, \frac{1}{\lambda_0}\right)$  is the corresponding eigenvalue of (5.1)–(5.3). Note that  $\mu_2$  and  $\lambda_0$  are always nonzero.

The eigenvalues of (5.1)–(5.3) form a countably infinite set of points without accumulation. There are only finitely many nonsemisimple eigenvalues and finitely many nonreal eigenvalues. Nonreal eigenvalues appear in conjugate pairs. See [8, Thms. 6.1 and 6.2] for these results.

The problem (5.1)–(5.3) for  $i = 1$  or  $i = 2$  has for a fixed value of  $\boldsymbol{\mu}$  at most a one-dimensional space of solutions. If  $\boldsymbol{\mu} \in \mathbb{R}^2$ , then  $W_i(\boldsymbol{\mu})$  is a selfadjoint operator. If  $\boldsymbol{\mu} \in \mathbb{C}^2 \setminus \mathbb{R}^2$ , then  $W_i(\boldsymbol{\mu})^* = W_i(\overline{\boldsymbol{\mu}})$ , where  $\bar{\phantom{x}}$  denotes complex conjugation. Since  $W_i(\boldsymbol{\mu})y_i = 0$  holds if and only if  $W_i(\overline{\boldsymbol{\mu}})\overline{y}_i = 0$ , it follows that  $\dim \mathcal{N}(W_i(\boldsymbol{\mu})) = \dim \mathcal{N}(W_i(\boldsymbol{\mu})^*) = 1$ . We are now ready to establish Assumption III.

**Lemma 5.4.** *Each eigenvalue  $\boldsymbol{\mu}$  of the problem (5.1)–(5.3) is geometrically simple of finite ascent, and both  $W_i(\boldsymbol{\lambda})$ ,  $i = 1, 2$ , are Fredholm operators of index 0.*

*Proof.* The remarks preceding this lemma together with [16, Example III.5.14] establish everything except finite ascent, and the latter follows from [8, Thm. 6.6].  $\square$

There are various differences between the approach of [8] and the one here. In [8] the eigenvalues  $\boldsymbol{\mu} = (\mu_1, \mu_2)$  are considered. We use instead the eigenvalues  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1)$ . In [8] root subspaces are considered only with respect to the pencil  $\Delta_2 - \mu_2 \Delta_0$ , while we consider joint root subspaces for pencils  $\Delta_i - \lambda_i \Delta_2$ ,  $i = 0, 1$ . When bases for root subspaces are constructed our approach significantly reduces technical difficulties since all the geometric eigenspaces are now one-dimensional, i.e., all the joint eigenvalues are geometrically simple.

In order to be able to compare our results with those of [8] we need the following auxiliary results:

**Lemma 5.5.** *If  $\lambda_0 \neq 0$  and  $z_1, z_2, \dots, z_m \in \mathcal{D}$  are such that*

$$(5.11) \quad (\Delta_0 - \lambda_0 \Delta_2)z_j = \Delta_2 z_{j-1}, \quad j = 1, 2, \dots, m,$$

*where  $z_0 = 0$ , then*

$$u_j = \sum_{k=0}^{j-1} (-1)^k \binom{j}{k} \lambda_0^{j+k} z_{k+1}$$

*are such that*

$$(5.12) \quad (\Delta_2 - \mu_2 \Delta_0)u_j = \Delta_0 u_{j-1}, \quad j = 1, 2, \dots, m,$$

*where  $u_0 = 0$  and  $\mu_2 = \lambda_0^{-1}$ .*

*Proof.* By induction on  $j$  we prove that (5.11) implies

$$(\Delta_2 - \mu_2 \Delta_0) z_j = \sum_{k=1}^{j-1} (-1)^k \mu_2^{k+1} \Delta_0 z_{j-k}, \quad j = 1, 2, \dots, m.$$

Then (5.12) follows by a straightforward calculation which we omit.  $\square$

We call a chain of vectors  $u_1, u_2, \dots, u_m$  such that (5.12) holds a *Jordan chain* for the pencil  $\Delta_2 - \mu_2 \Delta_0$ .

**Corollary 5.6.** *Let  $\mu_2 = \lambda_0^{-1}$ . Then  $\mu_2$  is a semi-simple eigenvalue for the pencil  $\Delta_2 - \mu_2 \Delta_0$  if and only if  $\lambda_0$  is a semi-simple eigenvalue for  $\Delta_0 - \lambda_0 \Delta_2$ .*

**Lemma 5.7.** *Suppose that  $\mathcal{R}_m$  is the  $m$ -th root subspace for (5.10) at an eigenvalue  $\lambda = (\lambda_0, \lambda_1)$  (see (1.7)). Then it follows that  $(\Delta_2 - \mu_2 \Delta_0) \mathcal{R}_m \subset \Delta_0 \mathcal{R}_{m-1}$ . In particular, if  $\mathcal{B}_m$  is the basis for  $\mathcal{R}_m$  given in Theorem 4.1, then*

$$\begin{aligned} & (\Delta_2 - \mu_2 \Delta_0) z_{mj} \\ &= \sum_{l=1}^{m-1} \left( (-1)^l \mu_2^l \sum_{k=1}^l \sum_{g=1}^{d_k} \gamma_{jg}^{mk} \Delta_0 z_{kg} \right), \end{aligned}$$

where  $\gamma_{jg}^{mk}$  are defined inductively by

$$\gamma_{jg}^{k+1,k} = c_{1gj}^{2,k,k+1} \quad \text{and} \quad \gamma_{jg}^{k+r,k} = \sum_{l=k+1}^{k+r-1} \sum_{h=1}^{d_l} c_{1hj}^{2,l,k+r} \gamma_{hg}^{lk}$$

for  $r = 2, 3, \dots$

*Proof.* If we multiply the relation  $(\Gamma_0 - \lambda_0 I) \mathcal{R}_m \subset \mathcal{R}_{m-1}$  by  $-\lambda_0^{-1} \Delta_2$  on the left, we have that  $(\Delta_2 - \mu_2 \Delta_0) \mathcal{R}_m \subset \Delta_2 \mathcal{R}_{m-1}$ . Then we prove by induction on  $m$  that  $(\Delta_2 - \mu_2 \Delta_0) \mathcal{R}_m \subset \Delta_0 \mathcal{R}_{m-1}$ . The second part of the lemma also follows by induction on  $m$ . We omit the details.  $\square$

Now we compare our results with those of [8]. First we consider the case of a *real eigenvalue*. We assume that  $\lambda \in \mathbb{R}^2$  is an eigenvalue of (5.10), that  $z_1 = x_{11} \otimes x_{21}$  is a corresponding eigenvector and that

$$S_2 = \begin{bmatrix} \langle A_{11} x_{11}, x_{11} \rangle & \langle A_{11} x_{11}, x_{11} \rangle \\ \langle A_{20} x_{21}, x_{21} \rangle & \langle A_{21} x_{21}, x_{21} \rangle \end{bmatrix}.$$

**Proposition 5.8.** *For  $\lambda \in \mathbb{R}^2$  the following are equivalent:*

- (i)  $\lambda$  is a semi-simple eigenvalue,
- (ii)  $\langle \Delta_0 z_1, z_1 \rangle \neq 0$ ,
- (iii) the rank of the matrix  $S_2$  is equal to 2.

*Proof.* Conditions (i) and (ii) are equivalent by [8, Thm. 6.15] and Corollary 5.6. The equivalence of (i) and (iii) follows from [5, Thm. 6.3].  $\square$

**Proposition 5.9.** *For  $\lambda \in \mathbb{R}^2$  the following are equivalent:*

- (i)  $\lambda$  is not a semi-simple eigenvalue,
- (ii)  $\lambda$  is a nonderogatory eigenvalue, but not algebraically simple,
- (iii)  $\langle \Delta_0 z_1, z_1 \rangle = 0$ ,
- (iv) the rank of the matrix  $S_2$  is less than or equal to 1.

*Proof.* If  $\text{rank } S_2$  is less than 2, then it has to be equal to 1. Namely, it follows from (iii') that  $\langle A_{i1}x_{i1}, x_{i1} \rangle \neq 0$  for  $i = 1, 2$ , and therefore  $S_2 \neq 0$ . Then if  $\lambda$  is not semi-simple, it has to be nonderogatory. The equivalence of conditions (i)–(iv) now follows from Proposition 5.8.  $\square$

By the above propositions it follows that a basis  $\mathcal{B}$  for a root subspace  $\mathcal{R}$  at a real eigenvalue  $\lambda$  is described by Theorem 2.2. Then a root subspace  $\mathcal{M}_{\mu_2}$  (in the notation of [8]) at a real eigenvalue  $\mu_2$  is equal to the direct sum of the root subspaces corresponding to the eigenvalues  $\lambda = (\lambda_0, \lambda_1)$  with  $\lambda_0 = \frac{1}{\mu_2}$ . Lemma 5.5 implies that a basis for  $\mathcal{M}_{\mu_2}$  is thus obtained by Theorem 2.2 as the union of all the bases of the root subspaces  $\mathcal{R}(\lambda)$  with  $\lambda_0 = \frac{1}{\mu_2}$ . This yields Theorems 3.5 and 6.18 of [8] and also the Sturm-Liouville case of [3, Thm. 3.1]. The interested reader will find a numerical implementation of the algorithm for computation of a basis at a nonderogatory eigenvalue in [23].

Assume now that  $\lambda = (\lambda_0, \lambda_1)$  is a *nonreal eigenvalue* of (5.10). Then  $\lambda_0$  is nonreal [8, p. 99]. We write  $z_1 = x_{11} \otimes x_{21}$  for an eigenvector corresponding to  $\lambda$  and

$$S_2 = \begin{bmatrix} \langle A_{10}x_{11}, \bar{x}_{11} \rangle & \langle A_{11}x_{11}, \bar{x}_{11} \rangle \\ \langle A_{20}x_{21}, \bar{x}_{21} \rangle & \langle A_{21}x_{21}, \bar{x}_{21} \rangle \end{bmatrix}.$$

**Proposition 5.10.** *For a nonreal eigenvalue  $\lambda$  the following are equivalent:*

- (i)  $\lambda$  is semi-simple,
- (ii)  $\langle \Delta_0 z_1, \bar{z}_1 \rangle \neq 0$ ,
- (iii) the rank of the matrix  $S_2$  is equal to 2.

*Proof.* Conditions (i) and (ii) are equivalent by [8, Thm. 6.14] and Corollary 5.6. The equivalence of (i) and (iii) follows from [5, Thm. 6.3].  $\square$

If  $\lambda$  is not semi-simple and  $\text{rank } S_2$  is equal to 1, then  $\lambda$  is nonderogatory and Theorem 2.2 gives a basis for the corresponding root subspace. By Lemma 5.7 it follows that the union of bases, of root subspace at the eigenvalues  $\lambda$  for which  $\lambda_0 = \frac{1}{\mu_2}$ , is a basis for the root subspace  $\mathcal{M}_{\mu_2}$  (in notation of [8]). If  $\mathcal{M}_{\mu_2}$  is a root subspace at a nonreal eigenvalue  $\mu_2$ , it is equal to the direct sum of the root subspaces corresponding to the eigenvalues  $\lambda = (\lambda_0, \lambda_1)$  with  $\lambda_0 = \frac{1}{\mu_2}$ . If all these eigenvalues  $\lambda$  are nonderogatory, then Lemma 5.5 implies that a basis for  $\mathcal{M}_{\mu_2}$  is obtained by Theorem 2.2. This yields Theorems 6.9 and 6.19 of [8] and it also gives a basis for those  $j$  in [8, pp. 118–122] for which  $s_2 < j \leq s_1$ . The eigenvalues  $\lambda$  that correspond to the remaining  $j$  (i.e., such that  $1 \leq j \leq s_2$ ) are derogatory, i.e.,  $S_2 = 0$ . Now we use Theorem 4.1 to obtain a basis  $\mathcal{B}$  for the root subspace at  $\lambda$ . The main result about bases for  $\mathcal{M}_{\mu_2}$  in [8] is Theorem 6.20. The basis that Faierman gives there is a union of bases of root subspaces  $\mathcal{R}(\lambda)$  over all  $\lambda$  with  $\lambda_0 = \frac{1}{\mu_2}$ . These are of two types: for  $s_2 < j \leq s_1$  the eigenvalues  $\lambda$  are nonderogatory and our Theorem 2.2 gives the required properties described in [8, Lemma 6.9]. The other eigenvalues with  $1 \leq j < s_2$  are derogatory. Bases for  $\mathcal{R}(\lambda)$  are described in Conjecture 6.1 of [8], which is established in [8] only when the ascent of  $\Gamma_0$  is at most 3. Our basis in Theorem 4.1 is constructed in all cases of finite ascent. We hope to discuss the precise relationship with [8, Conj. 6.1] in a separate publication.

Finally, we remark that bases for  $\mathcal{R}(\lambda)$  can be used to obtain completeness and expansion results as carried out in [8, Thms. 6.7 and 6.8] and [8, §6.7], respectively.

Then orthonormal bases are important. These can be constructed by a Gram-Schmidt type orthogonalization applied inductively for each  $m$  in the construction of a basis for  $\mathcal{R}(\lambda)$ . Additional care needs to be taken for the root subspaces at pairs of conjugate eigenvalues; for other pairs of eigenvalues the root subspaces are orthogonal by [8, Thm. 6.1].

## REFERENCES

- [1] F.V. Atkinson. *Multiparameter Eigenvalue Problems*. Academic Press, 1972. MR **56**:9291
- [2] P.A. Binding. Left Definite Multiparameter Eigenvalue Problems. *Trans. Amer. Math. Soc.*, 272:475–486, 1982. MR **83h**:47015
- [3] P.A. Binding. Multiparameter Root Vectors. *Proc. Edin. Math. Soc.*, 32:19–29, 1989. MR **90a**:47048
- [4] P.A. Binding and P.J. Browne. Asymptotics of Eigencurves for Second Order Ordinary Differential Equations I. *J. Diff. Eq.*, 88:30–45, 1990. MR **92b**:34101
- [5] P.A. Binding and T. Košir. Second Root Vectors for Multiparameter Eigenvalue Problems of Fredholm Type. *Trans. Amer. Math. Soc.*, 348:229–249, 1996. MR **96f**:35121
- [6] P.A. Binding and K. Seddighi. Elliptic Multiparameter Eigenvalue Problems. *Proc. Edin. Math. Soc.*, 30:215–228, 1987. MR **88g**:47036
- [7] G. Birkhoff. *Lattice Theory*, volume 25 of *Amer. Math. Soc. Colloq. Publ.*, 3rd edition, Amer. Math. Soc., Providence, 1973. MR **37**:2638
- [8] M. Faierman. *Two-parameter Eigenvalue Problems in Ordinary Differential Equations*, volume 205 of *Pitman Research Notes in Mathematics*. Longman Scientific and Technical, Harlow, U.K., 1991. MR **93b**:47095
- [9] G.A. Gadzhiev. *Introduction to Multiparameter Spectral Theory* (in Russian), Azerbaijan State University, Baku, 1987.
- [10] I.C. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications*. Wiley-Interscience, 1986. MR **88a**:15001
- [11] L. Grunenfelder and T. Košir. An Algebraic Approach to Multiparameter Eigenvalue Problems. *Trans. Amer. Math. Soc.*, 348:2983–2998, 1996. MR **96j**:47015
- [12] L. Grunenfelder and T. Košir. Coalgebras and Spectral Theory in One and Several Parameters. In the series *Operator Theory: Adv. and Appl.*, 87: 177–192, Birkhäuser-Verlag, 1996. MR **97k**:47019
- [13] L. Grunenfelder and T. Košir. Geometric Aspects of Multiparameter Spectral Theory, *Trans. Amer. Math. Soc.*, 350:2525–2546, 1998. MR **98h**:13032
- [14] H.(G.A.) Isaev. *Lectures on Multiparameter Spectral Theory*. Dept. of Math. and Stats., University of Calgary, 1985.
- [15] G.A. Isaev and A.S. Fainshtein. *The Taylor Spectrum and Multiparameter Spectral Theory for Systems of Operators*. *Soviet Math. Dokl.*, 36:413–417, 1988. MR **88m**:47028
- [16] T. Kato. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der math. Wiss.* Springer-Verlag, second edition, 1984. MR **53**:11389
- [17] T. Košir. *Commuting Matrices and Multiparameter Eigenvalue Problems*. PhD thesis, Dept. of Math. and Stats., University of Calgary, 1993.
- [18] T. Košir. On the Structure of Commutative Matrices. *Lin. Alg. Appl.*, 187:163–182, 1993. MR **94i**:15011
- [19] T. Košir. The Finite-Dimensional Multiparameter Spectral Theory: The Nonderogatory Case. *Lin. Alg. Appl.*, 212/213:45–70, 1994. MR **95k**:15020
- [20] T. Košir. On the Structure of Commutative Matrices II. *Lin. Alg. Appl.*, 261:293–305, 1997. MR **98k**:15018
- [21] T. Košir. Root Vectors for Geometrically Simple Multiparameter Eigenvalues, to appear in *Int. Equat. Oper. Theory*.
- [22] J. Meixner, F. W. Schäfke, and G. Wolf. *Mathieu Functions and Spheroidal Functions and Their Mathematical Foundations*, volume 873 of *Lect. Notes in Math.*, Springer-Verlag, 1980. MR **83b**:33013
- [23] B. Plestenjak. A Numerical Algorithm for Computing a Basis for the Root Subspace at a Nonderogatory Eigenvalue of a Multiparameter System, *Lin. Alg. Appl.*, 285:257–276, 1998. MR **2000a**:65048

- [24] B.D. Sleeman. *Multiparameter Spectral Theory in Hilbert Space*, volume 22 of *Pitman Research Notes in Mathematics*. Pitman Publ. Ltd., London U.K., Belmont U.S.A., 1978. MR **81h**:47004
- [25] A.E. Taylor and D.C. Lay. *Introduction to Functional Analysis*. Wiley, New York, second edition, 1980. MR **81b**:46001
- [26] H. Volkmer. *Multiparameter Eigenvalue Problems and Expansion Theorems*, volume 1356 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, New York, 1988. MR **90d**:47021

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF CALGARY, CALGARY, ALBERTA, CANADA T2N 1N4

*E-mail address:* `binding@ucalgary.ca`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LJUBLJANA, JADRANSKA 19, 1000 LJUBLJANA, SLOVENIA

*E-mail address:* `tomaz.kosir@fmf.uni-lj.si`