NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilming. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui ont déjà l'objet d'un droit d'auteur (articles, revue, tests publiés, etc.) ne sont pas microfilmmés.

La reproduction, même partelle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.
The Inverse Spectral Generalized Matrix Eigenvalue Problem

By

Moussa Boucekkine

Thesis Submitted
To the School of Graduate Studies
In partial fulfillment of the requirements for the degree of Master of Science in Mathematics

at the
University of Ottawa
MAY 1988

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-46717-7
To My Parents
Contents

Abstract .................................................. ii

Acknowledgements ......................................... iii

Introduction .............................................. 1

1 Preliminaries ........................................... 4

1.1 Definitions and Notations ............................ 5

1.2 Necessary and Sufficient Conditions for the Positive Definite-ness ........................................ 8

1.3 Reduction of Symmetric Matrices to Canonical Forms .......................................................... 12

1.4 The Generalized Symmetric Eigenvalue Problem ................................................................. 18

2 Orthogonal polynomials .................................. 24
2.1 Sturm Liouville Difference Equations With an Indefinite Weight 25

2.2 Lagrange - Type Identity .......................... 28

2.3 Orthogonality ........................................ 30

2.4 The Riemann-Stieltjes Integral .................. 33

2.5 The Spectral Function .............................. 35

3 The Real 2 by 2 Case ................................. 36

3.1 Necessary Condition for the Existence of a Solution ... 37

3.2 Construction of a Solution ......................... 40

4 The Real 3 by 3 Case ................................. 50

4.1 Necessary Conditions for Existence of a Solution ... 52

4.2 Construction of a Solution ......................... 55

Bibliography .............................................. 65
Abstract

We seek a positive definite Jacobi symmetric matrix $A$, such that the generalized eigenvalue problem

$$Ax = \lambda Bx$$  \hspace{1cm} (0.1)

holds for some $x \neq 0$, where the eigenvalues of (0.1) are distinct and $B$ is a real nonsingular diagonal indefinite matrix. In particular we discuss the 2 by 2 and the 3 by 3 cases.

We also present a proof for the non-uniqueness of the solution of this problem in the above cases.
Acknowledgements

I would like to express my sincere thanks to my thesis supervisor, Dr. Angelo B. Mingarelli, not only for suggesting the thesis topic but also for providing me with advice, encouragement and very constructive criticism. He always found the time for productive discussion.

I am also very grateful for his generous assistance and patient guidance throughout the preparation and writing of this thesis. I would also like to express my appreciation to Dr. Remi Vaillancourt for giving me an opportunity to discuss with him certain aspects of my thesis, and my wife who did everything conceivable to keep my morale high.

Finally, I would also like to express my gratitude to Dr. B. Tomiuk and Dr. L. Nel for their helpful criticisms.
Introduction

Inverse eigenvalue problems have always played a very important role in physics, geophysics and arise often in applied mathematics and they are treated by many authors; consult [14] for continuous problems and [9], [10] [5] for the matrix versions.

Among the most common inverse eigenvalue problems we find the following which are closely related to the spectral theory of symmetric matrices. Let A be a symmetric matrix.

a) Find a real diagonal matrix V such that the spectrum of A + V is given.
b) Find a diagonal matrix D such that the spectrum of AD is given (here A is assumed to be positive definite).

The first problem is called the inverse additive eigenvalue problem, while the second one the multiplicative eigenvalue problem. The additive problem is dealt with in [9] and the multiplicative problem in [10]. Both problems (a) and (b) are extended to the complex case that is, when A is a complex hermitian matrix, in (a) see [11] for (a) and [5] for (b).

The two problems are the discrete analogs of the following inverse eigen-
value problems. The simplest formulation is

a') find a potential \( q(x) \) such that the operator \( l(y) = -y'' + q(x)y \) with appropriate boundary conditions, possesses a prescribed spectrum.

b') find a density \( p(x) \) such that the operator \(-y''/p(x)\) with appropriate boundary conditions, possesses a prescribed spectrum.

The other matrix version is the following inverse generalized eigenvalue problem of

\[ Ax = \lambda B x \quad (0.2) \]

where \( A^T = A \in \mathbb{R}^{n \times n} \) and \( B^T = B \in \mathbb{R}^{n \times n} \).

When \( B = I \), the identity matrix, [8] gave the construction of a Jacobi matrix \( A \) from data satisfying (0.2).

For \( A \) a tridiagonal symmetric matrix and \( B \) diagonal (0.2) can be written as a three-term recurrence relation with appropriate boundary conditions (chapter 2). For a given \( B \) positive definite and a given spectral function \( \tau(\lambda) \) [2], that is, the spectrum of (0.2), and the normalization constants (2.13), [2] gave the construction of a tridiagonal symmetric matrix \( A \) satisfying (0.2) and the solution is then unique.

The intent of this work is to consider the indefinite case i.e., with \( B \) an indefinite diagonal matrix we seek a positive definite Jacobi symmetric matrix \( A \) satisfying (0.2). In this thesis we restrict ourselves to the 2 by 2 and the 3 cases. We will show in chapters 3 and 4 that the solution is generally not unique and we obtain a finite number of different matrices \( A \)
each satisfying (0, 2).

The introductory chapter 1 provides a survey of preliminary results on symmetric matrices and introduces the basic concept of the generalized matrix eigenvalue problem (0, 2). Chapter 2 is devoted to a special type of recurrence relation, but many of the results are more familiar in the context of orthogonal polynomials.

In chapters 3 and 4 the construction of the solution for the 2 by 2 and the 3 by 3 cases respectively is described.
Chapter 1

Preliminaries

There is a very extensive literature dealing with the subject of real symmetric matrices, see [3], [7], [17].

The purpose of this opening chapter is to give some preliminary results concerning real symmetric matrices and the symmetric generalized eigenvalue problem.

In section 1 we give some definitions and notations. In section 2 we give some results on the positive definiteness of real quadratic forms and their corresponding matrices, while in section 3 we shall study the reduction of real symmetric matrices to canonical forms.

In section 4 we shall be considering the generalized eigenvalue problem

\[ Ax = \lambda Bx \]

when \( A^T = A \in R^{n \times n} \) and \( B^T = B \in R^{m \times m} \).
1.1 Definitions and Notations

\( \mathbb{R}^{n \times n} \) denotes the vector space of all \( n \times n \) real matrices \( A \in \mathbb{R}^{n \times n} \)

\[
A = (a_{ij}) = \begin{bmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{bmatrix}
\]

We denote also by \( A^T = (a_{ji}) \) the transpose of \( A \). For \( A \in \mathbb{R}^{n \times n} \) we have

\[
det(A) = \sum_{j=1}^{n} (-1)^{j+i} a_{1j} det(A_{1j})
\]

where \( A_{1j} \) is an \((n - 1) \times (n - 1)\) matrix obtained by deleting the first row and \( j - th \) column of \( A \). Useful properties of the determinant include

for \( A, B \in \mathbb{R}^{n \times n} \)

- (i) \( \det(AB) = \det(A) \det(B) \)

- (ii) \( \det(A^T) = \det(A) \)

- (iii) \( \det(A) \neq 0 \iff A \text{ nonsingular} \)

There are two important subspaces associated with a matrix \( A \in \mathbb{R}^{n \times n} \). The range of \( A \) is defined by

\[
R(A) = \{ y \in \mathbb{R}^n \text{ such that } y = Ax, \text{ for each } x \in \mathbb{R}^n \}
\]
and the null space of $A$ by

$$N(A) = \{ x \in \mathbb{R}^n \text{ such that } Ax = 0 \}$$

The rank of a matrix $A$ is defined by

$$\text{rank}(A) = \dim(R(A)).$$

It can be shown that $\text{rank}(A) = \text{rank}(A^T)$, and thus, the rank of a matrix equals the maximal number of independent rows or columns.

For $A \in \mathbb{R}^{n \times n}$, $\dim(N(A)) + \text{rank}(A) = n$, and the following are equivalent:

- (i) $A$ is nonsingular

- (ii) $N(A) = 0$

- (iii) $\text{rank}(A) = n$.

We say that $A$ is

diagonal if $a_{ij} = 0$ whenever $i \neq j$

tridiagonal if $a_{ij} = 0'$ whenever $|i - j| > 1$

A special notation is convenient for diagonal matrices, if $A \in \mathbb{R}^{n \times n}$ and we write:

$$A = \text{diag}(\alpha_1, \ldots, \alpha_n)$$

then $A = (a_{ij})$ is diagonal and $a_{ii} = \alpha_i$ for $i = 1, \ldots, n$.

There are several important types of square matrices. We say that $A \in$
\( \mathbb{R}^{n \times n} \) is

- symmetric if \( A^T = A \)
- positive definite if \( x^T A x > 0, \; 0 \neq x \in \mathbb{R}^n \) and \( A = A^T \)
- nonnegative definite if \( x^T A x \geq 0, x \in \mathbb{R}^n \) and \( A = A^T \)
- indefinite if \( (x^T A x)(y^T A y) < 0 \), for some \( x, y \in \mathbb{R}^n \)
- orthogonal if \( A^T A = I_n \)
- positive if \( a_{ij} > 0 \) for all \( i \) and \( j \)
- non-negative if \( a_{ij} \geq 0 \) for all \( i \) and \( j \)
- diagonally dominant if \( |a_{ii}| > \sum_{i \neq j} |a_{ij}| \) for all \( i \).

An eigenvector \( x \) of a matrix \( A \) is a nonzero vector satisfying the equation

\[
Ax = \lambda x
\]

where \( \lambda \) is called the eigenvalue of \( A \).

**Definition**

Two matrices \( A, B \) are said to be similar if there exists a nonsingular matrix \( D \), such that \( B = D^{-1} A D \).

Similar matrices represent the same linear operator but refer to different basis systems. This is equivalent to saying that every linear operator in an \( n \)-dimensional vector space corresponds to a certain class of similar matrices. It is important to note that the eigenvalues of a matrix remain invariant under similarity transformations, on the other hand, the eigenvectors are transformed in accordance with the new basis.

To every symmetric matrix \( A \) belongs a corresponding quadratic form
\[(Ax, x) = \sum_{i,j=1}^{n} a_{ij}x_ix_j \quad x = (x_1, \ldots, x_n).\]

The quadratic form is called positive (negative) definite if for every \(x \neq 0\)

\[(Ax, x) = \sum_{i,j=1}^{n} a_{ij}x_ix_j > 0 \quad (< 0).\]

The corresponding matrix is termed positive (negative) definite.

If

\[(Ax, x) \geq 0, \quad (\leq 0)\]

we say that the quadratic form and \(A\) are positive (negative) semidefinite.

1.2 Necessary and Sufficient Conditions for the Positive Definiteness

This section examines necessary and sufficient conditions for the positive definiteness of a given real symmetric matrix, i.e., (a given real quadratic form). First we state some results concerning necessary conditions.

Theorem 1.1 A positive definite matrix must have nonvanishing positive diagonal elements [17. p.16].
A matrix with one or more vanishing or negative diagonal elements cannot be positive definite.

**Theorem 1.2** The elements of a positive definite matrix $A = (a_{ij})$ must fulfill the relation

$$a_{ij}^2 < a_{ii}a_{jj} \quad \text{for all } i \neq j.$$  

[17. p.16]

**Theorem 1.3** The largest element in absolute value of a positive definite matrix $A$ must lie on the diagonal.

**Proof**

The converse statement, that the largest element in absolute value of a positive definite matrix lies outside the diagonal, is contradictory to Theorem 1.2.

**Theorem 1.4** If for a given symmetric matrix $A = (a_{ij})$—with positive diagonal elements, its elements fulfill the \( \binom{n}{2} \) conditions

$$a_{ii}a_{kk} > \left( \sum_{i=1, i \neq j}^{n} |a_{ij}| \right) \left( \sum_{j=1, j \neq k}^{n} |a_{kj}| \right), \quad i \neq k, \ i, k = 1, 2, \ldots, n$$  

then $A$ is positive definite [17. p.120].
From the relation (1.7) in the third section it follows that

**Theorem 1.5** A necessary and sufficient condition that A be positive definite is that all the eigenvalues of A be positive [3. p.34].

Although this result (Theorem 1.5) is of theoretical value, it is relatively difficult to verify. The reason why it is not useful in applications is that the numerical determination of the eigenvalues of a matrix of large dimension is a very difficult matter and any direct attempt based upon a straight forward expansion of the determinant $|A - \lambda I|$ is surely destined for failure because of the extraordinarily large number of terms appearing in the expansion of the determinant. A determinant of order $n$ has $n!$ terms in its complete expansion, for example $20! = 2,433 \times 10^{15}$, so it is clear that direct methods cannot be applied. For this purpose we give the following theorem.

**Theorem 1.6** A necessary and sufficient set of conditions that A be positive definite is that the following relations hold:

$$D_k > 0, \quad k = 1, 2, \ldots, n.$$  \hspace{1cm} (1.1)

where

$$D_k = |a_{ij}|, \quad i, j = 1, 2, \ldots, k.$$  \hspace{1cm} (1.2)

For the proof see [3. pp.74-75], [6. p.306].

We see that we can represent a quadratic form as a sum of squares by the following theorem.

10
Theorem 1.7 Provided that no $D_k$ is equal to zero, we may write

\[ (Ax, x) = \sum_{k=1}^{n} \left( \frac{D_k}{D_{k-1}} \right) y_k^2 \quad D_0 = 1 \]

where

\[ y_k = x_k + \sum_{j=k+1}^{n} c_{kj}x_j, \quad k = 1, 2, \ldots, n - 1 \]

\[ y_n = x_n \]

The $c_{ij}$ are rational functions of the $a_{ij}$ [3, p.75].

Also, as a corollary of theorem 1.7 we have the

Corollary In a positive definite quadratic form $(Ax, x) = \sum_{i,j=1}^{n} a_{ij}x_i x_j$, all the principal minors of the coefficient matrix are positive:

\[ A \left( \begin{array}{cccc} i_1 & i_2 & \cdots & i_p \\ i_1 & i_2 & \cdots & i_p \end{array} \right) > 0 \quad (1 \leq i_1 < i_2 < \ldots < i_p \leq n, \ p = 1, 2, \ldots, n). \]

NOTE: If the successive principal minors are non-negative i.e.,

\[ D_1 \geq 0, \ D_2 \geq 0, \ \ldots, D_n \geq 0 \quad (1.3) \]

it does not follow that $(Ax, x)$ is positive semi-definite. For, the form

\[ a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \]

with

\[ a_{11} = a_{12} = 0, \ a_{22} < 0 \]
satisfy (1.3) but it is not positive semi-definite. However we have the following result by which we end this section.

**Theorem 1.8** A quadratic form \( (Ax, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \) is positive semi-definite if and only if all the principal minors of its coefficient matrix are non-negative:

\[
A \begin{pmatrix}
i_1 & i_2 & \cdots & i_p \\
i_1 & i_2 & \cdots & i_p
\end{pmatrix} \geq 0 \quad (1 \leq i_1 < i_2 < \ldots < i_p \leq n, \; p = 1, 2, \ldots, n).
\]

[6. p.307]

### 1.3 Reduction of Symmetric Matrices to Canonical Forms

First let us give two fundamental results of symmetric matrices upon which the entire analysis of them hinges.

**Theorem 1.9** The eigenvalues of a symmetric matrix \( A \) are real

**Proof**

If \( \lambda \) satisfies

\[
Ax = \lambda x
\]

for some \( x \neq 0 \)

then

\[
(Ax, x) = \lambda (x, x)
\]
Taking imaginary parts, we see that
\[ \text{Im}(\lambda)(x, x) = 0 \]
Since \( x \neq 0 \), this implies that
\[ \text{Im}(\lambda) = 0 \]
which completes the proof.

**Theorem 1.10** The eigenvectors associated with distinct eigenvalues of a symmetric matrix \( A \) are orthogonal

**Proof**

From

\[ Ax = \lambda x \]
\[ Ay = \mu y \]

\( \lambda \neq \mu \), we obtain

\[ (y, Ax) = \lambda(y, x) \]
\[ (x, Ay) = \mu(x, y) \]

Since

\[ (x, Ay) \overline{=} (Ax, y) = (y, Ax) \]

subtraction yields

\[ 0 = (\lambda - \mu)(x, y) \]
whence \((x, y) = 0\). This result has a great importance. In fact, its generalization to more general operators is one of the cornerstones of classical analysis.

1. Reduction to Diagonal Form: Distinct Eigenvalues

Assume that \(A\) has distinct eigenvalues \(\lambda_1, \ldots, \lambda_n\). Let \(x_1, \ldots, x_n\) be the associated eigenvectors, normalized by the condition that

\[
(x_i, x_i) = 1 \quad i = 1, 2, \ldots, n.
\]

We form the matrix \(T\) by using \(x_i\) as columns, i.e.,

\[
T = (x_1, x_2, \ldots, x_n).
\]  \hspace{1cm} (1.4)

Then \(T^T\) is the matrix obtained using the \(x_i\) as rows:

\[
T^T = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

Since \(((x_i, x_j)) = (\delta_{ij})\), we see that \(T\) is an orthogonal matrix. The product \(AT\) has the form

\[
AT = (\lambda_1 x_1, \ldots, \lambda_n x_n).
\]

It follows that

\[
T^T AT = (\lambda_i (x_i, x_j)) \\
= (\lambda_i \delta_{ij}) \\
= \text{diag}(\lambda_1, \ldots, \lambda_n)
\]  

14
Thus multiplying on the right by $T$ and on the left by $T^T$, and using the fact that $TT^T = I$, we obtain the important result

$$A = T \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} T^T.$$ (1.5)

In the following we use the notation

$$A = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}.$$ (1.6)

So far we have only established the result for the case where the $\lambda_i$ are all distinct. This result is true in the general case [3. chapter 4], that is, a general symmetric matrix can be reduced to diagonal form by means of an orthogonal transformation.

We state this result in the following theorem.

**Theorem 1.11** If $A \in \mathbb{R}^{n \times n}$ is symmetric, then there exists an orthogonal matrix $T$ such that

$$T^T A T = \text{diag}(\lambda_1, \ldots, \lambda_n)$$

**2. Reduction of Quadratic Form to Canonical Form**: We consider an arbitrary real quadratic form

$$(Ax, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$
where $A^T = A \in \mathbb{R}^{n \times n}$. Therefore (Theorem 1.11) there exists a real orthogonal matrix $T$ such that

$$
\Lambda = T^T A T = (\lambda_i \delta_{ij}),
$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A$.

Setting $x = Ty$ we have

$$(Ax, x) = (ATy, Ty) = (T^T A Ty, y) = (\Lambda y, y),$$

or,

$$(Ax, x) = \sum_{i=1}^{n} \lambda_i y_i^2. \quad (1.7)$$

Since $T$ is orthogonal we see that $x = Ty$ implies that

$$T^T x = T^T Ty = y.$$  

Hence to each value of $x$ corresponds precisely one value of $y$ and conversely.

**Theorem 1.12**  Every real quadratic form $(Ax, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ where $A^T = A \in \mathbb{R}^{n \times n}$, can be reduced to the canonical form (1.7) by an orthogonal transformation, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A$.

3 - Reduction to Tridiagonal Form:

In the calculation of the eigenvalues and eigenvectors of real symmetric matrices, several methods have been considered based upon an initial reduction of the matrix to tridiagonal form using orthogonal transformations. One
of the advantages of this initial reduction is that a tridiagonal form readily
yields a Sturm sequence of polynomials terminating with the characteristic
polynomial of the matrix [7, S-2]. We have

**Theorem 1.13** Given a symmetric matrix \( A \in \mathbb{R}^{n \times n} \), then there exists an
orthogonal matrix \( U \) such that

\[
T = U^T A U
\]

is tridiagonal.

Further results on the tridiagonalization consult [7, S-2], [7, 9-1], [17, 4-5].

4 - **Simultaneous Reduction to Diagonal Form:**

One can ask whether or not we can simultaneously reduce two real symmetric matrices \( A, B \) to diagonal form. The answer is given in the following theorem

**Theorem 1.14** A necessary and sufficient condition that there exist an
orthogonal matrix \( T \) such that

\[
T^T A T = diag(\lambda_1, \ldots, \lambda_n), \quad T^T B T = diag(\mu_1, \ldots, \mu_n)
\]

(1.8)

is that \( A \) and \( B \) commute.

For the proof see [3, pp.56-57].

In theorem 1.13 we used an orthogonal transformation to reduce simultaneously the two matrices \( A \) and \( B \) to diagonal form. In many cases however
it is sufficient to reduce $A$ and $B$ simultaneously to diagonal form using a nonsingular matrix. We state the result in the following theorem [3. pp.58-59].

**Theorem 1.15** Given two symmetric matrices, $A$ and $B$, with $A$ positive definite, there exists a nonsingular matrix $T$ such that

$$T^T AT = I$$  \hspace{1cm} (1.9)

$$T^T BT = \text{diag}(\mu_1, \ldots, \mu_n).$$  \hspace{1cm} (1.10)

### 1.4 The Generalized Symmetric Eigenvalue Problem

Let $A$ and $B$ be two $n$ by $n$ matrices. The set of all matrices of the form $A - \lambda B$ with $\lambda \in \mathbb{C}$ is said to be a pencil. The eigenvalues of the pencil are the elements of the set

$$\lambda(A, B) = \{ \lambda \in \mathbb{C} \mid \det(A - \lambda B) = 0 \}$$

If $\lambda \in \lambda(A, B)$ and

$$Ax = \lambda Bx, \quad x \neq 0$$  \hspace{1cm} (1.11)

then $x$ is referred to as an eigenvector of $A - \lambda B$.

In this section we briefly give some of the mathematical properties of the generalized eigenvalue problem (1. 11).
The important case is when $A$ and $B$ are symmetric with one of them positive definite.

The first thing we may observe about the generalized eigenvalue problem is that there are $n$ eigenvalues if and only if $\text{rank}(B) = n$ [7. p.252]. If $B$ is rank deficient then $\lambda(A, B)$ may be finite, empty, or infinite.

Examples:

(i) If $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ then $\lambda(A, B) = \{-1/3\}$.

(ii) If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ then $\lambda(A, B) = C$.

(iii) If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then $\lambda(A, B) = \emptyset$.

Note that if $0 \neq \lambda \in \lambda(A, B)$ then $1/\lambda \in \lambda(B, A)$.

When $A$ and $B$ are symmetric and one of them is positive definite, say $B$, the matrix equation $Ax = \lambda Bx$ has been extensively studied (see, e.g. [7. 8-6]).

The generalized symmetric eigenvalue problem can be reduced to a special eigenvalue problem

\[ Dz = \lambda z \]  

(1.12)

where $D = B^{-1}A$ provided that one of the matrices, say $B$ is nonsingular. The eigenvectors of the special eigenvalue problem are exactly the same as those of the generalized eigenvalue problem. $D$ is no longer symmetric,
despite the symmetry of \( A, B \) and \( B^{-1} \). We have

\[
D = B^{-1}A \quad \text{and} \quad D^T = (B^{-1}A)^T = A^T(B^{-1})^T = AB^{-1}.
\]

(1.13)

So it is clear that the matrix \( D \) is symmetric if and only if \( A \) and \( B^{-1} \) or \( A \) and \( B \) commute; in fact

if \( AB = BA \) then \( A = BAB^{-1} \) and \( B^{-1}A = AB^{-1} \) so that \( D = D^T \) by (1.13), and, conversely, if \( D = D^T \) then \( AB = BA \).

In the case where \( A \) and \( B \) do not commute, the symmetry is lost. From (1.12-13) we derive the determinant of \( A \). Since

\[
det(D) = det(B^{-1}A) = \lambda_1 \lambda_2 \ldots \lambda_n \quad (\lambda_i \in \lambda(A, B))
\]

and

\[
det(B^{-1}A) = det(B^{-1})det(A) = (detB)^{-1}det(A),
\]

we obtain

\[
det(A) = det(B)\lambda_1 \lambda_2 \ldots \lambda_n.
\]

(1.14)

In the special case where \( B \) is diagonal, say \( B = \text{diag}(a_1, a_2, \ldots, a_n) \), then we have for later use the following result

\[
detA = \lambda_1 \ldots \lambda_n a_1 \ldots a_n.
\]

(1.15)

In the rest of this section we assume that \( A^T = A \) is positive definite and \( B = B^2 \).

20
We note that if $\lambda = 0$ is not an eigenvalue of (1.11), it is therefore possible, and it will turn out to be useful to write the problem (1.11) in the form

$$Bx = \mu Ax,$$  \hspace{1cm} (1.16)

where $\mu = 1/\lambda$. It is possible that this problem has the eigenvalue $\mu = 0$ if $B$ has a nontrivial null space. In this case, we simply say that $\lambda = \infty$ is an eigenvalue of (1.11), and the following theorems hold.

\textbf{Theorem 1.16} Every eigenvalue of the problem (1.16), and hence also of the problem (1.11), is real.

\textbf{proof}

If (1.16) is satisfied then

$$(Bx, x) = \mu (Ax, x).$$

Taking imaginary parts, we see that

$$\text{Im}(\mu)(Ax, x) = 0.$$  

Since $x \neq 0$ and $A$ positive definite, this implies that

$$\text{Im}(\mu) = 0,$$

which completes the proof.
Theorem 1.17 Let \( \mu_1, \mu_2 \) be two different eigenvalues of the problem (1.16) and let \( x_1, x_2 \) be the corresponding eigenvectors. Then \( (Ax_1, x_2) = 0 \), that is, \( x_1 \) and \( x_2 \) are orthogonal in this sense.

proof

We have

\[
Bx_1 = \mu_1 Ax_1 \\
Bx_2 = \mu_2 Ax_2
\]

so that

\[
(Bx_1, x_2) = \mu_1 (Ax_1, x_2) \\
(Bx_2, x_1) = \mu_2 (Ax_2, x_1).
\]

Moreover

\[
(Bx_1, x_2) = (x_1, Bx_2) = (Bx_2, x_1)
\]

and

\[
(Ax_1, x_2) = (x_1, Ax_2) = (Ax_2, x_1).
\]

Therefore we obtain

\[
(\mu_1 - \mu_2)(Ax_1, x_2) = 0.
\]
Since $\mu_1 \neq \mu_2$, this implies $(Ax_1, x_2) = 0$, which completes the proof.

If none of the two matrices is positive definite then the eigenvalues of (1.11) may be complex. For example the eigenvalue problem

\[
\begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} = \lambda
\begin{bmatrix}
1 & 0 \\
0 & -1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

whose coefficient matrices are real and symmetric, has the eigenvalues $\lambda = \pm i$.

For further results on the number of non-real eigenvalues of (1.11) see [12]. We end this section with the following theorem

**Theorem 1.18** Let $A^T = A \in \mathbb{R}^{n \times n}$ be positive definite and $B^T = B \in \mathbb{R}^{n \times n}$ then by theorem 1.14 there exist an orthogonal matrix $T$ such that

\[T^TAT = I\]

and

\[T^TB = \text{diag}(\mu_1, \ldots, \mu_n)\].

Moreover

\[Ax_i = \lambda_i Bx_i \text{ for } i = 1, 2, \ldots, n \text{ where } \lambda_i = 1/\mu_i.\]
Chapter 2

Orthogonal polynomials

The study of orthogonal polynomials is a subject closely related to many important branches of analysis. They furnish comparatively general and instructive illustrations of certain situations in differential and integral equations. Here we discuss certain special orthogonal polynomials, which are taken with respect to an indefinite-weight function τ(λ). We shall be dealing with the three-term recurrence relations with an "indefinite-weight function."

Though, in section 1 we define the Sturm-Liouville difference equations with an indefinite weight-function, and write it in matrix notation in the form, \( AY = \lambda BY \), where \( A = A^T \), tridiagonal and \( B \) diagonal real matrices. In section 2 we give some results of the type of Green's theorem or Lagrange identity for differential equations. In section 3 we establish
some orthogonality properties while in section 4 we try to give a very brief survey of results on the Riemann-Stieltjes integral, and we end in section 5 by defining the spectral function.

2.1 Sturm Liouville Difference Equations With an Indefinite Weight

We consider here the boundary problems of Sturm-Liouville type associated with the recurrence formula

\[ c_n y_{n+1} = (a_n \lambda - b_n) y_n - c_{n-1} y_{n-1} \]  \hspace{1cm} (2.1)

where \( b_n, n = 0, 1, \ldots, m-1 \), is an arbitrary finite sequence of real numbers, and \( a_n, c_n \), are real scalars such that, \( a_n \neq 0, c_n > 0, n = -1, 0, 1, \ldots, m-1 \).

If we introduce the boundary conditions

\[ y_{-1} = 0, \quad y_m = 0 \]  \hspace{1cm} (2.2)

(2.1-2) defines an eigenvalue problem. It is clear that if we construct a typical solution, that is to say, a sequence, satisfying (2.1) and the first boundary condition (2.2), we have to impose at least \( y_0 \neq 0 \) since otherwise by (2.1), \( y_1 = 0, y_2 = 0, \ldots \), and the sequence vanishes identically. So fixing the initial conditions

\[ y_{-1}(\lambda) = 0, y_0(\lambda) = 1/c_{-1} > 0, \]  \hspace{1cm} (2.3)
the values of \( y_1(\lambda), y_2(\lambda), \ldots \), can be found successively from (2.1). For \( n \geq 0 \), it is obvious that \( y_n(\lambda) \) is a polynomial of degree precisely \( n \). The second boundary condition (2.2) is satisfied if

\[
y_m(\lambda) = 0.
\]

The roots of this equation, the eigenvalues, are thus the zeros of a polynomial of degree \( m \). For if (2.3) holds, the sequence

\[
y_{-1}(\lambda), y_0(\lambda), y_1(\lambda), \ldots, y_m(\lambda)
\]

satisfies the conditions (2.1), (2.2) of the boundary problem, without vanishing identically. A solution of (2.1-2) is therefore an \( m \)-vector \( Y(\lambda) \),

\[
Y(\lambda) = [y_0(\lambda), y_1(\lambda), \ldots, y_{m-1}(\lambda)],
\]

where \( y_{-1}(\lambda) = y_m(\lambda) = 0 \).

From (2.1) and (2.2) we have

\[
\begin{align*}
n &= 0, & c_0 y_0 + b_0 y_0 &= a_0 \lambda y_0 \\
n &= 1, & c_1 y_2 + b_1 y_1 + c_0 y_0 &= a_1 \lambda y_1 \\
 & \quad \vdots \\
n &= m-1, & c_{m-1} y_m + b_{m-1} y_{m-1} + c_{m-2} y_{m-2} &= a_{m-1} \lambda y_{m-1} \\
\end{align*}
\]

subject to

\[
y_{-1} = 0, y_m = 0.
\]
This is equivalent to

\[
\begin{bmatrix}
    b_0 & c_0 & 0 & \cdots & 0 & 0 \\
    c_0 & b_1 & c_1 & \cdots & 0 & 0 \\
    0 & c_1 & b_2 & \cdots & 0 & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & c_{m-2} & y_{m-2} \\
    0 & 0 & \cdots & c_{m-2} & b_{m-1}
\end{bmatrix}
\begin{bmatrix}
    y_0 \\
    y_1 \\
    y_2 \\
    \vdots \\
    y_{m-2} \\
    y_{m-1}
\end{bmatrix}
= \lambda
\begin{bmatrix}
    a_0 & 0 & 0 & \cdots & 0 & 0 \\
    0 & a_1 & 0 & 0 & \cdots & 0 \\
    0 & 0 & a_2 & 0 & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & a_{m-2} & 0 \\
    0 & 0 & \cdots & 0 & a_{m-1}
\end{bmatrix}
\begin{bmatrix}
    y_0 \\
    y_1 \\
    y_2 \\
    \vdots \\
    y_{m-2} \\
    y_{m-1}
\end{bmatrix}
\]

i.e.,

\[ A\tilde{Y} = \lambda B\tilde{Y} \]

where

\[
A = A^T =
\begin{bmatrix}
    b_0 & c_0 & 0 & \cdots & 0 & 0 \\
    c_0 & b_1 & c_1 & \cdots & 0 & 0 \\
    0 & c_1 & b_2 & \cdots & 0 & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & c_{m-2} \\
    0 & 0 & \cdots & c_{m-2} & b_{m-1}
\end{bmatrix}
\quad B =
\begin{bmatrix}
    a_0 & 0 & 0 & \cdots & 0 & 0 \\
    0 & a_1 & 0 & 0 & \cdots & 0 \\
    0 & 0 & a_2 & 0 & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & a_{m-2} & 0 \\
    0 & 0 & \cdots & 0 & a_{m-1}
\end{bmatrix}
\]

are \( m \times m \) real matrices with

\[ \tilde{Y} = [y_0, y_1, \ldots, y_{m-1}]^T. \]

Conversely, if we have

\[ AY = \lambda BY \quad (2.7) \]

with \( A = A^T \), tridiagonal, and \( B \) is diagonal \( m \times m \) real matrices, we can write (2.7) of the form (2.1). Then \( y_m = 0 \) is equivalent to \( \det(A-\lambda B) = 0 \),
that is; the zeros of \( y_n(\lambda) \) are the eigenvalues of the generalized eigenvalue problem, \( AY = \lambda BY \), for some \( Y \neq 0 \).

Concerning the separation properties of \( y_n \) in the case where, \( a_n > 0, c_n > 0 \), see [2. 4-3].

### 2.2 Lagrange - Type Identity

We prove here, for later use, some results on Green's theorem, or Lagrange identity for recurrence relations.

We consider the orthogonality only in the finite discrete case; these results may be used to establish the reality of the spectrum. The orthogonality of the eigenvectors is considered first.

**Theorem 2.1** For \( 0 \leq n < m \).

\[
(\lambda - \mu) \sum_{r=0}^{n} a_r y_r(\lambda) y_r(\mu) = c_n \begin{vmatrix} \begin{array}{cc} y_{n+1}(\lambda) & y_{n+1}(\mu) \\ y_n(\lambda) & y_n(\mu) \end{array} \end{vmatrix}
\]

**(2.8)**

**Proof.** By induction. We write (2.1) for the two arguments \( \lambda, \mu \) in full giving

\[
c_n y_{n+1}(\lambda) = (a_n \lambda - b_n) y_n(\lambda) - c_{n-1} y_{n-1}(\lambda)
\]

\[
c_n y_{n+1}(\mu) = (a_n \mu - b_n) y_n(\mu) - c_{n-1} y_{n-1}(\mu).
\]

(2.9)
Multiplying respectively by \( y_n(\mu), y_n(\lambda) \) and subtracting we have
\[
c_n \{y_{n+1}(\lambda)y_n(\mu) - y_{n+1}(\mu)y_n(\lambda)\} = a_n(\lambda - \mu)y_n(\lambda)y_n(\mu) + c_{n-1} \{y_n(\lambda)y_{n-1}(\mu) - y_n(\mu)y_{n-1}(\lambda)\}. \tag{2.10}
\]

Putting \( n = 0 \), and recalling that \( y_{-1}(\lambda) = y_{-1}(\mu) = 0 \), we derive (2.8) for \( n = 0 \). Induction over \( n \) then yields (2.8) from (2.10), in the general case.

Using L'Hopital's rule, we deduce two special cases obtained by dividing (2.8) by \( \lambda - \mu \) and making \( \mu \to \lambda \) for fixed \( \lambda \). We obtain,

\[
\text{Theorem 2.2 For } 0 \leq n < m,
\sum_{r=0}^{n} a_r \{y_r(\lambda)\}^2 = c_n \begin{vmatrix} y_{n+1}(\lambda) & y_{n+1}(\lambda) \\ y_{n}(\lambda) & y_{n}(\lambda) \end{vmatrix} \tag{2.11}
\]

for real \( \lambda \).

The other special case is:

\[
\text{Theorem 2.3 For } 0 \leq n < m, \text{ and complex } \lambda,
\sum_{r=0}^{n} a_r |y_r(\lambda)|^2 = (2i \text{Im} \lambda)^{-1} c_n \begin{vmatrix} y_{n+1}(\lambda) & y_{n+1}(\lambda) \\ y_{n}(\lambda) & y_{n}(\lambda) \end{vmatrix} \tag{2.12}
\]

This results immediately on putting \( \mu = \bar{\lambda} \) in (2.8).

\textbf{Definition} We call \( \rho_r \) the normalization constant associated with the eigenvalue \( \lambda_r \), where
\[
\rho_r = \sum_{\rho=0}^{m-1} a_r \{y_{\rho}(\lambda_r)\}^2. \tag{2.13}
\]
In the following sections, we assume that the normalization constants are
different from zero for \( r = 0, 1, \ldots, m - 1 \).

## 2.3 Orthogonality

We derive here two types of orthogonality; first the orthogonality of the
eigenvectors, that is to say, of certain sequences of the form \((2.5)\), and
secondly a dual orthogonality, which has a great importance, since in fact,
it establishes that the polynomials \( y_n \) are orthogonal in the usual sense.

We denote by \( \lambda_0, \lambda_1, \ldots, \lambda_{m-1} \), the roots of \((2.4)\), and we assume that these
zeros are all real and distinct. the first type of orthogonality is given by:

**Theorem 2.4** The sequences

\[
y_0(\lambda_r); y_1(\lambda_r); \ldots; y_{m-1}(\lambda_r), \quad r = 0, 1, \ldots, m - 1, \tag{2.14}
\]

are orthogonal according to

\[
\sum_{p=0}^{m-1} a_p y_p(\lambda_r) y_p(\lambda_s) = \rho_r \delta_{rs} \tag{2.15}
\]

where

\[
\rho_r = \sum_{p=0}^{m-1} \frac{a_p}{\rho} y_p(\lambda_r) = c_{m-1} y_{m-1}(\lambda_r) y_m(\lambda_r) \quad \tag{2.16}
\]

and \( \delta_{rs} \) denotes the Kronecker \( \delta \).
Proof. For \( r \neq s \), we take \( \lambda = \lambda_r, \mu = \lambda_s \), and \( n = m - 1 \) in (2.8), to find
\[
(\lambda_r - \lambda_s) \sum_{p=0}^{m-1} a_p y_p(\lambda_r)y_p(\lambda_s) = c_{m-1} \begin{vmatrix} y_m(\lambda_r) & y_m(\lambda_s) \\ y_{m-1}(\lambda_r) & y_{m-1}(\lambda_s) \end{vmatrix}
\]
The determinant on the right vanishes since \( y_m(\lambda) = 0 \) if \( \lambda = \lambda_r, \lambda_s \), therefore
\[
\sum_{p=0}^{m-1} a_p y_p(\lambda_r)y_p(\lambda_s) = 0
\]
since \( \lambda_r \neq \lambda_s \).

For \( r = s \), put, \( \lambda = \lambda_r \) and \( n = m - 1 \), in (2.11), to find
\[
\sum_{p=0}^{m-1} a_p \{y_p(\lambda_r)\}^2 = c_{m-1} \begin{vmatrix} y_m(\lambda_r) & y_m(\lambda_r) \\ y_{m-1}(\lambda_r) & y_{m-1}(\lambda_r) \end{vmatrix}
\]
\[
= c_{m-1} y_m(\lambda_r)y_m(\lambda_r).
\]
(2.18)
Since \( y_m(\lambda) = 0 \), note also that \( y_m'(\lambda) \neq 0 \) since all the roots of \( y_m(\lambda) \) are distinct and therefore, simple by assumption. The normalization constants \( \rho_r \) are essentially the reciprocal of the Christoffel numbers; see [19. 3. 4 7-8] in the case where \( \rho_r > 0 \).

The dual orthogonality, which is a consequence of (2.15), is

**Theorem 2.5** For \( 0 \leq p, q \leq m - 1 \).
\[
\sum_{r=0}^{m-1} y_p(\lambda_r)y_q(\lambda_r)\rho_r^{-1} = a_p^{-1} \delta_{pq}
\]
(2.19)

Proof It follows from the orthogonality property (2.15) that the \( m \) vectors \( y_{0,r} = y_0(\lambda_r); \ldots ; y_{m-1,r} = y_{m-1}(\lambda_r) \), for \( r = 0; 1; \ldots ; m - 1 \), are
linearly independent. Thus an arbitrary vector \( u_0, u_1, \ldots, u_{m-1} \), may be expressed in the form

\[
u_n = \sum_{r=0}^{m-1} v_r y_{nr} \rho_r^{-1}, \quad n = 0, 1, \ldots, m - 1.
\] (2.20)

where \( \rho_r \) are the normalization constants, and the Fourier coefficients \( v_r \) are to be found. Multiplying by \( a_n y_{np} \), summing over \( n \), and using (2.15) we get

\[
\sum_{n=0}^{m-1} a_n u_n y_{np} = \sum_{n=0}^{m-1} a_n y_{np} \sum_{r=0}^{m-1} v_r y_{nr} \rho_r^{-1} \\
= \sum_{r=0}^{m-1} v_r \rho_r^{-1} \sum_{n=0}^{m-1} a_n y_{np} y_{nr} \\
= v_p.
\]

Therefore, substituting for \( v_r \) in (2.20) we have

\[
u_n = \sum_{r=0}^{m-1} y_{nr} \rho_r^{-1} \sum_{q=0}^{m-1} a_q u_q y_{qr} \\
= \sum_{q=0}^{m-1} u_q a_q \sum_{r=0}^{m-1} y_{nr} y_{qr} \rho_r^{-1}.
\] (2.21)

Here the \( u_n \) are arbitrary, and we may derive (2.19) from (2.21) by comparing the coefficients of the \( u_q \). (2.19) establishes that the polynomials \( y_n \) are orthogonal with respect to the distribution of weights \( \rho_r^{-1} \) at the points \( \lambda_r \).
2.4 The Riemann-Stieltjes Integral

Before we state the definition of this integral and some of its properties we give the definition of functions of bounded variation.

**Definition** Let \( \sigma \) be a function defined on a finite real interval \([a, b]\). If \( \{x_r\} \) is a partition of \([a, b]\), and if there exists a positive number \( B \) such that

\[
\sum_{r=0}^{n-1} |\sigma(x_{r+1}) - \sigma(x_r)| \leq B.
\] (2.22)

For all partitions of \([a, b]\), then \( \sigma \) is said to be of *bounded variation on* \([a, b]\). The least common upper bound of all such sums is called the *total variation* of \( \sigma \) over \([a, b]\) and is denoted by \( V_\sigma(a, b) \) or \( \text{var}\{\sigma(x); a, b\} \) or if there is no danger of misunderstanding by \( V_\sigma \). Note that \( V_\sigma \) is a nonnegative real number with \( V_\sigma(a, b) = 0 \) if and only if \( \sigma \) is constant. For more information about functions of bounded variation see [1, chapter 6] and [13, pp.156-157].

**Definition** For a pair of functions \( f, \sigma \) defined on a finite interval \([a, b]\) and for any partition \( \{x_r\} \) of \([a, b]\) with

\[ a = x_0 < x_1 < \ldots < x_n = b \] (2.23)

we form the sum

\[ S = \sum_{r=0}^{r=n-1} f(\xi_r)(\sigma(x_{r+1}) - \sigma(x_r)) \] (2.24)

where \( \xi_r \in [x_r, x_{r+1}] \). If as \( n \to \infty \), and \( \max|x_{r+1} - x_r| \to 0 \), the sum \( S \) tends to a unique limit for all partitions \( \{x_r\} \) satisfying (2.24) and for
all choices of the \( \xi_r \in [x_r, x_{r+1}] \), the limit is called the Stieltjes integral of \( f \) with respect to \( \sigma \), written
\[
\int_a^b f(x) d\sigma(x).
\] (2.25)

If \( \sigma \) is a step-function, the Stieltjes integral reduces to a sum [1, p.148].

**Theorem 2.6** If \( f \) is continuous in the finite interval \([a,b]\) and \( \sigma \) is of bounded variation over \([a,b]\), then the integral
\[
\int_a^b f(x) d\sigma(x)
\]
exists [1, p.150].

**Theorem 2.7** Under the conditions of theorem (2.6), we have [2, p.421]
\[
|\int_a^b f(x) d\sigma(x)| \leq \max |f(x)| \text{Var}\{\sigma(x), a, b\},
\]
where the maximum is taken over \([a,b]\).

Integrals over infinite intervals are understood in the improper sense so that, for finite \(a\)
\[
\int_a^\infty f(x) d\sigma(x) = \lim_{b \to \infty} \int_a^b f(x) d\sigma(x)
\]
if the limit exists, and likewise
\[
\int_{-\infty}^a f(x) d\sigma(x) = \lim_{b \to -\infty} \int_a^b f(x) d\sigma(x).
\]

For more on the existence and other properties of the Stieltjes integral consult [2, Appendix I], [1, chapter 7], [13, Appendix I].
2.5 The Spectral Function

We define this function on the real axis, as a step-function, having jumps of amount $1/\rho_r$ at the eigenvalues $\lambda_r$ (points of the spectrum), and being constant in between the eigenvalues, where $\rho_r$ is the normalization constant defined by (2.13). More precisely we set

$$\tau(\lambda) = \sum_{\lambda_r < \lambda} 1/\rho_r,$$

for all $\lambda \in R$.

$\tau(\lambda)$ is called the spectral function. This function unifies the eigenvalues $\lambda_r$ and the normalization constants $\rho_r$ into a single concept. Thus the spectral function $\tau(\lambda)$ is a real-valued step function which, in this thesis, is not monotone since $\rho_r$ can take real negative and positive values.

When expressed in terms of the spectral function, the dual orthogonality (2.19) assumes the form

$$\int_{-\infty}^{\infty} y_p(\lambda)y_q(\lambda)d\tau(\lambda) = a_p^{-1}\delta_{pq}, \quad 0 \leq p, q \leq m - 1.$$

(2.27) is simply the Stieltjes integral of the functions $y_p(\lambda)y_q(\lambda)$ for $0 \leq p, q \leq m - 1$ with respect to the spectral function $\tau(\lambda)$, sometimes also called the distribution function.
Chapter 3

The Real 2 by 2 Case

The problem of solving the generalized matrix problem

$$AY = \lambda BY$$

(3.1)

has been extensively investigated using a variety of techniques see [7, 8-6], [18], [16].

We wish to consider a type of problem associated with this equation which has not been investigated to any extent - the inverse problem. Instead of assuming that the matrix $A$ is known and trying to find the solution, that is, the eigenvalues and the corresponding eigenvectors, we shall suppose that the eigenvalues of (3.1) and the matrix $B$ are known along with other conditions which we have to seek for them to try to determine the matrix $A$. We restrict ourselves to the 2 by 2 real case, with $B$ a nonsingular matrix.
of the form

$$B = \begin{bmatrix} a_0 & 0 \\ 0 & a_1 \end{bmatrix}$$

where either $a_0$ or $a_1$ is negative, but not both, i.e., $B$ is indefinite, and we look for a matrix $A$ of the form

$$\begin{bmatrix} b_0 & c_0 \\ c_0 & b_1 \end{bmatrix}$$

such that

- $C_1$ \((3.1)\) holds for some $Y \neq 0$;

- $C_2$ A symmetric Jacobi positive definite matrix

In section 1 we will prove a lemma on the existence of the solution of the inverse problem, and in section 2 we give our main result by constructing the solution and determining the total number of such solutions.

### 3.1 Necessary Condition for the Existence of a Solution

In this section we shall give a basic tool by proving a necessary condition for the existence of the solution. As shown in section 2.1, \((3.1)\) can be
written in the form (2.1-2) by taking \( m = 2 \):

\[
\begin{align*}
c_0 y_1 &= (a_0 \lambda - b_0) y_0 \\
c_1 y_2 &= (a_1 \lambda - b_1) y_1 - c_0 y_0
\end{align*}
\]

subject to \( y_2 = 0 \).  \hspace{1cm} (3.2)

Since seeking for \( A \) is the same as reconstructing the boundary problem (3.2), it is obvious that the knowledge of the eigenvalues only is insufficient [2, 1.7]. We get a problem which is soluble if we give not only the eigenvalues \( \{ \lambda_r, r = 0, 1 \} \) (the zeros of \( y_2 \)) but also the normalization constants \( \rho_r \) where \( \rho_r \) is defined in (2.13) or (2.15), with \( m = 2 \). From the definition (2.26), given \( \lambda \) and \( \rho_r \) is equivalent to saying that we are given the spectral function \( \tau(\lambda) \).

Now the question is: What is the necessary condition which \( \lambda_r \) and \( \rho_r \) must satisfy in order for the inverse problem to have a solution? That is, when can we find a symmetric positive definite Jacobi matrix \( A \)? The answer is given by the lemma below which also holds true in the general case i.e., \( A = A^T \), and \( B \) diagonal \( m \) by \( m \) real matrices.

**Lemma**

Given \( B = \text{diag}(a_0, a_1, \ldots, a_{m-1}) \in \mathbb{R}^{m \times m} \), \( \lambda(A, B) = \{ \lambda_0, \lambda_1, \ldots, \lambda_{m-1} \} \), the spectrum of (3.1), and the corresponding normalization constants \( \rho_r, r = 0, 1, \ldots, m - 1 \), defined by (2.13). A necessary condition for the existence of a positive definite symmetric Jacobi matrix \( A \) which solves the inverse problem is that

\[
\text{sign} \lambda_i = \text{sign} \rho_i, \quad (\lambda_i \rho_i > 0), \quad i = 0, 1, \ldots, m - 1. \hspace{1cm} (3.3)
\]
Proof
Suppose on the contrary i.e., there exist \( 0 \leq i_0 \leq m - 1 \) such that \( \text{sign} \lambda_{i_0} \neq \text{sign} \rho_{i_0} \). Then by (3.1) we have

\[
AY_{i_0} = \lambda_{i_0} BY_{i_0} \tag{3.4}
\]

for some \( Y_{i_0} = (y_0(\lambda_{i_0}), \ldots, y_{m-1}(\lambda_{i_0}))^T \). (3.4) implies that

\[
Y_{i_0}^* A Y_{i_0} = \lambda_{i_0} Y_{i_0}^* B Y_{i_0} \tag{3.5}
\]

and, on the other hand \( Y_{i_0}^* B Y_{i_0} = \sum_{r=0}^{m-1} \sigma_r \{ y_r(\lambda_{i_0}) \}^2 = \rho_{i_0} \), so that

\[
0 < Y_{i_0}^* A Y_{i_0} = \lambda_{i_0} \rho_{i_0} \Rightarrow 0.
\]

This contradiction completes the proof.

We end this section with the following definition.

Definition
The scalars

\[
\mu_j = \int_{-\infty}^{\infty} \lambda^j d\tau(\lambda), \quad j = 0, 1, \ldots, \tag{3.6}
\]

are called moments.

Remark
Note that (3.6) is equivalent to

\[
\mu_j = \sum_{r=0}^{m-1} \lambda_r^j / \rho_r \quad j = 0, 1, \ldots \tag{3.7}
\]

Hence by the above lemma the odd moments are all positive, indeed

\[
\mu_{2k+1} = \sum_{r=0}^{m-1} \lambda_r^{2k+1} / \rho_r = \sum_{r=0}^{m-1} \lambda_r^{2k} \lambda_r / \rho_r > 0
\]

for \( k = 0, 1, \ldots \).
3.2 Construction of a Solution

In this section we determine the matrix

\[
\begin{bmatrix}
  b_0 & a_0 \\
  a_0 & b_1
\end{bmatrix}
\]

satisfying \( C_1 \), and \( C_2 \) and therefore (3. 2).

We shall make the following assumptions: Given the matrix

\[
B = \begin{bmatrix}
  a_0 & 0 \\
  0 & a_1
\end{bmatrix}
\]

and the couples \((\lambda_0, \rho_0)\), \((\lambda_1, \rho_1)\) and also the moments \(\mu_j\), for \(j = 0, 1, 2\).

defined in (3. 6) such that.

\(H_1) \ a_0 \Delta_0 \Delta_1 > 0\)

\(H_2) \ \lambda_i \rho_i > 0\), for \(i = 0, 1\), and \(\lambda_0 \lambda_1 < 0\).

\(H_3) \ a_1 \Delta_1 \Delta_2 > 0\), where \(\Delta_i\) for \(i = 0, 1, 2\) is given as follows:

\[
\Delta_0 = 1
\]

\[
\Delta_1 = \mu_0
\]

\[
\Delta_2 = \text{det} \begin{bmatrix}
  \mu_0 & \mu_1 \\
  \mu_1 & \mu_2
\end{bmatrix}
\]

The solution may be derived in two steps, the first of which consists in constructing the two polynomials \(y_0(\lambda), y_1(\lambda)\).

**Theorem 3.1** Let \(\tau(\lambda)\) be the step function defined in (2. 26) and assume that \(H_1, H_2, H_3\) hold. Then there exist two real polynomials \(y_0(\lambda), y_1(\lambda)\) of
degrees 0 and 1, respectively, such that the coefficients of $\lambda^n$ ($n = 0, 1$) are
determined up to a change of sign.

proof
The proof is by "orthogonalization". We look for $y_0(\lambda)$ and $y_1(\lambda)$ of the
form:

$$
\begin{align*}
  y_0(\lambda) &= k_0 \\
  y_1(\lambda) &= k_1\{\lambda + \alpha_{1,0}\},
\end{align*}
$$

(3.5)

where $k_0 \neq 0$, $k_1 \neq 0$. First, consider the solution of (2.27) with $p \neq q$. For
these it is necessary and sufficient that

$$
  \int_{-\infty}^{\infty} y_p(\lambda)\lambda^q d\tau(\lambda) = 0 \quad q = 0, 1, \ldots, p - 1.
$$

(3.9)

For $p = 1$, $q = 0$, we have

$$
  \int_{-\infty}^{\infty} y_1(\lambda)\lambda^0 d\tau(\lambda) = 0
$$

$$
  \int_{-\infty}^{\infty} k_1\{\lambda + \alpha_{1,0}\} d\tau(\lambda) = 0.
$$

Since $k_1 \neq 0$,

$$
  \mu_1 + \alpha_{1,0} \mu_0 = 0
$$

so that

$$
  \alpha_{1,0} = -\mu_1/\mu_0.
$$

(3.10)

For $p = q = 1$ in (2.27) we have

$$
  \int_{-\infty}^{\infty} \{y_1(\lambda)\}^2 d\tau(\lambda) = a_1^{-1}
$$

$$
  \int_{-\infty}^{\infty} k_1^2\{\lambda^2 + 2\alpha_{1,0}\lambda + \alpha_{1,0}^2\} d\tau(\lambda) = a_1^{-1}
$$
these imply that
\[ k_1^2 = \mu_0/\alpha_1(\mu_2\mu_0 - \mu_1^2) \] (3.11)
which is positive by \( H_3 \).

For \( p = q = 0 \) in (2.27) we have
\[
\int_{-\infty}^{\infty} \{y_0(\lambda)\}^2 d\tau(\lambda) = a_0^{-1}
\]
\[
\int_{-\infty}^{\infty} k_0^2 d\tau(\lambda) = a_0^{-1}
\]
so that
\[ k_0^2 = 1/a_0\mu_0 \] (3.12)
which is positive by \( H_1 \).

Now since we have found \( y_0, y_1 \) we are able to find the numbers \( c_0, b_0, b_1 \)
i.e., the matrix \( A \).

**Theorem 3.2** Under the same assumptions as in theorem 3.1, there exist \( c_0, b_0, b_1 \) such that the boundary problem (3.2) has \( \tau(\lambda) \) as its spectral function, and the matrix \( A \) is positive definite.

**Proof**
(3.2) gives
\[ c_0y_1(\lambda) = (a_0\lambda - b_0)y_0(\lambda), \]
where \( y_0 \) and \( y_1 \) are known by theorem 3.1, and thus also \( c_0, b_0 \) by comparing coefficients. More precisely, since
\[ c_0(k_1\lambda + k_1\alpha_{1,0}) = (a_0\lambda - b_0)k_0 \]
we can choose
\[ c_0 = \frac{a_0 k_0}{k_1} \tag{3.13} \]
and
\[ b_0 = -a_0 \alpha_{1,0}. \]

By (3.10)
\[ b_0 = \frac{a_0 \mu_1}{\mu_0}. \tag{3.14} \]

To get \( b_1 \), first let us find \( c_1 \). Since the eigenvalues of (3.2) are the zeros of \( y_2(\lambda) \), then, without loss of generality, \( y_2(\lambda) \) can be written as
\[
y_2(\lambda) = (\lambda - \lambda_0)(\lambda - \lambda_1) = \lambda^2 - (\lambda_0 + \lambda_1)\lambda + \lambda_0\lambda_1.
\]

Thus, again by (3.2) we have
\[
c_1(\lambda^2 - (\lambda_0 + \lambda_1)\lambda + \lambda_0\lambda_1) = (a_1 \lambda - b_1)(k_1 \lambda + k_1 \alpha_{1,0}) - c_0 k_0
\]
and so we can choose
\[ c_1 = a_1 k_1 \]
and
\[ b_1 = a_1((\lambda_0 + \lambda_1) + \alpha_{1,0}) \]
or,
\[ b_1 = a_1((\lambda_0 + \lambda_1) - \mu_1/\mu_0). \tag{3.15} \]

We have thus obtained all the entries of \( A \):
\[
A = \begin{bmatrix}
a_0 \mu_1/\mu_0 & a_0 k_0/k_1 \\
\alpha_1((\lambda_0 + \lambda_1) - \mu_1/\mu_0)
\end{bmatrix}
\]
It is left only to check that the matrix $A$ is positive definite. By (1.1) it suffices to prove that $D_i > 0$ for $i = 1, 2$. But

$$D_1 = a_0 \mu_1 / \mu_0 > 0$$

by $H_1, H_2$ and the remark (see (3.7)) above. Also

$$D_2 = \det(A)$$

which can be computed directly from (3.16) or simply by (1.32):

$$D_2 = \det(A)$$
$$= \lambda_0 \lambda_1 \det(B)$$
$$= a_0 a_1 \lambda_0 \lambda_1 > 0.$$

since $\lambda_0 \lambda_1 < 0$ and $B$ is indefinite i.e., $a_0 a_1 < 0$.

The solution above is not unique. In fact we will prove that we can find two positive definite symmetric Jacobi matrices satisfying (3.2) and not more than 2 matrices. To see this we rewrite (3.2) with $\lambda = \lambda_0, \lambda_1$, and recall that $y_2(\lambda_0) = y_2(\lambda_1) = 0$. Rewriting (3.1) with $\lambda = \lambda_0, \lambda_1$ and the corresponding eigenvectors $Y_0 = [y_0(\lambda_0), y_1(\lambda_0)]^T$, $Y_1 = [y_0(\lambda_1), y_1(\lambda_1)]^T$, respectively where $y_0(\lambda), y_1(\lambda)$ are given in (3.8), that is,

$$AY_0 = \lambda_0 BY_0$$
$$AY_1 = \lambda_1 BY_1$$

(3.17) holds if and only if

$$b_0 k_0 + c_0 k_1 (\lambda_0 + \alpha_{1,0}) = a_0 \lambda_0 k_0$$

44
\[
\begin{align*}
c_0k_0 + b_1k_1(\lambda_0 + \alpha_{1,0}) &= a_1\lambda_0 k_1(\lambda_0 + \alpha_{1,0}) \\
b_0k_0 + c_0k_1(\lambda_1 + \alpha_{1,0}) &= a_0\lambda_1 k_0 \\
c_0k_0 + b_1k_1(\lambda_1 + \alpha_{1,0}) &= a_1\lambda_1 k_1(\lambda_1 + \alpha_{1,0})
\end{align*}
\]

(3.18)

If we write \(\alpha_0 = \lambda_0 + \alpha_{1,0}\) and \(\alpha_1 = \lambda_1 + \alpha_{1,0}\) then (3.18) holds if and only if

\[
\begin{align*}
b_0k_0 + c_0k_1\alpha_0 &= a_0\lambda_0 k_0 \\
c_0k_0 + b_1k_1\alpha_0 &= a_1\lambda_0 k_1\alpha_0 \\
b_0k_0 + c_0k_1\alpha_1 &= a_0\lambda_1 k_0 \\
c_0k_0 + b_1k_1\alpha_1 &= a_1\lambda_1 k_1\alpha_1.
\end{align*}
\]

(3.19)

This is a linear system of 4 equations and 3 unknowns, which we write in matrix notation in the form

\[
\begin{bmatrix}
k_1\alpha_0 & k_0 & 0 \\
k_0 & k_1\alpha_0 & 0 \\
k_1\alpha_1 & k_0 & 0 \\
k_0 & k_1\alpha_1 & 0
\end{bmatrix}
\begin{bmatrix}
c_0 \\
b_0 \\
b_1
\end{bmatrix}
=
\begin{bmatrix}
a_0\lambda_0 k_0 \\
a_1\lambda_0 k_1\alpha_0 \\
a_0\lambda_1 k_0 \\
a_1\lambda_1 k_1\alpha_1
\end{bmatrix}
\]

(3.20)

or \(DX = C\), where

\[
D =
\begin{bmatrix}
k_1\alpha_0 & k_0 & 0 \\
k_0 & k_1\alpha_0 & 0 \\
k_1\alpha_1 & k_0 & 0 \\
k_0 & k_1\alpha_1 & 0
\end{bmatrix}
\]
\[
C = \begin{bmatrix}
  a_0 \lambda_0 k_0 \\
  a_1 \lambda_0 k_1 a_0 \\
  a_0 \lambda_1 k_0 \\
  a_1 \lambda_1 k_1 a_1 \\
\end{bmatrix}
\]

and

\[
X = \begin{bmatrix}
  c_0 \\
  b_0 \\
  b_1 \\
\end{bmatrix}
\]

Again we write (3.20) in the form

\[
\begin{bmatrix}
  k_1 a_0 & k_0 & 0 & : & a_0 \lambda_0 k_0 \\
  k_0 & 0 & k_1 a_0 & : & a_1 \lambda_0 k_1 a_0 \\
  k_1 a_1 & k_0 & 0 & : & a_0 \lambda_1 k_0 \\
  k_0 & 0 & k_1 a_1 & : & a_1 \lambda_1 k_1 a_1 \\
\end{bmatrix}
\]

\[
\rightarrow
\begin{bmatrix}
  k_0 & 0 & k_1 a_0 & : & a_1 \lambda_0 k_1 a_0 \\
  k_1 a_0 & k_0 & 0 & : & a_0 \lambda_0 k_0 \\
  k_1 a_1 & k_0 & 0 & : & a_0 \lambda_1 k_0 \\
  k_0 & 0 & k_1 a_1 & : & a_1 \lambda_1 k_1 a_1 \\
\end{bmatrix}
\]

\[
\rightarrow
\begin{bmatrix}
  1 & 0 & k_1 a_0/k_0 & : & a_1 \lambda_0 k_1 a_0/k_0 \\
  0 & k_0 & -(k_1 a_0)^2/k_0 & : & a_0 \lambda_0 k_0 - a_1 \lambda_0 (k_1 a_0)^2/k_0 \\
  0 & k_0 & -k_1^2 a_1 a_0/k_0 & : & a_0 \lambda_1 k_0 - a_1 \lambda_0 k_1^2 a_1 a_0/k_0 \\
  0 & 0 & k_1 a_1 - k_1 a_0 & : & a_1 \lambda_1 k_1 a_1 - a_1 \lambda_0 k_1 a_0 \\
\end{bmatrix}
\]

46
\[
\begin{bmatrix}
1 & 0 & \frac{k_1 a_0}{k_0} & : & a_1 \lambda_0 k_1 a_0 / k_0 \\
0 & 1 & -(k_1 a_0 / k_0)^2 & : & a_0 \lambda_0 - a_1 \lambda_0 (k_1 a_0 / k_0)^2 \\
0 & 0 & k_1^2 a_0 (a_0 - \alpha_1) / k_0 & : & a_0 k_0 (\lambda_1 - \lambda_0) + (a_0 - \alpha_1) a_1 \lambda_0 k_1^2 a_0 / k_0 \\
0 & 0 & k_1 (\lambda_1 - \lambda_0 a_0) & : & a_1 k_1 (\lambda_1 a_1 - \lambda_0 a_0)
\end{bmatrix}
\]

Solving \(DX = C\) is equivalent to solving \(M X = N\) where,

\[
M = \begin{bmatrix}
1 & 0 & \frac{k_1 a_0}{k_0} \\
0 & 1 & -(k_1 a_0 / k_0)^2 \\
0 & 0 & \frac{k_1^2 a_0 (a_0 - \alpha_1) / k_0} \\
0 & 0 & \frac{k_1 (\lambda_1 - \lambda_0 a_0)}
\end{bmatrix}
\]

and

\[
N = \begin{bmatrix}
a_1 \lambda_0 k_1 a_0 / k_0 \\
a_0 \lambda_0 - a_1 \lambda_0 (k_1 a_0 / k_0)^2 \\
a_0 k_0 (\lambda_1 - \lambda_0) + (a_0 - \alpha_1) a_1 \lambda_0 k_1^2 a_0 / k_0 \\
a_1 k_1 (\lambda_1 a_1 - \lambda_0 a_0)
\end{bmatrix}
\]

In order to have a consistent system we must have \(\beta / \alpha - \delta / \gamma = 0\), where \(\alpha = m_{33}, \gamma = m_{43}\), and \(\beta, \delta\), are the third and the fourth components of the vector \(N\). We have \(\beta / \alpha = a_1 \lambda_0 - a_0 k_0^2 / a_0 k_1^2\) and \(\delta / \gamma = a_1 (\lambda_0 + \lambda_1 + \alpha_{1,0})\) so that

\[
\delta / \gamma - \beta / \alpha = a_1 (\lambda_1 + \alpha_{1,0}) + a_0 k_0^2 / a_0 k_1^2
\]

(3.21)

\[
= a_1 (-\mu_1 / \mu_0 + \lambda_1) + a_1 (\mu_2 \mu_0 - \mu_1) / \mu_0 (\mu_0 \lambda_0 - \mu_1)
\]

(3.22)

\[
= 0 \times a_1 / \rho_0^2 \rho_1^2 (\lambda_0 \mu_0 - \mu_1) = 0,
\]

(3.23)
where we used (3. 10-11-12) in deriving (3. 22) and (3. 7) for (3. 23). Therefore the system is consistent and thus the solution exists. Solving the equation $MX = N$ we derive

$$c_0 = a_0 k_0 / k_1$$  \hspace{1cm} (3.24)
$$b_0 = a_0 \mu_1 / \mu_0$$  \hspace{1cm} (3.25)
$$b_1 = a_1 (\lambda_0 + \lambda_1 - \mu_1 / \mu_0).$$  \hspace{1cm} (3.26)

Thus we can find at most two matrices, one with $c_0 > 0$ and the other with $c_0 < 0$ since $k_0, k_1$ are determined up to a change of sign (theorem 3.1). We now give an example.

Example

Given

$$B = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

and $\lambda(A, B) = \{-3, 1\}$ and $\rho_0 = -2, \rho_1 = 1$ then

- $\mu_0 = 1 / \rho_0 + 1 / \rho_1 = -1/2 + 1 = 1/2$

- $\mu_1 = \lambda_0 / \rho_0 + \lambda_1 / \rho_1 = 3/2 + 1 = 5/2$

- $\mu_2 = \lambda_0^2 / \rho_0 + \lambda_1^2 / \rho_1 = -9/2 + 1 = -7/2$
therefore $H_1, H_2, H_3$ are satisfied and so

\[ A_1 = \begin{bmatrix} 10 & 8 \\ 8 & 7 \end{bmatrix} \]

and

\[ A_2 = \begin{bmatrix} 10 & -8 \\ -8 & 7 \end{bmatrix}, \]

which are positive definite since

\[ D_1 = 10, \; D_2 = \text{det}(A_i) = 6 \]

and each satisfy (3. 1) with $Y = [y_0(\lambda), y_1(\lambda)]^T$, where

- $y_0(\lambda) = \pm 1$ and

- $y_1(\lambda) = \pm 1/4(\lambda - 5)$.

For example, the eigenvectors corresponding to $A_1 x = \lambda B x$ are given by $x = [1, -2]^T$ and $x = [1, 1]^T$ corresponding, respectively, to the generalized eigenvalues $-3$ and $1$.  

49
Chapter 4

The Real 3 by 3 Case

In this chapter we consider again the inverse generalized eigenvalue problem in the 3 by 3 case of the equation

$$Ax = \lambda Bx.$$  \hfill (4.1)

Necessary and sufficient conditions are given for the problem to have a solution.

In this chapter it is shown that all the principles used in 2 by 2 case carry over to the 3 by 3 real case; so we obtain results that are similar to those obtained in the 2 by 2 real case. Throughout this chapter we shall be working with a fixed nonsingular indefinite diagonal real 3 by 3 matrix B.
Let

\[
B = \begin{bmatrix}
  a_0 & 0 & 0 \\
  0 & a_1 & 0 \\
  0 & 0 & a_2 \\
\end{bmatrix}
\]

be nonsingular indefinite and \( \lambda(A, B) = \lambda_0, \lambda_1, \lambda_2 \) the spectrum of (4. 1) be given. One seeks a real symmetric matrix \( A \) of the form

\[
\begin{bmatrix}
  b_0 & c_0 & 0 \\
  c_0 & b_1 & c_1 \\
  0 & c_1 & b_2 \\
\end{bmatrix}
\]

such that:

- \( \hat{C}_1 \) (4. 1) holds for some \( x \neq 0 \);

- \( \hat{C}_2 \) \( A \) is a symmetric Jacobi positive definite matrix

In section 1 we formulate preliminary existence lemmas for the solution of the inverse problem, while in section 2 we give the construction of the solution. In section 2 we also give the proof of the maximum number of matrices (different solutions) we can find, and give a numerical example.
4.1 Necessary Conditions for Existence of a Solution

In addition to the lemma proved in chapter 3 we try to give another necessary condition concerning the eigenvalues of (4.1) and those of B for the inverse problem to have a solution. We note that since rank(B) = 3 and the eigenvalues of A - λB are distinct by assumption, then their corresponding eigenvectors are linearly independent. We denote by x_1, x_2, x_3 the eigenvectors of (4.1) corresponding to λ_1, λ_2, λ_3 respectively.

Lemma 4.1 If B is indefinite and all λ ∈ λ(A, B) of (4.1) are positive then A is not positive definite.

Proof
Suppose on the contrary, i. e., A is positive definite. Then

\[(Ax, x) = λ(Bx, x) > 0, \quad \text{for } x ≠ 0\]  \hspace{1cm} (4.2)

an eigenvector of (4.1). Since B is indefinite then there exist z ≠ 0 such that

\[(Bz, z) < 0.\]  \hspace{1cm} (4.3)

By (4.2), (Bx, x) > 0 for every eigenvector of (4.1) and there exist real scalars c_i, i = 1, 2, 3 not all zero such that

\[z = c_1x_1 + c_2x_2 + c_3x_3\]
\[(Bz, z) = (B(\sum_{i=1}^{3} c_{i}x_{i}), \sum_{j=1}^{3} c_{j}x_{j}) = \sum_{i=1}^{3} c_{i}^{2}(Bx_{i}, x_{i}) > 0\]

in contradiction with (4.3). Thus A is not positive definite. In other words the above lemma says that there exists at least one negative eigenvalue in \(\lambda(A, B)\) in order for the generalized inverse problem to have a positive definite solution matrix A.

**Lemma 4.2** If B is indefinite and all \(\lambda \in \lambda(A, B)\) of (4.1) are negative then A is not positive definite.

**Proof**

If A is positive definite then by the above argument

\[(Ax, x) = \lambda(Bx, x) > 0, \quad \text{for } x \neq 0 \quad (4.4)\]

an eigenvector of (4.1). Then \((Bx, x) < 0\), for every eigenvector of (4.1).

Since for every \(z \neq 0\) we can write

\[z = \sum_{i=1}^{3} c_{i}x_{i},\]

where \(c_{i}, \ i = 1, 2, 3\) are real scalars not all zero, we have

\[(Bz, z) = \sum_{i=1}^{3} c_{i}^{2}(Bx_{i}, x_{i}) < 0\]

whence

\[(Bz, z) < 0 \quad \text{for all } z \neq 0.\]
Therefore $B$ is negative definite, which leads to a contradiction. Thus $A$ is not positive definite. This lemma tells us that there exists at least one positive eigenvalue in $\lambda(A, B)$ in order for our problem to have a solution.

Now by the two above lemmas, (1. 15) and (1. 1) we have the following important result:

**Lemma 4.3** Let $B$ be an indefinite diagonal $3$ by $3$ real matrix and let $\lambda(A, B)$, the spectrum of (4. 1), be given. A necessary condition for the existence of a positive definite symmetric Jacobi matrix $A$ which solves the inverse problem is that the number of negative eigenvalues of (4. 1) equal the number of negative eigenvalues of $B$.

As in chapter 3, by section 2.1, (4. 1) can be written in the form (2. 1-2) by taking $m = 3$, viz.,

\[
\begin{align*}
c_0 y_1 &= (a_0 \lambda - b_0) y_0 \\
c_1 y_2 &= (a_1 \lambda - b_1) y_1 - c_0 y_0 \\
c_2 y_3 &= (a_2 \lambda - b_2) y_2 - c_1 y_1
\end{align*}
\]

subject to $y_3 = 0$ \hfill (4.5)
4.2 Construction of a Solution

Throughout this section we seek a matrix $A$ of the form

\[
\begin{bmatrix}
  b_0 & c_0 & 0 \\
  c_0 & b_1 & c_1 \\
  0 & c_1 & b_2
\end{bmatrix}
\]

satisfying $\hat{C}_1$, $\hat{C}_2$, and therefore (4.5). We will repeat the same procedure as in 2 by 2 case. In the following we assume:

$\hat{H}_1$) $a_0 \Delta_0 \Delta_1 > 0$

$\hat{H}_2$) Lemma 4.3 and (3.3) hold

$\hat{H}_3$) $a_1 \Delta_1 \Delta_2 > 0$

$\hat{H}_4$) $a_2 \Delta_2 \Delta_3 > 0$,

where we define $\Delta_i$ for $i = 0, 1, 2, 3$ as follows:

\[
\Delta_0 = 1
\]

\[
\Delta_1 = \mu_0
\]

\[
\Delta_2 = \text{det} \begin{bmatrix}
  \mu_0 & \mu_1 \\
  \mu_1 & \mu_2
\end{bmatrix}
\]

\[
\Delta_3 = \text{det} \begin{bmatrix}
  \mu_0 & \mu_1 & \mu_2 \\
  \mu_1 & \mu_2 & \mu_3 \\
  \mu_2 & \mu_3 & \mu_4
\end{bmatrix}
\]

Under these assumptions, there exists at least one positive definite symmetric Jacobi matrix $A$ which solves the inverse problem. The proof is given
in the two following theorems.

**Theorem 4.4** Let \( \tau(\lambda) \) be the step function defined in (2.26) and assume that \( \tilde{H}_1, \tilde{H}_2, \tilde{H}_3, \tilde{H}_4 \) hold. Then there exist three real polynomials \( y_0(\lambda), y_1(\lambda), y_2(\lambda) \) of degrees 0, 1, and 2, respectively, and such that the coefficients of \( \lambda^n \) \( (n = 0, 1, 2) \) are determined up to a change of sign.

Again, the proof is by "orthogonalization". We seek polynomials of the form:

\[
\begin{align*}
y_0(\lambda) &= k_0 \\
y_1(\lambda) &= k_1(\lambda + \alpha_{1,0}) \\
y_2(\lambda) &= k_2(\lambda^2 + \alpha_{2,1}\lambda + \alpha_{2,0}) \tag{4.6}
\end{align*}
\]

where \( k_i \neq 0, \ i = 0, 1, 2 \). First consider the solution of (2.27) with \( p \neq q \); for this it is necessary and sufficient that

\[
\int_{-\infty}^{\infty} y_0(\lambda) \lambda^q d\tau(\lambda) = 0, \quad q = 0, 1, \ldots, p - 1. \tag{4.7}
\]

For \( p' = 2 \), and \( q = 0, 1 \) we have

\[
\begin{align*}
\int_{-\infty}^{\infty} y_2(\lambda) d\tau(\lambda) &= 0 \\
\int_{-\infty}^{\infty} y_2(\lambda) \lambda d\tau(\lambda) &= 0.
\end{align*}
\]

Since \( \alpha_2 \neq 0 \) we have

\[
\alpha_{2,1} = \frac{\mu_3 \mu_0 - \mu_2 \mu_1}{\mu_1^2 - \mu_2 \mu_0} \tag{4.8}
\]
and
\[ \alpha_{2,0} = \frac{\mu_2^2 - \mu_1 \mu_3}{\mu_1^2 - \mu_2 \mu_0}. \] (4.9)

For \( p = 1, \ q = 0 \) we have
\[ \int_{-\infty}^{\infty} y_1(\lambda)d\tau(\lambda) = 0 \]
since \( k_1 \neq 0, \) we have
\[ \alpha_{1,0} = -\mu_1/\mu_0 \] \hfill (4.10)

For \( p = q = 2, \) (2.1) gives
\[ k_2^2 = a_2^{-1}[\int_{-\infty}^{\infty} (\lambda^2 + \alpha_{2,1} \lambda + \alpha_{2,0})^2d\tau(\lambda)]^{-1} \]
\[ = [a_2(\mu_4 + \mu_3 \alpha_{2,1} + \mu_2 \alpha_{2,0})]^{-1} \] \hfill (4.11)

which is positive by \( \tilde{H}_4. \)

For \( p = q = 1, \) we have
\[ k_1^2 = a_1^{-1}[\int_{-\infty}^{\infty} (\lambda + \alpha_{1,0})^2d\tau(\lambda)]^{-1} \]
\[ \tilde{\omega}_0^2 = \frac{\mu_0}{a_1(\mu_0 \mu_2 - \mu_3^2)} \] \hfill (4.12)

which is positive by \( \tilde{H}_3. \)

For \( p = q = 0 \) we get
\[ k_0^2 = a_0^{-1}[\int_{-\infty}^{\infty} d\tau(\lambda)]^{-1} \]
\[ = 1/a_0 \mu_0 > 0 \] \hfill (4.13)

by \( \tilde{H}_1. \)

Now since we have found the three polynomials we are able to determine the matrix \( A. \)
Theorem 4.5 Under the same assumptions as in theorem 4.4 there exist \( c_0, c_1, b_0, b_1, b_2 \) such that the boundary problem (4.5) has \( \tau(\lambda) \) as its spectral function.

Proof
Since \( y_0 \) and \( y_1 \) are known, so are \( c_0 \) and \( b_0 \) from (4.5) for we can choose

\[
c_0 = a_0 k_0 / k_1
\] (4.14)

and

\[
b_0 = -a_0 \alpha_{1,0}.
\]

By (4.10) we have

\[
b_0 = a_0 \mu_1 / \mu_0.
\] (4.15)

To find \( c_1 \), we consider the expansion

\[
\lambda_n a_n y_n(\lambda) = \sum_{r=0}^{n+1} \gamma_{n,r} y_r(\lambda) \quad \text{for } 1 \leq n \leq m - 2.
\] (4.16)

It is not difficult to determine \( \gamma_{n,r} \) and see that \( \gamma_{n,n+1}, \gamma_{n,n-1} \) have the form \( c_n, c_{n-1} \), respectively, and \( \gamma_{n,r} = 0, \ r < n - 1. \)

So we may use the Fourier process to determine the \( \gamma_{n,r} \), since

\[
\int_{-\infty}^{\infty} y_s(\lambda) a_n y_n(\lambda) \lambda d\tau(\lambda) = \sum_{r=0}^{n+1} \gamma_{n,r} \int_{-\infty}^{\infty} y_s(\lambda) y_r(\lambda) d\tau(\lambda)
\]

Thus by (2.27), we have

\[
\gamma_{n,r} = a_n a_r \int_{-\infty}^{\infty} \lambda y_n(\lambda) y_r(\lambda) d\tau(\lambda)
\] (4.17)

hence \( \gamma_{n,n+1} = c_n, \gamma_{n,n-1} = c_{n-1} \) where

\[
c_n = a_n a_{n+1} \int_{-\infty}^{\infty} \lambda y_n(\lambda) y_{n+1}(\lambda) d\tau(\lambda).
\] (4.18)
Now from (4.18) we see that

$$\lambda y_n(\lambda) = (k_n/k_{n+1})y_n(\lambda) + \text{lower powers of } \lambda$$

and so by the orthogonality (4.7) we deduce that

$$\int_{-\infty}^{\infty} \lambda y_n(\lambda)y_{n+1}(\lambda)d\tau(\lambda) = i(k_n/k_{n+1}) \int_{-\infty}^{\infty} \{y_{n+1}(\lambda)\}^2d\tau(\lambda), \quad (4.19)$$

whence (4.18), (4.19), and (2.27) give

$$c_n = a_n k_n/k_{n+1} \quad 1 \leq n \leq m - 2. \quad (4.20)$$

It remains only to prove that $\gamma_{n,r} = 0$, for $r < n - 1$. This results immediately from (4.17), since $\lambda y_r(\lambda)$ is then of lower degree than $y_n(\lambda)$, and so orthogonal to it by (4.7).

Since $m = 3$ by (4.20), we have

$$c_1 = a_1 k_1/k_2. \quad (4.21)$$

For $n = 0$, (4.18) and hence (4.20) are still in force, that is, $c_0$ can be obtained from (4.18).

Furthermore, since $c_1$ is known, by (4.5) and (4.21), we see by comparing the coefficients that we can choose

$$b_1 = \frac{a_1 k_1^2 \alpha_{2,0} + a_0 k_0^2}{k_1^2 \alpha_{1,0}}, \quad (4.22)$$

where $\alpha_{2,0}$ and $\alpha_{1,0}$ are defined in (4.9) and (4.10), respectively. It is left only to determine $b_2$; to do this let us first find $c_2$. The method of finding $c_2$ is to be modified to that used in finding $c_0$ and $c_1$. We have to determine
$y_3(\lambda)$ and $c_2$, $b_2$ such that $y_3(\lambda)$ has $\{\lambda_0, \lambda_1, \lambda_2\}$, the points of the spectrum of (4.1) as its zeros, and such that

$$c_2y_3(\lambda) = (a_2\lambda - b_2)y_2(\lambda) - c_1y_1(\lambda). \quad (4.23)$$

We define in this case

$$y_3(\lambda) = (\lambda - \lambda_0)(\lambda - \lambda_1)(\lambda - \lambda_2)$$
$$= \lambda^3 - \sigma_2\lambda^2 + \sigma_1\lambda - \sigma_0 \quad (4.24)$$

so that $y_3(\lambda)$ will have the correct zeros, where $\sigma_0 = \lambda_0\lambda_1\lambda_2$, $\sigma_1 = \lambda_0\lambda_1 + \lambda_0\lambda_2 + \lambda_1\lambda_2$, $\sigma_2 = \lambda_0 + \lambda_1 + \lambda_2$. There still holds an identity

$$\lambda c_2y_2(\lambda) = \sum_{r=0}^{3} \gamma_2 y_r(\lambda), \quad (4.25)$$

where $\gamma_{2r}$ is now to be determined by comparing the coefficients in (4.25). This shows that $\gamma_{23} = c_2$; more precisely,

$$c_2 = a_2k_2. \quad (4.26)$$

From (4.23-25) we may choose

$$b_2 = \frac{a_2k_2^2\sigma_2 - a_1k_1^2\sigma_{1,0}}{k_2^2\sigma_{2,0}} \quad (4.27)$$

Therefore:

$$A = \begin{bmatrix}
  a_0\mu_1/\mu_0 & a_0k_0/k_1 & 0 \\
  a_0k_0/k_1 & \frac{a_1k_1^2\sigma_{2,0}+a_0k_2^2}{k_2^2\sigma_{1,0}} & a_1k_1/k_2 \\
  0 & a_1k_1/k_2 & \frac{a_2k_2^2\sigma_2-a_1k_2^2\sigma_{1,0}}{k_2^2\sigma_{2,0}}
\end{bmatrix} \quad (4.28)$$

60
Finally, we have to prove that $A$ is positive definite. By (1.1) it suffices to prove that

$$D_i > 0, \quad i = 1, 2, 3.$$  

By $\tilde{H}_1$ and (3.3),

$$D_1 = a_0 \mu_1 / \mu_0 > 0.$$ 

$$D_2 = \begin{vmatrix} a_0 \mu_1 / \mu_0 & a_0 k_0 / k_1 \\ a_0 k_0 / k_1 & -a_1 k_0^2 a_2 \sigma_3 (\sigma_0 k_0^2) / k_1^2 \end{vmatrix}$$

$$= a_0 a_1 \frac{\mu_1 \mu_3 - \mu_2^2}{\mu_2 \mu_0 - \mu_1^2}.$$ 

Since

$$\mu_1 \mu_3 - \mu_2^2 = \rho_2^2 \rho_0 \rho_1 \lambda_0 \lambda_1 (\lambda_1 - \lambda_0)^2 + \rho_1^2 \rho_0 \rho_2 \lambda_0 \lambda_2 (\lambda_0 - \lambda_2)^2 + \rho_0^2 \rho_1 \rho_2 \lambda_1 \lambda_2 (\lambda_1 - \lambda_2)^2 > 0,$$

by (3.3), then by $\tilde{H}_1$ and $\tilde{H}_3$, $D_2 > 0$.

Note that $D_3 = \text{det}(A)$, which can be computed directly from (4.27) or simply by (1.15):

$$D_3 = a_0 a_1 a_2 \sigma_0$$

$$= a_0 a_1 a_2 \lambda_0 \lambda_1 \lambda_2 > 0$$

by lemma (4.3).

The same argument used in chapter 3 in order to calculate the number of different positive definite symmetric Jacobi matrices works here, but since it is too technical in the 3 by 3 case, it will be easier to use the following proof.

61
Since the polynomials $y_0(\lambda)$, $y_1(\lambda)$, $y_2(\lambda)$, $y_3(\lambda)$ are uniquely determined apart from constant factors $"k_n"," n = 0, 1, 2$ which are determined up to a change of sign and all the diagonal entries of $A$ do not depend on the sign of $"k_n"," n = 0, 1, 2$, the number of matrices depends only on the different $c_i's$, $i = 0, 1$ that we have. Hence, there exist at most four different positive definite symmetric Jacobi matrices $A_i$ satisfying (4.1), i.e.,

$A_1$ with $c_0 > 0$ and $c_1 > 0$

$A_2$ with $c_0 > 0$ and $c_1 < 0$

$A_3$ with $c_0 < 0$ and $c_1 > 0$

$A_4$ with $c_0 < 0$ and $c_1 < 0$.

**Example**

Let

$$B = diag(1, -1, 1/2),$$

$\lambda(A, B) = \{2, 5, -1\}$, $\rho_0 = 1/2$, $\rho_1 = 1$ and $\rho_2 = -2$. Then by (3.6) we have

$$\mu_0 = 5/2$$

$$\mu_1 = 19/2$$

$$\mu_2 = 65/2$$

$$\mu_3 = 283/2$$

$$\mu_4 = 1313/2$$

so that $\tilde{H}_1, \tilde{H}_2, \tilde{H}_3, \tilde{H}_4$ are satisfied. Thus by (4.27) we obtain
\[
A_1 = \begin{bmatrix}
19/5 & 6/5 & 0 \\
6/5 & S36/95 & -3\sqrt{5} \\
0 & -3\sqrt{5} & 11/2 \\
19/5 & -6/5 & 0 \\
\end{bmatrix}
\]

\[
A_2 = \begin{bmatrix}
-6/5 & S36/95 & 3\sqrt{5} \\
0 & 3\sqrt{5} & 11/2 \\
19/5 & 6/5 & 0 \\
\end{bmatrix}
\]

\[
A_3 = \begin{bmatrix}
6/5 & S36/95 & 3\sqrt{5} \\
0 & 3\sqrt{5} & 11/2 \\
19/5 & -6/5 & 0 \\
\end{bmatrix}
\]

\[
A_4 = \begin{bmatrix}
-6/5 & S36/95 & -3\sqrt{5} \\
0 & -3\sqrt{5} & 11/2 \\
\end{bmatrix}
\]

which are all positive definite since

\[
D_i^i = 3.8 \\
D_i^2 = 32 \\
D_i^3 = \text{det}(A_i) \\
= 5
\]

for \(i = 1, 2, 3, 4\), and satisfy (4.1) with \(Y = (y_0(\lambda), y_1(\lambda), y_2(\lambda))^T\), where

\[
y_0(\lambda) = \pm \frac{\sqrt{2}}{\sqrt{5}}
\]
\[ y_1(\lambda) = \pm \sqrt{5}/3 \sqrt{2}(\lambda - 19/5) \]
\[ y_2(\lambda) = \pm 1/3 \sqrt{2}(\lambda^2 + 5\lambda - 32). \]

For example, the eigenvectors corresponding to \( Ax = \lambda Bx \) are given by:
\[ x = (\sqrt{2}/\sqrt{5}, -3/\sqrt{10}, -\sqrt{2})^T, \]
\[ x = (\sqrt{2}/\sqrt{5}, \sqrt{2}/\sqrt{5}, \sqrt{2})^T \]
and
\[ x = (\sqrt{2}/\sqrt{5}, -4\sqrt{5}, -2\sqrt{2})^T \]
corresponding to the generalized eigenvalues 2, 5 and -1, respectively.
Bibliography


[18] Stewart, G. W. *On the sensitivity of the eigenvalue problem* $Ax = \lambda Bx$