

Geometric and function analytic approaches to the study of Sturm-Liouville theory.¹

Thomas Roth

0505606N

Under the supervision of Professor Bruce A. Watson

School of Mathematics

University of the Witwatersrand

Private Bag 3, P O WITS 2050, South Africa

February 7, 2012

¹Submitted to the University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Abstract

Ordinary second order linear differential boundary value problems are of great interest in the study of physical systems, despite this and the length of their history they have many aspects which have only recently been understood. The aim of this work is to examine the geometric and functional analytical approaches to the study of Sturm-Liouville eigenvalue problem with separated boundary conditions. In particular, the following sub-problems will be investigated: Self-adjointness, oscillation theory, variational formulations and the completeness of eigenfunction expansions. The construction of Self-adjoint operators will be studied through deficiency indices and through symplectic geometry. After the classical variational problem is identified and studied, Lusternik-Schnirelmann theory which is a topologically invariant variational principle will be examined.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Thomas Roth

This the 30th of August 2011 at Johannesburg, South Africa.

Acknowledgements

I would like to thank my family and friends, especially my mother and father for their support. I would like to thank my supervisor, Bruce Watson, for his patience and dedication to helping me along my mathematical journey. Last but not least, I would like to thank my cat, Simba, for keeping my papers in order and making sure they didnt blow away in the wind.

Contents

1	Introduction	1
1.1	Basic Overview	1
1.2	History	2
1.3	Chapter Structure	4
2	Functional Analysis Approaches	6
2.1	Introduction	6
2.2	Operator Theory	7
2.3	Spectral Theory	8
2.4	Deficiency Indices	14
2.5	Self-Adjoint Extensions	19
3	Eigenvalue asymptotics	28
3.1	Introduction	28
3.2	The Sturm-Liouville Expansion	29
3.3	Liouville Transformation	37

3.4	Eigenvalue Asymptotics	39
3.5	Properties of the Sturm-Liouville expansion	45
4	The Sturm theorems	48
4.1	Introduction	48
4.2	Ince	49
4.3	Coddington and Levinson	55
5	Variational approach	61
5.1	Introduction	61
5.2	Sobolev space and distributions	61
5.3	Variational Formulation	64
5.4	Lusternik-Schnirelmann method	70
5.4.1	Lusternik-Schnirelmann Category	71
5.4.2	The projective Space	73
5.4.3	Lusternik-Schnirelmann Theorems	76
5.5	Courant Nodal Domain Theorem	87
6	Classes of Self-Adjoint operators	89
6.1	Introduction	89
6.2	Symplectic Structure	90
6.3	Complete Lagrangian subspaces	93

<i>CONTENTS</i>	vi
6.4 Boundary space and the GKN-theorem	96
6.5 A classification of self-adjoint boundary conditions	99
7 Conclusion	104
7.1 Summary	104
7.2 Future Directions	105
A Grönwall's Lemma	107

Chapter 1

Introduction

1.1 Basic Overview

The focus of this dissertation is the interplay between geometric and function analytic approaches to the study of the differential equation

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + (\lambda - q(x))y = 0, \quad x \in [a, b], \quad (1.1)$$

with boundary conditions

$$y(a) \cos \alpha - p(a)y'(a) \sin \alpha = 0, \quad (1.2)$$

$$y(b) \cos \beta - p(b)y'(b) \sin \beta = 0. \quad (1.3)$$

Here p, q and p' are real and continuous on the closed interval $[a, b]$, $-\infty < a < b < \infty$, with $p(x) > 0$ and α and β are real. The associated differential operator, L , is given by

$$L = -\frac{d}{dx} \left[p(x) \frac{d}{dx} \right] + q(x), \quad (1.4)$$

with domain

$$\mathcal{D}(L) = \{f \in \mathcal{L}^2([a, b]) : f, Lf \in AC, f \text{ satisfies (1.2) and (1.3)}\}. \quad (1.5)$$

Finding the values of λ for which (1.1), (1.2), (1.3) has a non-identically-zero solution, will be referred to as the Sturm-Liouville eigenvalue problem. Prominent methods for studying this problem will be reviewed and an attempt made to lay down a unified framework that shows the connection between the different approaches.

1.2 History

It has been shown, see chapter 3, that a second order linear ordinary differential equation (1.1 with $p(x) > 0$ and $p(x)$, $p'(x)$ and $q(x)$ continuous on (a, b) , can be represented in the Sturm-Liouville normal form

$$-y'' + qy = \lambda y,$$

by suitable transformation of the dependent and independent variables (the so called Liouville transformation). Many differential equations appearing in physics can be transformed or decomposed into Sturm-Liouville problems, thus making Sturm-Liouville problems an important area of study for mathematical physics.

The first developments in this area were carried out by Jacques Sturm, see [39] and [38]. Here he investigated the qualitative properties of the eigenfunctions. This work was expanded on by Joseph Liouville, see [25], [26] and [27], in which he studied eigenfunction expansions and their convergence. For a more detailed history of Sturm-Liouville theory see [29].

The problem (1.1)-(1.3) can be reposed as an operator eigenvalue problem

$$Lx = \lambda x, \quad x \in D(L), \quad (1.6)$$

where L , as defined in (1.4)-(1.5), is a closed densely defined operator in a suitable Hilbert space. Here λ is an eigenvalue and x is an eigenfunction if $x \neq 0$ and (1.6) is obeyed. This formulation will be studied in greater detail in chapter 2.

An important part of the the theory of regular Sturm-Liouville operators is that the eigenvalues form a countable set of real numbers which is bounded below and has no

finite accumulation point and can thus be listed as an increasing sequence,

$$\lambda_0 < \lambda_1 < \dots < \lambda_n < \dots < \infty. \quad (1.7)$$

The eigenspace of the eigenvalue, λ_n , is one-dimensional and consists of scalar multiples of an eigenfunction, $f_n(x)$, which has n zeros in the interval (a, b) . The sequence (f_n) can be normalized so as to form an orthonormal basis for $\mathcal{L}^2([a, b])$.

The relative location of the zeros of these eigenfunctions was studied by Sturm and is formulated in what is now known as Sturm's separation, comparison and oscillation theorems, see Ince, [18].

These results have been reproved in various different contexts and by various methods, see Ince, [18], Coddington and Levinson, [3]. These proofs are more function analytic by nature. Ince's proof makes heavy use of the Picone's formula, which can be seen as Green's formula in a different form, see chapter 4. In Coddington and Levinson, [3], a Prüfer angle approach is taken. A direct geometric formulation of the above eigenvalue problem was carried out in terms of the, so called, Prüfer angle by Prüfer and Kamke, [21], and studied via algebraic topology by Arnold, [1]. Edwards, in [9], connected the classical analytic approaches mentioned above with such geometric concepts as bundles, homotopy classes and orbits.

The development of linear operator theory by Hilbert and von Neumann, gave rise to a general spectral theory, discussed in [28]. This is applicable to (1.1)-(1.3) when formulated as an operator eigenvalue problem (1.6). As a consequence apparently dissimilar expansion formulae can be unified as eigenfunction expansions corresponding to different second-order differential equations, see [41]. The spectral theorem for self-adjoint (unbounded) operators in a Hilbert space, see [17, pp. 257], can be seen as a generalization of the Hilbert-Schmidt theorem for compact self-adjoint operators, see [17, pp. 202]. We note that the spectra of operators such as the Sturm-Liouville operator above are of interest in quantum mechanics where the eigenvalues represent the energy levels of the quantum system which the operator represents, see [40, chapter 9].

To use the spectral theorem to study generalized eigenfunction expansions associated with an ordinary differential equation, first an appropriate self-adjoint operator, L , must be found, which represents the boundary eigenvalue problem for the differential equation and the imposed boundary conditions, see (1.4)-(1.5). This relation is examined by Kodaira, [22], with special attention to the singular case. A more geometric point of view is investigated by Markus and Everitt, [10], with the use of symplectic geometry.

1.3 Chapter Structure

In chapter 1.2 we give a brief history of the subject with references to some of the important contributions that relate to the work presented.

Chapter 2 focuses on the functional analytic background for Sturm-Liouville operators on a finite interval. The self-adjointness and domain of the operator (1.4) will be given detailed consideration here. Also considered are eigenvalue problems and the existence and uniqueness of solutions for their associated non-homogeneous problems. These developments will be used in chapters 3 and 6.

In chapter 3 we introduce the Sturm-Liouville expansion and study its convergence properties. Also considered are some common eigenvalue asymptotic approximations.

We examine in chapter 4 the qualitative properties of the eigenfunctions in the form of the Sturm Oscillation theorem. This theorem is proved using analytic and geometric arguments. These classical results prepare the reader for the variational approach to eigenvalue problems considered in chapter 5.2.

In chapter 5 the calculus of variations and topological considerations are used to obtain a general minimax principle. This also yields a partial proof for the Sturm Oscillation theorem known as the Courant nodal domain theorem.

The geometry of the boundary symplectic space is the focus of chapter 6. Here

self-adjoint boundary conditions are characterized as Lagrangian subspaces of the boundary space. The Group of transformations of this space is also considered.

Chapter 7 gives an overview of the work presented as well as possible future research directions.

Chapter 2

Functional Analysis Approaches

2.1 Introduction

Functional analysis in its modern form has seen a fairly rapid development in the last century. The acceleration of its development began in 1900 with pioneering work on integral equations carried out by Fredholm, see [12]. Fredholm's insights were embraced by David Hilbert who saw that integral equations were an abstraction of systems of linear equations. In 1912 he published his famous papers, see [15], where he extended the principle axis theorem to integral equations and showed that arbitrary functions could be represented linear combinations of eigenfunctions. Further rigour was introduced through abstraction by Fréchet with the introduction of functional calculus and function spaces which were endowed with a suitable metric. The theory of function spaces reached its maturity with Banach's thesis published in 1922 on abstract treatment of normed function spaces which are complete, see [2]. Although not mentioned here in detail, other mathematicians that played an important role in the formulation of functional analysis are Lebesgue, Riesz, Fourier and Titchmarsh. Functional analysis has become an indispensable tool in studying differential equations. In this section we will only look at those aspects of functional analysis relevant to the study of the regular boundary value problems described above.

2.2 Operator Theory

A linear operator L densely defined in a Hilbert space H is bounded if and only if it is continuous on its domain. In this case the operator L has a unique continuous extension to the whole of H and, as is customary, it will be assumed that this extension has been made, giving $\mathcal{D}(L) = H$. For a bounded linear operator L in H the image, under L , of the closed unit ball in H is norm-bounded and the infimum of the norm-bounds is denoted $\|L\| < \infty$ and is called the norm of L .

An operator L is said to be closed if for each convergent sequence (x_n) in $\mathcal{D}(L)$ with (Lx_n) convergent, we have that $x = \lim_{n \rightarrow \infty} x_n \in \mathcal{D}(L)$ and $Lx = \lim_{n \rightarrow \infty} Lx_n$.

If L is bounded and $\mathcal{D}(L)$ is closed then the operator L is also closed, see [17, pp. 103]. The graph $G(L)$ of L is the linear subspace of $H \times H$ consisting of all elements of the form $[f, Lf]$, where $f \in \mathcal{D}(L)$. The inverse graph, $G'(L)$, consists of the elements of $H \times H$ of the form $[Lf, f]$, where $f \in \mathcal{D}(L)$. The Closed Graph Theorem states that if L is closed and $\mathcal{D}(L) = H$ then L is bounded, see [17, pp. 106].

The dual space H^* of the Hilbert space H is the linear space consisting of all continuous linear functionals on H . The Riesz representation theorem, see [17, pp. 168], gives a conjugate linear isometry between the dual space H^* and H . In particular to each element $g \in H$ there is a unique $g^* \in H^*$ given by $g^*(f) = (f, g)$ for each $f \in H$. The map given by $g \rightarrow g^*$ between H and H^* is isometric and surjective. Thus the points of H and H^* can be identified with one another making each Hilbert space self-dual.

We recall, see [24], that the closed unit ball in a Hilbert space is compact if and only if the Hilbert space is finite dimensional.

The adjoint of the densely defined operator L is defined to be the operator L^* with domain

$$\mathcal{D}(L^*) = \{g \in H : (Lf, g) = (f, h) \text{ for all } f \in \mathcal{D}(L), \text{ for some } h \in H\}$$

and, for $g \in \mathcal{D}(L^*)$, $L^*(g)$ is uniquely specified by

$$(Lf, g) = (f, L^*g) \quad \text{for all} \quad f \in \mathcal{D}(L). \quad (2.1)$$

It is essential that L is densely defined as otherwise the adjoint is not uniquely defined.

The adjoint of a densely defined linear operator in a Hilbert space is a closed operator. This follows from the fact that $G'(-L^*) = G(L)^\perp$, see [17, pp. 182]. Here $G(L)^\perp$ is the orthogonal complement of $G(L)$ in $H \times H$ which is a closed set. A set S is said to be compact if each open cover of S has a finite subcover. The set S is relatively compact if $\bar{S}$ is compact. In the setting of metric spaces (and thus for normed space) compactness and sequentially compact are equivalent. A set S is sequentially compact if every sequence in S has a convergent subsequence with limit in S . An operator L is compact if it maps bounded sets to relatively compact sets. It should be noted that the adjoint of a compact operator on a Hilbert space is again compact, see [17, pp. 195].

A set S of real valued functions on $[a, b]$ is said to be equicontinuous if, for every $\epsilon > 0$, there is a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $f \in S$ and $x, y \in [a, b]$ with $|x - y| < \delta$. In finite dimensional normed spaces closed and bounded sets are compact, but this is not in general true in infinite dimensional normed spaces. A useful characterization of compact sets of continuous functions on an interval $[a, b]$ is given by the Arzelà-Ascoli theorem. This states that a set $S \subset \mathcal{C}([a, b])$ is uniformly bounded and equicontinuous if and only if S is relatively compact, see [17, pp. 154].

2.3 Spectral Theory

Spectral theory is concerned with studying the properties of the equation

$$(\lambda I - L)f = g, \quad (2.2)$$

as λ and g vary. Our interest is in the particular case where L is as defined in (1.4), $f \in \mathcal{D}(L)$ and $g \in \mathcal{L}^2([a, b])$. It is of significance to know for which λ values the

above equation has a solution and whether that solution is a unique solution. If the map $\lambda I - L$ is surjective and one-to-one on its domain then $(\lambda I - L)^{-1}$ exists. If, in addition, $(\lambda I - L)^{-1}$ is continuous, denote

$$R(\lambda; L) = (\lambda I - L)^{-1}, \quad (2.3)$$

and call this operator the resolvent of L . The set $\rho(L)$ of $\lambda \in \mathbb{C}$ for which the resolvent map $R(\lambda; L)$ exists (and is bounded), is called the resolvent set of L . The set $\sigma(L)$ which is the complement of the resolvent set in $\mathbb{C}$ is called the spectrum of L . The spectrum can be broken down into three parts:

(a) the point spectrum $\sigma_p(L)$, which consists of eigenvalues $\lambda \in \mathbb{C}$, that is λ for which the equation,

$$\lambda f = Lf, \quad (2.4)$$

has a non-zero solution f , called an eigenfunction;

(b) the continuous spectrum $\sigma_c(L)$, which consists of $\lambda \in \sigma(L)$ for which $(\lambda I - L)$ is one-to-one on its domain with $\mathcal{R}(\lambda I - L)$ dense in H ;

(c) the residual spectrum, $\sigma_r(L)$, consists of those $\lambda \in \mathbb{C}$ for which $(\lambda I - L)$ is one-to-one on its domain but for which $\mathcal{R}(\lambda I - L)$ is not dense in H .

It will be shown later that for the linear operator L , see (1.4)-(1.5), in the Hilbert space $\mathcal{L}^2[0, 1]$ associated with the problem (1.1), the resolvent operator is compact. This will assist with the characterization of the spectrum of L and the study of the expansion of an arbitrary function from $\mathcal{L}^2[0, 1]$ in terms of the eigenfunctions of L .

Theorem 2.1 The Fredholm Alternative Theorem¹

Let T be a compact linear operator in a Hilbert space H . If $\lambda \neq 0$, then one of the two alternatives holds:

(i) The homogeneous equation, $(\lambda I - T)f = 0$ has only the zero solution. In this case $\lambda \in \rho(T)$, and the resolvent operator $R(\lambda; T)$ exists on $\mathcal{H}$.

(ii) The homogeneous equation, $(\lambda I - T)f = 0$, has a non-zero solution and thus

¹cf. [17, pp. 198]

$\lambda \in \sigma_p(T)$. In this case there is a solution of $(\lambda I - T)f = g$ if and only if $g \in N(\lambda I - T^*)^\perp$.

Theorem 2.2 *Let T be a compact operator on a Hilbert space H . If $\lambda \neq 0$ then $N(\lambda I - T)$ is finite dimensional.*

Proof: Let S be a closed unit ball in $N(\lambda I - T)$. Since T is compact, $T(S)$, is relatively compact, so for each sequence $\{f_n\}$ in S , there is a subsequence of $\{f_n\}$, which we again denote by $\{f_n\}$, so that $T(f_n)$ is convergent. However $Tf_n = \lambda f_n$, thus $\{\lambda f_n\}$ is convergent, giving that $\{f_n\}$ is convergent. So each sequence in S has a convergent subsequence in H and thus in S (as S is closed). Hence the closed unit ball in $N(\lambda I - T)$ is compact, making $N(\lambda I - T)$ finite dimensional. ■

Theorem 2.3 *Let T be a compact operator defined on $H = \mathcal{L}^2([a, b])$, then the spectrum of T consists only of discrete spectrum and the point 0. Here each non-zero eigenvalue has a finite dimensional eigenspace.*

Proof: Firstly, suppose $0 \notin \sigma(T)$, hence T^{-1} exists and is bounded, furthermore $\mathcal{R}(T) = H$. Thus T is a homeomorphism from H to H . Let S be a closed unit ball in H , then $T(S)$ is closed. Also as T is compact, $T(S)$ is relatively compact making $T(S)$ compact. Thus $S = T^{-1}T(S)$ is the continuous image of a compact set and is thus compact, making H finite dimensional, contradiction. Thus $0 \in \sigma(T)$.

Now for $\lambda \neq 0$, if $\lambda \notin \rho(T)$, then by theorem 2.1, we see that the only alternative is for λ to be an eigenvalue, i.e. in the point spectrum. Finally, theorem 2.2 completes the proof. ■

Theorem 2.4 *The set of eigenvalues of a compact operator T , is at most countably infinite with the only possible limit point being zero.*

Proof: Suppose on the contrary that for some $\epsilon > 0$ there is an infinite sequence $\{\lambda_n\}$ of distinct eigenvalues of T such that $|\lambda_n| \geq \epsilon$ for all n . Let $\{f_n\}$ be a sequence of eigenfunctions of T associated with the sequence λ_n of eigenvalues. Let $\mathcal{M}_n$ be the linear span of $f_1, \dots, f_n$. The eigenfunctions $f_1, \dots, f_n$ are linearly independent. To see this observe that if $f_1, \dots, f_n$ are linearly independent but $f_1, \dots, f_{n+1}$ are linearly dependent, then there are scalars $\alpha_1, \dots, \alpha_n$ so that

$$f_{n+1} = \sum_{i=1}^n \alpha_i f_i. \quad (2.5)$$

Multiplying (2.5) by λ_{n+1} gives

$$\lambda_{n+1} f_{n+1} = \sum_{i=1}^n \alpha_i \lambda_{n+1} f_i. \quad (2.6)$$

Applying T to (2.5) gives

$$\lambda_{n+1} f_{n+1} = \sum_{i=1}^n \alpha_i \lambda_i f_i. \quad (2.7)$$

Subtracting (2.7) from (2.6) yields

$$0 = \sum_{i=1}^n \alpha_i (\lambda_{n+1} - \lambda_i) f_i. \quad (2.8)$$

Since $f_1, \dots, f_n$ are linearly independent from (2.8) we conclude that

$$\alpha_i (\lambda_{n+1} - \lambda_i) = 0, \quad \text{for } i = 1, \dots, n.$$

By assumption $\lambda_{n+1} \neq \lambda_i, i = 1, \dots, n$, giving $\alpha_i = 0$ for all $i = 1, \dots, n$, which with (2.5) yields $f_{n+1} = 0$, a contradiction.

Thus $\mathcal{M}_n$ is an n dimensional space and is a proper subspace of $\mathcal{M}_{n+1}$. There is a sequence² (g_n) such that for each $n \geq 1$, $g_n \in \mathcal{M}_n$, $\|g_n\| = 1$ and

$$\|g_n - f\| \geq 1/2 \quad \text{for all } f \in \mathcal{M}_{n-1}. \quad (2.9)$$

²cf. [17, Hutson,Pym and Cloud, page 213], problem 7.7. This can be solved using the orthogonal decomposition of $\mathcal{M}$.

For each $n \geq 0$, we may represent $g_n = \sum_{i=1}^n \alpha_i f_i$ for some $\alpha_i \in \mathbb{C}$. Since g_n is a sum of the eigenfunctions $f_1, \dots, f_n$ we have that $Tf_n \in \mathcal{M}_n$ as well as

$$(\lambda_n I - T)g_n = \sum_{j=1}^{n-1} \alpha_j (\lambda_n - \lambda_j) f_j, \quad (2.10)$$

hence $(\lambda_n I - T)g_n \in \mathcal{M}_{n-1}$. Let $m, 1 \leq m < n$ be an integer and set $f = (\lambda_n I - T)g_n + Tg_m$, from (2.10) we see that $\lambda_n^{-1}f \in \mathcal{M}_{n-1}$. Here

$$Tg_n - Tg_m = \lambda_n g_n - (\lambda_n g_n - Tg_n + Tg_m) = \lambda_n (g_n - \lambda_n^{-1}f). \quad (2.11)$$

Now, using (2.9) and the above gives $\|Tg_n - Tg_m\| \geq \frac{1}{2}|\lambda_n| \geq \frac{1}{2}\epsilon$. Hence the operator T maps a bounded sequence (g_n) to a sequence (Tg_n) with no convergent subsequence, contradicting the compactness of T . ■

We now give the well know Hilbert-Schmidt Theorem, see [17].

Theorem 2.5 Hilbert - Schmidt

The eigenfunctions of a compact self-adjoint operator, T , on a Hilbert space $\mathcal{H}$, form an orthonormal basis for $\mathcal{H}$.

Proof: Let $\mathcal{M}$ be the closure of the linear span of the eigenfunctions, $\{\psi_n\}$, of T . Then, from a well known projection theorem, $\mathcal{H} = \mathcal{M}^\perp \oplus \mathcal{M}$. It is clear that $T\mathcal{M} \subset \mathcal{M}$ and for $f \in \mathcal{M}^\perp$, we have,

$$(Tf, h) = (f, Th) = 0, \quad \text{for all } h \in \mathcal{M}, \quad (2.12)$$

as $Th \in \mathcal{M}$. Thus $T(\mathcal{M}^\perp) \subset \mathcal{M}^\perp$.

Denote the restriction of T to $\mathcal{M}^\perp$ by T_0 , hence $T_0\mathcal{M}^\perp \subset \mathcal{M}^\perp$ and T_0 is compact and self-adjoint. Here T_0 has no eigenvalues, since otherwise T would have eigenvalues in $\mathcal{M}^\perp$. This is not possible by the definition of $\mathcal{M}$.

From theorem 2.3, the spectrum, $\sigma(T_0)$, may only contain zero and eigenvalues of T , giving $\sigma(T_0) \subset \{0\}$. Hence $r_\sigma(T_0) = 0$ and, since $\|T_0\| = r_\sigma(T_0)$, [17, page 100], we

see that $T_0 = 0$. This gives $T\mathcal{M}^\perp = 0$, which implies that every element of $\mathcal{M}^\perp$ is an eigenfunction of T to the eigenvalue 0. Thus $\mathcal{M}^\perp \subset \mathcal{M}$, giving $\mathcal{M}^\perp = 0$ and $\mathcal{H} = \mathcal{M}$.

■

With L defined as (1.4) and (1.5) from the Sturm-Liouville problem, although L is not compact or even bounded, $(\lambda I - L)^{-1}$ is compact if $\lambda I - L$ is one-to-one. Since L can be shown to be self-adjoint, it follows that the spectrum of L is contained in the reals. The spectrum of L consists of a sequence of real eigenvalues with no finite accumulation point. Let μ be a real number in the resolvent set of L . Now $(\mu I - L)^{-1}$ is self-adjoint and compact. Thus by the Hilbert-Schmidt Theorem the eigenfunctions, $\{\phi_n\}$, of $(\mu I - L)^{-1}$ form an orthonormal basis for $\mathcal{L}^2[a, b]$. A simple computation shows that each of the eigenfunctions ϕ_n is also an eigenfunction of L and visa versa. Consequently the eigenfunctions of L form an orthonormal basis for $\mathcal{L}^2[a, b]$. Hence for each $f \in \mathcal{L}^2[a, b]$ we have $f = \sum_{n=1}^{\infty} (f, \phi_n) \phi_n$ and

$$R(\lambda; T)f = \sum_{n=1}^{\infty} \frac{(f, \phi_n)}{\lambda - \lambda_n} \phi_n. \quad (2.13)$$

The spectral theorem³ is a generalization of the Hilbert-Schmidt theorem to closed unbounded operators. It gives a constructive way of obtaining a generalization of the orthonormal basis expansion, discussed above, for a closed self-adjoint operator, provided the spectrum is known and that there is an explicit formula for the resolvent operator. The spectral theorem asserts that the expansion of an arbitrary function in terms of generalized eigenfunctions (or more precisely projectors) can be calculated by integrating the resolvent around a contour containing the spectrum. For self-adjoint operators the spectral theorem yields the following formula,

$$\lim_{\lambda \rightarrow \infty} P(-\lambda, \lambda)f = f \quad (2.14)$$

³cf. [17, page.257]

where

$$(P(\lambda_1, \lambda_2)f, g) = \lim_{\delta \rightarrow 0^+} \lim_{\epsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_{\lambda_1 + \delta}^{\lambda_2 - \delta} (R(\lambda - i\epsilon : L) - R(\lambda + i\epsilon : L)f, g)d\lambda. \quad (2.15)$$

2.4 Deficiency Indices

In this section we outline the theory of deficiency indices and its application to constructing self-adjoint extensions of symmetric operators. A closed linear operator L_0 in $\mathcal{H}$ is symmetric if and only if its domain is dense in $\mathcal{H}$ and it satisfies the following equivalent conditions:

$$\begin{aligned} (i) \quad & (L_0 f, g) = (f, L_0 g), \quad f, g \in \mathcal{D}(L_0); \\ (ii) \quad & L_0 \subset L_0^*. \end{aligned}$$

When defining a self-adjoint operator we first consider a symmetric operator with a severely restricted domain. We then search for a linear subspace satisfying $\mathcal{D}(L_0) \subset M \subset \mathcal{D}(L_0^*)$ to which the operator L_0 has a self-adjoint extension. To determine whether this is possible it is advantageous to study the closed linear spaces

$$N_+ = \mathcal{N}(L_0^* - iI) \quad \text{and} \quad N_- = \mathcal{N}(L_0^* + iI).$$

The spaces N_+ and N_- are called the deficiency spaces of L_0 . The difference between N_+ and N_- measures the extent to which the operator L_0 fails to be self-adjoint, in that, if L_0 were self-adjoint then $N_+ = N_-$. Their dimensions, denoted by $n_{\pm}$ respectively, are called the deficiency indices of L_0 , and for simplicity we write them as (n_+, n_-) .

The next lemma, based on [17], clearly shows how the adjoint of L_0 is related to the deficiency spaces. The starting point is a well known decomposition of the Hilbert space $\mathcal{H}$ via the operator L_0 , see [31]. Let L_0 be a symmetric operator on $\mathcal{H}$ and let $\text{Im} \lambda \neq 0$ then $\mathcal{H}$ can be decomposed into a direct sum of closed orthogonal subspaces as follows,

$$\mathcal{H} = \mathcal{R}(L_0 - \lambda I) \oplus \mathcal{N}(L_0^* - \lambda I). \quad (2.16)$$

The remainder of this section is based on a reworking of parts of [17].

Lemma 2.6 *Let L_0 be a closed symmetric operator, then*

$$\mathcal{D}(L_0^*) = \mathcal{D}(L_0) \oplus N_+ \oplus N_-.$$

Proof: Let $f \in \mathcal{D}(L_0^*)$, then using the decomposition (2.16) we have

$$L_0^*f - if = (L_0g - ig) + h_- \quad \text{for some } g \in \mathcal{D}(L_0), h_- \in N_-. \quad (2.17)$$

Rearranging the above equation and adding $L_0^*(-\frac{1}{2}ih_-)$ to both sides gives

$$\begin{aligned} L_0^*(f - g - \frac{1}{2}ih_-) &= i(f - g) + h_- + L_0^*(-\frac{1}{2}ih_-) \\ &= i(f - g) + h_- - \frac{1}{2}h_- \quad \text{since } L_0^*h_- = -ih_-. \\ &= i(f - g - \frac{1}{2}ih_-). \end{aligned}$$

It is now clear that $f - g - \frac{1}{2}ih_- \in N_+$. Hence setting $\phi_- = \frac{1}{2}ih_-$ and $\phi_+ = f - g - \frac{1}{2}ih_-$ we have

$$f = g + \phi_- + \phi_+ \quad \text{for some } g \in \mathcal{D}(L_0), \phi_{\pm} \in N_{\pm}. \quad (2.18)$$

Since $\mathcal{D}(L_0)$ and $N_{\pm}$ are contained in $\mathcal{D}(L_0^*)$, we will be done if we can show that the representation is unique, that is if $f = 0$ implies $g = \phi_- = \phi_+ = 0$. Apply $L_0^* - iI$ to (2.18) and note that $(L_0^* - iI)\phi_+ = 0$ and $L_0^*\phi_- = -i\phi_-$. Thus

$$(L_0^* - iI)g - 2i\phi_- = 0.$$

The decomposition in (2.16) is as a direct summation yielding $g = \phi_- = 0$. Similarly, applying $L_0^* - iI$ to (2.18) we have $\phi_+ = 0$ which completes the proof. ■

A direct consequence of the last lemma is that the domain of any linear extension, L_1 , of L_0 satisfying $L_0 \subset L_1 \subset L_0^*$ can be described as $D(L_1) = D(L_0) \oplus N$ where N is a unique linear subspace of $N_- \oplus N_+$.

Definition 2.7 Let the bilinear form $\langle \cdot, \cdot \rangle$ on $\mathcal{D}(L_0^*) \times \mathcal{D}(L_0^*)$ be defined as follows

$$\langle f, g \rangle := (L_0^* f, g) - (f, L_0^* g) \quad \text{for } f, g \in \mathcal{D}(L_0^*). \quad (2.19)$$

Let the adjoint N^* of N be the linear subspace such that

$$N^* := \{f \in N_- \oplus N_+ : \langle f, g \rangle = 0 \text{ for } g \in N\}.$$

The space N is said to be symmetric if $N \subset N^*$. The space N is said to be self-adjoint if $N = N^*$.

Lemma 2.8 Let L_0 be a closed symmetric operator on a Hilbert space. A function $f \in \mathcal{D}(L_0^*)$ is also in $\mathcal{D}(L_0)$ if and only if the following equivalent conditions hold:

$$(i) \quad \langle f, g \rangle = 0, \quad \text{for all } g \in \mathcal{D}(L_0^*); \quad (2.20)$$

$$(ii) \quad \langle f, \phi \rangle = 0, \quad \text{for all } \phi \in N_- \oplus N_+. \quad (2.21)$$

Proof: We begin by proving the necessity and sufficiency of (i). As L_0 is closed we have $L_0^{**} = L_0$. Thus from the definition of the adjoint, $f \in \mathcal{D}(L_0^*)$ lies in $\mathcal{D}(L_0^{**}) = \mathcal{D}(L_0)$ if and only if

$$\langle f, g \rangle = 0, \quad \text{for all } g \in \mathcal{D}(L_0^*). \quad (2.22)$$

Now we prove the necessity and sufficiency of (ii). Necessity is a fortiori true since $N_- \oplus N_+ \subset \mathcal{D}(L_0^*)$.

For sufficiency, using Lemma 2.6, every $g \in \mathcal{D}(L_0^*)$ may be written as $g = g_0 + \phi$ for $g_0 \in \mathcal{D}(L_0)$ and $\phi \in N_- \oplus N_+$, thus (2.22) becomes

$$\langle f, g \rangle = \langle f, g_0 \rangle + \langle f, \phi \rangle.$$

Now $\langle f, \phi \rangle = 0$ by assumption while $\langle f, g_0 \rangle = 0$ since L_0^* is an extension of L_0 . Thus in total we have $\langle f, g \rangle = 0$ for all $g \in \mathcal{D}(L_0^*)$, giving condition (i). ■

Lemma 2.9 *Let L_0 be a closed symmetric operator in a Hilbert space and let L_1 be a linear operator satisfying $L_0 \subset L_1 \subset L_0^*$. Then*

$$\mathcal{D}(L_1) = \mathcal{D}(L_0) \oplus N$$

and

$$\mathcal{D}(L_1^*) = \mathcal{D}(L_0) \oplus N^*$$

where N is a linear subspace of $N_- \oplus N_+$. Furthermore L_1 is symmetric if and only if N is symmetric, and L_1 is self-adjoint if and only if N is self-adjoint.

Proof: From above we have $L_0 \subset L_1 \subset L_1^* \subset L_0^*$. Hence, by Lemma 2.6, we can decompose the domains of L_1 and its adjoint as $\mathcal{D}(L_1) = \mathcal{D}(L_0) \oplus N$ and $\mathcal{D}(L_1^*) = \mathcal{D}(L_0) \oplus P$ where N, P are linear subspaces of $N_- \oplus N_+$.

For $f \in \mathcal{D}(L_1^*)$ and $g \in \mathcal{D}(L_0^*)$, Lemma 2.6 shows that we may write $f = f_0 + \phi$ and $g = g_0 + \psi$ respectively, where $f_0, g_0 \in \mathcal{D}(L_0)$ and $\phi, \psi \in N_- \oplus N_+$. Hence

$$\begin{aligned} \langle f, g \rangle &= (L_0^* f, g) - (f, L_0^* g) = (L_1^* f, g) - (f, L_0^* g) \\ &= \langle f_0, g_0 \rangle + \langle \phi, g_0 \rangle + \langle f_0, \psi \rangle + \langle \phi, \psi \rangle. \end{aligned}$$

Thus by Lemma 2.8

$$\langle f, g \rangle = \langle \phi, \psi \rangle. \tag{2.23}$$

For g as above, Lemma 2.8 and (2.23) give $g \in \mathcal{D}(L_1)$ if and only if $\langle \phi, \psi \rangle = 0$ for all $\phi \in P$, i.e. $\phi \in P^*$ and $N = P^*$. ■

Let L_0 have finite deficiency indices $n_{\pm}$. From methods of linear algebra we have that if N is a linear subspace of $N_+ \oplus N_-$ of dimension n , then N^* has dimension $n_+ + n_- - n$. We can thus show that a symmetric operator has a self-adjoint extension if and only if its deficiency indices are equal.

A set of functions $f_1, \dots, f_k \in \mathcal{D}(L_0^*)$ where $f_i = h_i + \phi_i$ for $h_i \in \mathcal{D}(L_0)$ and $\phi_i \in N_+ \oplus N_-$, are linearly independent relative to $\mathcal{D}(L_0)$ if the functions $\phi_1, \dots, \phi_k$ are linearly independent.

Theorem 2.10 *Let L_0 be a closed symmetric operator with equal finite deficiency indices, $n := n_- = n_+$. Let $f_1, \dots, f_n \in \mathcal{D}(L_0^*)$ be linearly independent relative to $\mathcal{D}(L_0)$ with*

$$\langle f_i, f_j \rangle = 0 \quad \text{for all } i, j = 1, 2, \dots, n. \quad (2.24)$$

Let $\mathcal{M}$ be a linear subspace of $\mathcal{H} = \mathcal{L}^2([a, b])$ given by

$$\mathcal{M} := \{f \in \mathcal{D}(L_0^*) : \langle f, f_i \rangle = 0, i = 1, 2, \dots, n\}. \quad (2.25)$$

Then the restriction of L_0^* to $\mathcal{M}$, denoted M , is a self-adjoint extension of L_0 . Conversely, every self-adjoint extension K of L_0 with domain $\mathcal{K}$ can be obtained in this way.

Proof: From Lemma 2.6, each f_i can be represented as $f_i = g_i + \phi_i$ where $g_i \in \mathcal{D}(L_0)$ and $\phi_i \in N_+ \oplus N_-$. Let $N \subset N_+ \oplus N_-$ be the linear span of the ϕ_i 's, then by definition of N^* , condition (2.24) guarantees that N is symmetric. Hence $N \subset N^*$. Since the dimension of N^* and N are equal it follows that $N = N^*$ and N is self-adjoint. From Lemma 2.9, the previous argument shows that the operator M , corresponding to L_0^* restricted to the domain of $\mathcal{M} = \mathcal{D}(L_0) \oplus N$, is self-adjoint. Conversely, if $\mathcal{K}$ is the domain of a self-adjoint extension K of L_0 , then we have $\mathcal{K} = \mathcal{D}(L_0) \oplus K$, where K is self-adjoint. The result follows on constructing an appropriate basis for $\mathcal{K}$ satisfying (2.24). ■

For this theory to be applicable, a concrete representation of the bilinear form $\langle \cdot, \cdot \rangle$ and the domains of $\mathcal{D}(L_0)$, $\mathcal{D}(L_0^*)$ must be found. This is studied in the next section.

2.5 Self-Adjoint Extensions

In this section we will use the theory of deficiency indices to find an extension of the operator L to a domain, on which this extension is self-adjoint. To do this we find a maximal domain for which L can be extended and remain symmetric. In what follows we let $l(y)$ be the formal differential expression given by,

$$l(y) := -\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + q(x)y.$$

Here we require,

$$1/p, q \in \mathcal{L}_{Loc}^1([a, b]). \quad (2.26)$$

We will consider the non-homogeneous equation $\lambda y - l(y) = g$ and the related homogeneous equation $\lambda y - l(y) = 0$. We first prove an existence result which will be necessary for calculating the deficiency indices. Consider the matrix equation,

$$\bar{y}' = A(x)\bar{y} + \bar{g}(x), \quad (2.27)$$

where $\bar{y} = (y_1, \dots, y_n)$ is a vector with components $y_m \in AC$ for $1 \leq m \leq n$. Let $\bar{g}(x) = (g_1(x), \dots, g_n(x))$ and $A(x) = (a_{ij}(x))$ be such that $a_{ij}, g_k \in \mathcal{L}_{Loc}^1([a, b])$ for $1 \leq i, j, k \leq n$.

Denote $|A(x)| := \sum_i \max_j |a_{ij}(x)|$. The following two theorems and their proofs are based on [41].

Theorem 2.11 *There exists a unique solution to (2.27) which satisfies the initial condition*

$$y(x_0) = \bar{y}_0 \quad \text{for} \quad \bar{y}_0 \in \mathbb{R}^n \text{ and } x_0 \in (a, b). \quad (2.28)$$

Proof: For any function $\bar{y}$ satisfying (2.27) and (2.28), using the initial condition (2.28) and integrating equation (2.27) gives

$$\bar{y}(x) = \bar{y}_0 + \int_{x_0}^x (A(\xi)\bar{y}(\xi) + \bar{g}(\xi)) d\xi. \quad (2.29)$$

Since $\bar{y}(\xi)$ is absolutely continuous, it is the integral of a locally integrable function on $[x_0, x]$. Hence it is continuous and bounded, so the integral on the right of (2.29) exists. We see that $\bar{y}$ is a solution of (2.27) satisfying (2.28) if and only if it is a solution of (2.29). It is therefore sufficient to show the existence of a solution to (2.29). We proceed by the method of successive approximations.

Define recursively

$$\bar{y}_0(x) \equiv \bar{y}_0 \quad (2.30)$$

$$\bar{y}_{n+1}(x) := \bar{y}_0 + \int_{x_0}^x A(\xi)\bar{y}_n(\xi) + \bar{g}(\xi)d\xi \quad \text{for } n = 0, 1, 2, 3, \dots \quad (2.31)$$

For $x > x_0$ let,

$$M_n(x) := \max_{x_0 \leq t \leq x} |y_{n+1}(t) - y_n(t)|$$

$$q(x) := \int_{x_0}^x |A(t)|dt.$$

From equation (2.31) we have

$$\bar{y}_{n+1}(x) - \bar{y}_n(x) = \int_{x_0}^x A(\xi)[\bar{y}_n(\xi) - \bar{y}_{n-1}(\xi)]d\xi. \quad (2.32)$$

We would like to show that the series of elements of the form (2.32) converges, to this end we will prove by induction that

$$M_n(x) \leq \frac{q^n(x)}{n!} M_0(x). \quad (2.33)$$

For $n = 1$ we see that equation (2.32) gives

$$|\bar{y}_1(t) - \bar{y}_0(t)| \leq M_0(t) \int_{x_0}^t |A(\xi)|d\xi \leq M_0(x) \int_{x_0}^x |A(\xi)|d\xi = M_0(x)q(x).$$

Thus $M_1(t) \leq q(x)M_0(x)$ for $x_0 \leq t \leq x$. Now suppose statement (2.33) is true for $n - 1 \geq 0$, thus by induction we have,

$$|\bar{y}_{n+1}(t) - \bar{y}_n(t)| \leq \int_{x_0}^x |A(\xi)|M_{n-1}(\xi)d\xi \leq \frac{M_0(x)}{(n-1)!} \int_{x_0}^x |A(\xi)|q^{n-1}(\xi)d\xi, \quad (2.34)$$

for $x_0 \leq t \leq x$. However, since $|A(\xi)| = q'(\xi)$ almost everywhere,

$$\int_{x_0}^x |A(\xi)|q^{n-1}(\xi)d\xi = \frac{1}{n} \int_{x_0}^x \frac{d}{d\xi}(q^n(\xi))d\xi = \frac{q^n(x)}{n}. \quad (2.35)$$

Statements (2.34) and (2.35) complete the induction, thus we have the following bound

$$\sum_{i=0}^{\infty} |\bar{y}_{i+1}(t) - \bar{y}_i(t)| \leq |\bar{y}_0| + M_0(x) \sum_{i=0}^{\infty} \frac{q^i}{i!} = |\bar{y}_0| + M_0(x)e^{q(x)}.$$

for $t \in [x_0, x]$. In the above equation the limit of partial sums converges uniformly on $[x_0, x]$. Hence the sequence $(\bar{y}_n)$ converges absolutely uniformly on every closed interval to a function $\bar{y}$ which is a continuous solution to (2.29). As $\bar{y}$ is continuous and a solution of (2.27) satisfying the initial conditions (2.28) it also follows that $\bar{y}$ is absolutely continuous.

To prove uniqueness, we assume there are two absolutely continuous solutions $\bar{y}(x)$ and $\bar{z}(x)$ different from each other at some point greater than x_0 , then

$$\begin{aligned} \bar{y}(x) &= \bar{y}_0 + \int_{x_0}^x [A(\xi)\bar{y}(\xi) + \bar{g}(\xi)]d\xi, \\ \bar{z}(x) &= \bar{y}_0 + \int_{x_0}^x [A(\xi)\bar{z}(\xi) + \bar{g}(\xi)]d\xi. \end{aligned}$$

Subtracting the above equations gives

$$\bar{y}(x) - \bar{z}(x) = \int_{x_0}^x A(\xi)[\bar{y}(\xi) - \bar{z}(\xi)]d\xi \quad (2.36)$$

Let α be the largest number such that $\bar{z}(x)$ and $\bar{y}(x)$ coincide on the interval (x_0, α) . Set $p(x) := \max_{\alpha \leq \xi \leq x} |\bar{y}(\xi) - \bar{z}(\xi)|$, then from (2.36) we have the following inequality,

$$p(x) \leq \max_{\alpha \leq \xi \leq x} |\bar{y}(\xi) - \bar{z}(\xi)| \int_{\alpha}^x |A(\xi)|d\xi = p(x) \int_{\alpha}^x |A(\xi)|d\xi. \quad (2.37)$$

Hence for all x sufficiently close to α such that

$$\int_{\alpha}^x |A(\xi)|d\xi < 1,$$

we have $p(x) = 0$. Contradicting the maximality of α . ■

The above result can be applied to an arbitrary linear n -th ordered ordinary differential equation with suitable coefficients. In particular, our non-homogeneous equation, $\lambda y - l(y) = g$, can be posed in matrix form (2.27) by setting

$$A(x) := \begin{pmatrix} 0 & 1/p \\ q - \lambda & 0 \end{pmatrix}, \quad \bar{g}(x) = \begin{pmatrix} 0 \\ g \end{pmatrix} \quad \text{and} \quad \bar{y}(x) = \begin{pmatrix} y \\ py' \end{pmatrix}.$$

Definition 2.12 Let $\mathcal{A}$ be the set of all functions, y , on $[a, b]$ such that y and py' are absolutely continuous and $l(y) \in \mathcal{L}^2([a, b])$.

The above theorem guarantees the existence of a unique solution in $\mathcal{A}$ to the non-homogeneous equation $\lambda y - l(y) = g$ for each pair of initial conditions $y(a) = \alpha_1$ and $py'(a) = \alpha_2$. A fundamental system is a maximal linearly independent set of solutions of the homogeneous equation. For the existence of such a fundamental system see Naimark⁴.

The next theorem allows us to obtain the dimension of the deficiency spaces.

Theorem 2.13 The set of all solutions in $\mathcal{A}$ of the homogeneous equation,

$$l(y) - \lambda y = 0, \quad (2.38)$$

forms a linear space of dimension 2, i.e. each fundamental system consist of 2 functions.

Proof: Each solution y of (2.38) is uniquely determined by the initial conditions $\alpha_1 = y(a)$ and $\alpha_2 = py'(a)$. Let z_1 and z_2 be the solutions of (2.38) obeying the initial conditions $z_1(a) = 1 = pz_2'(a)$ and $pz_1'(a) = 0 = z_2(a)$. Then by the linearity and uniqueness of the initial value problem, $y = \alpha_1 z_1 + \alpha_2 z_2$. Hence every solution of (2.38) is a linear combination of z_1 and z_2 . It remains only to show that z_1 and z_2 are linearly independent. If $0 = \alpha_1 z_1 + \alpha_2 z_2$ then by evaluation of this equality and p times its derivative at a we get $0 = \alpha_1$ and $0 = \alpha_2$. Hence z_1 and z_2 are linearly independent. ■

The Lagrange identity simplifies many of the subsequent calculations. It can easily be checked via integration by parts. In particular, the Lagrange identity states that

$$(l(y), z) - (y, l(z)) = [y, z]_a^b \quad \text{where} \quad [y, z]_a^b = p(x)[y(x)\bar{z}'(x) - \bar{y}'(x)z(x)] \Big|_a^b \quad (2.39)$$

⁴cf. Naimark [31, pp. 58]

The bilinear form defined in (2.19) has a concrete representation as $\langle f, g \rangle = [f, g]_a^b$. Here $[\cdot, \cdot]_a^b$ is called the Langrange form associated with l .

Lemma 2.14 *For arbitrary numbers $\alpha_0, \alpha_1, \beta_0$ and β_1 , there exists a function y in $\mathcal{A}$ such that*

$$p^k y^{(k)}(a) = \alpha_k \quad \text{and} \quad p^k y^{(k)}(b) = \beta_k \quad \text{for } k = 0, 1.$$

Proof: Let $g = \gamma_0 z_0 + \gamma_1 z_1$, where z_0 and z_1 are a fundamental system of the homogeneous equation such that $(z_i, z_j) = \delta_{ij}$ and

$$\gamma_k = pz'_k(b)\beta_0 + z_k(b)\beta_1. \quad (2.40)$$

Let v be the solution of $l(v) = g$ which satisfies the initial conditions $pv'(a) = v(a) = 0$. Such a v exists by the note following theorem 2.11.

Using the Lagrange identity we obtain

$$(g, z_k) = (l(v), z_k) = (v, l(z_k)) + [v, z_k]_a^b \quad \text{for } k = 0, 1. \quad (2.41)$$

Here $(v, l(z_k)) = 0$ and $[v, z_k]_a = 0$. Thus we have,

$$\gamma_0(z_0, z_k) + \gamma_1(z_1, z_k) = (g, z_k) = [v, z_k]_a^b = p(b)[z'_k(b)v(b) + z_k(b)v'(b)].$$

Since $(z_i, z_j) = \delta_{ij}$,

$$\gamma_0\delta_{0k} + \gamma_1\delta_{1k} = p(b)[z'_k(b)v(b) + z_k(b)v'(b)],$$

giving

$$\begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} = \begin{bmatrix} pz'_0(b) & z_0(b) \\ pz'_1(b) & z_1(b) \end{bmatrix} \begin{bmatrix} v(b) \\ pv'(b) \end{bmatrix}$$

Thus

$$\begin{bmatrix} pz'_0(b) & z_0(b) \\ pz'_1(b) & z_1(b) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} pz'_0(b) & z_0(b) \\ pz'_1(b) & z_1(b) \end{bmatrix} \begin{bmatrix} v(b) \\ pv'(b) \end{bmatrix}$$

Here $\begin{bmatrix} pz'_0(b) & z_0(b) \\ pz'_1(b) & z_1(b) \end{bmatrix}$ is invertible as its determinant is the Wronskian of z_1 and z_0 which are linearly independent. Hence $p^k v^{(k)}(b) = \beta_k$ and $v^{(k)}(a) = 0$ for $k = 0, 1$. In

a similar fashion, we can construct a w such that $w^{(k)}(b) = 0$ and $w^{(k)}(a) = \alpha_k$. It is clear that $y = v + w$ is the required function. ■

From the above lemma, we can make the following deduction.

Corollary 2.15 *Let $g \in \mathcal{L}^2([a, b])$. There is a solution to $l(y) = g$ with*

$$p^k y^{(k)}(a) = p^k y^{(k)}(b) = 0 \quad \text{for } k = 0, 1, \quad (2.42)$$

if and only if g is orthogonal to all solutions of $l(z) = 0$.

Proof: Let $g \in \mathcal{L}^2([a, b])$ and z_1, z_2 be a fundamental system of solutions to $l(z) = 0$, with $(z_i, z_j) = \delta_{ij}$. Let y be a solution of $l(y) = g$ which satisfies (2.42). Now from (2.41) we see that

$$(g, z_k) = (l(y), z_k) = (y, l(z_k)) + [y, z_k]_a^b = 0.$$

Since every solution of $l(y) = 0$ is a linear combination of z_1 and z_2 , it follows from the above equation that g is orthogonal to all solutions of $l(z) = 0$.

For the converse, suppose that $g \in \mathcal{L}^2([a, b])$ is orthogonal to all solutions of $l(z) = 0$. From the note prior to definition 2.12 it follows that $l(y) = g$ has a unique solution with $y(a) = 0 = py'(a)$. Now from (2.41) we see that

$$0 = (g, z_k) = (l(y), z_k) = (y, l(z_k)) + [y, z_k]_a^b = [y, z_k]_a^b = [y, z_k]^b,$$

where z_1, z_2 is a fundamental system of solutions to $l(z) = 0$. Thus

$$0 = \begin{bmatrix} z_1(b) & pz_1'(b) \\ z_2(b) & pz_2'(b) \end{bmatrix} \begin{bmatrix} \bar{y}(b) \\ p\bar{y}'(b) \end{bmatrix}.$$

Since z_1, z_2 is a fundamental system $\det \begin{bmatrix} z_1(b) & pz_1'(b) \\ z_2(b) & pz_2'(b) \end{bmatrix} \neq 0$ giving $\begin{bmatrix} y(b) \\ py'(b) \end{bmatrix} = 0$.
■

We define L_0 and L to be the minimal and maximal operators associated with l . The domains $\mathcal{D}(L_0)$ and $\mathcal{D}(L)$ respectively of these operators are given by

$$\begin{aligned}\mathcal{D}(L_0) &= \{y \in \mathcal{A} \mid y(a) = py'(a) = y(b) = py'(b) = 0\}, \\ \mathcal{D}(L) &= \mathcal{A}.\end{aligned}$$

We call L the maximal operator since its domain is the largest linear space of functions in $\mathcal{L}^2([a, b])$ on which the formal operator $l(y)$ can be meaningfully defined. The minimal operator L_0 has domain, $\mathcal{D}(L_0)$, so severely restricted that the domain of any symmetric operator formed from the formal operator has domain containing $\mathcal{D}(L_0)$. It will be shown that L_0 and L are adjoint to one another and are hence closed. In fact, $\mathcal{D}(L_0)$ is the closure of the space of functions in $\mathcal{A}$ with compact support contained in $[a, b]$. An important decomposition for the space $\mathcal{L}^2([a, b])$, which will be used to show the uniqueness of the adjoint, can now be formulated from Corollary 2.15 as

$$\mathcal{L}^2([a, b]) = \mathcal{R}(L_0) \oplus \mathcal{N}(L). \quad (2.43)$$

The adjoint of the operator L_0 is defined uniquely if the domain of L_0 is dense in $\mathcal{L}^2([a, b])$.

Lemma 2.16 *The domain of L_0 is dense in $\mathcal{H} = \mathcal{L}^2([a, b])$.*

Proof: We show that every element in $\mathcal{H}$ which is orthogonal $\mathcal{D}(L_0)$ is zero. Suppose $h \in \mathcal{H}$ is orthogonal to $\mathcal{D}(L_0)$ then we have

$$(h, y) = 0 \quad \text{for all } y \in \mathcal{D}(L_0). \quad (2.44)$$

By corollary 2.15 the equation $l(z) = h$ has a solution $z \in \mathcal{D}(L_0)$. It can be seen directly for the Lagrange identity and the definition of L_0 that $L_0 \subset L$ and

$$(z, L_0(y)) = (L_0(z), y) = (l(z), y) = (h, y) = 0 \quad \text{for all } y \in \mathcal{D}(L_0). \quad (2.45)$$

Thus $z \in \mathcal{R}(L_0)^\perp$ and from (2.43) we see that $z \in \mathcal{N}(L)$. Thus $h = L(z) = 0$. ■

We will now show that L and L_0 are adjoint to one another.

Theorem 2.17 *Let L and L_0 be as defined above, then $L^* = L_0$ and $L_0^* = L$.*

Proof: We will first prove $L_0^* = L$ by a double inclusion. Firstly we see that from symmetry and $L_0 \subset L$ that $(L_0x, y) = (x, Ly)$ for $x \in \mathcal{D}(L_0)$, $y \in \mathcal{D}(L)$. Now since the adjoint is the maximal extension we have $L \subset L_0^*$. We will be done if we can show $L \supset L_0^*$. Let $g \in \mathcal{D}(L_0^*)$ and set $L_0^*g = h$. From the existence theorem (2.11) there is a solution $f \in \mathcal{D}(L)$ of $Lf = h$. For each $k \in \mathcal{D}(L_0)$, using symmetry, we have

$$(k, h) = (k, Lf) = (L_0k, f) \quad \text{but also} \quad (k, h) = (k, L_0^*g) = (L_0k, g).$$

From subtraction we see that $(L_0k, f - g) = 0$. This shows that $f - g \in \mathcal{R}(L_0)^\perp$. Hence, by (2.43), we have $f - g \in \mathcal{N}(L)$. Since $f \in \mathcal{D}(L)$, therefore $g \in \mathcal{D}(L)$ and $Lg = Lf = h = L_0^*g$. Hence $L_0^* = L$.

We now show $L^* = L_0$. We already know that $L^* = L_0^{**}$. Since L_0^{**} is the closure of L_0 we have $L^* = L_0^{**} \supset L_0$. It suffices, therefore, to prove $L^* \subset L_0$. From symmetry we have $L = L_0^* \supset L^*$ and

$$(Lg, f) = (g, L^*f) = (g, Lf) \quad \text{for} \quad f \in \mathcal{D}(L^*), g \in \mathcal{D}(L). \quad (2.46)$$

Hence, by the Lagrange identity, $[g, f]_a^b = 0$, for all $g \in \mathcal{A}$. So, from Lemma 2.14, $pf'(a) = f(b) = pf'(b) = 0$, giving $f \in \mathcal{D}(L_0)$. Thus $\mathcal{D}(L^*) \subset \mathcal{D}(L_0)$, completing the proof. ■

We may now employ the theory of deficiency indices. Since $L_0^* = L$, which was shown in the previous theorem, we have that $\mathcal{N}(L_0^* - \lambda I) = \mathcal{N}(L - \lambda I)$. Thus theorem 2.13, setting $\lambda = i$, shows that the dimension of N_+ is 2. Similarly, the dimension of N_- is also 2. Hence the deficiency indices are $(2, 2)$, and we may apply theorem 2.10.

From Corollary 2.15 there exist functions f_1, f_2 satisfying the following conditions,

$$\begin{aligned} f_1(a) &= -\sin \alpha, & pf_1'(a) &= -\cos \alpha, & f_2(a) &= 0, & pf_2'(a) &= 0, \\ f_1(b) &= 0, & pf_1'(b) &= 0, & f_2(b) &= \sin \beta, & pf_2'(b) &= \cos \beta. \end{aligned}$$

It is clear that f_1, f_2 defined above, satisfies condition (2.24), that is, $\langle f_i, f_j \rangle = 0$ for $i, j = 1, 2$. Suppose $\gamma f_1(x) + \delta f_2(x) \in \mathcal{D}(L_0)$, then from imposing the boundary conditions we have the system of equations $\gamma \sin \alpha = \delta \sin \beta = \gamma \cos \alpha = \delta \cos \beta = 0$. Solving the system shows that $\gamma = \delta = 0$. Thus f_1, f_2 are linearly independent relative to $\mathcal{D}(L_0)$, and equality (2.25) of Theorem 2.10 gives that the restriction of L to the domain

$$\mathcal{D}(M) = \{f \in \mathcal{A} : f \text{ satisfies boundary conditions (1.2) and (1.3)}\},$$

is a self-adjoint extension of L_0 . All self-adjoint extension of L_0 may be obtained in this way, see [31].

Chapter 3

Eigenvalue asymptotics

3.1 Introduction

In this section we apply the functional analysis developed in chapter 2 to Sturm-Liouville problems. We investigate a slightly different equation to (1.1), nevertheless, we show through the Liouville transformation that (1.1) can be transformed to the form under consideration. A concrete representation of the resolvent operator is given in terms of the Green's kernel. Following this, the spectral theorem is used to obtain the completeness of the eigenfunction and hence eigenfunction expansions. Asymptotic estimates of the eigenvalues are established through an approach studied by Hochstadt [16], with these approximants, it is shown that the above eigenfunction expansions converge in a manner similar to that of Fourier series. The method of proof presented was developed by Titchmarsh [41, pp. 12].

3.2 The Sturm-Liouville Expansion

Toward our goal of studying the operator (1.4) related to (1.1), we will study the Sturm-Liouville problem

$$\frac{d^2 y}{dx^2} + (\lambda - q(x))y = 0, \quad x \in [a, b], \quad (3.1)$$

with boundary conditions

$$y(a) \cos \alpha - y'(a) \sin \alpha = 0, \quad (3.2)$$

$$y(b) \cos \beta - y'(b) \sin \beta = 0, \quad (3.3)$$

and the associated differential operator

$$L = -\frac{d^2}{dx^2} + q(x), \quad (3.4)$$

with domain

$$\mathcal{D}(L) = \{y \in \mathcal{L}^2(a, b) \mid y, y' \in AC, l(y) \in \mathcal{L}^2(a, b), y \text{ obeys (3.2), (3.3)}\}. \quad (3.5)$$

Here $q \in \mathcal{L}^1(a, b)$ and $-\infty < a < b < \infty$. We will show later that under an appropriate transformation (and smoothness conditions), the original problem, (1.1), can be converted to (3.1) and that it is thus sufficient to study the operator (3.4).

In particular we study the non-homogeneous operator equation,

$$(\lambda I - L)f = g. \quad (3.6)$$

What will follow can be seen as a particular application of the Hilbert-Schmidt theorem. The advantage of studying the problem (3.1) in this functional analytic way is that it gives concrete representations of the eigenfunction expansion, the resolvent and the Green's function. In addition it provides a theoretical foundation from which the eigenvalue asymptotics can be studied.

Let the wronskian of the solutions ϕ and χ of (3.1) be denoted by $\omega(\lambda)$ and defined by,

$$\omega(\lambda) = \phi(x)\chi'(x) - \phi'(x)\chi(x). \quad (3.7)$$

By differentiating $\omega(\lambda)$ with respect to x and applying equation (3.1), we see that $\omega'(\lambda) = 0$, and thus $\omega(\lambda)$ is independent of x .

Next we construct the Green's function for (3.1)-(3.3). To this end solutions obeying the boundary conditions at each endpoint are constructed.

Lemma 3.1 *For $\lambda \notin \sigma(L)^1$ there exist non-trivial solutions $\phi(x, \lambda)$ and $\chi(x, \lambda)$ to (3.1) satisfying the boundary conditions (3.2), (3.3), respectively.*

Proof: We have shown the existence² of solutions of (3.6) satisfying arbitrary initial conditions. Now let $\phi_0(x, \lambda)$, $\chi_0(x, \lambda)$ be a pair of such functions forming a fundamental system of the homogenous equation (3.1). We know that the solution space of the homogenous equation is two dimensional³, and may thus write each solution to the homogeneous equation in the form $\phi(x, \lambda) = c\phi_0(x, \lambda) + d\chi_0(x, \lambda)$. These initial conditions may be set out as,

$$\begin{pmatrix} \phi(a, \lambda) \\ \phi'(a, \lambda) \end{pmatrix} = \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} = \begin{pmatrix} \phi_0(a, \lambda) & \chi_0(a, \lambda) \\ \phi'_0(a, \lambda) & \chi'_0(a, \lambda) \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix}.$$

Let the 2×2 matrix in the above be denoted by $A(a, \lambda)$, then

$$A^{-1}(a, \lambda) = \frac{1}{\omega(\lambda)} \begin{pmatrix} \chi'_0(a, \lambda) & -\chi_0(a, \lambda) \\ -\phi'_0(a, \lambda) & \phi_0(a, \lambda) \end{pmatrix}.$$

Solving for c and d , we thus have,

$$\begin{pmatrix} c \\ d \end{pmatrix} = \frac{1}{\omega(\lambda)} \begin{pmatrix} \chi'_0(a, \lambda) \sin \alpha - \chi_0(a, \lambda) \cos \alpha \\ -(\phi'_0(a, \lambda) \sin \alpha + \phi_0(a, \lambda) \cos \alpha) \end{pmatrix}. \quad (3.8)$$

Substituting c and d back into $\phi(a, \lambda)$ we obtain

$$\phi(x, \lambda) = \frac{(\chi'(a, \lambda) \sin \alpha - \chi_0(a, \lambda) \cos \alpha)\phi_0(x, \lambda) - (\phi'_0(a, \lambda) \sin \alpha + \phi_0(a, \lambda) \cos \alpha)\chi_0(x, \lambda)}{\omega(\lambda)} \quad (3.9)$$

¹This is to ensure $\omega(\lambda) \neq 0$.

²cf. Theorem 2.11

³cf. Theorem 2.13

In a similar fashion, replacing the a, α by b, β respectively, we can obtain the expression for $\chi(x, \lambda)$ in terms of ϕ_0 and χ_0 . By construction $\phi(x, \lambda)$ and $\chi(x, \lambda)$ satisfy the boundary conditions at a and b respectively. ■

Following the approach of Coddington and Levinson [3, page 192], we define the Green's function of L for $\lambda \notin \sigma(L)$ as the function $k(x, y, \lambda)$ such that:

- (i) the function is continuous in x and y ,
- (ii) for $x \neq y$, its derivative with respect to x exists and is continuous,
- (iii) for $a < y < b$, as a function of x it satisfies the boundary conditions (3.2) and (3.3),
- (iv) $\frac{dk}{dx}(y+0, y, \lambda) - \frac{dk}{dx}(y-0, y, \lambda) = 1$ for $a < y < b$,
- (v) as a function of x , it satisfies the equation $lk - \lambda k = 0$ at $x \neq y$.

We now give a concrete representation of the resolvent operator of L in terms of its Green's function.

Theorem 3.2 (The Resolvent)

Let L be the self-adjoint extension (3.4) with domain (3.5) of l . Suppose $\text{Im}\lambda = 0$, $\lambda \notin \sigma(L)$, and $a < c < b$. Let $\phi \in \mathcal{L}_2(a, c)$ and $\chi \in \mathcal{L}_2(c, b)$ be non-zero solutions of $ly = \lambda y$ satisfying the boundary conditions (3.2) at a , and (3.3) at b respectively. The resolvent operator $R(\lambda : L) : \mathcal{L}_2(a, b) \rightarrow \mathcal{L}_2(a, b)$ is compact and given by

$$R(\lambda : L)g = \int_a^b k(x, y, \lambda)g(y)dy \quad (g \in \mathcal{L}_2(a, b)), \quad (3.10)$$

where $k(x, y, \lambda)$ is the Green's function

$$k(x, y, \lambda) = \begin{cases} \gamma\phi(x, \lambda)\chi(y, \lambda) & \text{for } a \leq x < y \\ \gamma\phi(y, \lambda)\chi(x, \lambda) & \text{for } y < x \leq b \end{cases} \quad \text{and} \quad \gamma = \frac{1}{\omega(\lambda)}. \quad (3.11)$$

Proof: The existence of the functions $\phi(x, \lambda)$ and $\chi(x, \lambda)$ is guaranteed by lemma 3.1. From the definition of the Green's function $k(x, y, \lambda)$ we may assume it to be of

the form,

$$k(x, y, \lambda) = \begin{cases} A(y)\phi(x, \lambda) & \text{for } a \leq x < y, \\ B(y)\chi(x, \lambda) & \text{for } y < x \leq b. \end{cases}$$

We require that $k(x, y, \lambda)$ satisfies the conditions (i) and (iv) and we may write these conditions via the matrix equation

$$\begin{pmatrix} -\phi(y, \lambda) & \chi(y, \lambda) \\ -\phi'(y, \lambda) & \chi'(y, \lambda) \end{pmatrix} \begin{pmatrix} A(y) \\ B(y) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (3.12)$$

Solving this system we find $A(y)$ and $B(y)$ to be

$$A(y) = \chi(y, \lambda)/\omega(\lambda) \quad \text{and} \quad B(y) = \phi(y, \lambda)/\omega(\lambda), \quad (3.13)$$

thus giving equation (3.11).

We need to verify equation (3.10). Let $K : \mathcal{L}_2(a, b) \rightarrow \mathcal{L}_2(a, b)$ be given by

$$Kg = \int_a^b k(x, y, \lambda)g(y) dy, \quad g \in \mathcal{L}_2(a, b). \quad (3.14)$$

It remains to show that $K = R(\lambda : L)$. To this end, we begin by showing that K is bounded on $\mathcal{L}_2(a, b)$. By hölders inequality we have

$$\|Kf\| \leq \left[\int_a^b \int_a^b |k(x, y, \lambda)|^2 dy \int_a^b |f(z)|^2 dz dx \right]^{1/2}. \quad (3.15)$$

Let $\|\cdot\|_2$ be defined as

$$\|k\|_2 = \sqrt{\int_a^b \int_a^b |k(x, y, \lambda)|^2 dy dx}. \quad (3.16)$$

From equation (3.15) and the continuity of $k(x, y, \lambda)$, we have

$$\|K\| \leq \|k\|_2 < \infty, \quad (3.17)$$

and thus K is bounded.

Writing

$$Kg(x) = \gamma\chi(x) \int_a^x \phi g dy + \gamma\phi(x) \int_x^b \chi g dy,$$

for $g \in \mathcal{L}_2(a, b)$, it follows that

$$(Kg)'' - qKg + \lambda Kg = g.$$

Furthermore a direct calculation shows that $Kg(x)$ satisfies the boundary conditions (3.2) and (3.3). Thus $\mathcal{R}(K) \subset \mathcal{D}(L - \lambda I)$ and

$$K = R(L : \lambda). \quad (3.18)$$

We now show that $R(L : \lambda)$ is compact. The Green's kernel $k(x, y, \lambda)$ is continuous, thus by Weierstrass' approximation theorem, we may approximate $k(x, y, \lambda)$ by a sequence of degenerate polynomial kernels $k_n(x, y, \lambda)$ that converge to $k(x, y, \lambda)$ in the supremum norm. Define the operator $K_n : \mathcal{L}^2(a, b) \rightarrow \mathcal{L}^2(a, b)$ by

$$K_n f = \int_a^b k_n(x, y, \lambda) f(y) dy. \quad (3.19)$$

Since $k_n(x, y, \lambda)$ is a polynomial of finite order n , the range of K_n is finite dimensional, hence K_n is compact. We now have

$$\|K - K_n\| \leq \|k - k_n\|_2 \leq \sup_{x \in [a, b]} \|k - k_n\| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (3.20)$$

thus K is the limit of a sequence of compact operators K_n showing that K is compact, and hence so is $R(L : \lambda)$. ■

Towards our goal of obtaining the Sturm-Liouville expansion and its completeness theorem, we apply the spectral theorem and integrate around the poles of the resolvent. Here the terms of the expansion are the residues of these poles. We show that these poles are simple. From equation (3.10) the poles of the resolvent are the zero's of $\omega(\lambda)$. These are on the real axis since L is self-adjoint.

Theorem 3.3 *The poles of $\omega(\lambda)$ are all simple.*

Proof: Let $\phi(x, \lambda)$ and $\chi(x, \lambda)$ be as in (3.1). Suppose that λ_0 is an eigenvalue of the operator (3.4) and also a double zero of $\omega(\lambda)$. Let $\lambda = \lambda_0 + iv$, where $v \in \mathbb{R}$.

We see, from the Lagrange identity, that

$$\int_a^b \bar{\phi} L\phi - \phi \overline{L\phi} dx = [\phi, \phi]_a^b. \quad (3.21)$$

Hence we have,

$$(\bar{\lambda} - \lambda) \int_a^b \phi(x, \lambda) \phi(x, \bar{\lambda}) dx = -\phi(b, \lambda) \overline{\phi'(b, \lambda)} + \phi'(b, \lambda) \overline{\phi(b, \lambda)}.$$

Without loss of generality we may assume $\chi(b, \lambda) = \sin \beta$ and $\chi'(b, \lambda) = \cos \beta$. Now $\omega(\lambda) = \phi(b, \lambda) \cos \beta - \phi'(b, \lambda) \sin \beta$, using $\omega(\lambda)$ multiplying the above equation by $\sin \beta$ we get

$$(\bar{\lambda} - \lambda) \sin \beta \int_a^b \phi(x, \lambda) \overline{\phi(x, \lambda)} dx = \overline{\omega(\lambda)} \phi(b, \lambda) - \omega(\lambda) \overline{\phi(b, \lambda)}. \quad (3.22)$$

Considering that λ_0 is a double zero of $\omega(\lambda)$, using taylor expansion,

$$\omega(\lambda_0 + iv) = \omega(\lambda_0) + \omega'(\lambda_0)iv + \frac{\omega''(\lambda_0)(iv)^2}{2} + o(v^2) = \frac{\omega''(\lambda_0)(iv)^2}{2} + o(v^2).$$

As λ_0 is real the left hand side of (3.22) is equal to,

$$-2iv \sin \beta \int_a^b |\phi(x, \lambda)|^2 dx.$$

Hence from (3.22) we have

$$2iv \sin \beta = \overline{\omega(\lambda)} \frac{\phi(b, \lambda)}{\int_a^b |\phi(x, \lambda)|^2 dx} - \omega(\lambda) \frac{\overline{\phi(b, \lambda)}}{\int_a^b |\phi(x, \lambda)|^2 dx}.$$

Hence

$$\frac{4}{iv} \sin \beta = (\overline{\omega''(\lambda_0)} + o(1)) \frac{\phi(b, \lambda)}{\int_a^b |\phi(x, \lambda)|^2 dx} - (\omega''(\lambda_0) + o(1)) \frac{\overline{\phi(b, \lambda)}}{\int_a^b |\phi(x, \lambda)|^2 dx}.$$

Here for each $\epsilon > 0$ if $v \in [-\epsilon, \epsilon] := I$ then $\int_a^b |\phi(x, \lambda)|^2 dx$ is positive and continuous and is thus bounded below by some $\delta > 0$ on I . Also $\phi(b, \lambda)$ is continuous in λ and thus for some $\kappa > 0$ we have $|\phi(b, \lambda)| < \kappa$ for $v \in I$. Hence

$$\left| \frac{4}{iv} \sin \beta \right| \leq 2(|\omega''(\lambda_0)| + o(1)) \frac{\kappa}{\delta} = O(1)$$

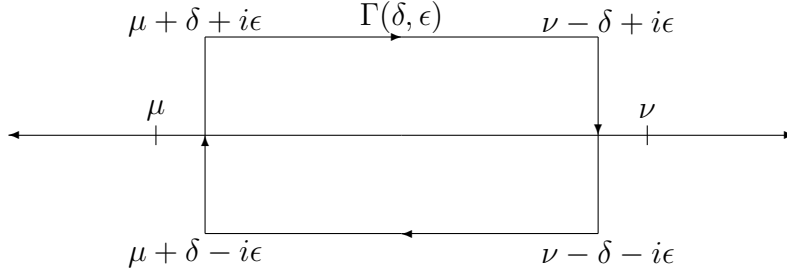
for $v \in I$. This is not possible as the left hand side tends to infinity as $v \rightarrow 0$.

The case of $\sin \beta = 0$ can be treated in a similar manner. ■

Label the eigenvalues in increasing order as $\lambda_1 < \lambda_2 < \dots < \lambda_n$. We may apply the spectral theorem and its related equations (2.14) and (2.15). For $\mu, \nu \in \mathbb{R}$, not eigenvalues, such that $\mu < \nu$, we may change the integration limits of (2.15) to obtain

$$(P(\mu, \nu)f, g) = \int_a^b \int_a^b \lim_{\delta, \epsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_{\mu+\delta}^{\nu-\delta} [k(x, y, \xi)]_{\xi=\lambda-i\epsilon}^{\xi=\lambda+i\epsilon} f(y) d\lambda \overline{g(x)} dy dx. \quad (3.23)$$

Let the path Γ be the clockwise oriented curve as defined in the diagram below.



As μ and ν are not poles of $k(x, y, \lambda)$ it follows that

$$(P(\mu, \nu)f, g) = \int_a^b \int_a^b \lim_{\delta, \epsilon \rightarrow 0^+} \frac{1}{2\pi i} \oint_{\Gamma(\delta, \epsilon)} k(x, y, \lambda) d\lambda f(y) \overline{g(x)} dy dx. \quad (3.24)$$

At an eigenvalue $\mu < \lambda_n < \nu$, the Wronskian is zero, thus functions $\phi(x, \lambda_n)$ and $\chi(x, \lambda_n)$ are linearly dependent non-trivial solutions, hence

$$\chi(x, \lambda_n) = k_n \phi(x, \lambda_n), \quad (3.25)$$

for some $k_n \in \mathbb{C} \setminus \{0\}$. Using residue calculus and l'Hôpital's rule we obtain

$$\begin{aligned} \lim_{\delta, \epsilon \rightarrow 0^+} \frac{1}{2\pi i} \oint_{\Gamma} k(x, y, \lambda) d\lambda &= \sum_{\mu < \lambda_n < \nu} k_n \phi(x, \lambda_n) \phi(y, \lambda_n) \lim_{z \rightarrow \lambda_n} \frac{(z - \lambda_n)}{\omega(z)} \\ &= \sum_{\mu < \lambda_n < \nu} \frac{k_n}{\omega'(\lambda_n)} \phi(x, \lambda_n) \phi(y, \lambda_n). \end{aligned}$$

It remains only to determine k_n . From (3.25)

$$k_n = \lim_{\lambda \rightarrow \lambda_n} \frac{\chi(x, \lambda)}{\phi(x, \lambda)}.$$

Now from (3.9), we have,

$$k_n = \frac{(\chi'_0(b, \lambda_n) \sin \beta - \chi_0(b, \lambda_n) \cos \beta) \phi_0(x, \lambda_n) - \chi_0(x, \lambda_n) (\phi'_0(b, \lambda_n) \sin \beta - \phi_0(b, \lambda_n) \cos \beta)}{(\chi'_0(a, \lambda_n) \sin \alpha - \chi_0(a, \lambda_n) \cos \alpha) \phi_0(x, \lambda_n) - \chi_0(x, \lambda_n) (\phi'_0(a, \lambda_n) \sin \alpha - \phi_0(a, \lambda_n) \cos \alpha)}$$

Since $\phi_0(x, \lambda)$ and $\chi_0(x, \lambda)$ are linearly independent, equating their coefficients in the above equation and solving the resulting two equations for k_n yields,

$$k_n = \frac{\chi'_0(b, \lambda_n) \sin \beta - \chi_0(b, \lambda_n) \cos \beta}{\chi'_0(a, \lambda_n) \sin \alpha - \chi_0(a, \lambda_n) \cos \alpha} = \frac{\phi'_0(b, \lambda_n) \sin \beta - \phi_0(b, \lambda_n) \cos \beta}{\phi'_0(a, \lambda_n) \sin \alpha - \phi_0(a, \lambda_n) \cos \alpha}. \quad (3.26)$$

By defining $\psi_n(x)$ and c_n as

$$\psi_n(x) := \left(\frac{k_n}{\omega'(\lambda_n)} \right)^{\frac{1}{2}} \phi(x, \lambda_n) \text{ and } c_n := \int_a^b f(y) \psi_n(y) dy,$$

we may write equation (3.24) as

$$\begin{aligned} (P(\mu, \nu) f, g) &= \sum_{\mu < \lambda_n < \nu} \int_a^b \int_a^b f(y) \frac{k_n}{\omega'(\lambda_n)} \phi(x, \lambda_n) \phi(y, \lambda_n) \overline{g(x)} dy dx \\ &= \left(\sum_{\mu < \lambda_n < \nu} c_n \psi_n(x), g \right). \end{aligned}$$

Here g is an arbitrary element of $\mathcal{L}^2(a, b)$ giving

$$P(\mu, \nu) f = \sum_{\mu < \lambda_n < \nu} c_n \psi_n(x).$$

Now by (2.14), we obtain

$$f(x) = \sum_{n=1}^{\infty} c_n \psi_n(x). \quad (3.27)$$

We see that $\{\psi_n(x)\}$ forms an orthonormal eigenfunction basis in $\mathcal{L}_2(a, b)$ which is complete as expected by the Hilbert-Schmidt theorem. The canonical form of the operator L is

$$Lf(x) = \sum_{n=1}^{\infty} \lambda_n (f, \psi_n) \psi_n(x) \quad \text{for } f \in \mathcal{D}(L). \quad (3.28)$$

Thus

$$f(x) = \sum_{n=0}^{\infty} \frac{(g, \psi_n) \psi_n(x)}{\lambda - \lambda_n}.$$

3.3 Liouville Transformation

It was stated previously that it was sufficient to consider the differential equation (3.1) in place of (1.1). In this section we review the Liouville transformation which allows for this correspondence. The only additional condition requires $p(x) > 0$ to have continuous derivatives of the first two orders. Two possible forms of this transformations will be examined where the one comes from Titchmarsh [41] and the other from Ried [34].

Theorem 3.4 *Consider the general second order differential equation,*

$$a(X)\frac{d^2Y}{dX^2} + b(X)\frac{dY}{dX} + (\lambda - c(X))Y = 0 \quad \text{for } X \in [a, b], \quad (3.29)$$

with $a(X) \in \mathcal{C}^1[a, b]$, $b(X) \in \mathcal{C}[a, b]$ and $a(X) > 0$, can be reduced to,

$$\frac{d^2y}{dx^2} + \left(\lambda - \frac{1}{4}\beta^2(x) - \frac{1}{2}\beta'(x) - \gamma(x)\right)y = 0 \quad (3.30)$$

which is in the form (3.1). Here

$$\begin{aligned} \beta(x) &:= (b(X) - \frac{1}{2}a'(X))/\sqrt{a(X)} \text{ and } \gamma(x) := c(X), \\ x(X) &:= \int_0^X \frac{dX}{\sqrt{a(X)}} dX, \\ Y &:= \exp\left[-\frac{1}{2}\int_0^x \beta(s)ds\right]y. \end{aligned}$$

Proof: Applying the chain rule with respect to the transformation gives,

$$\frac{d^2Y}{dX^2} = \frac{d}{dX^2} \left(\frac{dY}{dx} \frac{dx}{dX} \right) = \frac{dY}{dX dx} \frac{dx}{dX} + \frac{dY}{dx} \frac{d^2x}{dX^2}.$$

Using the above and the transformation equation (3.29) becomes

$$\frac{d^2Y}{dx^2} + \frac{dY}{dx} \beta(x) + (\lambda + \gamma(x))Y = 0,$$

from which (3.30) follows. ■

Theorem 3.5 Consider the general second order differential equation,

$$\frac{d}{dX}\left(r(X)\frac{dY}{dX}\right) + (\lambda k(X) - c(X))Y = 0 \quad \text{for } X \in [a, b], \quad (3.31)$$

with $r(X) > 0$, $k(x) > 0$, $r(X) \in \mathcal{C}^1([a, b])$ and $r(X)k(X) \in \mathcal{C}^2([a, b])$, then the above equation can be reduced to

$$\frac{d^2 y}{dx^2} + (\mu - P(x))y = 0, \quad (3.32)$$

which is in the form (3.1). Here $P(x) = \left(\frac{d^2 f}{dx^2}\right)/f - K^2 g$, $\mu = K^2 \lambda$,

$f(x) = (k(X)r(X))^{\frac{1}{4}}$, $g(x) = p(X)/k(X)$ and $K = \frac{1}{\pi} \int_a^b \left(\frac{k(s)}{r(s)}\right)^{\frac{1}{2}} ds$.

Proof: Define the following transformations

$$x(X) := \int_0^X \frac{1}{K} \left(\frac{k(s)}{r(s)}\right)^{\frac{1}{2}} ds \quad \text{and} \quad y(x) := (k(X)r(X))^{\frac{1}{4}} Y(X).$$

Now $y(x(X)) = f(x(X))Y(X)$, and by the chain rule,

$$\frac{d}{dX}(y(x(X))) = \frac{dy}{dx} \frac{dx}{dX} = \frac{df}{dx} \frac{dx}{dX} Y + \frac{dY}{dX} f.$$

Substituting the transformations into the above equation and simplifying gives,

$$\frac{dy}{dx} = \frac{df}{dx} Y + K f \frac{dY}{dX} \left(\frac{r}{k}\right)^{\frac{1}{2}}$$

Applying the chain rule again to the above gives,

$$\frac{d^2 y}{dx^2} \frac{dx}{dX} = \frac{d^2 f}{dx^2} \frac{dx}{dX} Y + \frac{df}{dx} \frac{dY}{dX} + K \left[\frac{d^2 Y}{dX^2} f \left(\frac{r}{k}\right)^{\frac{1}{2}} + \frac{dY}{dX} \frac{df}{dx} \frac{dx}{dX} \left(\frac{r}{k}\right)^{\frac{1}{2}} + f \frac{dY}{dX} \frac{d}{dX} \left(\frac{r}{k}\right)^{\frac{1}{2}} \right]$$

Simplifying the 2nd, 4th and 5th terms and substituting in again the transformations, we obtain,

$$\frac{d^2 y}{dx^2} = \frac{d^2 f}{dx^2} Y + K^2 f \frac{d^2 Y}{dX^2} \left(\frac{r}{k}\right) + \left(\frac{r}{k}\right)^{\frac{1}{2}} \left[2K \frac{df}{dx} + K^2 f \frac{d}{dX} \left(\frac{r}{k}\right)^{\frac{1}{2}} \right]$$

Simplifying further,

$$\frac{d^2 y}{dx^2} = \frac{d^2 f}{dx^2} Y + K^2 \frac{r^{\frac{1}{4}}}{k^{\frac{3}{4}}} \left[\frac{d^2 Y}{dX^2} + \frac{dY}{dX} r(X)' \right]$$

Substituting the second term on the right into (3.31), we obtain,

$$\frac{d^2}{dx^2} - \frac{d^2 f}{dx^2} Y + K^2 \frac{r^{\frac{1}{4}}}{k^{\frac{3}{4}}} \left[\lambda k(X) - p(X) \right] Y = 0.$$

A reordering of the above, reduces the equation to (3.32). ■

3.4 Eigenvalue Asymptotics

We now discuss asymptotic approximations of the eigenvalues. There are two constructive approaches to obtain these eigenvalue asymptotics, those being the approach of Titchmarsh, [41], and that of Hochstadt, [16]. Both of these approaches rely on the Gronwall's lemma which is outlined in the appendix. The asymptotics obtained here will be used in a later section to study properties of the Sturm-Liouville expansion. We now consider the approach of Hochstadt to obtain eigenvalue asymptotics to the equation (1.1). A variation on Hochstadt's proof will be given only for the case (3.1), since (1.1) can be transformed into (3.1) by the Liouville transformation, described in section 3.3.

Theorem 3.6 *Let y_1, y_2 be the solutions of the differential equation (3.1), with $q \in \mathcal{L}([0, 1])$, obeying the initial conditions*

$$y_1(0) = 1, \quad y_1'(0) = 0,$$

$$y_2(0) = 0, \quad y_2'(0) = 1.$$

Let $\eta = |\operatorname{Im} \sqrt{\lambda}|$, then for large $|\lambda|$,

$$y_1 = \cos \sqrt{\lambda}x + O\left(\frac{e^{\eta x}}{\sqrt{\lambda}}\right) \tag{3.33}$$

$$y_1' = -\sqrt{\lambda} \sin \lambda x + O(e^{\eta x}) \tag{3.34}$$

$$y_2 = \frac{\sin \sqrt{\lambda}x}{\sqrt{\lambda}} + O\left(\frac{e^{\eta x}}{\lambda}\right) \tag{3.35}$$

$$y_2' = \cos \sqrt{\lambda}x + O\left(\frac{e^{\eta x}}{\sqrt{\lambda}}\right) \tag{3.36}$$

Proof: Define the dependent variables w_1 and w_2 by

$$w_1 = y, \quad w_2 = \frac{y'}{\sqrt{\lambda}}. \tag{3.37}$$

Denote the corresponding vector by

$$W(x) = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{\lambda}} \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} = B \begin{pmatrix} y \\ y' \end{pmatrix}. \tag{3.38}$$

Setting

$$C = \begin{pmatrix} 0 & 1 \\ q - \lambda & 0 \end{pmatrix},$$

the differential equation (3.1) is equivalent to the first order system

$$\frac{d}{dx} \begin{pmatrix} y \\ y' \end{pmatrix} = C \begin{pmatrix} y \\ y' \end{pmatrix}. \quad (3.39)$$

Combining (3.38) and (3.39) gives,

$$\frac{dW}{dx} = BCB^{-1}W = \begin{pmatrix} 0 & \sqrt{\lambda} \\ \frac{q}{\sqrt{\lambda}} - \sqrt{\lambda} & 0 \end{pmatrix} W, \quad (3.40)$$

which can reformulated as,

$$\frac{dW}{dx} = \sqrt{\lambda}AW + \frac{q}{\sqrt{\lambda}}MW \quad (3.41)$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (3.42)$$

Multiplying the above equation by the factor $\exp(-\sqrt{\lambda}Ax)$ and integrating gives the Volterra integral equation

$$W(x) = [\exp(\sqrt{\lambda}Ax)]W(0) + \frac{1}{\sqrt{\lambda}} \int_0^x q[\exp(\sqrt{\lambda}A)(x-t)]MW(t)dt. \quad (3.43)$$

From the series expansion of the exponential function we obtain

$$\exp \sqrt{\lambda}Ax = (\cos \sqrt{\lambda}x)I + (\sin \sqrt{\lambda}x)A$$

Let $W(x) = e^{\eta x}V(x)$. In terms of V the above integral equation becomes

$$V(x) = [\exp(\sqrt{\lambda}A - \eta)x]W(0) + \frac{1}{\sqrt{\lambda}} \int_0^x q[\exp(\sqrt{\lambda}A - \eta)(x-t)]MV(t)dt. \quad (3.44)$$

Since $\pm i$ are the eigenvalues of A , $A = D \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} D^{-1} = DHD^{-1}$, where D is the matrix whose columns are eigenvectors of A . Hence

$$e^{A\sqrt{\lambda}x} = De^{H\sqrt{\lambda}x}D^{-1} = D((\cos \sqrt{\lambda}x)I + (\sin \sqrt{\lambda}x)H)D^{-1}. \quad (3.45)$$

Using (3.45), we see that there exist a constant C_1 independent of x and λ , such that,

$$\|\exp(\sqrt{\lambda}A - \eta)x\| \leq C_1. \quad (3.46)$$

Combining (3.44) and (3.46) we obtain,

$$\|V\| \leq C_1\|W(0)\| + \frac{C_2}{\sqrt{\lambda}} \int_0^x |q|\|V\|dt. \quad (3.47)$$

Applying Gronwall's lemma, (A.1), to the above inequality gives

$$\|V\| \leq C_1\|W(0)\| \exp\left(\frac{C_2}{\sqrt{\lambda}} \int_0^x |q|dt\right). \quad (3.48)$$

This shows for sufficiently large $|\sqrt{\lambda}|$ that $V(x) = O(\|W(0)\|)$, hence $W(x) = O(\|W(0)\|e^{\eta x})$. Substituting this estimate into (3.43) gives,

$$W(x) = \exp(\sqrt{\lambda}Ax)W(0) + O\left(\|W(0)\|\frac{e^{\eta x}}{\sqrt{\lambda}}\right). \quad (3.49)$$

The solution y_1 corresponds to the boundary condition $W(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Substituting $W(0)$ into (3.49) the asymptotics (3.33) and (3.34) are obtained. Similarly y_2 corresponds to the boundary condition $W(0) = \begin{pmatrix} 0 \\ 1/\sqrt{\lambda} \end{pmatrix}$ which when substituted into (3.49) yields (3.35) and (3.36). ■

Theorem 3.7 *For sufficiently large λ , the estimates (3.33)-(3.36) obtained in theorem 3.6 can be replaced by*

$$y_1 = \cos(\sqrt{\lambda}x) + \frac{\sin(\sqrt{\lambda}x)}{2\sqrt{\lambda}} \int_0^x q dt + o\left(\frac{1}{\sqrt{\lambda}}\right), \quad (3.50)$$

$$y_1' = -\sqrt{\lambda} \sin(\sqrt{\lambda}x) + \frac{\cos(\sqrt{\lambda}x)}{2} \int_0^x q dt + o(1), \quad (3.51)$$

$$y_2 = \frac{\sin(\sqrt{\lambda}x)}{\sqrt{\lambda}} - \frac{\cos(\sqrt{\lambda}x)}{2\lambda} \int_0^x q dt + o\left(\frac{1}{\lambda}\right), \quad (3.52)$$

$$y_2' = \cos(\sqrt{\lambda}x) + \frac{\sin(\sqrt{\lambda}x)}{2\sqrt{\lambda}} \int_0^x q dt + o\left(\frac{1}{\sqrt{\lambda}}\right). \quad (3.53)$$

Proof: Substituting (3.49) into (3.43) we obtain,

$$W(x) = e^{\sqrt{\lambda}Ax}W(0) + \frac{1}{\sqrt{\lambda}} \int_0^x q(t)e^{\sqrt{\lambda}A(x-t)}Me^{\sqrt{\lambda}At}W(0) dt + O\left(\frac{\|W(0)\|}{\lambda}\right), \quad (3.54)$$

Considering the solution y_2 with corresponding boundary conditions gives $W(0) = \begin{pmatrix} 0 \\ 1/\sqrt{\lambda} \end{pmatrix}$. Through direct calculation of the above we have

$$W(x) = \begin{pmatrix} \sin(\sqrt{\lambda}x)/\sqrt{\lambda} \\ \cos(\sqrt{\lambda}x)/\sqrt{\lambda} \end{pmatrix} + \frac{1}{\lambda} \int_0^x q(t) \begin{pmatrix} \sin(\sqrt{\lambda}(x-t)) \sin(\sqrt{\lambda}t) \\ \cos(\sqrt{\lambda}(x-t)) \sin(\sqrt{\lambda}t) \end{pmatrix} dt + O\left(\frac{1}{\lambda^{3/2}}\right). \quad (3.55)$$

Thus giving (3.52) and (3.53).

We now set $W(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and (3.54) simplifies to

$$W(x) = \begin{pmatrix} \cos(\sqrt{\lambda}x) \\ -\sin(\sqrt{\lambda}x) \end{pmatrix} + \frac{1}{\sqrt{\lambda}} \int_0^x q(t) \begin{pmatrix} \sin(\sqrt{\lambda}(x-t)) \cos(\sqrt{\lambda}t) \\ \cos(\sqrt{\lambda}(x-t)) \cos(\sqrt{\lambda}t) \end{pmatrix} dt + O\left(\frac{1}{\lambda}\right). \quad (3.56)$$

Using the product-to-sum formula and then applying the Riemann-Lebesgue⁴ theorem we obtain (3.50) and (3.51). ■

Theorem 3.8 *The asymptotic form of the eigenvalues of the problem (3.1) are,*

(i) $\sin \alpha = \sin \beta = 0$,

$$\sqrt{\lambda_n} = n\pi + \frac{1}{2n\pi} \int_0^1 qdt + o\left(\frac{1}{n}\right), \quad (3.57)$$

(ii) $\sin \alpha = 0, \sin \beta \neq 0$,

$$\sqrt{\lambda_n} = n\pi - \frac{\pi}{2} + \frac{1}{n\pi} \left[\frac{1}{2} \int_0^1 qdt - \cot \beta \right] + o\left(\frac{1}{n}\right), \quad (3.58)$$

(iii) $\sin \alpha \neq 0, \sin \beta = 0$,

$$\sqrt{\lambda_n} = n\pi - \frac{\pi}{2} + \frac{1}{n\pi} \left[\frac{1}{2} \int_0^1 qdt + \cot \alpha \right] + o\left(\frac{1}{n}\right), \quad (3.59)$$

⁴cf. Titchmarsh, [42] pp. 403

(iv) $\sin \alpha \sin \beta \neq 0$,

$$\sqrt{\lambda_n} = (n-1)\pi + \frac{1}{n\pi} \left[\frac{1}{2} \int_0^1 q dt + \cot \beta - \cot \alpha \right] + o\left(\frac{1}{n^2}\right), \quad (3.60)$$

as $n \rightarrow \infty$ for $n \in \mathbb{N}$.

Proof: A solution satisfying problem (3.1) and boundary condition (3.2) is

$$w(x) = \sin \alpha y_1(x) + \cos \alpha y_2(x). \quad (3.61)$$

For case (i) we substitute (3.61) into (3.3) and obtain $y_2(1) = 0$. Using theorem 3.7 we have $\sin \sqrt{\lambda} = O(1/n)$. For $n \in \mathbb{N}$ large let

$$\sqrt{\lambda_n} = \pi n + \delta_n \quad \text{for} \quad \delta_n \in (-\pi/2, \pi/2]. \quad (3.62)$$

However using the trigonometric difference identities we have

$$|\sin \sqrt{\lambda_n}| = |\sin \delta_n| > 1/\sqrt{2} \neq O\left(\frac{1}{n}\right) \quad \text{for } \delta_n \in (-\pi/2, -\pi/4] \cup [\pi/4, \pi/2]. \quad (3.63)$$

Thus $\delta_n \in (-\pi/4, \pi/4)$ and $1/\sqrt{\lambda_n} = O(1/n)$ giving

$$\sin \delta_n = O\left(\frac{1}{n}\right) \quad \text{for } \delta_n \in (-\pi/4, \pi/4). \quad (3.64)$$

From the limits of δ_n we have $1/\sqrt{2} \leq \cos \delta_n \leq 1$, as $\sin 0 = 0$,

$$|\delta_n| \leq \sqrt{2} \sin \delta_n = O\left(\frac{1}{n}\right). \quad (3.65)$$

Thus $\delta_n = O(1/n)$ and

$$\begin{aligned} \sin \sqrt{\lambda_n} &= (-1)^n \delta_n + O\left(\frac{1}{n^2}\right), \\ \cos \sqrt{\lambda_n} &= (-1)^n + O\left(\frac{1}{n^2}\right). \end{aligned}$$

The equation $y_2(1) = 0$ gives

$$\frac{\sin(\sqrt{\lambda_n})}{\sqrt{\lambda_n}} - \frac{\cos(\sqrt{\lambda_n})}{2\lambda_n} \int_0^1 q dt = o\left(\frac{1}{\lambda_n}\right).$$

Combining the above we have

$$\delta_n - \frac{1}{2(n\pi + \delta_n)} \int_0^1 q dt = o\left(\frac{1}{n}\right).$$

Thus

$$\delta_n = \frac{1}{2n\pi} \int_0^1 q \, dt + o\left(\frac{1}{n}\right),$$

from which case (i) follows.

For case (ii) we substitute (3.61) into (3.3) and obtain

$$y_2(1) \cos \alpha \cos \beta - y_2'(1) \sin \beta \cos \alpha = 0, \quad (3.66)$$

and proceed in an manner similar to that for case (i). For case (iii) we substitute (3.61) into (3.3) and obtain

$$y_1(1) + \cot \alpha y_2(1) = 0, \quad (3.67)$$

and proceed as before. Finally for case (iv) we substitute (3.61) into (3.3) and obtain

$$\cot \beta y_1(1) + \cot \alpha \cot \beta y_2(1) - y_1'(1) - \cot \alpha y_2'(1) = 0. \quad (3.68)$$

■

We will use the following results in the next section. Let u_1 and v_1 be the solutions of (3.1) obeying the following initial and terminal conditions respectively:

$$\begin{aligned} u_1(a) &= \sin \alpha, & u_1'(a) &= \cos \alpha, \\ v_1(b) &= \sin \beta & v_1'(b) &= \cos \beta. \end{aligned}$$

From theorem 3.6 with the variable x shifted by a we have

$$u_1 = \sin \alpha y_1 + \cos \alpha y_2,$$

giving

$$u_1(x) = \cos(\sqrt{\lambda}(x-a)) \sin \alpha + O(e^{|t|(x-a)}), \quad (3.69)$$

$$u_1'(x) = -\sqrt{\lambda} \sin(\sqrt{\lambda}(x-a)) \sin \alpha + O(e^{|t|(x-a)}). \quad (3.70)$$

Reversing the direction in which the interval is traversed we obtain

$$v_1(x) = \cos(\sqrt{\lambda}(b-x)) \sin \beta + O(e^{|t|(b-x)}), \quad (3.71)$$

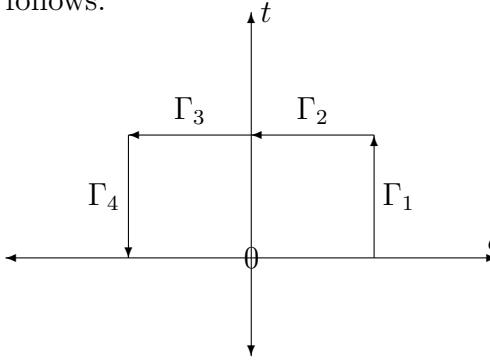
$$v_1'(x) = \sqrt{\lambda} \sin(\sqrt{\lambda}(b-x)) \sin \beta + O(e^{|t|(b-x)}). \quad (3.72)$$

3.5 Properties of the Sturm-Liouville expansion

In this section an important result pertaining to the Sturm-Liouville problem will be proven, namely, that the Sturm-Liouville expansion converges in a similar manner to that of an ordinary Fourier series, that is, to the mean value of the left and right limits of each point in $[a, b]$. This is summarized by the following theorem.

Theorem 3.9 *The Sturm-Liouville expansion of $f \in \mathcal{L}^1([a, b])$ for $a < x < b$, defined in (3.27), converges in the same way as a Fourier series, furthermore, it converges to $\frac{1}{2}\{f(x+0) + f(x-0)\}$ if $f(x)$ is of bounded variation in the neighborhood of x .*

Proof: Using the asymptotic estimates previously established we will consider the integral of the resolvent $\frac{1}{2\pi i} \int_{\Gamma} R(L : \lambda) i f d\lambda$, around the closed contour described as follows:



$$\begin{aligned} \Gamma_1 : s_1(t) &= \frac{(n+\frac{1}{2})\pi}{(b-a)} + it \\ \Gamma_2 : s_2(t) &= (i+1)\frac{(n+1/2)\pi}{b-a} - t \\ \Gamma_3 : s_3(t) &= i\frac{(n+1/2)\pi}{b-a} - t \\ \Gamma_4 : s_4(t) &= (i-1)\frac{(n+1/2)\pi}{b-a} - it \end{aligned}$$

Let $\lambda = s^2$ where $s = \sigma + it$, and $0 \leq t \leq \frac{(n+1/2)\pi}{b-a}$. The above contour is a closed λ -contour which will be denoted by Γ . We consider the case when $\sin \beta \neq 0, \sin \alpha \neq 0$, using equations (3.69) and (3.71), we calculate the Wronskian of u_1 and v_1 to be,

$$\omega(\lambda) = s \sin(s(b-a)) \sin \alpha \sin \beta + O(e^{t(b-a)}). \quad (3.73)$$

Along the contour, for each path Γ_i we have $|\sin(s(b-a))| > Ae^{t(b-a)}$, for some fixed $A > 0$. Thus

$$\frac{1}{|\sin(s(b-a))|} = O(e^{-t(b-a)}). \quad (3.74)$$

From (3.73) where $k := \sin \alpha \sin \beta$, using (3.74) gives,

$$\frac{1}{\omega(\lambda)} = \frac{1}{ks \sin(s(b-a))} \left(1 + \frac{O(e^{t(b-a)}|s|^{-1})}{\sin(s(b-a))} \right) = \frac{1}{ks \sin(s(b-a))} + O\left(\frac{1}{|s|}\right). \quad (3.75)$$

Combining (3.75) and (3.70), we obtain that the first term (3.11) of the resolvent is

$$\int_a^x \frac{\cos(s(b-x)) \cos(s(y-a))}{s \sin(s(b-a))} f(y) dy + O\left(\frac{1}{|s|^2} \int_a^x e^{t(y-x)} |f(y)| dy\right). \quad (3.76)$$

Let $0 < \delta < x - a$, then looking at the asymptotic term, we can split the integral as follows

$$\int_a^x e^{t(y-x)} |f(y)| dy = \left(\int_a^{x-\delta} + \int_{x-\delta}^x \right) e^{t(y-x)} |f(y)| dy.$$

Here, as $t \geq 0$, we have $e^{t(y-x)} \leq 1$ for $y \in [x - \delta, x]$ giving

$$\int_{x-\delta}^x e^{t(y-x)} |f(y)| dy \leq \int_{x-\delta}^x |f(y)| dy.$$

For the other term we have

$$\int_a^{x-\delta} e^{t(y-x)} |f(y)| dy \leq e^{-\delta t} \int_a^{x-\delta} |f(y)| dy.$$

Using the bounds above, the asymptotic term in (3.76), reduces to,

$$O\left(\frac{1}{|s|^2} \int_a^x e^{t(y-x)} |f(y)| dy\right) = O\left(\frac{1}{|s|^2} \int_{x-\delta}^x |f(y)| dy\right) + O\left(\frac{e^{-\delta t}}{|s|^2}\right).$$

We now integrate the above asymptotic term around the λ -contour given by $\lambda = s^2$ with $s \in \Gamma$ for large $n \in \mathbb{N}$. To do this we change variables from the λ to s . Here $d\lambda = 2s ds$ and the integral of the above gives,

$$\begin{aligned} \int O\left(\frac{1}{|s|^2} \int_a^x e^{t(y-x)} |f(y)| dy\right) d\lambda &= O\left(\int_{x-\delta}^x |f(y)| dy \int_{\Gamma} \left|\frac{ds}{s}\right|\right) + O\left(\int_{\Gamma} e^{-\delta t} \left|\frac{ds}{s}\right|\right) \\ &= O\left(\int_{x-\delta}^x |f(y)| dy\right) + O\left(\int_{\Gamma} e^{-\delta t} \left|\frac{ds}{s}\right|\right). \end{aligned}$$

The first asymptotic term in the above tends to zero as δ tends to zero. The second term integrated around Γ gives,

$$O\left(\int_{\Gamma} e^{-\delta t} \left|\frac{ds}{s}\right|\right) = O\left(\frac{1}{n\delta}\right).$$

For large $n \in \mathbb{N}$ let $\delta_n = 1/\sqrt{n}$, now

$$\int O\left(\frac{1}{|s|^2} \int_a^x e^{t(y-x)} |f(y)| dy\right) d\lambda = O\left(\int_{x-\delta_n}^x |f(y)| dy\right) + O\left(\frac{1}{\sqrt{n}}\right) \rightarrow 0$$

as $n \rightarrow \infty$.

Finally, a similar argument yields the second part of the resolvent where $x \leq y \leq b$.

In summary we have,

$$\begin{aligned} \frac{1}{2\pi i} \int R(L : \lambda) f d\lambda &= \frac{1}{2\pi i} \int \left\{ \int_a^x \frac{\cos(s(b-x)) \cos(s(y-a))}{s \sin(s(b-a))} f(y) dy \right. \\ &\quad \left. + \int_x^b \frac{\cos(s(b-y)) \cos(s(x-a))}{s \sin(s(b-a))} f(y) dy \right\} d\lambda + o(1). \end{aligned} \quad (3.77)$$

The right hand side is the sum of two Fourier cosine expansions thus completing the proof. ■

Chapter 4

The Sturm theorems

4.1 Introduction

In this chapter, classical proofs of the Sturm comparison and oscillation theorems are considered from perspectives varying through a function analytic perspective to a more geometric argument at the end of the chapter. The first section, 4.2, gives a series of proofs as considered by Ince [18]. In section 4.3 a Prüfer angle argument is used, while retaining some functional analysis as seen in [3]. For the final section a geometric proof described by Kamke [21] is given. Throughout this chapter the following slightly more general problem is considered

$$\frac{d}{dx} \left[p(x, \lambda) \frac{dy}{dx} \right] + g(x, \lambda)y = 0, \quad x \in [a, b], \quad (4.1)$$

satisfying boundary conditions

$$\alpha'y(a) - \alpha y'(a) = 0, \quad (4.2)$$

$$\beta'y(b) + \beta y'(b) = 0. \quad (4.3)$$

Where p, p' and g are continuous, $p, g > 0$, p and g are monotonically decreasing functions of λ on the interval (Λ_1, Λ_2) . When the oscillation theorem is applied to (1.1) with boundary conditions (1.2)-(1.3), we obtain the eigenfunctions corresponding to the n th eigenvalue have n zeros on (a, b) , $n = 0, 1, \dots$

4.2 Ince

In this section the Sturm Comparison and Oscillation theorems will be proved in a completely function analytic manner. The Sturm comparison theorems are instrumental in the proof of the Oscillation theorems, and will be proved first. Under suitable conditions, the comparison theorems relate the distribution of zeros of solutions of a Sturm-Liouville equation to the distribution of zeros of solutions of a second Sturm-Liouville equation.

For the comparison theorems consider the two equations

$$\frac{d}{dx} \left[p_1(x) \frac{d\psi}{dx} \right] + g_1(x)\psi = 0, \quad (4.4)$$

$$\frac{d}{dx} \left[p_2(x) \frac{d\varphi}{dx} \right] + g_2(x)\varphi = 0, \quad (4.5)$$

with $p_1 \geq p_2 > 0$, $g_2 \geq g_1$ on $[a, b]$. We say that φ **oscillates more rapidly** than ψ if it has more zeros on the interval $[a, b]$. We first derive Picone's formula.

Lemma 4.1 (Picone's formula) *For solutions φ and ψ of (4.4) and (4.5) with φ having no zeros in the interval $[x_1, x_2]$, the following equation holds,*

$$\left[\frac{\psi}{\varphi} (p_1 \psi' \varphi - p_2 \psi \varphi') \right]_{x_1}^{x_2} = \int_{x_1}^{x_2} (g_2 - g_1) \psi^2 dx + \int_{x_1}^{x_2} (p_1 - p_2) \psi'^2 dx + \int_{x_1}^{x_2} p_2 \left(\psi' - \frac{\psi \varphi'}{\varphi} \right)^2 dx.$$

Proof: Applying the product rule gives

$$\frac{d}{dx} \left[\frac{\psi}{\varphi} (p_1 \psi' \varphi - p_2 \psi \varphi') \right] \quad (4.6)$$

$$= \frac{\psi}{\varphi} [\varphi(p_1 \psi')' - \psi(p_2 \varphi')' + (p_1 - p_2) \varphi' \psi'] + \left(\frac{\psi' \varphi - \psi \varphi'}{\varphi^2} \right) [p_1 \psi' \varphi - p_2 \psi \varphi']. \quad (4.7)$$

Expanding the second term of (4.7), yields,

$$\left(\frac{\psi' \varphi - \psi \varphi'}{\varphi^2} \right) [p_1 \psi' \varphi - p_2 \psi \varphi'] = p_1 \psi'^2 - (p_1 + p_2) \frac{\psi \varphi' \psi'}{\varphi} + \frac{p_2 \psi^2 \varphi'^2}{\varphi^2}, \quad (4.8)$$

while using equations (4.4) and (4.5), the first term of (4.7) gives,

$$\frac{\psi}{\varphi} [\varphi(p_1 \psi')' - \psi(p_2 \varphi')' + (p_1 - p_2) \varphi' \psi'] = \frac{\psi}{\varphi} [(g_2 - g_1) \psi \varphi + (p_1 - p_2) \psi' \varphi']. \quad (4.9)$$

Combining the terms (4.8) and (4.9) with the term $p_2\psi'^2 - p_2\psi'^2$, (4.7) becomes,

$$\begin{aligned} & (g_2 - g_1)\psi^2 + p_1\psi'^2 + p_2\psi'^2 - p_2\psi'^2 - 2p_2\frac{\psi'\varphi'\psi}{\varphi} + p_2\frac{\psi^2\varphi'^2}{\varphi^2} \\ &= (g_2 - g_1)\psi^2 + (p_1 - p_2)\psi'^2 + p_2\left(\psi' - \frac{\psi\varphi'}{\varphi}\right)^2. \end{aligned}$$

Integrating the above and (4.6) gives the result. ■

Theorem 4.2 (First Comparison theorem) *Let ψ and φ be non-trivial solutions of (4.5) and (4.4) obeying the initial conditions,*

$$\alpha'_1\psi(a) - \alpha_1\psi'(a) = 0, \quad (4.10)$$

$$\alpha'_2\varphi(a) - \alpha_2\varphi'(a) = 0. \quad (4.11)$$

In addition assume:

(i) α_i and α'_i are not both zero;

(ii) if $\alpha_1 \neq 0$ then

$$\frac{p_1(a)\alpha'_1}{\alpha_1} \geq \frac{p_2(a)\alpha'_2}{\alpha_2};$$

(iii) there is no non-empty open subinterval of (a, b) on which $g_1 - g_2 = 0$.

Then φ oscillates more rapidly than ψ , with the i^{th} zero of φ less than the i^{th} zero of ψ .

Proof: Let x_i be the zeros of ψ on (a, b) ordered as follows,

$$a < x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n \leq b.$$

We first show that between consecutive zeros, x_i and x_{i+1} , of ψ , there is at least one zero of φ . Assume that this is not the case. Then $\varphi \neq 0$ on (x_i, x_{i+1}) .

We first consider the case when φ is not zero at the end points. Without loss of generality, we assume that $\psi \geq 0$ and $\varphi > 0$ on the interval $[x_i, x_{i+1}]$. The right side of Picone's formula is strictly positive and thus,

$$0 = \left[\frac{\psi}{\varphi} (p_1 \psi' \varphi - p_2 \psi \varphi') \right]_{x_i}^{x_{i+1}} > 0.$$

This contradiction shows that φ must vanish on $[x_i, x_{i+1}]$.

If ψ vanishes at an endpoints, say x_i ,

$$\lim_{x \rightarrow x_i} \left[\frac{\psi}{\varphi} (p_1 \psi' \varphi - p_2 \psi \varphi') \right] = (p_1 \psi' \varphi - p_2 \psi \varphi')(x_i) \lim_{x \rightarrow x_i} \frac{\psi'}{\varphi'} = 0. \quad (4.12)$$

Thus in the limit the left hand side of Picone's formula is zero while the right hand side is positive, yielding a contradiction. It follows that φ vanishes between consecutive zeros of ψ . In the above argument we may replace φ and ψ with ψ and φ , respectively, concluding that the zeros of φ and ψ separate each other.

All that is left to show is that φ vanishes once on the interval $[a, x_1]$. If $\psi(a) = 0$ then by the above reasoning φ vanishes on $[a, x_1]$, furthermore φ cannot vanish more than once on $[a, x_1]$ otherwise ψ would have a zero less than x_1 . Suppose φ is non-zero on $[a, x_1]$ and $\alpha_1 \neq 0$ so Picone's formula can be applied. Assuming without loss of generality that φ is positive then the right hand side of Picone's formula is positive while by assumption (ii), the left hand side becomes

$$0 \geq -\psi^2(a) \left(p_1(a) \frac{\alpha'_1}{\alpha_1} - p_2(a) \frac{\alpha'_2}{\alpha_2} \right).$$

From this contradiction it follows that φ has a zero in $[a, x_1]$. ■

Theorem 4.3 (Second Comparison theorem)

Consider an interior point c of the interval (a, b) such that ψ and φ have the same number of zeros on the interval (a, c) . Let all the assumptions be preserved as in the first comparison theorem furthermore let $\varphi(c) \neq 0$ then $\psi(c) \neq 0$ and

$$\frac{p_1(c) \psi'(c)}{\psi(c)} > \frac{p_2(c) \varphi'(c)}{\varphi(c)}.$$

Proof: Suppose there are i zeros of ψ and φ on the interval considered. Let x_i be the greatest zero of ψ before c . Then by the first comparison theorem, the i th zero of φ is necessarily less than the i th zero of ψ on $a < x \leq x_i$. Thus φ is non-zero on $[x_i, c]$ and ψ is non-zero on $(x_i, c]$. Considering the Picone formula taken between the limits x_i and c , as the right hand side of Picone's formula is positive, we get

$$\left[\frac{p_1 \psi'}{\psi} - \frac{p_2 \varphi'}{\varphi} \right]_{x_i}^c > 0.$$

This gives the desired inequality. ■

The first comparison theorem tells us something important about the nature of the zeros of the solutions, y , of equation (4.1). In particular, as λ increases so $g(x, \lambda)$ diminishes continuously with λ and the zeros of y move left towards a in the interval $[a, b]$. As $g(x, \lambda)$ and $p(x, \lambda)$ are continuous on $[a, b]$, they are bounded for each fixed λ . Since $g(x, \lambda)$ and $p(x, \lambda)$ are monotonically decreasing functions of λ on (Λ_1, Λ_2) and bounds on them can be chosen to be monotonically decreasing as follows:

$$\mathbf{P}(\lambda) \geq p(x, \lambda) \geq \mathbf{p}(\lambda) > 0 \quad \text{for} \quad x \in [a, b],$$

$$\mathbf{G}(\lambda) \geq g(x, \lambda) \geq \mathbf{g}(\lambda) \quad \text{for} \quad x \in [a, b].$$

Theorem 4.4 (Oscillation theorem)

Consider the equation (4.1) with boundary conditions (4.2 and (4.3). Furthermore assume

$$\mathbf{g}/\mathbf{P} \rightarrow +\infty \quad \text{as} \quad \lambda \rightarrow \Lambda_2, \quad (4.13)$$

$$\mathbf{G}/\mathbf{p} \rightarrow -\infty \quad \text{as} \quad \lambda \rightarrow \Lambda_1, \quad (4.14)$$

then the eigenvalues are bounded below and form a monotonic increasing sequence,

$$\lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_n < \dots, \quad (4.15)$$

with accumulation point at infinity, furthermore the eigenfunction y_n corresponding to the eigenvalue λ_n has n zeros in the interval (a, b) .

Proof: Consider a non-trivial solution, y of (4.1) satisfying the boundary condition (4.2). From the first comparison theorem, the solution $z = \sin(\sqrt{\mathbf{g}/\mathbf{P}}x)$ of

$$\frac{d^2 z}{dx^2} + \frac{\mathbf{g}}{\mathbf{P}} z = 0$$

oscillates at most as rapidly as y , the solution of (4.1). From this we see that the assumption (4.13) guarantees an arbitrarily large number of zeros of the solution of (4.1) for λ sufficiently near Λ_2 . We know that $p(\lambda, x)$ and $g(\lambda, x)$ are continuous monotonic decreasing functions of λ . Thus from the first comparison theorem as λ increases, zeros of y will move left towards a . Let μ_i be the largest value for λ such that the solutions of (4.1) have no more than i zeros on the interval (a, b) . We have an increasing sequence of numbers $\{\mu_n\}$ where on the interval

$$\mu_i < \lambda < \mu_{i+1},$$

the solutions have exactly $i + 1$ zeros on (a, b) . Applying the second comparison theorem for $\lambda_1 < \lambda_2$ on the interval above we have

$$p(\lambda_1, b)y'(b)/y(b) > p(\lambda_2, b)y'(b)/y(b). \quad (4.16)$$

Thus inequality (4.16) shows that

$$p(\lambda, b)y'(b)/y(b) \quad (4.17)$$

is a monotonic decreasing function of λ on (μ_i, μ_{i+1}) . When $\lambda = \mu_{i+1}$ and $\lambda = \mu_i$ we see that $y(b) = 0$ and $y'(b) \neq 0$, thus (4.17) must decrease from $+\infty$ to $-\infty$. Since β and β' are constants and $p(\lambda, x)$ is monotonically decreasing in λ , we see that

$$-p(\lambda, b)\beta'/\beta \quad (4.18)$$

is a monotonically increasing function of λ on (μ_i, μ_{i+1}) . Using a similar argument as above we see that it increasing from $-\infty$ to ∞ as λ increases from μ_i to μ_{i+1} . At some point λ_{i+1} the two function values (4.17) and (4.18) must be equal yielding following equality

$$-\frac{p(\lambda, b)y'(b)}{y(b)} = \frac{p(\lambda, b)\beta'}{\beta},$$

where the solution y satisfying the first boundary condition, (4.2), also satisfies the above which is equivalent to the second boundary condition, (4.3). Thus when $\lambda = \lambda_{i+1}$, we see that the corresponding solution y_n is an eigenfunction which has precisely $i + 1$ zeros on the interval (a, b) . All that is left to show is that there is a λ_0 below which there are no consistent solutions to the system (4.1). We first investigate the nature of the zeros of the solutions as λ tends to Λ_1 . Consider the equation,

$$\mathbf{p} \frac{d^2 u}{dx^2} + \mathbf{G}u = 0, \quad (4.19)$$

with solution u satisfying the boundary conditions (4.2). For λ sufficiently near Λ_1 we may rewrite the above equation as

$$\frac{d^2 u}{dx^2} - \frac{|\mathbf{G}|}{\mathbf{p}} u = 0. \quad (4.20)$$

Setting $s := \sqrt{|\mathbf{G}|/\mathbf{p}}$, we see that (4.20) has the solution

$$u(x) = \frac{1}{2} \left(\alpha + \frac{\alpha'}{s} \right) e^{s(x-a)} + \frac{1}{2} \left(\alpha - \frac{\alpha'}{s} \right) e^{-s(x-a)}. \quad (4.21)$$

Consequently for λ near Λ_1 , s becomes large and the above solution is approximated by $\alpha \cosh s(x - a)$, which has no zeros. Applying the first comparison theorem to equations (4.1) and (4.20), we see that y oscillates at most as rapidly as u . Therefore there is a least μ_i , namely μ_0 , below which the solutions of (4.1) have no zeros. The second comparison theorem shows that for every λ on the interval (Λ_1, μ_0) the following holds

$$\frac{p(\lambda, b)y'(b)}{y(b)} > \frac{\mathbf{p}(\lambda)u'(b)}{u(b)}. \quad (4.22)$$

As $\lambda \rightarrow \Lambda_1$ it follows that $|s| \rightarrow +\infty$, thus we may use the hyperbolic approximation for $u(x)$ and the above inequality to show that

$$p(\lambda, b)y'(b)/y(b) \rightarrow +\infty.$$

This shows that as λ increases on (Λ_1, μ_0) , $p(b)y'(b)/y(b)$ decreases from $+\infty$ to $-\infty$. Since β' and β are constants, we see that $-p(b)\beta'/\beta$ increases on the same interval. Therefore there must exist a single eigenvalue λ_0 in (Λ_1, μ_0) corresponding to a solution of (4.1) with no zeros in (a, b) . ■

We can now apply the above oscillation theorem to (1.1) taking (Λ_1, Λ_2) as the interval $(-\infty, \infty)$ where we see that $p(x)$ is bounded and not dependent on λ , also

$$\begin{aligned} \mathbf{G} &= \max_{x \in (a,b)} (\lambda - q(x)) \rightarrow +\infty & \text{as } \lambda \rightarrow \infty, \\ \mathbf{g} &= \min_{x \in (a,b)} (\lambda - q(x)) \rightarrow -\infty & \text{as } \lambda \rightarrow -\infty. \end{aligned}$$

The Sturm theorem thus follows in this case.

4.3 Coddington and Levinson

We now study a more geometric proof of the comparison and oscillation theorems. This approach involves considering the solutions of (1.1) in the phase plane and converting the co-ordinate system into polar co-ordinates. In this way we only need to consider the angle part of the solutions and not the length part, furthermore a criteria for a solution obtaining a zero can be clearly seen. Let $p(x)y' = \tilde{y}$, then (1.1) gives

$$y' = \frac{\tilde{y}}{p(x)} \quad \text{and} \quad \tilde{y}' = -g(x)y, \quad (4.23)$$

Let

$$y = r \sin \theta \quad \text{and} \quad \tilde{y} = r \cos \theta, \quad (4.24)$$

applying the chain rule to (4.24), we can write (1.1) as a system, thus

$$\begin{pmatrix} \sin \theta & r \cos \theta \\ \cos \theta & -r \sin \theta \end{pmatrix} \begin{pmatrix} r' \\ \theta' \end{pmatrix} = \begin{pmatrix} 0 & 1/p(x) \\ -g(x) & 0 \end{pmatrix} \begin{pmatrix} r \sin \theta \\ r \cos \theta \end{pmatrix}.$$

By multiplying the above equation by the inverse of the matrix on the left of the equality, we obtain the following equations,

$$r' = \left(\frac{1}{p} - g\right) r \sin \theta \cos \theta, \quad (4.25)$$

$$\theta' = \frac{1}{p} \cos^2 \theta + g \sin^2 \theta, \quad (4.26)$$

thus for φ to be a solution of (1.1), there are also solutions $r = \rho(x)$ and $\omega = \theta(x)$ of (4.25) and (4.26) respectively such that $\varphi(x) = \rho(x) \sin \omega(x)$. Using pythagorus and

trigonometric identities we have the following two identities,

$$\rho^2 = (p\varphi')^2 + \varphi^2 \quad \text{and} \quad \omega = \tan^{-1} \left(\frac{\varphi}{p\varphi'} \right).$$

Since there is no nontrivial solution that satisfies $\varphi(c) = 0$ and $\varphi'(c) = 0$ for some $c \in [a, b]$, thus we may assume that $\rho > 0$, therefore the zero's of φ are multiples of π . We can also see that the boundary condition, (1.2) and (1.3), using (4.23) and (4.24), can be replaced by simpler conditions, namely $\theta(a) = \alpha(\text{mod } \pi)$ and $\theta(b) = \beta(\text{mod } \pi)$ respectively. We can now restate the comparison theorems as follows.

Theorem 4.5 *Consider equations defined as (4.4) and (4.5), assuming that $\omega_2(a) \geq \omega_1(a)$, then*

$$\omega_2(x) \geq \omega_1(x) \quad \text{for } x \in [a, b], \quad (4.27)$$

Furthermore, if $g_2 > g_1$ on (a, b) , then

$$\omega_2(x) > \omega_1(x) \quad \text{for } x \in (a, b], \quad (4.28)$$

Proof: To each of the equations we perform the above procedure to obtain

$$\omega'_i = \frac{1}{p} \cos^2 \omega_i + g \sin^2 \omega_i \quad \text{for } i = 1, 2. \quad (4.29)$$

Subtracting the above two equations and adding the term $g_1 \sin^2 \omega_2 - g_1 \sin^2 \omega_1$ we have

$$(\omega_2 - \omega_1)' = \left(g_1 - \frac{1}{p_1}\right)(\sin^2 \omega_2 - \sin^2 \omega_1) + h \quad (4.30)$$

where

$$h = \left(\frac{1}{p_2} - \frac{1}{p_1}\right) \cos^2 \omega_2 + (g_2 - g_1) \sin^2 \omega_2. \quad (4.31)$$

We can see from first assumptions that $h \geq 0$. Let $u := \omega_2 - \omega_1$, then setting

$$f := \left(g_1 - \frac{1}{p_1}\right)(\sin \omega_2 + \sin \omega_1) \left(\frac{\sin \omega_2 - \sin \omega_1}{\omega_2 - \omega_1}\right), \quad (4.32)$$

we may rewrite (4.30) as

$$u' = fu + h. \quad (4.33)$$

Since p_i and g_i are point wise continuous hence bounded on a compact interval we see that f is uniformly bounded. Since $h \geq 0$, (4.33) gives,

$$u' - fu \geq 0. \quad (4.34)$$

We now can multiply this equation by an integrating factor, set $F(x) := \int_t^b f(s)ds$, we now multiply (4.34) by e^F , yielding

$$e^F u' + F' e^F u = (e^F u)' \geq 0. \quad (4.35)$$

Integrating the above equation on (a, t) gives

$$e^{F(x)} u(x) \geq e^{F(a)} u(a) \geq 0, \quad (4.36)$$

this proves (4.27). We will prove (4.28) by contradiction, suppose that the statement is false, then we will show that in this case the only possibility is that there exists a $c > a$ such that

$$\omega_2(t) = \omega_1(t) \quad \text{for} \quad t \in [a, c]. \quad (4.37)$$

To show this, we suppose it is not the case, then there is a sequence of points t_j with a as its cluster point such that for each t_j we have $\omega_2(t_j) > \omega_1(t_j)$. We then consider (4.36) with a replaced with t_j , we see that for each $t > t_j$, we have $\omega_2(t) > \omega_1(t)$, but t_j has its limit at a , this shows (4.28) is true, thus so is (4.37). Using (4.37), we see that

$$(\omega_2(t) - \omega_1(t))' = 0 \quad \text{for} \quad t \in (a, c).$$

But for the above to be true with $g_2 > g_1$, (4.30) requires that $p_1 = p_2$ on (a, c) and

$$\omega_1 = \omega_2 = 0(\text{mod}\pi) \quad \text{on} \quad (a, c),$$

but (4.29) show this to be impossible, thus (4.28) is proven. ■

The above theorem thus shows that $\omega(t, \lambda)$ is a strictly monotonic increasing function of λ .

Theorem 4.6 *Consider equation (1.1) with boundary conditions (1.2)-(1.3), then the eigenvalues are bounded below and form a monotonic increasing sequence,*

$$\lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_n < \dots, \quad (4.38)$$

with accumulation point at infinity, furthermore, the eigenfunction y_n corresponding to the eigenvalue λ_n has n zeros in the interval (a, b) .

Proof: Without loss of generality we may assume $\alpha \in [0, \pi)$, $\beta \in (0, \pi]$. From existence theorem 2.11, we may find a non-trivial solution φ of the equation (1.1) which satisfies the boundary condition (1.2). Equivalently for $\omega(t, \lambda)$, the Prüfer angle corresponding to φ , the condition $\omega(a, \lambda) = \alpha$ is obeyed. If λ is such that the solution $\varphi = \varphi(t, \lambda)$ satisfies both boundary conditions (1.2) and (1.3), then λ is an eigenvalue. We see by the comparison theorem, that for any fixed x in $(a, b]$, ω is a monotonic increasing function of λ . Furthermore, the zeros of φ occur when $\omega = 0 \pmod{\pi}$. Considering equation (4.29), we see that $\omega' > 0$ when $\omega = 0 \pmod{\pi}$. This shows that ω is strictly increasing at the zeros of φ . Thus the zeros may be labeled consecutively as $x_1 < x_2 < \dots$ and $\omega(x_k, \lambda) = k\pi$. This also shows that when $x > x_k$, we have $\omega(x, \lambda)$ is strictly greater than $k\pi$. From the chain rule we obtain,

$$\frac{d}{dx}\omega(x_k(\lambda), \lambda) = \omega'(x_k, \lambda)\frac{dx_k}{d\lambda} + \frac{\partial\omega(x_k, \lambda)}{\partial\lambda} = 0.$$

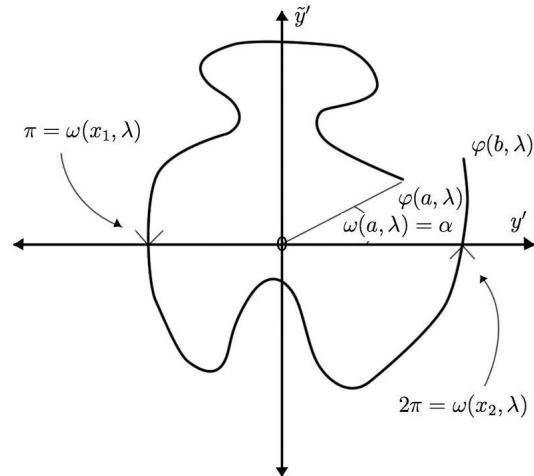
Combining the above with the fact that ω is monotonic increasing in λ , we have that the zeros of φ decrease monotonically in λ , that is $x_k = x_k(\lambda)$ is monotonically decreasing function in λ .

For a fixed $t = c$ in $(a, b]$ we will prove the following,

$$\omega(c, \lambda) \rightarrow \infty \quad \text{as} \quad \lambda \rightarrow \infty, \quad (4.39)$$

$$\omega(c, \lambda) \rightarrow 0 \quad \text{as} \quad \lambda \rightarrow -\infty. \quad (4.40)$$

To prove (4.39) we first see that since $\alpha \geq 0$ we must have that $\omega(x, \lambda) \geq 0$, as otherwise $\omega(x, \lambda)$ would have to decrease through 0, which from the above considerations is impossible. Consider the equation



$$Py'' + (\lambda - Q)y = 0, \quad (4.41)$$

where $p(x) \leq P$ and $q(x) \geq Q$, with solution $\tilde{\varphi}$ satisfying $\tilde{\varphi}(x_0, \lambda) = \varphi(x_0, \lambda)$ and $\tilde{\varphi}'(x_0, \lambda) = \varphi'(x_0, \lambda)$. We may thus take $\tilde{\omega}(x_0, \lambda) = \omega(x_0, \lambda)$. From the comparison theorem we have

$$\omega(c, \lambda) \geq \tilde{\omega}(c, \lambda). \quad (4.42)$$

The zeros of (4.41) have a spacing of $\sqrt{\frac{P\pi^2}{\lambda-Q}}$. Thus, as $\lambda \rightarrow \infty$, the number of zeros in the interval (a, b) grows without bound. Now

$$\frac{\tilde{\omega}(c, \lambda)}{\pi} > \frac{c-a}{\sqrt{\frac{P\pi^2}{\lambda-Q}}} - 2,$$

so from (4.42),

$$\omega(c, \lambda) \rightarrow \infty,$$

proving (4.39).

To prove (4.40), we let $0 < P \leq p(x)$ and $Q \geq q(x)$ for all x in $[a, b]$. For sufficiently negative λ we may consider the equation

$$\frac{d^2u}{dx^2} - \frac{|\lambda - Q|}{P}u = 0, \quad (4.43)$$

with the initial condition (1.2). Setting $s := \sqrt{\frac{|\lambda - Q|}{P}}$ the solution u is given by (4.21). The angular part of this solution ω is given by

$$\omega(x, \lambda) = \tan^{-1} \left(\frac{u(x, \lambda)}{Pu'(x, \lambda)} \right).$$

We will now compute the limit of $\omega(x, \lambda)$ as $\lambda \rightarrow -\infty$. To this end we see that the term inside $\tan^{-1}$ is of an indeterminate form, thus using L'Hôpital's rule and (4.43) we have

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \left(\frac{u(x, \lambda)}{Pu'(x, \lambda)} \right) &= \lim_{\lambda \rightarrow \infty} \left(\frac{u'(x, \lambda)}{Ps^2u(x, \lambda)} \right) \\ &= \lim_{\lambda \rightarrow \infty} \left(\frac{e^{2s(x-a)}(\cos \alpha + Ps \sin \alpha)}{-s(\cos \alpha + Ps \sin \alpha + e^{2s(x-a)}(\cos \alpha + Ps \sin \alpha))} \right). \end{aligned}$$

Using L'Hôpital's rule again we see that

$$\lim_{\lambda \rightarrow \infty} \left(\frac{u(x, \lambda)}{Pu'(x, \lambda)} \right) = \lim_{\lambda \rightarrow \infty} \left(\frac{1}{s} \right).$$

From the above calculation we deduce that the limit of $\omega(x, \lambda)$ as $\lambda \rightarrow -\infty$ is zero for every $x \in (a, b]$. ■

We can now see from (4.39) and (4.40) that since $\beta > 0$, we can choose a λ sufficiently small such that $0 < \omega(b, \lambda) < \beta$ and as we increase λ , there will be a point λ_0 such that $\omega(b, \lambda_0) = \beta$ and at this point condition (1.3) will be satisfied. Furthermore, $0 < \omega(b, \lambda_0) < \pi$ there for the solution $\varphi(t, \lambda_0)$ has no zeros. As λ increase beyond λ_0 there will be another point λ_1 such that $\omega(b, \lambda_1) = \beta + \pi$. We can see using (4.23) and (4.24) that it satisfies (1.3) with exactly one zero on (a, b) , thus the n th eigenvalue λ_n satisfies

$$\omega(b, \lambda_n) = \beta + n\pi$$

and has exactly n zeros on (a, b) . Finally by the comparison theorem there is only one eigenvalue corresponding to a given oscillation count. ■

There are a few other approaches involving the Prüfer angle, such as the approach by Bôcher¹, by converting (1.1) into the corresponding Riccati differential equation,

$$u' + p(x)u^2 + (\lambda - q(x)) = 0, \quad \text{where} \quad h = 1/p(x), \quad yu = p(x)y',$$

Although interesting they give a similar feel as the above while retaining the main principles of this section.

¹cf. Transactions of the American Mathematical Society, vol. 1, 1900, pp. 414-420.

Chapter 5

Variational approach

5.1 Introduction

In this section we introduce a method of finding the eigenvalues of the differential operator (3.4). The Calculus of variations has been known to various mathematicians and physicists throughout history. In particular Lagrange, Morse, Hilbert and Jakobi have been credited with its modern formulation and use. The variational characterization of eigenvalues, as so far as it is understood withing the framework of spectral theory, was originally proposed by Hilbert and Courant, see [4]. It is based on reformulating the eigenvalue problem as a variational problem in the calculus of variations where the eigenvalues correspond to critical points of a certain functional, this formulation is considered in Watson and Currie, [35]. The variational principle is futher abstracted by Lusternik and Schnirelmann to a topologically invariant principle.

5.2 Sobolev space and distributions

The Sobolev Space $\mathcal{H}^1(\mathcal{S})$ and the introduction of the weak derivative on distributions gives an appropriate space on which to define the variational functional. We note

without proof the following important facts; $\mathcal{C}_0^\infty(\mathcal{S})$ is dense in $\mathcal{L}^p(\mathcal{S})$ and the Sobolev space $\mathcal{H}^1(\mathcal{S})$ is compactly imbedded in $\mathcal{L}^p(\mathcal{S})$, see [17, chapter 11].

Definition 5.1 *A distribution T on $\mathcal{S}$ is a conjugate-linear functional defined on the dual space of $\mathcal{C}_0^\infty(\mathcal{S})$, thus for $f \in \mathcal{L}_{loc}^1(\mathcal{S})$, we assign the distribution $T_f \in \mathcal{C}_0^\infty(\mathcal{S})^*$ defined by,*

$$T_f(\psi) = \int_{\mathcal{S}} f \bar{\psi} dt \quad \psi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.1)$$

As each f is identified with a distribution, T_f , we make no distinction between the function and its distribution, thus $f(\psi) = T_f(\psi)$. For the derivative of a distribution to make sense, it is important that on some level it agrees with the derivative of a continuously differentiable function. The definition is inspired by integration by parts.

Definition 5.2 *The α^{th} derivative of the distribution T is the distribution $\partial^\alpha T$ defined by,*

$$\partial^\alpha T(\psi) = (-1)^{|\alpha|} T(D^\alpha \psi), \quad \psi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.2)$$

Thus we can see when $f \in \mathcal{C}^\alpha(\mathcal{S})$ we have,

$$\partial^\alpha f = (-1)^{|\alpha|} \int_{\mathcal{S}} f \overline{(D^\alpha \psi)} dt = \int_{\mathcal{S}} D^\alpha f \bar{\psi} dt = D^\alpha f, \quad \psi \in \mathcal{C}_0^\infty(\mathcal{S}).$$

Hence the derivative of a distribution agrees with that of a derivative of continuously differentiable function and, more importantly, functions in $\mathcal{L}^2(\mathcal{S})$, considered as distributions, have derivatives of all orders.

Definition 5.3 *A function $f \in \mathcal{L}_{loc}^1(\mathcal{S})$ is said to have a α^{th} weak derivative if there exists a $g \in \mathcal{L}_{loc}^1(\mathcal{S})$ such that*

$$\partial^\alpha f(\psi) = g(\psi), \quad \psi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.3)$$

We call g the α^{th} weak derivative of f .

Definition 5.4 *The Sobolev Space of order m , denoted, $\mathcal{H}^m(\mathcal{S})$, is the subspace of $\mathcal{L}^2(\mathcal{S})$ consisting of the functions with weak derivatives that exist up to and including order m which are in $\mathcal{L}^2(\mathcal{S})$, thus,*

$$\mathcal{H}^m(\mathcal{S}) = \{f \in \mathcal{L}^2(\mathcal{S}) | \partial^\alpha f \in \mathcal{L}^2(\mathcal{S}) \text{ for } \alpha \leq m\}. \quad (5.4)$$

Definition 5.5 *A distribution is **constant** if and only if it depends only on the mean value of each $\psi \in \mathcal{C}_0^\infty(\mathcal{S})$, thus*

$$T(\psi) = c \int_{\mathcal{S}} \bar{\psi} dt, \quad \psi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.5)$$

Note that for any distribution T with $\partial T = 0$, it follows that T is constant, i.e., $T(\psi) = T_c(\psi)$. The next lemma is taken from Showalter, see [36].

Lemma 5.6 *If f is absolutely continuous on $\mathbb{R}$, then we can define a $g \in \mathcal{L}_{loc}^1(\mathbb{R})$ by $Df = g$ and $\partial f = g$ on $\mathcal{C}_0^\infty(\mathbb{R})$. Conversely, there exists a distribution T in $\mathcal{C}_0^\infty(\mathbb{R})^*$ such that $\partial T \in \mathcal{L}_{loc}^1(\mathbb{R})$ and $\partial T = \partial f$.*

Proof: Defining f and g as above, since an absolutely continuous function can be represented as an integral of a locally integrable function, we have that $f(x) = \int_0^x g(t)dt + f(0)$, where $g \in \mathcal{L}_{loc}^1(\mathbb{R})$. Using integrating by parts and representing f as a derivation gives

$$\partial f(\psi) = - \int_{\mathbb{R}} f(D\bar{\psi}) dt = \int_{\mathbb{R}} g\bar{\psi} dt = g(\psi), \quad \psi \in \mathcal{L}_{loc}^1(\mathbb{R}). \quad (5.6)$$

Furthermore, let $g = \partial T \in \mathcal{L}_{loc}^1(\mathbb{R})$ and let $h(x) := \int_0^x g(t)dt$. It is clear that $h \in AC$ and by above $\partial h = g$. By linearity, this shows that $\partial(T - h) = 0$, and hence $T - h$ is constant, and thus is of the form $T = h(x) + c$, for some $c \in \mathbb{R}$. This shows that $\partial T = \partial f$. ■

5.3 Variational Formulation

What follows is an exposition and reworking of the mathematics in Watson and Currie [35].

Definition 5.7 *The variational form F of the Sturm-Liouville problem (1.1), is the sesquilinear functional denoted by*

$$F(x, y) = \int_{\partial S} f x \bar{y} d\sigma + \int_S (x' \bar{y}' + q x \bar{y}) dt. \quad (5.7)$$

For brevity we write $F(x, x) = F(x)$ and note that F is continuous on its domain, $\mathcal{D}(F) = \mathcal{H}^1$.

The Sturm-Liouville problem with boundary conditions (3.2)-(3.3) can be converted to a variational problem as follows. Let L be the differential operator defined by (3.4) on in the compact interval $[a, b]$, set $\mathcal{S} := [a, b]$. Define the sesquilinear form $F(x, y)$ as $F(x, y) := (Lx, y)$ which may be rewritten as

$$F(x, y) = (-x'', y) + (qx, y).$$

Rewriting the inner product as an integral and integrating the first term by parts gives,

$$F(x, y) = [-x' \bar{y}]_a^b + \int_S x' \bar{y}' dt + \int_S q x \bar{y} dt.$$

By subjecting x to the boundary conditions (3.2)-(3.3), we see that $x'(a) \sin(\alpha) = x(a) \cos(\alpha)$ and $x'(b) \sin(\beta) = x(b) \cos(\beta)$. Defining

$$\int_{\partial S} y d\sigma := y(b) - y(a), \quad \text{for } y \in \mathcal{D}(F).$$

Defining f on ∂S by $f(a) := -\cot(\alpha)$ and $f(b) := -\cot(\beta)$, gives the required form.

Definition 5.8 *The boundary conditions are **co-normal** if there exists an f defined on ∂S , such that $x \in \mathcal{D}(F)$ has*

$$\int_{\partial S} f x \bar{y} d\sigma = \int_{\partial S} x' \bar{y} d\sigma, \quad \text{for } y \in \mathcal{D}(F).$$

if and only if x obeys the boundary conditions.

From above it is clear that the boundary conditions (3.2)-(3.3) are co-normal. The next theorem shows that L has a least eigenvalue and that the functional $\frac{F(u,u)}{\|u\|^2}$ is bounded below hence the infimum over its domain exists. This important fact will be used to give a min-max characterization of the eigenvalues of (3.1).

Theorem 5.9 *The operator L is lower semibounded.*

Proof: Since L is self-adjoint, using [43, p.247, Corollary 2], we only need to show that L is lower semibounded on a dense subset of the domain of L . To see that is the case, observe that

$$(Lf, f) = \|f'\|^2 + \int_{\mathcal{S}} q|f|^2 dx \geq -\|f\|^2 \operatorname{esssup} |q| \quad \text{for } f \in \mathcal{C}_0^\infty(\mathcal{S}). \quad \blacksquare$$

Definition 5.10 *The form $F(x,y)$ is said to have a **variational eigenvalue** λ with associated **variational eigenfunction** u if $u \in \mathcal{D}(F)$ such that $F(u, v) = \lambda(u, v)$ for all $v \in \mathcal{D}(F)$.*

The next theorem shows that each variational eigenfunction of the form F is also an eigenfunction of the operator eigenvalue problem associated with (3.1).

Lemma 5.11 *The function $u \in \mathcal{D}(F)$ satisfies $F(u, v) = \lambda(u, v)$ for all $v \in \mathcal{D}(F)$ if and only if $u \in \mathcal{H}^2(\mathcal{S})$, u obeys (3.1) and the boundary conditions (3.2)-(3.3).*

Proof: We begin with necessity. Let $u \in \mathcal{H}^2(\mathcal{S})$ be an eigenfunction of (3.1) satisfying boundary conditions (3.2)-(3.3) with associated eigenvalue $\lambda \in \mathbb{R}$. We have

$$F(u, v) = \int_{\partial\mathcal{S}} fu\bar{v} d\sigma + \int_{\mathcal{S}} (u'\bar{v}' + qu\bar{v}) dt \quad \text{for } v \in \mathcal{D}(F).$$

Since u obeys the differential equation (3.1), we have

$$F(u, v) = \int_{\partial\mathcal{S}} fu\bar{v} d\sigma + \int_{\mathcal{S}} (u'\bar{v})' dt + \lambda(u, v).$$

Integrating the second term on the right and imposing boundary conditions gives,

$$F(u, v) = \int_{\partial S} (fu + u')\bar{v} d\sigma + \lambda(u, v).$$

Since the the boundary conditions where shown to be co-normal, the first term on the right is zero, thus proving the necessity of the condition. For sufficiency, assume $u \in \mathcal{D}(F)$ satisfies $F(u, v) = \lambda(u, v)$ for all $v \in \mathcal{D}(F)$. Since $\mathcal{C}_0^\infty(\mathcal{S})$ is a dense subset of $\mathcal{D}(F)$, we have

$$F(u, v) = \lambda(u, v) \quad \text{for } v \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.8)$$

Using Cauchy-Schwartz inequality we obtain $|F(u, v)| \leq |\lambda| \|u\| \|v\|$, hence $F_u(\cdot)$ is bounded on $\mathcal{L}^2(\mathcal{S})$ and can be continuously extended to a bounded functional on $\mathcal{L}^2(\mathcal{S})$. By the Riesz representation theorem, there exists an $h \in \mathcal{L}^2(\mathcal{S})$ such that

$$F(u, v) = (h, v) \quad \text{for } v \in \mathcal{L}^2(\mathcal{S}).$$

Restricting the above equation to the dense subset $\mathcal{C}_0^\infty(\mathcal{S})$ of $\mathcal{L}^2(\mathcal{S})$ we have

$$\int_{\mathcal{S}} u' \bar{v}' dx = - \int_{\mathcal{S}} (qu - h) \bar{v} dt \quad \text{for } v \in \mathcal{C}_0^\infty(\mathcal{S}),$$

thus $\partial u' = qu - h \in \mathcal{L}^2(\mathcal{S}) \subset \mathcal{L}_{loc}^2(\mathcal{S}) = \mathcal{L}_{loc}^1(\mathcal{S})$. Lemma 5.6 shows that $u' \in AC$ and $u'' \in \mathcal{L}_{loc}^1(\mathcal{S})$, thus $lu = -u'' + qu \in \mathcal{L}_{loc}^1(\mathcal{S})$. However $u' \in AC$, therefore u' can be represented as $u' = \int_a^x f ds$ where $f \in \mathcal{L}_{loc}^1(\mathcal{S})$. Integration by parts on the second term of (5.7) is possible giving

$$F(u, v) = \int_{\mathcal{S}} (-u'' + qu) \bar{v} dt \quad \text{for all } v \in \mathcal{C}_0^\infty(\mathcal{S}).$$

Combining the above equation and (5.8) we have $(lu, v) = (\lambda u, v)$ for all $v \in \mathcal{C}_0^\infty(\mathcal{S})$, which is dense in $\mathcal{L}^2(\mathcal{S})$. This shows that $lu = \lambda u$ and $u'' \in \mathcal{L}^2(\mathcal{S})$, therefore showing that $u \in \mathcal{H}^2(\mathcal{S})$. ■

Lemma 5.12 *Let F be the variational form defined in (5.7) and $c \in \mathbb{R}$. Then every sequence $(u_k) \subset \mathcal{H}^1$ satisfying*

$$F(u_k, u_k) \rightarrow c, \quad (5.9)$$

has a convergent subsequence.

Proof: Since $\mathcal{H}^1(\mathcal{S})$ is compactly embedded in $\mathcal{L}^2(\mathcal{S})$, there exists a subsequence of (u_k) which we will also denote by (u_k) , that converges in $\mathcal{L}^2(\mathcal{S})$. Let us denote this limit by $u \in \mathcal{L}^2(\mathcal{S})$. To show that $u \in \mathcal{H}^1(\mathcal{S})$ we only need to show that its weak derivative is in $\mathcal{L}^2(\mathcal{S})$. Notice that

$$\lim_{k \rightarrow \infty} |(u, \psi') - (u_k, \psi')| \leq \lim_{k \rightarrow \infty} \int_{\mathcal{S}} \|u - u_k\| \|\psi'\| dt = 0. \quad (5.10)$$

Thus for every $\varphi \in \mathcal{C}_0^\infty(\mathcal{S}) \subset \mathcal{D}(F)$, taking the limit and integrating u_k by parts gives,

$$\partial u(\varphi) = -(u, \varphi') = - \lim_{k \rightarrow \infty} (u_k, \varphi') = \lim_{k \rightarrow \infty} (u'_k, \varphi). \quad (5.11)$$

Since $\mathcal{C}_0^\infty(\mathcal{S})$ is dense in $\mathcal{L}^2(\mathcal{S})$, we have that for each u_k there exists a sequence $(u_{k_m}) \subset \mathcal{C}_0^\infty(\mathcal{S})$ convergent to u_k in $\mathcal{L}^2(\mathcal{S})$ as $m \rightarrow \infty$. Taking the cross sequence (u_{k_k}) we have that $\limsup F(u_{k_k}, u_{k_k}) = c$ and the boundary term in $F(u_{k_k}, u_{k_k})$ is zero, thus

$$\limsup \|u'_{k_k}\| \leq (c + \text{esssup}|q|)^{\frac{1}{2}}. \quad (5.12)$$

Furthermore

$$\limsup \|u'_k\| \leq \limsup \|u'_k - u'_{k_k}\| + \limsup \|u'_{k_k}\| \quad (5.13)$$

$$\leq [c + \text{esssup}|q|]^{\frac{1}{2}}. \quad (5.14)$$

Combining (5.11) and (5.14) gives,

$$|\partial u(\varphi)| \leq \limsup \|u'_k\| \|\varphi\| \leq [c + \text{esssup}|q|]^{1/2} \|\varphi\|.$$

Thus $\partial u(\cdot)$ is bounded on $\mathcal{C}_0^\infty(\mathcal{S})$, hence it can be extended to a continuous linear functional on $\mathcal{L}^2(\mathcal{S})$. By the Riesz representation theorem, there exists an $h \in \mathcal{L}^2(\mathcal{S})$ such that,

$$\partial u(\varphi) = \int_{\mathcal{S}} h \varphi dx \quad \text{for } \varphi \in \mathcal{C}_0^\infty(\mathcal{S})$$

this shows that $u \in \mathcal{H}^1(\mathcal{S})$, furthermore by Lemma 5.6 we see that $u \in AC$. We may now integrate by parts and obtain

$$(u'_k, \psi) = -(u_k, \psi') \rightarrow -(u, \psi') = (u', \psi) \text{ for all } \psi \in \mathcal{C}_0^\infty(\mathcal{S}),$$

thus u'_k converges to u' in $\mathcal{L}^2(\mathcal{S})$. ■

The next theorem gives a characterization of the eigenvalues for the boundary value problem (3.1) in terms of the max-min property. Assume the eigenvalues λ_n are labeled in increasing order and repeated according to multiplicity. The eigenfunctions y_n form a complete orthonormal basis in $\mathcal{L}^2(\mathcal{S})$. We will denote $\{v_0, \dots, v_{n-1}\}^\perp$ to be the orthogonal complement in $\mathcal{L}^2(\mathcal{S})$ of $\text{span}\{v_0, \dots, v_{n-1}\}$.

Definition 5.13 *The Gateaux derivative of $J \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ at the point $x \in \mathcal{H}$ defines a linear functional $J'(x)$ in the dual space of $\mathcal{H}$ defined by,*

$$\langle J'(x), \varphi \rangle := dJ(x, \varphi) = \left. \frac{d}{d\epsilon} J(x + \epsilon\varphi) \right|_{\epsilon=0}, \quad (5.15)$$

for all $\varphi \in \mathcal{H}$.

Theorem 5.14 *For $v_j \in \mathcal{L}^2(\mathcal{S})$ such that $j = 0, 1, 2, \dots$, let*

$$d_n(v_0, \dots, v_{n-1}) = \inf \left\{ \frac{F(\varphi, \varphi)}{\|\varphi\|^2} \mid \varphi \in \{v_0, \dots, v_{n-1}\}^\perp \cap D(F) \setminus \{0\} \right\}, \quad (5.16)$$

then

$$\lambda_n = \sup \{d_n(v_0, \dots, v_{n-1}) \mid v_0, \dots, v_{n-1} \in \mathcal{L}^2(\mathcal{S})\} \quad \text{for } n = 0, 1, \dots$$

The max-min is attained for $\varphi = y_n$ and $v_i = y_i$ for $i = 1, \dots, n-1$.

Proof: Let $v_0, \dots, v_{n-1} \in \mathcal{L}^2(\mathcal{S})$. Notice that $\text{span}\{y_0, \dots, y_n\}$ is $n+1$ dimensional and that $\text{span}\{v_0, \dots, v_{n-1}\}$ is at most n dimensional. Therefore there exists a $\varphi \in \text{span}\{y_0, \dots, y_n\} \setminus \{0\}$ with $\varphi \in \{v_0, \dots, v_{n-1}\}^\perp$. That is,

$$(\varphi, v_i) = 0 \quad \text{for } i = 0, 1, \dots, n-1. \quad (5.17)$$

Since each $y_i \in D(F)$ it follows that $\varphi \in D(F)$, and we may denote $\varphi = \sum_{i=0}^n c_i y_i$, thus

$$F(\varphi, \varphi) = \sum_{i,j=0}^n c_i \overline{c_j} F(y_i, y_j) = \sum_{i,k=0}^n c_i \overline{c_k} \lambda_i \delta_{i,k} = \sum_{i=0}^n |c_i|^2 \lambda_i \leq \lambda_n \sum_{i=0}^n |c_i|^2 = \lambda_n \|\varphi\|^2. \quad (5.18)$$

Considering the above equation and (5.16) we see that

$$d_n(v_0, \dots, v_{n-1}) \leq \lambda_n \text{ for all } v_0, \dots, v_{n-1} \in \mathcal{L}^2(\mathcal{S}).$$

For simplicity we denote $m := \sup\{d_n(v_0, \dots, v_{n-1}) | v_0, \dots, v_{n-1} \in \mathcal{L}^2(\mathcal{S})\}$, therefore the above becomes $m \leq \lambda_n$. We will show by induction that there exists a $\varphi \in D(F)$ with $\|\varphi\| = 1$ and $(\varphi, v_i) = 0$ for all $i = 0, 1, \dots, n-1$ such that $F(\varphi, \varphi) = d_n(v_0, \dots, v_{n-1})$. If $n = 0$ then m does not depend on any v_i and we have $m = d_0$. Since F is bounded the infimum exists and by Lemma 5.12 this infimum is attained in $\{v_0, \dots, v_{n-1}\}^\perp \cap D(F) \setminus \{0\}$. Inequality (5.18) shows that d_0 is an eigenvalue less than or equal to the least eigenvalue, thus we must have $m = d_0 = \lambda_0$ where $\varphi = y_0$. Supposing all eigenvalues are characterized by the above method up to and including λ_{n-1} with corresponding eigenfunction y_{n-1} . We will show that the above method holds true for the next eigenvalue. From the definition of $d_n(y_0, \dots, y_{n-1})$ as an infimum and since F is continuous on the subspace $D(F) \cap \{y_0, \dots, y_{n-1}\}^\perp \setminus \{0\}$, there exists a sequence $(u_k) \subset D(F)$ with $\|u_k\| = 1$ and $(u_k, y_i) = 0$ for all $i = 0, 1, \dots, n-1$ and $k \in \mathbb{N}$ such that

$$F(u_k) \rightarrow d_n(y_0, \dots, y_{n-1}) \quad \text{and} \quad \left\| \left(\frac{F(u_k)}{\|u_k\|^2} \right)' \right\| \rightarrow 0. \quad (5.19)$$

Lemma 5.12 shows that there exists a subsequence also denoted by (u_k) converging to a $u \in D(F) \cap \{y_0, \dots, y_{n-1}\}^\perp \setminus \{0\}$ such that $F(u) = d_n(y_0, \dots, y_{n-1})$ and $\|u\| = 1$. We now show that u is an eigenfunction of (3.1) satisfying boundary conditions (3.2)-(3.3) with eigenvalue $\lambda_n = d_n(y_0, \dots, y_{n-1})$.

The second statement of (5.19) gives that the Gâteaux derivative is zero in all directions at the point u , thus

$$\left\langle \left(\frac{F(u)}{\|u\|^2} \right)', \varphi \right\rangle = 0, \quad \text{for } \varphi \in \mathcal{C}_0^\infty(\mathcal{S}) \cap \{v_0, \dots, v_{n-1}\}^\perp. \quad (5.20)$$

A simple calculation shows that

$$\langle F'(u), \varphi \rangle = 2\Re F(\varphi, u) \text{ and } \langle (\|u\|^2)', \varphi \rangle = 2\Re(\varphi, u).$$

It is clear that (5.20) becomes

$$2\Re\{F(\varphi, u) - d_n(v_0, \dots, v_{n-1})(\varphi, u)\} = 0 \quad \text{for all } \varphi \in \mathcal{C}_0^\infty(\mathcal{S}) \cap \{v_0, \dots, v_{n-1}\}^\perp. \quad (5.21)$$

Consider $\varphi \in \{v_0, \dots, v_{n-1}\}$, we may write $\varphi = \sum_{i=0}^{n-1} c_i y_i$, where the c_i 's are appropriately chosen, then the right side of (5.21) gives

$$\sum_{i=0}^{n-1} c_i [\lambda_i(y_i, u) - d_n(y_1, \dots, y_{n-1})(y_i, u)] = 0. \quad (5.22)$$

Thus we have

$$2\Re\{F(\varphi, u) - d_n(v_0, \dots, v_{n-1})(\varphi, u)\} = 0 \quad \text{for } \varphi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.23)$$

Setting $z(\psi) := F(\psi, u) - d_n(v_0, \dots, v_{n-1})(\psi, u)$ for $\psi \in \mathcal{C}_0^\infty(\mathcal{S})$ and $\alpha = e^{-i \arg(z)}$. From (5.21) we have,

$$\Re\{z(\alpha\psi)\} = \Re\{\alpha z(\psi)\} = |z(\psi)| = 0 \quad \text{for } \psi \in \mathcal{C}_0^\infty(\mathcal{S}).$$

Thus u is a variational eigenfunction with variational eigenvalue $\lambda = d_n(v_0, \dots, v_{n-1})$, hence lemma 5.11 shows that $u \in \mathcal{H}^2(\mathcal{S})$ and $u = y_n$ is an eigenfunction of (3.1) satisfying boundary conditions (3.2)-(3.3) with eigenvalue $\lambda_n = d_n(v_0, \dots, v_{n-1})$. ■

5.4 Lusternik-Schnirelmann method

Lusternik-Schnirelmann theory yields a topological characterization of the eigenvalues of the differential equation (3.1) with boundary conditions (3.2)-(3.3). Furthermore the Lusternik-schnirelmann theorems which yield the general Min-max principle relies heavily on the Quantitative deformation Lemma. This lemma is proven on a compact manifold of class $\mathcal{C}^2$ and requires the existence of a so called tangent pseudogradient vector field. In order to obtain such a characterization the notion of the Lusternik-schnirelmann catagory must first be made clear. This section is a reworking of the results obtained in Drabek, see [6].

5.4.1 Lusternik-Schnirelmann Category

Definition 5.15 $\mathcal{A}$ is said be **contractable** to $\mathcal{B}$ denoted, $\mathcal{A} \prec \mathcal{B}$, if $\mathcal{A}$ and $\mathcal{B}$ are closed sets contained in a topological space $\mathcal{V}$ and there exists a homotopy $h \in \mathcal{C}([0, 1] \times \mathcal{V}, \mathcal{V})$ with the property that $h(0, u) = u$ for all $u \in \mathcal{A}$ and $h(1, \mathcal{A}) \subseteq \mathcal{B}$.

Proposition 5.16 (*Transitivity*) Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be closed sets contained in a topological space $\mathcal{V}$. If $\mathcal{A} \prec \mathcal{B}$ and $\mathcal{B} \prec \mathcal{C}$ then $\mathcal{A} \prec \mathcal{C}$.

Proof: Let $h, g \in \mathcal{C}([0, 1] \times \mathcal{V}, \mathcal{V})$ be the homotopy maps contracting $\mathcal{A}$ into $\mathcal{B}$ and $\mathcal{B}$ into $\mathcal{C}$ respectively. Then define the mapping,

$$f(t, u) = \begin{cases} h(2t, u) & \text{if } 0 \leq t \leq \frac{1}{2} \\ g(2t - 1, h(1, u)) & \text{if } \frac{1}{2} < t \leq 1 \end{cases}$$

thus $f \in \mathcal{C}([0, 1] \times \mathcal{V}, \mathcal{V})$ is the homotopy contracting $\mathcal{A}$ into $\mathcal{C}$ and the result follows.

■

Definition 5.17 The **Lusternik-Schnirelmann category** denoted, $cat_{\mathcal{V}}(\mathcal{A})$, of a closed non-empty subset $\mathcal{A}$ of a topological space $\mathcal{V}$ is the least number of closed sets in $\mathcal{V}$ covering $\mathcal{A}$ each of which contractable to a point in $\mathcal{V}$.

In a nutshell, the Lusternik-Schnirelmann method asserts the following:

A C^1 -functional F on a Compact manifold $\mathcal{V}$ has at least $cat_{\mathcal{V}}(\mathcal{V})$ critical points.

The critical values corresponding to the critical points found by the Lusternik - Schnirelmann method are given by

$$c_k := \inf_{\mathcal{A} \in \mathcal{A}_k} \sup_{u \in \mathcal{A}} F(u) \quad \text{where } \mathcal{A}_k := \{\mathcal{A} \subset \mathcal{V} : \mathcal{A} \text{ closed, } cat_{\mathcal{V}}(\mathcal{A}) \geq k\}. \quad (5.24)$$

It is important to note that the method described above in general does not yield all critical points of F on $\mathcal{V}$. It does however give a lower bound on the number of critical points and gives the values at these points. We now prove some properties of the Lusternik-Schnirelmann category.

Proposition 5.18 *Let $\mathcal{A}, \mathcal{B}$ be closed subsets of $\mathcal{V}$, then we have*

- (i) $\text{cat}_{\mathcal{V}}(\emptyset) = 0$,
- (ii) if $\mathcal{A} \subset \mathcal{B}$ then $\text{cat}_{\mathcal{V}}(\mathcal{A}) \leq \text{cat}_{\mathcal{V}}(\mathcal{B})$,
- (iii) $\text{cat}_{\mathcal{V}}(\mathcal{A} \cup \mathcal{B}) \leq \text{cat}_{\mathcal{V}}(\mathcal{A}) + \text{cat}_{\mathcal{V}}(\mathcal{B})$,
- (iv) if $\mathcal{A} \prec \mathcal{B}$ then $\text{cat}_{\mathcal{V}}(\mathcal{A}) \leq \text{cat}_{\mathcal{V}}(\mathcal{B})$.

Proof: To prove (i) notice that the empty set contains no points and cannot be contracted to a point. For (ii), if $\mathcal{A} \subset \mathcal{B}$ then any cover of $\mathcal{B}$ is also a cover of $\mathcal{A}$, thus $\text{cat}_{\mathcal{V}}(\mathcal{A})$ is bounded above by $\text{cat}_{\mathcal{V}}(\mathcal{B})$. Proving (iii), let $\{\mathcal{A}_i\}_{i=1}^n$ and $\{\mathcal{B}_i\}_{i=1}^m$ be the coverings of $\mathcal{A}$ and $\mathcal{B}$ associated with $n := \text{cat}_{\mathcal{V}}(\mathcal{A})$ and $m := \text{cat}_{\mathcal{V}}(\mathcal{B})$ respectively, then $\mathcal{A} \cup \mathcal{B} \subset \{\cup_{i=1}^n \mathcal{A}_i\} \cup \{\cup_{i=1}^m \mathcal{B}_i\}$, which is at most $\text{cat}_{\mathcal{V}}(\mathcal{A}) + \text{cat}_{\mathcal{V}}(\mathcal{B})$ sets. Let us prove (iv). Suppose $\mathcal{A} \prec \mathcal{B}$, then $\mathcal{A}$ is contractable to $\mathcal{B}$ by homotopy h . Let $\{\mathcal{B}_i\}_{i=1}^m$ be the family of coverings associated with $m := \text{cat}_{\mathcal{V}}(\mathcal{B})$. We define the sets,

$$\mathcal{A}_i = \{u \in \mathcal{A} : h(1, u) \in \mathcal{B}_i\}, \quad \text{for } j = 1, \dots, m. \quad (5.25)$$

This gives,

$$\mathcal{A} = \bigcup_{i=1}^n \mathcal{A}_i \text{ and } \mathcal{A}_i \prec \mathcal{B}_i \text{ for } i = 1, \dots, m. \quad (5.26)$$

Each $\mathcal{A}_i$ is contractable to $\mathcal{B}_i$ which is contractable to a point, thus $\text{cat}_{\mathcal{V}}(\mathcal{A}) \leq m$. ■

Definition 5.19 *A metric space $\mathcal{V}$ is an **absolute neighborhood extensor** if for every metric space E , every closed subset $\mathcal{F} \subset E$ and every continuous mapping $f : \mathcal{F} \rightarrow \mathcal{V}$ there exists a continuous extension of f on a neighborhood of $\mathcal{F}$ in E .¹*

Proposition 5.20 *Let $\mathcal{A}$ be a closed set of an absolute neighborhood extensor $\mathcal{V}$, then there exists a closed neighborhood $\mathcal{B}$ of $\mathcal{A}$ in $\mathcal{V}$ such that*

$$\text{cat}_{\mathcal{V}}(\mathcal{A}) = \text{cat}_{\mathcal{V}}(\mathcal{B}). \quad (5.27)$$

¹This definition varies in the literature, the object is also known as an “absolute neighborhood retract” in Deimling [5] and Dugundji [8].

Proof: Let us assume that $\text{cat}_{\mathcal{V}}(\mathcal{A}) = 1$, and let h be the homotopy that contracts $\mathcal{A}$ to a point in $\mathcal{V}$. We define the following set

$$\mathcal{N} := ([0, 1] \times \mathcal{A}) \cup (\{0, 1\} \times \mathcal{V}) \quad \text{is closed in} \quad M := [0, 1] \times \mathcal{V}. \quad (5.28)$$

Let $p \in \mathcal{A}$ be fixed and define the continuous map $f : \mathcal{N} \rightarrow \mathcal{V}$ by

$$f(t, u) = \begin{cases} h(t, u) & \text{if } t \in [0, 1], \quad u \in \mathcal{A}, \\ u & \text{if } t = 0, \quad u \in \mathcal{V} \setminus \mathcal{A}, \\ h(1, p) & \text{if } t = 1, \quad u \in \mathcal{V} \setminus \mathcal{A}. \end{cases}$$

Since $\mathcal{V}$ is an absolute neighborhood extensor there is a continuous extension g of f defined on a neighborhood $\mathcal{U}$ of $\mathcal{N}$ in M . From the compactness of $[0, 1]$ we may assume $\mathcal{U}$ to be of the form $[0, 1] \times \mathcal{C} = \mathcal{U}$ where $\mathcal{C}$ is a neighborhood of $\mathcal{A}$. Consider the closed sets $\mathcal{A}$ and $\overline{\mathcal{V} \setminus \mathcal{C}}$. Since $\mathcal{V}$ is a metric space, it is also a normal Hausdorff space thus there exists open disjoint sets $\mathcal{D}$ and $\mathcal{E}$ such that $\mathcal{A} \subset \mathcal{D}$ and $\overline{\mathcal{V} \setminus \mathcal{C}} \subset \mathcal{E}$. Set $\mathcal{B} := \mathcal{V} \setminus \mathcal{E}$, then $\mathcal{B}$ is closed and $\mathcal{A} \subset \mathcal{B} \subset \text{int}(\mathcal{C})$. By definition of g as a continuous extension of f , $\mathcal{B}$ is contractable to a point, thus $\text{cat}_{\mathcal{V}}(\mathcal{B}) = 1$. Now suppose that $\{\mathcal{A}_i\}_{i=1}^n$ is the covering associated with $n := \text{cat}_{\mathcal{V}}(\mathcal{A})$, then applying the above procedure we obtain closed neighborhoods $\mathcal{B}_i$ of each $\mathcal{A}_i$. Define $\mathcal{B} := \cup_{i=1}^n \mathcal{B}_i$, hence $\mathcal{A} \subset \mathcal{B}$. Using Proposition 5.18 (ii) and (iii) we have,

$$\text{cat}_{\mathcal{V}}(\mathcal{A}) \leq \text{cat}_{\mathcal{V}}(\mathcal{B}) \leq \sum_{i=1}^n \text{cat}_{\mathcal{V}}(\mathcal{B}_i) = n, \quad (5.29)$$

thus $\text{cat}_{\mathcal{V}}(\mathcal{B}) = n$. ■

5.4.2 The projective Space

Definition 5.21 The $(N - 1)$ -dimensional **unit sphere** $\mathcal{S}^{N-1} \subset \mathbb{R}^N$ is defined as

$$\mathcal{S}^{N-1} = \{x \in \mathbb{R}^N : \|x\| = 1\}. \quad (5.30)$$

We define the projective plane by identifying antipodal points in terms of an equivalence relation, that is

$$\mathcal{A} \in \mathcal{P}^{N-1} \quad \text{if and only if} \quad \mathcal{A} = \tilde{\mathcal{A}} \cup (-\tilde{\mathcal{A}}) \quad \text{for some} \quad \tilde{\mathcal{A}} \in \mathcal{S}^{N-1},$$

where $-\tilde{\mathcal{A}}$ is the set of antipodal points of $\tilde{\mathcal{A}}$.

Definition 5.22 *The $(N - 1)$ -dimensional **projective space** is defined as*

$$\mathcal{P}^{N-1} = \{\{x, -x\} : x \in \mathcal{S}^{N-1}\}. \quad (5.31)$$

*The **infinite dimensional projective space** defined on an infinite dimensional Hilbert space $\mathcal{H}$ is given by*

$$\mathcal{P}^\infty = \{\{x, -x\} : x \in \partial B(0, 1) \subset \mathcal{H}\}, \quad (5.32)$$

Where $\partial B(0, 1)$ is the boundary of a unit ball centered at 0.

We state the following lemmas without proof from Drabek, [6, Corollary 4.3.132].

Lemma 5.23 (Borsuk-Ulam). *Let $f : \mathcal{S}^{N-1} \rightarrow \mathbb{R}^{N-1}$ be a continuous mapping, then there is a point $x_0 \in \mathcal{S}^{N-1}$ such that*

$$f(x_0) = f(-x_0). \quad (5.33)$$

Lemma 5.24 (Lusternik-Schnirelmann) *Let $\mathcal{A}_1, \dots, \mathcal{A}_N$ be a closed covering of $\mathcal{S}^{N-1}$, then at least one $\mathcal{A}_i$ contains a set of antipodal points.*

Proof: Suppose that

$$\mathcal{A}_i \cap (-\mathcal{A}_i) = \emptyset \quad \text{for} \quad i = 1, \dots, (N - 1). \quad (5.34)$$

From the Tietze Theorem we can define continuous functions $f^i : \mathcal{S}^{N-1} \rightarrow [0, 1]$ such that

$$f^i(\mathcal{A}_i) = \{0\} \quad \text{and} \quad f^i(-\mathcal{A}_i) = \{1\} \quad \text{for} \quad i = 1, \dots, (N - 1). \quad (5.35)$$

We set $(f^1, \dots, f^{N-1}) = f$ and apply the Borsuk-Ulam theorem, thus there exists $x_0 \in \mathcal{S}^{N-1}$ such that $f(x_0) = f(-x_0)$. However this implies that $f^i(x_0) = f^i(-x_0)$ for $i = 1, \dots, N-1$. Therefore $x_0 \in \mathcal{A}_i$ or $x_0 \in (-\mathcal{A}_i)$ for $i = 1, \dots, N-1$, which is not possible by (5.35). Thus the only possibility is that $x_0 \in (-\mathcal{A}_N) \cap \mathcal{A}_N$. ■

We make use of the following identity.

Proposition 5.25 *For the projective plane defined above, the following holds,*

$$\text{cat}_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1}) = N. \quad (5.36)$$

Proof: We first show that $\text{cat}_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1}) \leq N$ inductively. Firstly it is clear $\mathcal{S}^0 = \{1, -1\}$, so $\mathcal{P}^0 = \{\{1, -1\}\}$ thus $\text{cat}_{\mathcal{P}^0}(\mathcal{P}^0) = 1$. Considering $\mathcal{P}^{M-1}$, we note that a cross section of the equator of $\mathcal{S}^{M-1}$ is homeomorphic to $\mathcal{S}^{M-2}$. Hence this cross section can be covered by $\text{cat}_{\mathcal{P}^{M-2}}(\mathcal{P}^{M-2})$ sets. To completely cover $\mathcal{S}^{M-1}$ all that is needed is to cap the north and south poles by one symmetric closed set, thus we have,

$$\text{cat}_{\mathcal{P}^{M-1}}(\mathcal{P}^{M-1}) \leq \text{cat}_{\mathcal{P}^{M-2}}(\mathcal{P}^{M-2}) + 1 \leq (M-1) + 1 = M.$$

All that is needed is $\text{cat}_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1}) \geq N$, the proof of this uses homological algebra, see [20]. ■

Corollary 5.26 *For the projective space defined above, with $0 \leq N-1 \leq M$, the following holds*

$$\text{cat}_{\mathcal{P}^M}(\mathcal{P}^{N-1}) = N.$$

Proof: We first show that $\text{cat}_{\mathcal{P}^M}(\mathcal{P}^{N-1}) \leq N$, for each $0 \leq N-1 \leq M$, $\mathcal{P}^{N-1}$ is a subspace of $\mathcal{P}^M$ which can be covered by $\text{cat}_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1})$ sets, thus

$$\text{cat}_{\mathcal{P}^M}(\mathcal{P}^{N-1}) \leq \text{cat}_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1}) \leq N.$$

To show $\text{cat}_{\mathcal{P}^M}(\mathcal{P}^{N-1}) \geq N$, suppose that $\mathcal{P}^{N-1}$ can be covered by $N-K$ sets in $\mathcal{P}^M$ contractable to a point, where $0 < K < N$. Then these $N-K$ sets are homeomorphic

to $N - K$ convex closed sets which cover $\mathcal{P}^{N-1}$. The intersection of these convex sets with $\mathcal{P}^{N-1}$ covers $\mathcal{P}^{N-1}$ by $N - K$ sets, thus

$$cat_{\mathcal{P}^{N-1}}(\mathcal{P}^{N-1}) \leq N - K,$$

which contradicts proposition 5.25. ■

The above shows an important results that will be used in the sequel, namely

$$cat_{\mathcal{P}^\infty}(\mathcal{P}^\infty) = \infty,$$

$$cat_{\mathcal{P}^\infty}(\mathcal{P}^{N-1}) = N.$$

5.4.3 Lusternik-Schnirelmann Theorems

We now outline the manifold structure in Banach spaces and the existence of a tangent pseudogradient vector field thereon. Following this, using the Lusternik-schnirelmann category we can state and prove the Quantitative Deformation Lemma.

Definition 5.27 ² Let $\mathcal{H}$ be a real separable Banach space. Define $\mathcal{V}$ and $\psi \in \mathcal{C}^2(\mathcal{H}, \mathbb{R})$ such that

$$\mathcal{V} = \{v \in \mathcal{H} : \psi(v) = 1\} \neq \emptyset \quad \text{and} \quad \psi'(v) \neq 0 \quad \text{for all} \quad v \in \mathcal{V}, \quad (5.37)$$

then $\mathcal{V}$ is a **Differentiable Manifold of class $\mathcal{C}^2$** .

The norm on $\mathcal{H}$ induces a metric on $\mathcal{V}$ making it a metric space and thus a metric manifold. It has also been shown that $\mathcal{V}$ is an absolute neighborhood extensor.³ The Manifold $\mathcal{V}$ was obtained under the assumption of separability of $\mathcal{H}$, however this condition can be relaxed by assuming second countability and hence paracompactness, making $\mathcal{V}$ of class $\mathcal{C}^1$. Similar results as those presented in this section have been proved by Ghossoub, see [13].

²cf. Deimling[5] , Zeidler[46, Chapter 43]

³cf. Deimling[5, Prop. 27.6]

Definition 5.28 The **tangent space** at $v \in \mathcal{V}$ is denoted by $T_v\mathcal{V}$ and defined as

$$T_v\mathcal{V} := \{y \in \mathcal{H} : \langle \psi'(v), y \rangle = 0\}. \quad (5.38)$$

Definition 5.29 Let $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$. The **norm of the restriction of the derivative** $F'(v)$ to $T_v\mathcal{V}$ is defined as

$$\|F'(v)\|_* := \sup_{\substack{y \in T_v\mathcal{V} \\ \|y\|_X = 1}} \langle F'(v), y \rangle. \quad (5.39)$$

The point v is a **critical point of the restriction of F to $\mathcal{V}$** if the restriction of $F'(v)$ to $T_v\mathcal{V}$ is equal to the zero operator.

The norm (5.39) can simply be described as the operator norm restricted to $T_v\mathcal{V}$. This gives the maximum gradient between any two points in $T_v\mathcal{V}$. We will use the notation

$$F^d := \{v \in \mathcal{V} : F(v) \leq d\}. \quad (5.40)$$

Lemma 5.30 (Duality Lemma).⁴ If $f, g \in \mathcal{H}^*$, then

$$\sup_{\substack{\langle g, y \rangle = 0 \\ \|y\|_X = 1}} \langle f, y \rangle = \min_{\lambda \in \mathbb{R}} \|f - \lambda g\|.$$

It is clear from the Duality Lemma and the norm of the restriction of F' to the tangent bundle that x is a critical point of $F|_{\mathcal{V}}$ if and only if there exists a $\lambda \in \mathbb{R}$ such that

$$F'(u) = \lambda \psi'(u). \quad (5.41)$$

Definition 5.31 Let $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$. A **tangent pseudogradient vector field** for F on

$$\mathcal{M} := \{u \in \mathcal{V} : \|F'(u)\|_* \neq 0\} \quad (5.42)$$

is a locally Lipschitz continuous vector field $g : \mathcal{M} \rightarrow \mathcal{H}$ such that $g(u) \in T_u\mathcal{V}$ and

$$\|g(u)\|_{\mathcal{H}} \leq 2\|F'(u)\|_* \text{ and } \langle F'(u), g(u) \rangle \geq \frac{2}{3}\|F'(u)\|_* \text{ for every } u \in \mathcal{M}. \quad (5.43)$$

⁴Drabek [6, Proposition 6.4.35]

Proposition 5.32 *Let $f \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ and $f(v) > 0$ for some $v \in \mathcal{H}$, then there is a neighborhood of v denoted $\mathcal{N}(v)$ such that*

$$f(v) > 0 \quad \text{for all} \quad u \in \mathcal{N}(v). \quad (5.44)$$

Proof: Suppose the conclusion is false, then for every neighborhood $\mathcal{N}(v)$ of v there exists a $u \in \mathcal{N}(v)$ such that $f(u) \leq 0$. Let $\epsilon := f(v)$, then from continuity there exists a neighborhood $\mathcal{U}(v)$ of v such that,

$$\|f(v) - f(u)\| < \epsilon \quad \text{for every} \quad u \in \mathcal{U}(v), \quad (5.45)$$

but then there is a $\tilde{u} \in \mathcal{U}(v)$ such that $f(\tilde{u}) \leq 0$ giving,

$$\|f(v) - f(\tilde{u})\| \geq \|f(v)\| = \epsilon. \quad (5.46)$$

This contradiction proves the result. ■

Definition 5.33 *Let $f : X \rightarrow Y$, then f is said to be **locally Lipschitz continuous** if for every $x \in X$ there exists a $K \in \mathbb{R}$ and a neighborhood $\mathcal{U}$ of x such that*

$$\|f(u) - f(v)\|_Y \leq K\|u - v\|_X \quad \text{for every} \quad u, v \in \mathcal{U}.$$

Lemma 5.34 *Let $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$. Then there exists a tangent pseudogradient vector field for F on $\mathcal{M}$.*

Proof: Since $\|F'(v)\|_* \neq 0$ for every $v \in \mathcal{M}$, the definition of the operator norm implies that there exists $x \in T_v\mathcal{V}$ such that

$$\langle F'(v), x \rangle > \frac{2}{3}\|F'(v)\|_* \quad \text{with} \quad \|x\|_{\mathcal{H}} = 1. \quad (5.47)$$

From the definition of $\mathcal{M}$ and linearity of $\psi'(v)$ there exists a $z \in \mathcal{H}$ with $\langle \psi'(v), z \rangle = 1$. Set $y := \frac{3}{2}\|F'(v)\|_*$ and for $u \in \mathcal{V}$ such that $\langle F'(u), z \rangle \neq 0$ define

$$g_v(u) := y - \frac{\langle \psi'(u), y \rangle}{\langle \psi'(u), z \rangle} z. \quad (5.48)$$

Since $x \in T_v \mathcal{V}$ we have $\langle \psi'(v), x \rangle = 0$ giving $\langle \psi'(v), y \rangle = 0$ and $g_v(v) = y$, also

$$\|g_v(v)\|_{\mathcal{H}} < 2\|F'(v)\|_* \quad \text{and} \quad \langle F'(v), g_v(v) \rangle > \|F'(v)\|_*^2. \quad (5.49)$$

Since F' and g_v are continuous, by proposition 5.32, there exists an open neighborhood $\mathcal{N}(v)$ of v such that

$$\|g_v(u)\| < 2\|F'(u)\|_* \quad \text{and} \quad \langle F'(u), g_v(u) \rangle > \|F'(u)\|_*^2 \quad \text{for every} \quad u \in \mathcal{N}(v). \quad (5.50)$$

The family

$$\mathcal{N} := \{\mathcal{M}_i : i \in \mathbb{N}\} \quad (5.51)$$

is a open covering of $\mathcal{M}$. Since $\mathcal{M}$ is a metric space, there exists a locally finite open covering

$$\mathcal{M} = \{\mathcal{M}_i : i \in \mathbb{N}\} \quad (5.52)$$

which is subordinate to $\mathcal{N}$, that is, the closure of every $\mathcal{M}_i$ of $\mathcal{M}$ is contained in the neighborhood $\mathcal{N}(v)$ of at least one $v \in \mathcal{V}$. For every $i \in \mathbb{N}$ chose one such $v := v_i$ and define

$$g_i(u) := \begin{cases} g_{v_i}(u) & \text{if } u \in \mathcal{N}(v_i) \\ 0 & \text{if } u \notin \mathcal{N}(v_i) \end{cases}$$

and

$$\rho_i(u) := \text{dist}(u, \mathcal{H} \setminus \mathcal{M}_i)^5, \quad g(u) := \sum_{i \in \mathbb{N}} \frac{\rho_i(u) g_i(u)}{\sum_{j \in \mathbb{N}} \rho_j(u)} \quad (5.53)$$

Since $\psi \in \mathcal{C}^2(\mathcal{H}, \mathbb{R})$ we have that each $g_{v_i} \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ and are linear thus $g(u)$ is locally Lipschitz continuous on $\mathcal{M}$. Finally, $g(u)$ is the required tangent pseudogradient vector field. ■

For the sequel if $\mathcal{S} \subset \mathcal{H}$ we define,

$$\mathcal{S}_{2\delta} := \{u \in \mathcal{H} : \text{dist}(u, \mathcal{S}) \leq 2\delta\}. \quad (5.54)$$

Lemma 5.35 (*Quantitative Deformation Lemma*). *Let $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$, $\mathcal{S} \subset \mathcal{V}$, $c \in \mathbb{R}$, $\epsilon, \delta > 0$ be such that*

$$\|F'(u)\|_* \geq \frac{8\epsilon}{\delta} \quad \text{for any} \quad u \in F_{-1}([c - 2\epsilon, c + 2\epsilon]) \cap \mathcal{S}_{2\delta} \cap \mathcal{V}. \quad (5.55)$$

⁵Here $\text{dist}(u, \mathcal{H} \setminus \mathcal{M}_i) = \inf\{\|u - b\| : b \in \mathcal{H} \setminus \mathcal{M}_i\}$

⁶The local finiteness property guarantees only a finite number of $\rho'_i s$ are nonzero.

Then there exists a homotopy $\eta \in \mathcal{C}([0, 1] \times \mathcal{V}, \mathcal{V})$ such that

- (i) $\eta(t, u) = u$ if $t = 0$ or $u \notin F_{-1}([c - 2\epsilon, c + 2\epsilon]) \cap \mathcal{S}_{2\delta} \cap \mathcal{V}$,
- (ii) $\eta(1, F^{c+\epsilon} \cap \mathcal{S}) \subset F^{c-\epsilon}$,
- (iii) $F(\eta(\cdot, u))$ is decreasing for any $u \in \mathcal{V}$,
- (iv) $\eta(t, \cdot)$ is a homeomorphism on $\mathcal{H}$, for each fixed t .

Proof: By Lemma 5.34 there exists a tangent pseudogradient vector field g for F on $\mathcal{M}$. Define the following sets

$$\mathcal{A} := F_{-1}([c - 2\epsilon, c + 2\epsilon]) \cap \mathcal{S}_{2\delta} \cap \mathcal{V} \quad \text{and} \quad \mathcal{B} := F_{-1}([c - \epsilon, c + \epsilon]) \cap \mathcal{S}_{2\delta} \cap \mathcal{V},$$

and the functional

$$\phi(u) := \frac{\text{dist}(u, \mathcal{H} \setminus \mathcal{A})}{\text{dist}(u, \mathcal{H} \setminus \mathcal{A}) + \text{dist}(u, \mathcal{B})}. \quad (5.56)$$

We will now show that ϕ is locally Lipschitz continuous. For $x_1, x_2 \in \mathcal{H}$ from a bounded set, without loss of generality we may assume $\text{dist}(x_1, \mathcal{M}) \geq \text{dist}(x_2, \mathcal{M})$, and for any $\epsilon > 0$ there exists a $x_\epsilon \in \mathcal{M}$ such that

$$\|x_2 - x_\epsilon\|_{\mathcal{H}} < \text{dist}(x_2, \mathcal{M}) + \epsilon. \quad (5.57)$$

Applying the triangle inequality to $\|x_1 - x_\epsilon\|_{\mathcal{H}}$ yields

$$\|x_1 - x_\epsilon\|_{\mathcal{H}} < \text{dist}(x_2, \mathcal{H}) + \epsilon + \|x_1 - x_2\|, \quad (5.58)$$

but since $x_\epsilon \in \mathcal{M}$ we have $\|x_1 - x_\epsilon\|_{\mathcal{H}} \geq \text{dist}(x_1, \mathcal{M})$. Hence

$$|\text{dist}(x_1, \mathcal{M}) - \text{dist}(x_2, \mathcal{M})| \leq \|x_1 - x_2\|_{\mathcal{H}}. \quad (5.59)$$

Noticing that

$$\text{dist}(x_i, \mathcal{H} \setminus \mathcal{A}) + \text{dist}(x_i, \mathcal{B}) \geq \beta, \quad i = 1, 2 \quad \text{for some } \beta \geq 0. \quad (5.60)$$

Using the triangle inequality we have from (5.60) and (5.59) that

$$|\phi(x_1) - \phi(x_2)| \leq \frac{1}{\beta} \|x_1 - x_2\|_{\mathcal{H}} \quad \text{for } x_1, x_2 \in \mathcal{M}. \quad (5.61)$$

Thus ϕ is locally Lipschitz continuous on $\mathcal{M}$ with the property

$$\phi(u) := \begin{cases} 1 & \text{on } \mathcal{B} \\ 0 & \text{on } \mathcal{H} \setminus \mathcal{A}. \end{cases} \quad (5.62)$$

Define the following vector field

$$f(u) := \begin{cases} -\phi(u) \frac{g(u)}{\|g(u)\|_{\mathcal{H}}^2} & \text{on } u \in \mathcal{A} \\ 0 & \text{on } u \in \mathcal{H} \setminus \mathcal{A}, \end{cases}$$

where g is a pseudogradient vector field for F on $\mathcal{M}$ as given by Lemma 5.34. By definition g was defined to be continuous and nowhere zero, else $\mathcal{M}_i$ would not be a covering of $\mathcal{M}$, and g is locally Lipschitz continuous. The product of locally Lipschitz functions is locally Lipschitz, thus we see that f is locally Lipschitz continuous. Using definition (5.43) we have

$$\|F'(u)\|_* \|g(u)\|_{\mathcal{H}} \geq \langle F'(u), g(u) \rangle \geq \|F'(u)\|_*^2, \quad (5.63)$$

hence using (5.55), we have

$$\|g(u)\|_{\mathcal{H}} \geq \frac{8\epsilon}{\delta} \quad \text{so} \quad \|f(u)\|_{\mathcal{H}} \leq \frac{\delta}{8\epsilon}. \quad (5.64)$$

We now consider the Cauchy problem $\dot{\sigma} = f(\sigma)$ with initial condition $\sigma(0) = u$. Since f is locally Lipschitz continuous the above problem has a unique solution $\sigma(\cdot, u)$ which is defined on the whole $\mathbb{R}$ and σ is continuous on $\mathbb{R} \times \mathcal{H}$. We now define $\eta : [0, 1] \times \mathcal{H} \rightarrow \mathcal{H}$ by

$$\eta(t, u) := \sigma(8\epsilon t, u). \quad (5.65)$$

From (5.55), (5.64) and the Fundamental Theorem of Calculus we have for $t \geq 0$,

$$\|\sigma(t, u) - u\|_{\mathcal{H}} = \left\| \int_0^t f(\sigma(\tau, u)) d\tau \right\|_{\mathcal{H}} \leq \int_0^t \|f(\sigma(\tau, u))\|_{\mathcal{H}} d\tau \leq \frac{\delta t}{8\epsilon}. \quad (5.66)$$

From the chain rule and (5.49) we have

$$\frac{d}{dt} F(\sigma(t, u)) = \langle F'(\sigma(t, u)), \frac{d}{dt} \sigma(t, u) \rangle = \langle F'(\sigma(t, u)), f(\sigma(t, u)) \rangle \leq -\frac{1}{4} \phi(\sigma(t, u)) \leq 0. \quad (5.67)$$

Thus (iii) holds. We see that (i) holds for $t = 0$ from the initial conditions and also when $u \notin \mathcal{A}$, $f(u) = 0$ so σ is constant and the initial condition yields $\sigma(t, u) = u$

for all $u \in \mathcal{H} \setminus \mathcal{A}$. (iv) is true since $\eta(t, u)$ is continuous and one-to-one, else this would contradict the uniqueness of solution of the Cauchy problem. Hence $\eta(t, u)$ is continuously invertible. All that remains is (ii). Let $u \in \mathcal{F}^{c+\epsilon} \cap \mathcal{S}$, there are two possible cases, the first case, assuming that $F(\sigma(t, u)) < c - \epsilon$ for some $t \in [0, 8\epsilon]$, but since $F(\sigma(t, u))$ is decreasing we have $F(\sigma(t, 8\epsilon)) < c - \epsilon$ and (ii) holds. For the second case assume that

$$\sigma(t, u) \in F_{-1}([c - \epsilon, c + \epsilon]) \quad \text{for any } t \in [0, 8\epsilon]. \quad (5.68)$$

From (5.66) we know $\sigma(t, u) \in \mathcal{B}$, (5.67) yields

$$F(\sigma(8\epsilon, u)) = F(u) + \int_0^{8\epsilon} \frac{d}{dt} F(\sigma(t, u)) dt \leq F(u) - \frac{1}{4} \int_0^{8\epsilon} \phi(\sigma(t, u)) dt. \quad (5.69)$$

Since $F(u) = F(\sigma(0, u)) \leq c - \epsilon$ and $\phi(\sigma(t, u)) = 1$, the above becomes,

$$F(\sigma(8\epsilon, u)) \leq c + \epsilon - 2\epsilon = c - \epsilon, \quad (5.70)$$

satisfying (ii) and the theorem is complete. ■

The preparation is now complete. We proceed with the Lusternik-Schnirelmann theorems.

Theorem 5.36 (General Minimax Principle) *Assume that F and c_j are as in (5.24). If*

$$c := c_k = c_{k+1} = \cdots = c_{k+m}, \quad (5.71)$$

we have for every $\epsilon, \delta > 0$, $\mathcal{A} \in \mathcal{A}_{k+m}$ and $\mathcal{B} \subset \mathcal{V}$ closed with

$$\sup_{u \in \mathcal{A}} F(u) \leq c + \epsilon \quad \text{and} \quad \text{cat}_{\mathcal{V}}(\mathcal{B}) \leq m, \quad (5.72)$$

then there exists a $u_0 \in \mathcal{V}$ such that

$$(i) \quad c - 2\epsilon \leq F(u_0) \leq c + 2\epsilon,$$

$$(ii) \quad \text{dist}(u_0, \mathcal{A} \setminus \text{int}(\mathcal{B})) \leq 2\delta,$$

$$(iii) \quad \|F(u_0)\|_* \leq \frac{8\epsilon}{\delta}.$$

Proof: Suppose $\delta, \epsilon > 0$ and that there exist closed sets $\mathcal{A} \in \mathcal{A}_{k+m}$ and $\mathcal{B} \subset \mathcal{V}$ satisfying (5.72). Since the category of $\mathcal{A}$ is greater than that of $\mathcal{B}$, by Proposition 5.18 (ii), $\mathcal{A}$ is not a subset of $\mathcal{B}$, thus (ii) holds. From (5.24) we see that from the definition of c_{k+m} , we have the bound

$$c < \sup_{u \in \mathcal{A}} F(u) \leq c + \epsilon, \quad (5.73)$$

thus the subset of $\mathcal{V}$ satisfying (i) and (ii) is nonempty. We prove (iii), by contradiction assuming (i) and (ii) hold while (iii) does not. From (5.72) using the Quantitative Deformation Lemma, we have

$$\mathcal{A} \setminus \text{int} \mathcal{B} \prec F^{c-\epsilon} \quad (5.74)$$

From lemma 5.18 (iii) and (iv), noticing that $\mathcal{A} = \mathcal{A} \setminus \text{int}(\mathcal{B}) \cup \mathcal{B}$ giving

$$k + m \leq \text{cat}_{\mathcal{V}}(\mathcal{A} \setminus \text{int}(\mathcal{B})) + \text{cat}_{\mathcal{V}}(\mathcal{B}) \leq \text{cat}_{\mathcal{V}}(F^{c-\epsilon}) + m. \quad (5.75)$$

Notice that

$$\text{cat}_{\mathcal{V}}(F^{c-\epsilon}) \leq k - 1, \quad (5.76)$$

otherwise supposing $\text{cat}_{\mathcal{V}}(F^{c-\epsilon}) \geq k$ then $F^{c-\epsilon} \in \mathcal{A}_k$ but then we have

$$\inf_{\mathcal{A} \in \mathcal{A}_k} \sup_{u \in \mathcal{A}} F(u) \leq \sup_{u \in F^{c-\epsilon}} F(u) = c - \epsilon,$$

a contradiction. Thus using (5.75) and (5.76) we obtain

$$k + m \leq \text{cat}_{\mathcal{V}}(F^{c-\epsilon}) + m \leq k + m - 1.$$

This contradiction proves the result. ■

Definition 5.37 A functional F satisfies the **Palais-Smale Condition**, $(PS)_c$, at level c on $\mathcal{V}$ if any sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{V}$ satisfying

$$F(x_n) \rightarrow c, \quad \|F'(x_n)\|_* \rightarrow 0, \quad (5.77)$$

has a convergent subsequence.

This condition on a functional is important because you could have the situation for a sequence $\{x_n\}_{n=1}^\infty$ satisfying

$$F(x_n) \rightarrow c, \quad \|F'(x_n)\|_* \rightarrow 0 \quad (5.78)$$

but the sequence is unbounded, in this case the conditions on the manifold for a critical point are satisfied however the critical point does not exist. An example of such a functional given by Brezis and Nirenberg.⁷

Theorem 5.38 *Let F be bounded below on $\mathcal{V}$ and satisfy $(PS)_c$ on $\mathcal{V}$. Let*

$$\mathcal{K}_c := \{u \in \mathcal{V} : F(u) = c, \|F'(u)\|_* = 0\},$$

denote the critical points at level c , as in (5.71), then

$$\text{cat}_{\mathcal{V}}(\mathcal{K}_c) \geq m + 1.$$

Proof: Assume that $\text{cat}_{\mathcal{V}}(\mathcal{K}_c) \leq m$, since $\mathcal{V}$ is an absolute neighborhood extensor there exists a closed $\mathcal{B}$ such that $\text{cat}_{\mathcal{V}}(\mathcal{B}) \leq m$ and $\mathcal{K}_c \subset \text{int}(\mathcal{B})$. By the General Minimax Principle setting $\mathcal{A} = \mathcal{V}$ there exists a sequence $\{x_n\}_{n=1}^\infty \subset \mathcal{V}$ such that

$$F(x_n) \rightarrow c, \quad \text{dist}(x_n, \mathcal{V} \setminus \text{int}(\mathcal{B})) \rightarrow 0, \quad \|F'(x_n)\|_* \rightarrow 0.$$

But F satisfies $(PS)_c$ on $\mathcal{V}$ thus a subsequence of x_n converges to a point in $\mathcal{K}_c$ and in $\mathcal{B} \setminus \text{int}(\mathcal{B})$, contradicting the definition of $\mathcal{B}$ being a closed neighborhood of $\mathcal{A}$. ■

Theorem 5.39 *Let $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ be bounded below on $\mathcal{V}$ let d be an upper bound for $\inf_{u \in \mathcal{V}} F(u)$ and let F satisfy the $(PS)_c$ on the levels $c \in [\inf_{x \in \mathcal{V}} F(u), d]$. Then $F|_{\mathcal{V}}$ has a minimum and F^d contains at least $\text{cat}_{\mathcal{V}}(F^d)$ critical points of $F|_{\mathcal{V}}$.*

Proof: Supposing $n := \text{cat}_{\mathcal{V}}(F^d)$ and that c_i , $i = 1, \dots, n$, are given by (5.24) such that

$$\inf_{u \in \mathcal{V}} F(u) = c_1 \leq c_2 \leq c_3 \leq \dots \leq c_n \leq d. \quad (5.79)$$

⁷cf. Drabek [6, page 428].

If the critical points correspond to different critical values which are not equal then we are done. If this is not the case then we have the scenario as in (5.71) and we may apply theorem 5.38, hence the result follows. ■

Theorem 5.40 *Let $\mathcal{H}$ be a infinite dimensional real Hilbert space and let the functional $F \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ be bounded below, even and satisfying $(PS)_c$ on $\partial B(0, 1) \subset \mathcal{H}$. Then $F|_{\partial B(0,1)}$ has infinitely many distinct pairs of critical points.*

Proof: We define

$$X = \{x = \{u, -u\} : u \in \mathcal{H}\} \quad \mathcal{V} := \{x \in X : \|x\|_{\mathcal{H}}^2 = 1\}. \quad (5.80)$$

We note that $\|x\|_{\mathcal{H}}^2 \in \mathcal{C}^2(\mathcal{H}, \mathbb{R})$ and is even, thus we have that $\mathcal{V} = \mathcal{P}^\infty$ and applying proposition 5.25, we have

$$cat_{\mathcal{V}}(\mathcal{V}) = \infty.$$

Since theorem 5.38 shows that the number of critical points of $F|_{\mathcal{V}}$ is bounded below by $cat_{\mathcal{V}}(\mathcal{V})$, the theorem follows. ■

The final theorem of this section characterizes the critical points of the functional (5.7) related to the Sturm-Liouville variational problem, however the adjustment is made to only consider the real Hilbert spaces. We consider the functional $F(x) := F(x, x)$ defined in (5.7) as

$$F(x, y) = \int_{\partial \mathcal{S}} fxy \, d\sigma + \int_{\mathcal{S}} x'y' + qxy \, dt. \quad (5.81)$$

Here $\mathcal{S} = [a, b]$ with $0 < a < b < \infty$, and

$$X := \{x = \{u, -u\} : u \in \mathcal{H}^1\}, \quad \mathcal{V} := \{x \in X : \|x\|_{\mathcal{L}^2}^2 = 1\}. \quad (5.82)$$

Theorem 5.41 *The functional $F(u)$ defined above on X has infinitely many critical points on $\mathcal{V}$ defined as,*

$$\lambda_i = \inf_{\mathcal{A} \in \mathcal{A}_i} \sup_{x \in \mathcal{A}} F(x)$$

where $\mathcal{A}_i = \{\mathcal{A} \subset \mathcal{V} : \mathcal{A} \text{ is closed, } \text{cat}_{\mathcal{V}}(\mathcal{A}) \geq i\}$.

Proof: Theorem 5.9 shows that F is bounded below, now to show that $(PS)_c$ at level $c \geq \inf_{x \in \mathcal{V}} F(x)$ is satisfied, let

$$F(x_n) \rightarrow c, \quad \|F'(x_n)\|_* \rightarrow 0 \quad \text{for } x_n \in \mathcal{V}. \quad (5.83)$$

From Lemma 5.12, there exists a subsequence also denoted x_n such that $x_n \rightarrow x$ for some $x \in \mathcal{V}$.⁸ We now know from Theorem 5.41 that $\mathcal{V}$ has infinitely many critical points $x_i \in \mathcal{V}$. Using the Gateaux derivative we have

$$\langle F'(x_i), \varphi \rangle = 2F(\varphi, x_i) \text{ and } \langle (\|x_i\|_{\mathcal{L}^2}^2)', \varphi \rangle = 2(\varphi, x_i), \quad \text{for } \varphi \in \mathcal{H}^1.$$

From the Duality Lemma 5.30 we obtain

$$F(x_i, \varphi) = \lambda_i(x_i, \varphi), \quad \varphi \in \mathcal{H}^1. \quad (5.84)$$

We know from Lemma 5.11 that x_i are eigenvectors and λ_i are the eigenvalues of the Sturm-Liouville problem. We show that the above scheme obtains all the variational eigenvalues of F . We see that the eigenvalues λ_i are obtained in increasing order of magnitude, thus the infimum over $\mathcal{H}^1$ obtains the first eigenvalue, λ_1 . Suppose now that λ_k is not obtained by the above scheme and all eigenvalues are obtained using the above method up to and excluding λ_k , for $1 < k < \infty$. Consider the set,

$$\mathcal{A} := \left\{ \sum_{i=1}^k \gamma_i x_i : \sum_{i=1}^k |\gamma_i|^2 = 1 \right\},$$

noticing $F(x) \leq \lambda_k \quad \text{for } x \in \mathcal{A}.$

It is clear that $\mathcal{A}$ is a symmetric set in $\mathcal{V}$, furthermore, $\mathcal{A}$ is homeomorphic to $\mathcal{P}^{k-1}$.

From Corollary 5.26 we have,

$$\text{cat}_{\mathcal{V}}(\mathcal{A}) = n,$$

thus $\mathcal{A} \in \mathcal{A}_k$, and since above scheme does not obtain λ_k we have

$$\inf_{\mathcal{A} \in \mathcal{A}_k} \sup_{x \in \mathcal{A}} F(x) < \sup_{x \in \mathcal{A}} F(x).$$

⁸This x is obtained by normalising this subsequence.

The above shows that there exists an eigenvalue between two consecutive eigenvalues. This contradiction proves the result. ■

This characterization can also be used for non-linear problems such as the p-Laplacian eigenvalue problem considered by Drábek and Robinson [7].

5.5 Courant Nodal Domain Theorem

In this section we will give a variational proof of the Courant Nodal domain theorem, this theorem gives an upper bound for the number of zero's of each eigenfunction. This is can be seen as the variational analogue of the upper bound of the Sturm oscillation theorem. For the next proof we define the following Unique Continuity Property, (UCP), of the Sturm-Liouville problem. In the sequel we refer to F as defined in (5.81) with domain $\mathcal{H}^1$. The results in this section are a reworking of those obtained by Drábek and Robinson, see [7].

Definition 5.42 (UCP). *Let u_{λ_n} be an eigenfunction of F , then u_{λ_n} has no non-empty open interval on which it is equal to zero, that is, the set $\{x : u_{\lambda_n}(x) = 0\}$ has a empty interior.*

Definition 5.43 *A nodal domain of $u \in \mathcal{C}(\mathcal{S})$ is a maximal open connected subset of $\{x : u(x) \neq 0\}$, that is, it is a maximal set whose Lusternik-Schnirelmann category is 1.*

Theorem 5.44 (Courant). *Let u_{λ_n} be an eigenfunction associated with λ_n of F , then u_{λ_n} has at most n nodal domains.*

Proof: Let us assume that u_{λ_n} has $n + k$ nodal domains, with $k \geq 1$. We shall call these nodal domains $\{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{n+k}\}$. Define the following functions $u_i := u_{\lambda_n} \cdot \chi_{\mathcal{S}_i}$,

and $u_i \in \mathcal{H}^1$ since

$$\partial u_i = - \int_S u_{\lambda_n} \cdot \chi_{S_i} \varphi dt = - \int_{S_i} u_{\lambda_n} \varphi dt \in \mathcal{H}^1 \quad \text{for } \varphi \in \mathcal{C}_0^\infty(\mathcal{S}). \quad (5.85)$$

Define

$$\mathcal{A} := \left\{ \sum_{i=1}^n \gamma_i u_i : \sum_{i=1}^n \gamma_i^2 \int_{S_i} u_i^2 dt = 1 \right\},$$

also for $u \in \mathcal{A}$,

$$\begin{aligned} \|u\|^2 &= \int_S u^2 dt \\ &= \sum_{i=1}^n \int_S \gamma_i u_i^2 dt \\ &= \sum_{i=1}^n \gamma_i^2 \int_{S_i} u_{\lambda_n}^2 dt = 1, \end{aligned}$$

hence we see that $\mathcal{A} \subset \mathcal{V}$ and is symmetric thus homeomorphic to $\mathcal{P}^{n-1}$. Corollary 5.26 shows that $\text{cat}_{\mathcal{V}}(\mathcal{A}) = n$, thus $\mathcal{A} \in \mathcal{A}_n$. Now for $u \in \mathcal{A}$, integrating by parts we have,

$$\begin{aligned} F(u) &= \sum_{i=1}^n \left[\int_{\partial S_i} u^2 d\sigma + \int_{S_i} (u')^2 + qu^2 dt \right] \\ &= \sum_{i=1}^n \gamma_i^2 \left[\int_{\partial S_i} u_i^2 d\sigma + \int_{S_i} [(u'_i)^2 + qu_i^2] dt \right] \\ &= \sum_{i=1}^n \gamma_i^2 \left[-f(a)u(a)_{\lambda_n}^2 + \int_{S_i} (u'_{\lambda_n})^2 + qu_{\lambda_n}^2 dt \right] \\ &= \sum_{i=1}^n \gamma_i^2 \left[-f(a)u(a)_{\lambda_n}^2 + f(a)u(a)_{\lambda_n}^2 + \int_{S_i} (-u''_{\lambda_n} + qu_{\lambda_n})u dt \right] \\ &= \lambda_n \sum_{i=1}^n \gamma_i^2 \int_{S_i} u_{\lambda_n}^2 dt = \lambda_n. \end{aligned}$$

Thus $F \equiv \lambda_n$ on $\mathcal{A}$. However for $u \in \mathcal{A}$, $u \equiv 0$ on $\mathcal{S}_{n+1}$, therefore no eigenfunction is contained in $\mathcal{A}$, as (UCP) would be contradicted. Thus F has no critical points on $\mathcal{A}$. Since $\mathcal{A}$ is compact, there exists numbers $\delta, \epsilon > 0$ such that,

$$\|F'(u)\|_* \geq \frac{8\epsilon}{\delta} \quad \text{for } u \in \mathcal{A}.$$

From the Quantitative Deformation Lemma 5.35, there exists a homeomorphism $\eta(t, u)$ as given in the statement. We can now define a set $\tilde{\mathcal{A}} := \eta(t, \mathcal{A})$, with the property $\tilde{\mathcal{A}} \in \mathcal{A}_n$ and

$$\sup_{u \in \tilde{\mathcal{A}}} F(u) < \lambda_n.$$

This contradiction proves the result. $\blacksquare$

Chapter 6

Classes of Self-Adjoint operators

6.1 Introduction

Marcus and Everitt, see [10, 30], used methods of symplectic geometry to yield a theory which can be used to study which boundary conditions lead to self-adjoint domains of the formal operator L defined in (1.4). This novel approach uses the Glazman-Krein- Naimark theorem (GKN) which shows that there is a bijection between self-adjoint boundary conditions and complete Lagrangian subspaces of the operator L . Using this and a prescribed symplectic product on a symplectic space S , the problem of finding self-adjoint boundary conditions can be converted to an entirely algebraic problem of finding the complete Lagrangian subspaces of S . Furthermore, a symplectic algebra gives a criterion for the existence of these self-adjoint boundary conditions, namely, the balanced intersection principle. The deficiency indices can be determined through the symplectic invariants of the symplectic structure imposed by the symplectic product.

6.2 Symplectic Structure

In this section we will consider finite dimensional complex symplectic spaces with dimension $D \geq 0$. In this work a *complex symplectic space* S is a complex linear space with a prescribed form $[\cdot : \cdot]$ which satisfies the following properties:

- (i) $[c_1u + c_2v : w] = c_1[u : w] + c_2[v : w]$, (Linearity in the first variable)
- (ii) $[u : v] = -\overline{[v : u]}$, (Skew-Hermitian)
- (iii) $[u : S] = 0$ implies $u = 0$ (non-degeneracy),

for all $c_1, c_2 \in \mathbb{C}$ and $u, v, w \in S$.

A *Lagrangian subspace* $L \subset S$ is a subspace of S satisfying $[u : v] = 0$ for all $u, v \in L$.

A *symplectic isomorphism* is a linear bijective map $\mathcal{V}$ between two symplectic spaces S_1 and S_2 with symplectic forms $[\cdot : \cdot]_1$ and $[\cdot : \cdot]_2$ respectively, satisfying

$$[u, v]_1 = [\mathcal{V}u : \mathcal{V}v]_2. \quad (6.1)$$

In view of applications, if there is a symplectic isomorphism $\mathcal{V}$, say, between a symplectic space of interest S with symplectic product $[\cdot : \cdot]$ and the symplectic space $\mathbb{C}^D$, then we may represent the elements $\mathcal{V}u, \mathcal{V}v \in \mathbb{C}^D$ as the D -tuples $\mathcal{V}u = (u_1, \dots, u_D)$ and $\mathcal{V}v = (v_1, \dots, v_D)$. The symplectic isomorphism $\mathcal{V}$ induces a symplectic form on $\mathbb{C}^D$ specified by a $D \times D$ skew-Hermitian non-singular matrix H and we may represent the symplectic form on S relative to the isomorphism $\mathcal{V}$ as

$$[u : v] = (u^1, \dots, u^D)H(v^1, \dots, v^D)^*. \quad (6.2)$$

The isomorphism $\mathcal{V}$ takes bases in S into bases in $\mathbb{C}^D$. A basis in $\mathbb{C}^D$ is changed by congruent matrices which represent automorphisms on $\mathbb{C}^D$ which may or may not change the symplectic structure defined by H , but either way, through composition of maps this change of basis defines a new isomorphism of S onto $\mathbb{C}^D$, dependent on $\mathcal{V}$. Consequently, bases in $\mathbb{C}^D$ determining a symplectic structure (through the matrix H) are unique up to isomorphism $\mathcal{V}$. We will examine the symplectic invariants of S relative to the induced matrix H under an arbitrary fixed symplectic isomorphism $\mathcal{V}$, although the invariants are independent of $\mathcal{V}$. Since the spaces S and $\mathbb{C}^D$ are

symplectically isomorphic, for brevity we shall drop the subscript and denote their symplectic products by $[\cdot : \cdot]$. Unless otherwise stated, when we refer to a basis of S , we are referring to the basis in $\mathbb{C}^D$ isomorphic through $\mathcal{V}$ to the corresponding basis in S .

A *positive (negative) Lagrangian subspace* $P_{\pm} \subset S$ is a Lagrangian subspace which satisfies $[u, v] \geq 0 (\leq 0)$ for all $u, v \in P_{\pm}$. The *symplectic invariants* of S are $(p, q), \Delta, Ex$, determined by the symplectic structure of S . They are defined as follows:

$$\begin{aligned} p &= \max\{\text{complex dimension of positive linear subspaces}\}, \\ q &= \max\{\text{complex dimension of negative linear subspaces}\}, \end{aligned}$$

we call (p, q) the *signature* of S . The *Lagrangian index* Δ of S ,

$$\Delta = \max\{\text{complex dimension of Lagrangian subspaces of } S\},$$

and the *excess* Ex of S , $Ex = p - q$. The next theorem relates the diagonal form of the matrix H with the symplectic invariants.

Theorem 6.1 *Fix an isomorphism $\mathcal{V}$ and choose a basis for the symplectic space S with dimension $D \geq 0$ and corresponding skew-Hermitian matrix H determined by the symplectic product $[\cdot : \cdot]$, H is congruent to some diagonal matrix $G = \text{diag}\{i, i, \dots, i, -i, \dots, -i\}$, furthermore the symplectic invariants are related to H by p being the number of $+i$'s on the diagonal of G and q being the number of $-i$'s on the diagonal of G , also,*

$$D = p + q \quad \text{and} \quad \Delta = \min\{p, q\}.$$

Proof: Considering the symplectic space S , we may define a fixed isomorphism $\mathcal{V}$ from an arbitrary basis of S to a basis of $\mathbb{C}^D$, then for some associated matrix H we may write the symplectic form as

$$[u, v] = (\mathcal{V}u)H(\mathcal{V}v)^* \quad \text{for} \quad u, v \in S, \quad (6.3)$$

where $*$ represents the adjoint. Performing an eigenvector decomposition of H , yields $H = Q\Lambda Q^*$, where Q is a matrix with columns of orthonormal eigenvectors and $\Lambda = \text{diag}\{i, i, \dots, i, -i, \dots, -i\}$. Hence we see that H is congruent to Λ with Λ having $\pi \geq 0$ terms $(+i)$ and $\nu \geq 0$ terms $(-i)$. We may now write the symplectic form as

$$[u, v] = (\mathcal{V}uQ)\text{diag}\{i, i, \dots, i, -i, \dots, -i\}(\mathcal{V}vQ)^* \quad \text{for} \quad u, v \in S. \quad (6.4)$$

We first note that the skew-symmetric property of the symplectic form guarantees that $\text{Re}[v : v] = 0$ for all $v \in S$, thus the only possibilities for any element $v \in S$ is $\text{Im}[v : v] > 0$ (positive vector), $\text{Im}[v : v] < 0$ (negative vector) or $\text{Im}[v : v] = 0$ (neutral vector). Furthermore from (6.4) we have $D = \pi + \nu$. We will prove that

$$p = \pi \quad \text{and} \quad q = \nu. \quad (6.5)$$

We first consider the positive subspace $P_+ \subset S$ such that $\dim P_+ = \pi$, which is spanned by the first π basis vectors of S taken componentwise. It is clear from (6.4) that $\text{Im}[v : v] \geq 0$ for all $v \in P_+$. Since p is the maximal dimension of such a subspace, we have $p \geq \pi$. To show the reverse inequality we will show that every linear positive subspace $P \subset S$ has $\dim P \leq \pi$. Let the subspace $N_- \subset S$ be spanned by the basis vectors which are not a basis for P , thus N_- is a negative Lagrangian subspace with dimension ν . Since $P \cap N_- = \{0\}$ we have that $\dim P + \nu \leq D$, hence $\dim P \leq D - \nu = \pi$. We can prove $q = \nu$ in a similar fashion. We now prove the last inequality $\Delta = \min\{p, q\}$. Without loss of generality we may assume that $p \geq q$. The justification above is enough to define $\mathcal{V}$ so that the symplectic structure is determined by the matrix

$$\Lambda = \begin{pmatrix} iI_q & 0 & 0 \\ 0 & -iI_q & 0 \\ 0 & 0 & iI_{|Ex|} \end{pmatrix}, \quad (6.6)$$

where $|Ex|$ represents the magnitude of the excess. Let us denote our componentwise basis for S , relative to the symplectic isomorphism $\mathcal{V}$, by

$$\{e^1, e^2, \dots, e^q, f^1, f^2, \dots, f^q, g^1, \dots, g^{Ex}\}.$$

We may form a Lagrangian subspace L_1 of S defined by

$$L_1 = \text{span}\{e^1 + f^1, e^2 + f^2, \dots, e^q + f^q\}. \quad (6.7)$$

It is clear that $\dim L_1 = q$, hence $\Delta \geq q$. Every vector $v \in P_+$, except $v = 0$, satisfies $\text{Im}[v : v] > 0$, thus every Lagrangian subspace $L \subset S$ satisfies $L \cap P_+ = \{0\}$. Hence $\dim L + p \leq D$, thus $\dim L \leq D - p = q$. This shows that $\Delta = q$. We may use a similar argument for when $q \geq p$ to complete the proof. ■

It is of particular interest for applications when $Ex = 0$, in this case, $p = q$ and $D = 2p$. The deficiency indices are given by p and q . From the above theorem $D = 2\Delta$ and the symplectic space S is isomorphic to $\mathbb{C}^{2\Delta}$ and can be seen as the complexification¹ of a real symplectic space $\mathbb{R}^{2\Delta}$, in this case the canonical basis may be chosen with an appropriate symplectic isomorphism $\mathcal{V}$ and the associated matrix determining the symplectic structure can take the form

$$K = \begin{pmatrix} 0 & I_\Delta \\ -I_\Delta & 0 \end{pmatrix}. \quad (6.8)$$

6.3 Complete Lagrangian subspaces

A *Lagrangian subspace* $\hat{L}$ is a subspace of the symplectic space S where every vector is symplectically orthogonal to every other vector in $\hat{L}$. A *complete* Lagrangian subspace L is a subspace of S which satisfies,

$$[u, L] = 0 \text{ and } u \in S \quad \text{implies} \quad u \in L. \quad (6.9)$$

Thus a complete Lagrangian subspace is the maximal subspace of all self orthogonal Lagrangian subspaces. Two subspaces $S_1, S_2 \subset S$ are *symplectic direct sum decompositions* of S denoted by

$$S = S_1 \oplus S_2, \quad (6.10)$$

¹cf. Markus and Everitt [10, pp, 4909]

in the case S_1 and S_2 are the symplectic ortho-complements of one another. There exists a basis, unique up to isomorphism $\mathcal{V}$, which can be decomposed into the direct sum

$$S = S^{2\Delta} \oplus S^{|Ex|} \quad \text{with} \quad [S^{2\Delta} : S^{|Ex|}] = 0, \quad (6.11)$$

with dimension 2Δ and $|Ex|$, and excess 0 and $|Ex|$, respectively. Depending on the choice of $\mathcal{V}$, an appropriate basis can be chosen in $\mathbb{C}^D$ so that the $D \times D$ skew-hermitian matrix preserving the symplectic structure is

$$\hat{K} = \begin{pmatrix} 0 & I_{\Delta} & 0 \\ -I_{\Delta} & 0 & 0 \\ 0 & 0 & \pm i I_{|Ex|} \end{pmatrix}. \quad (6.12)$$

The basis in $\mathbb{C}^D$ corresponding to the above skew-hermitian matrix is constructed by first starting with the direct sum decomposition

$$S = P_+ \oplus N_-, \quad (6.13)$$

where P_+ and N_- are as defined in Theorem 6.1. Taking any Lagrangian δ -space ($\delta \leq \Delta$) defined by $\hat{L} = \text{span}\{e^1, e^2, \dots, e^{\delta}\}$ we can decompose each basis vector in terms of its components in the N_- and P_+ spaces. If need be we may recursively add orthogonal basis vectors to $\hat{L}$, each of which are the sum of orthogonal vectors in the subspaces P_+ and N_- not in $\hat{L}$. The resulting complete Lagrangian Δ -space, $S^{2\Delta} = \{e^1, e^2, \dots, e^{\delta}, \dots, e^{\Delta}\}$, is constructed with its ortho-complement being $S^{|Ex|}$. The associated matrix for the symplectic space $S^{2\Delta}$ is the first 2Δ diagonal elements of (6.6). The basis vectors of $S^{2\Delta}$ can be redefined to get the basis for $S^{2\Delta}$ in canonical form and thus the associated matrix becomes (6.12). This full construction is seen in [10, p. 4921]. The construction is important because it shows how to construct larger Lagrangian subspaces from smaller ones, and it is used in the next theorem to obtain a necessary and sufficient condition for the existence of complete Lagrangian subspaces.

Theorem 6.2 *Let S be a complex symplectic space with symplectic form $[\cdot : \cdot]$ then,*
(i) there exists a complete Lagrangian subspace of S if and only if S has excess $Ex = 0$.
If $Ex = 0$ then $\dim S = 2\Delta$ and also,
(ii) the Lagrangian subspace L is complete if and only if $\dim L = \Delta$.

Proof: To prove necessity of (i), let $Ex = 0$, then S is the complexification of the real symplectic space $\mathbb{R}^{2\Delta}$. Thus there exists a canonical basis² $\{e^1, e^2, \dots, e^\Delta, \dots, e^{2\Delta}\}$ relative to $\mathcal{V}$ with associated skew-Hermitian matrix (6.8) of the symplectic form. Let us define the subspace $L = \text{span}\{e^1, \dots, e^\Delta\}$, then it is clear that L is a Lagrangian subspace, furthermore L is complete. To see this, consider an arbitrary element $u \in S$ with $u = c_1 e^1 + c_2 e^2 + \dots + c_{2\Delta} e^{2\Delta}$ satisfying $[u, L] = 0$. Then using the definition of the canonical basis $[e^i, e^{\Delta+j}] = \delta_{ij}$ for $1 \leq i, j \leq \Delta$, we have $c_{\Delta+1} = c_{\Delta+2} = \dots = c_{2\Delta} = 0$, hence $u \in L$. We prove sufficiency of (i) by contraposition. Suppose $Ex \neq 0$, then without loss of generality, we may assume that $Ex = p - q > 0$, so that $\Delta = \min\{p, q\} = q$. There exists a basis for S outlined previously to be $\{\hat{e}^1, \dots, \hat{e}^{2\Delta}, \epsilon^1, \dots, \epsilon^{Ex}\}$. For a given Lagrangian subspace $\hat{L}$ and the associated skew-hermitian matrix (6.12), we know that $\hat{L} = \text{Span}\{\hat{e}^1, \dots, \hat{e}^\Delta\}$, however $[\epsilon^1 : \hat{L}] = 0$ and $\epsilon^1 \notin \hat{L}$, thus $\hat{L}$ is not complete.

In the case $Ex = 0$, it is clear that $\dim S = 2\Delta$. For necessity of (ii) we know from the construction above that for any Lagrangian δ -space $\hat{L}$ with $\delta \leq \Delta$ there exists a canonical basis $\{e^1, e^2, \dots, e^\Delta, \dots, e^{2\Delta}\}$ such that $\hat{L} = \text{span}\{e^1, \dots, e^\delta\}$. From the argument used in the necessity of (i), $\hat{L}$ is complete in case $\delta = \Delta$. Finally, we prove sufficiency of (ii) by contraposition. If $\delta < \Delta$, then $\hat{L}$ is not complete, since we may choose $\hat{e}^{\delta+1} \in S$ such that $[\hat{e}^{\delta+1} : \hat{L}] = 0$ yet $\hat{e}^{\delta+1} \notin \hat{L}$. ■

²cf. [10, pp. 4922 (2.35) and (2.36)]

6.4 Boundary space and the GKN-theorem

The following results hold for any even ordered ordinary differential equation although we will just consider the case of equations of the second order. We will consider the same minimal and maximal operators defined in section 2.5, save some notational renaming. These domains are chosen so that any self-adjoint extension will be an extension of the minimal domain and a restriction of the maximal domain. Let the *boundary space* $\mathbf{S}$ be the quotient space defined as

$$\mathbf{S} = \mathcal{D}(T_1) \setminus \mathcal{D}(T_0). \quad (6.14)$$

This is a 4 complex dimensional space consisting of elements which are equivalence classes of functions from $\mathcal{D}(T_1)$ with the same function and derivative values at the boundary. That is for any $u \in \mathcal{D}(T_1)$,

$$\mathbf{u} \in \mathbf{S} \quad \text{implies} \quad \mathbf{u} = \{v \in \mathcal{D}(T_1) : u - v \in \mathcal{D}(T_0)\}.$$

We may define the symplectic form, making use of the Lagrange identity³, as

$$[\mathbf{f} : \mathbf{g}] = [f : g] = (T_1 f, g) - (f, T_1 g) = p(x)[f(x)\bar{g}'(x) - \bar{f}'(x)g(x)] \Big|_a^b. \quad (6.15)$$

The symplectic form defined above is non-degenerate since $[\mathbf{u} : \mathbf{S}] = 0$ implies that $u \in \mathcal{D}(T_0)$, hence $\mathbf{u} = 0$. Therefore $\mathbf{S}$ together with the symplectic form above, forms a complex symplectic space. The next theorem relates the Lagrangian subspaces of $\mathbf{S}$ to the domain of the self-adjoint operators which are restrictions of T_1 .

Theorem 6.3 GKN-theorem

There exists a bijection between the set of all self-adjoint operators T as restrictions of $\mathcal{D}(T_1)$ and the set of all complete Lagrangian subspaces $\mathbf{L}$ on $\mathbf{S}$. For each complete Lagrangian subspace $\mathbf{L}$ corresponding to the self-adjoint operator T , the Lagrangian subspace is defined as,

$$\mathbf{L} = \mathcal{D}(T) / \mathcal{D}(T_0). \quad (6.16)$$

³cf. statement of equation (2.39)

The domain of the corresponding self-adjoint extension is defined as,

$$\mathcal{D}(T) = a_1 f_1 + \cdots + a_d f_d + \mathcal{D}(T_0), \quad (6.17)$$

where $a_1, \dots, a_d \in \mathbb{R}$ are arbitrary and $f_1, \dots, f_d$ are representatives of the equivalence classes $\mathbf{f}_1, \dots, \mathbf{f}_d$, respectively, which form a basis of the Lagrangian subspace $\mathbf{L}$ in $\mathbf{S}$.

Proof: Consider any self-adjoint operator T whose domain is the restriction of the maximal operator T_1 and an extension of the minimal operator T_0 . Define the subset $\mathbf{L}$ of $\mathbf{S}$ by

$$\mathbf{L} = \mathcal{D}(T)/\mathcal{D}(T_0). \quad (6.18)$$

The elements of $\mathbf{S}$ form equivalence classes and hence have a fixed associated projection mapping $\Psi : \mathcal{D}(T_1) \rightarrow \mathbf{S}$ taking each element of $\mathcal{D}(T_1)$ into its equivalence class, hence

$$\Psi(f) = \mathbf{f} = f + \mathcal{D}(T_0) \quad \text{for all } f \in \mathcal{D}(T_1). \quad (6.19)$$

We will show that $\mathbf{L}$ is a Lagrangian subspace. It is clear that $\mathbf{L}$ is a linear subspace of $\mathbf{S}$, since $\mathcal{D}(T)$ is a linear subspace of $\mathcal{D}(T_1)$ and Ψ is a linear projective map. We compute the symplectic product for any $\mathbf{f}, \mathbf{g} \in \mathbf{L}$ considering that T is a symmetric restriction of T_1 ,

$$\begin{aligned} [\mathbf{f} : \mathbf{g}] &= (T_1 f, g) - (f, T_1 g) \quad \text{for all } f \in \mathbf{f}, g \in \mathbf{g}, \\ &= (T f, g) - (f, T g) = 0. \end{aligned}$$

It follows that $\mathbf{L}$ is a Lagrangian subspace of $\mathbf{S}$. To show that it is complete, consider any $\mathbf{h} \in \mathbf{S}$ satisfying $[\mathbf{h} : \mathbf{L}] = 0$, that is,

$$(Th, g) = (T_1 h, g) = (h, T_1 g) = (h, T g) \quad \text{for all } g \in \mathcal{D}(T), h \in \mathbf{h}. \quad (6.20)$$

However this is sufficient, by definition of the adjoint ⁴ for $h \in \mathcal{D}(T^*)$ to be in $\mathcal{D}(T)$ for all $h \in \mathbf{h}$. Thus $\mathbf{h} \in \mathbf{L}$ and $\mathbf{L}$ is complete.

Now for the converse, let $\mathbf{L}$ be any Lagrangian subspace of $\mathbf{S}$, we must show that

⁴cf. equation (2.1)

there exists a self-adjoint operator T . Moreover we must show that $\mathbf{L}$ takes the form (6.16). To this end we require $\Psi\mathcal{D}(T) = \mathbf{L}$, therefore we define $\mathcal{D}(T)$ to be the inverse set mapping of Ψ as

$$\mathcal{D}(T) = \Psi^{-1}\mathbf{L} = \{h \in \mathcal{D}(T_1) : \mathbf{h} \in \mathbf{L}\}. \quad (6.21)$$

From the above and $\mathcal{D}(T_0) = \mathbf{0} \in \mathbf{L}$, we see that $\mathcal{D}(T_0) \subset \mathcal{D}(T) \subset \mathcal{D}(T_1)$. Thus T is a restriction of T_1 . We will show that T is self-adjoint. Since $\mathbf{L}$ is a Lagrangian subspace of $\mathbf{S}$ we have

$$[\mathbf{f} : \mathbf{g}] = (Tf, g) - (f, Tg) = 0 \quad \text{for all } f, g \in \mathcal{D}(T).$$

From the definition of the domain⁵ of the adjoint $\mathcal{D}(T^*)$, we have that $\mathcal{D}(T) \subset \mathcal{D}(T^*)$. Furthermore T is a restriction of T_1 and since T_0 and T_1 are adjoint to one another,

$$\mathcal{D}(T_0) = \mathcal{D}(T_1^*) \subset \mathcal{D}(T^*) \subset \mathcal{D}(T_0^*) = \mathcal{D}(T_1).$$

Therefore T^* is also a restriction of T_1 . For $f \in \mathcal{D}(T^*)$ we know by definition,

$$(T^*f, g) = (f, Tg) \quad \text{for all } g \in \mathcal{D}(T). \quad (6.22)$$

Since T and T^* are restrictions of T_1 , fixing a $f \in \mathcal{D}(T^*)$ and using (6.22) we have,

$$[f : g] = (T_1f, g) - (f, T_1g) = 0 \quad \text{for all } g \in \mathcal{D}(T).$$

The above statement in the quotient space is equivalent to $[\mathbf{f} : \mathbf{L}] = 0$ but since $\mathbf{L}$ is complete $\mathbf{f} \in \mathbf{L}$ we have $f \in \mathcal{D}(T)$. We have shown that T is self-adjoint and from (6.21) we see that $\Psi\mathcal{D}(T) = \mathbf{L}$, thus $\mathbf{L}$ takes the form (6.18). From equation (6.21), it is clear that we can define a surjective mapping Ψ from the set of all self-adjoint operators $\{T\}$ onto the set of all Lagrangian subspaces $\{\mathbf{L}\}$ induced by the mapping Ψ and defined by,

$$\Psi : T \rightarrow \mathbf{L} = \mathcal{D}(T)/\mathcal{D}(T_0). \quad (6.23)$$

To complete the proof we need to show that Ψ is injective. Consider two different self-adjoint operators T_α and T_β which are restrictions of T_1 . Define the two corresponding

⁵cf. domain $\mathcal{D}(T^*)$ of equation (2.1)

Lagrangian subspaces as,

$$\mathbf{L}_\alpha = \mathcal{D}(T_\alpha)/\mathcal{D}(T_0) \quad \text{and} \quad \mathbf{L}_\beta = \mathcal{D}(T_\beta)/\mathcal{D}(T_0),$$

respectively. Without loss of generality we may assume that $\mathcal{D}(T_\beta) \subseteq \mathcal{D}(T_\alpha)$. Then there exists a $u \in \mathcal{D}(T_\alpha)$ not in $\mathcal{D}(T_\beta)$. This means that $\mathbf{u} = \{u + \mathcal{D}(T_0)\} \in \mathbf{L}_\alpha$ with $\mathbf{u} \notin \mathbf{L}_\beta$, otherwise $u \in \mathcal{D}(T_\beta)$. So $\mathbf{L}_\alpha \neq \mathbf{L}_\beta$, thus by contraposition Ψ is injective. ■

6.5 A classification of self-adjoint boundary conditions

The GKN-theorem gives a bijection between Lagrangian subspaces and self-adjoint operators T . With the use of symplectic algebra it is possible to obtain a classification of the class of self-adjoint operators in terms of the intrinsic properties of their corresponding Lagrangian subspaces. A brief exposition of the results obtained by Markus and Everitt, [10], will be given. One of the main results which gives a characterization of complete Lagrangian subspaces is known as the *balanced intersection principle*. This principle states that for every symplectic direct sum decomposition of $\mathbf{S}$ given by

$$\mathbf{S} = \mathbf{S}_+ \oplus \mathbf{S}_-, \tag{6.24}$$

with symplectic invariants $\Delta_\pm$ and $Ex_\pm$, respectively, every complete Lagrangian subspace $\mathbf{L} \subset \mathbf{S}$ satisfies

$$0 \leq \Delta_- - \dim(\mathbf{L} \cap \mathbf{S}_-) = \Delta_+ - \dim(\mathbf{L} \cap \mathbf{S}_+) \leq \min\{\Delta_-, \Delta_+\}. \tag{6.25}$$

To give an intuitive meaning to the characterization we will choose a particular decomposition of S . Define

$$D_-(T_1) := \{f \in D(T_1) : \text{some neighborhood of } b \text{ does not intersect } \text{supp}(f)\}$$

$$D_+(T_1) := \{f \in D(T_1) : \text{some neighborhood of } a \text{ does not intersect } \text{supp}(f)\}.$$

For Ψ defined as in (6.19), we set

$$\mathbf{S}_1 := \Psi D_-(T_1) \quad \text{and} \quad \mathbf{S}_2 := \Psi D_+(T_1).$$

The spaces $\mathbf{S}_1$ and $\mathbf{S}_2$ are called the left and right boundary spaces, respectively, see [11, p. 490]. For any non-zero vector $\mathbf{v} \in \mathbf{S}$ we say that $\mathbf{v}$ is separated at the left in case $\mathbf{v} \in \mathbf{S}_-$, separated at the right in case $\mathbf{v} \in \mathbf{S}_+$ and coupled otherwise. A basis for $\mathbf{L} \subset \mathbf{S}$ is said to be *minimally coupled* if no other basis for $\mathbf{L}$ has fewer coupled vectors. For each minimally coupled basis of $\mathbf{L}$, the number of coupled vectors is called the *necessary coupling* of $\mathbf{L}$, denoted by

$$\text{Nec-coupling}(\mathbf{L}) = \Delta - \dim(\mathbf{L} \cap \mathbf{S}_-) - \dim(\mathbf{L} \cap \mathbf{S}_+). \quad (6.26)$$

A Lagrangian subspace $\mathbf{L} \subset \mathbf{S}$ is said to be *strictly separated* if

$$\text{Nec-coupling}(\mathbf{L}) = 0, \quad (6.27)$$

and $\mathbf{L}$ is said to be *totally coupled* if

$$\text{Nec-coupling}(\mathbf{L}) = \Delta. \quad (6.28)$$

The *coupling grade* of $\mathbf{L}$ is defined as

$$\text{grade}(\mathbf{L}) := \Delta_- - \dim(\mathbf{L} \cap \mathbf{S}_-) = \Delta_+ - \dim(\mathbf{L} \cap \mathbf{S}_+). \quad (6.29)$$

It is known⁶ that there exists Lagrangian subspaces with all possible coupling grades $l = 0, \dots, \min\{\Delta_-, \Delta_+\}$. For any complete basis $\{\mathbf{f}_1, \dots, \mathbf{f}_\Delta\}$ of $\mathbf{L}$ there are at least $\dim(\mathbf{L} \cap \mathbf{S}_-)$ vectors which are separated at the left, at least $\dim(\mathbf{L} \cap \mathbf{S}_+)$ vectors separated at the right, and at most $\text{Nec-coupling}(\mathbf{L})$ coupled vectors, with equality when the basis of $\mathbf{L}$ is minimally coupled.

Suppose we have a basis for $\mathbf{S}$

$$\{f^1, \dots, f^\Delta, f^{\Delta+1}, \dots, f^{2\Delta}\}, \quad (6.30)$$

then it is said to be a *canonical basis* if the following conditions are satisfied:

$$[f^j : f^{\Delta+j}] = -[f^{\Delta+j} : f^j] = 1 \quad \text{for} \quad 1 \leq j \leq \Delta,$$

⁶ cf. Markus and Everitt [10, pp. 4934]

$$[f^i : f^j] = 0 \quad \text{for } i \neq \Delta + j.$$

We may choose for $\mathbf{S}$ the canonical basis

$$\begin{aligned} f^1(x) &= \frac{(x-a)(b-x)^2}{p(a)(b-a)^2}, & f^2(x) &= \frac{(a-3b+2x)(a-x)^2}{(a-b)^3}, \\ f^3(x) &= \frac{(3a-b-2x)(b-x)^2}{(a-b)^3}, & f^4(x) &= \frac{(a-x)^2(x-b)}{p(b)(b-a)^2}. \end{aligned}$$

We next introduce the isomorphism $\mathcal{V} : \mathbf{S} \rightarrow \mathbb{C}^4$ defined by

$$\mathbf{f} \rightarrow \{p(a)f'(a), f(b), p(a), p(b)f'(b)\}. \quad (6.31)$$

The symplectic isomorphism $\mathcal{V}$ is associated with the skew-Hermitian matrix K defined in (6.8), furthermore $\mathcal{V}$ maps the above canonical basis into a canonical basis for $\mathbb{C}^4$ defined by

$$e_1 = \{1, 0, 0, 0\}, e_2 = \{0, 1, 0, 0\}, e_3 = \{0, 0, 1, 0\} \text{ and } e_4 = \{0, 0, 0, 1\}.$$

We may now define the symplectic form as follows

$$[\mathbf{f} : \mathbf{g}] = \mathcal{V}(\mathbf{f})K\mathcal{V}(\mathbf{g}) \quad \text{for all } \mathbf{f}, \mathbf{g} \in \mathbf{S}. \quad (6.32)$$

Using direct algebraic manipulation to verify possible complete Lagrangian subspaces in $\mathbb{C}^4$, we may consider some complete Lagrangian subspaces given by:

$$L_2 = \{e_1, e_4\}, L_3 = \{e_2, e_3\} \text{ and } L_4 = \{e_1 + e_4, e_2 + e_3\}.$$

We see that L_2 and L_3 are strictly separated with $\text{grade}(L_2)$ equal to 0, while L_4 is totally coupled with $\text{grade}(L_4)$ equal to 1. We may now use the GKN-Theorem to yield the self-adjoint domain of the operators corresponding to each complete Lagrangian subspace mentioned above. We use the inverse mapping of $\mathcal{V}$ to obtain the corresponding complete Lagrangian subspaces in $\mathbf{S}$, and then we obtain the domain of each self-adjoint operator through formula (6.17) giving⁷:

$$\begin{aligned} \mathcal{D}(T_2) &= \{f \in \mathcal{A} : f(a) = 0 = f(b)\}, & \mathcal{D}(T_3) &= \{f \in \mathcal{A} : f'(a) = 0 = f'(b)\}, \\ \mathcal{D}(T_4) &= \{f \in \mathcal{A} : f(a) = f(b) \text{ and } p(a)f'(a) = p(b)f'(b)\}. \end{aligned}$$

⁷ $\mathcal{A}$ is defined in Section 2.5

Here T_2 , T_3 and T_4 correspond to L_2 , L_3 and L_4 , respectively. The operator T_2 is obtained by imposing Dirichlet boundary conditions, T_3 is obtained by imposing Neumann boundary conditions and T_4 is obtained by imposing limit-circle boundary conditions. We may also consider the following complete Lagrangian subspace of $\mathbb{C}^4$,

$$L_5 = \{e_3 \cos \alpha - e_1 \sin \alpha, e_2 \cos \beta - e_4 \sin \beta\}. \quad (6.33)$$

This Lagrangian subspace relative to $\mathcal{V}$, is associated with the boundary conditions (1.2) and (1.3). It is clear that a continuous transformation in the variables α and β can define a continuous transformation of L_2 into L_3 through intermediary subspaces which are also complete Lagrangian subspaces of $\mathbf{S}$.

It was illustrated above that transformation of the symplectic form by congruent matrices is also a change of basis in $\mathbb{C}^D$, see (6.2). Of all the transformations of K by congruent matrices the most important are those transformations which preserve the symplectic structure defined by K . We call the transformation B a *symplectic automorphism* of $\mathbb{C}^D$ if it satisfies the following formula

$$BKB^* = K. \quad (6.34)$$

The above subset of transformations is important because under these transformations the matrix K which determines the symplectic structure is held constant while the basis of $\mathbb{C}^D$ is changed. This has the effect through the isomorphism $\mathcal{V}$, that Lagrangian subspaces of $\mathbb{C}^D$ are mapped to the corresponding Lagrangian subspaces in the space $\mathbf{S}$.

The group of all symplectic automorphisms of $\mathbb{C}^D$ and thus $\mathbf{S}$ is called the *Sturm-Liouville group* denoted by $\Sigma\wedge(\mathbb{C})$. The topology of this group was studied by Markus, [30]. The Sturm-Liouville group is known to be a 16 dimensional noncommutative, noncompact topological group which is a closed, arcwise-connected submanifold of the general linear group $GL(4, \mathbb{C})$ of 4×4 nonsingular matrices.

It is conjectured by Markus and Everitt, [11], that since the Sturm-Liouville group is arcwise-connected there should be a transformation between strictly separated and

totally coupled complete Lagrangian subspaces.

Chapter 7

Conclusion

7.1 Summary

What has preceded has been a broad introduction to the interesting, vibrant subject area of second order differential equations. Various questions on the main problem outlined in Chapter 1 have been answered, moreover Sturm-Liouville theory as presented here serves as an introduction to the areas of Functional Analysis, Calculus of Variations and symplectic algebra.

Chapter 2 lays down the foundations of deficiency index theory and answers the question of whether a given domain of 1.1 is self-adjoint. Chapter 3 characterizes the spectrum of the Sturm-Liouville equation and gives a concrete representation of the Sturm-Liouville expansion using the spectral theorem. The asymptotics are calculated and are used to prove the completeness of the eigenfunctions. In chapter 4 the Sturm theorems are proven, which relates the number zeros of the eigenfunctions to the eigenvalue count. In chapter 5 we conclude that the Sturm-Liouville problem has a variational formulation and using this, we prove the Courant nodal domain theorem. In chapter 6 we verify the existence of an isomorphism between the set of all self-adjoint operators and the set of all complete Lagrangian subspaces. This gives

us an algebraic analogue to the deficiency index theory described in chapter 2.

The project offers a beneficial perspective to the reader because it outlines in the earlier chapters the classical results in the theory. It also used these results as a stepping stone in the later chapters for the more contemporary work, thus giving a good history while laying the foundations for research in the area.

7.2 Future Directions

Due to the broad nature of the project, various diverging research paths can be taken. In Section 3.4, asymptotic estimates were found for the eigenvalues. We may consider the equation (1.1) where p and r are allowed to change sign. This has been investigated by many authors, but most results require additional conditions on p and r . In addition, for this indefinite case there remain open questions regarding completeness of the eigenfunctions.

In Section 5.4, the Lusternik-Schirelmann method was outlined. This gave a variational method which used topological considerations to obtain information regarding critical values of a given even functional. It might be possible to apply this theory to differential equations on graphs to obtain new results. In Section 5.5, the Courant nodal domain theorem gives a variational proof of an upper bound for the number of zero's of each eigenfunction. It would be of benefit if a variational proof for a lower bound were found because this could be used in the theory of non-linear eigenvalue problems.

In Section 6, a transformation was found between two strictly separated complete Lagrangian subspaces of the boundary symplectic space in the form of a symplectic automorphism. It would be interesting to study the topology of the Sturm-Liouville group of symplectic automorphism as this would give insight into the existence of transformations between strictly separated and totally coupled complete Lagrangian subspaces, as conjectured by Markus and Everitt, [11]. This would be beneficial

because it would give insight into new ways of classifying the self-adjoint boundary conditions. However if there does not exist such a transformation then it would mean that separate analytic theories have been developed for operators with these two boundary conditions because their underlying symplectic spaces are fundamentally different.

Appendix A

Grönwall's Lemma

Lemma A.1 *Let ϕ, ψ and $\chi > 0$ be piecewise continuous functions on the interval $a \leq x \leq b$, if*

$$\phi(t) \leq \psi(t) + \int_a^t \chi(s)\phi(s)ds, \quad (1.1)$$

then

$$\phi(t) \leq \psi(t) + \int_a^t \chi(s)\psi(s)\exp\left(\int_s^t \chi(u)du\right)ds. \quad (1.2)$$

Proof: Let

$$R(t) := \int_a^t \chi(s)\phi(s)ds.$$

Now

$$R'(t) - \chi(t)R(t) = \chi(t)\left[\phi(t) - \int_a^t \chi(s)\phi(s)ds\right].$$

Using (1.1), we have,

$$R'(t) - \chi(t)R(t) \leq \chi(t)\psi(t).$$

Multiplying the above by a suitable integrating factor and integrating gives,

$$\int_a^t \left[R'(s) - \chi(s)R(s)\right]e^{\int_s^t \chi(u)du}ds \leq \int_a^t \chi(s)\psi(s)e^{\int_s^t \chi(u)du}ds. \quad (1.3)$$

The integral on the left hand side can be rewritten using the chain rule as follows,

$$\int_a^t \frac{d}{ds} \left(R(s)e^{-\int_t^s \chi(u)du} \right) ds.$$

By the Fundamental Theorem of Calculus, this term is precisely $R(t)$, so by inequality (1.3), we obtain,

$$\int_a^t \chi(s)\phi(s)ds \leq \int_a^t \chi(s)\psi(s)e^{\int_s^t \chi(u)du}ds. \quad (1.4)$$

On adding $\psi(t)$ to both sides of (1.4) and using (1.1) we obtain (1.2). ■

Bibliography

- [1] V. Arnold. On characteristic classes in quantization. *Functional Analysis and its applications*, **1**, 1969.
- [2] S. Banach. sur les opérations dans les ensembles abstraits et leur application aux équations intégrals. *Fundamenta Mathematicæ*, 3:133 – 181, 1922.
- [3] E. Coddington and N. Levinson. *Theory of differential equations*. McGraw-Hill Book Company, 1955.
- [4] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Springer-Berlin, 1937.
- [5] K. Deimling. *Nonlinear Functional Analysis*. Springer-Berlin, 1985.
- [6] P. Drábek and J. Milota. *Methods of nonlinear analysis : applications to differential equations*. Birkhäuser, 2007.
- [7] P. Drábek and S. Robinson. On the generalization of the courant nodal domain theorem. *Journal of Differential Equations*, 181:58–71, 2002.
- [8] J. Dugundji. *Topology*. Allyn and Bacon, 1966.
- [9] Edwards. Generalised sturm theorem. *The Annals of Mathematics*, 80:22–57, July 1964.

- [10] W.N. Everitt and L. Markus. Complex symplectic geometry with applications to ordinary differential operators. *Transactions of the American Mathematical Society*, **351**(12):4905–4945, 1999.
- [11] W.N. Everitt and L. Markus. Complex symplectic spaces and boundary value problems. *Bulletin of the American mathematical society*, 42(4):461–500, July 2005.
- [12] I. Fredholm. sur une classe déquations fonctionelles. *Acta Mathematica*, 27:365–390, 1903.
- [13] N. Goussoub. *Duality and perturbation methods in critical point theory*, volume 99. Cambridge University press, 1993.
- [14] M.R. Hestenes. Applications of the theory of quadratic forms in hilbert space to the calculus of variations. *Pacific Journal of Mathematics*, 1(525-581), 1951.
- [15] D. Hilbert. *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. Chelsea, 1952.
- [16] H. Hochstadt. On inverse problems associated with sturm-liouville operators. *Journal of differential Equations*, 17:220–235, november 1973.
- [17] V. Hutson and J. S. Pym. Applications of functional analysis and operator theory. *Academic Press*, 146, 1980.
- [18] E.L. Ince. *Ordinary differential equations*. New York: dover, 1956.
- [19] V.A. Jakubovič. Oscillation properties of solutions. *American Mathematical Society*, **42**:247–288, 1964.
- [20] I. M. James. On the category, in the sense of lusternik-schnirelmann. *Topology*, 17:331–348, 1978.
- [21] E. Kamke. A new proof of sturms comparison theorems. *American Mathematical Society*, 1938.

- [22] K. Kodaira. The eigenvalue problem for ordinary differential equations of the second order and heisenberg's theory of s-matrices. *American journal of Mathematics*, 71:921–945, Oct. 1949.
- [23] K. Kodaira. On ordinary differential equations of any even order and the corresponding eigenfunction expansions. *American journal of Mathematics*, 72:502–544, july. 1950.
- [24] E. kreyszig. *Introductory Functional analysis with Applications*. John Wiley and Sons Inc., 1978.
- [25] J. Liouville. Mémoire sur le développement des fonctions ou parites de fonctions en séries dont les divers termes sont assujettis à satusfaire à une même équation différentielle du seond ordre, contenant un paramètre variable. *Math. Pures Appl.*, 1:253–265, 1836.
- [26] J. Liouville. Second mémoire sur le développement des fonctions ... *Journ. Math. Pures Appl.*, 2:16–35, 1837.
- [27] J. Liouville. Troisième mémoire sur le développement des fonctions ... *Journ. Math. Pures Appl.*, 2:418–437, 1837.
- [28] E. R. Lorch. *Spectral Theory*. Oxford University Press(New York), 1962.
- [29] J. Lützen. Sturm and liouville's work on ordinary linear differential equations. the emergence of sturm-liouville theory. *Archive for History of Exact Sciences*, 29:309–376, 2004.
- [30] L. Markus. The sturm-liouville group. *Journal of Computational and Applied Mathematics*, **171**:335–365, 2004.
- [31] M. A. Naimark. *Linear differential operators*. Frederich Ungar Publishing co., 1968.

- [32] R.D. Richardson. Theorems of oscillation for two linear differential equations of the second order with two parameters. *Transactions of the American Mathematical Society*, **13**(1):22–34, 1911.
- [33] R.D. Richardson. Contributions to the study of oscillation properties of the solutions of linear differential equations of the second order. *The John Hopkins University Press*, **40**(3):283–316, 1918.
- [34] W. Ried. *Ordinary Differential Equations*. John Wiley and Sons, 1971.
- [35] S. Currie and B. Watson. Dirichlet-neumann bracketing for boundary-value problems on graphs. *Electronic Journal of Differential Equations*, 2005(93):1–11.
- [36] R. E. Showalter. *Hilbert Space Methods for Partial Differential Equations*. Electron. J. Diff. Eqns., 01 2001.
- [37] M. Spivak. *Introduction to differential geometry*, volume 1. Publish or Perish, 2005.
- [38] C. Sturm. Analyse d’un mémoire sur les propriétés générales des fonctions qui dépendent d’équations différentielles linéaires du second ordre. *L’Institut Journ. Acad. et Soc. Sci.*, pages 219–223, nov. 1836.
- [39] C. Sturm. Mémoire sur les Équations différentielles linéaires du second ordre. *Journ. Math. Pures Appl.*, 1:106–186, sept. 1836.
- [40] G. Teschl. *Mathematical Methods in Quantum Mechanics with applications to Schrödinger Operators*. American mathematical Society, Rhode Island, 2009.
- [41] E. Titchmarsh. *Eigenfunction Expansions: associated with second order differential equations*, volume 1. Oxford: Clarendon Press, 1946-1958.
- [42] E. Titchmarsh. *The Theory of Functions*. Oxford University Press, second edition edition, 1958.
- [43] J. Weidmann. *Linear Operators in Hilbert Spaces*. Springer-Verlag.

- [44] A. Weinstein. Symplectic manifolds and their lagrangian submanifolds. *Advances in Mathematics*, **6**:329–346, 1971.
- [45] A. Weinstein. Symplectic geometry. *Bulletin of the American mathematical society*, **5** , 1981.
- [46] E. Zeidler. *Non-Linear analysis and its applications*, volume 1. Springer-Verlag, 1986.