Generalization of the Limit-Point & Limit-Circle Classification to Equations With Distributional Coefficients

Varun Bhardwaj
University Of Alabama At Birmingham

GENERALIZATION OF THE LIMIT-POINT & LIMIT-CIRCLE CLASSIFICATION TO EQUATIONS WITH DISTRIBUTIONAL COEFFICIENTS

by

VARUN BHARDWAJ

RUDI WEIKARD, COMMITTEE CHAIR
SHANGBING AI
IAN KNOWLES
ILIAS PERAKIS
YUANZHEN SHAO

A DISSERTATION

Submitted to the faculty of the University of Alabama at Birmingham,
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2023

GENERALIZATION OF THE LIMIT-POINT & LIMIT-CIRCLE CLASSIFICATION TO EQUATIONS WITH DISTRIBUTIONAL COEFFICIENTS

VARUN BHARDWAJ

APPLIED MATHEMATICS

ABSTRACT

This thesis aims to generalize the limit-point & limit-circle classification to a system with complex valued distributional coefficients and re-establish the relationship between the number of linearly independent square integrable solutions and the convergence of the circle consisting of Titchmarsh-Weyl m functions to a point or a circle in the complex plane.

We begin by discussing linear relations, distributions of orders 0 and measures, and ordinary differential equations. Afterwards, we present some preliminary results that will help us generalize the classification of limit-points and limit-circles. With the preliminary results in hand, we develop the limit-point and limit-circle classification, following Weyl and Titchmarsh's original approach.

We conclude by examining the limit-point and limit-circle classification in the absence of existence and uniqueness theorem.

DEDICATION

To my parents who gave me the freedom and opportunity to grow and explore. My eternal gratitude goes out to them for their love and dedication. This is also dedicated to the mathematicians of the past who laid the foundations for modern mathematics. Without their immense contributions during their lifetimes, my research would not have been possible. I owe them a great debt of gratitude and I am humbled by the opportunity to contribute to their legacy.

ACKNOWLEDGEMENTS

My deepest gratitude goes out to Dr. Rudi Weikard, my Ph.D. advisor who has been an exemplary mentor. Through my research journey, he has been an invaluable source of guidance and support. Without his mentorship, I would not be able to achieve what I have achieved. I will always be grateful to him for the opportunities he has given me.

I would like to express my sincere gratitude to the members of my committee for their guidance, patience, and support throughout the process. Finally, I wish to thank the Department of Mathematics at the University of Alabama at Birmingham for providing me with the opportunity to expand my knowledge and skills in mathematics and pursue my research interests.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Jacques Charles-François Sturm and Joseph Liouville published a series of papers in 1830s on second-order linear ordinary differential equation

$$-(p(x)u'(x))' + q(x)u(x) = \lambda w(x)u(x) \tag{1.1}$$

on the compact interval $[a, b]$, with following boundary conditions imposed:

$$\begin{aligned} p(a)u'(a) - hu(a) &= 0, \\ p(b)u'(b) - Hu(b) &= 0. \end{aligned} \tag{1.2}$$

Here the coefficients p, w, q are positive on the compact interval $[a, b]$, h, H are non negative constants, and λ is a real-valued parameter. This problem, in the above generality was first studied by Sturm in a memoir presented to the *Paris Academy of sciences* in September 1833. This was followed by a detailed memoir in the Journal *de Mathématiques Pures et Appliquées* in 1836. In 1837, Sturm and Liouville jointly published another memoir that expanded on Sturm's earlier work. In his initial work, Sturm worked on determining properties of eigenvalues and qualitative behaviour of the eigenfunctions to the boundary value problem (1.1), (1.2). Later, Sturm and Liouville investigated how the family of solutions to the boundary value problem can be used to represent, in a Fourier expansion type manner, a large class of arbitrary functions. Today we call this an eigenfunction expansion. Paper by Jesper Lützen [6] discusses the history of Sturm-Liouville theory in detail.

Searching for qualitative behavior of solutions was one of Sturm and Liouville's most innovative ideas. Prior to 1820, the theory of differential equations was primarily concerned with obtaining analytic solutions. Sturm and Liouville realized the limitations of this approach and realized that properties of solutions could be found directly from equations, even in the absence of analytic expressions. This transformed the theory of differential equation and opened up a whole new field of study concerned with the qualitative properties of solutions to a given differential equation.

Hermann Klaus Hugo Weyl's work is unquestionably responsible for making the most significant progress during the first decade of the twentieth century in regards to Sturm-Liouville theory. In his paper [1], Weyl discussed

$$-(p(x)u'(x))' + q(x)u(x) = \lambda w(x)u(x) \text{ for } x \in [0, \infty) \tag{1.3}$$

with the following conditions on the coefficients and parameter. The functions p, q and w are real-valued continuous functions with $p(x) > 0$ and $w(x) = 1$ for all $x \in [0, \infty)$ and parameter λ is complex-valued. In Weyl's paper, no assumptions have been made about the differentiability of p. Due to this assumption, we must work with the quasi-derivative pu' rather than u' since it is possible that the derivative of u may not exist at any point in the interval under consideration. In addition, Weyl assumes that 0 is a regular endpoint and impose the boundary condition

$$\cos(\alpha)u(0) + \sin(\alpha)p(0)u'(0) = 0, \alpha \in [0, \pi).$$

Here are some key findings from the paper that are relevant to the dissertation. In Chapter I, Weyl introduces the concepts of *limit-point* and *limit-circle* and then demonstrates that, given λ with non-zero imaginary part, there exists at least one non-trivial square integrable solution to the differential equation (1.3), i.e., there exists a solution to (1.3) that is in $\mathcal{L}^2([0, \infty))$. In Chapter II, he shows that in the limit-circle case all the solutions to the differential equation (1.3) are square integrable and in the limit-point case there is only one solution, unique up to a constant multiple, to the differential equation (1.3) that is square integrable. Moreover, he establishes that the classification of limit-point and limit-circle is independent of the parameter λ.

In other words, he shows that if for some particular λ all the solutions are square integrable then, same is going to be true for any other parameter λ. There are some other interesting results in the paper related to spectrum and eigenfunction expansion that we are omitting here as they are not relevant to our discussion. Note that Weyl points out at the end of the paper that the main conclusions of the paper still hold if we assume that w is an positive-valued continuous function on $[0, \infty)$.

Later in the year 1912, Alfred Cardew Dixon in his paper [2], discussed the existence of solutions of (1.1) under the following three assumptions.

(1) The coefficients p, q, and w are assumed to be real-valued functions on the compact interval $[a, b]$.
(2) Functions p^{-1}, q, and w are assumed to be Lebesgue integrable, i.e., $p^{-1}, q, w \in L^1([a, b])$.
(3) p and w are positive almost everywhere on $[a, b]$ with respect to the Lebesgue measure.

The next important development to Sturm-Liouville theory came from Edward Charles Titchmarsh in a series of papers in 1941, see [3], [4], and [5]. He studied (1.3) under the conditions that p, q and w are real-valued continuous functions on $[0, \infty)$ such that $p(x) = w(x) = 1$ for all $x \in [0, \infty)$, and $\lambda \in \mathbb{C}$. It is important to note that Titchmarsh studied both the regular problem, i.e., on the compact interval $[a, b]$ and the singular problem, i.e., on the interval $[0, \infty)$. Titchmarsh focused primarily on the eigenfunction expansions, which we will not discuss here. We will, however, discuss the *Titchmarsh-Weyl m-coefficient* here. It is one of the most important concepts introduced in these papers.

For Titchmarsh to apply the methods he developed for the regular case to the singular case, he had to find a solution χ to (1.3) that satisfies a boundary condition at the right endpoint of the interval under consideration. To resolve this issue Titchmarsh used the existence of Weyl square integrable solution. Consider two solutions $\phi(\cdot, \lambda)$

and $\psi(\cdot, \lambda)$ to (1.3) for $\lambda \in \mathbb{C}$ such that

$$\phi(0, \lambda) = \cos(\alpha), \phi'(0, \lambda) = \sin(\alpha)$$
$$\psi(0, \lambda) = -\sin(\alpha), \psi'(0, \lambda) = \cos(\alpha)$$

for some $\alpha \in [0, \pi)$. It follows that every solution can be written as a linear combination of these two solutions. Weyl proved in [1] that in the limit-circle case both of these solutions will be square integrable and in the limit-point case neither one of these solutions will be square integrable. Titchmarsh in [3] showed that in both of these cases there exists an $m(\lambda)$-coefficient, a Nevanlinna function, such that for all complex λ with non-zero imaginary part, the solution

$$\chi(\cdot, \lambda) = \phi(\cdot, \lambda) + m(\lambda)\psi(\cdot, \lambda)$$

is square integrable. Weyl and Titchmarsh also discovered that the Titchmarsh-Weyl m functions form a circle in the complex plane. The circle converges to a circle when two linearly independent solutions to (1.3) are square integrable. The circle converges to a point when only one linearly independent solution is square integrable. We can therefore use this circle to determine whether or not all the solutions to a given problem are square integrable. Although we could discuss many papers, books, and articles on the Sturm-Liouville theory here, we have covered enough relevant historical developments to meet our current needs. This discussion is heavily influenced by W. Norrie Everitt's paper [10], which is an invaluable resource for anyone interested in learning more about the historical development of these ideas.

In this dissertation we will study the 2×2 system $Ju' + qu = \lambda wu$ on the interval $[0, b_\infty)$ with $0 < b_\infty \leq \infty$. Here u is a vector-valued function such that each entry of u is a *balanced* function of *locally bounded variation* on $(0, b_\infty)$, q and w are Hermitian and non-negative matrices, respectively with entries that are complex-valued Lebesgue-Stieltjes measures (distributions of order 0), and $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Our main objective is to follow in the footsteps of Weyl and Titchmarsh and develop an intuitive understanding of limit-point and limit-circle for our system. Specifically, we want to re-establish, for our system, the relationship between the number of linearly independent square integrable solutions and the convergence of the disk consisting of Titchmarsh-Weyl m functions to a point or a disk in the complex plane.

One might ask following two questions.

(1) Is there a reason to work with a system rather than a second order differential equation?

(2) Is there a purpose behind using distributional coefficients?

The first question can be answered by looking at the Schrödinger equation

$$-\psi'' + v\psi = g$$

on the interval (a, b) with $-\infty \leq a < 0 < b \leq \infty$. Here $v = \delta_0$, the Dirac delta distribution ("function") concentrated at 0. Let

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } u = \begin{pmatrix} \psi \\ \psi' + s\psi \end{pmatrix}.$$

Then

$$Ju' = \begin{pmatrix} -\psi'' - s'\psi - s\psi' \\ \psi' \end{pmatrix}.$$

We can use the differential equation to get

$$Ju' = \begin{pmatrix} g - v\psi - s'\psi - s\psi' \\ \psi' \end{pmatrix}.$$

If we define

$$q = \begin{pmatrix} v + s' - s^2 & s \\ s & -1 \end{pmatrix}, w = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } f = \begin{pmatrix} g \\ 0 \end{pmatrix}$$

Then $Ju' + qu = wf$ is the system representation of the Schrödinger equation. If we replace s with the negative of the Heaviside function with jump at 0, then $s' = -v$. Using this in the expression for q gives us that

$$q = \begin{pmatrix} -s^2 & s \\ s & -1 \end{pmatrix}.$$

Observe that this system still represents the Schrödinger equation with a δ potential, but now all the entries of both q and w are bona fide functions. Therefore, we can look for the solutions to $Ju' + qu = wf$ to get solutions to the Schrödinger equation without having to deal with the Dirac delta distribution.

Here is another such example where system can help us simplify the situation. Consider

$$-(p[\psi' + s\psi])' + sp(\psi' + s\psi) = rg$$

on the interval (a, b) with $-\infty \leq a < b \leq \infty$. If we select

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, u = \begin{pmatrix} \psi \\ p(\psi' + s\psi) \end{pmatrix}, q = \begin{pmatrix} 0 & s \\ s & -1/p \end{pmatrix}, w = \begin{pmatrix} r & 0 \\ 0 & 0 \end{pmatrix}.$$

Then the system $Ju' + qu = wu$ represents the differential equation and every solution of the differential equation is also a solution of the system $Ju' + qu = wu$. However, if s represents a distribution of order 0, such as Dirac delta distribution, then in order to solve the differential equation we must make sense of $(p[\psi' + s\psi])'$, which may now be a distribution of order 1; however, in system form, we can solve such problems without worrying about introducing distributions of order 1. These examples were based on [12].

To answer the second question, we can look at the Kronig-Penney model with Dirac delta potentials. The model describes how electrons behave in crystalline solids. The model takes the form $-\psi(x)'' + V(x)\psi(x) = E\psi(x)$, where ψ is the wave function of electron, E is the energy, and V is the potential that consists of a periodic arrangement of Dirac delta “functions”, Dirac delta distributions. It is a particular example of a Sturm-Liouville equation as can be seen from the model.

An outline of the dissertation follows. First two sections of chapter 2 introduce the reader to linear relations and measures and distributions. Both sections begin with some basic definitions and then move on to discuss advanced topics. The main goal of the linear relations section is to familiarize the reader with linear relations and state some important results. The primary objectives of the section on measures and distributions are to introduce the reader to distributions and describe the relation between Lebesgue-Stieltjes measures and distributions of order 0. As we move into the third section of chapter 2, we will look at how we can use the ideas covered in the first two sections of the chapter to study ordinary differential equations. Our results regarding generalizations of the limit-point and limit-circle classification will be presented in chapter 3.

CHAPTER 2

Background Material

The purpose of this chapter is to provide the reader with some background information that will enhance their understanding of the next chapter. We will discuss linear relations, measures, and distributions of order 0, as well as ordinary differential equations in this chapter. Finally, we will examine how we can apply the ideas presented in the chapter to the Kronig-Penney model with periodic Dirac delta potentials.

1. Linear Relations

We start by describing our Hilbert space. For any Hilbert space, $\mathcal{H}$, we will denote by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ the *scalar product* associated with it. The scalar product has the following three properties:

$$\langle u, \lambda v + w \rangle_{\mathcal{H}} = \lambda \langle u, v \rangle_{\mathcal{H}} + \langle u, w \rangle_{\mathcal{H}}.$$

$$\langle u, v \rangle_{\mathcal{H}} = \overline{\langle v, u \rangle_{\mathcal{H}}}.$$

$$\langle u, u \rangle_{\mathcal{H}} \geq 0 \text{ with equality only if } u = 0.$$

where $u, v, w \in \mathcal{H}$ and $\lambda \in \mathbb{C}$. It is worth pointing out that our scalar products will be *linear* in the second argument. The *norm* that we will consider is the *canonical norm* defined as $||u||_{\mathcal{H}} = \sqrt{\langle u, u \rangle_{\mathcal{H}}}$ for $u \in \mathcal{H}$ and if $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert spaces with the scalar products $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$, respectively, then $\mathcal{H}_1 \times \mathcal{H}_2$ will denote the external direct sum of $\mathcal{H}_1$ and $\mathcal{H}_2$, i.e., a Hilbert space with scalar product $\langle (u, f), (v, g) \rangle_{\mathcal{H}_1 \times \mathcal{H}_2} = \langle u, v \rangle_{\mathcal{H}_1} + \langle f, g \rangle_{\mathcal{H}_2}$. Before we start our discussion of linear relations it is worth mentioning that most of the results mentioned in this section are taken from [14], [9], and [16].

We will now discuss *linear relations.* A linear relation T from $\mathcal{H}_1$ to $\mathcal{H}_2$ is a linear subspace of $\mathcal{H}_1 \times \mathcal{H}_2$. The set of all linear relations from $\mathcal{H}_1$ to $\mathcal{H}_2$ will be denote by $\mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ and if $\mathcal{H}_1 = \mathcal{H}_2$, then to shorten the notation $\mathcal{LR}(\mathcal{H}_1)$ will be used instead of $\mathcal{LR}(\mathcal{H}_1, \mathcal{H}_1)$. The *domain, range, multi-valued part,* and *null space (or kernel)* of $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ are defined by

$$\mathcal{D}_T = \{u \in \mathcal{H}_1 : (u, v) \in T \text{ for some } v \in \mathcal{H}_2\},$$

$$\mathcal{R}_T = \{v \in \mathcal{H}_2 : (u, v) \in T \text{ for some } u \in \mathcal{H}_1\},$$

$$\mathcal{M}_T = \{v \in \mathcal{H}_2 : (0, v) \in T\},$$

$$\mathcal{N}_T = \{u \in \mathcal{H}_1 : (u, 0) \in T\},$$

respectively. Clearly, a linear relation T is an operator precisely if $\mathcal{M}_T$ is trivial. Furthermore, for simplicity, we write $v = Tu$ if T is an operator and $(u, v) \in T$.

If $S, T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$, $K \in \mathcal{LR}(\mathcal{H}_2, \mathcal{H}_3)$, and $\lambda \in \mathbb{C}$, then we define

$$\lambda T = \{(u, \lambda v) : (u, v) \in T\},$$

$$S + T = \{(u, v_S + v_T) : (u, v_S) \in S \text{ and } (u, v_T) \in T\},$$

$$K \circ S = \{(u, w) \in \mathcal{H}_1 \times \mathcal{H}_3 : \exists v \in \mathcal{H}_2 : (u, v) \in S \text{ and } (v, w) \in K\}.$$

It follows that $\mathcal{D}_{\lambda T} = \mathcal{D}_T$ and $\mathcal{D}_{S+T} = \mathcal{D}_S \cap \mathcal{D}_T$. The composition $K \circ S$ will mostly be denoted by KS. The *inverse* of a linear relation $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ will be defined as $T^{-1} = \{(v, u) \in \mathcal{H}_2 \times \mathcal{H}_1 : (u, v) \in T\}$. It is clear from the definition of inverse that every linear relation T has an inverse and that $T^{-1} \in \mathcal{LR}(\mathcal{H}_2, \mathcal{H}_1)$, i.e., the inverse is also a linear relation. The following two propositions give us some elementary properties of linear relations.

Proposition 2.1. Let $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ and $S \in \mathcal{LR}(\mathcal{H}_2, \mathcal{H}_3)$ then

(1) $\mathcal{N}_T \subset \mathcal{N}_{ST} \subset \mathcal{D}_{ST} \subset \mathcal{D}_T$, and
(2) $\mathcal{M}_S \subset \mathcal{M}_{ST} \subset \mathcal{R}_{ST} \subset \mathcal{R}_S$.

(3) $\mathcal{D}_T = \mathcal{R}_{T^{-1}}$ and $\mathcal{R}_T = \mathcal{D}_{T^{-1}}$.

(4) $\mathcal{N}_T = \mathcal{M}_{T^{-1}}$ and $\mathcal{M}_T = \mathcal{N}_{T^{-1}}$.

PROOF. For the first inclusion in claim (1), consider $u \in \mathcal{N}_T$ then $(u, 0) \in T$. As $(0, 0) \in S$, we have that $u \in \mathcal{N}_{ST}$. The other inclusions in the first claim follow directly from definitions. As with claim (1), only the first inclusion needs to be proven in claim (2), since the rest follows from the definitions. Let $v \in \mathcal{M}_S$ then $(0, v) \in S$ and as $(0, 0) \in T$, it follows that $(0, v) \in ST$ and hence $v \in \mathcal{M}_{ST}$. Claims three and four follow directly from definitions. □

PROPOSITION 2.2. Let $R, S, T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$, and $K \in \mathcal{LR}(\mathcal{H}_2, \mathcal{H}_3)$. Then we have following properties:

(1) $S + T = T + S$.

(2) $(R + S) + T = R + (S + T)$.

(3) $(KS)^{-1} = S^{-1}K^{-1}$.

(4) $\lambda(KS) = (\lambda K)S = K(\lambda S)$, where $\lambda \in \mathbb{C} \setminus \{0\}$.

(5) $\lambda(S + T) = \lambda S + \lambda T$, where $\lambda \in \mathbb{C} \setminus \{0\}$.

PROOF. Claim one follows by observing that if $(u, v_S) \in S$ and $(u, v_T) \in T$, then $(u, v_S + v_T) = (u, v_T + v_S)$, a similar argument proves the second claim. For the third claim, observe that $(w, u) \in (KS)^{-1}$ if and only if $(u, w) \in KS$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(u, v) \in S$ and $(v, w) \in K$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(v, u) \in S^{-1}$ and $(w, v) \in K^{-1}$ if and only if $(w, u) \in S^{-1}K^{-1}$.
For the fourth claim, take note that $(u, w) \in (\lambda K)S$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(u, v) \in S, (v, w) \in \lambda K$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(u, v) \in S, (v, \lambda^{-1}w) \in K$ if and only if $(u, \lambda^{-1}w) \in KS$ if and only if $(u, w) \in \lambda(KS)$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(u, \lambda^{-1}v) \in S$, $(\lambda^{-1}v, \lambda^{-1}w) \in K$ if and only if there exists a $v \in \mathcal{H}_2$ such that $(u, v) \in \lambda S, (v, w) \in K$ if and only if $(u, w) \in K(\lambda S)$.

The final claim follows from the fact that $(u, v) \in (S+T)$ if and only if $v = v_S + v_T$ such that $(u, v_S) \in S$ and $(u, v_T) \in T$. This completes the proof. $\square$

As we continue to discuss linear relations, we will introduce two more concepts. First, $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ is called *bounded* if there is a constant $C \geq 0$ such that $||v||_{\mathcal{H}_2} \leq C||u||_{\mathcal{H}_1}$ for all $(u, v) \in T$ and second, a linear relation $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ is called *closed* if it is closed as a linear subspace of $\mathcal{H}_1 \times \mathcal{H}_2$. It is important to note that every linear relation has a closure, which is also a linear relation, but the same is not true for operators, i.e., the closure of an operator may not be an operator. For this reason, we will define a *closable* operator. An operator T is called closable (as an operator) if $\mathcal{M}_T = \{0\}$ implies $\mathcal{M}_{\overline{T}} = \{0\}$, i.e., an operator is closable, if the closure of the operator is also an operator. Here is an example of a non-closable operator.

EXAMPLE 2.3. Consider the operator $T : \mathcal{C}_0^\infty(\mathbb{R}) \subset L^2(\mathbb{R}) \to \mathbb{C}$ defined by $T\varphi = \varphi(0)$, where $\mathcal{C}_0^\infty(\mathbb{R})$ represents the set of compactly supported and infinitely differentiable functions in $\mathbb{R}$. Then T is non-closable.

Consider some function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$ such that $\varphi(0) = 1$ and define $u_n(x) := \varphi(nx)$. By construction, we have $Tu_n = 1 \neq 0$ for all n, whereas

$$||u_n||^2 = \int_{-\infty}^{\infty} |\varphi(nx)|^2 dx = \frac{1}{n}\int_{-\infty}^{\infty} |\varphi(x)|^2 dx = \frac{1}{n}||u_1||^2 \to 0 \text{ as } n \to \infty$$

Therefore, the operator is non-closable.

PROPOSITION 2.4. Let $T : \mathcal{D}_T \subset \mathcal{H}_1 \to \mathcal{H}_2$ be an operator, then

(1) T is *closable* if and only if the following holds: If $n \mapsto u_n$ is a sequence in $\mathcal{D}_T$ such that $u_n \to 0$, and the sequence $Tu_n \in \mathcal{H}_2$ is convergent, then we have $Tu_n \to 0$.

(2) T is *closed* if and only if $n \mapsto u_n \in \mathcal{D}_T$ converges to $u \in \mathcal{H}_1$ and $n \mapsto Tu_n$ converges to $v \in \mathcal{H}_2$, then $u \in \mathcal{D}_T$ and $Tu = v$.

PROOF. It is clear that the first claim is just a restatement of the definition. As for the second claim, we note that $n \mapsto (u_n, Tu_n)$ is a sequence in T that converges to

(u, v) if and only if $u_n \to u$ and $Tu_n \to v$. Therefore, $(u, v) \in T$ if and only if $u \in \mathcal{D}_T$ and $Tu = v$. □

Let us switch gears for a little while and take a closer look at linear operators, beginning with a few definitions. The set of all bounded linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ will be denoted by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. The notation $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_1)$ will be abbreviated as $\mathcal{L}(\mathcal{H}_1)$. If $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then the norm of the operator T is denoted by $||T||$ and is defined as the infimum of all $C \geq 0$ such that $||Tu||_{\mathcal{H}_2} \leq C||u||_{\mathcal{H}_1}$ for all $u \in \mathcal{H}_1$. Next, we will discuss the convergence of operators. Assume that T and T_n for each $n \in \mathbb{N}$ are elements of $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then

- The sequence $n \mapsto T_n$ of linear operators is said to converge *uniformly* or in norm to T if

  $$\lim_{n\to\infty} ||T_n - T|| = 0$$

  we will denoted this by $T_n \rightrightarrows T$.
- The sequence $n \mapsto T_n$ of linear operators is said to converge *strongly* to T if

  $$\lim_{n\to\infty} ||(T_n - T)u||_{\mathcal{H}_2} = 0, \text{ for each } u \in \mathcal{H}_1$$

  we will denoted this by $T_n \to T$.
- The sequence $n \mapsto T_n$ of linear operators is said to converge *weakly* to T if

  $$\lim_{n\to\infty} \langle T_n u, v\rangle_{\mathcal{H}_2} = \langle Tu, v\rangle_{\mathcal{H}_2}, \text{ for each } u \in \mathcal{H}_1 \text{ and } v \in \mathcal{H}_2$$

  we will denote this by $T_n \rightharpoonup T$.

It follows from, $|\langle (T_n - T)u, v\rangle_{\mathcal{H}_2}| \leq ||v||_{\mathcal{H}_2}||(T_n - T)u||_{\mathcal{H}_2} \leq ||v||_{\mathcal{H}_2}||u||_{\mathcal{H}_1}||T_n - T||$ that

$$\text{uniform convergence} \;\Rightarrow\; \text{strong convergence} \;\Rightarrow\; \text{weak convergence}.$$

In general, the reverse implications may not be true unless we are working with finite dimensional Hilbert spaces, as illustrated below.

EXAMPLE 2.5. Throughout this example $\mathbf{0}$ will denote the zero operator in ℓ^2, $\{e_n\}_{n=1}^{\infty}$ will denote the canonical basis for ℓ^2, and $u = \sum_{n=1}^{\infty} u_n e_n, v = \sum_{n=1}^{\infty} v_n e_n \in \ell^2$.

(1) Strong convergence $\not\Rightarrow$ uniform convergence: Consider a sequence of operators $n \mapsto T_n \in \mathcal{L}(\ell^2)$ given by $T_n(u) = u_n e_n$. Then for all $u \in \ell^2, ||T_n(u) - \mathbf{0}(u)||_{\ell^2} = |u_n| \to 0$ as $n \to \infty$. Therefore, T_n converges strongly to $\mathbf{0}$. In contrast, for all $n \in \mathbb{N}, ||T_n - \mathbf{0}|| = ||T_n|| = 1$ and as a result, T_n doesn't uniformly converge to $\mathbf{0}$.

(2) Weak convergence $\not\Rightarrow$ strong convergence: Consider the right shift operator $S \in \mathcal{L}(\ell^2)$, i.e., $S(u) = 0e_1 + \sum_{n=2}^{\infty} u_{n-1} e_n$ and define $T_n = S^n$, then

$$\forall u, v \in \ell^2, |\langle T_n u, v\rangle_{\ell^2}| = \left| \sum_{k=n+1}^{\infty} \overline{u}_{k-n} v_k \right| \leq \sum_{k=n+1}^{\infty} |u_{k-n}||v_k|.$$

By the Cauchy–Schwarz inequality, we get

$$|\langle T_n u, v\rangle_{\ell^2}| \leq \left(\sum_{k=n+1}^{\infty} |u_{k-n}|^2 \right)^{1/2} \left(\sum_{k=n+1}^{\infty} |v_k|^2 \right)^{1/2} \leq ||u||_{\ell^2} \left(\sum_{k=n+1}^{\infty} |v_k|^2 \right)^{1/2} \to 0$$

as $n \to \infty$. Therefore $\langle T_n u, v\rangle_{\ell^2} = \langle \mathbf{0}u, v\rangle_{\ell^2}$, i.e., the sequence weakly converges to $\mathbf{0}$. However,

$$\forall u \in \ell^2, ||(T_n - \mathbf{0})u||_{\ell^2} = ||T_n u||_{\ell^2} = ||u||_{\ell^2}$$

but this precisely means that the sequence $n \mapsto T_n$ fails to strongly converge to $\mathbf{0}$.

The example above and others can be found in [18]. For anyone interested in operator theory, it is a valuable resource. The following proposition offers some conditions that can be checked to determine if a given linear operator is closed.

PROPOSITION 2.6. Let T be a linear operator from $\mathcal{H}_1$ to $\mathcal{H}_2$. Then the following statements hold:

(1) Let T be closable. If $\mathcal{D}_T$ is closed, then T is closed.

(2) Let T be bounded. Then T is closable.

(3) Let T be bounded. Then $\mathcal{D}_T$ is closed if and only if T is closed.

Proof. In order to prove the first claim, consider a Cauchy sequence $n \mapsto u_n$ in $\mathcal{D}_T$. Since $\mathcal{D}_T$ is closed, the sequence converges to some $u \in \mathcal{D}_T$. By definition of $u \in \mathcal{D}_T$, there exists a $v \in \mathcal{H}_2$ such that $Tu = v$. Clearly, the sequence $n \mapsto u_n - u$ is a Cauchy sequence in $\mathcal{D}_T$ converging to 0. Since T is closable, it follows from Proposition 2.4 part (1) that $Tu_n - v = T(u_n - u) \to 0$ or $Tu_n \to v$. This proves the first claim.

As for the second claim, if T is a bounded operator, then $||Tu_n||_{\mathcal{H}_2} \leq C||u_n||_{\mathcal{H}_1}$. Therefore, $u_n \to 0$ implies $Tu_n \to 0$.

Finally, if T is bounded and $\mathcal{D}_T$ is closed, then by the previous two claims, we get that T is closed. Conversely, let T be closed and bounded. Moreover, let $n \mapsto u_n \in \mathcal{D}_T$ be a sequence converging to some $u \in \mathcal{H}_1$. As the operator T is bounded $n \mapsto Tu_n$ is Cauchy and converges to some $v \in \mathcal{H}_2$, but as T is also closed the Cauchy sequence $(u_n, Tu_n) \in T$ converge to $(u, v) \in T$ and thus $Tu = v$ or $u \in \mathcal{D}_T$. □

We conclude our short discussion of linear operators with this proposition and move on to discuss linear relations once more. Let us continue our discussion by providing a few additional definitions. The *adjoint* of a linear relation $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ is defined as

$$T^* = \{(v, g) \in \mathcal{H}_2 \times \mathcal{H}_1 : \forall (u, f) \in T : \langle g, u\rangle_{\mathcal{H}_1} = \langle v, f\rangle_{\mathcal{H}_2}\}.$$

It is important to mention that every $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ has an adjoint and that $T^* \in \mathcal{LR}(\mathcal{H}_2, \mathcal{H}_1)$, i.e., the adjoint is also a linear relation. The *orthogonal complement* of a subset M of the Hilbert space $\mathcal{H}$ is defined as

$$M^\perp = \{u \in \mathcal{H} : \langle u, v\rangle = 0 \text{ for all } v \in M\}.$$

It should be noted that the orthogonal complement of a set is closed, i.e., $M^\perp = \overline{M^\perp}$ and we can write the Hilbert space as a direct sum of $\overline{M}$ and its orthogonal complement, i.e., $\mathcal{H} = \overline{M} \oplus M^{\perp}$. We also define the *boundary operator*

$$\mathcal{U} : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_2 \times \mathcal{H}_1 \text{ by } \mathcal{U}(u, f) = (-if, iu).$$

The adjoint of the boundary operator is given by $\mathcal{U}^* : (v, u) \mapsto (-iu, iv)$. With these definitions it is easy to see that $\mathcal{U}^*\mathcal{U} = \mathbb{1}_{\mathcal{H}_1}$, where $\mathbb{1}_{\mathcal{H}_1}$ denotes the *identity operator* in $\mathcal{H}_1$. The following two propositions give us some properties of the adjoint of a linear relation.

PROPOSITION 2.7. If $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$ then $T^* = \mathcal{U}(T^{\perp}) = (\mathcal{U}T)^{\perp} = (-T^{-1})^{\perp}$.

PROOF. To show the first equality, consider $(v, g) \in T^{\perp}$, then $\mathcal{U}(v, g) = (-ig, iv) \in \mathcal{U}(T^{\perp})$ and for all $(u, f) \in T$, we have that $\langle (v, g), (u, f)\rangle_{\mathcal{H}_1 \times \mathcal{H}_2} = 0$ which is equivalent to $(-ig, iv) \in T^*$. Conversely, consider $(v, g) \in T^*$, then for all $(u, f) \in T, \langle g, u\rangle_{\mathcal{H}_1} = \langle v, f\rangle_{\mathcal{H}_2}$ which implies $\langle (-ig, iv), (u, f)\rangle_{\mathcal{H}_1 \times \mathcal{H}_2} = 0$ or $(v, g) \in \mathcal{U}(T^{\perp})$. This proves $T^* = \mathcal{U}(T^{\perp})$. To prove the second equality, we will show that $T^* = (\mathcal{U}T)^{\perp}$. Let $(u, f) \in T$ then $(-if, iu) \in \mathcal{U}T$ and for any $(v, g) \in T^*$ we have $\langle (v, g), (-if, iu)\rangle_{\mathcal{H}_1 \times \mathcal{H}_2} = i(\langle g, u\rangle_{\mathcal{H}_1} - \langle v, f\rangle_{\mathcal{H}_2}) = 0$, thus $T^* \subset (\mathcal{U}T)^{\perp}$. Conversely, assume $(v, g) \in (\mathcal{U}T)^{\perp}$, then for all $(u, f) \in T, \langle (v, g), (-if, iu)\rangle_{\mathcal{H}_2 \times \mathcal{H}_1} = 0$ which implies $\langle g, u\rangle_{\mathcal{H}_1} = \langle v, f\rangle_{\mathcal{H}_2}$ and this is equivalent to $(v, g) \in T^*$. This proves the second equality. To prove the third and final equality, first observe that $(v, g) \in T^*$ if and only if for all $(u, f) \in T, \langle v, f\rangle_{\mathcal{H}_2} - \langle g, u\rangle_{\mathcal{H}_1} = 0$ which is same as $\langle (v, g), (f, -u)\rangle_{\mathcal{H}_2 \times \mathcal{H}_1} = 0$, but this is exactly true if $(v, g) \in (-T^{-1})^{\perp}$ as $(u, f) \in T$ if and only if $(f, -u) \in -T^{-1}$, and second that if $(v, g) \in (-T^{-1})^{\perp}$ then for all $(u, f) \in T, \langle (v, g), (f, -u)\rangle_{\mathcal{H}_2 \times \mathcal{H}_1} = 0$ as $(f, -u) \in -T^{-1}$, which by the definition of the adjoint means that $(v, g) \in T^*$. □

PROPOSITION 2.8. Let $T, S \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$, then

(1) If $T \subset S$, then $S^* \subset T^*$.
(2) T^* is closed and $T^{**} = \overline{T}$.
(3) $(T^{-1})^* = (T^*)^{-1}$.

Proof. To prove the first claim, consider $(v, g) \in S^*$, then for all $(u, f) \in T \subset S$, we have $\langle g, u\rangle_{\mathcal{H}_1} = \langle v, f\rangle_{\mathcal{H}_2}$ and thus, $(v, g) \in T^*$. The second claim follows from observing that T^* is closed by proposition 2.7 and the orthogonal complement of a set is closed.Therefore, $T^{**} = (\mathcal{U}T^*)^\perp = (\mathcal{U}^*(\mathcal{U}T^\perp))^\perp = (T^\perp)^\perp = \overline{T}$. The final claim follows directly from the definitions. $\square$

The next proposition describes some interesting ways to write a Hilbert space as a direct sum utilizing linear relations.

Proposition 2.9. If $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$, then

(1) $\mathcal{H}_1 \times \mathcal{H}_2 = \overline{T} \oplus T^\perp$.
(2) $\mathcal{H}_2 \times \mathcal{H}_1 = \overline{\mathcal{U}T} \oplus T^*$
(3) $\mathcal{H}_1 = \overline{D_T} \oplus \mathcal{M}_{T^*}$.
(4) $\mathcal{H}_2 = \overline{\mathcal{N}_T} \oplus \mathcal{R}_{T^*}$

Proof. The first claim is a direct consequence of the fact that orthogonal complements are closed. The second claim follows from Proposition 2.7. Claim number three follows immediately after noting that $\mathcal{M}_{T^*} = (\mathcal{D}_T)^\perp$. To see this observe that $g \in \mathcal{M}_{T^*}$ if and only if $\langle g, u\rangle = \langle 0, f\rangle$ for all $(u, f) \in T$ if and only if $g \in (\mathcal{D}_T)^\perp$. The proof of the last claim is analogous to the proof of the previous claim. $\square$

Proposition 2.10. Consider $T \in \mathcal{LR}(\mathcal{H}_1, \mathcal{H}_2)$, then T^* is an operator if and only if $\mathcal{D}_T$ is dense in $\mathcal{H}_1$. Furthermore, if T is an operator, then T is closeable if and only if $\mathcal{D}_{T^*}$ is dense in $\mathcal{H}_2$.

Proof. The first claim is a direct consequence of Proposition 2.9 (3). For the second claim, note that the operator T is closable if and only if $\mathcal{M}_{\overline{T}} = \mathcal{M}_{T^{**}} = \{0\}$ if and only if T^{**} is an operator, i.e., $\mathcal{D}_{T^*}$ is dense in $\mathcal{H}_2$ by the previous claim. $\square$

Here are two important definitions we will use throughout the remainder of this section. A linear relation $T \in \mathcal{LR}(\mathcal{H}_1)$ is called *symmetric* if $T \subset T^*$ and is called *self-adjoint* if $T = T^*$. Given $T \in \mathcal{LR}(\mathcal{H}_1)$, $\mathcal{H}_T$ will denote the closure of $\mathcal{D}_T$. Moreover, if T is closed, then we define $T_0 = T \cap (\mathcal{H}_T \times \mathcal{H}_T)$.

Proposition 2.11. If $T \in \mathcal{LR}(\mathcal{H}_1)$ is self-adjoint, then $T_0 = T \cap (\mathcal{H}_T \times \mathcal{H}_T)$ is a densely defined self-adjoint operator in $\mathcal{H}_T \times \mathcal{H}_T$ with domain $\mathcal{D}_T$.

Proof. We will prove this in two steps. The first step is to demonstrate that T_0 is an operator, and the second step is to demonstrate that T_0 is self-adjoint.

First step, if $(0, v) \in T_0$, then $v \in \mathcal{H}_T \cap \mathcal{M}_T = \overline{\mathcal{D}_T} \cap \mathcal{M}_{T^*} = \{0\}$ by Proposition 2.9 part (3). Therefore, T_0 is an operator. All that remains to prove is that T_0 is self-adjoint. In this respect, consider $(v, g) \in T_0$ then $\mathcal{U}(v, g) \in \mathcal{U}T_0 \subset \mathcal{U}T$. Recall that by Proposition 2.7, $T^* = (\mathcal{U}T)^\perp$, then for all $(u, f) \in T_0 \subset T$ we have $\langle \mathcal{U}(v, g), (u, f)\rangle_{\mathcal{H}_1 \times \mathcal{H}_1} = 0$ if and only if for all $(u, f) \in T_0, \langle g, u\rangle_{\mathcal{H}_1} = \langle v, f\rangle_{\mathcal{H}_1}$. Therefore $T_0 \subset T_0^*$. Now consider some $(u, f) \in T_0^*$ and $(v, g) \in T$. By Proposition 2.9 (3), we can write $g = g_0 + g_\infty$, where $g_0 \in \mathcal{H}_T$ and $g_\infty \in \mathcal{M}_T$. Consequently, $(v, g_0) \in T_0$ and $\langle g_0, u\rangle_{\mathcal{H}_1} - \langle v, f\rangle_{\mathcal{H}_1} = \langle g_0 + g_\infty, u\rangle_{\mathcal{H}_1} - \langle v, f\rangle_{\mathcal{H}_1} = 0$. As $(v, g) \in T$ was arbitrary, we can conclude that $T_0^* \subset T^* = T$. Intersecting both sides by $\mathcal{H}_T \times \mathcal{H}_T$ will give us that $T_0^* \subset T_0$ and our claim has thus been proven. □

The rest of the section is dedicated to closed linear relations and to make our life easier, we will assume, for the rest of this section, that T is a closed linear relation in $\mathcal{H}_1$. Here are some definitions we need for the upcoming discussion.

A complex number λ is called an *eigenvalue* of T if $(u, \lambda u) \in T$ for some non-zero $u \in \mathcal{H}_1$ and the set $\{u \in \mathcal{H}_1 : (u, \lambda u) \in T \text{ or } u = 0\}$ is called the *eigenspace* corresponding to λ. For a complex number λ, we define the *deficiency space*, denoted D_λ, of T at λ as $D_\lambda = \{(u, \lambda u) \in T^*\}$ and the *solvability space*, denoted S_λ, of T at λ as $S_\lambda = \{v \in \mathcal{H}_1 : (u, \lambda u + v) \in T \text{ for some } u \in \mathcal{H}_1\}$. It is clear that $S_\lambda = \mathcal{R}_{T - \lambda \mathbb{1}}$. The following is an important result that pertains to symmetric relations.

PROPOSITION 2.12. If T is a symmetric relation, then all of its eigenvalues are real.

PROOF. Consider $(u, \lambda u) \in T$ for some non-zero u. Then $(u, \lambda u) \in T^*$ because T is symmetric. It follows that

$$\langle u, \lambda u\rangle_{\mathcal{H}_1} = \langle \lambda u, u\rangle_{\mathcal{H}_1}.$$

Since our scalar product is linear in the second argument, we have

$$\lambda ||u||_{\mathcal{H}_1} = \overline{\lambda} ||u||_{\mathcal{H}_1}.$$

For this equality to hold, either $||u||_{\mathcal{H}_1}$ has to be zero or λ has to be real. Since $||u||_{\mathcal{H}_1}$ cannot be zero as u is non-zero. Therefore, we must have that λ is real. □

Going forward, we will not provide proofs of the propositions because they are more technical and long in nature.

The next proposition is Corollary B.3 in [15]. It tells us that the dimensions (D_λ) will not change in the upper and the lower halves of the complex plane. Therefore, we can use $\lambda = \pm i$ to determine the dimensions of (D_λ) in the upper and lower half planes. The *deficiency indices* of T, denoted $n_\pm$, are defined as the dimensions of $D_{\pm i}$ respectively. Here are two more very useful propositions

PROPOSITION 2.13. The dimensions of the deficiency space D_λ is independent of λ as long as λ remains in either the upper or the lower half plane, i.e., $\dim(D_\lambda) = \dim(D_\mu)$ as long as $\text{Im}(\lambda)$ and $\text{Im}(\mu)$ are either both positive or negative.

PROPOSITION 2.14. If $T \in \mathcal{LR}(\mathcal{H}_1)$ is a closed symmetric relation, then $T^* = T \oplus D_i \oplus D_{-i}$.

PROPOSITION 2.15. Suppose $T \in \mathcal{LR}(\mathcal{H}_1)$ is a closed symmetric relation with $d = \dim(D_i \oplus D_{-i}) < \infty$ and that $m \leq d/2$ is a natural number or 0. If $A : T^* \to \mathbb{C}^{d-m}$ is a surjective linear operator such that $T \subset \mathcal{N}_A$ and $A\mathcal{U}A^*$ has rank $d - 2m$, then

$\mathcal{N}_A$ is a closed symmetric extension of T for which $\dim(\mathcal{N}_A \ominus T) = m$. Conversely, every proper closed symmetric extension of T is the kernel of such a linear operator A. Finally, $\mathcal{N}_A$ is self-adjoint if and only if $A\mathcal{U}A^* = 0$, i.e., $m = d/2$.

Proposition 2.14 is taken from [14], see Theorem 2.4.1, and Proposition 2.15 is taken from [15], see Theorem B.5.

For T, a closed linear relation, we define the *resolvent set* as the set of all complex numbers λ for which $\mathcal{N}_{T-\lambda\mathbb{1}} = \{0\}$ and $\mathcal{R}_{T-\lambda\mathbb{1}} = \mathcal{H}_1$. The resolvent set will be denoted by $\rho(T)$.

The complement of the resolvent set, i.e., $\mathbb{C} \setminus \rho(T)$, is called the *spectrum* of T and is denoted by $\sigma(T)$. Clearly, $(u, v) \in (T - \lambda\mathbb{1})^{-1}$ if and only if $(v, \lambda v + u) \in T$, i.e., $u \in S_\lambda$.

Moreover, if $\lambda \in \rho(T)$, then $(T - \lambda\mathbb{1})^{-1}$ is an everywhere defined closed linear operator. We will called it the *resolvent* and denote it by $R_\lambda(T)$ or by R_λ, if there is no room for confusion.

Proposition 2.16. Following are some properties of the resolvent set of T and the resolvent of T.

(1) For $\lambda, \mu \in \rho(T), R_\lambda - R_\mu = (\lambda - \mu)R_\lambda R_\mu = (\lambda - \mu)R_\mu R_\lambda$.
(2) The nullspace $\{u \in \mathcal{H}_1 : R_\lambda u = 0\}$ is independent of $\lambda \in \rho(T)$ and equals $\mathcal{M}_T$. Therefore, the nullspace is $\{0\}$ if and only if T is an operator.
(3) The resolvent set $\rho(T)$ is open.
(4) For $\lambda \in \rho(T)$, the function $\lambda \mapsto R_\lambda$ is analytic in the uniform operator topology as a $\mathcal{L}(\mathcal{H}_1)$-valued function.
(5) If T is self-adjoint, then all the complex numbers λ such that $\mathrm{Im}(\lambda) \neq 0$ are in the resolvent set and for $\lambda \in \rho(T), (R_\lambda)^* = R_{\overline{\lambda}}$.

A proof of the proposition can be found in [14], see Theorem 2.3.3, Theorem 2.3.4, and Theorem 2.3.7. It follows from the last proposition that the spectrum $\sigma(T)$ is a closed set. The spectrum $\sigma(T)$ can be decomposed into three mutually disjoint components: *point spectrum* $\sigma_p(T)$, *continuous spectrum* $\sigma_c(T)$, and *residual spectrum* $\sigma_r(T)$.

- $\sigma_p(T) = \{\lambda \in \mathbb{C} : \mathcal{N}_{T-\lambda\mathbb{1}} \neq \{0\}\}$. This precisely is the set of eigenvalues of T.
- $\sigma_c(T) = \{\lambda \in \mathbb{C} : \mathcal{N}_{T-\lambda\mathbb{1}} = \{0\}$ and $\mathcal{R}_{T-\lambda\mathbb{1}}$ is dense but not closed in $\mathcal{H}_1\}$. In other words, this the set of all $\lambda \in \mathbb{C}$, which are not eigenvalues of T and for which $\overline{S_\lambda} = \mathcal{H}_1$, but S_λ is not closed.
- $\sigma_r(T) = \{\lambda \in \mathbb{C} : \mathcal{N}_{T-\lambda\mathbb{1}} = \{0\}$ and $\mathcal{R}_{T-\lambda\mathbb{1}}$ is not dense in $\mathcal{H}_1\}$. This is the set of all $\lambda \in \mathbb{C}$ which are not eigenvalues of T and for which S_λ is not dense in $\mathcal{H}_1$.

Finally, we conclude our discussion of linear relations with the *Spectral Theorem* for self-adjoint linear relations and Spectral Theorem for densely defined, self-adjoint operators. The theorems have been taken from [14] see Theorems 2.2.9 and 3.1.1.

Before we can state the theorems, we need two more definitions. An operator $P : \mathcal{H}_1 \to \mathcal{H}_1$ is called an *orthogonal projection* if $P^2 = P$ and $P^* = P$. Furthermore, we say that $\{P_{(-\infty,t)}\}_{t\in\mathbb{R}}$ is an increasing family of projection if the function $t \mapsto \langle P_{(-\infty,t)}u, u\rangle_{\mathcal{H}_1}$ is increasing for every $u \in \mathcal{H}_1$.

Proposition 2.17. Let T be a selfadjoint linear relation in $\mathcal{LR}(\mathcal{H}_1)$, then $T \cap (\mathcal{H}_T \times \mathcal{H}_T)$ is a densely defined selfadjoint operator in $\mathcal{H}_T$ with domain $\mathcal{D}_T$.

Proposition 2.18. Let T be a densely defined, selfadjoint operator in $\mathcal{H}_1$. Then there exists a unique family of orthogonal projection operators $\{E_{(-\infty,t)}\}_{t\in\mathbb{R}}$ that are increasing and left-continuous with the following properties:

(1) $E_{(-\infty,-t)} \to \mathbf{0}$ and $E_{(-\infty,t)} \to \mathbb{1}$ as $t \to \infty$, i.e., $E(\emptyset) = \mathbf{0}$ and $E(\mathbb{R}) = \mathbb{1}$.
(2) $T = \int_{-\infty}^{\infty} t dE_{(-\infty,t)}$ in the sense that

$$u \in \mathcal{D}_T \iff \int_{-\infty}^{\infty} t^2 d\langle E_{(-\infty,t)}u, u\rangle_{\mathcal{H}_1} < \infty,$$

and then

$$\langle Tu, v\rangle_{\mathcal{H}_1} = \int_{-\infty}^{\infty} td\langle E_{(-\infty,t)}u, v\rangle_{\mathcal{H}_1} \text{ and } ||Tu||^2_{\mathcal{H}_1} = \int_{-\infty}^{\infty} t^2 d\langle E_{(-\infty,t)}u, u\rangle_{\mathcal{H}_1}.$$

The integrals converges absoloutley.

(3) $TE_{(-\infty,t)} = \overline{E_{(-\infty,t)}T}$.

Recall that the notation $T_n \to T$ represents strong convergence for operators. It is worth mentioning that the operator $E_{(-\infty,t)}$ is called the *spectral projector* for the interval $(-\infty, t)$. If $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$, then for each $u, v \in \mathcal{H}_1$ the function $t \mapsto \langle E_{(-\infty,t)}u, v\rangle_{\mathcal{H}_1}$ is a complex measure on $(\mathbb{R}, \mathcal{B})$. Our next section will discuss some properties of distributions and explore the relationship between measures and distributions of order zero.

2. Distributions and measures

Note: Throughout this section $I = (a, b)$, where $-\infty \leq a < b \leq \infty$, will denote an open interval in $\mathbb{R}$ and K will denote a compact subset of I.

The main resources used to write this section are [14], [8], and [11].

The differential calculus can be extended to include a larger class of objects (called distributions or generalized functions) to enable the handling of functions that are non-differentiable or include sharp discontinuities, such as the Heaviside function, as well as functions whose derivatives do not exist at specific points. In this way, a more robust form of calculus can be developed that can better handle boundary values and discontinuities.

This section has two objectives: To define distributions and discuss some of their properties, and to connect certain distributions to measures. Our particular interest is in *distributions of order* 0 and *Lebesgue-Stieltjes measures.* To begin, we will provide some preliminary definitions, followed by the definition of a distribution. Let u be a complex-valued function on I, then the *support* of u, denoted supp u, is defined as the complement of the largest open set where u vanishes.

$$\text{supp } u = I \setminus \cup\{o \subset I : 0 \text{ is open and } u|_o = 0\}.$$

The space of complex-valued functions defined on I which have derivatives of all orders and are supported on compact subsets of I, is denoted by $\mathcal{D}(I)$. These functions are called *test functions.* With these two preliminary definitions in hand, we can now define a distribution. A linear functional $q : \mathcal{D}(I) \to \mathbb{C}$ is called a *distribution* on I if, for every K, there exists a $C \geq 0$ and a non-negative integer N such that

$$|q(\varphi)| \leq C \sum_{j=0}^{N} \sup\{|\varphi^{(j)}(x)| : x \in K\}$$

whenever the test function φ has its support in K. The set of all distributions on I is denoted by $\mathcal{D}'(I)$. If the non-negative integer N in the above definition can be chosen uniformly for every K, then the distribution q is said to have *finite order* and the smallest such non-negative integer is called the order of q. The set of all distributions on I with order at most N will be denoted by $\mathcal{D}'^N(I)$.

It is obvious that $\mathcal{D}(I)$ is a linear space under pointwise addition and multiplication by scalars as $\text{supp}(\varphi+\psi) \subset (\text{supp } \varphi) \cup (\text{supp } \psi)$ for $\varphi, \psi \in \mathcal{D}(I)$. But a close inspection of the definition of $\mathcal{D}(I)$ forces us to ask whether the combination of the requirements "compactly supported" and "infinitely differentiable" is so restrictive as to only allow the zero function to be in $\mathcal{D}(I)$. Fortunately, this is not the case. In fact, there are infinitely many functions that satisfy both requirements. This means that the definition is broad enough to encompass a wide range of functions. There is a temptation to believe that perhaps analytic functions (functions defined by a converging Taylor expansion) are in $\mathcal{D}(\mathbb{R})$, but an analytic function must be identically zero in order to be in $\mathcal{D}(\mathbb{R})$. Therefore, it becomes necessary to go beyond elementary functions such as sine, cosine, polynomials, etc., in order to discover nontrivial $\mathcal{D}(\mathbb{R})$-functions. In the following example, adapted from Lemma 2.1 in [11], we will construct a non-trivial test function in $\mathcal{D}(I)$ and provide some ways in which known test functions can be used to construct a variety of new test functions.

EXAMPLE 2.19. Consider the function

$$f(t) = \begin{cases} e^{-1/t} & \text{for } t > 0, \\ 0 & \text{for } t \leq 0 \end{cases}$$

then clearly f has derivatives of all orders for $t \neq 0$. At the point $t = 0$, we have that $\lim_{t\downarrow 0} f(t) = 0$ and the n^{th} derivative of f for $t \neq 0$ is given by

$$f^{(n)}(t) = \begin{cases} p_n(t)e^{-1/t} & \text{for } t > 0, \\ 0 & \text{for } t < 0 \end{cases}$$

where $p_n(t)$ is a rational function with $t = 0$ being a pole of order $n+1$. Since $e^{-1/t}$ trumps the pole of the rational function for all n, we have $\lim_{t \downarrow 0} f^{(n)}(t) = 0$ for every non-negative $n \in \mathbb{Z}$, but this means that f has continuous derivatives of all orders. Using the function f, we can finally construct a test function.

Consider $[c, d] \subset I$, then the function

$$\varphi(t) = f(t-c)f(d-t)$$

is a test function with $\operatorname{supp} \varphi \subseteq [c, d]$. With this we have proven the existence of a non-trivial test function. We can construct even more test functions by

- Shifting left or right: $\phi(t + t_0)$.
- Adjusting the vertical scale: $\alpha\phi(t)$.
- Adjusting the horizontal scale: $\phi(\alpha t)$.
- Combining test functions: $\alpha_1\phi_1(t) + \alpha_2\phi_2(t)$.

A function $f : I \to \mathbb{C}$ is called *locally Lebesgue integrable* if f is Lebesgue measurable and $f|_K$ (restriction of f to K), is Lebesgue integrable for every K. The space of all *locally Lebesgue integrable* functions on I is denoted as $L^1_{\text{loc}}(I)$. Now we are ready to present two examples of distributions.

Example 2.20.

- $q_f : \varphi \mapsto \int_I f\varphi$ defines a distribution for $f \in L^1_{\text{loc}}(I)$. To see this, consider $\varphi, \psi \in \mathcal{D}(I)$ such that their supports are in K, and $\alpha, \beta \in \mathbb{C}$. Then it is obvious that

  $$q_f(\alpha\varphi + \beta\psi) = \alpha q_f(\varphi) + \beta q_f(\psi)$$

  and if we let $C = \int_K |f|$, then

  $$|q_f(\varphi)| \leq C \sup\{|\varphi(x)| : x \in K\}.$$

  One may wonder whether the requirement of locally integrability on f can be relaxed. To answer this consider the function $f(x) = 1/|x|$ for $x \in (-1, 0) \cup$

$(0, 1)$ and zero otherwise. Then f is not integrable on any neighborhood of 0. Therefore, we cannot find a constant $C \geq 0$ that satisfies the inequality in the definition for any closed interval containing 0 so f does not define a distribution on $\mathbb{R}$.

- The linear functional $\delta_{x_0} : \varphi \mapsto \varphi(x_0)$, where $x_0 \in I$ and $\varphi \in \mathcal{D}(I)$. The distribution δ_{x_0} is known as the *Dirac delta distribution.*

Having established the existence of non-trivial test functions and examined a few examples of distributions, it is now time to shift gears and further explore the topic; starting with a definition. We say that a sequence $n \mapsto \varphi_n$ in $\mathcal{D}(I)$ tends to zero precisely if the following is satisfied:

(1) There is a fixed compact subset K_0 of I such that for all $n \in \mathbb{N},$ supp $\varphi_n \subseteq K_0$,

(2) For every $m \in \mathbb{N}_0, \varphi_n^{(m)} \to 0$ uniformly.

Using this definition, we say that a sequence $n \mapsto \varphi_n$ of test functions in $\mathcal{D}(I)$ converges to a test function $\varphi \in \mathcal{D}(I)$ if $\varphi_n - \varphi \to 0$ as $n \to \infty$. A linear functional $q : \mathcal{D}(I) \to \mathbb{C}$ is said to be continuous if the sequence $n \mapsto q(\varphi_n)$ converges to $q(\varphi)$ whenever the sequence $n \mapsto \varphi_n$ in $\mathcal{D}(I)$ converges to $\varphi \in \mathcal{D}(I)$.

Proposition 2.21. A linear functional $q : \mathcal{D}(I) \to \mathbb{R}$ is a distribution if and only if it is continuous.

Proof. It is obvious from the definition of a distribution and the definition of convergence in $\mathcal{D}(I)$ that a distribution is a continuous functional. In order to prove that a continuous linear functional is a distribution, we will use the contrapositive. Suppose q is a linear functional on $\mathcal{D}(I)$ which is not a distribution. This means that for some fixed K and every N there exists a test function φ_N with support in K such that

$$|q(\varphi_N)| > N \sum_{j=0}^{N} \sup\{|\varphi_N^{(j)}(x)| : x \in K\}$$

We may assume that $q(\varphi_N) = 1$ for all N since if it is not, we can always multiply φ_N by a constant to ensure that this is true. Thus, $\sup\{|\varphi_N^{(j)}(x)| : x \in K\} < 1/N$ for all

$j \leq N$. It follows that the sequence $N \mapsto \varphi_N$ converges to 0 in $\mathcal{D}(I)$ but $q(\varphi_N) = 1$ for all N and thus q is not continuous. $\square$

We can turn $\mathcal{D}'(I)$ into a vector space by defining $\alpha q + \beta r$ as $(\alpha q + \beta r)(\varphi) = \alpha q(\varphi) + \beta r(\varphi)$ whenever $\alpha, \beta \in \mathbb{C}$, $q, r \in \mathcal{D}'(I)$, and $\varphi \in \mathcal{D}(I)$ and using a similar definition we can turn $\mathcal{D}'^N(I)$ into a vector space for any non-negative integer N. The support of $q \in \mathcal{D}'(I)$, denoted supp q, is defined as

$$\text{supp } q = I \setminus \cup\{o \subset I : o \text{ is open and } q(\varphi) = 0 \text{ for every } \varphi \text{ that has support in } o\}.$$

For any distribution $q \in \mathcal{D}'(I)$ its derivative $q' \in \mathcal{D}'(I)$ is given by

$$q'(\varphi) = -q(\varphi')$$

It should be noted that φ' is also a test function. Clearly, q' is a distribution. To see this consider a test function φ such that supp $\varphi \subset K$. Then

$$|q'(\varphi)| = |q(\varphi')| \leq C_K \sum_{j=0}^{N} \sup\{|\varphi^{(j+1)}(x)| : x \in K\} \leq C_K \sum_{j=0}^{N+1} \sup\{|\varphi^{(j)}(x)| : x \in K\}.$$

From the above inequality, we conclude that $q' \in \mathcal{D}'(I)$. Using induction one can define the n^{th}-derivative of q by $q^{(n)}(\varphi) = (-1)^n q(\varphi^{(n)})$. The following proposition and its proof are adapted from Theorem C.10 in [14]. We also define $||\varphi||_\infty = \sup\{|\varphi(x)| : x \in \text{ supp } \varphi\}$ for $\varphi : D \subseteq I \to \mathbb{R}$.

Proposition 2.22. Every distribution q on I has an antiderivative.

Proof. Fix $\psi \in \mathcal{D}(I)$ with $\int_I \psi = 1$ and define

$$\Phi(x) = \int_a^x \varphi - \left(\int_a^b \varphi\right) \int_a^x \psi$$

Clearly, Φ has derivatives of all orders. To see that Φ is compactly supported in I, note that because ψ and φ are compactly supported in I, there exist $c, d \in I$ such that $\varphi(x) = \psi(x) = 0$ if $x < c$ or if $x > d$. This immediately implies that $\Phi(x) = 0$

whenever $x < c$ and for $x > d$, $\Phi(x) = (\int_I \varphi) - (\int_I \varphi) = 0$. Thus, $\Phi \in \mathcal{D}(I)$ whenever $\varphi \in \mathcal{D}(I)$. If $q \in \mathcal{D}^{'}(I)$, define $Q(\varphi) = -q(\Phi)$. Now, consider a compact subset $K_0 \subset I$ and $\varphi \in \mathcal{D}(I)$ such that supp φ is in K_0, then

$$||\Phi||_\infty \leq m(K_0)(1 + \int_a^b |\psi|)||\varphi||_\infty$$

where $m(K_0)$ is the Lebesgue measure of K_0. Then for all $j > 0$,

$$||\Phi^{(j)}||_\infty \leq ||\varphi^{(j-1)}||_\infty + m(K_0)(||\psi^{(j-1)}||_\infty)||\varphi||_\infty.$$

Based on the previous two inequalities, for each positive integer j, we can find $\alpha_j, \beta_j > 0$ such that $||\Phi^{(j)}||_\infty \leq \alpha_j ||\varphi||_\infty + \beta_j ||\varphi^{(j-1)}||_\infty$ and a $\gamma > 0$ such that $||\Phi||_\infty \leq \gamma ||\varphi||_\infty$. Therefore, for any $\varphi \in \mathcal{D}(I)$ with support in K, we have

$$|Q(\varphi)| = |q(\Phi)| \leq C_K \sum_{j=0}^{N} ||\Phi^{(j)}|| \leq C_k' \sum_{j=0}^{N-1} ||\varphi^{(j)}||.$$

Where $C_k' = \max\{C_k(\gamma + \sum_{j=1}^N \alpha_j), C_k \sum_{j=1}^N \beta_j\}$. It follows that Q is a distribution. Furthermore, $Q'(\varphi) = -Q(\varphi') = q(\varphi)$ because

$$-Q(\varphi') = q\left(\int_a^x \varphi' - \left(\int_a^b \varphi'\right)\int_a^x \psi\right) = q(\varphi).$$

Here we have used the facts that $\int_a^b \varphi' = 0$ as φ is compactly supported and that $\varphi(x) = \int_a^x \varphi'$. □

It follows from the proof that if q is a distribution of order N, then the order of Q is $\max\{0, N-1\}$. Before we proceed any further we need another definiton. A distribution $q \in \mathcal{D}^{'}(I)$ is called a *constant distribution* if there exists a constant c such that $q(\varphi) = c\int_a^b \varphi$ for all $\varphi \in \mathcal{D}(I)$. Now we will prove *The lemma of du Bois-Reymond* or *The fundamental lemma of the calculus of variations*. The version of the lemma presented here and its proof have been adapted from [15].

PROPOSITION 2.23. Suppose $q \in \mathcal{D}^{'}(I)$ be such that the derivative of q is zero. Then q is a constant distribution.

PROOF. Let $\varphi \in \mathcal{D}(I)$ and Φ be defined as in the proof of Proposition 2.22. Then

$$0 = q'(\Phi) = -q(\Phi') = -q\left(\varphi - \psi \int_a^b \varphi\right) = q(\psi)\int_a^b \varphi - q(\varphi)$$

It follows that $q(\varphi) = q(\psi)\int_a^b \varphi = c\int_a^b \varphi$ where $c = q(\psi)$. Therefore q is a constant distribution. □

It is actually possible to extend the result of this proposition. If q is a distribution such that the n^{th} derivative of q is zero then q is a polynomial of degree strictly less than n. We say a distribution q is a polynomial of degree n if for every $\varphi \in \mathcal{D}(I)$ we have

$$q(\varphi) = \int_I \left(\sum_{k=0}^{n} c_k x^k\right) \varphi(x)dx.$$

Here $c_k \in \mathbb{C}$ for every $0 \leq k \leq n-1$ and $c_n \in \mathbb{C} \setminus \{0\}$. To prove this extension observe that if $q^{(2)} = 0$, then $q'(\varphi) = c_1 \int_I \varphi(x)dx$. Now consider $r \in \mathcal{D}'(I)$ given by $r(\varphi) = c_1 \int_I x\varphi(x)dx$. Consequently,

$$\begin{aligned}
(q-r)'(\varphi) &= -q(\varphi') + r(\varphi') \\
&= q'(\varphi) + r(\varphi') \\
&= c_1 \int_I \varphi(x)dx + c_1 \int_I x\varphi'(x)dx \\
&= c_1 \int_I \varphi(x)dx - c_1 \int_I \varphi(x)dx = 0
\end{aligned}$$

Therefore, $(q-r)' = 0$ or $q-r$ is a constant distribution and it follows that $(q-r)(\varphi) = \int_I c_0\varphi(x)dx$ or

$$q(\varphi) = \int_I (c_1x + c_0)\varphi(x)dx.$$

Induction provides the rest of the proof.

Our next step will be to examine how distributions of order 0 relate to Lebesgue-Stieltjes measures. We start by defining some terms. The set of compactly supported and n-times continuously differentiable functions on I will be denoted by $C_c^n(I)$, where n is a non-negative integer. A *Radon functional* on $C_c^0(I)$ is a function $q : C_c^0(I) \to \mathbb{C}$

which is linear and has the property that for every K there exists a constant C such that for every $\varphi \in C_c^0(I)$ with support in K we have

$$|q(\varphi)| \leq C \sup\{|\varphi(x)| : x \in K\}.$$

The set of all Radon functionals on $C_c^0(I)$ can be made into a complex vector space by defining $\alpha q_1 + \beta q_2$ as $(\alpha q_1 + \beta q_2)(\varphi) = (\alpha q_1)(\varphi) + (\beta q_2)(\varphi) = \alpha q_1(\varphi) + \beta q_2(\varphi)$, where q_1 and q_2 are Radon functionals on $C_c^0(I)$, $\varphi \in C_c^0(I)$, and $\alpha, \beta \in \mathbb{C}$. Given a Radon functional q we define its conjugate $\overline{q}$ by $\overline{q}(\varphi) = \overline{q(\overline{\varphi})}$. A Radon functional q is called real if $\overline{q} = q$. We also define $\operatorname{Re}(q)$ and $\operatorname{Im}(q)$ respectively by

$$(\operatorname{Re}(q))(\varphi) = \frac{q(\varphi) + \overline{q}(\varphi)}{2}, (\operatorname{Im}(q))(\varphi) = \frac{q(\varphi) - \overline{q}(\varphi)}{2i}.$$

Hence, $q = \operatorname{Re}(q) + i \operatorname{Im}(q)$ both $\operatorname{Re}(q)$ and $\operatorname{Im}(q)$ are real Radon functionals.

A linear function $q : C_c^0(I) \to \mathbb{C}$ is called a *positive linear functional* if $q(\varphi) \geq 0$ whenever $\varphi \geq 0$. Theorem 2.14 and 2.18 from [7] tells us that if q is a positive linear functional on $C_c^0(I)$, then there exists a unique positive measure μ with following properties:

(1) $q(\varphi) = \int_I \varphi\mu$.
(2) μ is regular and σ-finite.
(3) $\mu(K) < \infty$ for every compact subset K of I.

In the next proposition, we will extend this definition to any linear functional on $C_c^0(I)$.

PROPOSITION 2.24. If q is a Radon functional on $C_c^0(I)$, then there exist a unique positive Radon functionals $|q|$ on $C_c^0(I)$ such that

(1) $|q(\varphi)| \leq |q|(|\varphi|)$ for all $\varphi \in C_c^0(I)$.
(2) $|q|(|\varphi|) \leq \lambda(|\varphi|)$, if $\lambda : C_c^0(I) \to \mathbb{C}$ is any positive linear functional such that $|q(\varphi)| \leq \lambda(|\varphi|)$ for all $\varphi \in C_c^0(I)$.

PROOF. Given a non-negative $\Phi \in C_c^0(I)$, we define $|q|$ as

$$|q|(\Phi) = \sup\{|q(\varphi)| : |\varphi| \leq \Phi\}.$$

First we show that $|q|(\Phi)$ is finite. Suppose that the support of Φ is contained in K. It follows that the support of every φ such that $|\varphi| \leq \Phi$, is also contained in K. Therefore, for all such φ, we have

$$|q(\varphi)| \leq C||\varphi||_\infty \leq C||\Phi||_\infty < \infty.$$

Consequently, $\sup(|q(\varphi)|) \leq C||\Phi||_\infty < \infty$. In other words $|q|(\Phi)$ is finite and $|q|(\Phi) \leq C||\Phi||_\infty$ for all the $\Phi : C_c^0(I) \to [0, \infty)$ that are supported in K. If we can show that $|q|$ is linear then we would have that $|q|$ is a positive Radon functional. The next step is to demonstrate that $|q|$ is linear. It is clear that

$$|q|(c\Phi) = c|q|(\Phi) \text{ for } c \in [0, \infty).$$

Also, given $\epsilon > 0$ we can find $|\varphi_k| \leq \Phi_k$ for $k = 1, 2$ such that

$$|q|(\Phi_1) + |q|(\Phi_2) \leq |q(\varphi_1)| + |q(\varphi_2)| + \epsilon.$$

Furthermore, we can find $\alpha_1, \alpha_2 \in \mathbb{C}$, with the property that $|\alpha_1| = |\alpha_2| = 1$, such that $|q(\varphi_1)| + |q(\varphi_2)| = \alpha_1 q(\varphi_1) + \alpha_2 q(\varphi_2)$. It is clear from the linearity of q and the definition of $|q|$ that $|q|(\Phi_1) + |q|(\Phi_2) \leq |q|(|\alpha_1\varphi_1 + \alpha_2\varphi_2|) + \epsilon$. As $\epsilon > 0$ was arbitrary, we get $|q|(\Phi_1) + |q|(\Phi_2) \leq |q|(\Phi_1 + \Phi_2)$. For the other inequality assume that $|\varphi| \leq \Phi_1 + \Phi_2$ and set

$$\varphi_k(x) = \begin{cases} \Phi_k(x)\varphi(x)/(\Phi_1(x) + \Phi_2(x)) & \text{if } \Phi_1(x) + \Phi_2(x) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

for $k = 1, 2$. Clearly, φ_k is continuous whenever $\Phi_1(x) + \Phi_2(x) \neq 0$ and whenever $\Phi_1(x) + \Phi_2(x) = 0$, then by the assumption that $|\varphi| \leq \Phi_1 + \Phi_2$ we get that $\varphi(x) = 0$.

Since φ is continuous at x and $|\varphi_k| \leq |\varphi|$ we get that φ_k is continuous at x for $k = 1, 2$. Therefore, functions φ_k are in $C_c^0(I)$. Furthermore, $|\varphi_1 + \varphi_2| = |\varphi| \leq \Phi_1 + \Phi_2$ and thus $|q|(\Phi_1 + \Phi_2) \leq |q|(\Phi_1) + |q|(\Phi_2)$. Combining the two inequalities give us

$$|q|(\Phi_1 + \Phi_2) = |q|(\Phi_1) + |q|(\Phi_2).$$

We extend $|q|$ to $C_c^0(I)$ by

$$|q|(\varphi) = |q|((\mathrm{Re}(\varphi))_+) - |q|((\mathrm{Re}(\varphi))_-) + i|q|((\mathrm{Im}(\varphi))_+) - i|q|((\mathrm{Im}(\varphi))_-)$$

where $f_\pm$ are positive and negative parts of the real valued function f. Therefore, $|q|$ is a positive Radon functional such that $|q(\varphi)| \leq |q|(|\varphi|)$.

Now all that is left to do is to prove the uniqueness of $|q|$. If λ is a positive linear functional such that $|q(\varphi)| \leq \lambda(|\varphi|)$, for any $\varphi \in C_c^0(I)$, then we have

$$|q(\varphi)| \leq \lambda(|\varphi|) \leq \lambda(|\Phi|)$$

for any $\varphi \in C_c^0(I)$ such that $|\varphi| \leq |\Phi|$. This means that, $\lambda(|\Phi|)$ is an upper bound of the set for which $|q|(|\Phi|)$ is the supremum. This proves the uniqueness of $|q|$. $\square$

The functional $|q|$ defined in the previous proposition is called the *total variation functional* of q.

Proposition 2.25. Let q be a Radon functional on $C_c^0(I)$, then there exist a unique positive regular Borel measure ν and a measurable function h of absolute value 1 such that

$$q(\varphi) = \int_I \varphi h \nu$$

Additionally, ν is finite on the compact subsets of I.

PROOF. Given a Radon functional q on $C_c^0(I)$, define

$$q_0 = \frac{|\operatorname{Re}(q)| + \operatorname{Re}(q)}{2}, q_1 = \frac{|\operatorname{Im}(q)| + \operatorname{Im}(q)}{2}$$
$$q_2 = \frac{|\operatorname{Re}(q)| - \operatorname{Re}(q)}{2}, q_3 = \frac{|\operatorname{Im}(q)| - \operatorname{Im}(q)}{2}.$$

It follows from Proposition 2.24 that q_k are positive Radon functionals on $C_c^0(I)$. q can be written as a linear combination of q_k, i.e., $q = \sum_{k=0}^{3} i^k q_k$.

In light of what was discussed previously, each q_k is associated with a unique positive σ-finite regular Borel measure μ_k. Define $\mu = \sum_{k=0}^{3} \mu_k$ then μ is also regular σ-finite and finite on compact subsets of I. The measure μ_k is absolutely continuous with respect to the measure μ for each k. Denote the Radon-Nikodym derivatives of μ_k with respect to μ by g_k then by Theorem 1.40 in [7], we have $0 \leq g_k \leq 1$ for each k. Define $g = \sum_{k=0}^{3} i^k g_k$, then

$$q(\varphi) = \sum_{k=0}^{3} i^k q_k(\varphi) = \sum_{k=0}^{3} i^k \int_I \varphi \mu_k = \sum_{k=0}^{3} i^k \int_I \varphi g_k \mu = \int_I \varphi g \mu.$$

Define $h(x) = g(x)/|g(x)|$, whenever $g(x) \neq 0$ and 1 otherwise and define $\nu(E) = \int_E |g|\mu$ for $E \subset I$. Then

$$q(\varphi) = \int_I f g \mu = \int_I f h |g| \mu = \int_I f h \nu.$$

The measure ν and the function h have the desired properties. □

We have shown so far that every Radon functional has a unique positive regular Borel measure associated with it. Proposition C.6 in [14] extend the domain of distributions of order 0 from $D'(I)$ to $C_c^0(I)$. In other words, Proposition C.6 in [14] tells us that distributions of order 0 are the same as Radon functionals. Therefore by Proposition 2.25, if q is a distribution of order 0 then there exists a unique positive regular Borel measure μ and a unique (up to μ-measure zero) function h with $|h| = 1$ such that

$$q(\varphi) = \int_I \varphi h \mu.$$

The last step is to show that the measure μ is a Lebesgue-Stieltjes measure. There is, however, some work to be done before we can establish this relation.

Let $F : I \to \mathbb{R}$ be a non-decreasing function. Let $\mathcal{S} = \{(c,d) \subset I : a < c \leq d < b\}$ and define $L_F : \mathcal{S} \to [0,\infty]$ as $L_F((c,d)) = F^-(d) - F^+(c)$ whenever $c < d$ and $L_F(\emptyset) = L_F((c,c)) = 0$. Then

$$dF^*(A) = \inf\{\sum_{j=1}^{\infty} L_F((c_j,d_j)) : (c_j,d_j) \in \mathcal{S}, A \subset \cup(c_j,d_j)\}$$

is an outer measure and by Caratheodory's Theorem (Theorem 1.11, [17]) induces a complete positive measure dF. The measure dF is called a *positive Lebesgue-Stieltjes measure* on I. It is defined on the Borel σ-algebra regular and finite on compact subsets of I. Positive Lebesgue-Stieltjes measures form a vector space. To see this consider another such non-decreasing function G, then for $\alpha, \beta \in [0,\infty)$ we have

$$\begin{aligned} d(\alpha F + \beta G)((c,d)) &= (\alpha F + \beta G)^-(d) + (\alpha F + \beta G)^+(c) \\ &= \alpha(F^-(d) - F^+(c)) + \beta(G^-(d) - G^+(c)) \\ &= \alpha dF((c,d)) + \beta dG((c,d)). \end{aligned}$$

This argument along with the outer regularity gives us the desired result on the Borel σ-algebra on I. It is worth noting that if the function F is the identity function on I, then the associated Lebesgue-Stieltjes measure is the Lebesgue measure.

For a function $Q : I \to \mathbb{C}$ the *variation* of Q over I is

$$\mathrm{Var}_Q(I) = \sup\left\{\sum_{j=1}^{n} |Q(x_j) - Q(x_{j-1})| : x_j \in I \text{ and } x_{j-1} < x_j \text{ for } 1 \leq j \leq n\right\}.$$

If $\mathrm{Var}_Q(I)$ is finite, we say that Q is of bounded variation on I, denoted $\mathrm{BV}(I)$. For $Q \in \mathrm{BV}(I)$, we define the *variation function* as

$$V_Q(x) = \begin{cases} \mathrm{Var}_Q([c, x]), & x > c \\ 0, & x = c \\ -\mathrm{Var}_Q([x, c]), & x < c \end{cases}$$

According to the Lemma 3.26, Theorem 3.27, and the discussion immediately before Lemma 3.26 in [17], the functions $V_Q, V_Q \pm Q$ are non-decreasing and $Q \in \mathrm{BV}(I)$ if and only if $\mathrm{Re}\, Q, \mathrm{Im}\, Q \in \mathrm{BV}(I)$. As a result, a function $Q \in \mathrm{BV}(I)$ can be written as a linear combination of four non-decreasing functions, i.e., $Q_k = \sum_{k=0}^{3} i^k Q_k$, where each Q_k is non-decreasing. Each Q_k being non-decreasing is associated with a finite positive Lebesgue-Stieltjes measure dQ_k. The measure is finite because $d_{Q_k}(I) = Q_k^-(b) - Q_k^+(a) < \infty$ as each Q_k is bounded on I. Therefore, we can associate $Q \in \mathrm{BV}(I)$ with a complex Lebesgue-Stieltjes measure defined as

$$dQ = \sum_{k=0}^{3} i^k dQ_k.$$

If $\mathrm{Var}_Q(K) < \infty$ for every compact subintervals K of I, then Q is said to be of *locally bounded variation* on I. The set of functions of locally bounded variation on I is denoted by $\mathrm{BV}_{\mathrm{loc}}(I)$, which is a vector space. As every function of locally bounded variation is a function of bounded variation on compact subintervals of I, so every $Q \in \mathrm{BV}_{\mathrm{loc}}(I)$ generates a finite complex Lebesgue-Stieltjes measure dQ on any compact subinterval K of I. While dQ is a bonafide measure on compact subsets of I, it may not necessarily be a measure on I; for this reason, it is called a *local measure*.

Suppose that μ is a Borel measure on I such that $\mu(K) < \infty$ for every K. Let $c \in I$ be fixed, then the function

$$F_\mu(x) = \begin{cases} \mu([c,x)), & x > c \\ 0, & x = c \\ -\mu([x,c)), & x < c \end{cases}$$

is called a *cumulative distribution function* (cdf). The function F_μ is of locally bounded variation on I. If $|\mu(I)| < \infty$, then it is customary to define $F_\mu(x) = \mu((a,x))$ and if μ is a positive measure, then F_μ is a non-decreasing function. By Theorem 1.16, [17], if $E \subset K$ is in the Borel σ-algebra on I, i.e., if E is a Borel set that is compactly contained in I, then

$$\mu(E) = dF_\mu(E) = \int_I \chi_E dF_\mu$$

where χ_E is the *characteristic function* or the *indicator function* of E.

Now for a function $Q \in$ BV$_{loc}(I)$ define $q_Q : \mathcal{D}(I) \to \mathbb{C} : \varphi \mapsto \int_I \varphi dQ$. Consider, $\varphi \in \mathcal{D}(I)$ such that the support of φ is in K, then

$$|q_Q(\varphi)| = \left| \int_I \varphi dQ \right| \leq C||\varphi||_\infty, \text{where } C = |dQ(K)|.$$

In other words, a function of locally bounded variation can be associated with a distribution of order 0. Therefore, we have shown the following theorem.

Proposition 2.26. Given $q \in \mathcal{D}'^0(I)$, there exist a function of locally bounded variation Q such that

$$q(\varphi) = \int_I \varphi dQ.$$

Conversely, a function of locally bounded variation Q on I give rise to a $q \in \mathcal{D}'^0(I)$.

We will conclude this section with following two important results. First, if q is a distribution of order 0 and Q is the corresponding $\mathrm{BV_{loc}}(I)$ function, then for any

$f \in L^1_{\text{loc}}(|dQ|)$ the functional

$$fq = qf : \mathcal{D}(I) \to \mathbb{C} : \varphi \mapsto \int_I f\varphi dQ$$

is also a distribution of order 0 and $\overline{qf} = \overline{q}\overline{f}$. Second, Lemma A.3 in [15], tells us that if F and G are $\mathrm{BV}_{\text{loc}}(I)$, then

$$\int_{[c,d)} (F^+ dG + G^- dF) = (FG)^-(d) - (FG)^-(c)$$

whenever $[c, d) \subset I$. Note that we can use the above expression to write the product rule for functions of locally bounded variation:

$$(FG)' = F^+G' + F'G^-. \tag{2.1}$$

Therefore, for a $Q \in \mathrm{BV}_{\text{loc}}(I)$, the distribution $Q(\varphi) = \int_I \varphi Q$ is an antiderivative of the distribution $q(\varphi) = \int_I \varphi dQ$. In order to see this, consider

$$Q(\varphi) = \int_I \varphi Q = \int_I \varphi(x)Q(x)dx.$$

Then,

$$Q'(\varphi) = -Q(\varphi') = -\int_I \varphi'(x)Q(x)dx = \int_I \varphi(x)Q'(x)dx = \int_I \varphi dQ.$$

This concludes our section on distributions and measures. In the next section, we will explore how the concepts that we have discussed so far can be applied to the study of ordinary differential equations.

3. Ordinary differential equations

Note: Throughout this section $I = (a, b)$, where $-\infty \le a < b \le \infty$, will denote an open interval in $\mathbb{R}$.

We will begin the section by introducing the reader to the system of ordinary differential equations, whose coefficients are distributions of order 0. Then we will discuss some properties of such a system, such as the existence and uniqueness of solutions to initial value problems. Finally, we will look at an example to conclude the section.

We start by discussing *matrix-valued* and *vector-valued* distributions. Let M be a set of numbers, functions, or operators, then $M^{m\times n}$ will denote the set of $m \times n$ matrices whose entries are elements of M. If $n = 1$, then we will write M^m instead of $M^{m\times 1}$. If $u \in \mathbb{C}^m$ is a column vector, then $u^* \in \mathbb{C}^{1\times m}$ will denote the row vector whose entries are complex conjugates of the entries of u.

If M is a normed vector space, then we will equip $M^{m\times n}$ with 1-norm, i.e., for $U \in M^{m\times n}$,

$$|U|_1 = \sum_{i=1}^{m}\sum_{j=1}^{n} |U_{i,j}|.$$

Using this definition we can define the variation, Var_Q and variation function V_Q for a matrix-valued function U by replacing absolute values in the original definitions by $|\cdot|_1$. For example, the variation of a matrix-valued function U is given by

$$\begin{aligned}\mathrm{Var}_U(I) &= \sup\{\sum_{j=1}^{k} |U(x_j) - U(x_{j-1})|_1 : x_0 < x_1 < \cdots < x_k, x_j \in I\} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \mathrm{Var}_{U_{ij}}(I).\end{aligned}$$

Therefore, the definitions of bounded and locally bounded variation for such matrix-valued functions are analogous to those for scalar-valued functions. Furthermore, if

$Q \in \mathrm{BV}_{\mathrm{loc}}(I)^{m\times n}$ is left-continuous, then

$$\sup_{x\in I}\{|Q^+(x) - Q^-(x)|_1\} \le \sum_{x\in I} |Q^+(x) - Q^-(x)|_1 \le \mathrm{Var}_Q(I) = \int_I dV_Q.$$

Let $q \in D'^0(I)^{m\times n}$, therefore each $q_{i,j} \in D'^0(I)$. Then by Proposition 2.26 each $q_{i,j}$ is associated with a $f_{i,j} \in \mathrm{BV}_{\mathrm{loc}}(I)$ such that

$$q_{i,j}(\varphi) = \int_I \varphi df_{i,j}$$

The equality will remain if we change $f_{i,j}$ up to a constant because $df_{i,j} = d(f_{i,j} + c)$. Therefore, for $q \in D'^0(I)^{m\times n}$, define $Q \in \mathrm{BV}_{\mathrm{loc}}(I)^{m\times n}$ by $Q_{i,j} = f_{i,j}$ and $(dQ)_{i,j} = dQ_{i,j} = df_{i,j}$. It follows that

$$q(\varphi) = \int_I \varphi dQ$$

where $\int_I \varphi dQ$ is defined componentwise, i.e., $(\int_I \varphi dQ)_{i,j} = \int_I \varphi dQ_{i,j}$. As the distribution $Q(\varphi) = \int_I \varphi Q$ is an antiderivative of the distribution $q(\varphi) = \int_I \varphi dQ$, therefore, we will call Q an antiderivative of q. If $Q \in \mathrm{BV}_{\mathrm{loc}}(I)^{m\times n}$ is an antiderivative of $q \in D'^0(I)^{m\times n}$, then we will not distinguish between dQ, Q', and q, i.e.,

$$q(\varphi) = \int_I \varphi dQ = \int_I \varphi Q' = \int_I \varphi q.$$

We will define

$$\Delta_q(x) = Q^+(x) - Q^-(x).$$

For $q \in D'^0(I)^{m\times n}$, we define $q^* \in D'^0(I)^{n\times m}$ by $(q^*)_{i,j} = \overline{q_{j,i}}$. Using this definition we can define a *Hermitian distribution*. A distribution $q \in D'^0(I)^{n\times n}$ is called Hermitian if $q = q^*$. The distribution q is called a non-negative distribution, if u^*qu is a non-negative distribution for every $u \in \mathbb{C}^n$.

By Appendix A.3 of [15], every non-negative distribution q must also be a Hermitian distribution. Its diagonal elements and therefore, its trace must also be non-negative distributions, and its antiderivative Q must be non-decreasing, i.e., u^*Qu is non-decreasing for every $u \in \mathbb{C}^n$. Each $dQ_{i,j}$ is absolutely continuous with respect to

$d\,\mathrm{tr}(Q)$, where $\mathrm{tr}(Q)$ is the trace of Q. Let $\tilde{Q}$ denote the matrix of Radon-Nikodym derivative of the entries of Q with respect to $\mathrm{tr}(Q)$, i.e., $\tilde{Q}_{i,j} = (dQ_{i,j}/d\,\mathrm{tr}(Q))$. Then $q = \tilde{Q}d\,\mathrm{tr}(Q)$ and $\mathbf{0} \le \tilde{Q} \le \mathrm{tr}(\tilde{Q})\mathbb{1} = \mathbb{1}$, where the equality is true almost everywhere with respect to $d\,\mathrm{tr}(Q)$.

If $f, g : I \to \mathbb{C}^n$ and $w \in D'^0(I)^{n\times n}$ is non-negative, then

$$\langle f, g\rangle = \int_I f^* w g = \int_I f^* \tilde{W} g d\,\mathrm{tr}(W)$$

defines a semi-inner product. This semi-inner product give rise to a pseudo-norm $||f|| = \sqrt{\langle f, f\rangle}$. If we denote by $\mathcal{L}^2(w)$ the set of all $\mathbb{C}^n$-valued functions on I such that $\int_I f^* w f = \int_I f^* \tilde{W} f d\,\mathrm{tr}(W) < \infty$ for all $f : I \to \mathbb{C}^n$. $\mathcal{L}^2(w)$ can be turned into a Hilbert space $L^2(w)$ by identifying functions whose differences have norm 0. Thus, $L^2(w)$ consist of equivalent classes of functions in $\mathcal{L}^2(w)$ and f, g belong to the same equivalent class if $||f - g|| = 0$.

We now have all the information that we need to study the following system of differential equations on the interval I.

$$Ju' + qu = wf \tag{2.2}$$

where J is a constant, invertible and skew-Hermitian $n\times n$-matrix and $q, w \in D'^0(I)^{n\times n}$ are such that w is non-negative, and q Hermitian.

It is important to note here that each term in the system is a distribution of order 0. Let us go through each term to understand this. We start with the term Ju'. Ju' is a column vector and each entry of Ju' is a linear combination of entries of the vector u'. The solution u that we are looking for a vector-valued function such that each component of u is of locally bounded variation on I. Therefore, each component of Ju' can be thought of as a distribution of order 0 in the following sense.

$$(Ju')_i(\varphi) = \sum_{j=1}^{n} J_{i,j} \int_I \varphi du_j$$

The next is the term qu, which is also a column vector. We already know that each entry of q is in $D'^0(I)$. The i^{th} entry of qu looks like

$$(qu)_i = \sum_{j=1}^{n} q_{i,j} u_j.$$

Each entry of u is $\mathrm{BV}_{\mathrm{loc}}(I)$ and therefore for every $1 \leq i, j, k \leq n, u_k \in L^1_{\mathrm{loc}}(|dQ_{i,j}|)$. Therefore each entry of qu is in $D'^0(I)$ by using the definition

$$(qu)_i(\varphi) = \sum_{j=1}^{n} \int_I u_j \varphi dQ_{i,j}.$$

Similarly, each entry of wf can be thought of as a distribution of order 0.

Now we can go back to existence and uniqueness of solutions. If the matrices

$$B_{\pm}(x) = J \pm \frac{1}{2}\Delta_q(x) \tag{2.3}$$

are invertible for every $x \in I$, then by Theorem 2.2 in [15], the initial value problem

$$\begin{cases} Ju' + qu = wf \\ u(x_0) = u_0 \in \mathbb{C}^n, \quad x_0 \in I \end{cases} \tag{2.4}$$

has a unique *balanced* solution of locally bounded variation on I. We call a function $u : I \to \mathbb{C}^n$ balanced if $u(x) = (u^+(x) + u^-(x))/2$ for all $x \in I$. The set of all such functions of locally bounded variation on I is denoted by $\mathrm{BV}^{\#}_{\mathrm{loc}}(I)^n$. At a point of discontinuity of Q or W, the differential equation requires that

$$B_+(x)u^+(x) - B_-(x)u^-(x) = \Delta_w(x)f(x).$$

Next, we describe the notion of *definiteness condition*. We say that the definiteness condition holds for the problem if

$$\mathcal{L}_0 = \{u \in \mathrm{BV}^{\#}_{\mathrm{loc}}(I)^n : Ju' + qu = 0, wu = 0\}$$

is trivial.

Now finally, we will look at the Dirac-Kronig-Penney model to see how some of the ideas we have discussed so far can help us solve differential equations.

EXAMPLE 2.27. **Dirac-Kronig-Penney model:** In 1931, Ralph Kronig and William Penney developed a one dimensional mathematical model based on the work of Felix Bloch to study the behavior of electrons in periodic potentials. The model is commonly used in condensed matter physics to study the behavior of electrons in solids, such as metals, semiconductors, and insulators. The Dirac-Kronig-Penney model is a special case of the Kronig-Penney model obtained by replacing rectangular potentials with Dirac delta potentials.

Schrödinger's equation corresponding to the Dirac-Kronig-Penney model is

$$-\psi'' + \sum_{n\in\mathbb{N}} V_n \delta_n \psi = E\psi$$

where $V_n, E \in \mathbb{R}$ for each n, the interval under consideration here is $(0, \infty)$, and we impose a boundary condition at the origin

$$\cos(\alpha)\psi(0) + \sin(\alpha)\psi'(0) = 0, \alpha \in [0, \pi).$$

Here $\delta_n : \mathcal{D}(\mathbb{R}) \to \mathbb{C} : \varphi \mapsto \varphi(n)$ is the Dirac delta distribution concentrated at n. The solution ψ also has to satisfy the normalization condition

$$\int_{(0,\infty)} |\psi|^2 = 1.$$

Our potentials are located on all the positive integer along the half-line. We first convert the second order differential equation along with the boundary condition to a boundary value problem (BVP) for a 2×2 system:

$$J\Psi' + q\Psi = Ew\Psi$$

$$\begin{pmatrix}\cos(\alpha) & \sin(\alpha)\end{pmatrix} \Psi(0) = 0$$

where

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \Psi = \begin{pmatrix} \psi \\ \psi' \end{pmatrix}, q = \begin{pmatrix} \sum_{n\in\mathbb{N}} V_n\delta_n & 0 \\ 0 & -1 \end{pmatrix}, \text{ and } w = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

The normalization condition changes to

$$\int_{(0,\infty)} \Psi^* w \Psi = 1.$$

Our goal is to find a $\Psi \in \mathrm{BV}^{\#}_{loc}((0,\infty))^2$ that solves the BVP.

We will first see how our solution Ψ looks like on the intervals $(k, k+1)$ for $k \in \mathbb{N}_0$. If $\varphi_k \in \mathcal{D}((0,\infty))$ with supp $(\varphi_k) \subset (k, k+1)$, then

$$\begin{aligned} 0 &= J\Psi'(\varphi_k) + (q\Psi)(\varphi_k) - (Ew\Psi)(\varphi_k) \\ &= \begin{pmatrix} -\Psi_2' + \sum_{n\in\mathbb{N}_0} V_n\Psi_1\delta_n - E\Psi_1 \\ \Psi_1' - \Psi_2 \end{pmatrix} (\varphi_k) \end{aligned}$$

Observe that the distribution $\sum_{n\in\mathbb{N}} V_n\Psi_1\delta_n$ and the function φ_k have disjoint supports. Therefore, $(\sum_{n\in\mathbb{N}_0} V_n\Psi_1\delta_n)(\varphi_k) = 0$ and thus, we must have that

$$\begin{pmatrix} -\Psi_2' - E\Psi_1 \\ \Psi_1' - \Psi_2 \end{pmatrix} (\varphi_k) = \begin{pmatrix} \int_k^{k+1} (-\Psi_2' - E\Psi_1)\varphi_k \\ \int_k^{k+1} (\Psi_1' - \Psi_2)\varphi_k \end{pmatrix} = 0.$$

In order for this equality to hold for every $\varphi_k \in \mathcal{D}([0,\infty))$ with supp $(\varphi_k) \subset (k, k+1)$, it is essential that $\Psi_2' = -E\Psi_1$ and $\Psi_1' = \Psi_2$ on any interval $(k, k+1)$. Clearly, the second equality holds. The first equality gives us that the solution on the interval $(k, k+1)$ is given by

$$\Psi(x) = \begin{pmatrix} A_k e^{i\sqrt{E}x} + B_k e^{-i\sqrt{E}x} \\ i\sqrt{E}(A_k e^{i\sqrt{E}x} - B_k e^{-i\sqrt{E}x}) \end{pmatrix}.$$

Using the boundary condition at zero, we can find the relation between A_0 and B_0.

$$B_0 = -A_0 \frac{\cos(\alpha) + i\sqrt{E}\sin(\alpha)}{\cos(\alpha) - i\sqrt{E}\sin(\alpha)}.$$

The next step is to find the relation between A_{k+1}, B_{k+1} and A_k, B_k, so that we can find the coefficients A_{k+1}, B_{k+1}, provided we already know all the coefficients on previous intervals. To find this relation, notice that $x = k+1$ is a point of discontinuity of Q for $k \in \mathbb{N}$. Thus, at $x = k+1$, we must have that

$$B_+(k+1)\Psi^+(k+1) = B_-(k+1)\Psi^-(k+1).$$

To find the expressions for $B_\pm(k+1)$ in this case, it is useful to rewrite our differential equation as $J\Psi' + (q - Ew)\Psi = 0\Psi$ and compare it to (2.2). Equation (2.3) can now be used to get the modified expressions for $B_\pm(k+1)$, which are given by

$$B_\pm(k+1) = J \pm \frac{1}{2}(\Delta_q(k+1) - E\Delta_w(k+1)) = \frac{1}{2}\begin{pmatrix} \pm V_{k+1} & -2 \\ 2 & 0 \end{pmatrix}$$

where

$$\Delta_q(k+1) = \begin{pmatrix} V_{k+1} & 0 \\ 0 & 0 \end{pmatrix}, \Delta_w(k+1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Solving this equation gives us

$$A_{k+1} = \frac{2\sqrt{E}A_k - i(A_k + B_k e^{-2i\sqrt{E}(k+1)})V_{k+1}}{2\sqrt{E}}$$

$$B_{k+1} = \frac{2\sqrt{E}B_k + i(A_k e^{2i\sqrt{E}(k+1)} + B_k)V_{k+1}}{2\sqrt{E}}.$$

The example we have just presented brings the chapter to a close. The next chapter will discuss generalizing the concept of limit-point and limit-circle to a 2×2 system.

CHAPTER 3

Generalization of the limit-point and limit-circle classification

In this chapter we will generalize the concept of limit-point and limit-circle classification to the first-order system $Ju' + qu = \lambda wu$ of differential equations on the real interval $[0, b_\infty), 0 < b_\infty \leq \infty$, where $q, w \in \mathcal{D}'^0((0, b_\infty))^{2\times 2}$. J is the constant skew-Hermitian matrix given by

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The following definitions will also be used throughout this chapter.

Definition 3.1.

- $U(\cdot, \lambda)$ is a fundamental matrix for the system $Ju' + qu = \lambda wu$.
- $B_\pm(x, \lambda) = J \pm \frac{1}{2}(\Delta_q(x) - \lambda \Delta_w(x))$, where $\Delta_r(x) = R^+(x) - R^-(x)$ when R is an antiderivative of r.
- $\Lambda = \{\lambda \in \mathbb{C} : \det(B_+(x, \lambda)) \det(B_-(x, \lambda)) = 0 \text{ for some } x \in (0, b_\infty)\}$.
- If m is a measure, then m_c and m_d, respectively, will denote its continuous and discrete parts.
- If $u : (0, b_\infty) \to \mathbb{C}^2$ is a w-measurable function, $f \in \mathbb{C}^2$, and $S \in \mathbb{C}^{2\times 2}$, then
  - $||u||_x = \sqrt{\int_{[0,x)} u^* wu}$, $||u|| = \sqrt{\int_{[0,b_\infty)} u^* wu}$.
  - $|f| = \sqrt{f^* f}$.
  - $S_{\cdot,j} =$ column j of S, $S_{j,\cdot} =$ row j of S.
- $\mathcal{L}^2(w) = \{u \in \mathrm{BV}_{\mathrm{loc}}((0, b_\infty))^2 : ||u|| < \infty\}$.

We start by defining the terms *definiteness condition* and *regular endpoint*. We say that the definiteness condition holds if the set

$$\mathcal{L}_0 = \{u \in \mathrm{BV}^{\#}_{\mathrm{loc}}((0, b_\infty))^2 : Ju' + qu = 0\}$$

is trivial, i.e., $\mathcal{L}_0 = \{0\}$. If we are looking at the system $Ju' + qu = \lambda wu$ on the interval (a, b), then we say that a is regular endpoint if the antiderivatives Q and W of q and w are of bounded variation on (a, c) for some $c \in (a, b)$. Similarly, b is a regular endpoint if the antiderivatives Q and W are of bounded variation on (c, b) for some $c \in (a, b)$.

Unless otherwise stated, the remainder of this chapter will require the following hypothesis.

HYPOTHESIS 3.2. w is non-negative and q is Hermitian, $\lambda \notin \Lambda$, 0 is a regular endpoint and a point of continuity of antiderivatives of q and w, and the definiteness condition holds.

1. Preliminary results

We will present some preliminary results in this section that are required to generalize the limit-point and limit-circle classification, which we will cover in the next section. There are two main results in this section. First is Theorem 3.5, which describes the relation between $u(\cdot, \lambda)$ a balanced solution to $Ju' + qu = \lambda wu$ on $(0, b_\infty)$ and $u(\cdot, \overline{\lambda})$ a balanced solution to $Ju' + qu = \overline{\lambda} wu$. Second is Theorem 3.6 that tells us that, as long as a certain condition is met, if every solution of

$$Ju' + qu = \lambda_0 wu$$

is square integrable for some $\lambda_0 \notin \Lambda$, then for any $\lambda \notin \Lambda$ every solution of

$$Ju' + qu = \lambda wu$$

will also be square integrable.

LEMMA 3.3. Let $n \mapsto x_n$ be a sequence of distinct points in $(0, b_\infty)$ then

$$P(x,\lambda) = \prod_{x_n \in (0,x)} \frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)}$$

exists for every $x \in [0, b_\infty)$ and this product converges absolutely.

It should be noted that the empty product is defined as 1.

PROOF. It is important to note that the definition of $P(x,\lambda)$ make sense because $\det B_+(x,\lambda) \neq 0$ for any $x \in (0, b_\infty)$ as $\lambda \notin \Lambda$. The cases when $x = 0$ or when the set $\{x_n : x_n \in (0,x)\}$ is empty are trivial because 0 is assumed to be a point of continuity of antiderivatives of q and w and the empty product is defined as 1. Otherwise, let $x \in (0, b_\infty)$ be fixed. Throughout the following analysis, we will exclusively focus on x_n, which are strictly smaller than x. By Theorem 15.4 [7] it is sufficient to show that $\sum_{x_n \in [0,x)} \left| \frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)} - 1 \right|$ is convergent. Clearly, we only need to deal with those x_n which are in the interval $(0,x)$ and for which at least one of the $\Delta_q(x_n)$ or $\Delta_w(x_n)$ is non-zero, because if $\Delta_q(x_n) = \Delta_w(x_n) = 0$, then $\det B_\pm(x_n,\lambda) = 1$. Define $A(x_n) := \frac{1}{2} J(\Delta_q(x_n) - \lambda \Delta_w(x_n))$, then $\det B_\pm(x_n,\lambda) = \det(\mathbb{1} \mp A(x_n))$ and

$$\left| \frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)} - 1 \right| = 2 \left| \frac{A_{11}(x_n) + A_{22}(x_n)}{\det(\mathbb{1} - A(x_n))} \right|.$$

The entries of Q and W, the antiderivatives of q and w are of locally bounded variation and therefore, $|A(x_n)|_1 \to 0$ as $n \to \infty$ and there exists a $N \in \mathbb{N}$ such that for all $n > N, \det(\mathbb{1} - A(x_n)) > 1/2$. Therefore, for $n > N$ we have

$$\left| \frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)} - 1 \right| < 4 |A_{11}(x_n) + A_{22}(x_n)|$$

and

$$\sum_{x_n \in [0,x)} \left| \frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)} - 1 \right| < 4 \sum_{x_n \in [0,x)} (|A_{11}(x_n)| + |A_{22}(x_n)|)$$

but $|A_{11}(x_n)| = |\Delta_{\overline{q_{12}}}(x_n) - \lambda\Delta_{\overline{w_{12}}}(x_n)|/2$ and $|A_{22}(x_n)| = |\Delta_{q_{12}}(x_n) - \lambda\Delta_{w_{12}}(x_n)|/2$. It follows that

$$\sum_{x_n\in[0,x)} \left|\frac{\det B_-(x_n,\lambda)}{\det B_+(x_n,\lambda)} - 1\right| < \sum_{x_n\in[0,x)} 4(|\Delta_{q_{12}}(x_n)| + |\lambda||\Delta_{w_{12}}(x_n)|)$$
$$\leq 4(\operatorname{Var}_{Q_{12}}([0,x)) + |\lambda| \operatorname{Var}_{W_{12}}([0,x))) < \infty.$$

The last inequality is again true because Q_{12} and W_{12}, antiderivatives q_{12} and w_{12}, are of locally bounded variation. Convergence of the series is independent of the order of the factors, thus, the product converges absolutely. □

The lemma that we are about to present is the core of this section. It is this lemma that does all the heavy lifting for the Theorem 3.5.

Lemma 3.4. Let v be a balanced solution of $Jv' + qv = \overline{\lambda}wv$ such that it satisfies some initial condition. If $u^-(x) = \tau(x,\lambda)\overline{v}^-(x,\overline{\lambda})$ where

$$\tau(x,\lambda) = P(x,\lambda)\exp\left(2i\int_{(0,x)} (\operatorname{Im}(q_{c,12}) - \lambda \operatorname{Im}(w_{c,12}))\right),$$

then τ is left-continuous,

$$\tau^+(x,\lambda) = \frac{\det B_-(x,\lambda)}{\det B_+(x,\lambda)}\tau(x,\lambda),$$

the balanced version of u^- satisfies $Ju' + qu = \lambda wu$, a.e. and $u(0) = \overline{v(0)}$.

Proof. We will first prove the claims about τ. The function P is left-continuous and the exponential part in τ is continuous, therefore τ is left-continuous. Also, given $x \in (0, b_\infty)$, we have

$$\tau^+(x,\lambda) = P^+(x,\lambda)e(x,\lambda) = \frac{\det B_-(x,\lambda)}{\det B_+(x,\lambda)}P(x,\lambda)e(x,\lambda),$$

where $e(x,\lambda)$ is the exponential part in the definition of τ. Therefore, we have proven our claims for τ.

To prove that u satisfies the differential equation, we will show that the measure $Ju' + qu - \lambda wu$ is a zero measure. We will do this in two steps. First we will show that it is a continuous measure, i.e. $(Ju' + qu - \lambda wu)(\{x\}) = 0$ and then complete the proof by showing that $(Ju' + qu - \lambda wu)((0, x)) = 0$. As v satisfies the equation $Jv' + qv = \overline{\lambda} wv$, we have

$$(Jv' + qv - \overline{\lambda} wv)(\{x\}) = 0 \iff \overline{B_+(x, \overline{\lambda})}\overline{v}^+(x) = \overline{B_-(x, \overline{\lambda})}\overline{v}^-(x).$$

To prove $(Ju' + qu - \lambda wu)(\{x\}) = 0$ we have to show that

$$B_+(x, \lambda)\tau^+(x)\overline{v}^+(x) = B_-(x, \lambda)\tau(x)\overline{v}^-(x),$$

which is true if and only if

$$\frac{\det B_-(x, \lambda)}{\det B_+(x, \lambda)} B_+(x, \lambda)\overline{B_+(x, \overline{\lambda})}^{-1}\overline{B_-(x, \overline{\lambda})}\tau(x)\overline{v}^-(x) = B_-(x, \lambda)\tau(x)\overline{v}^-(x).$$

As the above equality must hold for all $\tau(x)\overline{v}^-(x) \in \mathbb{C}^2$, we have to show that

$$\frac{\det B_-(x, \lambda)}{\det B_+(x, \lambda)} B_+(x, \lambda)\overline{B_+(x, \overline{\lambda})}^{-1}\overline{B_-(x, \overline{\lambda})} = B_-(x, \lambda)$$

which holds if and only if

$$B_+(x, \lambda)((\det B_-(x, \lambda))B_-(x, \lambda)^{-1})^T = B_-(x, \lambda)((\det B_+(x, \lambda))B_+(x, \lambda)^{-1})^T$$

because $\overline{B_\pm(x, \overline{\lambda})} = -B_\mp(x, \lambda)^T$. We can simplify the expression by observing that $(\det M)M^{-1} = -JM^T J$. Therefore, we need to show that

$$B_+(x, \lambda)(JB_-(x, \lambda)^T J)^T = B_-(x, \lambda)(JB_+(x, \lambda)^T J)^T.$$

The last expression can be simplified to

$$B_+(x, \lambda)JB_-(x, \lambda) = B_-(x, \lambda)JB_+(x, \lambda)$$

which is true as $(J+M)J(J-M) = (J-M)J(J+M)$ for any $M \in \mathbb{C}^{2\times 2}$. This shows that our measure is continuous. Next, we will show that the measure of the interval $(0,x)$ is zero. As u is the balanced version of $u^- = \tau\overline{v}^-$, thus

$$\begin{aligned}
&2\int_{(0,x)} (Ju' + (q-\lambda)u) \\
&= \int_{(0,x)} \left(J(\tau\overline{v}^- + \tau^+\overline{v}^+)' + (q-\lambda w)(\tau\overline{v}^- + \tau^+\overline{v}^+)\right).
\end{aligned}$$

Using the product rule for functions of locally bounded variation (2.1) and adding and subtracting $\tau\overline{v}^+$ in the second term, we get

$$\begin{aligned}
&2\int_{(0,x)} (Ju' + (q-\lambda)u) \\
&= \int_{(0,x)} \left(J(\tau'(\overline{v}^- + \overline{v}^+) + (\tau^+ + \tau)\overline{v}') + (q-\lambda w)(2\tau\overline{v} + (\tau^+ - \tau)\overline{v}^+)\right).
\end{aligned}$$

Let $\theta = \{y \in (0,x) : \Delta_q(y) \neq 0 \text{ or } \Delta_w(y) \neq 0\}$, then $\tau = \tau^+$ outside of the set θ. Also note that θ is at most countable and as our measure, $Ju' + (q-\lambda w)u$, is continuous therefore θ is a set of measure zero and hence

$$\begin{aligned}
&2\int_{(0,x)} (Ju' + (q-\lambda)u) \\
&= \int_{(0,x)\setminus\theta} (2J(\tau'\overline{v} + \tau\overline{v}') + 2(q-\lambda w)\tau\overline{v}).
\end{aligned}$$

Observe that for a 2×2 Hermitian matrix A, $A = A^T - 2i\,\mathrm{Im}(A_{12})J$, consequently

$$\begin{aligned}
&\int_{(0,x)} (Ju' + (q-\lambda)u) \\
&= \int_{(0,x)\setminus\theta} \left(J\tau'\overline{v} + \tau(J\overline{v}' + (q^T - \lambda w^T)\overline{v}) - 2i[\mathrm{Im}(q_{12}) - \lambda\,\mathrm{Im}(w_{12})]J\tau\overline{v}\right).
\end{aligned}$$

It should be noted if v satisfy $Jv' + (q - \overline{\lambda}w)v = 0$, then $\overline{v}$ satisfies $J\overline{v}' + (q^T - \lambda w^T)\overline{v} = 0$. It follows that the middle term in the integral is zero. Furthermore, $\tau' = P'e + P^-e' = P'e + 2i(\mathrm{Im}(q_{c,12}) - \lambda\,\mathrm{Im}(w_{c,12}))\tau$, using this expression in the previous equality gives us

$$\begin{aligned}&\int_{(0,x)} (Ju' + (q-\lambda)u)\\ &= \int_{(0,x)\setminus\theta} (JP'e\overline{v} - 2i[\mathrm{Im}(q_{d,12}) - \lambda\,\mathrm{Im}(w_{d,12})]J\tau\overline{v}) = 0.\end{aligned}$$

The last integral is zero because the discrete measures P', q_d, and w_d are zero outside of the set θ. It is also clear that $u(0) = \overline{v(0)}$ because $\tau(0) = \tau^+(0) = 1$. This proves our claim. Furthermore, by the uniqueness of first-order system, we have that for every v there exists a unique u and vice versa. □

THEOREM 3.5. If u and v are as defined in the Lemma 3.4, then

$$\overline{v(x,\overline{\lambda})} = \overline{v}(x,\overline{\lambda}) = \left(\frac{1}{\tau(x,\lambda)}\left(\overline{B_+(x,\overline{\lambda})}\right)^{-1} B_+(x,\lambda)\right) u(x,\lambda).$$

Additionally, if q and w are real, then $\tau(x,\lambda) = 1$ and $\left(\overline{B_+(x,\overline{\lambda})}\right)^{-1} B_+(x,\lambda) = \mathbb{1}$ for all $x \in [0, b_\infty)$.

PROOF. As u is a balanced solution, it follows that

$$u^+(x,\lambda) + u^-(x,\lambda) = 2u(x,\lambda).$$

Furthermore, u satisfies

$$B_+(x,\lambda)u^+(x,\lambda) - B_-(x,\lambda)u^-(x,\lambda) = 0$$

for every $x \in [0, b_\infty)$. As a result of these two equations, we have

$$u^\pm(x,\lambda) = J^{-1}B_\mp(x,\lambda)u(x,\lambda).$$

Using the Lemma 3.4 along with the facts that $\overline{v}$ is a balanced solution, we get

$$2\overline{v}(x,\overline{\lambda}) = \frac{1}{\tau(x,\lambda)}\left(\mathbb{1} + \frac{\det B_+(x,\lambda)}{\det B_-(x,\lambda)}(B_+(x,\lambda))^{-1}B_-(x,\lambda)\right)u^-(x,\lambda).$$

The expression can be simplified to

$$\overline{v}(x,\overline{\lambda}) = \frac{1}{\tau(x,\lambda)(\det B_-(x,\lambda))} J^{-1} B_-(x,\lambda) u^-(x,\lambda).$$

Using the relation between u^- and u that we found before, we get

$$\overline{v}(x,\overline{\lambda}) = \left(\frac{-1}{\tau(x,\lambda)} (B_-(x,\lambda)^T)^{-1} B_+(x,\lambda) \right) u(x,\lambda).$$

We can use $\overline{B_\pm(x,\overline{\lambda})} = -B_\mp(x,\lambda)^T$ in the expression above to get

$$\overline{v}(x,\overline{\lambda}) = \left(\frac{1}{\tau(x,\lambda)} \left(\overline{B_+(x,\overline{\lambda})} \right)^{-1} B_+(x,\lambda) \right) u(x,\lambda).$$

The second claim, follows directly from observing that $\det B_+(x,\lambda) = \det B_-(x,\lambda)$ and $\left(\overline{B_+(x,\overline{\lambda})} \right)^{-1} = B_+(x,\lambda)^{-1}$ whenever q and w are real. □

The following theorem will serve as the foundation for the definition of our limit-point/limit-circle dichotomy in the next section.

THEOREM 3.6. Let $\lambda_0 \notin \Lambda$ and let each entry of

$$\frac{1}{\tau(x,\lambda_0)} (B_-(x,\lambda_0)^T)^{-1} B_+(x,\lambda_0)$$

be bounded for all $x \in (0, b_\infty)$. Then if every solution of $Ju' + qu = \lambda_0 wu$ is of class $\mathcal{L}^2(w)$, then for any $\lambda \notin \Lambda$ we have that every solution of $Ju' + qu = \lambda wu$ is of class $\mathcal{L}^2(w)$.

PROOF. Consider the system $Ju' + qu = \lambda wu$ for some arbitrary $\lambda \in \mathbb{C}$ such that $\lambda \notin \Lambda$. Then we can write this system as $Ju' + qu = \lambda_0 wu + w(\lambda - \lambda_0)u$. By the variation of parameters the general solution for this non-homogeneous system is given by

$$u(x,\lambda) = U(x,\lambda_0) \left[c + (\lambda - \lambda_0) f(x, \overline{\lambda_0}, \lambda) \right],$$

where $c \in \mathbb{C}^2$, $f(x,\overline{\lambda_0},\lambda) = \int_{[0,x)} J^{-1} U(\cdot,\overline{\lambda_0})^* wu(\cdot,\lambda)$, and $U(\cdot,\lambda_0)$ is a balanced fundamental matrix for $Ju' + (q - \lambda_0 w)u = 0$ such that $U(x_0,\lambda_0) = \mathbb{1}$ for some

$x_0 \in (0, b_\infty)$. In order to prove the claim we need to show that

$$||u(\cdot,\lambda)||_x^2 = \int_{[0,x)} u(\cdot,\lambda)^* w u(\cdot,\lambda)$$

is bounded for all $x > 0$. Our first step will be to find an inequality for the vector f. The inequality will then be used to prove that $u(\cdot,\lambda)$ is square integrable. Observe that

$$|f(x,\overline{\lambda_0},\lambda)| = \left| \int_{[0,x)} \Big((U(\cdot,\overline{\lambda_0})J^{-1})^*_{\cdot,1} w u(\cdot,\lambda), \quad (U(\cdot,\overline{\lambda_0})J^{-1})^*_{\cdot,2} w u(\cdot,\lambda) \Big)^T \right|.$$

Using the Cauchy-Bunyakovsky-Schwarz inequality, we obtain

$$|f(x,\overline{\lambda_0},\lambda)| \le \sqrt{2}\max\{||(U(\cdot,\overline{\lambda_0})J^{-1})_{\cdot,1}||_x, ||(U(\cdot,\overline{\lambda_0})J^{-1})_{\cdot,2}||_x\}||u(\cdot,\lambda)||_x.$$

Note that when J^{-1} is multiplied to a matrix from the right, the two columns are interchanged and the first column becomes negative. Therefore,

$$\max\{||(U(\cdot,\overline{\lambda_0})J^{-1})_{\cdot,1}||_x, ||(U(\cdot,\overline{\lambda_0})J^{-1})_{\cdot,2}||_x\} = \max\{||U(\cdot,\overline{\lambda_0})_{\cdot,1}||_x, ||U(\cdot,\overline{\lambda_0})_{\cdot,2}||_x\}.$$

Theorem 3.5 tells us that for $k = 1, 2$

$$||U(\cdot,\overline{\lambda_0})_{\cdot,k}||_x = ||r(\cdot,\lambda_0)U(\cdot,\lambda_0)_{\cdot,k}||_x,$$

where $r(\cdot,\lambda_0) = (-(B_-(\cdot,\lambda_0)^T)^{-1}B_+(\cdot,\lambda_0))/\tau(\cdot,\lambda_0)$. Using the assumption that each entry of r is bounded, we get

$$||U(\cdot,\overline{\lambda_0})_{\cdot,k}||_x \le \tilde{r}||U(\cdot,\lambda_0)_{\cdot,k}|| < \infty$$

for some $\tilde{r} > 0$. It follows from this inequality that

$$|f(x,\overline{\lambda_0},\lambda)| \le \tilde{c}||u(\cdot,\lambda)||_x \tag{3.1}$$

where $\tilde{c} = \sqrt{2}\tilde{r}\max\{||U(\cdot,\lambda_0)_{\cdot,1}||, ||U(\cdot,\lambda_0)_{\cdot,2}||\}$. Now we will show that $u(\cdot,\lambda)$ is square integrable.

$$||u(\cdot,\lambda)||_x^2 = \int_{[0,x)} (U(\cdot,\lambda_0)c)^* w (U(\cdot,\lambda_0)c) + |\lambda-\lambda_0|^2 \int_{[0,x)} F^*(\cdot) w F(\cdot)$$
$$+ \int_{[0,x)} ((U(\cdot,\lambda_0)c)^* w((\lambda-\lambda_0)F(\cdot)) + ((\lambda-\lambda_0)F(\cdot))^* w(U(\cdot,\lambda_0)c))$$

where $F(\cdot) = U(\cdot,\lambda_0)f(\cdot,\overline{\lambda_0},\lambda)$. By applying the linearity of the integral and the Cauchy-Bunyakovsky -Schwarz inequality to the third integral, we can obtain

$$||u(\cdot,\lambda)||_x^2 \le 2||U(\cdot,\lambda_0)c||_x^2 + 2|\lambda-\lambda_0|^2 \int_{[0,x)} F^*(\cdot) w F(\cdot) \tag{3.2}$$

The first integral is finite because columns of $U(\cdot,\lambda_0)$ are in $\mathcal{L}^2(w)$. Therefore, we only have to worry about the second integral. Now we will work with the second integral and to make the analysis legible, we will suppress the arguments. In addition, f_1, f_2 will denote the first and second entries of the vector $f(\cdot,\overline{\lambda_0},\lambda)$, $U_{\cdot,1}, U_{\cdot,2}$ will denote the first and second columns of $U(\cdot,\lambda_0)$, and matrix $\tilde{W}$ will denote Radon-Nikodym derivatives of w with respect to $\mathrm{tr}(w)$. Recall that $\mathrm{tr}(w)$ is a positive measure on $(0,b_\infty)$ and

$$\int f^* w g = \int f^* \tilde{W} g \,\mathrm{tr}(w).$$

Keeping these things in mind, observe that

$$\left|\int_{[0,x)} (Uf)^* w (Uf)\right| \le \int_{[0,x)} \left|f^* U^* \tilde{W} U f\right| \mathrm{tr}(w)$$
$$= \int_{[0,x)} \left|(U_{\cdot,1}^* \tilde{W} U_{\cdot,1})(f_1^* f_1) + (U_{\cdot,2}^* \tilde{W} U_{\cdot,2})(f_2^* f_2)\right| \mathrm{tr}(w).$$
$$+ \int_{[0,x)} \left|(U_{\cdot,1}^* \tilde{W} U_{\cdot,2})(f_1^* f_2) + (U_{\cdot,2}^* \tilde{W} U_{\cdot,1})(f_2^* f_1)\right| \mathrm{tr}(w)$$

The first two integrands are positive and thus we can drop the absolute value bars in the first integral. As for the second integral, observe that f_1 and f_2 are scalars and therefore, we have

$$|(U_{\cdot,1}^{*}\tilde{W}U_{\cdot,2})(f_1^*f_2)+(U_{\cdot,2}^{*}\tilde{W}U_{\cdot,1})(f_2^*f_1)|=|(f_2^*U_{\cdot,1})^*\tilde{W}(f_1^*U_{\cdot,2})|+|(f_1^*U_{\cdot,2})^*\tilde{W}(f_2^*U_{\cdot,1})|.$$

We use this inequality in the previous integral and then use the Cauchy–Bunyakovsky –Schwarz inequality to get

$$\begin{aligned}\left|\int_{[0,x)}(Uf)^*w(Uf)\right| &\leq \int_{[0,x)}\left((U_{\cdot,1}^{*}\tilde{W}U_{\cdot,1})(f_1^*f_1)+(U_{\cdot,2}^{*}\tilde{W}U_{\cdot,2})(f_2^*f_2)\right)\operatorname{tr}(w)\\ &\quad+\int_{[0,x)}\left((U_{\cdot,1}^{*}\tilde{W}U_{\cdot,1})(f_2^*f_2)+(U_{\cdot,2}^{*}\tilde{W}U_{\cdot,2})(f_1^*f_1)\right)\operatorname{tr}(w)\,.\\ &=\int_{[0,x)}\left(((U_{\cdot,1}^{*}\tilde{W}U_{\cdot,1})+(U_{\cdot,2}^{*}\tilde{W}U_{\cdot,2}))|f|^2\right)\operatorname{tr}(w)\end{aligned}$$

Applying generalized triangle inequality to (3.2) and then using the above inequality, we get

$$\begin{aligned}&||u(\cdot,\lambda)||_x^2\leq 2||U(\cdot,\lambda_0)c||_x^2\\ &\quad+2|\lambda-\lambda_0|^2\int_{[0,x)}((U(\cdot,\lambda_0)_{\cdot,1}^{*}wU(\cdot,\lambda_0)_{\cdot,1})+(U(\cdot,\lambda_0)_{\cdot,2}^{*}wU(\cdot,\lambda_0)_{\cdot,2}))|f(\cdot,\overline{\lambda_0},\lambda)|^2\end{aligned}.$$

Using equation (3.1) in above, we get

$$\begin{aligned}&||u(\cdot,\lambda)||_x^2\leq 2||U(\cdot,\lambda_0)c||_x^2\\ &+2(\tilde{c})^2|\lambda-\lambda_0|^2\int_{[0,x)}((U(\cdot,\lambda_0)_{\cdot,1}^{*}wU(\cdot,\lambda_0)_{\cdot,1})+(U(\cdot,\lambda_0)_{\cdot,2}^{*}wU(\cdot,\lambda_0)_{\cdot,2}))||u(\cdot,\lambda)||_x^2\end{aligned}.$$

Using the Gronwall's inequality gives us

$$||u(\cdot,\lambda)||_x^2\leq 2||U(\cdot,\lambda_0)c||_x^2\exp\left(2(\tilde{c})^2|\lambda-\lambda_0|^2(||U(\cdot,\lambda_0)_{\cdot,1}||_x^2+||U(\cdot,\lambda_0)_{\cdot,2})||_x^2\right).$$

Taking the limit $x\to b_\infty$ gives us that $u\in\mathcal{L}^2(w)$ because the expression on the right finite. □

If (v,g) and (u,f) satisfies $Jv'+qv=wg$ and $Ju'+qu=wf$ respectively, and $\int_{[0,b)}v^*wf<\infty$ and $\int_{[0,b)}g^*wu<\infty$, then by equation (4.3) in [15] the following equality holds:

$$(v^*Ju)^-(b) - (v^*Ju)^+(0) = \int_{[0,b)} (v^*wf - g^*wu). \tag{3.3}$$

Equation (3.3) is called *Green's formula* or *Lagrange's identity*.

We now have the preliminary results we need to tackle the dissertation's core idea.

2. Generalization of the limit-point and limit-circle classification

Throughout this section we will assume that the Hypothesis 3.2 holds and $\mathrm{Im}(\lambda) \neq 0$.

To accomplish our goal in this section, we will use the ideas presented in [13]. We begin this section by defining the boundary value problem (BVP). Consider the following BVP on the interval $[0, b_\infty)$ where $0 < b_\infty \leq \infty$.

$$\begin{gathered} Ju' + qu = \lambda wu, \mathrm{Im}(\lambda) \neq 0 \\ \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \end{pmatrix} \begin{pmatrix} u_1(0) & u_2(0) \end{pmatrix}^T = 0, \alpha \in [0, \pi). \end{gathered} \tag{3.4}$$

Our aim is to determine how many linearly independent solutions to this differential equation are in $\mathcal{L}^2(w)$. In order to accomplish this, we will first study the problem on the interval $[0, b)$, where b is strictly smaller than b_∞ and then observe how the solutions behave as b tend towards b_∞. Note that b is always chosen as a point of continuity of the antiderivates of both q and w. A condition such as this can be imposed on b because the antiderivates of q and w can only have, at worst, countably many points of discontinuity.

For the rest of the section, $U(\cdot, \lambda)$ will denote the fundamental matrix for the system $Ju' + qu = \lambda wu$ such that

$$U(0, \lambda) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}. \tag{3.5}$$

Moreover, let $\phi(\cdot,\lambda)$ and $\psi(\cdot,\lambda)$ be the first and second columns of U respectively. Then it follows that $\psi(\cdot,\lambda)$ is a solution to the boundary value problem (3.4). Next, we define a solution to the differential equation χ given by

$$\chi(\cdot,\lambda) = U(\cdot,\lambda)\begin{pmatrix}1 & m\end{pmatrix}^T = \phi(\cdot,\lambda) + m\psi(\cdot,\lambda).$$

In the expression above, m is selected such that χ satisfies the following boundary condition.

$$\begin{pmatrix}\cot(\beta) & 1\end{pmatrix}\chi(b,\lambda) = 0, \beta \in (0,\pi). \tag{3.6}$$

The solution χ is known as the *Titchmarsh-Weyl solution* and m is known as the *Titchmarsh-Weyl m-function*. As m satisfies (3.6), therefore we must have that m is given by

$$m = -\frac{\begin{pmatrix}\cot(\beta) & 1\end{pmatrix}\phi(b,\lambda)}{\begin{pmatrix}\cot(\beta) & 1\end{pmatrix}\psi(b,\lambda)} = -\frac{\cot(\beta)\phi_1(b,\lambda) + \phi_2(b,\lambda)}{\cot(\beta)\psi_1(b,\lambda) + \psi_2(b,\lambda)}.$$

$\phi_1(b,\lambda), \phi_2(b,\lambda), \psi_1(b,\lambda)$, and $\psi_2(b,\lambda)$ are all constants, for a fixed b. Therefore, we can rewrite the above equation as

$$m = -\frac{Az+B}{Cz+D},$$

where $z = \cot(\beta), A = \phi_1(b,\lambda), B = \phi_2(b,\lambda), C = \psi_1(b,\lambda)$, and $D = \psi_2(b,\lambda)$. This is a Möbius transformation.

Note that as β varies in $(0,\pi)$, z varies in $\mathbb{R}$. Using the fact that z is real, we get

$$(C\overline{D} - \overline{C}D)|m|^2 + (\overline{B}C - \overline{A}D)m + (A\overline{D} - B\overline{C})\overline{m} + (A\overline{B} - \overline{A}B) = 0. \tag{3.7}$$

To simplify the expression above, we will use the Green's formula. Letting $(u,f) = (v,g) = (\psi(\cdot,\lambda), \lambda\psi(\cdot,\lambda))$ in (3.3) and observe that as $\psi(\cdot,\lambda) \in BV_{\text{loc}}((0,b_\infty))^2$ therefore $\psi(\cdot,\lambda) \in \mathcal{L}^2(w,[0,b))$. Combined, these two give us

$$\psi(b,\lambda)^* J\psi(b,\lambda) - \psi(0,\lambda)^* J\psi(0,\lambda) = 2i\,\text{Im}(\lambda)\int_{[0,b)} \psi(\cdot,\lambda)^* w\psi(\cdot,\lambda).$$

Note that $\psi(0,\lambda)^* J\psi(0,\lambda) = 0$ because the entries of the vector $\psi(0,\lambda)$ are real. Additionally, by assumption $\text{Im}(\lambda) \neq 0$ and the integral on the right is not zero because the definiteness condition holds, therefore the expression on the right is not zero. As a consequence of this, we can divide (3.7) by $C\overline{D} - \overline{C}D = \psi(b,\lambda)^* J\psi(b,\lambda) \neq 0$.

$$|m|^2 - \frac{\overline{A}D - \overline{B}C}{C\overline{D} - \overline{C}D} m - \frac{B\overline{C} - A\overline{D}}{C\overline{D} - \overline{C}D} \overline{m} - \frac{\overline{A}B - A\overline{B}}{C\overline{D} - \overline{C}D} = 0. \tag{3.8}$$

The equation above represents the boundary of a disk C_b in $\mathbb{C}$. To see this consider a closed disk in $\mathbb{C}$ with radius r_b and center $\tilde{m}_b$. A point m is in the disk if and only if $|m - \tilde{m}_b|^2 \leq r_b^2 \iff |m|^2 - \overline{\tilde{m}_b} m - \tilde{m}_b \overline{m} + |\tilde{m}_b|^2 - r_b^2 \leq 0$. It is clear from this expression that the boundary of this disk is given by

$$|m|^2 - \overline{\tilde{m}_b} m - \tilde{m}_b \overline{m} + |\tilde{m}_b|^2 - r_b^2 = 0.$$

Comparing this with (3.8), we see that the radius r_b and the center $\tilde{m}_b$ of the disk C_b are given by

$$r_b = \left| \frac{AD - BC}{C\overline{D} - \overline{C}D} \right|, \tilde{m}_b = \frac{B\overline{C} - A\overline{D}}{C\overline{D} - \overline{C}D}.$$

Furthermore, note that $\chi(b,\lambda)$ can be written in terms of A, B, C, D, and m as

$$\chi(b,\lambda) = \begin{pmatrix} A + Cm \\ B + Dm \end{pmatrix}.$$

This information allows us to rewrite equation (3.8) as

$$\frac{\chi(b,\lambda)^* J\chi(b,\lambda)}{\psi(b,\lambda)^* J\psi(b,\lambda)} = 0.$$

The equation above tells us that a point m is on the boundary of disk C_b if and only if

$$\chi(b,\lambda)^* J\chi(b,\lambda) = 0, \tag{3.9}$$

or in general, a point m is in C_b if and only if

$$\chi(b,\lambda)^* J \chi(b,\lambda) \over \psi(b,\lambda)^* J \psi(b,\lambda) \le 0. \tag{3.10}$$

Letting $(v,g) = (u,f) = (\chi(\cdot,\lambda), \lambda\chi(\cdot,\lambda))$ in the Green's formula (3.3) and making an observation that $\chi(\cdot,\lambda) \in BV_{\text{loc}}((0,b_\infty))^2$ therefore $\chi(\cdot,\lambda) \in \mathcal{L}^2(w,[0,b))$ and $\chi(0,\lambda)^* J\chi(0,\lambda) = -2i\,\text{Im}(m)$ give us

$$\chi(b,\lambda)^* J\chi(b,\lambda) = 2i\,\text{Im}(\lambda)\int_{[0,b)} \chi(\cdot,\lambda)^* w\chi(\cdot,\lambda) - 2i\,\text{Im}(m). \tag{3.11}$$

Equations (3.9) and (3.11) tell us that a point m is on the boundary of disk C_b if and only if

$$\int_{[0,b)} \chi(\cdot,\lambda)^* w\chi(\cdot,\lambda) = \frac{\text{Im}(m)}{\text{Im}(\lambda)}.$$

Combining (3.10) and (3.11), we get that a point m is in C_b if and only if

$$\frac{\chi(b,\lambda)^* J\chi(b,\lambda)}{\psi(b,\lambda)^* J\psi(b,\lambda)} \le 0,$$

or

$$\int_{[0,b)} \chi(\cdot,\lambda)^* w\chi(\cdot,\lambda) \le \frac{\text{Im}(m)}{\text{Im}(\lambda)}, \text{Im}(\lambda) \ne 0.$$

Note that if m is a point in C_b, then by the monotonicity of the integral we have that for $0 < a \le b < b_\infty$,

$$||\chi||_a^2 \le ||\chi||_b^2 \le \frac{\text{Im}(m)}{\text{Im}(\lambda)}.$$

This means that the point m is also in C_a for $0 < a \le b$. In other words, C_b is contained in C_a for $a \le b$. Thus for a given λ with $\text{Im}(\lambda) \ne 0$, as $b \to b_\infty$ the disks C_b can shrink down to a disk C_∞ or to a point $\tilde{m}_\infty$.

In case when C_b shrinks down to $\tilde{m}_\infty$, we see by the argument presented earlier that the point $\tilde{m}_\infty$ is in C_b for every $0 < b < b_\infty$. Therefore, for every $0 < b < b_\infty$, we have

$$\int_{[0,b)} \chi(\cdot, \tilde{m}_\infty, \lambda)^* w \chi(\cdot, \tilde{m}_\infty, \lambda) \leq \frac{\mathrm{Im}(\tilde{m}_\infty)}{\mathrm{Im}(\lambda)} < \infty.$$

Note that we are adding a third argument here to emphasis that $\chi(\cdot, \tilde{m}_\infty, \lambda)$ is associated with $\tilde{m}_\infty$, i.e., $\chi(\cdot, \tilde{m}_\infty, \lambda) = \phi(\cdot, \lambda) + \tilde{m}_\infty \psi(\cdot, \lambda)$. As the right hand side in the inequality above is independent of b, taking limit $b \to b_\infty$ tells us that

$$\chi(\cdot, \tilde{m}_\infty, \lambda) \in \mathcal{L}^2(w).$$

In case when C_b shrinks down to C_∞, we have that for any $\tilde{m}_1, \tilde{m}_2 \in C_\infty$,

$$\chi(\cdot, \tilde{m}_1, \lambda), \chi(\cdot, \tilde{m}_2, \lambda) \in \mathcal{L}^2(w).$$

An immediate consequence of this is that all solutions are in $\mathcal{L}^2(w)$. In light of the argument presented above, we can conclude that there is always a Titchmarsh-Weyl solution that is square-integrable.

We will now rewrite the expression for r_b to make it more useful and use it to define the terms *limit-point* and *limit-circle*. In order to accomplish this, let $(v, g) = (\psi(\cdot, \overline{\lambda}), \overline{\lambda}\psi(\cdot, \overline{\lambda}))$ and $(u, f) = (\phi(\cdot, \lambda), \lambda\phi(\cdot, \lambda))$ in the Green's formula. This informs us that

$$\psi(b, \overline{\lambda})^* J \phi(b, \lambda) - \psi(0, \overline{\lambda})^* J \phi(0, \lambda) = \int_{[0,b)} (\psi(\cdot, \overline{\lambda})^* w(\lambda\phi(\cdot, \lambda)) - (\overline{\lambda}\psi(\cdot, \overline{\lambda}))^* w \phi(\cdot, \lambda)).$$

Clearly, the integral on the right is zero as the integrand is zero and as 0 and b are points of continuity of the antiderivatives of q and w, it follows by Lemma 3.4 that $\psi(0, \lambda)/\tau(0, \lambda) = \overline{\psi}(0, \overline{\lambda})$ and $\psi(b, \lambda)/\tau(b, \lambda) = \overline{\psi}(b, \overline{\lambda})$ since b is a point of continuity of the antiderviatives of q and w. As a result of using these in the last expression, we obtain:

$$\frac{\psi(b, \lambda)^T J \phi(b, \lambda)}{\tau(b)} = \frac{\psi(0, \lambda)^T J \phi(0, \lambda)}{\tau(0)}.$$

Since $\psi(0,\lambda)^T J\phi(0,\lambda) = \tau(0) = 1$ therefore, $\tau(b) = \psi(b,\lambda)^T J\phi(b,\lambda) = AD - BC$. With this information, we can find an expression for the radius r_b.

$$r_b = \left|\frac{AD - BC}{\overline{C}D - C\overline{D}}\right| = \frac{|\psi(b,\lambda)^T J\phi(b,\lambda)|}{|\psi^*(b,\lambda) J\psi(b,\lambda)|} = \frac{|\tau(b,\lambda)|}{2|\operatorname{Im}(\lambda)|||\psi(\cdot,\lambda)||_b^2}.$$

Since r_b is a monotone decreasing function of b that is bounded below by zero, it converges to some non-negative value. We thus have a dichotomy for r_b as $b \to b_\infty$.

The first possibility is that r_b tends to zero as b tends to b_∞. Due to the fact that the disk C_b shrinks down to a point, this is called the limit-point case. In the classical setting, $r_b \to 0$ if and only if $\psi(\cdot,\lambda) \notin \mathcal{L}^2(w)$. The other possibility is that r_b tends to some $r_\infty > 0$ as b tends to b_∞. Since the disk C_b shrinks down to a disk, it is known as the limit-circle case. In the classical setting, $r_b \to r_\infty > 0$ if and only if $\psi(\cdot,\lambda) \in \mathcal{L}^2(w)$. Recall that we already proved that there is always a Titchmarsh-Weyl solution that is square-integrable. As a result, the number of linearly independent square integrable solutions correlates one-to-one with the convergence of the disk C_b to a point or disk in the classical case. In the classical case, the limit-point case is equivalent to the existence of only one linearly independent solution that is square integrable and the limit-circle case is equivalent to the existence of two linearly independent solutions that are square integrable or in other words, every solution is square integrable.

The next lemma shows that the classical relation between convergence of C_b to a disk or a point and the number of linearly independent square integrable solutions still holds in our setting, if we impose some conditions on τ.

THEOREM 3.7. Let $\tau(\cdot,\lambda)$ be bounded and bounded away from 0, i.e., there exist a $\gamma > 0$ such that $1/\gamma < |\tau(x,\lambda)| < \gamma$ for all $x \in (0, b_\infty)$. Then

(1) Deficiency indices are the same, i.e., $n_+ = n_-$.
(2) $\psi(\cdot,\lambda) \notin \mathcal{L}^2(w)$ if and only if the disk C_b converges to a point as b tends to b_∞ and $\psi(\cdot,\lambda) \in \mathcal{L}^2(w)$ if and only if the disk C_b converges to a disk as b tends to b_∞.

PROOF. We will start by showing that $\psi(\cdot,\lambda)\in\mathcal{L}^2(w)$ if and only if $\psi(\cdot,\overline{\lambda})\in\mathcal{L}^2(w)$. Observe that since b is a point of continuity of the antiderivatives of q and w, we have

$$(3.12)\qquad ||\psi(\cdot,\overline{\lambda})||_b^2=\frac{\overline{|\psi^*(b,\overline{\lambda})J\psi(b,\overline{\lambda})|}}{2|\operatorname{Im}(\overline{\lambda})|}=\frac{|\psi^*(b,\lambda)J\psi(b,\lambda)|}{2|\operatorname{Im}(\overline{\lambda})||\tau(b,\lambda)|^2}=\frac{||\psi(\cdot,\lambda)||_b^2}{|\tau(b,\lambda)|^2}.$$

Similarly, one can get

$$(3.13)\qquad ||\psi(\cdot,\lambda)||_b^2=\frac{||\psi(\cdot,\overline{\lambda})||_b^2}{|\tau(b,\overline{\lambda})|^2}.$$

Substitute, the expression above in (3.12), we get

$$||\psi(\cdot,\overline{\lambda})||_b=\frac{||\psi(\cdot,\overline{\lambda})||_b}{|\tau(b,\lambda)||\tau(b,\overline{\lambda})|}.$$

As the definiteness condition holds, we can make b large enough so that $||\psi(\cdot,\overline{\lambda})||_b$ is not zero and therefore, we have

$$|\tau(b,\lambda)||\tau(b,\overline{\lambda})|=1.$$

Therefore, the boundedness condition along with (3.12) and (3.13) give us the desired result. For both λ and $\overline{\lambda}$, there exist a Titchmarsh-Weyl solution that is square integrable, therefore $\psi(\cdot,\lambda)\in\mathcal{L}^2(w)$ if and only if $\psi(\cdot,\overline{\lambda})\in\mathcal{L}^2(w)$ tells us that either $n_+=n_-=1$ or $n_+=n_-=2$. Deficiency indices are the same in both cases.

We can now prove the two implications. First, we will now show that r_b converges to 0 if and only if $\psi(\cdot,\lambda)\notin\mathcal{L}^2(w)$. Assume that $\psi(\cdot,\lambda)\notin\mathcal{L}^2(w)$, then r_b must converge to 0 as $\tau(\cdot,\lambda)$ is bounded. Conversely, assume that r_b converges to 0, then we must have that $\psi(\cdot,\lambda)\notin\mathcal{L}^2(w)$ because $\tau(\cdot,\lambda)$ is bounded away from 0.

The claim r_b converges to $r_\infty>0$ if and only if $\psi(\cdot,\lambda)\in\mathcal{L}^2(w)$ can be proved by the contrapositive of the argument presented above. □

THEOREM 3.8. Let $\tau(\cdot,\lambda)$ be bounded and bounded away from 0. If $\operatorname{Im}(\lambda)\neq 0$, then $\chi(\cdot,\lambda)$ satisfies the real boundary condition (3.6) if and only if $m(b,\lambda,\beta)$ lies on the boundary of disk $C_b\subset\mathbb{C}$ if and only if $\chi(b,\lambda)^*J\chi(b,\lambda)=0$. As $b\to b_\infty$

either $C_b \to C_\infty$, a disk, or $C_b \to \tilde{m}_\infty$, a point. The space of solutions that are in $\mathcal{L}^2(w)$ is two dimensional in the former case, and one dimensional in the latter case. Moreover, in the limit-circle case, a point m is on the boundary of disk C_∞ if and only if $\chi(b_\infty, \lambda)^* J \chi(b_\infty, \lambda) = 0$.

In the next section we will take a look at how things change when we remove the condition $\lambda \in \Lambda$.

3. Limit-point and limit-circle without the existence and uniqueness theorem

The rest of this section will assume that the Hypothesis 3.2 holds with one exception: we will consider $\lambda \in \Lambda$.

The main difference between this section and the previous section is that by removing the condition $\lambda \notin \Lambda$, we have lost the existence and uniqueness of initial value problem for the system. In other words, we can not use the approach presented in the previous section because the solutions $\phi(\cdot, \lambda)$ and $\psi(\cdot, \lambda)$, as defined in the previous section, may not exist on the interval $(0, b)$ for $0 < b < b_\infty$.

In order to overcome the problem of existence and uniqueness, we will look at two cases. The goal in both cases is to define the Titchmarsh-Weyl m function and the Titchmarsh-Weyl solution.

Also as before, $b \in (0, b_\infty)$ will be selected as a point of continuity of the antiderivates of q and w. Let Ξ_λ be the set of all points in $(0, b_\infty)$ where $\det(B_\pm(\cdot, \lambda))$ is zero, i.e.,

$$\Xi_\lambda = \{x \in (0, b_\infty) : \det(B_-(x, \lambda)) \det(B_+(x, \lambda)) = 0\}.$$

As the antiderivatives of q and w are of locally of bounded variation, therefore Ξ_λ is discrete and finite on compact subintervals of $(0, b_\infty)$.

In the last section we assumed that $\lambda \notin \Lambda$. As a result, we were able to use the existence and uniqueness theorem to get solutions ϕ and ψ that satisfy initial conditions at 0. Using these solutions, we were able to define Titchmarsh-Weyl m

function and extend the limit-point and limit-circle classifications to a more general system. Without the condition $\lambda \notin \Lambda$, we may end up in a situation where non-trivial solutions ϕ and ψ may only exist on a smaller interval $(0, x_0)$, $x_0 < b_\infty$ and not on the whole interval $(0, b_\infty)$.

The following lemma shows the relation between the complex entries of $\Delta_q(x)$ and $\Delta_w(x)$ and the invertibility of $B_\pm(x, \lambda)$. An interesting conclusion of this lemma is that in the case when q and w are real, either both matrices, $B_\pm$ are invertible or not invertible at a given point. The same may not be true in case when the off-diagonal entries of q or w have complex entries.

Lemma 3.9. Let the entries of $\Delta_q(x)$ and $\Delta_w(x)$ be real for some $x \in (0, b_\infty)$, then $\det(B_-(x, \lambda)) = 0$ if and only if $\det(B_+(x, \lambda)) = 0$.

Conversely, let $\det(B_-(x, \lambda)) = \det(B_+(x, \lambda)) = 0$ for some $x \in (0, b_\infty)$, then the entries of $\Delta_q(x)$ and $\Delta_w(x)$ are real.

Proof. We will start by writing the expressions for $\det(B_\pm(x, \lambda))$. Let $\Delta_q(x)$ and $\Delta_w(x)$ be given, respectively, by

$$\Delta_q(x) = 2 \begin{pmatrix} a & b + ic \\ b - ic & d \end{pmatrix} \text{ and } \Delta_w(x) = 2 \begin{pmatrix} \alpha & \beta + i\gamma \\ \beta - i\gamma & \delta \end{pmatrix}.$$

Here $a, b, c, d, \alpha, \beta, \gamma, \delta$ are all real numbers. A simple calculation shows that

$$\det(B_+(x, \lambda)) = K - 2\operatorname{Im}(\lambda)\gamma + i(L - 2(c - \operatorname{Re}(\lambda)\gamma)),$$

and

$$\det(B_-(x, \lambda)) = K + 2\operatorname{Im}(\lambda)\gamma + i(L + 2(c - \operatorname{Re}(\lambda)\gamma)).$$

where

$$K = 1 + \frac{\det(q)}{4} + \frac{\det(w)}{4}(\operatorname{Re}(\lambda)^2 - \operatorname{Im}(\lambda)^2) - (\alpha d + \delta a - 2(\beta b + \gamma c))\operatorname{Re}(\lambda),$$

and

$$L = \left(\frac{\det(w)}{2}\operatorname{Re}(\lambda) - \delta a - \alpha d + 2(\beta b + \gamma c)\right)\operatorname{Im}(\lambda).$$

If the entries of $\Delta_q(x)$ and $\Delta_w(x)$ are real for some $x \in (0, b_\infty)$, i.e., $c = \gamma = 0$, then the claim is trivial as in this case $\det(B_-(x,\lambda)) = \det(B_+(x,\lambda))$.

Conversely, let $\det(B_-(x,\lambda)) = \det(B_+(x,\lambda)) = 0$ for some $x \in (0, b_\infty)$. It follows that the real and the imaginary parts of these determinants must also be zero. First condition, real parts must be zero, informs us that $K \pm \operatorname{Im}(\lambda)\gamma = 0$. This forces both K and γ to be zero. Similarly, the second condition, imaginary parts must be zero, informs us that $L \pm 2c = 0$. This forces both L and c to be zero. $\square$

Case 1: $B_+(\cdot,\lambda)$ is always invertible.

As Ξ_λ is discrete and $b < b_\infty$, there can only be at worst finitely many points in the interval $(0, b)$ where $B_-(\cdot,\lambda)$ is not invertible. Let us denote these points by x_k for $k = 1, 2, \ldots, n$. Furthermore, let us denote 0 by x_0 and b by x_{n+1}. Note that as b is a point of continuity of the antiderivatives of q and w, $\det(B_-(b,\lambda)) \neq 0$.

On any interval of the form (x_{k-1}, x_k) for $k = 1, 2, \ldots, n+1$ there is no point where $B_\pm(x,\lambda)$ are not invertible. In particular, by the discussion that follows (2.4), we can find a unique balanced solution, ϕ_1 on the interval (x_0, x_1) that satisfy the initial conditions

$$\phi_1(0,\lambda) = \begin{pmatrix}\cos(\alpha) & \sin(\alpha)\end{pmatrix}^T \text{ for some } \alpha \in [0,\pi).$$

Similarly, for $k = 2, 3, \ldots, n+1$, we can find a unique balanced solution ϕ_k on the interval (x_{k-1}, x_k) using the initial condition

$$\phi_k^+(x_{k-1},\lambda) = B_+(x_{k-1},\lambda)^{-1}B_-(x_{k-1},\lambda)\phi_k^-(x_{k-1},\lambda).$$

We can now define a balanced solution ϕ on the whole interval $(0, b)$. Define

$$\phi(x,\lambda) = \phi_k(x,\lambda)$$

on the interval (x_{k-1}, x_k) for $k = 1, 2, \ldots, n+1$ and

$$\phi(x_k, \lambda) = \frac{\phi_k(x_k, \lambda) + \phi_{k+1}(x_k, \lambda)}{2}$$

for $k = 1, 2, \ldots, n$. Then ϕ is a balanced solution on the whole interval $(0, b)$ and satisfies the initial condition

$$\phi(0, \lambda) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \end{pmatrix}^T.$$

In a similar fashion we can define a balanced solution ψ on the whole interval $(0, b)$ that satisfies the initial condition

$$\psi(0, \lambda) = \begin{pmatrix} -\sin(\alpha) & \cos(\alpha) \end{pmatrix}^T.$$

An important point to remember is that the solutions ϕ and ψ are going to be multiple of one another after the point x_1 as Rank $(B_-(x_1, \lambda)) = 1$.

In this case, we can use the same definition as in the previous section to define the Titchmarsh-Weyl m.

$$m = -\frac{\begin{pmatrix} \cot(\beta) & 1 \end{pmatrix} \phi(b, \lambda)}{\begin{pmatrix} \cot(\beta) & 1 \end{pmatrix} \psi(b, \lambda)}.$$

As $B_-(x_1, \lambda)$ is not invertible, therefore for all $x \geq x_1, \phi(x, \lambda) = -m_0 \psi(x, \lambda)$ for some fixed constant m_0. Therefore, for all $b > x_1$,

$$m = m_0 \frac{\begin{pmatrix} \cot(\beta) & 1 \end{pmatrix} \psi(b, \lambda)}{\begin{pmatrix} \cot(\beta) & 1 \end{pmatrix} \psi(b, \lambda)} = m_0.$$

Furthermore, the Titchmarsh-Weyl solution in this case is given by

$$\chi(x, \lambda) = \phi(x, \lambda) + m\psi(x, \lambda) = \begin{cases} \phi(x, \lambda) + m_0 \psi(x, \lambda), & x \in (0, x_1) \\ 0, & x \in [x_1, b_\infty) \end{cases}.$$

The solution is zero on $[x_1, b_\infty)$ because $\phi(x, \lambda) = -m_0 \psi(x, \lambda)$ for all $x \geq x_1$. Note that χ is a square integrable solution.

In the definition of m above, we are dividing by $\begin{pmatrix}\cot(\beta) & 1\end{pmatrix}\psi$. It is important to ensure that this expression is not zero. By Theorem 6.3 in [15], if (u, f) satisfies $Ju' + qu = wf$ on an interval (c, d), with c and d being regular endpoints, such that u satisfies real separated boundary conditions at c and d, then the problem is self-adjoint on (c, d). Furthermore, by Proposition 2.12 such a problem can only have real eigenvalues.

Suppose $\psi(y, \lambda) = 0$ for some $y \in (x_{k-1}, x_k)$ for some $k \in \{1, 2, \ldots, n+1\}$. Let $\tilde{\psi}(x, \lambda) = \psi(x, \lambda)$ on the interval $(0, y)$ and $\tilde{\psi}(x, \lambda) = 0$ for all $x \in [y, b)$. It follows that $\tilde{\psi}$ solves $Ju' + qu = \lambda wu$ on $(0, b)$. Moreover, $\tilde{\psi}$ satisfies two boundary conditions given by

$$\begin{pmatrix}\cos(\alpha) & \sin(\alpha)\end{pmatrix}\tilde{\psi}(0, \lambda) = 0 \text{ and } \begin{pmatrix}\cos(\beta) & \sin(\beta)\end{pmatrix}\tilde{\psi}(b, \lambda) = 0.$$

These are two real separated boundary conditions at two regular endpoints. Therefore, it follows from the argument presented in the paragraph above that this is a self-adjoint problem on $(0, b)$, but this cannot happened as $\mathrm{Im}(\lambda) \neq 0$. Similar reasoning shows that the numerator is not zero as well.

Case 2: $B_+(\cdot, \lambda)$ is not always invertible.

We have two possible sub-cases here First, $B_-(x, \lambda)$ is invertible for all $x \in (0, b_\infty)$ and second $B_-(x, \lambda)$ is not invertible for some $x \in (0, b_\infty)$.

In either case, we will first define a solution η and then use η to define Titchmarsh-Weyl m function and solution.

If $B_-(x, \lambda)$ is invertible for all $x \in (0, b)$, then let $x_1, x_2, \ldots, x_n$ be the points in the interval $(0, b)$ such that $\det(B_+(x_k, \lambda)) = 0$ for $k = 1, 2, \ldots, n$. An argument very similar to the one presented in the previous case tells us that we can find a solution η such that

(1) $\eta(b, \lambda) = \begin{pmatrix}-\sin(\beta) & \cos(\beta)\end{pmatrix}^T$ for some $\beta \in [0, \pi)$.

(2) η satisfies $B_-(x_k, \lambda)\eta^-(x_k, \lambda) = B_+(x_k, \lambda)\eta^+(x_k, \lambda)$ for $k = 1, 2, \ldots, n$.

For this sub-case, the Titchmarsh-Weyl m function and a constant c are selected such that

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} 1 \\ m \end{pmatrix} = c\eta(0,\lambda) \text{ for some } \alpha \in [0,\pi). \tag{3.14}$$

It follows from equation above that m must be given by

$$m = \frac{\cos(\alpha)\eta_2(0,\lambda) - \sin(\alpha)\eta_1(0,\lambda)}{\cos(\alpha)\eta_1(0,\lambda) + \sin(\alpha)\eta_2(0,\lambda)}. \tag{3.15}$$

On the other hand, if $B_-(x,\lambda)$ is not invertible for some $x \in (0,b)$, then let y be the minimum of the set $\{x \in (0,b) : \det(B_-(x,\lambda)) = 0\}$. Now there are two possibilities.

First, $B_+(x,\lambda)$ is invertible for all $x \in (0,y)$. In this case we can find a solution η on the interval $(0,b)$ such that $\eta^-(y,\lambda)$ is in the kernel of $B_-(y,\lambda)$ and $\eta(\cdot,\lambda)$ is zero on the interval (y,b).

Second, $B_+(x,\lambda)$ is not invertible for some $x \in (0,y)$. If $x_1, x_2, \ldots, x_n$ are the points in $(0,y)$ where $B_+(\cdot,\lambda)$ is not invertible, then as before we can find a solution η on $(0,b)$ such that

(1) η satisfies $B_-(x_k,\lambda)\eta^-(x_k,\lambda) = B_+(x_k,\lambda)\eta^+(x_k,\lambda)$ for all $k = 1,2,\ldots,n$.
(2) $\eta^-(y,\lambda)$ is in the kernel of $B_-(y,\lambda)$.
(3) $\eta(\cdot,\lambda)$ is zero on the interval (y,b).

We define the Titchmarsh-Weyl m by using (3.14). Note that in both of these cases the Titchmarsh-Weyl solution is $\chi = \frac{1}{c}\eta$. The solution is square integrable in the case when $B_-(\cdot,\lambda)$ is not always invertible.

Theorem 3.10. If $\Xi_\lambda = \emptyset$, then (3.6) and (3.14) give us the same Titchmarsh-Weyl m function.

Proof. To distinguish between the two m functions we will denote the new m function, given by (3.15), will be denoted by $\hat{m}$. If Ξ_λ is empty, then $\frac{1}{c}\eta(x,\lambda)$ is the same solution as Titchmarsh-Weyl solution $\chi(x,\lambda) = \phi(x,\lambda) + m\psi(x,\lambda)$ from the previous section, where ϕ and ψ are two solutions that satisfy initial conditions at 0. Using this information in (3.15) gives us

$$\hat{m} = \frac{\cos(\alpha)\frac{1}{c}(\sin(\alpha) + m\cos(\alpha)) - \sin(\alpha)\frac{1}{c}(\cos(\alpha) - m\sin(\alpha))}{\cos(\alpha)\frac{1}{c}(\cos(\alpha) - m\sin(\alpha)) + \sin(\alpha)\frac{1}{c}(\sin(\alpha) + m\cos(\alpha))} = m.$$

This proves our last claim. □

The theorem establishes that the two different approaches to defining the m function converge whenever λ is in Λ. This new definition of m can be viewed as an extension of the classical definition of m.

Bibliography

[1] H. Weyl. "Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen. (Mit 1 Figur im Text)". In: *Mathematische Annalen* 68 (1910), pp. 220–269. URL: http://eudml.org/doc/158437.

[2] A. C. Dixon and J. W. L. Glaisher. "On the singular solutions of simultaneous ordinary differential equations and the theory of congruencies". In: *Royal Society* 56 (1912), pp. 277–278. URL: https://doi.org/10.1098/rspl.1894.0110.

[3] E. C. TITCHMARSH. "ON EXPANSIONS IN EIGENFUNCTIONS (IV)". In: *The Quarterly Journal of Mathematics* os-12.1 (Jan. 1941), pp. 33–50. ISSN: 0033-5606. DOI: 10.1093/qmath/os-12.1.33. eprint: https://academic.oup.com/qjmath/article-pdf/os-12/1/33/4506975/os-12-1-33.pdf. URL: https://doi.org/10.1093/qmath/os-12.1.33.

[4] E. C. TITCHMARSH. "ON EXPANSIONS IN EIGENFUNCTIONS (V)". In: *The Quarterly Journal of Mathematics* os-12.1 (Jan. 1941), pp. 89–107. ISSN: 0033-5606. DOI: 10.1093/qmath/os-12.1.89. eprint: https://academic.oup.com/qjmath/article-pdf/os-12/1/89/4507050/os-12-1-89.pdf. URL: https://doi.org/10.1093/qmath/os-12.1.89.

[5] E. C. TITCHMARSH. "ON EXPANSIONS IN EIGENFUNCTIONS (VI)". In: *The Quarterly Journal of Mathematics* os-12.1 (Jan. 1941), pp. 154–166. ISSN: 0033-5606. DOI: 10.1093/qmath/os-12.1.154. eprint: https://academic.oup.com/qjmath/article-pdf/os-12/1/154/4506871/os-12-1-154.pdf. URL: https://doi.org/10.1093/qmath/os-12.1.154.

[6] Jesper Lützen. "Sturm and Liouville's Work on Ordinary Linear Differential Equations. The Emergence of Sturm-Liouville Theory". In: *Archive for History of Exact Sciences* 29.4 (1984), pp. 309–376. ISSN: 00039519, 14320657. URL: http://www.jstor.org/stable/41133712 (visited on 06/17/2023).

[7] Walter Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill, May 1986. ISBN: 0070542341.

[8] M.A. Al-Gwaiz. *Theory of Distributions*. 1992, p. 276. ISBN: 9780367579968.

[9] Ronald Cross. *Multivalued Linear Operators*. 1998, p. 352. ISBN: 9781439807132.

[10] W. Norrie Everitt. "Charles Sturm and the Development of Sturm-Liouville Theory in the Years 1900 to 1950". In: *Sturm-Liouville Theory: Past and Present*. Ed. by Werner O. Amrein, Andreas M. Hinz, and David P. Pearson. Basel: Birkhäuser Basel, 2005, pp. 45–74. ISBN: 978-3-7643-7359-7. DOI: 10.1007/3-7643-7359-8_3. URL: https://doi.org/10.1007/3-7643-7359-8_3.

[11] Gerd Grubb. *Distributions and Operators*. Oct. 2008, p. 464. ISBN: 9780387848945. DOI: 10.1007/978-0-387-84895-2.

[12] Jonathan Eckhardt et al. "Weyl-Titchmarsh theory for Sturm-Liouville operators with distributional potentials". In: *Opuscula Mathematica* 33.3 (2013), p. 467. DOI: 10.7494/opmath.2013.33.3.467. URL: https://doi.org/10.7494%2Fopmath.2013.33.3.467.

[13] Earl A. Coddington and Norman Levinson. *Theory of ordinary differential equations*. 2016, p. 429. ISBN: 9780070992566.

[14] Christer Bennewitz, Rudi Weikard, and Malcolm Brown. *Spectral and scattering theory for ordinary differential equations*. Vol. I. Springer, 2020.

[15] Ahmed Ghatasheh and Rudi Weikard. "Spectral theory for systems of ordinary differential equations with distributional coefficients". In: *Journal of Differential Equations* 268.6 (Mar. 2020), pp. 2752–2801. DOI: 10.1016/j.jde.2019.09.042. URL: https://doi.org/10.1016%2Fj.jde.2019.09.042.

[16] Aymen Ammar and Aref Jeribi. *Spectral Theory of Multivalued Linear Operators*. Feb. 2021, p. 287. ISBN: 9781003131120. DOI: 10.1201/9781003131120.

[17] G. B. Folland. *Real analysis: Modern, techniques and their applications*. 2nd. Wiley, 2021.

[18] Mohammed Hichem Mortad. *Counterexamples in Operator Theory*. May 2022. ISBN: ISBN 978-3-030-97813-6; ISBN 978-3-030-97814-3 (eBook). DOI: 10.1007/978-3-030-97814-3.