8-25-2016

Jost Functions for Jacobi Operators with Super-exponentially Decaying Parameters

Fred Kaul II

Fred Robert Kaul, II

Candidate

Department of Mathematics and Statistics

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Maxim Zinchenko, Chair

Dr. Anna Skripka, Member

Dr. Maria Cristina Pereyra, Member

Jost Functions for Jacobi Operators with Super-exponentially Decaying Parameters

by

Fred Robert Kaul, II

B.A., Fort Lewis College, 2010

B.S., Mathematics, University of New Mexico, 2012

THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Mathematics

The University of New Mexico

Albuquerque, New Mexico

August, 2016

Dedication

To my parents, Susan and Emil Kaul, for all of their love and support.

Acknowledgments

I would like to thank my advisor, Dr. Maxim Zinchenko, for his support and guidance, and his willingness to work with me by correspondence. I would not have had this wonderful opportunity otherwise. I would also like to thank Dr. Maria Cristina Pereyra for several years of support and mentorship, during both my undergraduate and graduate studies at UNM. I would like to give a special thanks to Dr. Anna Skripka for her cooperation as a committee member, and for being the one who introduced me to functional analysis.

I would also like to extend my deepest gratitude to Ana Parra Lombard, for all of her help with anything and everything, and for her patience with all of my many questions. I also acknowledge everyone else in the Department of Mathematics and Statistics for several years worth of mentoring and support.

Many, many thanks to my friends and family, both here in Albuquerque and elsewhere, for their endless support and encouragement.

Jost Functions for Jacobi Operators with Super-exponentially Decaying Parameters

by

Fred Robert Kaul, II

B.A., Fort Lewis College, 2010

B.S., Mathematics, University of New Mexico, 2012

M.S., Mathematics, University of New Mexico, 2016

Abstract

The decay of the parameters $\{a_n, b_n\}$ for a Jacobi operator J on $\ell^2(\mathbb{N})$ is related to the analyticity of the Jost function $u(z; J)$ associated with J, which is in turn related to the spectral measure $d\mu$ of J. Damanik and Simon demonstrated the equivalence between the exponential decay of these parameters and the analyticity of the Jost function on a disk whose radius is given by the rate of decay R. In this paper, these equivalences are summarized, and an additional equivalence is shown in the case when the parameters $\{a_n, b_n\}$ decay super-exponentially, so that $|a_n - 1| + |b_n| \leq 1/n^{\gamma n}$. In this case, the Jost function will be an entire function with finite growth order no greater than $2/\gamma$.

Contents

Chapter 1

Introduction

The purpose of this thesis is to demonstrate several important results in the spectral theory of Jacobi operators, and to demonstrate an important extension of these results. A Jacobi matrix J is a tri-diagonal symmetric matrix with parameters $\{a_n, b_n\}_{n=1}^{\infty}$ which defines an operator on the Hilbert space $\ell^2(\mathbb{N}, \mathbb{C})$. It is well-known that there is a one-to-one correspondence $J \longleftrightarrow d\mu$ between bounded Jacobi operators and compactly supported probability measures on the the real line. The forward direction is a direct application of the spectral theorem for self-adjoint operators on a Hilbert space, while the reverse direction relies on the classical theory of orthogonal polynomials on the real line.

There is also a correspondence between bounded Jacobi matrices and the Jost functions associated with their solutions. The free Jacobi matrix J_0 is defined with $a_n \equiv 1$ and $b_n \equiv 0$. This gives the exactly solvable system

$$u_{n+1} + u_{n-1} = \lambda u_n$$

for $n \geq 1$ which has solutions $\{z^n\}_{n=1}^{\infty}$, provided that $\lambda = z + z^{-1}$. If the parameters obey $\sum |a_n - 1| + |b_n| < \infty$, then the Jost solutions $\{u_n\}_{n=1}^{\infty}$ of J are those solutions that asymptotically look like the free ones for small z: $z^{-n}u_n(z) \to 1$ for $|z| < 1$. The Jost function $u_0(z)$ is defined by $(z + \frac{1}{z} - J)\{u_n(z)\}_{n=1}^{\infty} = u_0(z)e_0$ and is connected to the measure $d\mu$ by

$$|u_0(e^{i\theta})|^2 \mathrm{Im}\left(\int_{\mathrm{supp}(d\mu)} \frac{d\mu(s)}{2\cos\theta - s}\right) = \sin\theta,$$

where $u_0(z) = 0$ if and only if $z + z^{-1}$ is a point mass of $d\mu$. If J has eigenvalues outside of the interval $[-2, 2]$—$i.e.$, $u_0(z)$ has zeros $E_j = z_j + z_j^{-1}$ inside the unit disk—then $u_0(z)$ does not uniquely determine $d\mu$ unless the weights $w_j = \mu(\{E_j\})$ are taken into account.

It has long been known that the rate of decay at which $a_n \to 1$ and $b_n \to 0$ controls the analyticity of the Jost function associated with J. Damanik and Simon demonstrated in [1] that the reverse implication also holds: If $u_0(z)$ is analytic on a disk of radius R, then the Jacobi parameters decay as R^{-2}. The method of proof depends on whether or not there exist bound states of J. In this paper, these results are extended to the case where the Jacobi parameters decay super-exponentially. It will be shown that the condition $|a_n - 1| + |b_n| \leq \dfrac{1}{n^{\gamma n}}$ implies that the Jost function is entire of finite growth order, and more remarkably, the growth order is no greater that $\dfrac{2}{\gamma}$. In other words, the faster the super-exponential decay of the parameters, the smaller the growth order of u_0. A partial inverse to this statement will also be demonstrated in the case of no bound states. These results are similar to those derived by my thesis advisor Maxim Zinchenko in [12].

There is an interesting connection between these results and the classical theory of Fourier transforms. There is a correspondence $f \longleftrightarrow \hat{f}$ between L^2 functions on the real line and their Fourier transforms. The Paley-Wiener theorems link the analyticity of a function to the decay of its Fourier transform. If $a > 0$, then $e^{b|k|}\hat{f}(k) \in L^2(\mathbb{R})$ for all $b < a$ if and only if f is analytic on $|\mathrm{Im}(z)| < a$ and $\sup_{|y|<b} \int_{-\infty}^{\infty} |f(x + iy)|^2 < \infty$. This is analogous to the case of exponential decay presented in this paper. Moreover, if $a > 0$ and $f \in L^2(\mathbb{R})$, then f has compact support $\mathrm{supp}(f) \subset [-a, a]$, if and only if $\hat{f}$ is entire and $|\hat{f}(z)| \leq ce^{a|\mathrm{Im}(z)|}$ for some $c > 0$ and all $z \in \mathbb{C}$. This is analogous to the case where $J - J_0$ has finite range, so that $u_0(z)$ is a polynomial. For the Fourier transform on the unit circle, we have the following version of the Paley-Wiener theorem: If $\rho \geq 0$, then for all $\alpha < 1/\rho$ there is a c_α such that $|\hat{f}_n| \leq c_\alpha |n|^{-\alpha|n|}$ for all $n \in \mathbb{Z}$ if and only if f is analytic on $\mathbb{C} \setminus \{0\}$ and for all $\alpha > \rho$ there is a c_α such that $|f(z)| \leq c_\alpha e^{|z|^\alpha}$ for all $|z| > 1$. This is analogous to the case where the Jacobi parameters decay super-exponentially with the growth order of u_0 related to the decay rate, and as such lies somewhere in between the case of finite range (or finite support in the case of the Fourier transforms) and exponential decay. In this way, the transform $J \mapsto u_0$ can be considered a non-linear analogue of the Fourier transform.

In chapter 2, the definition and essential properties of Jacobi operators on ℓ^2 are given, and a detailed proof of the spectral theorem for these operators is given. We then demonstrate the fundamental importance of the Free Jacobi Matrix J_0 and how it provides the setting for the discussion that follows. The construction of the Jost functions are then given via the Geronimo-Case equations, which can be found in [4] and [3]. Other more classical methods for computing the Jost function exist ($e.g.$, variation of parameters), but for consistency with the presentation of Damanik and Simon, this approach is not pursued (although variation of parameters is used in a different context to derive the Weyl m-function). The method of Geronimo and Case is in any event an elegant presentation, and these equations will provide the necessary estimates needed to demonstrate the analyticity of u_0 in chapter 3.

In chapter 3, the decay of the Jacobi parameters will be used to demonstrate the analyticity of the Jost functions on various domains. In particular, the case of exponential decay is demonstrated in detail, as the case of super-exponential decay follows readily from these results.

In chapter 4, the reverse direction is demonstrated in detail in the case of no bound states. It requires a detailed analysis of the zeros of the Jost function and its related M-function, and is interesting in its own right. The case of super-exponential decay follows from the machinery developed for these results.

In chapter 5, future directions for research are discussed. Most obviously, the case of bound states is considered in some detail.

Important results from functional analysis and measure theory are included in two appendices at the end of the paper. These are summaries of topics that were independently researched by the author as preliminaries to the main body of the paper, so they are included for the sake of completeness, and as a reference for more esoteric results.

Chapter 2

Jacobi Operators

2.1 Definition and Properties

Take the square summable sequences

$$\ell^2(\mathbb{N}, \mathbb{C}) = \{f : \mathbb{N} \to \mathbb{C} : \sum_{n=1}^{\infty} |f_n|^2 < \infty\}.$$

with the standard inner product given by

$$\langle f, g \rangle = \sum_{n=1}^{\infty} f_n \overline{g_n}.$$

This is a separable Hilbert space, abbreviated as just ℓ^2 that is complete in the norm $\|f\| = \langle f, f \rangle^{1/2}$. The standard orthonormal basis for ℓ^2 is given by the vectors $e_k = (\delta_{kn})$, where $\delta_{kn} = 1$ if $k = n$ and is equal to zero otherwise, so that $\langle e_m, e_n \rangle = \delta_{mn}$.

Definition 2.1.1. *A* **Jacobi operator** *J on ℓ^2 is a real symmetric linear operator defined via the equations*

$$[Jf]_1 = b_1 f_1 + a_1 f_2$$

and

$$[Jf]_n = a_{n-1} f_{n-1} + b_n f_n + a_n f_{n+1}$$

for $n > 1$*, where* $(b_n)_{n=1}^{\infty} \subset \mathbb{R}$ *and* $(a_n)_{n=1}^{\infty} \subset (0, \infty)$.

J has the matrix representation

$$J = \begin{pmatrix} b_1 & a_1 & 0 & \dots \\ a_1 & b_2 & a_2 & \ddots \\ 0 & a_2 & b_3 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

with respect to the standard orthonormal basis. It is clear from this that J is self-adjoint, and we note that we can write $Je_k = a_{k-1}e_{k-1} + b_k e_k + a_k e_{k+1}$, so we have that $\langle Je_k, e_k\rangle = b_k$ and $\langle Je_k, e_{k+1}\rangle = a_k$. We will always assume that $\|a\|_\infty = \sup_n |a_n| < \infty$ and $\|b\|_\infty = \sup_n |b_n| < \infty$, so that J is a bounded linear operator:

$$\begin{aligned} \|J\| = \sup_{\|f\|\le 1} \|Jf\| &\le \sup_{\|f\|\le 1} \sup_{n\ge 1} |a_{n-1}f_{n-1} + b_n f_n + a_n f_{n+1}| \\ &\le \sup_{\|f\|\le 1} \sup_{n\ge 1} (\|a\|_\infty |f_{n-1}| + \|b\|_\infty |f_n| + \|a\|_\infty |f_{n+1}|) \\ &\le \|b\|_\infty + 2\|a\|_\infty. \end{aligned}$$

Note that it follows that the spectrum $\sigma(J)$ of J is compact.

2.1.1 Eigenvalue Problems and τ-operators

We now want to look at solutions of $(J - \lambda)f = 0$ for $f = (f_n)_{n=1}^{\infty}$ in $\ell^2(\mathbb{N}, \mathbb{C})$. The lemmas and proofs in this section and the next are mostly standard; good overviews are found in [9] and [6]. For reasons that will become clear, it is convenient to work initially in a more generalized setting: let τ be an operator on the set of complex-valued sequences $S = \{(f_n)_{n\ge 0}\}$, not necessarily ℓ^2, defined by

$$[\tau f]_n = a_{n-1}f_{n-1} + b_n f_n + a_n f_{n+1},$$

for $n \ge 1$, where we take a_0 to be some positive real number, while $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are as in Definition 2.1.1. From the recursion relations in Definition 2.1.1, we see that if $f \in \ell^2$, then

$[(\tau-\lambda)f]_n = [(J-\lambda)f]_n$ for $n \geq 2$, but

$$[(\tau-\lambda)f]_1 = [(J-\lambda)f]_1 + a_0 f_0. \tag{2.1}$$

We shall enjoy the freedom of extending our ℓ^2 sequences by specifying the values at $n = 0$. First we show a few elementary facts about solutions to eigenvalue problems.

Lemma 2.1.1. *Given $a, b \in \mathbb{C}$, the initial value problem given by $(\tau-\lambda)f = 0$, $f_0 = a$ and $f_1 = b$, has a unique solution $f \in S$.*

Proof. Since $(\tau-\lambda)f = 0$, we have

$$a_{n-1}f_{n-1} + (b_n - \lambda)f_n + a_n f_{n+1} = 0$$

for all n. Since $a_n > 0$ for all n, we can write this as the recursion relation

$$f_{n+1} = -\frac{1}{a_n}(a_{n-1}f_{n-1} + (b_n-\lambda)f_n), \tag{2.2}$$

so the value of f at any given n is determined completely by specifying its value at two consecutive integers, and clearly this solution is unique. □

We call the unique sequences $s = (s_n(\lambda))_{n=0}^{\infty}$ and $c = (c_n(\lambda))_{n=0}^{\infty}$ that solve $(\tau-\lambda)f = 0$ with initial conditions

$$a_0 s_0 = 0, \quad s_1 = 1, \quad a_0 c_0 = -1, \quad c_1 = 0 \tag{2.3}$$

the **fundamental solutions** of $(\tau-\lambda)f = 0$. We shall see in section 2.1.3 that if we take p_k, $k \geq 0$, to be the orthogonal polynomials associated with J via $e_k = p_{k-1}(J)e_1$, then the sequence $p_k(\lambda)$ satisfies $(\tau-\lambda)f = 0$ by the three-term recurrence relation for orthogonal polynomials, provided we take $a_0 = 1$. Now equation (2.1) prompts us to define the **transfer matrix** T_n to be

$$T_n = \frac{1}{a_n}\begin{pmatrix} -b_n+\lambda & -a_{n-1} \\ a_n & 0 \end{pmatrix}.$$

for $n > 0$. Note that $\det(T_n) = a_{n-1}/a_n \neq 0$, so that T_n is non-singular for all n with inverse

$$T_n^{-1} = \frac{1}{a_{n-1}}\begin{pmatrix} 0 & a_{n-1} \\ -a_n & -b_n+\lambda \end{pmatrix}.$$

These transfer matrices provide a convenient way to characterize the general solutions to the IVP from Lemma 2.1.1. Another important quantity is the **Wronskian**: if $g = (g_n)_{n=0}^{\infty}$ and $h = (h_n)_{n=0}^{\infty}$ are two sequences in S, then their Wronskian is defined to be

$$W_n(g,h) = a_n[g_n h_{n+1} - g_{n+1}h_n].$$

Perhaps unsurprisingly, we have the following:

Lemma 2.1.2. *If g and h both solve $(\tau - \lambda)f = 0$, then their Wronskian is constant for all n. Moreover, g and h are linearly independent if $W(g,h) \neq 0$.*

Proof. First note that If $f \in S$ solves the homogeneous equation $(\tau - \lambda)f = 0$, then we have that

$$T_n \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = \frac{1}{a_n} \begin{pmatrix} -b_n + \lambda & -a_{n-1} \\ a_n & 0 \end{pmatrix} \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix}$$

and

$$T_n^{-1} \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \frac{1}{a_{n-1}} \begin{pmatrix} 0 & a_{n-1} \\ -a_n & -b_n + \lambda \end{pmatrix} \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}.$$

From this it is clear that

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = T_n T_{n-1} \dots T_2 T_1 \begin{pmatrix} f_1 \\ f_0 \end{pmatrix}.$$

It follows that for any two solutions of g and h of $(\tau - \lambda)f = 0$, their Wronskian is constant for all n, since

$$\begin{aligned} \det \begin{pmatrix} g_{n+1} & h_{n+1} \\ g_n & h_n \end{pmatrix} &= \det(T_n T_{n-1} \dots T_2 T_1) \det \begin{pmatrix} g_1 & h_1 \\ g_0 & h_0 \end{pmatrix} \\ &= \prod_{i=1}^{n} \frac{a_{i-1}}{a_i} \det \begin{pmatrix} g_1 & h_1 \\ g_0 & h_0 \end{pmatrix} \\ &= \frac{a_0}{a_n} \det \begin{pmatrix} g_1 & h_1 \\ g_0 & h_0 \end{pmatrix}, \end{aligned}$$

for all n, so that $W_n(g,h) = W_0(g,h)$, and we refer to the Wronskian of g and h as just $W(g,h)$. If $g = ch$ for some $c \in \mathbb{C}$, then

$$W(g,h) = a_n(g_n h_{n+1} - g_{n+1} h_n) = c a_n(h_n h_{n+1} - h_{n+1} h_n) = 0,$$

so if $W(g,h) \neq 0$, then g and h are linearly independent. $\square$

Lemma 2.1.3. *The null space $N_\lambda(\tau) = \{f \in S : (\tau - \lambda)f = 0\}$ has dimension 2 for all $\lambda \in \mathbb{C}$.*

Proof. Define the map $L : \mathbb{C}^2 \to S$ by $L(a,b) = f$, where $(\tau - \lambda)f = 0$ with $f_0 = a$ and $f_1 = b$. This map is well-defined since f exists and is unique by Lemma 2.1.1. It is linear since if $(\tau - \lambda)f = 0$ and $(\tau - \lambda)g = 0$ with $(f_0, f_1) = (a,b)$ and $(g_0, g_1) = (c,d)$, then for any $\alpha \in \mathbb{C}$, we have $(\tau - \lambda)(\alpha f + g) = 0$ with $(\alpha f + g)_0 = \alpha a + c$ and $(\alpha f + g)_1 = \alpha b + d$. This shows that

$$L(\alpha(a,b) + (c,d)) = L(\alpha a + c, \alpha b + d) = \alpha f + g = \alpha L(a,b) + L(c,d).$$

Moreover, if $L(a,b) = f = L(c,d)$, then $(a,b) = (f_0, f_1) = (c,d)$, so that L is injective. Moreover, for any $f \in N_\lambda(J)$, we have that $L(f_0, f_1) = f$, so that L is surjective. This shows that L is a vector space isomorphism between $\mathbb{C}^2$ and $N_\lambda(J)$, so in particular we have that $N_\lambda(J)$ has dimension 2. $\square$

Using the Wronskian, we now have a convenient characterization of the general solutions to the IVP of Lemma 2.1.1 in terms of the fundamental solutions:

Lemma 2.1.4. *If f solves $(\tau - \lambda)f = 0$ and s and c are the fundamental solutions, we have*

$$f_n = f_1 s_n - a_0 f_0 c_n$$

for all n.

Proof. If $(\tau - \lambda)f = 0$, then we can write

$$\begin{aligned} W(f,h)g_n - W(f,g)h_n &= a_n g_n[f_n h_{n+1} - f_{n+1} h_n] \\ &\quad - a_n h_n [f_n g_{n+1} - f_{n+1} g_n] \\ &= f_n a_n [g_n h_{n+1} - h_n g_{n+1}] \\ &= f_n W(g,h), \end{aligned}$$

whence

$$f_n = \frac{W(f,h)}{W(g,h)} g_n - \frac{W(f,g)}{W(g,h)} h_n.$$

By equations (2.3), we see that $W(c, s) = a_0[c_0s_1 - c_1s_0] = -1$, while $W(f, c) = a_0[f_0c_1 - c_0f_1] = f_1$ and $W(f, s) = a_0[f_0s_1 - s_0f_1] = a_0f_0$, so that

$$f_n = f_1 s_n - a_0 f_0 c_n$$

for all n. □

Note that in particular, this shows that every $\lambda \in \mathbb{C}$ is an eigenvalue of τ whose null space has dimension 2, and we can take the fundamental solutions $\{c, s\}$ as a basis. Further, the recurrence relation gives

$$s_2(\lambda) = -a_1^{-1}[a_0s_0(\lambda) + (b_1 - \lambda)s_1(\lambda)] = a_1^{-1}\lambda + O(1).$$

Similarly, we have

$$s_3(\lambda) = a_2^{-1}[a_1s_1(\lambda) + (b_2 - \lambda)s_2(\lambda)] = -(a_1a_2)^{-1}\lambda^2 + O(\lambda).$$

Inductively, we can see that s_n is a polynomial of degree $n - 1$:

$$s_n(\lambda) = (-1)^{n-1}(a_1 \dots a_n)^{-1}\lambda^{n-1} + O(\lambda^{n-2}).$$

If we consider the finite matrix J_{n-1} given by taking the $\{1, \dots, n-1\} \times \{1, \dots, n-1\}$ block of J, and suppose that $s_n(\lambda) = 0$, then since $s_0(\lambda) = 0$, we see that $(J_{n-1} - \lambda)(s_1(\lambda), \dots, s_{n-1}(\lambda))^T = 0$, so that λ is an eigenvalue of J_{n-1}. The converse holds as well, so this shows that the determinant of $J_{n-1} - \lambda$ and $s_n(\lambda)$ differ by a constant multiple:

$$s_n(\lambda) = (-1)^{n-1} \frac{\det(J_{n-1} - \lambda)}{\prod_{i=1}^{n-1} a_i}.$$

The expression for c_n can be computed similarly.

2.1.2 Solutions of $(J-\lambda)f=0$

Now suppose that $\lambda \in \mathbb{C}$ with $\mathrm{Im}(\lambda) \neq 0$. Since J is self-adjoint, we have that its spectrum is real, $\sigma(J) \subset \mathbb{R}$, so $(J-\lambda)^{-1}$ is a bounded linear operator on ℓ^2. We have the following:

Lemma 2.1.5. *For $\lambda \in \mathbb{C}$ with Im$(\lambda) \neq 0$, the equation $(\tau-\lambda)f=0$ has precisely one linearly independent solution in ℓ^2.*

Proof. Define the sequence $u=(J-\lambda)^{-1}e_1$. Clearly this sequence belongs to ℓ^2, and $(J-\lambda)u = e_1$. We have that $[(\tau-\lambda)u]_n = [(J-\lambda)u]_n = 0$ for $n \geq 2$, and by extending u to $u_0 = -1/a_0$, we have by equation (2.2) that

$$[(\tau-\lambda)u]_1 = [(J-\lambda)u]_1 + a_0u_0 = [e_1]_1 - 1 = 0.$$

This shows that u is an ℓ^2 solution of $(\tau-\lambda)f=0$. Since the solution space of $(\tau-\lambda)f=0$ is two-dimensional, if we had a second linearly independent solution in ℓ^2, then all solutions would belong to ℓ^2. However, this would imply the existence of an ℓ^2 solution w with $w_0=0$, but in this case we would have $(J-\lambda)w = (\tau-\lambda)w = 0$, which would imply that J has an eigenvalue with non-zero imaginary part, a contradiction. This shows that $(\tau-\lambda)f=0$ has precisely one linearly independent solution in ℓ^2. □

This now allows us to define the **Titchmarsh-Weyl m-function**: for all $\lambda \in \mathbb{C}^+ = \{\lambda \in \mathbb{C} : \mathrm{Im}(\lambda) > 0\}$, there exists a unique number $m(\lambda)$ such that $w(\lambda) = c(\lambda) + m(\lambda)s(\lambda) \in \ell^2(\mathbb{N},\mathbb{C})$. To see this is well-defined, note that by Lemma 2.1.5, there is a unique (up to a constant) ℓ^2 solution f of $(\tau-\lambda)f=0$. Since $\{c,s\}$ is a basis for the entire solution space by Lemmas 2.1.3 and 2.1.4, we must have that there exist unique $a(\lambda)$ and $b(\lambda)$ such that $f(\lambda) = a(\lambda)c(\lambda) + b(\lambda)s(\lambda)$. If $a(\lambda)$ were zero, then $s(\lambda)$ would have to be ℓ^2, but since $s_0 = 0$, we would have that $(J-\lambda)s = 0$, contradicting the fact that $\lambda \in \mathbb{C}^+$. So we may define w by $w(\lambda) = f(\lambda)/a(\lambda)$ and take $m(\lambda) = b(\lambda)/a(\lambda)$. The m-function has two important properties:

Theorem 2.1.6. *For all $\lambda \in \mathbb{C}^+$, if $(\tau-\lambda)u = 0$, $u \neq 0$, and $u \in \ell^2$, then*

i) $m(\lambda) = -\dfrac{u_1(\lambda)}{a_0u_0(\lambda)}$;

ii) $m(\lambda) = \langle (J-\lambda)^{-1} e_1, e_1 \rangle$.

Proof. To see (i), note that for $w(\lambda) = c(\lambda) + m(\lambda)s(\lambda)$, we have immediately that $w_1(\lambda) = c_1(\lambda) + m(\lambda)s_1(\lambda) = m(\lambda)$ and $a_0 w_0(\lambda) = a_0 c_0(\lambda) = -1$ by equations (2.3). For any u with $(\tau - \lambda)u = 0$ and $u \in \ell^2$, we must have that $u(\lambda) = cw(\lambda)$ for some $c \in \mathbb{C} \setminus \{0\}$, so that

$$-\frac{u_1(\lambda)}{a_0 u_0(\lambda)} = -\frac{c w_1(\lambda)}{a_0 c w_0(\lambda)} = -\frac{w_1(\lambda)}{a_0 w_0(\lambda)} = m(\lambda).$$

Property (ii) is more involved. We first note that if p is a particular solution of the inhomogeneous equation $(\tau - \lambda)p = q$, $q = (q_n)_{n=0}^{\infty}$ a complex-valued sequence, then the general solution is given by $y = \alpha s + \beta c + p$ for $\alpha, \beta \in \mathbb{C}$. This follows since clearly $(\tau - \lambda)(y - p) = 0$, and so $y - p$ is a linear combination of the fundamental solutions s and c for $(\tau - \lambda)f = 0$. Now suppose that u and v solve $(\tau - \lambda)y = 0$ with $W(u, v) = 1$. We claim that

$$y_n = \sum_{j=1}^{n} K_n^j q_j$$

solves $(\tau - \lambda)y = q$, where $K_n^j = u_j v_n - u_n v_j$. To see this, we calculate

$$\begin{aligned}
[(\tau - \lambda)y]_n &= a_n \sum_{j=1}^{n+1} K_{n+1}^j q_j + (b_n - \lambda) \sum_{j=1}^{n} K_n^j q_j + a_{n-1} \sum_{j=1}^{n-1} K_{n-1}^j q_j \\
&= \sum_{j=1}^{n} (a_n K_{n+1}^j + (b_n - \lambda) K_n^j + a_{n-1} K_{n-1}^j) q_j - a_{n-1} K_{n-1}^n q_n \\
&= \sum_{j=1}^{n} (u_j [(\tau - \lambda)v]_n - [(\tau - \lambda)u]_n v_j) q_j + a_{n-1}(u_{n-1} v_n - u_n v_{n-1}) q_n \\
&= W_n(u, v) q_n = q_n.
\end{aligned}$$

This method is known as **variation of parameters**. Define the **Green function** by $G(m, n; \lambda) = \langle (J - \lambda)^{-1} e_n, e_m \rangle$. Fix $n \in \mathbb{N}$ and take $g^{(n)} = (J - \lambda)^{-1} e_n$ with $g_0^{(n)} = 0$. Then $g^{(n)}$ solves the inhomogeneous equation $(\tau - \lambda)g = e_n$, and we see that $G(m, n; \lambda) = \langle g^{(n)}, e_m \rangle = g_m^{(n)}$. Applying variation of parameters with $u = s$ and $v = w$, we have that

$$K_m^j = \sum_{j=1}^{m} (s_j w_m - s_m w_j) \delta_{jn} = \begin{cases} 0 & m \leq n \\ s_n w_m - s_m w_n & m > n \end{cases}$$

solves the inhomogeneous equation. It follows that $g^{(n)}$ can be written in the form

$$g_m^{(n)} = \alpha c_m + \beta s_m + \begin{cases} 0 & m \leq n \\ s_n w_m - s_m w_n & m > n \end{cases}$$

for $\alpha, \beta \in \mathbb{C}$. Since $g_0 = \alpha c_0 + \beta s_0 = \alpha c_0 = 0$, we have immediately that $\alpha = 0$. To determine β, we note that since w and $g^{(n)}$ are ℓ^2,

$$g^{(n)} - s_n w_m = (\beta - w_n) s_m \to 0$$

as $m \to \infty$. Since s_m is a polynomial, we cannot have $s_m \to 0$, so we must have $\beta - w_n = 0$. This shows that

$$G(m, n; \lambda) = \begin{cases} s_m w_n & m \leq n \\ s_n w_m & m > n \end{cases}.$$

Taking $n = m = 1$ gives

$$\langle (J - \lambda)^{-1} e_1, e_1 \rangle = G(1, 1; \lambda) = s_1 w_1 = s_1 c_1 + m s_1^2 = m(\lambda),$$

as desired. □

2.1.3 Spectral Theorem

In this section, we present a proof of a version of the spectral theorem specifically for Jacobi operators. In particular, we will see that there is a one-to-one correspondence between Jacobi operators and compactly supported probability measures on the real line. The presentation follows the outline presented in chapter 2 of [2], but can also be found in [6]. We see that since $J^0 e_1 = e_1$ and $J^1 e_1 = b_1 e_1 + a_1 e_2$, we have

$$J^2 e_1 = b_1^2 e_1 + a_1 b_1 e_2 + a_1(a_1 e_1 + b_2 e_2 + a_2 e_3) = a_1 a_2 e_3 + u_2,$$

where $u_2 \in \text{span}(e_1, e_2)$. Inductively, we see that $J^k e_1 = a_1 \cdots a_k e_{k+1} + u_k$, where $u_k \in \text{span}(e_1, \ldots, e_k)$. It follows that $\text{span}(e_1, \ldots, e_{k+1}) = \text{span}(e_1, J e_1, \ldots, J^k e_1)$, so that e_1 is a cyclic vector. If we take $\mu = \mu_{e_1}$, then the spectral theorem (Theorem B.2.1) shows that for $z \in \rho(J)$, we have

$$\langle (J-z)^{-1}e_1, e_1\rangle = \int_{\sigma(J)} \frac{1}{\lambda - z} d\mu(\lambda),$$

where supp$(\mu) = \sigma(J)$, since J has a simple spectrum (see appendix B.3). Moreover, Theorem B.1.1 shows that $\mu(\mathbb{R}) = \|e_1\|^2 = 1$, so that μ is a Borel probability measure on the real line, and since μ is uniquely defined by (J, e_1), we have that the map Φ given by $J \overset{\Phi}{\mapsto} \mu$ from the set of bounded Jacobi matrices to the set of probability measures with compact support is well-defined. We will refer to μ as the **spectral measure** of J. Note that if we restrict our attention to finite Jacobi matrices on $\mathbb{C}^n$ for some n, then supp(μ) is a finite set with at most n elements, and so μ is a finite sum of point measures. We wish to show that Φ gives a one-to-one correspondence.

We first note that for each e_k, there is a unique polynomial p_k with positive leading coefficient $(a_1 \cdots a_k)^{-1}$ such that $e_k = p_k(J)e_1$. We see that

$$\langle e_m, e_n\rangle = \langle p_m(J)e_1, p_n(J)e_1\rangle = \int_{\sigma(J)} p_m(\lambda)\overline{p_n(\lambda)} d\mu(\lambda) = \delta_{mn},$$

so the polynomials p_n, p_m are orthogonal with respect to μ. If T is another Jacobi operator with diagonal entries (d_n) and off-diagonal entries (c_n) such that $\Phi(T) = \mu$, then clearly $e_k = q_k(T)e_1$ implies that $p_k = q_k$ by the uniqueness of orthogonal polynomials, whence

$$\begin{aligned} b_k &= \langle Je_k, e_k\rangle = \langle Jp_k(J)e_1, p_k(J)e_1\rangle \\ &= \int_{\sigma(J)} \lambda |p_k(\lambda)|^2 d\mu = \int_{\sigma(J)} \lambda |q_k(\lambda)|^2 d\mu = d_k, \end{aligned}$$

and similarly $a_k = \langle Je_k, e_{k+1}\rangle = c_k$. This shows that $T = J$, and so Φ is injective.

To show that Φ is onto, we construct explicitly a map Θ from the set of probability measures with compact support to the set of bounded Jacobi matrices such that $\Phi \circ \Theta$ is the identity. This can be done with orthonormal polynomials. Let μ be a probability measure with compact support $\Sigma = \text{supp}(\mu)$. First note that if $\{1, \lambda, \ldots, \lambda^k\}$ is a linearly dependent set in $L^2(\mathbb{R}, d\mu)$, then there exist a non-zero polynomial p with deg$(p) \leq k$ such that $\int_\Sigma p\, d\mu = 0$. Since $p \not\equiv 0$, we must have that Σ is the zero set of the polynomial p. Since p has at most k zeroes, say $\{x_i\}_{i=1}^k$, we have that μ must be a finite sum of point measures: $\mu = \sum_{i=1}^k \alpha_i \delta_{x_i}$, with $\sum_{i=1}^k \alpha_i = 1$. This gives one direction of the correspondence between such measures and finite Jacobi matrices on $\mathbb{C}^k$. Otherwise, we must have $\{1, \lambda, \dots, \lambda^k\}$ is linearly independent for all k.

Applying the Gram-Schmidt procedure to $\{1, \lambda, \dots, \lambda^k\}$, we find a sequence $(p_k(\lambda))_{k=0}^{\infty}$, $\deg(p_k) = k$, of orthonormal polynomials with positive leading coefficients such that $p_0 \equiv 1$ and

$$\int_\Sigma p_m(\lambda) p_n \, d\mu(\lambda) = \delta_{mn}.$$

We can take $\lambda p_k(\lambda) = \sum_{i=0}^{k+1} c_i p_i(\lambda)$, so we have

$$\int_\Sigma (\lambda p_k(\lambda)) p_l(\lambda) \, d\mu(\lambda) = \sum_{i=0}^{k+1} c_i \int_\Sigma p_i(\lambda) p_l(\lambda) \, d\mu(\lambda) = c_l$$

by orthonormality. But since

$$\int_\Sigma (\lambda p_k(\lambda)) p_l(\lambda) \, d\mu(\lambda) = \int_\Sigma p_k(\lambda)(\lambda p_l(\lambda)) \, d\mu(\lambda) = 0$$

for $l < k-1$, we have that $c_l = 0$ for $l < k-1$, and the remaining coefficients yield the three-term recurrence relation

$$\lambda p_k(\lambda) = c_{k-1} p_{k-1}(\lambda) + c_k p_k(\lambda) + c_{k+1} p_{k+1}(\lambda),$$

for $k > 0$ and $\lambda p_0(\lambda) = c_1 p_1(\lambda) + c_0 p_0(\lambda)$. We can define the sequences

$$b_k = \int_\Sigma \lambda p_{k-1}(\lambda)^2 \, d\mu(\lambda) \quad \text{and} \quad a_k = \int_\Sigma \lambda p_k(\lambda) p_{k-1}(\lambda) \, d\mu(\lambda),$$

so that our recurrence relation becomes

$$\begin{aligned} \lambda p_0(\lambda) &= b_1 p_0(\lambda) + a_1 p_1(\lambda) \\ \lambda p_k(\lambda) &= a_{k+1} p_{k+1}(\lambda) + b_{k+1} p_k(\lambda) + a_k p_{k-1}(\lambda), \end{aligned}$$

for $k > 1$. Comparing the $(k+1)$-st coefficients shows that a_k is the ratio of the leading terms for $p_k(\lambda)$ and $p_{k+1}(\lambda)$, and so $a_k > 0$ for all k. This yields a Jacobi matrix J. We have the bound

$$|b_k| \leq \sup_{\lambda \in \Sigma} |\lambda| \int_\Sigma |p_{k-1}(\lambda)|^2 \, d\mu(\lambda) = \sup_{\lambda \in \Sigma} |\lambda| < \infty$$

for all k, and the same bound of course also works for each a_k. Hence, J is a bounded Jacobi matrix, and we take $\Theta(\mu(\lambda)) = J$. From the definition of J, we have that $e_k = p_k(J)e_1$ for each k. Writing $\lambda^k = \sum_{i=0}^k \gamma_{ki} p_i(\lambda)$, we see that $J^k = \sum_{i=0}^k \gamma_{ki} p_i(J)$, and so orthonomality gives

$$\langle J^k e_1, e_1 \rangle = \sum_{i=0}^{k} \gamma_{ki} \langle p_i(J) e_1, e_1 \rangle = \gamma_{k0} = \sum_{i=0}^{k} \gamma_{ki} \int_\Sigma p_0(\lambda) p_i(\lambda)\, d\mu(\lambda) = \int_\Sigma \lambda^k \, d\mu(\lambda).$$

This in turn implies that for all $z \in \rho(J)$ we have

$$\langle (J-z)^{-1} e_1, e_1 \rangle = \int_\Sigma \frac{d\mu(\lambda)}{\lambda - z},$$

so that $\mu = \Phi(J) = (\Phi \circ \Theta)(\mu)$, which was to be shown. We have shown the following:

Theorem 2.1.7. (Spectral Theorem for Jacobi Operators) *There is a one-to-one correspondence between bounded Jacobi matrices and probability measures on the real line with compact support. If J is a Jacobi matrix and μ the corresponding spectral measure, then*

$$\langle (J-z)^{-1} e_1, e_1 \rangle = \int_{\sigma(J)} \frac{d\mu(\lambda)}{\lambda - z}, \quad \forall z \in \rho(J).$$

Moreover, e_1 is cyclic under J in ℓ^2, and so there is a unitary map U such that $(UJU^{-1}\psi)(\lambda) = \lambda\psi(\lambda)$ for all $\psi \in L^2(d\mu)$.

If we consider finite Jacobi matrices over $\mathbb{C}^n$, then μ will be a sum of point measures corresponding to the eigenvalues of J. It is notable that the classical theory of orthonormal polynomials was essential to the construction of the correspondence in the infinite case.

2.2 Free Jacobi Operator

We define the **free Jacobi matrix** J_0 by taking $a_n = 1$ and $b_n = 0$ for all n, so that J_0 has the matrix representation

$$J_0 = \begin{pmatrix} 0 & 1 & 0 & \cdots & \\ 1 & 0 & 1 & \ddots & \\ 0 & 1 & 0 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \end{pmatrix}.$$

This matrix represents the discrete Laplacian and can be used to model a discrete version of a free (zero potential) quantum mechanical particle on the half line.

2.2.1 The Spectrum of J_0

Taking the equation $J_0 u = \lambda u$ gives the recurrence relation

$$u_{n+1} + u_{n-1} = \lambda u_n,$$

for $n > 1$ and $u_2 = \lambda u_1$. We shall use the associated τ-operator τ_0 to show that J_0 has no eigenvalues. Many of these results are also summarized in [9] and [6]. First we show:

Lemma 2.2.1. *There exist $z_\pm(\lambda)$ such that the sequence $u_n(\lambda) = [z_\pm(\lambda)]^n$ solve $(\tau_0 - \lambda)u = 0$ with $z_+(\lambda)z_-(\lambda) = 1$. Moreover, we have:*

i) if $\lambda \in \mathbb{C} \setminus [-2, 2]$, then we can choose $z_\pm$ such that $|z_+(\lambda)| > 1$ and $|z_-(\lambda)| < 1$;

ii) if $\lambda \in (-2, 2)$, then $|z_\pm| = 1$ and $z_-(\lambda) = \overline{z_+(\lambda)}$;

iii) if $\lambda = \pm 2$, then the fundamental solutions are $\{(\pm 1)^n, (\pm 1)^n n\}$.

Proof. Using the ansatz $u_n = z^n$, we have that $u_{n+1}+u_{n-1} = \lambda u_n$ is equivalent to $\lambda = z+z^{-1}$, or $z^2-\lambda z+1 = 0$. Defining $u_0 \equiv 1$, we have that $u(z_\pm(\lambda)) = ([z_\pm(\lambda)]^n)_{n=1}^\infty$ solves $(\tau-\lambda)u(\lambda) = 0$, provided that we pick $a_0 = 1$, where

$$z_\pm(\lambda) = \frac{\lambda \pm \sqrt{\lambda^2 - 4}}{2}.$$

To verify this for $n = 1$, we note that

$$u_2 + u_0 = \left(\frac{\lambda \pm \sqrt{\lambda^2 - 4}}{2}\right)^2 + 1 = \frac{2\lambda^2 \pm 2\lambda\sqrt{\lambda^2 - 4} - 4}{4} + 1 = \lambda u_1.$$

We see that

$$z_+(\lambda)z_-(\lambda) = \frac{\lambda + \sqrt{\lambda^2 - 4}}{2} \cdot \frac{\lambda - \sqrt{\lambda^2 - 4}}{2} = \frac{\lambda^2 - (\lambda^2 - 4)}{4} = 1,$$

and

$$\begin{aligned} W(u_+, u_-) &= z_+(\lambda)^n z_-(\lambda)^{n+1} - z_+(\lambda)^{n+1} z_-(\lambda)^n \\ &= (z_+(\lambda) z_-(\lambda))^n z_-(\lambda) - (z_+(\lambda) z_-(\lambda))^n z_+(\lambda) \\ &= z_-(\lambda) - z_+(\lambda) \\ &= \sqrt{\lambda^2 - 4}, \end{aligned}$$

so that u_+ and u_- are linearly independent by Lemma 2.1.2 since $\lambda \neq \pm 2$. To prove (i), we see that $|z_+(\lambda)| = 1$ if and only if $z_+(\lambda) = e^{i\theta}$ if and only if

$$\lambda = e^{i\theta} + e^{-i\theta} = 2\cos\theta$$

if and only if $\lambda \in [-2, 2]$, and similarly for $z_-(\lambda)$. So if $\lambda \in \mathbb{C} \setminus [-2, 2]$, then we can take $z_+(\lambda)$ such that $|z_+(\lambda)| > 1$ and $z_-(\lambda)$ such that $|z_-(\lambda)| < 1$. To prove (ii), we see that if $-2 < \lambda < 2$, then

$$z_\pm(\lambda) = \frac{1}{2}(\lambda \pm i\sqrt{4 - \lambda^2}),$$

so that $|z_\pm(\lambda)|^2 = \dfrac{1}{4}(\lambda^2 + (4 - \lambda^2)) = 1$ and $z_-(\lambda) = \overline{z_+(\lambda)}$. To prove (iii), we note that if $\lambda = \pm 2$, then $z = \pm 1$, so that $(\pm 1)^n$ is one solution. We can then take $(\pm 1)^n n$ as our second solution, since we have

$$W(u_+(\pm 2), u_-(\pm 2)) = (\pm 1)^n (\pm 1)^{n+1}(n+1) - (\pm 1)^{n+1}(\pm 1)^n n = (\pm 1)^{2n}(\pm(n+1) \mp n) = \pm 1.$$

□

Now we have

Lemma 2.2.2. *J_0 has no eigenvalues.*

Proof. Since τ_0 extends J_0, parts (ii) and (iii) of Lemma 2.2.1 show that there are no possible ℓ^2 solutions for $\lambda \in [-2, 2]$. If $\lambda \in \mathbb{C} \setminus [-2, 2]$, then there is exactly one solution $u(\lambda)$ in the unit disk, which is clearly ℓ^2. By Lemma 2.1.5, this solution is unique (up to a constant). Take $z(\lambda) = z_+(\lambda)$ in the unit disk, and denote the solution by $u_n(\lambda) = [z(\lambda)]^n$. This is the only option for an eigenvalue of J_0, but since $u_0 = 1$, we have that $u_2 = \lambda u_2 - 1 \neq \lambda u_1$, so that this solution does not solve $(J_0 - \lambda)u = 0$. □

From Theorem 2.1.6(i), we have that

$$m(\lambda) = -\frac{u_1(\lambda)}{a_0 u_0(\lambda)} = -u_1(\lambda) = -\frac{\lambda + \sqrt{\lambda^2 - 4}}{2}.$$

Then Theorem 2.1.6(ii) and Theorem 2.1.7 show that

$$m(\lambda) = \langle (J - \lambda)^{-1} e_1, e_1 \rangle = \int_{\sigma(J_0)} \frac{d\mu(\zeta)}{\zeta - \lambda},$$

where μ is the spectral measure of J_0. We see that for $x \in \mathbb{R}$:

$$\begin{aligned} m(x) &= \lim_{\varepsilon \to 0^+} m(x + i\varepsilon) = \lim_{\varepsilon \to 0^+} -\frac{x + i\varepsilon \pm \sqrt{(x + i\varepsilon)^2 - 4}}{2} \\ &= \frac{1}{2} \begin{cases} -x + i\sqrt{4 - x^2} & x \in [-2, 2] \\ -x - \sqrt{x^2 - 4} & x < -2 \\ -x + \sqrt{x^2 - 4} & x > 2 \end{cases}, \end{aligned}$$

so that $\mathrm{Im}(m) = \frac{1}{2}\chi_{[-2,2]}\sqrt{4 - x^2}$. Since $m(x)$ has no singularities, part (i) of Theorem B.3.3 shows that $\mu_s = 0$. Furthermore, part (ii) shows that

$$\mu_{ac} = \frac{1}{\pi}\mathrm{Im}(m)dx = \frac{1}{2\pi}\chi_{[-2,2]}\sqrt{4 - x^2}\, dx.$$

This shows that $\sigma(J_0) = [-2, 2]$ and must be purely continuous. Under the transformation $\lambda = 2\cos\theta$, this can also be expressed as

$$d\mu = \frac{2}{\pi}\chi_{[0,\pi]} \sin^2\theta \, d\theta.$$

2.2.2 Uniformization

Consider the transformation $f(z) = z + z^{-1}$. Using $z = re^{i\theta}$, we see that

$$\begin{aligned} f(z) = u + iv &= re^{i\theta} + \frac{1}{r}e^{-i\theta} \\ &= \left(r + \frac{1}{r}\right)\cos\theta + i\left(r - \frac{1}{r}\right)\sin\theta. \end{aligned}$$

This gives

$$\frac{u^2}{(r + 1/r)^2} + \frac{v^2}{(r - 1/r)^2} = 1.$$

This shows that circles of the form $|z| = r < 1$ in $\mathbb{D}$ are mapped to ellipses in $\mathbb{C} \setminus [-2, 2]$, and evidently $0 \mapsto \infty$. This is exactly the resolvent set of J_0. We see that f has the derivative

$$f'(z) = 1 - \frac{1}{z^2} = \frac{z^2 - 1}{z^2},$$

so it is conformal everywhere except $z = \pm 1$. We conclude that $z \mapsto \lambda = z + z^{-1}$ is a conformal map from $\mathbb{D}$ to $(\mathbb{C} \cup \{\infty\}) \setminus [-2, 2]$. It is important to note that in the case of J_0, we have from the proof of Lemmas 2.2.1 and 2.2.2 the existence of a $z \in \mathbb{D}$ such that $u_n(f(z)) = u_n(\lambda) = [z_\pm(\lambda)]^n$ is ℓ^2 and solves $(\tau_0 - \lambda)u_n(f(z))_{n=1}^\infty = 0$.

2.3 Jost Solutions and Jost Functions

We want to consider ℓ^2 sequences $u_n(z)$ that solve

$$(J - \lambda)(u_n(z))_{n=1}^\infty = -u_0(z)e_1$$

with $\lambda = z + z^{-1}$. This is equivalent to finding ℓ^2 solutions $(u_n)_{n=0}^\infty$ of the homogeneous equation $(\tau - \lambda)(u_n)_{n=0}^\infty = 0$. The solutions $(u_n(z; J))_{n=1}^\infty$ are called the **Jost solutions** of J and the function $u(z; J) := u_0(z; J)$ on $\mathbb{D}$ is called the **Jost function** of J. An equivalent characterization is

$$\lim_{n\to\infty} \frac{z^n}{[(\lambda - J)e_1]_n} = u(z; J),$$

so that $(u_n(z; J))_{n=1}^\infty = -u_0(z; J)(J - \lambda)^{-1}e_1$. In other words, we are interested in Jost solutions that asymptotically look like those for J_0. In the free case, we have the Jost solutions $u_n(z) = z^n$ for $n \geq 0$ and the Jost function is $u(z; J_0) \equiv 1$, and we shall see that $u(z; J)$ satisfies the asymptotic

$$\frac{u_n(z; J))}{z^n} \to 1$$

for $z \in \mathbb{D}$.

2.3.1 Truncated Jacobi Operators

For a given Jacobi matrix J and fixing $\lambda = z + z^{-1}$, we define $\tilde{J}_l$ by taking $\tilde{J}_l = J$ for $a_n, b_n \leq l$ and $a_n = 1, b_n = 0$ for $n \geq l+1$. So $\tilde{J}_l$ agrees with J up to index l and agrees with J_0 for $n \geq l+1$.

Note that this implies that $\tilde{J}_l - J_0$ is finite range. We denote the Jost solutions of $(\tilde{J}_l - \lambda)u = 0$ by $u_n(z;\tilde{J}_l)$, $n \geq 1$. Clearly since $\tilde{J}_l$ and J_0 agree for $n \geq l+1$, we have that $u_n(z;\tilde{J}_l) = z^n$ for $n \geq l+1$. Considering these operators is standard (see [1] and [10]) since the particular Jacobi matrices that we will consider are compact perturbations of J_0. Let $p_{n-1}(z+z^{-1};\tilde{J}_l)$ be the orthogonal polynomials associated with $\tilde{J}_l$. To simplify notation, we write $u_n(z) = u_n(z;\tilde{J}_l)$ and $p_{n-1}(\lambda) = p_{n-1}(z+z^{-1};\tilde{J}_l)$ unless otherwise stated. We first have

Lemma 2.3.1. *The Jost function of $\tilde{J}_l$ is the Wronskian of the Jost solutions and the orthogonal polynomials associated with $\tilde{J}_l$.*

Proof. The Wronskian of u_n and p_{n-1} is

$$u(z;\tilde{J}_l) = W_n(u_n(z), p_{n-1}(\lambda); \tilde{J}_l) = a_n(u_n(z)p_n(\lambda) - u_{n+1}(z)p_{n-1}(\lambda)).$$

Since these are solutions of $\tilde{J}_l$, we have that the Wronskian is constant with respect to n. Taking $n = 1$ and using the relation

$$a_1u_2(z) + (b_1 - \lambda)u_1(z) = u_0,$$

we see that

$$\begin{aligned} u(z;\tilde{J}_l) &= a_1(u_1(z)p_1(\lambda) - u_2(z)p_0(\lambda)) \\ &= u_1(z)(a_1p_1(\lambda)) - a_1u_2(z) \\ &= u_1(z)(\lambda - b_1) - (\lambda - b_1)u_1(z) + u_0(z) \\ &= u_0(z;\tilde{J}_l), \end{aligned}$$

since $p_0(\lambda) = 1$ and $a_1p_1 = (\lambda - b_1)$ (as in the proof of Theorem 2.1.7). □

We would like to characterize the Jost function of J, so since the truncated matrices $\tilde{J}_l$ approach J as $l \to \infty$, we can use the Jost functions/solutions for the $\tilde{J}_l$ to extract information about $u_0(z;J)$.

2.3.2 The Geronimo-Case Equations

For the truncated Jacobi operators $\tilde{J}_l$, we can derive a nice set of equations to describe the evolution of their Jost functions. These were introduced in [4] and [3], and were also used in the presentation of Simon and Damanik (see [1]). From the recurrence relation and the fact that $a_n = 1, b_n = 0$ for $n \geq l+1$, we see that

$$\begin{aligned} a_{l+2}u_{l+2}(z) + (b_{l+1} - \lambda)u_{l+1}(z) + a_l u_l(z) &= z^{l+2} - (z+z^{-1})z^{l+1} + a_l u_l(z) \\ &= a_l u_l(z) - z^l = 0, \end{aligned}$$

so that $u_l(z) = a_l^{-1} z^l$. Now taking the Wronskian at $n = l$, we see that

$$\begin{aligned} u(z; \tilde{J}_l) &= a_l(u_l(z)p_l(\lambda) - u_{l+1}(z)p_{l-1}(\lambda)) \\ &= a_l(a_l^{-1} z^l p_l(\lambda) - z^{l+1} p_{l-1}(\lambda)) \\ &= z^l(p_l(\lambda) - a_l z p_{l-1}(\lambda)). \end{aligned}$$

Since we must have that $p_m(\lambda; J) = p_m(\lambda; \tilde{J}_l)$ for all $m \leq l$, it is natural to define the following sequences:

$$g_n(z) = z^n(p_n(z+z^{-1}) - a_n z^2 p_{n-1}(z+z^{-1}))$$

and

$$c_n(z) = z^n p_n(z+z^{-1}),$$

for the orthogonal polynomials associated with J. Clearly g_n and c_n are polynomials, with $\deg(g_n) \leq 2n$ and $\deg(c_n) = 2n$, and by definition, we see that g_n is the Jost function for $\tilde{J}_n$, and clearly $g_0(z) = c_0(z) \equiv 1$. We see that

$$g_{n+1}(z) = c_{n+1}(z) - a_{n+1} z^2 c_n(z).$$

The recurrence relation for p_n shows that

$$a_{n+1}p_{n+1}(z+z^{-1}) = (z+z^{-1} - b_{n+1})p_n(z+z^{-1}) - a_n p_{n-1}(z+z^{-1}).$$

Multiplying through by z^{n+1}, we see that

$$a_{n+1}z^{n+1}p_{n+1}(z+z^{-1}) = (z^2 + 1 - zb_{n+1})z^n p_n(z+z^{-1}) - a_n z^{n+1} p_{n-1}(z+z^{-1}),$$

or

$$\begin{aligned}
a_{n+1}c_{n+1}(z) &= (z^2+1-zb_{n+1})c_n(z) - a_n z^2 c_{n-1}(z) \\
&= (z^2-zb_{n+1})c_n(z) + c_n(z) - a_n z^2 c_{n-1}(z) \\
&= (z^2-zb_{n+1})c_n(z) + g_n(z).
\end{aligned}$$

Using this equation and the definition of $g_n(z)$, we see that

$$\begin{aligned}
a_{n+1}g_{n+1}(z) &= a_{n+1}c_{n+1}(z) - a_{n+1}^2 z^2 c_n(z) \\
&= (z^2-zb_{n+1})c_n(z) + g_n(z) - a_{n+1}^2 z^2 c_n(z) \\
&= [(1-a_{n+1}^2)z^2 - zb_{n+1}]c_n(z) + g_n(z).
\end{aligned}$$

The two recurrence relations

$$\begin{aligned}
a_{n+1}c_{n+1}(z) &= (z^2-zb_{n+1})c_n(z) + g_n(z) \\
a_{n+1}g_{n+1}(z) &= [(1-a_n^2)z^2 - zb_{n+1}]c_n(z) + g_n(z)
\end{aligned}$$

are known as the **Geronimo-Case equations** (or GC equations for short). Defining the update matrix

$$U_n(z) = \begin{pmatrix} z^2 - zb_{n+1} & 1 \\ (1-a_n^2)z^2 - zb_n & 1 \end{pmatrix},$$

we see that

$$\begin{pmatrix} c_{n+1} \\ g_{n+1} \end{pmatrix} = \frac{1}{a_{n+1}} U_{n+1} \begin{pmatrix} c_n \\ g_n \end{pmatrix}.$$

Taking $T_n = U_n \ldots U_1$, we have that

$$\begin{pmatrix} c_{n+1} \\ g_{n+1} \end{pmatrix} = \left(\prod_{i=1}^{n+1} a_i\right)^{-1} T_{n+1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For what follows, it is more convenient to work with the functions

$$\begin{aligned}
C_n(z) &= \left(\prod_{i=1}^{n+1} a_i\right)^{-1} c_n(z) \\
G_n(z) &= \left(\prod_{i=1}^{n+1} a_i\right)^{-1} g_n(z),
\end{aligned}$$

with $G_0(z) = C_0(z) \equiv 1$, so that

$$\begin{pmatrix} C_n \\ G_n \end{pmatrix} = T_n \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

In the next chapter, we will extract information about the analyticity of the Jost function from assumptions regarding the decay rates of the Jacobi parameters. As an example, we first prove

Lemma 2.3.2. *If $J - J_0$ is finite range, then the Jost function $u(z; J)$ is a polynomial with real coefficients.*

Proof. If $J - J_0$ is finite range, then $J = J_0$ for all indices greater than some index l, so that $a_n = 1$ and $b_n = 0$ for all $n \geq l + 1$. In this case, we have that

$$u(z; J) = g_l(z) = a_l^{-1}[(1 - a_l^2)z^2 - b_l z]c_{l-1}(z) + a_l^{-1} g_{l-1}(z)$$

is a real polynomial. □

We can even determine the degree of u: We know that deg$(g_{l-1}) \leq 2l - 2$, and we know that $c_{l-1} = (\prod_{i=1}^{l+1} a_i)^{-1} z^{2l-2} +$ lower order. So if $a_l \neq 1$, then deg$(u) = 2l$, and if $a_l = 1$ and $b_l \neq 0$, then deg$(u) = 2l - 1$.

Chapter 3

Determining the Jost Function from the Jacobi Parameters

In this section, we will use a series of successively stronger hypotheses on the Jacobi parameters $(b_n)_{n=1}^{\infty}$ and $(a_n)_{n=1}^{\infty}$ to determine the behavior of the Jost function $u(z;J)$. This is essentially a summary of the work of Damanik and Simon in [1].

3.1 Decay Rates of Jacobi Parameters

Now we make a series of hypotheses on the rate of convergence of the Jacobi parameters of J. First we assume that

$$\sum_{n=1}^{\infty} |a_n^2 - 1| + |b_n| < \infty. \tag{3.1}$$

Note first that since $\sum_{n=1}^{\infty} |a_n - 1|$ is finite, we have that $\prod_{n=1}^{\infty} a_n$ converges. This means that the partial products are uniformly bounded, so it suffices to look at C_n and G_n instead of c_n and g_n. In this case we have

Lemma 3.1.1. *If the Jacobi parameters satisfy (3.1), then $|C_n(z)| + |G_n(z)|$ is uniformly bounded over n on compact subsets of $z \in \mathbb{D} \setminus \{\pm 1\}$.*

Proof. Since we know that

$$\begin{pmatrix} C_n(z) \\ G_n(z) \end{pmatrix} = T_n \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

it suffices to show that $\sup_n \|T_n\|_\infty < \infty$. First note that we can write

$$\begin{aligned} U_n(z) &= \begin{pmatrix} z^2 - zb_n & 1 \\ (1-a_n^2)z^2 - zb_n & 1 \end{pmatrix} \\ &= \begin{pmatrix} z^2 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -zb_n & 0 \\ (1-a_n^2)z^2 - zb_n & 0 \end{pmatrix} \\ &= U^{(0)}(z) + A_n(z). \end{aligned}$$

As $z \neq \pm 1$, we see that $U^{(0)}$ is diagonalized by

$$L(z) = \begin{pmatrix} 1 & \frac{-1}{1-z^2} \\ 0 & 1 \end{pmatrix},$$

so we take

$$K_0(z) = L(z)U^{(0)}(z)L(z)^{-1} = \begin{pmatrix} z^2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Taking $B_n(z) = L(z)A_n(z)L(z)^{-1}$, we see that since

$$\begin{aligned} K_0(z) + B_n(z) &= L(z)U^{(0)}L(z)^{-1} + L(z)A_n(z)L(z)^{-1} \\ &= L(z)[U^{(0)} + A_n(z)]L(z)^{-1} \\ &= L(z)U_n(z)L(z)^{-1}, \end{aligned}$$

we have that

$$\begin{aligned} L(z)T_n(z)L(z)^{-1} &= [L(z)U_n(z)L(z)^{-1}] \dots [L(z)U_1(z)L(z)^{-1}] \\ &= [K_0(z) + B_n(z)] \dots [K_0(z) + B_1(z)]. \end{aligned}$$

Since $|z| \leq 1$, we see that

$$\begin{aligned} \|K_0(z) + B_n(z)\|_\infty &\leq 1 + \|L(z)\|_\infty \|L(z)^{-1}\|_\infty \|A_n(z)\|_\infty \\ &\leq 1 + \left(1 + \frac{1}{|1-z^2|}\right)^2 (|a_n^2 - 1| + |b_n|). \end{aligned}$$

This shows that

$$\begin{aligned}\|T_n\|_\infty &\leq \|L(z)\|_\infty^2 \prod_{i=1}^{n}(\|K_0(z)+B_i(z)\|_\infty) \\ &\leq \left(1+\frac{1}{|1-z^2|}\right)^2 \prod_{i=1}^{n}\left(1+\left(1+\frac{1}{|1-z^2|}\right)^2(|a_n^2-1|+|b_n|)\right)\end{aligned}$$

for all n, so that

$$\sup_n \|T_n\|_\infty \leq \left(1+\frac{1}{|1-z^2|}\right)^2 \prod_{i=1}^{\infty}\left(1+\left(1+\frac{1}{|1-z^2|}\right)^2(|a_i^2-1|+|b_i|)\right)$$

Since $\sum_{n=1}^\infty(|a_n^2-1|+|b_n|)<\infty$ by hypothesis, we must have that the product on the right converges, and taking z on compact sets ensures that $(1+|1-z^2|^{-1})^2$ is uniformly bounded as well. This gives a uniform bound $A(z)$ and thus establishes the result. $\square$

Now suppose we have the slightly stronger condition that

$$\sum_{n=1}^{\infty} n(|a_n^2-1|+|b_n|)<\infty. \tag{3.2}$$

In this case we have

Lemma 3.1.2. *If the Jacobi parameters satisfy (3.2), then there exists $M<\infty$ such that*

$$\sup_{n,z\in\bar{\mathbb{D}}}|G_n(z)|\leq M$$

and

$$\sup_{n,z\in\bar{\mathbb{D}}}\frac{|C_n(z)|}{n+1}\leq M.$$

Proof. By the hypothesis, we know that

$$M=\prod_{i=1}^{\infty}(1+i(|a_i^2-1|+|b_i|))<\infty,$$

so it suffices to establish that

$$|G_n(z)|\leq \prod_{i=1}^{n}(1+i(|a_i^2-1|+|b_i|))$$

and

$$|C_n(z)| \leq (n+1)\prod_{i=1}^{n}(1+i(|a_i^2-1|+|b_i|))$$

for all n. We see that $|G_0(z)| = |C_0(z)| \leq 1$, so the case $n = 0$ is trivial. If the inequalities hold for $n = k$, then since $|z| \leq 1$we have that

$$\begin{aligned}
|G_{k+1}(z)| &\leq |(1-a_{k+1}^2)z^2 - b_{k+1}z||C_k(z)| + |G_k(z)| \\
&= (|(1-a_{k+1}^2| + |b_{k+1}|)(k+1)+1)\prod_{i=1}^{k}(1+i(|a_i^2-1|+|b_i|) \\
&= \prod_{i=1}^{k+1}(1+i(|a_i^2-1|+|b_i|).
\end{aligned}$$

Similarly, we see that

$$|C_{k+1}(z)| \leq [(k+1)(1+|b_{k+1}|)+1]\prod_{i=1}^{k}(1+i(|a_i^2-1|+|b_i|),$$

which is certainly less than $(k+2)\prod_{i=1}^{k+1}(1+i(|a_i^2-1|+|b_i|))$, since

$$\begin{aligned}
(k+1)(1+|b_{k+1}|)+1 &= (k+2)+(k+1)|b_{k+1}| \\
&\leq (k+2)+(k+2)(k+1)(|a_{k+1}^2-1|+|b_{k+1}|),
\end{aligned}$$

as all quantities are positive. □

Now suppose that there is an N such that for all $n \geq N$, we have the estimate

$$|a_n^2-1|+|b_n| \leq CR^{-2n}, \tag{3.3}$$

for some $C > 0$ and $R > 1$, so the Jacobi parameters decay exponentially. Then we have

Lemma 3.1.3. *If the Jacobi parameters satisfy 3.3, then there exists $K < \infty$ such that*

$$|G_n(z)| + |C_n(z)| \leq K[\max(1,|z|)]^{2n}$$

for all z such that $|z| < R$.

Proof. If $|z| < 1$, then the estimate in the proof of Lemma 3.1.3 gives the bound. For $|z| > 1$, we note that $|z| < |z|^2$, so that

$$\begin{aligned}\|K_0(z) + B_i(z)\|_\infty &\leq |z|^2 + \left(1 + \frac{1}{|1 - z^2|}\right)^2 (|a_i^2 - 1||z|^2 + |b_i||z|) \\ &\leq |z|^2 \left(1 + \left(1 + \frac{1}{|z|^2 - 1}\right)^2 (|a_i^2 - 1| + |b_i|)\right).\end{aligned}$$

Since $C_n(z)$ and $G_n(z)$ are analytic, the maximum modulus principle allows us to only check the estimate for $|z| = R - \varepsilon > 1$. Taking products from 1 to n, we see that

$$\begin{aligned}\|T_n(z)\|_\infty &\leq |R - \varepsilon|^{2n} \left(1 + \frac{1}{|R - 1|^2 - 1}\right)^2 \prod_{i=1}^{n} \left(1 + \left(1 + \frac{1}{|R - 1|^2 - 1}\right)^2 [|a_i^2 - 1| + |b_i|]\right) \\ &= |R - \varepsilon|^{2n} \beta \prod_{i=1}^{n} (1 + \beta[|a_i^2 - 1| + |b_i|]).\end{aligned}$$

For $n \geq N$, this becomes

$$\|T_n(z)\|_\infty \leq |R - \varepsilon|^{2n} \beta \prod_{i=1}^{N} (1 + \beta[|a_i^2 - 1| + |b_i|]) \prod_{i=N}^{n} (1 + \beta C R^{2i}).$$

But since $\sum_{i=1}^\infty R^{-2i}$ is geometric, it must converge for $R > 1$, so the product on the right is uniformly bounded by some $B > 0$. Taking $\alpha = \prod_{i=1}^\infty (1 + \beta[|a_i^2 - 1| + |b_i|])$, we get

$$\|T_n(z)\|_\infty \leq K|R - \varepsilon|^{2n},$$

where $K = \alpha\beta B$. □

3.2 Analyticity of the Jost Function

Evidently, the Jost solutions for J and $\tilde{J}_n$ agree for $j \leq n$. So we have that the Jost function for J satisfies

$$u(z; J) = \lim_{n\to\infty} u(z; \tilde{J}_n) = \lim_{n\to\infty} g_n(z),$$

provided that the limit exists. We show now that on the regions given for the bounds in the previous theorems, we have that $g_n(z) \to u(z; J)$ locally uniformly, which will imply that $u(z; J)$ is analytic on those regions.

Theorem 3.2.1. *If the Jacobi parameters satisfy any of (3.1), (3.2), or (3.3), then*

$$u(z;J) = \lim_{n\to\infty} g_n(z)$$

exists for all z on the regions given in the conclusions of Lemmas 3.1.1, 3.1.2, and 3.1.3, respectively.

Proof. By the remarks preceeding the statement of the theorem and the fact that $g_n(z)$ and $G_n(z)$ differ by a uniform bound over n, it suffices to prove the convegence of $G_n(z)$ for all z in the appropriate regions. Then the limit function $G_\infty(z)$ will differ from $u(z;J)$ by that same bound. We note that $G_n(z)$ converges locally uniformly if $|G_{n+1}(z) - G_n(z)| \to 0$ locally uniformly. So it is suffcient to prove the convergence of $\sum_{n=1}^\infty |G_{n+1}(z) - G_n(z)|$ on compact sets. The GC equations give the bound

$$\begin{aligned}|G_{n+1}(z) - G_n(z)| &= |([(1-a_{n+1}^2)z^2 - b_{n+1}z]C_n(z) + G_n(z)) - G_n(z)| \\ &\le (|1-a_{n+1}^2||z|^2 + |b_{n+1}||z|)|C_n(z)|.\end{aligned}$$

If z is contained in any compact set in $z \in \bar{\mathbb{D}} \setminus \pm 1$ and $\sum_{n=1}^\infty |1-a_n^2| + |b_n| < \infty$, then Lemma 3.1.1 shows that there is a uniform bound A for $|C_n(z)|$ over n, so that

$$|G_{n+1}(z) - G_n(z)| \le (|1-a_{n+1}^2| + |b_{n+1}|)A.$$

Summing over n gives the result. If $z \in \bar{\mathbb{D}}$ and $\sum_{n=1}^\infty n(|1-a_n^2| + |b_n|) < \infty$, then Lemma 3.1.2 shows that

$$|G_{n+1}(z) - G_n(z)| \le (|1-a_{n+1}^2| + |b_{n+1}|)(n+1)M$$

for all n. Summing over n gives the result again. If $|z| \le R - \varepsilon$ for $0 < \varepsilon << 1$ and $|a_n^2 - 1| + |b_n| \le CR^{-2n}$ for all n, then Lemma 3.1.3 shows that

$$\begin{aligned}|G_{n+1}(z) - G_n(z)| &\le (R-\varepsilon)^2(|1-a_{n+1}^2| + |b_{n+1}|)K[\max(1,|z|)]^{2n} \\ &\le CK(R-\varepsilon)^2\left(\frac{\max(1,|z|)}{R}\right)^{2n}.\end{aligned}$$

As $|z| < R$, summing over n gives a convergent geometric series on the right, which shows the result. $\square$

3.3 The Case of Super-exponential Decay

This section presents the first half of our contribution to the topic. We suppose that the Jacobi parameters satisfy

$$|a_n^2 - 1| + |b_n| \leq \frac{\alpha}{n^{\gamma n}} \tag{3.4}$$

for all n and some $\gamma > 0$. These results are analogous to those proved for CMV operators found in [12]. We wish to demonstrate the following

Theorem 3.3.1. *If the Jacobi parameters satisfy (3.4), then $u(z; J)$ is an entire function of finite growth order no greater than $\frac{2}{\gamma}$; that is, $u(z; J)$ satisfies*

$$|u(z; J)| \leq Ae^{BR^{2/\gamma}},$$

for $|z| > R$, where A and B are constants that do not depend on R.

Proof. For any R, there is an N such that

$$\frac{\alpha}{n^{\gamma n}} \leq \frac{1}{R^{2n}}$$

for $n \geq N$. But then we have that

$$|a_n^2 - 1| + |b_n| \leq \frac{1}{R^{2n}}$$

for all $n \geq N$. By Lemma 3.1.4, we must have that $u(z; J)$ is analytic for $|z| < R$. Since this holds for all $R > 1$, we must in fact have that $u(z; J)$ is entire. To compute the growth order ρ of u, let $|z| = R$. By Lemma 3.1.4 and the definition of growth order, it suffices to show that

$$|G_{n+1}(z)| < Ae^{BR^{\frac{2}{\gamma}+\varepsilon}} + C$$

for all n and $\varepsilon > 0$, where A, B, and C are positive constants that do not depend on R. For ease of notation, we note that it suffices to derive this bound with $\frac{2}{\beta}$ in place of $\frac{2}{\gamma} + \varepsilon$, where $\beta < \gamma$. Then since we have

$$\frac{\alpha}{n^{\gamma n}} \leq \frac{\alpha}{n^{\beta n}},$$

we see that from the proof of Theorem (4) that

$$
\begin{aligned}
|G_{n+1}(z) - G_0(z)| &\leq \sum_{k=1}^{n+1} |G_k(z) - G_{k-1}(z)| \\
&\leq \sum_{k=1}^{n+1} KR^{2k}(|a_{k+1}^2 - 1| + |b_{k+1}|) \\
&\leq K_0 \sum_{k=1}^{n+1} \frac{R^{2k}}{k^{\beta k}},
\end{aligned}
$$

where $K_0 = K\alpha$. We see that if take $N = \left\lfloor (2R^2)^{\beta^{-1}} \right\rfloor$, then

$$
\frac{R^{2k}}{k^{\beta k}} \leq \frac{1}{2^k},
$$

for all $k \geq N$, so we get

$$
\begin{aligned}
|G_{n+1}(z))| &\leq K_0 \sum_{k=0}^{N} \frac{R^{2k}}{k^{\beta k}} + K_0 \sum_{k=N+1}^{n} \frac{1}{2^k} + |G_0(z)| \\
&\leq K_0 \sum_{k=0}^{N} \frac{R^{2k}}{k^{\beta k}} + (K_0 + 1),
\end{aligned}
$$

since $G_0(z) \equiv 1$. Maximizing the function $\dfrac{R^{2x}}{x^{\beta x}}$ for $x > 0$ gives $x = e^{-1}R^{\frac{2}{\beta}}$. This gives the bound

$$
\begin{aligned}
\sum_{k=1}^{N} \frac{R^{2k}}{k^{\beta k}} &\leq N \frac{R^{2e^{-1}R^{\frac{2}{\beta}}}}{\left(e^{-1}R^{\frac{2}{\beta}}\right)^{\beta e^{-1}R^{\frac{2}{\beta}}}} \\
&= N \frac{R^{2e^{-1}R^{\frac{2}{\beta}}}}{e^{-\beta e^{-1}R^{\frac{2}{\beta}}} R^{2e^{-1}R^{\frac{2}{\beta}}}} \\
&= N e^{\frac{\beta}{e}R^{\frac{2}{\beta}}}.
\end{aligned}
$$

However, since $N \leq (2R^2)^{\beta^{-1}}$ and $\ln x < x$ for $x > 0$, we get

$$
\begin{aligned}
K_0 \sum_{k=1}^{N} \frac{R^{2k}}{k^{\beta k}} &\leq K_0 (2R^2)^{\frac{1}{\beta}} e^{\frac{\beta}{e}R^{\frac{2}{\beta}}} \\
&= K_0 2^{\frac{1}{\beta}} e^{\frac{\beta}{e}R^{\frac{2}{\beta}} + \ln R^{\frac{2}{\beta}}} \\
&< K_0 2^{\frac{1}{\beta}} e^{\frac{\beta}{e}R^{\frac{2}{\beta}} + R^{\frac{2}{\beta}}} \\
&= Ae^{BR^{\frac{2}{\beta}}},
\end{aligned}
$$

where $A = K_0 2^{\frac{1}{\beta}}$ and $B = 1 + \dfrac{\beta}{e}$. Letting $n \to \infty$ and taking $C = K_0 + 1$, we get

$$|u(z;J)| \leq Ae^{BR^{\frac{2}{\beta}}} + C.$$

This for all $\beta < \gamma$, so it follows that $u(z;J)$ has growth order no greater than $\dfrac{2}{\gamma}$. $\square$

Chapter 4

Controlling Decay of Jacobi Parameters through Analyticity of Jost Functions

In the previous section, we demonstrated how assumptions on the decay rates of the Jacobi parameters allows one to draw conclusions about the analyticity of the Jost functions. In this section, we shall prove a partial converse. To this end, we first state and prove a number of lemmas that help to characterize the roots of $u(z;J)$ for a given Jacobi matrix J, and relate it to the function

$$M(z;J) := -m\left(z+\frac{1}{z}\right) = -\int_{\sigma(J)} \frac{d\mu(\lambda)}{\lambda-(z+z^{-1})}$$

defined on $\mathbb{D}$. Note that this is the Weyl m-function composed with the uniformization discussed in section 2.2.2. We wish to construct partial inverses to Theorems 3.2.1 and 3.3.1. These results were summarized in [1], but we follow more closely the presenation in [8], since it is a bit more streamlined.

4.1 Characterization of the M-function and the Zeros of the Jost Function

Since the m-function is related to the Jost function by Theorem 2.1.6, there is a relationship to explore between $M(z;J)$ and $u(z;J)$. In particular, we shall see that constructing an analytic continuation of u is—at least under certain assumptions—equivalent to extending $M(z;J)$ meromorphically. In what follows, we define $J^{(l)}$ to be the Jacobi operator defined by $a_n^{(l)} = a_{n+l}$ and $b_n^{(l)} = a_{n+l}$, and in turn define the Jost solutions

$$u_n(z;J) = a_n^{-1} z^n u(z;J^{(n)}).$$

We will show that these solutions coincide with the Jost solutions for $J = \tilde{J}_l$ defined at the beginning in section 2.3. Evidently, if the parameters of J satisfy any of the hypotheses of Lemmas 3.1.1 - 3.1.3, then so do the parameters of each of the $J^{(l)}$, and if all of the $J^{(l)}$ satisfy those hypotheses, then so does J. So in the following theorems, when a statement is made about the behavior of the $u_n(z;J)$, it is understood to hold in the regions given in the statements of Lemmas 3.1.1 - 3.1.3, depending on which hypothesis is satisfied. First we have

Lemma 4.1.1. *The Jost solutions $u_n(z;J)$ satisfy the Jacobi relation*

$$a_n u_{n+1}(z;J) + (b_n - \lambda) u_n(z;J) + a_{n-1} u_{n-1}(z;J) = 0$$

on the appropriate region.

Proof. It suffices to prove that the $u_n(z;\tilde{J}_l)$ coincide with the Jost solutions defined in section 2.3, so first let the original Jost solutions be written as $v_n(z;\tilde{J}_l)$. For $n \geq 1$ and $k \geq l+1$, we have

$$v_n(z;\tilde{J}_l^{(k)}) = z^n = z^{-k} z^{n+k} = z^{-k} v_{n+k}(z;\tilde{J}_l).$$

Since $J^{(n)}$ shifts by n steps, this in fact holds for $k \geq 1$. For $n = 0$, we see that $v_n(z;\tilde{J}_l^{(k)})$ uses $a_0 = a_k$, while $v_n(z;\tilde{J}_l)$ uses $a_0 = 1$, so we must have

$$v_0(z;\tilde{J}_l^{(k)}) = a_k z^{-k} v_k(z;\tilde{J}_l),$$

which implies

$$v_k(z;\tilde{J}_l) = a_k^{-1} z^k v(z;\tilde{J}_l^{(k)}) = u_k(z;\tilde{J}_l).$$

Now, taking $l \to \infty$ we know that $v(z;\tilde{J}_l) = g_l(z;J) \to u(z;J)$, so we have

$$v_k(z;J) = a_k^{-1} z^k v(z;J^{(k)}) = u_k(z;J),$$

as desired. $\square$

The following lemma verifies the intuitive idea that the Jost solutions of J should asymptotically look like the free ones:

Lemma 4.1.2. *We have*

$$\lim_{n\to\infty} z^{-n} u_n(z;J) = 1$$

on the appropriate region.

Proof. Since $z^{-n}u_n(z;J) = a_n^{-1}u(z;J^{(n)})$ and $a_n \to 1$, it suffices to show that $G_\infty(z;J^{(n)}) \to 1$ as $n \to \infty$. In all cases, we have by the proof of Lemma 3.1.4 that

$$\begin{aligned} |G_\infty(z;J^{(n)}) - 1| &\leq \sum_{k=0}^{\infty} (|(a_{k+1}^{(n)})^2 - 1| \cdot |z|^2 + |b_{k+1}^{(n)}| \cdot |z|)|C_{k+1}(z,J^{(n)})| \\ &= \sum_{k=n+1}^{\infty} (|a_k^2 - 1| \cdot |z|^2 + |b_k| \cdot |z|)|C_{k-n}(z,J^{(n)})|. \end{aligned}$$

We have a uniform bound for $|C_{k-n}(z;J^{(n)})|$ over k. Since the sum converges for fixed z, its remainder tends to zero. $\square$

Since we will consider extensions of u beyond the unit disk, we must investigate its possible behavior in these regions. First we have

Lemma 4.1.3. *The only possible zeros of $u(z;J)$ on $\partial\mathbb{D}$ are $z = \pm 1$. If this is the case, then they must be simple.*

Proof. Suppose $z \in \partial\mathbb{D}$. Since $u_n(z;J)$ and $u_n(z^{-1};J)$ both satisfy $Ju = (z+z^{-1})u$, we must have that their Wronskian

$$W[u_n(z;J), u_n(z^{-1};J)] = a_n[u_{n+1}(z;J)u_n(z^{-1};J) - u_n(z;J)u_{n+1}(z^{-1};J)]$$

is constant. But since $a_n \to 1$ and $u_n(z;J)$ tends to z^n, we have that this expression tends to

$$z^{n+1}z^{-n} - z^n z^{-n-1} = z - z^{-1}.$$

Since $W[u_n(z;J), u_n(z^{-1};J)] = u_1(z;J)u_0(z^{-1};J) - u_1(z;J)u_0(z^{-1};J)$, we must have

$$u_1(z;J)u_0(z^{-1};J) - u_0(z;J)u_1(z^{-1};J) = z - z^{-1}.$$

Since $u(z;J)$ is real analytic on the real line, the Schwartz Reflection Principle implies that $u(\bar{z};J) = \overline{u(z;J)}$ on $\partial\mathbb{D}$. Since $z^{-1} = \bar{z}$, we get

$$u_1(z;J)\overline{u_0(z;J)} - \overline{u_1(z;J)\overline{u_0(z;J)}} = -2\mathrm{Im}(u_1(z;J)\overline{u_0(z;J)}) = -2\mathrm{Im}(z),$$

or taking $z = e^{i\theta}$:

$$\mathrm{Im}(u_1(e^{i\theta};J)\overline{u_0(e^{i\theta};J)}) = \sin\theta.$$

It follows that unless $\theta = 0$ or $\theta = \pi$, we must have $u_0(e^{i\theta};J) \neq 0$. If $u_0(\pm 1;J) = 0$, then we see that

$$-\mathrm{Re}\left(u_1(e^{i\theta};J)\frac{u_0(e^{-i\theta};J)}{-i\theta}\right) = \frac{\sin\theta}{\theta} \to 1 \text{ (or } -1),$$

as $\theta \to 0$ (or π), since $u_1(\pm 1;J)$ and $u_0(\pm 1;J)$ cannot both be zero. Then $u_0'(\pm 1;J) \neq 0$, so $z = \pm 1$ must be simple. □

Note that it follows as an immediate corollary of $u_1(z;J)u_0(z^{-1};J) - u_0(z;J)u_1(z^{-1};J) = z - z^{-1}$ that if $u_0(z;J) = 0$ for $0 < |z| < 1$, then $u_0(z^{-1};J) \neq 0$. We now show that $M(z)$ can be extended meromorphically:

Lemma 4.1.4. *$M(z)$ has a continuation to $\partial\mathbb{D} \setminus \{\pm 1\}$ that is finite and non-zero on $\partial\mathbb{D} \setminus \{\pm 1\}$, and*

$$|u(e^{i\theta};J)|^2\mathrm{Im}(M(z;J)) = \sin\theta.$$

Proof. From the previous lemma, we have that

$$|u_0(e^{i\theta};J)|^2\text{Im}\left(\frac{u_1(e^{i\theta};J)\overline{u_0(e^{i\theta};J)}}{|u_0(e^{i\theta};J)|^2}\right) = |u_0(e^{i\theta};J)|^2\text{Im}\left(\frac{u_1(e^{i\theta};J)}{u_0(e^{i\theta};J)}\right) = \sin\theta,$$

but then Lemma 4.1.3 shows that

$$|u(e^{i\theta};J)|^2\text{Im}(M(z;J)) = \sin\theta,$$

which suffices to define a continuation of M on $\partial\mathbb{D}$. By Lemma 4.1.3, this continuation has at worst simple poles at $z = \pm 1$, since Theorem 2.1.6(a) shows that the poles of M coinide with the zeroes of u, and the orders must be the same. □

This formula will allow us to extend M beyond the unit disk in the next section.

4.2 The Case of Exponential Decay with No Bound States

Since $\tilde{J}_n - J_0$ is finite range for all n, and since clearly $\tilde{J}_n$ converges in norm to J, we must have that J is a compact perturbation of J_0. Then by Weyl's theorem, it follows that since $J = (J - J_0) + J_0$ and $J - J_0$ is compact, we must have

$$\sigma_{ess}(J) = [-2, 2] = \sigma_{ess}(J_0).$$

However, even though J_0 has no isolated (or indeed, any) eigenvalues, it is possible that the discrete spectrum of J intersects $\mathbb{R} \setminus [-2, 2]$. Eigenvalues of J on this interval are known as **bound states** of J. As per Appendix A.3, the spectral measure for J can be decomposed as

$$d\mu = d\mu_{ac} + d\mu_s = \frac{1}{\pi}\text{Im}(m)dx + d\mu_s,$$

where m is the Weyl m-function for J. Since the support for $d\mu_s$ corresponds to the discrete spectrum of J, we have that if λ is a bound state for J, then $d\mu_s(\{\lambda\}) > 0$. In what follows, we shall assume that J has no bound states. This case is much simpler than the general case, where appropriate weights for the point masses must be considered. This implies that $u(z; J)$ has no zeroes in $\mathbb{D}$, and we will see that then $M(z)$ has no singularities. In the case of bound states, these singularities would have to be appropriately weighted for the following proof to extend to this case. In what follows, we are going to assume that u is in fact the Jost function for a Jacobi matrix J with parameters $\{a_n, b_n\}$ and that $M(z)$ satisfies

$$\left.\frac{M(z)}{z}\right|_{z=0} = 1,$$

but it is possible to prove that u does in fact correspond to a unique Jacobi that gives this condition, using an appropriate normalization of the spectral measure $d\mu$. This is outside the scope of this paper, so we shall take this as given. So our goal in this section is to prove the following partial inverse to Theorem 3.2.1:

Theorem 4.2.1. *Let $R > 1$. Suppose $u(z)$ is analytic on $\{z \mid |z| < R\}$ and real analytic on the real line, such that u is non-zero on $\overline{\mathbb{D}} \setminus \{\pm 1\}$, and if $u(\pm 1) = 0$, then these zeros are simple. Then we have for each $\varepsilon > 0$ that*

$$|a_n - 1| + |b_n| \leq \frac{K_\varepsilon}{(R-\varepsilon)^{2n}},$$

for some K_ε not depending on n.

This is the inverse of Lemma 3.1.4 under the hypothesis of Lemma 3.1.3. If u is entire with finite growth order ρ, then this theorem will give a similar inverse to Theorem 3.3.1. We define

$$u^{(n)}(z) = u(z; J^{(n)})$$

and

$$M^{(n)}(z) = M(z; J^{(n)})$$

with $u(z) = u(z;J) = u^{(0)}(z)$ and $M(z) = M(z;J) = M^{(0)}(z)$. Then Lemma 4.1.1 and Theorem 2.1.6(a) give

$$M^{(n)}(z) = \frac{a_{n+1}^{-1} z u(z; J^{(n+1)})}{u(z; J^{(n)})} = \frac{a_{n+1}^{-1} z (a_{n+1} z^{-n-1} u_{n+1}(z; J))}{a_n z^n u_n(z; J)} = \frac{u_{n+1}(z)}{a_n u_n(z)}.$$

This gives

$$\begin{aligned} u^{(n+1)}(z) = u(z; J^{(n+1)}) &= a_{n+1} z^{-n-1} u_{n+1}(z) \\ &= a_{n+1} z^{-1} (a_n z^{-n} u_n(z) M^{(n)}(z)) \\ &= a_{n+1} z^{-1} u^{(n)}(z) M^{(n)}(z), \end{aligned}$$

as well as

$$
\begin{aligned}
[M^{(n)}(z)]^{-1} = \frac{a_n u_n(z)}{u_{n+1}(z)} &= \frac{(z+z^{-1}-b_{n+1})u_{n+1}(z) - a_{n+1}u_{n+2}(z)}{u_{n+1}(z)} \\
&= z+z^{-1}-b_{n+1} - a_{n+1}^2 \frac{u_{n+2}(z)}{a_{n+1}u_{n+1}(z)} \\
&= z+z^{-1}-b_{n+1} - a_{n+1}^2 M^{(n+1)}(z),
\end{aligned}
$$

by the recurrence relation for the u_n. For small z, we see that

$$(z+z^{-1}-J)^{-1} = z(1-z(J-z))^{-1} = z(1+z(J-z)+z^2(J-z)^2+\dots),$$

so that near $z=0$ we have

$$\frac{M^{(n)}(z)}{z} = 1 + O(z).$$

Combining this with the above update equation we get

$$\left(\frac{M(z)}{z}\right)^{-1} = 1 - zb_{n+1} + z^2 - z^2 a_{n+1}^2(1+O(z)) = 1 - b_{n+1}z - (a_{n+1}^2-1)z^2 + O(z^3),$$

or

$$\frac{M(z)}{z} = 1 + b_{n+1}z - [(a_{n+1}^2-1)^2 - b_{n+1}]z^2 + O(z^3).$$

Note that this shows that for small $|z| \le 1/2$, we must have

$$\frac{1}{4}(|b_{n+1}| + |a_{n+1}^2-1|^2) \le \sup_{|z|\le 1/2} \left|\frac{M(z)}{z} - 1\right|,$$

so since $|a_n - 1| \le |a_n^2 - 1| \le |a_n^2-1|^2$, it suffices to find an appropriate bound on $z^{-1}M^n(z) - 1$. Lemma 4.1.4 shows that $M(z)$ can be extended to the boundary of the unit disk minus $\{\pm 1\}$. If we wish to extend M outside of the boundary, we define $f^{\#}(z) = \overline{f(1/\bar{z})}$. Then $f : R \to \mathbb{C}$ is analytic on an open set $R \subset \mathbb{C}$ if and only if $f^{\#} : \{\bar{z}^{-1} : z \in R\} \to \mathbb{C}$ is analytic as well. It now makes sense to define

$$M(z) - M^{\#}(z) = [u(z)u^{\#}(z)]^{-1}(z - z^{-1}).$$

This coincides with the extension of M on the unit disk given in Lemma 4.1.4, and in fact this extends M outside the unit disk wherever u is analytic. In general, we even have:

Lemma 4.2.2. *For each $R > 1$, M has a continuation to $\{z : |z| < R\}$ if and only if u does, and the above holds on the annulus $R^{-1} < 1 < R$.*

Proof. The function

$$g(z) = \frac{z - z^{-1}}{M(z) - M^{\#}(z)}$$

satisfies $g^{\#} = g$, and since Lemma 4.1.4 shows that $M(z) - M^{\#}(z) \neq 0$ on $\partial\mathbb{D} \setminus \{\pm 1\}$, g must be meromorphic and real there. Since $g(e^{i\theta}) = |u(e^{i\theta})|^2$ and u is uniformly bounded on $\overline{\mathbb{D}}$, we must in fact have that g is analytic there, and so in a neighborhood of the annulus. It follows that in this neighborhood, an analytic continuation of u is given by $\tilde{u}(z) = g(z)/u^{\#}(z)$ and the Schwarz reflection principle, since $\tilde{u}(e^{i\theta}) = u(e^{i\theta})$, and the above formula continues this extension to the annulus. Conversely, if u is analytic in $\{z : |z| < R\}$, then since u does not vanish on $\partial\mathbb{D} \setminus \{\pm 1\}$, we have that

$$\tilde{M}(z) = [u(z)u^{\#}(z)]^{-1}(z - z^{-1}) + M^{\#}(z)$$

is analytic near $\partial\mathbb{D}$ and bounded away from $z = \pm 1$. So since again $\tilde{M}(e^{i\theta}) = M(e^{i\theta})$, the Schwarz reflection principle gives the continuation to a neighborhood of $\partial\mathbb{D}$, and the above formula continues this extension to the annulus. □

We wish to find a bound on $M(z)z^{-1}$ in terms of the $u^{(n)}$, so we must verify that they behave nicely on the same region as u. This is the content of the following lemma. Note in the proof that it is critical that u have no zeroes in the unit disk.

Lemma 4.2.3. *If $u(z)$ is real analytic on an open disk of radius $R > 1$ with no zeros on $\overline{\mathbb{D}} \setminus \{\pm 1\}$, and at most simple zeros at $z = \pm 1$, then the same is true for each $u^{(n)}(z)$. Similarly, $M^{(n)}(z)$ is meromorphic wherever M is.*

Proof. Inductively, it suffices to prove this for $u^{(1)}(z)$ and $M^{(1)}(z)$. We see that M can be extended to $\{z \mid 1 < |z| < R\}$ since $u(z)$ is analytic for $|z| < R$. Since

$$u^{(1)}(z) = a_1 z^{-1} u(z) M(z),$$

and we are assuming $z^{-1}M(z)|_{z=0} = 1$, the only possible singularities of $u^{(1)}$ are at ± 1. However, u has at most simple zeros at these points, and the extension of M above shows that any poles must be simple, so in fact $u^{(1)}$ must be analytic on $\overline{\mathbb{D}}$ by the factor $u(z)M(z)$. Since we also have

$$u^{(1)}(z) = a_1 z^{-1} u(z) M(z) = a_1(1 - z^{-2})(u^{\#}(z))^{-1} + a_1 z^{-} M^{\#}(z),$$

and since $u^{\#}(z)$ and $M^{\#}(z)$ are analytic for $1 < |z| < R$, we must in fact have that $u^{(1)}$ is analytic on the same region as u. Moreover, $u^{(1)}$ is non-vanishing on $\overline{\mathbb{D}}$ by our hypothesis that u has no zeros (and so M has no poles) in the unit disk, and Lemma 4.1.4 takes care of the boundary. Since

$$[M(z)]^{-1} = z + z^{-1} - b_1 - a_1^2 M^{(1)}(z),$$

we see that $M^{(1)}$ must be meromorphic wherever M is. □

We require one final lemma. The proof of this statement depends on a theorem of Killip and Simon in [5], the proof of which is very non-trivial and requires machinery far beyond the scope of this paper, so we state the lemma here without proof:

Lemma 4.2.4. *If $u(z; J)$ has finitely many zeros in $\mathbb{D}$ and the only zeros of u on the boundary are simple ones at $z = \pm 1$, then*

$$|a_n - 1| + |b_n| \to 0$$

and

$$M^{(n)}(z) \to z$$

uniformly on compact subsets of the unit disk.

Note that this implies

$$\sup_{|z| \leq \rho} \left| \frac{M^{(n)}(z)}{z} \right| \to 1$$

as $n \to \infty$ for all $\rho < 1$. We can now combine the formula for $u^{(n+1)}(z)$ and the extension of $M(z)$ to get

$$\begin{aligned}
u^{(n+1)}(z) &= a_{n+1} z^{-1} u^{(n)}(z) M^{(n)}(z) \\
&= a_{n+1} z^{-1} u^{(n)}(z)([u^{(n)}(z) u^{(n)\#}(z)]^{-1}(z - z^{-1}) + M^{(n)\#}(z)) \\
&= a_{n+1}(1 - z^{-2}) u^{(n)\#}(z)^{-1} + a_{n+1} z^{-2} u^{(n)}(z) N_n^{\#}(z),
\end{aligned}$$

where $N_n(z) = M^{(n)}(z)/z$. By Lemma 4.2.3, if $u(z)$ is analytic for $|z| < R$, then so is $u^{(n)}(z)$, and so the Laurent series of $(1 - z^{-2}) u^{(n)\#}(z)$ about $|z| = R - \varepsilon$, for $\varepsilon > 0$ such that $1 < R - \varepsilon < R$, contains only non-positive powers. Let $R_1 = R - \varepsilon$. Consider the space $L^2(R_1 \partial\mathbb{D}, d\theta/2\pi)$. If f has the Laurent series

$$\sum_{n=-\infty}^{\infty} a_n z^n = \sum_{n=-\infty}^{\infty} b_n e^{in\theta}$$

around $|z| = R_1$, then the same series taken from $n = 1$ defines an L^2 function f_+ on $R_1 \partial\mathbb{D}$, and these functions clearly form a closed subspace of L^2. We can now define a projection P_+ onto positive powers $\{e^{in\theta}\}$. Since $u^{(n)}$ is analytic about $|z| = R_1$, this projection acts as identity on $u^{(n+1)}(z) - u^{(n+1)}(0)$, and since $(1 - z^{-2}) u^{(n)\#}(z)$ has only non-positive powers, using the above equation and applying the projection P_+ gives

$$u^{(n+1)}(z) - u^{(n+1)}(0) = a_{n+1} P_+ \left[(R_1 e^{i\theta})^{-2} [u^{(n)}(R_1 e^{i\theta}) - u^{(n)}(0)] N_n^{\#}(z) \right].$$

Defining

$$|||g|||_{R_1} = \left(\int_0^{2\pi} |g(R_1 e^{i\theta}) - g(0)|^2 \frac{d\theta}{2\pi} \right)^{1/2},$$

we see that since $\|P_+\| = 1$ we must have

$$\begin{aligned}
\left|\left|\left| u^{(n+1)} \right|\right|\right|_{R_1} &= \left(\int_0^{2\pi} \left| a_{n+1} P_+ \left[(R_1 e^{i\theta})^{-2} [u^{(n)}(R_1 e^{i\theta}) - u^{(n)}(0)] N_n^{\#}(z) \right] \right|^2 \frac{d\theta}{2\pi} \right)^{1/2} \\
&\le a_{n+1} R_1^{-2} \|N_n^{\#}(z)\|_\infty \left(\int_0^{2\pi} \left| u^{(n)}(R_1 e^{i\theta}) - u^{(n)}(0) \right|^2 \frac{d\theta}{2\pi} \right)^{1/2} \\
&= a_{n+1} R_1^{-2} \|N_n^{\#}(z)\|_\infty \left|\left|\left| u^{(n)} \right|\right|\right|_{R_1}.
\end{aligned}$$

We are now ready to prove Theorem 4.2.1.

Proof of Theorem 4.2.1:

We see that

$$\sup_\theta |N_n^\#(R_1 e^{i\theta})| = \sup_\theta |R_1 e^{i\theta} M^{(n)}(R_1^{-1}e^{i\theta})| = \sup_\theta \left| \frac{M^{(n)}(R_1^{-1}e^{i\theta})}{R_1^{-1}} \right| \leq \sup_{|z|\leq R_1^{-1}} \left| \frac{M^{(n)}(z)}{z} \right|.$$

By Lemma 4.2.4, both this quantity and the a_n approach 1 as $n \to \infty$. This shows that

$$\lim_{n\to\infty} \left(\prod_{j=0}^{n-1} a_{j+1} \|N_j^\#(R_1 e^{i\theta})\|_\infty \right) < \infty,$$

so the nth roots tend to 1:

$$\lim_{n\to\infty} \left(\prod_{j=0}^{n-1} a_{j+1} \|N_j^\#(R_1 e^{i\theta})\|_\infty \right)^{1/n} = 1.$$

Now we see that

$$\begin{aligned} \prod_{j=0}^{n-1} \left\|\left| u^{(j+1)} \right\|\right|_{R_1} &\leq \prod_{j=0}^{n-1} \left(a_{j+1} R_1^{-2} \|N_j^\#(R_1 e^{i\theta})\|_\infty \left\|\left| u^{(j)} \right\|\right|_{R_1} \right) \\ &= R_1^{-2n} \prod_{j=0}^{n-1} \left(a_{j+1} \|N_j^\#(R_1 e^{i\theta})\|_\infty \right) \prod_{j=0}^{n-1} \left(\left\|\left| u^{(j)} \right\|\right|_{R_1} \right), \end{aligned}$$

so that

$$\left\|\left| u^{(n+1)} \right\|\right|_{R_1}^{1/n} \leq R_1^{-2} \left(\prod_{j=0}^{n-1} \left(a_{j+1} \|N_j^\#(R_1 e^{i\theta})\|_\infty \right) \right)^{1/n} \left(\|| u \||_{R_1} \right)^{1/n}.$$

By Lemma 4.1.2. and the remarks above, the n-th roots on the right tend to 1 as $n \to \infty$, so we can pick a uniform bound C_ε over n such that these quantities are less than $C_\varepsilon(1+\varepsilon)^n$. Then we have

$$\left\|\left| u^{(n+1)} \right\|\right|_{R_1} \leq \tilde{C}_\varepsilon (R - \varepsilon)^{-2n},$$

where $\tilde{C}_\varepsilon = C_\varepsilon(1+\varepsilon)$ does not depend on n. Under the assumption that $M(z)/z|_{z=0} = 1$ and Lemma 4.2.2, we have that $M^{(n)}(z)/z|_{z=0} = 1$ for each n. Then the recurrence relation for $u^{(n)}$ shows that

$$u^{(n)}(0) = a_{n+1} u_n(0).$$

This shows that

$$\frac{M^{(n)}(z)}{z} = \frac{u^{(n+1)}(z)/u^{(n+1)}(0)}{u^{(n)}(z)/u^{(n)}(0)},$$

and that $u^n(0) = [a_n \dots a_1]u(0)$. Since Lemma 4.2.3 shows that $u^{(n)}(z)$ has no zeroes in $\mathbb{D}$, we see that $|u^{(n)}(z)| \geq \alpha$ for some positive α. Taking β as a uniform bound on $\prod_{i=1}^{\infty} a_i$, this and the above show that

$$\begin{aligned} \left| \frac{M^{(n)}(z)}{z} - 1 \right| &\leq \frac{1}{u(0)\alpha\beta} \left| [\beta u(0)] u^{(n+1)}(z) - u^{(n+1)}(0) \right| \\ &\leq |u^{(n+1)}(z) - u^{(n+1)}(0)| \begin{cases} \dfrac{1}{u(0)\alpha\beta} & \text{if } \beta u(0) \leq 1 \\ \dfrac{1}{\alpha} & \text{if } \beta u(0) > 1 \end{cases}. \end{aligned}$$

In either case, we call the bound constant A and note that it does not depend on n. By the Cauchy integral formula, for $|z| = R_1$ we must have

$$\begin{aligned} |u^{(n)}(z) - u^{(n)}(0)| &\leq \frac{1}{2\pi} \oint_{\partial R_1 \mathbb{D}} \frac{|u^{(n)}(w) - u^{(n)}(0)|}{|z-w|} dw \\ &= \int_0^{2\pi} \frac{|u^{(n)}(R_1 e^{i\theta}) - u^{(n)}(0)|}{|z - R_1 e^{i\theta}|} \frac{d\theta}{2\pi} \\ &\leq \sup_{|z|<R_1} \frac{1}{|z - R_1 e^{i\theta}|} \left|\left|\left| u^{(n)} \right|\right|\right|_{R_1}, \end{aligned}$$

for any $\varepsilon > 0$, so that

$$\sup_{|z|\leq 1} |u^{(n)}(z) - u^{(n)}(0)| \leq M \left|\left|\left| u^{(n)} \right|\right|\right|_{R_1},$$

where $M = \dfrac{1}{|R_0 - 1|}$. Combining this with the above and recalling the remark before the statement of Lemma 4.2.2, we see that

$$|b_n| + |a_n - 1| \leq 4MA \left|\left|\left| u^{(n+1)} \right|\right|\right|_R \leq K_\varepsilon (R - \varepsilon)^{-2n},$$

where $K_\varepsilon = 4MA\tilde{C}_\varepsilon$ does not depend on n, as was to be shown. $\square$

4.3 The Case of Super-exponential Decay with No Bound States

This section presents the second half of our contribution to the topic, and follows closely the presentation in [12], which allows for a very clean proof of the following theorem. If $u(z)$ is an entire function with growth order ρ, then by the definition of growth order we have for all $\beta > \rho$ that $|u(z)| \leq Ae^{BR^{\beta}}$ for $|z| = R$, where A and B are positive constants that do not depend on R. We wish to prove the following:

Theorem 4.3.1. *Suppose $u(z)$ is an entire function of finite growth order ρ that is real analytic on the real line, such that u is non-zero on $\overline{\mathbb{D}} \setminus \{\pm 1\}$, and if $u(\pm 1) = 0$, then these zeros are simple. Then we have for all $\beta > \rho$ that*

$$|a_n - 1| + |b_n| \leq \frac{C}{n^{\frac{2}{\beta}n}}$$

where C is a constant depending only on β (and not n). In other words, the Jacobi parameters decay super-exponentially at a rate no more than $2/\rho$.

Proof. For ease of notation, we take $\alpha_n = |a_n - 1| + |b_n|$. First we note that it is sufficient to prove

$$\limsup_{n\to\infty} \frac{n \ln n}{-\ln \alpha_n} \leq \frac{\beta}{2}$$

for all $\beta > \rho$. For simplicity, we will suppose that $|u(z) - u(0)| \leq Ce^{R^{\beta}}$, where $R = |z|$, $\beta > \rho$, and C is a constant depending on β (it will be clear how to modify the proof to accomodate the general case). We have from the proof of Theorem 4.2.1 that

$$\alpha_n \leq K \left|\left|\left| u^{(n+1)} \right|\right|\right|_R \leq R^{-2n} N_n |||u|||_R$$

for any disk of radius R, and N_n is uniformly bounded over n. As u is entire, we are free to choose $R = n^{1/\beta}$, so that a Cauchy estimate gives

$$\begin{aligned} \alpha_n &\leq n^{-\frac{2}{\beta}n} N_n \left(\int_0^{2\pi} |u(Re^{i\theta}) - u(0)|^2 \frac{d\theta}{\pi} \right)^{1/2} \\ &\leq MN_n n^{-\frac{2}{\beta}n} e^n, \end{aligned}$$

where $M = \sqrt{2}C$. Inverting and taking logs gives

$$-\ln \alpha_n \geq \frac{2}{\beta} n \ln n - \ln M - \ln N_n - n.$$

If we take $\gamma_n = \ln M + \ln N_n + n$, then since N_n is uniformly bounded over n, we must have that $\frac{\gamma_n}{n} \to 1$ as $n \to \infty$. Then we have

$$\begin{aligned} \frac{n \ln n}{-\ln \alpha_n} &\leq \frac{n \ln n}{\frac{2}{\beta}(n \ln n - \frac{\beta}{2}\gamma_n)} \\ &= \frac{n \ln n - \frac{\beta}{2}\gamma_n + \frac{\beta}{2}\gamma_n}{\frac{2}{\beta}(n \ln n - \frac{\beta}{2}\gamma_n)} \\ &= \frac{\beta}{2} + \frac{\frac{\beta}{2}\gamma_n}{\frac{2}{\beta} n \ln n - \gamma_n} \\ &= \frac{\beta}{2} + \frac{\frac{\beta}{2}\frac{\gamma_n}{n}}{\frac{2}{\beta}\ln n - \frac{\gamma_n}{n}}. \end{aligned}$$

Evidently the fraction on the right tends to 0 for large n thanks to the presence of the $\ln n$ term, so we must have that

$$\limsup_{n \to \infty} \frac{n \ln n}{-\ln \alpha_n} \leq \frac{\beta}{2}.$$

This for all $\beta > \rho$, so the result follows. $\square$

Chapter 5

Future Work

5.1 The Case of Bound States

The proof of Lemma 4.2.2 relies on the fact that u has no bound states. If $u(z) = 0$ for some $z \in \mathbb{D}$, then there is no guarantee that the extension of $M(z)$ to $|z| < R$ will agree with the weights w_j that are required to determine $d\mu$ from u. In the third section of [1] (as well as [8]), Damanik and Simon extend their result to the case where u has bound states. It should be possible to use their results in the case where u is entire of finite growth order. The possible values of the weights w_j are given by

$$\sum_j w_j + \frac{2}{\pi} \int_0^\pi \frac{\sin^2 \theta}{|u(e^{i\theta})|^2} d\theta = 1.$$

Suppose u is analytic on a neighborhood of $\bar{\mathbb{D}}$ and real analytic on the real line. If all of the zeros of u occur on $(\bar{\mathbb{D}} \cap \mathbb{R}) \setminus \{0\}$, and these are simple zeros, then there is a unique measure for which u is the Jost function and the w_j are the weights. In this case, $M(z)$ can be extended meromorphically to a neighborhood of $\bar{\mathbb{D}}$ in a way consistent with the weights for u. In particular, we may take

$$w_j = \frac{z_j - z_j^{-1}}{z_j u_0'(z_j) u_0^{\#}(z_j)}.$$

These are known as the canonical weights for u_0. The methods of section 4 can then be applied to show the result in the case of bound states. The relevant theorems can be found in [8, p. 918].

Appendix A

Measure Theory

A.1 Borel Measures and Distributions

The results in this section are fairly standard and summarized in [11]. The main result is an application of vague convergence of Borel measures. First we let μ be a measure on $\mathbb{R}$ and define the **Borel algebra** to be the smallest σ-algebra on $\mathbb{R}$ that contains all open intervals. These sets are called the **Borel sets** of $\mathbb{R}$. The measure μ is called a **Borel measure** if $\mu(C) < \infty$ for all compact sets C in $\mathbb{R}$. For any Borel measure μ, its **distribution** is defined to be the function

$$d(x) = \begin{cases} -\mu((x,0]) & x < 0 \\ 0 & x = 0 \\ \mu((0,x]) & x > 0 \end{cases} .$$

By the monotonicity of the measure, d must be non-decreasing. Suppose that $a > 0$. Then since μ is continuous from above, we have that $B_{n+1} \subset B_n$ implies that $\mu(B_n) \to \mu(\cap_n B_n)$, provided that $\mu(B_1)$ is finite. Since Borel measures assign finite measure to closed intervals, monotonicity implies that μ assigns finite measure to any finite interval. So if (a_n) is a monotone decreasing sequence converging to a, then we see that $(0,a] \subset (0,a_{n+1}] \subset (0,a_n]$ implies

$$\lim_{n\to\infty} d(a_n) = \lim_{n\to\infty} \mu((0,a_n]) = \mu((0,a]) = d(a),$$

which shows that d is right-continuous. The cases for $a \leq 0$ are shown similarly. Also note μ begin continuous from above and below implies that $d(x) \to \mu((-\infty, 0])$ as $x \to -\infty$ and $d(x) \to \mu((0, \infty))$ as $x \to \infty$. In the other direction, let $\mathcal{T}$ be the algebra of finite unions of disjoint intervals of the form $(a, b]$ together with the empty set. Take $A = \cup_{i=1}^{n} A_i$ where $A_i = (a_i, b_i]$ are disjoint. For every right-continuous and non-decreasing function d on the real line, we define

$$\mu_*(A) = \sum_{i=1}^{n} [d(b_i) - d(a_i)].$$

This is well-defined since any partition of A can be expressed in the form $A' = \cup_{i=1}^{n} A'_i$ where the a_i are distinct from the b_i. Then since the sum is telescoping at those points where $a_{i+1} = b_i$, we get $\mu_*(A) = \mu_*(A')$. Now we can show the following:

Lemma A.1.1. *Every right-continuous, non-decreasing function $d : \mathbb{R} \to \mathbb{R}$ defines a unique Borel measure μ such that $\mu = \mu_*$ on $\mathcal{T}$. Two functions d_1 and d_2 both generate μ if and only if $d_1 = d_2 + C$ for some constant C.*

Proof. Since extensions of pre-measures are unique, it will suffice to show that μ_* defines a pre-measure for μ, and then we will check regularity. We first need to show that $\mathcal{T}$ generates the Borel algebra on $\mathbb{R}$. Clearly $\mathcal{T}$ is contained in the Borel algebra, so by definition the σ-algebra $\sigma(\mathcal{T})$ generated by $\mathcal{T}$ must be contained in the Borel algebra. Since every open set in $\mathbb{R}$ can be expressed as a union of open intervals, and any open interval (a, b) has the form $\cup_{n \geq 1}(a, b - 1/n]$, the σ-algebra generated by $\mathcal{T}$ must contain all the open sets. But by definition, the Borel algebra is the smallest σ-algebra with this property, so we must in fact have that $\sigma(\mathcal{T})$ is exactly the Borel algebra. We now need to see that μ_* defines a pre-measure on $\mathcal{T}$. We see that $\mu_*(\cdot) \geq 0$ since d is a non-decreasing function, and $\mu_*(\phi) = \mu_*((a, a]) = d(a) - d(a) = 0$. It suffices (by taking countable unions of countable unions) to check σ-additivity in the case that $A = \cup_{n=1}^{\infty} A_n = (a, b]$, where $A_n = (a_n, a_{n-1}]$, so that $a_0 = b$ and $a_n \to a^+$ as $n \to \infty$. We see that

$$\begin{aligned}
\mu_*(A) - \mu_*(\cup_{n=1}^{N} A_n) &= d(b) - d(a) - \sum_{n=1}^{N} [d(a_{n-1}) - d(a_n)] \\
&= d(b) - d(a) + d(a_N) - d(a_0) \\
&= d(a_N) - d(a) \to 0
\end{aligned}$$

as $N \to \infty$ by the right-continuity of d. It follows that μ_* extends to a unique Borel measure μ. This extension preserves regularity, so it suffices check that μ_* is regular. For simplicity, we check inner and outer regularity for $A = (a, b]$. Since A is open on the left and closed on the right, we may take our open sets to be of the form $O_\varepsilon = (a, b + \varepsilon)$ and our compact sets to be of the form $C_\varepsilon = [a + \varepsilon, b]$. We see that monotonicity of the pre-measure μ_* gives

$$d(b) - d(a) = \mu_*(A) \leq \mu_*(O_\varepsilon) \leq \mu_*((a, b + \varepsilon)) = d(b + \varepsilon) - d(a)$$

and

$$d(b) - d(a + \varepsilon) = \mu_*((a + \varepsilon, b]) \leq \mu_*(C_\varepsilon) \leq \mu_*(A) = d(b) - d(a),$$

so since d is right-continuous, $\mu_*(O_\varepsilon) - \mu_*(C_\varepsilon)$ can be made as small as we like, so we have that μ_* is regular. The last statement of the lemma is clear from the definition of μ_*. □

If μ is finite, it is usually more convenient to work with the non-negative function F defined by

$$F(x) = \mu((-\infty, 0]) + d(x).$$

Clearly F is also right-continuous, and we see that

$$F(x) = \begin{cases} \mu((-\infty, 0]) - \mu((x, 0]) & x < 0 \\ \mu((-\infty, 0]) & x = 0 \\ \mu((-\infty, 0]) + \mu((0, x]) & x > 0 \end{cases},$$

so that $F(x) = \mu((-\infty, x])$. In particular, we have that the continuity of μ from above and below gives $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$ and $F(\infty) = \lim_{x\to\infty} F(x) = \mu(\mathbb{R})$. We refer to F as the **normalized distribution** of μ. Furthermore, we call x a **point of continuity** of the measure μ if $\mu(\{x\}) = 0$. We see that x is a point of continuity for μ if and only if F is continuous at x:

$$\begin{aligned} \mu(\{x\}) = \mu(\cap_{n\geq 1}(x - 1/n, x]) &= \lim_{n\to\infty} \mu((x - 1/n, x]) \\ &= \lim_{n\to\infty} (\mu(-\infty, x]) - \mu((-\infty, x - 1/n])) \\ &= \lim_{n\to\infty} (F(x) - F(x - 1/n)). \end{aligned}$$

In particular this means that since F can only have countably many discontinuities, the same is true for μ. This is also true in the general case with d in place of F. Now suppose that μ_n is a sequence of Borel measures and suppose that μ is another Borel measure such that

$$\int_{\mathbb{R}} f\,d\mu_n \to \int_{\mathbb{R}} f\,d\mu$$

for all compactly supported continuous functions f. Then we say that the μ_n **converge vaguely** to the measure μ. We have the following lemma:

Lemma A.1.2. *A sequence of Borel measures (μ_n) on $\mathbb{R}$ converges vaguely to a Borel measure μ on $\mathbb{R}$ if and only if the distributions of the μ_n converge to the distribution of μ at every point of continuity of μ.*

Proof. Suppose that the μ_n converge vaguely to μ. For any bounded interval $I = (a, b]$, we can find continuous functions f and g with compact supports satisfying $f \leq \chi_I \leq g$. Then

$$\int_{\mathbb{R}} f\,d\mu_n \leq \mu_n(I) \leq \int_{\mathbb{R}} g\,d\mu_n,$$

and similarly for μ. Then we have that

$$\int_{\mathbb{R}} f\,d\mu_n - \int_{\mathbb{R}} f\,d\mu - \int_{\mathbb{R}} (g-f)\,d\mu \leq \mu_n(I) - \mu(I) \leq \int_{\mathbb{R}} (g-f)\,d\mu + \int_{\mathbb{R}} g\,d\mu_n - \int_{\mathbb{R}} g\,d\mu,$$

so the vague convergence of the μ_n shows that

$$\limsup_{n\to\infty} |\mu_n(I) - \mu(I)| \leq \int_{\mathbb{R}} (g-f)\,d\mu.$$

This for any $f, g \in C_c(\mathbb{R})$, so in particular we can choose f_k and g_k with $|g_k - f_k| \leq \chi_{(a-\delta,a+\delta)}$, for $\delta > 0$ fixed, such that $f_k \to \chi_{(a,b]}$ from below and $g_k \to \chi_{[a,b]}$ from above pointwise. Then the Lebesgue dominated convergence theorem shows that $\int_{\mathbb{R}}(g_k - f_k)d\mu$ converges to $\mu(\{a\})$. Thus we have that

$$\limsup_{n\to\infty} |d_n(a) - d(a)| \leq \limsup_{n\to\infty} |\mu_n(I) - \mu(I)| \leq \mu(\{a\}),$$

so that $d_n(a) \to d(a)$, provided that a is a point of continuity of μ. Conversely, suppose that $d_n(x) \to d(x)$ wherever $\mu(\{x\}) = 0$, and let $f \in C_0(\mathbb{R})$. Since the support of f is compact and f is continuous, f must be uniformly continuous. Fix $\varepsilon > 0$. Then there is a $\delta > 0$ such that

$|f(x) - f(y)| < \varepsilon$ for $|x - y| < \delta$. Find an interval $I = (a_0, a_N)$ containing the support of f and a partition $a_0 < a_1 < \cdots < a_{N-1} < a_N$ of I such that $a_{i+1} - a_i < \delta$ for $0 \leq i \leq N$. Since the discontinuities of μ are countable, we can arrange things so that $\mu(\{a_i\}) = 0$ for each i. Then for large enough n, we have $|d_n(x_i) - d(x_i)| < \varepsilon/(2N)$, so that

$$\mu_n((a_{i-1}, a_i]) = |d_n(a_{i-1}) - d_n(a_i)| < |d(a_{i-1}) - d(a_i)| + \frac{\varepsilon}{N} = \mu((a_{i-1}, a_i]) + \frac{\varepsilon}{N}.$$

From this and supp$(f) \subset (a_0, a_N)$, we see that

$$\begin{aligned}\left|\int_{\mathbb{R}} f\,d\mu_n - \sum_{i=1}^{N} f(a_{i-1})\mu_n((a_{i-1}, a_i])\right| &\leq \sum_{i=1}^{N} \int_{(a_{i-1}, a_i]} |f(x) - f(a_{i-1})|\,d\mu_n \\ &< \varepsilon \sum_{i=1}^{N} \int_{(a_{i-1}, a_i]} d\mu_n \\ &= \varepsilon \sum_{i=1}^{N} \mu_n((a_{i-1}, a_i]) \\ &< \varepsilon(\mu((a_0, a_N]) + \varepsilon).\end{aligned}$$

So we have

$$\begin{aligned}\left|\int_{\mathbb{R}} f\,d\mu_n - \int_{\mathbb{R}} f\,d\mu\right| &\leq \sum_{i=1}^{N} \int_{(a_{i-1}, a_i]} |f(x) - f(a_{i-1})|\,d\mu_n \\ &\quad + \sum_{i=1}^{N} \int_{(a_{i-1}, a_i]} |f(x) - f(a_{i-1})|\,d\mu \\ &\quad + \sum_{i=1}^{N} |f(a_{i-1})| \cdot |\mu_n((a_{i-1}, a_i]) - \mu((a_{i-1}, a_i])| \\ &< \varepsilon \cdot (\mu((a_0, a_N]) + \varepsilon) + \varepsilon \cdot \mu((a_0, a_N]) + \varepsilon \cdot \max_{1 \leq i \leq N} |f(a_{i-1})| \\ &= \varepsilon \cdot [2\mu((a_0, a_N] + \max_{1 \leq i \leq N} |f(a_{i-1})| + \varepsilon],\end{aligned}$$

which establishes that $\mu_n \to \mu$ vaguely. $\square$

We are now in a position to prove the following theorem:

Theorem A.1.3. *If (μ_n) is a sequence of finite Borel measures on the real line with $\mu_n(\mathbb{R}) \leq M$ for all n, then there exists a subsequence (μ_k) that converges vaguely to a unique measure μ with*

$\mu(\mathbb{R}) \leq M$. Moreover, we have that

$$\int_{\mathbb{R}} f\, d\mu_n \to \int_{\mathbb{R}} f\, d\mu$$

for all $f \in C_0(\mathbb{R})$.

Proof. Since the μ_n are all finite measures, we can consider the normalized distributions $F_n(x) = \mu_n((-\infty, x])$. By lemma A.1.2, it suffices to show that there is a subsequence whose distributions converge pointwise to another distribution F, since then lemma A.1.1 then shows that $F(x) = \mu((-\infty, x])$ gives the desired measure, provided that the distributions of the subsequence converge to F wherever F is continuous. Now let $\{r_i\}_{i=1}^{\infty}$ be an enumeration of the rationals. Since $F_n(r_1) \in [0, M]$, by the Heine-Borel theorem there exists a subsequence $n_k^{(1)} \subset n$ such that that $F_{n_k}^{(1)}(r_1) \to H(r_1)$. However, since $F_{n_k}^{(1)}(r_2) \in [0, M]$, we have the existence of a subsequence $n_k^{(2)} \subset n_k^{(1)}$ such that $F_{n_k}^{(2)}(r_2) \to H(r_2)$. Moreover, since $n_k^{(2)} \subset n_k^{(1)}$, we must have that $F_{n_k}^{(2)}(r_1) \subset F_{n_k}^{(1)}(r_1) \to H(r_1)$ as well. In general, we can find subsequences $n_k^{(j)}$ satisfying

$$n_k^{(j)} \subset n_k^{(j-1)} \subset \cdots \subset n_k^{(1)} \subset n$$

and

$$F_{n_k}^{(j)}(r_i) \to H(r_i)$$

for $i \leq j$. It follows that the diagonal function $H(r_i) = \lim_{k\to\infty} F_{n_k}^{(k)}(r_i)$ is non-decreasing since $r_i \leq r_j$ implies that $F_{n_k}^{(k)}(r_i) \leq F_{n_k}^{(k)}(r_i)$ for all k. Now for any $x \in \mathbb{R}$ we can define the function F by

$$F(x) = \inf_{r_i \geq x} H(r_i).$$

Then F is non-decreasing, because $x \leq y$ implies that $\inf_{r_i \geq x} H(r_i) \leq \inf_{r_i \geq y} H(r_i)$ since H is non-decreasing. If we suppose that $\varepsilon > 0$, then it is clear that

$$\begin{aligned} \lim_{\varepsilon\to 0^+} F(x+\varepsilon) &= \inf\{H(r_i) : r_i \geq x + \delta\,,\, 0 < \delta < \varepsilon\} \\ &= \inf\{H(r_i) : r_i \geq x\} = F(x), \end{aligned}$$

so that F is continuous from the right. So F is a distribution, and we have the existence of a unique Borel measure μ defined by $\mu((-\infty, x]) = F(x)$. We see that

$$\mu(\mathbb{R}) = \lim_{x\to\infty} F(x) = \lim_{\substack{r\to\infty \\ r\in\mathbb{Q}}} H(r) \leq M.$$

Now we must verify that $F_{n_k^{(k)}}(x) \to F(x)$ for all x such that F is continuous. Let $\varepsilon > 0$ and suppose that F is continuous at x. Then we can find a $y < x$ such that $F(x) - \varepsilon < F(y)$. We can find rational r_1 and r_2 such that $y < r_1 < x < r_2$ and $F(x) + \varepsilon > H(r_2)$. Then

$$F(x) - \varepsilon < F(y) < H(r_1) < H(r_2) < F(x) + \varepsilon.$$

Since $F_{n_k^{(k)}}(r_i) \to H(r_i)$ as $k \to \infty$, for large enough k we have that $H(r_1) \leq F_{n_k^{(k)}}(x) \leq H(r_2)$, and then

$$|F_{n_k^{(k)}}(x) - F(x)| < \varepsilon,$$

which shows that $F_{n_k^{(k)}}(x) \to F(x)$ as $k \to \infty$. For the last statement of the theorem, assume that f is continuous and vanishes at infinity. Let $\varepsilon > 0$ and write $f = f_1 + f_2$, where f_1 has compact support $\sigma = \operatorname{supp}(f_1)$ and $f_2 \leq \varepsilon$. Then since μ_n converges vaguely to μ, we have that

$$\begin{aligned}
\left|\int_{\mathbb{R}} f\, d\mu_n - \int_{\mathbb{R}} f\, d\mu\right| &\leq \left|\int_\sigma f_1\, d\mu_n - \int_\sigma f_1\, d\mu\right| + \left|\int_{\mathbb{R}\setminus\sigma} f_2\, d\mu_n - \int_{\mathbb{R}\setminus\sigma} f_2\, d\mu\right| \\
&\leq \left|\int_\sigma f_1\, d\mu_n - \int_\sigma f_1\, d\mu\right| + 2\varepsilon \\
&\to 2\varepsilon
\end{aligned}$$

as $n \to \infty$. This for all $\varepsilon > 0$, so the claim follows. □

A.2 Herglotz Functions and the Stieltjes Inversion Formula

The results in this section follow from previous one, and can also be found in [11]. Theorem A.1.3 can now be used to prove the following important theorem:

Theorem A.2.1. *If $F : \mathbb{C}^+ \to \mathbb{C}^+$ is a holomorphic function on the upper-half plane (i.e., F is a* ***Herglotz function****) and satisfies*

$$|F(z)| \leq \frac{M}{Im(z)}$$

for all $z \in \mathbb{C}^+$, then there exists a unique Borel measure μ such that $\mu(\mathbb{R}) \leq M$ and

$$F(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu(\lambda)$$

(i.e., F is the **Borel transform** *of μ).*

Proof. Suppose that F is a Herglotz function, so that $F(z)\mathrm{Im}(z) \leq M$ for some $M > 0$. For a fixed $z = x + iy \in \mathbb{C}^+$, define the contour

$$\Gamma = \{x + i\varepsilon + t \,:\, t \in [-R, R]\} \cup \{x + i\varepsilon + Re^{i\theta} \,:\, \theta \in [0, \pi]\} = \Gamma_1 \cup \Gamma_2.$$

By construction, we have that z is interior to Γ while $\bar{z} + 2i\varepsilon$ is exterior to Γ. Then since $F(\zeta)/(\zeta - \bar{z} - 2i\varepsilon)$ is holomorphic on an interior to Γ, its integral over that contour is zero by the Cauchy integral theorem. Moreover, since F is holomorphic on and interior to Γ, we have by the Cauchy integral formula that

$$\begin{aligned}
F(z) &= \frac{1}{2\pi i} \int_\Gamma \frac{F(\zeta)}{\zeta - z} d\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma \frac{F(\zeta)}{\zeta - z} d\zeta - \frac{1}{2\pi i} \int_\Gamma \frac{F(\zeta)}{\zeta - \bar{z} - 2i\varepsilon} d\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma \left(\frac{1}{\zeta - z} - \frac{1}{\zeta - \bar{z} - 2i\varepsilon} \right) F(\zeta)\, d\zeta \\
&= \frac{1}{\pi} \int_\Gamma \frac{y - \varepsilon}{(\zeta - z)(\zeta - \bar{z} - 2i\varepsilon)} F(\zeta)\, d\zeta,
\end{aligned}$$

since $\zeta - \bar{z} - 2i\varepsilon - (\zeta - z) = 2iy - 2i\varepsilon$. Using the substitution $\zeta = x + i\varepsilon + t$ for $t \in [-R, R]$, we see that

$$\begin{aligned}
\frac{1}{\pi} \int_{\Gamma_1} \frac{y - \varepsilon}{(\zeta - z)(\zeta - \bar{z} - 2i\varepsilon)} F(\zeta)\, d\zeta &= \frac{1}{\pi} \int_{-R}^{R} \frac{y - \varepsilon}{(t + i(\varepsilon - y))(t - i(\varepsilon - y))} F(x + i\varepsilon + t)\, dt \\
&= \frac{1}{\pi} \int_{-R}^{R} \frac{y - \varepsilon}{t^2 + (y - \varepsilon)^2} F(x + i\varepsilon + t)\, dt.
\end{aligned}$$

Letting R tend to infinity and shifting by $s = x + t$ we see that

$$\frac{1}{\pi}\int_{\Gamma_1} \frac{y-\varepsilon}{(\zeta - z)(\zeta - \bar{z} - 2i\varepsilon)} F(\zeta)\, d\zeta \to \int_{-\infty}^{\infty} \frac{y-\varepsilon}{(s-x)^2 + (y-\varepsilon)^2} \cdot \frac{F(s+i\varepsilon)}{\pi}\, ds.$$

Using the substitution $\zeta = Re^{i\theta}$, we see that

$$\frac{1}{\pi}\int_{\Gamma_2} \frac{y-\varepsilon}{(\zeta - z)(\zeta - \bar{z} - 2i\varepsilon)} F(\zeta)\, d\zeta = \frac{i}{\pi}\int_0^{\pi} \frac{(y-\varepsilon)F(Re^{i\theta} + x + i\varepsilon)}{(Re^{i\theta} + x + i\varepsilon)(Re^{i\theta} - y - i\varepsilon)} Re^{i\theta}\, d\theta.$$

From $F(z) \leq M\mathrm{Im}(z)^{-1}$, we have the bound

$$\begin{aligned} \left| \frac{i}{\pi}\int_0^{\pi} \frac{(y-\varepsilon)F(Re^{i\theta} + x + i\varepsilon)}{(Re^{i\theta} + x + i\varepsilon)(Re^{i\theta} - y - i\varepsilon)} Re^{i\theta}\, d\theta \right| &\leq \frac{2R^2|y-\varepsilon|}{R|R - |\varepsilon + y||} \sup_{\theta\in[0,\pi]} F(Re^{i\theta} + x + i\varepsilon) \\ &\leq \frac{2MR|y-\varepsilon|}{|R - |\varepsilon + y|| \cdot |\varepsilon + R\sin\theta|}, \end{aligned}$$

so the integral over Γ_2 tends to zero as $R \to \infty$. Thus we arrive at the representation

$$F(z) = \int_{-\infty}^{\infty} \frac{y-\varepsilon}{(s-x)^2 + (y-\varepsilon)^2} \cdot \frac{F(s+i\varepsilon)}{\pi}\, ds,$$

Taking imaginary parts, we see that

$$\mathrm{Im}(F(z)) = \omega(z) = \int_{\mathbb{R}} \varphi_\varepsilon(s)\omega_\varepsilon(s)\, ds,$$

where

$$\varphi_\varepsilon(s) = \frac{y-\varepsilon}{(s-x)^2 + (y-\varepsilon)^2} \quad \text{and} \quad \omega_\varepsilon(s) = \frac{\omega(s+i\varepsilon)}{\pi}.$$

Now we can define

$$F_\varepsilon(\lambda) = \int_{-\infty}^{\lambda} \omega_\varepsilon(s)\, ds.$$

Since F is Herglotz, $\omega(x + i\varepsilon)$ is postive and continuous, so then F_ε is increasing and continuous. In particular, it is a distribution, so that $\mu_\varepsilon((-\infty, \lambda]) = F_\varepsilon(\lambda)$ gives a family of Borel measures. Our bound on F shows that

$$y\omega(s) = \int_{\mathbb{R}} \frac{y(y-\varepsilon)}{(s-x)^2 + (y-\varepsilon)^2} \omega_\varepsilon(s)\, ds \leq y|F(z)| \leq M,$$

so as $y \to \infty$ we have that

$$\mu_\varepsilon(\mathbb{R}) = \int_{\mathbb{R}} \omega_\varepsilon(s)\, ds \leq M.$$

Taking $\varepsilon = 1/n$, by Theorem A.1.3 we have that μ_ε converges vaguely to some unique Borel measure μ with $\mu(\mathbb{R}) \leq M$. Moreover, we see that for $A_\lambda = (-\infty, \lambda]$, we have that

$$\int_{\mathbb{R}} \chi_A(s) d\mu_\varepsilon(s) = \int_A d\mu_\varepsilon(s) = \int_A \omega_\varepsilon(s)\, ds,$$

so that $d\mu_\varepsilon = \omega_\varepsilon(s)\, ds$. Since

$$\begin{aligned} |\varphi_\varepsilon(s) - \varphi_0(s)| &= \left| \frac{y-\varepsilon}{(s-x)^2 + (y-\varepsilon)^2} - \frac{y}{(s-x)^2+y^2} \right| \\ &= \left| \frac{-\varepsilon(s-x)^2}{[(s-x)^2 + (y-\varepsilon)^2] \cdot [(s-x)^2 + y^2]} \right| \\ &= \varepsilon \cdot \alpha(s, \varepsilon), \end{aligned}$$

where $\alpha(s, \varepsilon)$ is bounded for all $s \in \mathbb{R}$ for fixed $\varepsilon > 0$. So for small ε we have $\alpha(s, \varepsilon) \leq K$ uniformly for some $K > 0$. Now we see that

$$\begin{aligned} \left| \omega(z) - \int_{\mathbb{R}} \varphi_0(s) d\mu_\varepsilon(s) \right| &= \left| \int_{\mathbb{R}} \varphi_\varepsilon(s)\omega_\varepsilon(s)\, ds - \int_{\mathbb{R}} \varphi_0(s) d\mu_\varepsilon(s) \right| \\ &= \left| \int_{\mathbb{R}} (\varphi_\varepsilon(s) - \varphi_0(s)) d\mu_\varepsilon(s) \right| \\ &\leq \sup_{\lambda \in \mathbb{R}} |\varphi_\varepsilon(s) - \varphi_0(s)| \int_{\mathbb{R}} d\mu_\varepsilon(s) \\ &\leq (K\varepsilon)\mu(\mathbb{R}) \\ &\leq MK\varepsilon. \end{aligned}$$

This shows that

$$\omega(z) = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}} \varphi_0(s) d\mu_\varepsilon(s).$$

But clearly $\varphi_0 \in C_0(\mathbb{R})$, so the last statement in Theorem A.1.3 shows that

$$\int_{\mathbb{R}} \varphi_0\, d\mu_\varepsilon \to \int_{\mathbb{R}} \varphi_0\, d\mu,$$

which implies

$$\omega(z) = \int_{\mathbb{R}} \varphi_0(s) d\mu(s) = \int_{\mathbb{R}} \frac{y^2}{(s-x)^2 + y^2} d\mu(s).$$

But since

$$\frac{1}{s-z} = \frac{s-\bar{z}}{|s-z|^2} = \frac{(s-x)+iy}{(s-x)^2+y^2} = \frac{s-x}{(s-x)^2+y^2} + i\frac{y^2}{(s-x)^2+y^2},$$

which implies

$$\text{Im}(F(z)) = \text{Im}\left(\int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu(\lambda)\right).$$

Since F and the Borel transform are both holomorphic on $\mathbb{C}^+$ with identical imaginary parts, we must have that they differ by a real constant A. But since

$$|A| \leq |F(z)| \leq \frac{M}{\text{Im}(z)} \to 0$$

as $\text{Im}(z) \to \infty$, we must in fact have that

$$F(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu(\lambda),$$

which was to be shown. $\square$

The next theorem shows that the measure μ can be explicitly computed:

Theorem A.2.2. *If F is the Borel transform of μ, then μ is given by the* ***Stieltjes inversion formula****:*

$$\frac{1}{2}(\mu((\lambda_1, \lambda_2)) + \mu([\lambda_1, \lambda_2])) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} Im(F(\lambda + i\varepsilon))\, d\lambda.$$

Proof. We see that

$$\frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \text{Im}(F(\lambda + i\varepsilon))\, d\lambda = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} d\mu(x)\, d\lambda.$$

Since the integrand is clearly continuous for all $(x, \lambda) \in \mathbb{R}^2$, Fubini's theorem implies that

$$\begin{aligned}
\frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} d\mu(x)\, d\lambda &= \frac{1}{\pi} \int_{\mathbb{R}} \int_{\lambda_1}^{\lambda_2} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2}\, d\lambda\, d\mu(x) \\
&= \frac{1}{\pi} \int_{\mathbb{R}} \int_{\lambda_1}^{\lambda_2} \frac{1}{[(\lambda - x)/\varepsilon^{-1}]^2 + 1}\, (\varepsilon^{-1} d\lambda)\, d\mu(x) \\
&= \frac{1}{\pi} \int_{\mathbb{R}} \left[\tan^{-1}\left(\frac{\lambda_2 - x}{\varepsilon}\right) - \tan^{-1}\left(\frac{\lambda_1 - x}{\varepsilon}\right)\right]\, d\mu(x).
\end{aligned}$$

As $\varepsilon \to 0^+$, we have

$$\frac{1}{\pi}\left[\tan^{-1}\left(\frac{\lambda_2 - x}{\varepsilon}\right) - \tan^{-1}\left(\frac{\lambda_1 - x}{\varepsilon}\right)\right] \to \frac{1}{\pi}\begin{cases} \tan^{-1}(\infty) - \tan^{-1}(-\infty) & \lambda_1 < x < \lambda_2 \\ \tan^{-1}(\infty) & x = \lambda_1 \\ -\tan^{-1}(-\infty) & x = \lambda_2 \\ \tan^{-1}(\infty) - \tan^{-1}(\infty) & x < \lambda_1 \\ \tan^{-1}(-\infty) - \tan^{-1}(-\infty) & x > \lambda_2 \end{cases}$$

$$= \begin{cases} 1 & x \in (\lambda_1, \lambda_2) \\ 1/2 & x = \lambda_1, \lambda_2 \\ 0 & x \in \mathbb{R} \setminus (\lambda_1, \lambda_2) \end{cases}$$

$$= \frac{1}{2}\left[\chi_{(\lambda_1,\lambda_2)}(x) + \chi_{[\lambda_1,\lambda_2]}(x)\right].$$

Since $0 \leq \dfrac{1}{\pi}(\tan^{-1}(\cdot) - \tan^{-1}(\cdot)) \leq 1$, the Lebesgue dominated convergence theorem gives

$$\begin{aligned} \lim_{\varepsilon \to 0^+} \frac{1}{\pi}\int_{\lambda_1}^{\lambda_2} \mathrm{Im}(F(\lambda + i\varepsilon))\, d\lambda &= \frac{1}{2}\int_{\mathbb{R}} \left[\chi_{(\lambda_1,\lambda_2)}(x) + \chi_{[\lambda_1,\lambda_2]}(x)\right] d\mu(x) \\ &= \frac{1}{2}(\mu((\lambda_1, \lambda_2)) + \mu([\lambda_1, \lambda_2])), \end{aligned}$$

which was to be shown. □

A.3 Lebesgue Decomposition of Finite Measures

We give here a brief summary of Lebesgue decomposition of Borel measures on the real line, which follows closely the presenation in [7]. First we recall that if μ and ν are measures on a common measure space, then we say that μ and ν are **mutually singular**, denoted $\mu \perp \nu$, if there is a measurable set Ω such that $\mu(\Omega) = 0$ and $\nu(X \setminus \Omega) = 0$. On the other hand, if $\mu(B) = 0$ implies $\nu(B) = 0$ for any measurable set B, then we say that ν is **absolutely continuous** with respect to μ, denoted $\nu \ll \mu$. We start with the following lemma:

Lemma A.3.1. *If μ and ν are σ-finite Borel measures on the real line, then there is a unique* *(modulo almost everywhere w.r.t. μ) non-negative function f and a set Ω such that $\mu(\Omega) = 0$ and*

$$\nu(B) = \nu(B \cap \Omega) + \int_{\mathbb{R}} f \, d\mu$$

for any measurable set B.

Proof. First assume that μ and ν are finite. Take $\alpha = \mu + \nu$. Then

$$\phi(h) = \int_{\mathbb{R}} h \, d\nu$$

is a bounded linear function on $L^2(\mathbb{R}, d\alpha)$ since

$$\begin{aligned} |\phi(h)|^2 = \left| \int_{\mathbb{R}} h \, d\nu \right|^2 &\leq \left(\int_{\mathbb{R}} d\nu \right) \left(\int_{\mathbb{R}} |h|^2 \, d\nu \right) \\ &\leq \nu(\mathbb{R}) \left(\int_{\mathbb{R}} |h|^2 \, d\alpha \right) \\ &= \nu(\mathbb{R}) \|h\|^2 \end{aligned}$$

by Cauchy-Schwarz. By the Riesz representation theorem for bounded linear functionals, we have the existence of a $g \in L^2(\mathbb{R}, d\alpha)$ such that

$$\phi(h) = \langle h, g \rangle = \int_{\mathbb{R}} hg \, d\alpha.$$

Then we have

$$\nu(B) = \int_{\mathbb{R}} \chi_B \, d\nu = \int_{\mathbb{R}} \chi_B g \, d\alpha = \int_B g \, d\alpha.$$

If we take $A = \{x : g(x) \leq 0\}$, then we see that

$$0 \leq \nu(A) = \int_B g \, d\alpha \leq 0,$$

so that $\nu(A) = 0$. This shows that g is non-negative almost everywhere. Now define $\Omega = \{x : g(x) \geq 1\}$. Then

$$\nu(\Omega) = \int_{\Omega} g \, d\alpha = \int_{\mathbb{R}} \chi_{\Omega} g \, d\alpha \geq \int_{\mathbb{R}} \chi_{\Omega} \, d\alpha = \alpha(\Omega) = \nu(\Omega) + \mu(\Omega),$$

so that $0 \leq \mu(\Omega) \leq 0$ or $\mu(\Omega) = 0$. Now we can define

$$f = \frac{g}{1-g} \chi_{\mathbb{R} \setminus \Omega},$$

which is also non-negative almost everywhere. Since $\int_{\mathbb{R}} \chi_B \, d\nu = \int_{\mathbb{R}} \chi_B g \, d\alpha$, we have that $d\nu = g \, d\alpha$ so that $d\mu = d\alpha - d\nu = (1-g) \, d\alpha$. Then

$$\begin{aligned}
\int_B f \, d\mu = \int_{\mathbb{R}} \chi_B f \, d\mu &= \int_{\mathbb{R}} \frac{g}{1-g} \chi_{\mathbb{R}\setminus\Omega} \chi_B \, d\mu \\
&= \int_{\mathbb{R}} \frac{g}{1-g} \chi_{(\mathbb{R}\setminus\Omega)\cap B} \, d\mu \\
&= \int_{\mathbb{R}} \chi_{(\mathbb{R}\setminus\Omega)\cap B} g \, d\alpha \\
&= \int_{\mathbb{R}} \chi_{(\mathbb{R}\setminus\Omega)\cap B} \, d\nu \\
&= \nu((\mathbb{R} \setminus \Omega) \cap B).
\end{aligned}$$

We arrive at

$$\nu(B) = \nu(\Omega \cap B) + \nu((\mathbb{R} \setminus \Omega) \cap B) = \nu(\Omega \cap B) + \int_B f \, d\mu,$$

where $\mu(\Omega) = 0$. If there is a second function f' satisfying this equation, then we would have $\int_B (f - f') \, d\mu = 0$ for any Borel set B, so that $f = f'$ almost everywhere, so f is unique modulo a set of μ measure zero. To extend to the σ-finite case, we first consider the restriction of the Borel algebra on $\mathbb{R}$ to the sets $X_n = (-n-1, n] \cap [n, n+1)$ for $n \geq 0$. Evidently $\cup_{n\geq 0} X_n = \mathbb{R}$. Since μ and ν are Borel algebras, $\mu(X_n), \nu(X_n) < \infty$ for each n. So for each X_n we have Ω_n and f_n satisfying

$$\nu(B) = \nu(\Omega_n \cap B) + \int_B f_n \, d\mu,$$

for all Borel sets $B \subset X_n$. Take $\Omega = \cup_{n\geq 1} \Omega_n$, so that $\mu(\Omega) = 0$ by additivity, and define f by $f(x) = f_n(x)$ for $x \in X_n$, or since the supports are disjoint, $f = \sum_{n=0}^{\infty} f_n$. Now for any Borel set

B on the real line we have

$$\begin{aligned}
\nu(B) = \nu\left(\bigcup_{n\geq 0}(B\cap X_n)\right) &= \sum_{n=0}^{\infty}\nu(B\cap X_n) \\
&= \sum_{n=0}^{\infty}\left[\nu(\Omega_n\cap(B\cap X_n)) + \int_{B\cap X_n} f_n\,d\mu\right] \\
&= \sum_{n=0}^{\infty}\nu(\Omega_n\cap(B\cap X_n)) + \sum_{n=0}^{\infty}\int_{B\cap X_n} f_n\,d\mu \\
&= \nu\left(\bigcup_{n\geq 0}\Omega_n\cap(B\cap X_n)\right) + \int_B f\,d\mu \\
&= \nu\left(\left[\bigcup_{n\geq 0}\Omega_n\right]\bigcap\left[\bigcup_{n\geq 0}(B\cap X_n)\right]\right) + \int_B f\,d\mu \\
&= \nu(\Omega\cap B) + \int_B f\,d\mu,
\end{aligned}$$

so since the uniqueness of the f_n gives f uniquely, this concludes the proof in the σ-finite case. □

Let μ and ν be σ-finite Borel measures and consider the representation given by Lemma A.3.1:

$$\nu(B) = \nu(B\cap\Omega) + \int_{\mathbb{R}} f\,d\mu,$$

where $\mu(\Omega) = 0$ and f is a non-negative μ-measurable function. If $\nu \ll \mu$, then $\nu(B\cap\Omega) \subset \nu(\Omega) = 0$, so that $d\nu = f\,d\mu$. Conversely, if there is a non-negative measurable function f such that $d\nu = f\,d\mu$, then clearly $\mu(B) = 0$ implies that

$$\int_B f\,d\mu = \nu(B) = 0.$$

So $\nu \ll \mu$ if and only if $d\nu = f\,d\mu$ for some non-negative measurable function f. In this case, the function f is written as $\dfrac{d\nu}{d\mu}$ and is called the **Radon-Nikodym derivative** of ν with respect to μ. We can now prove a particular case of the famous Lebesgue decomposition theorem:

Theorem A.3.2. (Lebesgue Decomposition of Borel Measures) *Let μ be a finite Borel measure on the real line, and let m be the Lebesgue measure. Then there are unique measures μ_{ac} and μ_s such that $\mu_{ac} \ll m$, $\mu_s \perp \mu_{ac}$, and $\mu = \mu_{ac} + \mu_s$.*

Proof. Since μ is finite and m is σ-finite, by Lemma A.3.1, there is a non-negative Lebesgue measurable function f and a set M of Lebesgue measure zero such that

$$\mu(B) = \mu(M \cap B) + \int_B f\, dm.$$

Take $\mu_s(\cdot) = \mu(M \cap \cdot)$ and $d\mu_{ac} = f\, dm$. Then since $\mu_{ac}(M) = \int_M f\, dm = 0$ and

$$\mu_s(\mathbb{R} \setminus M) = \mu((\mathbb{R} \setminus M) \cap M) = \mu(\phi) = 0,$$

we have that $\mu_{ac} \perp \mu_s$, and we have already seen that $\mu_{ac} \ll m$. To see that this decomposition is unique, let $\mu = \tilde{\mu}_{ac} + \tilde{\mu}_s$. Then $d\tilde{\mu}_{ac} = \tilde{f}\, dm$ for some non-negative Lebesgue measurable function $\tilde{f}$. So

$$\mu_s(A) - \tilde{\mu}_s(A) = \mu(A) - \mu_{ac}(A) - (\mu(A) - \tilde{\mu}_{ac}(A)) = \int_A (\tilde{f} - f)\, dm.$$

Since $\tilde{\mu}_{ac} \perp \tilde{\mu}_s$, there is a $\tilde{M}$ such that $m(\tilde{M}) = 0$ and $\tilde{\mu}_s(\mathbb{R} \setminus \tilde{M}) = 0$. Then $m(M \cup \tilde{M}) = 0$. Let A be a Lebesgue measurable set. We may assume WLOG that A is either contained in $M \cup \tilde{M}$ or its complement. If $A \subset (M \cup \tilde{M})$, then

$$\int_A (\tilde{f} - f)\, dm = 0.$$

If instead $A \subset [(\mathbb{R} \setminus M) \cap (\mathbb{R} \setminus \tilde{M})] \neq \phi$, we have that

$$\int_A (\tilde{f} - f)\, dm = \mu_s(A) - \tilde{\mu}_s(A) = 0$$

by monotonicity. This shows that $f = \tilde{f}$ almost everywhere, since $\int_A (\tilde{f} - f)\, dm = 0$ for all Lebesgue measurable sets A. Then $\mu_s = \tilde{\mu}_s$, so that

$$\mu_{ac} - \tilde{\mu}_{ac} = \mu - \mu_s - (\mu - \tilde{\mu}_s) = 0,$$

and so $\mu_{ac} = \tilde{\mu}_{ac}$ as well. $\square$

We would now like to characterize the supports of μ_{ac} and μ_s in terms of m. This is done via differentiation of measures. We begin by defining the **symmetric derivative** of a Borel measure μ to be

$$(D\mu)(x) = \lim_{r \to 0} \frac{\mu((x-r, x+r))}{2r} = \lim_{r \to 0} \frac{\mu((x-r, x+r))}{m(x-r, x+r)}$$

where m is again the Lebesgue measure. First we have

Lemma A.3.3. *The set $\{M\mu > \lambda\} = \{x \in \mathbb{R} : (M\mu)(x) > \lambda\}$ are open, where $M\mu$ is the maximal function*

$$(M\mu)(x) = \sup_{0<r<\infty} \frac{\mu((x-r, x+r))}{2r}.$$

In particular, $M\mu$ is Lebesgue measurable.

Proof. It suffices to show that the sets $\{M\mu > \lambda\} = \{x \in \mathbb{R} : (M\mu)(x) > \lambda\}$ are open. Fix $x \in \{M\mu > \lambda\}$ for some $\lambda > 0$. By the definition of $\{M\mu > \lambda\}$, we can find $r > 0$ and $t > \lambda$ such that

$$\mu((x-r, x+r)) = 2tr.$$

Since $t/\lambda > 1$, there exists a $\delta > 0$ such that $r + \delta < \dfrac{rt}{\lambda}$, so for any y with $|x - y| < \delta$ we have $y - \delta < x < y + \delta$ so that

$$(x-r, x+r) \subset (y-r-\delta, y+r+\delta),$$

which gives

$$\mu((y-r-\delta, y+r+\delta)) \geq \mu((x-r, x+r)) = 2tr > \lambda \cdot 2(r+\delta),$$

so that

$$\frac{\mu((y-r-\delta, y+r+\delta))}{2(r+\delta)} > \lambda,$$

which shows that $(x-r, x+r) \subset \{M\mu > \lambda\}$. This for all $x \in \{M\mu > \lambda\}$, so $\{M\mu > \lambda\}$ is open. $\square$

If $I_i = (x_i - r_i, x_i + r_i)$, $1 \leq i \leq n$, is a finite collection of intervals with $r_i \leq r_{i+1}$ for each $1 \leq i < n$, then we can find a subset $\{r_j\}_{j=1}^k \subset \{r_i\}_{i=1}^n$ such that the I_j are all disjoint and each I_i is contained in some interval of the form $(x_j - 3r_j, x_j + 3r_j)$. This can be seen by noting that if $I_i \cap I_j \neq \phi$ with $r_i \leq r_j$, then for $y \in I_i$, we have $|x_i - y| \leq |x_i - x_j| + |x_j - y| < 2r_j + r_j = 3r_j$. It follows that

$$m(\cup_{i=1}^n I_i) \leq \sum_{i=1}^k m(I_j) = 6 \sum_{j=1}^k r_j.$$

We can use this fact to prove:

Lemma A.3.4. *For any Borel measure μ on the real line and $\lambda > 0$, we have*

$$m(\{M\mu > \lambda\}) \leq \frac{3\mu(\mathbb{R})}{\lambda}.$$

Proof. Let K be any compact subset of $\{M\mu > \lambda\}$. Then for each $x \in K$, the definition of $M\mu$ gives an r_x such that $\mu((x - r_x, x + r_x)) > 2r_x\lambda$. Some finite collection of these open intervals $\{I_i = (x_i - r_i, x_i + r_i)\}_{i=1}^n$, must cover K since it is compact. The remarks above show that there are $\{r_j\}_{j=1}^k$ for some $k \leq n$ such that

$$m(K) \leq 3\sum_{j=1}^{k}(2r_j) < \frac{3}{\lambda}\sum_{j=1}^{k}\mu(I_i).$$

But m is a regular measure and $\{M\mu > \lambda\}$ is open, so taking the supremum over all compact subsets gives

$$m(\{M\mu > \lambda\}) \leq \frac{3}{\lambda}\mu(\mathbb{R}),$$

which concludes the proof. □

In particular this shows that if f is a Lebesgue measurable function and $d\mu = f\,dm$, then for the maximal operator

$$(Mf)(x) = \sup_{0<r<\infty} \frac{1}{2r}\int_{(x-r,x+r)} |f|\,dm$$

we get

$$m(\{M(|f|\,dm) > \lambda\}) \leq \frac{3}{\lambda}\int_{\mathbb{R}} |f|\,dm.$$

This leads us to define a **Lebesgue point** of a Lebesgue measurable function f to be any $x \in \mathbb{R}$ such that

$$\lim_{r\to 0}\frac{1}{2r}\int_{(x-r,x+r)} |f(y) - f(x)|\,dm(y) = 0.$$

Note that if f is a continuous function, then every point is a Lebesgue point. The following theorem establishes existence in the general case:

Theorem A.3.5. *If f is Lebesgue measurable, then almost every $x \in \mathbb{R}$ is a Lebesgue point of f.*

Proof. It suffices to show that

$$(Tf)(x) = \sup_{0<r<\infty} \frac{1}{2r} \int_{(x-r,x+r)} |f(y) - f(x)|\, dm(y) = 0,$$

for almost every x. For any $\varepsilon > 0$ we can find a $g \in C(\mathbb{R})$ such that $\int_{\mathbb{R}} |f - g|\, dm < \varepsilon$. Since g is continuous, taking $h = f - g$ shows that for all $r > 0$ we have

$$\begin{aligned}
\frac{1}{2r} \int_{(x-r,x+r)} |f - f(x)|\, dm &\le \frac{1}{2r} \int_{(x-r,x+r)} |f - g - (f(x) - g)|\, dm \\
&\le \frac{1}{2r} \int_{(x-r,x+r)} |h|\, dm + \frac{1}{2r} \int_{(x-r,x+r)} |f(x) - g|\, dm \\
&= \frac{1}{2r} \int_{(x-r,x+r)} |h|\, dm + |h(x)|.
\end{aligned}$$

Then $(Tf)(x) \le (Mh)(x) + |h(x)|$, so if x is such that $(Mf)(x) > 2y$ for some $y > 0$, then either $(Mf)(h) > y$ or $|h| > y$, so we have

$$\{Tf > 2y\} \subset \{Mh > y\} \cup \{|h| > y\}.$$

But $m(\{h > y\}) \le \dfrac{1}{y} \int_{\mathbb{R}} |h|\, dm$ by Markov's inquality, so Lemma A.3.4 gives

$$m(\{Tf > 2y\}) \le m(\{Mh > y\}) + m(\{|h| > y\} \le \frac{4}{\lambda} \int_{\mathbb{R}} |h|\, dm = \frac{4\varepsilon}{y}.$$

Since $\varepsilon > 0$ was arbitrary, we must have

$$m(\{Tf > 2y\}) = 0.$$

This for all $y > 0$, so that $(Tf)(x) = 0$ for almost every $x \in \mathbb{R}$. $\square$

Now we are ready to prove the main result:

Theorem A.3.6. *Let μ be a Borel measure on $\mathbb{R}$. We have*

i) $\mu \perp m$ implies that $(D\mu)(x) = \infty$ everywhere but a set of μ measure 0 (i.e., μ is supported on $T = \{x \in \mathbb{R} : (D\mu)(x) = \infty\}$, so that $\mu(\mathbb{R} \setminus T) = 0$);

ii) if $d\mu = f\, dm$, then $\dfrac{d\mu}{dm} = D\mu$ (i.e., if $\mu \ll m$, then the Radon-Nikodym derivative of μ is $D\mu$).

Proof. (i) If $\mu \perp m$, then there is a Borel set B such that $m(B) = 0$ and $\mu(\mathbb{R} \setminus B) = 0$. We must show that the set of all x such that $(D\mu)(x) = 0$ has μ measure zero. Since $\mu(\mathbb{R} \setminus B)$ has μ measure zero, we may restrict our attention to B. Since $m(S) = 0$, we can find V_j for $j \geq 1$ such that $B \subset V_j$ for all j and $m(V_j) \leq 1/j$. If $(D\mu)(x) < \infty$, then we have that

$$\lim_{r \to 0} \frac{\mu((x-r, x+r))}{2r} < N$$

for some $N \geq 1$. So we define E_N to be the set of all $x \in B$ such that there exists a sequence $r_x = r_x^{(i)}$ such that $r_x^{(i)} \to 0$ and

$$\mu((x - r_x^{(i)}, x + r_x^{(i)})) < 2Nr_x^{(i)}.$$

Evidently, x lies in some E_N if and only if $(D\mu)(x) < \infty$, so we must show that $\cup_{N \geq 1} E_N$ has μ measure zero. Fix N and j. For every x in E_N there is an r_x small enough such that $I_x = (x - r_x, x + r_x) \subset V_j$ and satisfies $\mu(I_x) < 2Nr_x$. Set $O_{j,N} = \cup_{x \in E_N} I_x$. Then $E_N \subset O_{j,N} \subset V_j$, and for any compact set $K \subset O_{j,N}$, there is a finite subset F of E_N such that I_x cover K for $x \in F$. Then by the remarks before Lemma A.3.4 we have the existence of x_k in F such that $\mathcal{I}_{x_k} = (x - 3r_{x_k}, x + 3r_{x_k})$ are disjoint and cover K. Then

$$\mu(K) \leq \sum_k \mu(\mathcal{I}_{x_k}) < N \sum_k m(\mathcal{I}_{x_k}) \leq 3N \sum_k m(I_{x_k}) = 3Nm(\cup_k I_{x_k}) \leq 3Nm(V_j) \leq \frac{3N}{j}.$$

Taking

$$\Omega_N = \bigcap_{j=1}^{\infty} O_{j,N},$$

we see that $E_N \subset \Omega_N$ and since $\mu(\Omega_N) \leq \dfrac{3N}{j}$ for all j, we have that $\mu(\Omega_N) = 0$ for all N. This shows that each E_N has μ measure zero, so that $\mu(\cup_{N \geq 1} E_N) = 0$ as well.

(ii) If x is a Lebesgue point of f, then since f is non-negative we have that

$$f(x) = \lim_{r \to 0} \frac{1}{2r} \int_{(x-r,x+r)} f\, dm = \lim_{r \to 0} \frac{\mu((x-r, x+r)}{2r} = (D\mu)(x),$$

since $d\mu = f\, dm$. □

Appendix B

Functional Analysis

B.1 Spectral Measures

This section and the next give a concise but detailed presenation of the spectral theorem for self-adjoint operators, and follows closely the presentation in [11]. Take $\mathfrak{B}$ to be the Borel sets on the real line and $\mathfrak{P}(H)$ to be the space of projections on a separable complex Hilbert space H. The inner product $\langle \cdot, \cdot \rangle$ of H is taken to be conjugate linear in the second argument.

Definition B.1.1. *A map $P : \mathfrak{B} \to \mathfrak{P}(H)$ is a* **spectral measure** *if*

i) $P(\mathbb{R}) = I$

ii) $P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n)$

for any disjoint collection of Borel sets $\{B_n\}$. The sum is understood to converge in the strong operator topology.

A spectral measure is not a genuine measure, as it does not take values in $[0, \infty]$, but it does satisfy many of the same properties. Moreover, it is possible to use the inner product on H to define Borel measures for each f in H, as follows:

Theorem B.1.1. *For each $f \in H$, the function*

$$\mu_f(B) = \langle P(B)f, f\rangle = \|P(B)f\|^2$$

defines a finite Borel measure on the real line.

Proof. Fix an $f \in H$. We wish to see that $\mu_f(\phi) = 0$. Since $\mu_f(\phi) = \|P(\phi)f\|^2$, it is sufficient to prove that $P(\phi)$ is the zero operator. Since $\mathbb{R} \cup \phi$ is disjoint, we have from the properties of the spectral measure that

$$I = P(\mathbb{R}) = P(\mathbb{R} \cup \phi) = P(\mathbb{R}) + P(\phi) = I + P(\phi).$$

These operators can only be equal if $P(\phi)f = 0$ for all $f \in H$. To see the additivity of μ_f, let $\varepsilon > 0$. It is clear that $\mu_0 \equiv 0$, so we may take $f \neq 0$. Then by the definition of a spectral measure, there exists an $N \geq 1$ such that

$$\left\| P\left(\bigcup_{n=1}^{\infty} B_n\right) f - \sum_{n=1}^{N} P(B_n)f \right\| < \frac{\varepsilon}{\|f\|}$$

for all . Then the continuity and linearity of the inner product on H implies that

$$\begin{aligned} \left| \mu_f\left(\bigcup_{n=1}^{\infty} B_n\right) - \sum_{n=1}^{N} \mu_f(B_n) \right| &= \left| \left\langle P\left(\bigcup_{n=1}^{\infty} B_n\right) f, f \right\rangle - \sum_{n=1}^{N} \langle P(B_n)f, f\rangle \right| \\ &= \left| \left\langle P\left(\bigcup_{n=1}^{\infty} B_n\right) f - \sum_{n=1}^{N} P(B_n)f, f \right\rangle \right|, \end{aligned}$$

so by Cauchy-Schwarz we have

$$\left| \mu_f\left(\bigcup_{n=1}^{\infty} B_n\right) - \sum_{n=1}^{N} \mu_f(B_n) \right| \leq \left\| P\left(\bigcup_{n=1}^{\infty} B_n\right) f - \sum_{n=1}^{N} P(B_n)f \right\| \cdot \|f\| < \varepsilon.$$

Since each $P(\mathbb{R})f = If = f$, we have that $\mu_f(\mathbb{R}) = \langle P(\mathbb{R})f, f\rangle = \|P(\mathbb{R})f\|^2 = \|f\|^2$, and μ_f is finite. ◻

Since we have

$$\|P(B)f + g\|^2 = \mu_f(B) + \|g\|^2 + 2\,\mathrm{Re}(\langle P(B)f, g\rangle)$$

and

$$\|P(B)f + ig\|^2 = \mu_f(B) + \|g\|^2 + 2\,\mathrm{Im}(\langle P(B)f, g\rangle),$$

we can use the polarization identity to define the complex measures

$$\mu_{fg}(B) = \langle P(B)f, g\rangle = \frac{1}{4}(\mu_{f+g}(B) - \mu_{f-g}(B) - i\mu_{f-ig}(B) + i\mu_{f+ig}(B))$$

for each pair of f and g in H. We see that for any Borel set B

$$\overline{\mu_{fg}(B)} = \overline{\langle P(B)f, g\rangle} = \langle g, P(B)f\rangle = \langle P(B)g, f\rangle = \mu_{gf}(B)$$

so that $\mu_{gf} = \overline{\mu_{fg}}$. Let Simp($\mathbb{R}$) denote the set of complex-valued simple functions, regarded as a subspace of the space of complex-valued bounded Borel functions equipped with the sup norm, and let $B(H)$ denote the space of bounded linear operators on H. Denote the indicator function of $A \subset \mathbb{R}$ by χ_A. For any $\varphi = \sum_{n=1}^{N} \alpha_n \chi_{B_n} \in$ Simp$^+$($\mathbb{R}$), we define the operator $P_* : \mathrm{Simp}(\mathbb{R}) \to B(H)$ by the formula

$$P_*(\varphi) = \sum_{n=1}^{N} \alpha_n P(B_n) = \int_{\mathbb{R}} \varphi(\lambda) dP(\lambda).$$

This gives

$$\langle P_*(\varphi)f, g\rangle = \sum_{n=1}^{N} \alpha_n \langle P(B_n)f, g\rangle = \sum_{n=1}^{N} \alpha_n \mu_{fg}(B_n) = \int_{\mathbb{R}} \varphi \, d\mu_{fg},$$

and in particular

$$\|P_*(\varphi)f\|^2 = \sum_{n=1}^{N} |\alpha_n|^2 \langle P(B)f, f\rangle = \int_{\mathbb{R}} |\varphi|^2 d\mu_f.$$

By way of the Hahn-Banach theorem, we have the following:

Theorem B.1.2. (Integral Representation for Borel Functions) *If P is a spectral measure on a Hilbert space H, then every bounded Borel function $\psi : \mathbb{R} \to \mathbb{C}$ has the representation*

$$P_*(\psi) = \int_{\mathbb{R}} \psi(\lambda) dP(\lambda).$$

Moreover, we have that P_* satisfies the following:

i) P_* has norm one;

ii) the adjoint of $P_*(\psi)$ is $P_*(\overline{\psi})$ for all ψ;

iii) $P_*(\psi_1\psi_2) = P_*(\psi_1)P_*(\psi_2)$ for all ψ_1 and ψ_2;

iv) $P_*(\psi)$ is normal for all ψ.

Proof. Clearly P_* is linear on $\mathrm{Simp}^+(\mathbb{R})$, so since the set of simple functions is dense in the space of complex-valued bounded Borel functions, these properties will follow immediately from the Hahn-Banach theorem, once they have been shown to hold on $\mathrm{Simp}^+(\mathbb{R})$. So we let $\varphi = \sum_{n=1}^N \alpha_n \chi_{B_n} \in \mathrm{Simp}^+(\mathbb{R})$ and proceed as follows:

(i) Let $\|P_*(\varphi)\|_{op}$ denote the operator norm of $P_*(\varphi)$. We wish to see that $\sup_{|\varphi|_\infty=1} \|P_*(\varphi)\|_{op} = 1$. If $|\varphi|_\infty = 1$, then we must have

$$\|P_*(\varphi)f\|^2 = \int_{\mathbb{R}} |\varphi|^2 d\mu_f \leq |\varphi|_\infty \int_{\mathbb{R}} d\mu_f = \mu_f(\mathbb{R}) = \|f\|^2.$$

So if $\|f\| = 1$, then $\|P_*(\varphi)\|_{op} \leq 1$, and so P_* has norm no greater than 1. To see that equality holds, we take $\varphi(x) \equiv 1$, so that $\|P_*(\varphi)f\| = 1$ for $\|f\| = 1$.

(ii) We can compute the adjoint of $P_*(\varphi)$ from the properties of the spectral measure:

$$\langle f, P_*(\overline{\varphi})g\rangle = \overline{\int_{\mathbb{R}} \overline{\varphi}\mu_{gf}} = \int_{\mathbb{R}} \varphi\overline{\mu_{gf}} = \int_{\mathbb{R}} \varphi\mu_{fg} = \langle P_*(\varphi)f, g\rangle.$$

(iii) If B_1 and B_2 are disjoint, then we have $P(B_1 \cup B_2) = P(B_1) + P(B_2)$. Squaring this and using the fact that $P(\cdot)$ is a projection, we have

$$\begin{aligned} P(B_1) + P(B_2) &= P(B_1 \cup B_2) = P(B_1 \cup B_2)^2 \\ &= P(B_1)^2 + P(B_1)P(B_2) + P(B_2)P(B_1) + P(B_2)^2 \\ &= P(B_1) + P(B_1)P(B_2) + P(B_2)P(B_1) + P(B_2), \end{aligned}$$

or $P(B_1)P(B_2) + P(B_2)P(B_1) = 0$. Multiplying on the left by $P(B_1)$ gives $P(B_1)P(B_2) + P(B_1)P(B_2)P(B_1) = 0$. Multiplying on the right by $P(B_1)$ then gives $2P(B_1)P(B_2)P(B_1) = 0$. Substituting into the previous equation now gives $P(B_1)P(B_2) = 0$ for any pair of disjoint

Borel sets. Now note that for any two simple functions φ_1 and φ_2, we can refine their supports using complements and intersections to write

$$\varphi_1(x) = \sum_{i=1}^{N} \alpha_i \chi_{B_i} \quad \text{and} \quad \varphi_2(x) = \sum_{i=1}^{N} \beta_i \chi_{B_i},$$

where the α_i and β_i may not be distinct and non-zero, but the B_i are all disjoint. This is also a simple function since the B_i are also Borel sets. But since $\chi_{B_i}\chi_{B_j} = 0$ for $B_i \cap B_j = \phi$, this shows that

$$\varphi_1(x)\varphi_2(x) = \left(\sum_{i=1}^{N} \alpha_i \chi_{B_i}\right)\left(\sum_{i=1}^{N} \beta_i \chi_{B_i}\right) = \sum_{i=1}^{N} \alpha_i \beta_i \chi_{B_i},$$

so that

$$P_*(\varphi_1(x)\varphi_2(x)) = \sum_{i=1}^{N} \alpha_i \beta_i P(B_i).$$

But since $P(B_i)P(B_j) = 0$ for disjoint sets, this shows that

$$P_*(\varphi_1(x))P_*(\varphi_2(x)) = \sum_{i=1}^{N} \alpha_i \beta_i P(B_i)^2 = \sum_{i=1}^{N} \alpha_i \beta_i P(B_i),$$

since each $P(B_i)$ is a projection. Thus $P_*(\varphi_1\varphi_2) = P_*(\varphi_1)P_*(\varphi_2)$ for simple functions, as desired.

(iv) From (ii) and (iii), we have that $P_*(\varphi)[P_*(\varphi)]^* = P_*(\varphi)P_*(\overline{\varphi}) = P_*(|\varphi|^2) = [P_*(\varphi)]^* P_*(\varphi)$ for all $f \in H$. It follows that $P_*(\varphi)$ is normal. □

For any $f \in H$, define the subspace H_f of H to be the images of f under $P_*(\psi)$ for each $\psi \in L^2(\mathbb{R}, d\mu_f) = L^2(d\mu_f)$. In other words,

$$H_f = \{P_*(\psi)f : \int_{\mathbb{R}} |\psi|^2 d\mu_f < \infty\}.$$

This is clearly a subspace since P_* is linear and $\alpha\psi_1 + \psi_2 \in L^2(d\mu_f)$ if $\psi_1, \psi_2 \in L^2(d\mu_f)$. The space $L^2(d\mu_f)$ for the Borel measure μ_f is closed for any f, so if $|\psi_n - \psi|_\infty \to 0$, we have that

$$\|P_*(\psi_n)f - P_*(\psi)f\|^2 = \|P_*(\psi_n - \phi)f\|^2 = \int_{\mathbb{R}} |\psi_n - \psi|^2 d\mu_f \le |\psi_n - \psi|_\infty \|f\|^2 \to 0.$$

This shows that H_f is closed, so we can consider the projection P_f of H onto H_f.

Lemma B.1.3. *For each $f \in H$ and $\psi \in L^2(d\mu_f)$, H_f reduces $P_*(\psi)$; i.e., $P_f P_*(\psi) = P_*(\psi)P_f$.*

Proof. Any $h \in H$ has the form $h = P_*(\psi_h)f + g$ for some $\psi_h \in L^2(\mathbb{R}, \mu_f)$ and $g \in H_f^\perp$. We see that

$$P_*(\psi)P_f h = P_*(\psi)P_f(P_*(\psi_h)f + g) = P_*(\psi)P_*(\psi_h)f.$$

Since $g \in H_f^\perp$, we have that $\langle P_*(\psi_0)f, P_*(\psi)g\rangle = \langle P_*(\psi_0\overline{\psi})f, g\rangle = 0$, so that $P_*(\psi)g \in H_f^\perp$ as well. Then

$$P_f P_*(\psi)h = P_f P_*(\psi)(P_*(\psi_h)f + g) = P_f(P_*(\psi)P_*(\psi_h)f + P_*(\psi)g) = P_*(\psi_h)P_*(\psi)f,$$

so since $P_*(\psi)$ and $P_*(\psi_h)$ commute, we have that $P_f P_*(\psi) = P_*(\psi)P_f$. □

From this lemma, we can prove the following theorem:

Theorem B.1.4. *If $H_f = H$, then there exists a unitary map from H to $L^2(\mathbb{R}, d\mu_f)$, and $P_*(\psi)$ is unitarily equivalent with the multiplication operator Ψ defined by $\Psi(\varphi) = \psi\varphi$ for $\varphi \in L^2(d\mu_f)$.*

Proof. Define $U_f : H_f \to L^2(\mathbb{R}, d\mu_f)$ by

$$U_f(P_*(\psi)f) = \psi.$$

Clearly this map is onto by the definition of H_f, and it is well-defined by Lemma B.1.3. We have by the properties of P_* that

$$\begin{aligned}
\langle U_f(P_*(\psi_1)f), U_f(P_*(\psi_2)f)\rangle_{L^2} &= \langle \psi_1, \psi_2\rangle_{L^2} \\
&= \int_{\mathbb{R}} \psi_1\overline{\psi_2}\, d\mu_f \\
&= \langle P_*(\overline{\psi_2}\psi_1)f, f\rangle_H \\
&= \langle P_*(\overline{\psi_2})P_*(\psi_1)f, f\rangle_H \\
&= \langle P_*(\psi_1)f, P_*(\psi_2)f\rangle_H.
\end{aligned}$$

We have that if $g = P_*(\psi')f \in H_f$, then

$$(U_f P_*(\psi))g = U_f(P_*(\psi)P_*(\psi')f) = U_f(P_*(\psi\psi')f) = \psi\psi' = (\Psi U_f)(g),$$

so that $U_f P_*(\psi) = \Psi U_f$. So Ψ and $P_*(\psi)$ are unitarily equivalent for each ψ. □

We just saw that if $H_f = H$, then U_f is a unitary map from H onto $L^2(d\mu_f)$. If $H_f \neq H$, then since H_f reduces $P_*(\psi)$, we can write

$$P_*(\psi) = P_*(\psi)|_{H_f} \oplus P_*(\psi)|_{H_f^\perp}$$

for any $\psi \in L^2(\mathbb{R}, d\mu_f)$. We call a sequence $\{\delta_n\}_{n=1}^\infty \subset H$ a **spectral basis** for H if $\|\delta_n\| = 1$ for all n and $H_{\delta_i} \perp H_{\delta_j}$ for $i \neq j$. If a spectral basis exists, we can write

$$H = \bigoplus_{n=1}^{\infty} H_{\delta_n},$$

and we have that $U = \oplus_{n=1}^\infty U_{\delta_n}$ is a unitary map from H to $\oplus_{n=1}^\infty L^2(\mathbb{R}, d\mu_{\delta_n})$, and $UP_*(\psi) = \Psi U$ for all ψ. Since H is separable, an at most countable spectral basis can always be constructed by applying the Gram-Schmidt procedure to a total set in H. The detailed proof is omitted, as this result is not used in the main paper, but it is mentioned for the sake of completeness.

B.2 The Spectral Theorem for Self-Adjoint Operators

For every spectral measure P on H, we can assign the operator $A = P_*(x) = \displaystyle\int_{\mathbb{R}} \lambda\, dP(\lambda)$. Since the identity function $\psi(x) = x$ is real-valued, we have $[P_*(x)]^* = P_*(\overline{x}) = P_*(x)$, so A is self-adjoint. Recall that the **resolvent** of A is the map $R_A(z) : \rho(A) \to B(H)$ defined by $z \mapsto (A - zI)^{-1}$. For a fixed $z \in \mathbb{C}$, we denote the simple function $z\chi_{\mathbb{R}}$ by just z. Then we have

$$P_*(z) = \int_{\mathbb{R}} z\, dP(\lambda) = zP(\mathbb{R}) = zI.$$

In particular, $P_*(1) = I$. But since $P_*(\psi_1\psi_2) = P_*(\psi_1)P_*(\psi_2)$, we must have

$$\begin{aligned}(A - zI)P_*\left(\frac{1}{x-z}\right) &= (P_*(x) - P_*(z))P_*\left(\frac{1}{x-z}\right)\\ &= (P_*(x-z))P_*\left(\frac{1}{x-z}\right)\\ &= P_*(1) = I.\end{aligned}$$

This shows that $R_A(z) = \displaystyle\int_{\mathbb{R}} \frac{1}{\lambda - z} dP(\lambda)$. We then have that

$$F_f(z) = \langle R_A(z)f, f\rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_f(\lambda)$$

is a Borel transform. Since

$$\operatorname{Im}(F_f(z)) = \operatorname{Im} \int_{\mathbb{R}} \frac{\lambda - \overline{z}}{|\lambda - z|^2} d\mu_f(\lambda) = \operatorname{Im}(z) \int_{\mathbb{R}} \frac{1}{|\lambda - z|^2} d\mu_f(\lambda),$$

we see that F_f is a Herglotz function on the upper half plane, and so from Theorem A.1.1, the measure μ_f is can be recovered via the Stieltjes inversion formula:

$$\mu_f(\lambda) = \lim_{\delta \to 0^+} \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{-\infty}^{\lambda+\delta} \operatorname{Im}(F_f(t + i\varepsilon))dt.$$

We can now use the results from Appendix A and the above to prove the following theorem:

Theorem B.2.1. *(***Spectral Theorem for Self-Adjoint Operators)** *If A is a self-adjoint operator on a Hilbert space H, then there exists a unique spectral measure P_A such that*

$$A = \int_{\mathbb{R}} \lambda dP_A(\lambda).$$

Conversely, for every spectral measure P, the operator A defined by the above formula is a self-adjoint operator.

Proof. The reverse direction has already been shown in the remarks before the statement of the theorem, so we focus here on the forward direction. Let A be any self-adjoint operator on H, and define the function $F_f(z) = \langle R_A(z)f, f\rangle$, which is clearly holomorphic on $\rho(A)$. It is important to note that since A is self-adjoint, we have that $\rho(A) \subset \mathbb{R}$ and that $(R_A(z))^* = R_A(\overline{z})$. From this it is clear that F_f satisfies $\overline{F_f(z)} = F_f(\overline{z})$, and

$$|F_f(z)| \leq \|R_A(z)f\|\|f\| \leq \|R_A(z)\|\|f\|^2 \leq \frac{\|f\|^2}{\operatorname{Im}(z)},$$

by Cauchy-Schwarz and the bound $\|R_A(z)\| \leq \operatorname{Im}(z)^{-1}$ for the resolvent. Since $R_A(z) - R_A(z') = (z - z')R_A(z')R_A(z)$, we have that

$$2i \operatorname{Im}(F_f(z)) = \langle R_A(z)f, f\rangle - \langle R_A(\overline{z})f, f\rangle = 2i \operatorname{Im}(z)\|R_A(z)f\|^2,$$

so F_f is a Herglotz function for any f. Then from Theorem A.1.1 there is a unique Borel measure μ_f given by the Stieltjes inversion formula such that F_f is the Borel transform for μ_f:

$$F_f(z) = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_f(\lambda).$$

By the polarization identity, we recover a complex measure μ_{fg} for each pair of f and g in H such that $\langle R_A(z)f, g\rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{fg}(\lambda)$. For a fixed Borel set B, we define

$$s_B(f,g) = \int_{\mathbb{R}} \chi_B d\mu_{fg} = \mu_{fg}(B).$$

It is clear from the definition of μ_{fg} that s_B is linear in the first variable and conjugate linear in the second. We also see that

$$s_B(f,f) = \int_{\mathbb{R}} \chi_B d\mu_f = \mu_f(B) \geq 0,$$

so s_B is a positive-definite sesquilinear form. We now have by Cauchy-Schwarz that

$$|s_B(f,g)|^2 \leq s_B(f,f)s_B(g,g) = \mu_f(B)\mu_g(B) \leq \mu_f(\mathbb{R})\mu_g(\mathbb{R}) = \|f\|^2\|g\|^2.$$

So by the Riesz Representation Theorem, we have the existence of a unique operator $P_A(B)$ satisfying $s_B(f,g) = \langle P_A(B)f, g\rangle$ with $\|P_A(B)\| = \|s_B\|$. Since $|s_B(f,g)| \leq \|f\| \cdot \|g\|$, we have that $\|s_B\| \leq 1$ for all B. So we have constructed a unique family of operators $\{P_A(B) : B \text{ a Borel set}\}$ such that $0 \leq \langle P_A(B)f, g\rangle \leq 1$ and

$$\langle P_A(B)f, g\rangle = \int_{\mathbb{R}} \chi_B \mu_{fg}.$$

We wish to show that the map $P : P_A \to A$ is in fact a spectral measure. First we fix a Borel set B and show that $P_A(B)$ is a projection. We know that $P_A(B)$ is self-adjoint by construction, so it is sufficient to show that $[P_A(B)]^2 = P_A(B)$. If we can prove $P_A(B_1 \cap B_2) = P_A(B_1)P_A(B_2)$, this will follow from taking $B_1 = B_2$. We first use the formula $R_A(z) - R_A(z') = (z - z')R_A(z')R_A(z)$ to compute:

$$\begin{aligned}
\int_{\mathbb{R}} \frac{1}{\lambda - z'} d\mu_{f,R_A(\overline{z})g}(\lambda) &= \langle R_A(z')f, R_A(\overline{z})g\rangle \\
&= \langle R_A(z)R_A(z')f, g\rangle \\
&= \frac{1}{z' - z}\langle [R_A(z') - R_A(z)]f, g\rangle \\
&= \frac{1}{z' - z}\int_{\mathbb{R}} \left(\frac{1}{\lambda - z'} - \frac{1}{\lambda - z}\right) d\mu_{fg}(\lambda) \\
&= \int_{\mathbb{R}} \frac{1}{\lambda - z'} \frac{d\mu_{fg}(\lambda)}{\lambda - z},
\end{aligned}$$

so that $d\mu_{f,R_A(\overline{z})g}(\lambda) = (\lambda - z)^{-1}d\mu_{fg}(\lambda)$ by the uniqueness of Borel measures. We can now compute

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{P_A(B)f,g}(\lambda) &= \langle R_A(z)P_A(B)f, g\rangle \\ &= \langle P_A(B)f, R_A(\overline{z})g\rangle \\ &= \int_{\mathbb{R}} \chi_B(\lambda) d\mu_{f,R_A(\overline{z})g}(\lambda) \\ &= \int_{\mathbb{R}} \frac{1}{\lambda - z}\chi_B(\lambda) d\mu_{fg}(\lambda),\end{aligned}$$

from which it follows that $d\mu_{P_A(B)f,g}(\lambda) = \chi_B(\lambda)d\mu_{fg}(\lambda)$. As $\chi_{B_1}\chi_{B_2} = \chi_{B_1\cap B_2}$, we arrive at

$$\begin{aligned}\langle P_A(B_1 \cap B_2)f, g\rangle &= \int_{\mathbb{R}} \chi_{B_1\cap B_2}(\lambda) d\mu_{fg}(\lambda) \\ &= \int_{\mathbb{R}} \chi_{B_1}\chi_{B_2}(\lambda) d\mu_{fg}(\lambda) \\ &= \int_{\mathbb{R}} \chi_{B_1} d\mu_{P_A(B)f,g}(\lambda) \\ &= \langle P_A(B_1)P_A(B_2)f, g\rangle.\end{aligned}$$

This holds for all $f, g \in H$, so it follows that $P_A(B_1 \cap B_2) = P_A(B_1)P_A(B_2)$. In order to compute $P_A(\mathbb{R})$, suppose that f is in the kernel of $P_A(\mathbb{R})$. Then $\langle P_A(\mathbb{R})f, f\rangle = \mu_f(\mathbb{R}) = 0$, which implies that μ_f is the zero measure. But then $\langle R_A(z)f, f\rangle = \displaystyle\int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_f = 0$ as well. But this can only hold if $f = 0$, so since $P_A(\mathbb{R})$ is a projection, $\mathrm{Ker}(P_A(\mathbb{R})) = \{0\}$ implies that $P_A(\mathbb{R}) = I$. Finally, to prove additivity we let $B = \displaystyle\bigcup_{n=1}^{\infty} B_n$, where $\{B_n\}_{n=1}^{\infty}$ is a collection of pairwise disjoint Borel sets. But then the additivity of μ_f shows that

$$\begin{aligned}\left\| P_A(B)f - \sum_{n=1}^{N} P_A(B_n)f \right\|^2 &= \left\langle \left(P_A(B) - \sum_{n=1}^{N} P_A(B_n) \right) f, f \right\rangle \\ &= \langle P_A(B)f, f\rangle - \left\langle \sum_{n=1}^{N} P_A(B_n)f, f \right\rangle \\ &= \mu_f(B) - \sum_{n=1}^{N} \mu_f(B_n) \to 0\end{aligned}$$

as $N \to \infty$. This establishes the existence of a spectral measure for A, and uniqueness follows from the fact that the Borel measures μ_{fg} were uniquely determined by the Stieltjes inversion formula and the projections of the spectral measure were uniquely determined by the Riesz represenation theorem. □

This theorem allows us to characterize any self-adjoint operator on H as an integral over a spectral measure that is unitarily equivalent to a class of multiplication operators. Moreover, we see that $(P_A)_*(x) = A$, so Theorem B.1.4 shows that if e is a cyclic vector for A, then for all $f \in H$ we have

$$f = \sum_k \alpha_k [(P_A)_*(x)]^k e = (P_A)_* \left(\sum_k \alpha_k x^k \right) e.$$

This shows that $f \in H_e$, since $\sum_k \alpha_k x^k$ is a polynomial and so is in $L^2(d\mu_e)$. This gives a unitary map $U : H \to L^2(d\mu_e)$, such that $UAU^{-1}\psi(x) = x\psi(x)$. Otherwise, we may take an orthonormal basis for H and construct a spectral basis by a Gram-Schimdt procedure to produce a unitary map onto a direct sum of L^2 spaces such that $UAU^{-1}\psi(x) = x\psi(x)$. We summarize this as:

Lemma B.2.2. *If e is a cyclic vector for A, then $H_e = H$.*

B.3 The Spectrum of a Self-Adjoint Operator

Here we will demonstrate a characterization of the spectrum $\sigma(A)$ of a self-adjoint operator A in terms of its spectral measure P_A, which is also presented in [11]. We also present a theorem relating these results to the material in the previous appendix (a brief summary can also be found in [6]). We will need the following lemma:

Lemma B.3.1. *Let $z \in \mathbb{C}$. If $\{f_n\}_{n=1}^{\infty}$ is a sequence in H such that $\|f_n\| = 1$ for all n and $\|(A - z)f_n\| \to 0$, then $z \in \sigma(A)$.*

Proof. Suppose $\|f_n\| = 1$ for all n and $(A - z)f_n \to 0$. If $z \in \rho(A)$, then $R_A(z)$ exists and is bounded, so we have

$$\|f_n\| = \|R_A(z)(A - z)^{-1} f_n\| \leq \|R_A(z)\| \cdot \|(A - z)f_n\| \to 0,$$

a contradiction. □

These sequences are called **Weyl sequences**. Using them we may prove:

Theorem B.3.2. *For a self-adjoint operator A with spectral measure P_A, we have*

$$\sigma(A) = \{\lambda \in \mathbb{R} : P_A((\lambda - \varepsilon, \lambda + \varepsilon)) \neq 0, \forall \varepsilon > 0\}.$$

In particular, we have:

i) $P_A((\lambda_1, \lambda_2)) = 0$ *if and only if* $(\lambda_1, \lambda_2) \subset \rho(A)$*;*

ii) $P_A(\sigma(A)) = I$*;*

iii) $P_A(\mathbb{R} \cap \rho(A)) = 0$*.*

Proof. First note that since $\lambda \in \mathbb{R}$, $A - \lambda$ is also self-adjoint. Moreover, $A - \lambda = (P_A)_*(t - \lambda)$. Now suppose that $B_n = (\lambda - 1/n, \lambda + 1/n)$ and $P_A(B_n) \neq 0$ for all n. By Riesz's Lemma, for each n we can find $f_n \in P_A(B_n)(H)$ with $\|f_n\| = 1$. Then since $d\mu_{P_A(B_n)f} = \chi_{B_n} d\mu_f$ and $f_n \in P_A(B_n)$, we must have

$$\begin{aligned}
\|(A - \lambda)f_n\|^2 &= \|(A - \lambda)P_A(B_n)f_n\|^2 \\
&= \int_{\mathbb{R}} (t - \lambda)^2 d\mu_{P_A(B_n)f_n}(t) \\
&= \int_{\mathbb{R}} (t - \lambda)^2 \chi_{B_n}(t) d\mu_{f_n}(t) \\
&\leq \sup_{t \in B_n} (t - \lambda)^2 \mu_{f_n}(B_n) \\
&\leq \frac{1}{n^2},
\end{aligned}$$

since $\mu_{f_n}(\cdot) \leq \|f_n\|^2 = 1$. This shows that (f_n) is a Weyl sequence, and so Lemma B.3.1 implies that $\lambda \in \sigma(A)$. Conversely, take $\varepsilon > 0$ with $B_\varepsilon = (\lambda - \varepsilon, \lambda + \varepsilon)$, and suppose $P_A((\lambda - \varepsilon, \lambda + \varepsilon)) = 0$. Define $\psi_\varepsilon(t) = \chi_{\mathbb{R} \setminus ((\lambda - \varepsilon, \lambda + \varepsilon)}(t - \lambda)^{-1}$. We note that if $B_1 \subset B_2$, then B_2 is the disjoint union of B_1 and $B_2 \setminus B_1$, so that $P_A(B_2) = P_A(B_1) + P_A(B_2 \setminus B_1)$, or $P_A(B_2 \setminus B_1) = P_A(B_2) - P_A(B_1)$.

We now have that

$$
\begin{aligned}
(A-\lambda)(P_A)_*(\psi_\varepsilon) &= (P_A)_*(t-\lambda)(P_A)_*(\psi_\varepsilon) \\
&= (P_A)_*((t-\lambda)\psi_\varepsilon) \\
&= (P_A)_*(\chi_{\mathbb{R}\setminus B_\varepsilon}) \\
&= P_A(\chi_{B_{\mathbb{R}\setminus\varepsilon}}) \\
&= P_A(\mathbb{R}) - P_A(B_\varepsilon) \\
&= I,
\end{aligned}
$$

since $P_A(B_\varepsilon) = 0$ by assumption. The same reasoning shows that $(P_A)_*(\psi_\varepsilon)(A-\lambda) = I$, so that $A-\lambda$ is invertible. Thus $\lambda \in \rho(A)$.

To see (i), we need the following fact: If B_1 and B_2 are Borel sets with $B_1 \subset B_2$, then $P_A(B_1) \leq P_A(B_2)$ in the sense that

$$\langle P_A(B_1)f, f\rangle \leq \langle P_A(B_2)f, f\rangle$$

for all $f \in H$ (this is the analogue of monotonicity for a spectral measure). In particular, if $P_A(B_2)$ is the zero operator, then so is $P_A(B_1)$. To see this, we note that since the characteristic functions satisfy $\chi_{B_1} \leq \chi_{B_2}$ for $B_1 \subset B_2$, we have

$$\int_{\mathbb{R}} \chi_{B_1} d\mu_f \leq \int_{\mathbb{R}} \chi_{B_2} d\mu_f,$$

which is equivalent to the above. Now if $\lambda \in (\lambda_1, \lambda_2)$ and $P_A((\lambda_1, \lambda_2)) = 0$, we find $\varepsilon > 0$ such that $B_\varepsilon \subset (\lambda_1, \lambda_2)$. Then the above shows that $P_A(B_\varepsilon) = 0$ as well. So the first part of the theorem implies that $\lambda \in \rho(A)$. Conversely, if $\lambda \in \sigma(A)$, then $P_A(B_\varepsilon) \neq 0$ for all $\varepsilon > 0$, so monotonicity again implies that $P_A((\lambda_1, \lambda_2)) \neq 0$.

To prove (ii) and (iii), we first note that since A is self-adjoint, $\rho(A)$ must be open, since $\sigma(A)$ must be closed. For each $\lambda \in \rho(A)$, we find an $\varepsilon > 0$ such that $B_\varepsilon \subset \rho(A)$. From this we get an open cover $\mathcal{O}$ of $\rho(A)$, and we may pass to a countable open subcover $\{O_n\}$ of $\rho(A)$. We can define $O_N^* = O_N \setminus \cup_{n<N} O_n$. These sets are all disjoint by construction, and part (i) shows that $P_A(O_N^*) = 0$ for all N. Then additivity of the spectral measure gives

$$P_A(\mathbb{R} \cap \rho(A)) = P_A(\rho(A)) = \sum_{N=1}^{\infty} O_N^* = 0.$$

This shows (ii), and now (iii) follows from

$$P_A(\sigma(A)) + P_A(\rho(A)) = P_A(\sigma(A) \cup \rho(A)) = P_A(\mathbb{R}) = I,$$

completing the proof. □

Remarkably, this theorem shows that if A has a cyclic vector, then we have

$$\mu_f(\mathbb{R}) = \langle f, f\rangle = \langle P_A(\sigma(A))f, f\rangle = \int_{\mathbb{R}} \chi_{\sigma(A)} d\mu_f = \mu_f(\sigma(A)),$$

while

$$\mu_f(\rho(A)) = 0.$$

Moreover, $\lambda \in \rho(A)$ if and only if there is an open neighborhood N of λ such that $P_A(N) = 0$, so that $\mu_f(N) = 0$. This shows that the support of μ_f is exactly $\sigma(A)$: supp$(\mu_f) = \sigma(A)$ for any f. We refer to this invariance by saying that A has **simple spectrum**. We note that in particular, we must have

$$\langle R_A(z)f, f\rangle = \int_{\sigma(A)} \frac{1}{\lambda - z} d\mu_f(\lambda)$$

for any choice of $f \in H$. It is often also beneficial to split the spectrum into two pieces: the **discrete spectrum**, denoted by $\sigma_d(A)$, includes all isolated points in the spectrum of A, and the **essential spectrum**, denoted by $\sigma_{ess}(A)$, includes all the points in the spectrum that are not isolated ($\sigma_{ess}(A) = \sigma(A) \setminus \sigma_d(A)$. Note that if we suppose that $P_A(\{\lambda\}) \neq 0$, then for $f \in P_A(\{\lambda\})(H)$, we have that

$$(A - \lambda I)f = (A - \lambda I)P_A(\{\lambda\})f = \int_{\mathbb{R}} (t - \lambda)\chi_{\{\lambda\}} d\mu_f(t) = \int_{\{\lambda\}} (t - \lambda) d\mu_f(t) = 0,$$

so that λ is an eigenvalue of A. This shows that if $\mu_f(\{\lambda\}) > 0$, then λ is an eigenvalue of A, so it follows from Theorem B.3.2 that the discrete spectrum is contained in the point spectrum, $i.e.$, $\sigma_d(A)$ can only contain eigenvalues of A. We have that

Lemma B.3.3. *If there is a Weyl sequence f_n converging to λ that also conveges weakly to 0, then $\lambda \in \sigma_{ess}(A)$.*

Proof. Suppose that $\lambda \in \sigma_d(A)$. Then λ is an eigenvalue of A, so again by Theorem B.3.2, there is an $\varepsilon > 0$ such that $P_A(\lambda - \varepsilon, \lambda + \varepsilon)$ is finite rank for $0 < \delta < \varepsilon$. In particular, E_λ is compact. Take $g_n = E_\lambda f_n$. Then the weak convergence of the f_n shows that $g_n \to 0$. Then we have that

$$\begin{aligned}
\|f_n - g_n\|^2 &= \|(I - P_A(\lambda - \varepsilon, \lambda + \varepsilon))f_n\|^2 \\
&= \|(P_A(\mathbb{R}) - P_A(\lambda - \varepsilon, \lambda + \varepsilon))f_n\|^2 \\
&= \|(P_A(\mathbb{R} \setminus (\lambda - \varepsilon, \lambda + \varepsilon))f_n\|^2 \\
&= \int_{\mathbb{R}\setminus(\lambda-\varepsilon,\lambda+\varepsilon)} d\mu_{f_n} \\
&\leq \frac{1}{\varepsilon^2} \int_{\mathbb{R}\setminus(\lambda-\varepsilon,\lambda+\varepsilon)} (t-\lambda)^2 d\mu_{f_n}(t) \\
&\leq \frac{1}{\varepsilon^2} \|(A - \lambda I)f_n\|^2.
\end{aligned}$$

This tends to 0, so that $\|g_n\| \to \|f_n\| = 1$, a contradiction. Hence, we must have $\lambda \in \sigma_{ess}(A)$. $\square$

An immediate consequence of this lemma is that since the spectrum of A is simple, the point masses of the measures $d\mu_f$ coincide with the discrete spectrum of A. We can also use this lemma to provide a quick proof of **Weyl's Theorem**: If A is self-adjoint and K is compact, then $\sigma_{ess}(A)$ and $\sigma_{ess}(A+K)$. First suppose that (f_n) is a sequence with $\|f_n\| = 1$ for all n. If $\|(A+K)f_n\| \to 0$, then clearly $\|Af_n\| \to 0$. If $\|Af_n\| \to 0$ and the f_n converge weakly to 0, then we have that $\|(A+K)f_n\| \leq \|Af_n\| + \|Kf_n\| \to 0$, since K is compact. Thus, A and $A+K$ share the same Weyl sequences that are weakly convergent to 0, so they must share the same essential spectrum. Now that we have this characterization of the spectrum in terms of the spectral measure, we can finally prove the following theorem:

Theorem B.3.4. *Let $F(z)$ be the Borel transform of a finite Borel measure μ and $dx = dm$ be the Lebesgue measure. Then*

i) $\mu(\{x \in \mathbb{R} : \lim_{\varepsilon \to 0^+} \mathrm{Im}(F(x + i\varepsilon)) = \infty\}) = 0$*;*

ii) $d\mu_{ac} = \frac{1}{\pi}\mathrm{Im}(F)\, dx$*;*

iii) $\mu(\{x\}) = \lim_{\delta \to 0^+} \lim_{\varepsilon \to 0^+} \frac{2}{\pi} \int_x^{x+\delta} \mathrm{Im}(F(\lambda + i\varepsilon)) d\lambda$*, for all* $x \in \mathbb{R}$*.*

Proof. We see that if $\lambda \in (x-\varepsilon, x+\varepsilon)$, then $(\lambda - x)^2 + \varepsilon^2 < \varepsilon^2$, so that

$$\frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} > \frac{1}{2\varepsilon}.$$

Then

$$\begin{aligned}
\operatorname{Im}(F(x+i\varepsilon)) &= \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} d\mu(\lambda) \\
&\geq \int_{(x-\varepsilon, x+\varepsilon)} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} d\mu(\lambda) \\
&> \frac{1}{2\varepsilon} \int_{(x-\varepsilon, x+\varepsilon)} d\mu \\
&= \frac{\mu(x-\varepsilon, x+\varepsilon)}{2\varepsilon}.
\end{aligned}$$

Then the monotonicity of μ proves (i) since the set T in Theorem A.3.6(i) must contain $\{x \in \mathbb{R} : \lim_{\varepsilon \to 0^+} \operatorname{Im}(F(x+i\varepsilon)) = \infty\}$ by the above inequality. To verify (ii), we note that since μ is finite, it suffices to show that if f is continuous and bounded, then

$$\frac{1}{\pi} \int_{\mathbb{R}} f \operatorname{Im}(F)\, dm = \int_{\mathbb{R}} f\, d\mu.$$

So suppose that f is continuous and bounded and denote $dx = dm(x)$. Then by Fubini's theorem we have

$$\begin{aligned}
\frac{1}{\pi} \int_{\mathbb{R}} f(x) \operatorname{Im}(F(x+i\varepsilon))\, dx &= \frac{1}{\pi} \int_{\mathbb{R}} f(x) \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} d\mu(\lambda)\, dx \\
&= \int_{\mathbb{R}} \frac{1}{\pi} \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} f(x)\, dx\, d\mu(\lambda).
\end{aligned}$$

Since f is bounded, we may take $|f(x)| \leq M$ for all $x \in \mathbb{R}$, and then Holder's equality gives

$$\begin{aligned}
\frac{1}{\pi} \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2} f(x)\, dx &\leq \sup_{x \in \mathbb{R}} |f(x)| \cdot \frac{1}{\pi} \int_{\mathbb{R}} \frac{\varepsilon}{(\lambda - x)^2 + \varepsilon^2}\, dx \\
&= \sup_{x \in \mathbb{R}} |f(x)| \cdot \frac{1}{\pi} \tan^{-1}\left(\frac{x-\lambda}{\varepsilon}\right)\Bigg|_{-\infty}^{\infty} \leq M.
\end{aligned}$$

So taking $\varepsilon = 1/n$ and applying the Lebesgue dominated convergence with respect to the measure μ as $n \to \infty$ shows that the integral on the left converges to some limit function $G(\lambda)$. We claim that $G = f$. We have

$$\begin{aligned}\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2} f(x)\,dx - f(\lambda) =& \frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2} f(x)\,dx \\ &- f(\lambda)\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}\,dx \\ =& \frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}(f(x)-f(\lambda))\,dx,\end{aligned}$$

but since f is continuous, we can find a $\delta_n > 0$ such that $|f(x) - f(\lambda)| < 1/n$ for $|x-\lambda| < \delta_n$. Then we have for each fixed n that

$$\begin{aligned}\left|\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}[f(x)-f(\lambda)]\,dx\right| \leq& \frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}|f(x)-f(\lambda)|\,dx \\ <& \frac{1}{n\pi}\int_{|x-\lambda|<\delta_n} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}\,dx \\ &+ 2M\cdot\frac{1}{\pi}\int_{|x-\lambda|\geq\delta_n} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}\,dx \\ \leq& \frac{1}{n} + \frac{2M}{\pi}\int_{|x-\lambda|\geq\delta_n} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2}\,dx.\end{aligned}$$

For each n, the second integral tends to zero as $\varepsilon \to 0^+$. This shows that

$$\lim_{\varepsilon\to 0^+}\left[\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2} f(x)\,dx - f(\lambda)\right] < \frac{1}{n}.$$

This for all n, so that

$$\lim_{\varepsilon\to 0^+}\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2} f(x)\,dx = f(\lambda).$$

This shows that

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\frac{1}{\pi}\int_{\mathbb{R}} f(x)\,\mathrm{Im}(F(x+i\varepsilon))\,dm =& \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}}\frac{1}{\pi}\int_{\mathbb{R}} \frac{\varepsilon}{(\lambda-x)^2+\varepsilon^2} f(x)\,dx\,d\mu(\lambda) \\ =& \int_{\mathbb{R}} f(\lambda)\,d\mu(\lambda),\end{aligned}$$

which establishes (ii). (iii) is just a restatement of the Stieltjes inversion formula. □

References

[1] D. Damanik and B. Simon, *Jost functions and Jost solutions for Jacobi matrices. II. Decay and analyticity*, Int. Math. Res. Not. (2006), Art. ID 19396, 32. MR2219207

[2] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics, vol. 3, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999. MR1677884

[3] J. S. Geronimo and K. M. Case, *Scattering theory and polynomials orthogonal on the real line*, Trans. Amer. Math. Soc. **258** (1980), no. 2, 467–494. MR558185

[4] J. S. Geronimo, *Scattering theory, orthogonal polynomials, and q-series*, SIAM J. Math. Anal. **25** (1994), no. 2, 392–419. MR1266566

[5] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. (2) **158** (2003), no. 1, 253–321. MR1999923

[6] C. Remling, *Jacobi matrices and orthogonal polynomials*, 2009. `http://www2.math.ou.edu/~cremling/teaching/lecturenotes/fa-new/ln13.pdf`.

[7] W. Rudin, *Real and complex analysis*, Third, McGraw-Hill Book Co., New York, 1987. MR924157

[8] B. Simon, *Orthogonal polynomials on the unit circle. Part 2*, American Mathematical Society Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005. Spectral theory. MR2105089

[9] G. Stolz and R. Weikard, *Seminar on Jacobi matrices*, 2004. `http://people.cas.uab.edu/~weikard/teaching/Jacobi_matrices.pdf`.

[10] G. Teschl, *Jacobi operators and completely integrable nonlinear lattices*, Mathematical Surveys and Monographs, vol. 72, American Mathematical Society, Providence, RI, 2000. MR1711536

[11] G. Teschl, *Mathematical methods in quantum mechanics*, Second, Graduate Studies in Mathematics, vol. 157, American Mathematical Society, Providence, RI, 2014. With applications to Schrödinger operators. MR3243083

[12] M. Zinchenko, *CMV matrices with super exponentially decaying Verblunsky coefficients*, Math. Model. Nat. Phenom. **9** (2014), no. 5, 282–294. MR3264323