On Random Schrödinger Operators

Mark Richard

Submitted under the supervision of Wei-Kuo Chen to the University Honors Program at the University of Minnesota-Twin Cities in partial fulfillment of the requirements for the degree of Bachelor of Science, *magna cum laude* in Mathematics.

May 5, 2018

Abstract

This report develops the mathematical properties of random operators of the Schrödinger equation, with the ultimate goal of proving the phenomenon of localization in a material that has high disorder. Localization is when a wave stops propagating due to random potential wells and destructive interference. It has long been of interest to determine under what conditions such localization will occur.

First, some functional analysis and measure theory is presented in the general setting of a self-adjoint operator T in a separable Hilbert space $\mathcal{H}$. Emphasis is put on the properties of the spectrum $\sigma(T)$, which is the set of λ such that $T - \lambda I$ is not invertible. When T is self-adjoint, $\sigma(T)$ is a closed subset of $\mathbb{R}$, and has physical meaning as the energy a particle can take. In addition, the measure-theoretic structure of the spectrum determines when localization will occur.

Once the mathematical theory is developed in general, specific operators are considered over the space of square-summable complex sequences, $\ell^2(\mathbb{Z}^d)$. The discrete Laplacian H_0 acts as a kinetic energy operator in $\ell^2(\mathbb{Z}^d)$, and its spectrum is found to be $[0, 4d]$. Then, a random potential multiplication operator V_ω is considered, where $V_\omega(n)$ is an IID random variable for each $n \in \mathbb{Z}^d$, with distribution P_0. It is found that $\sigma(V_\omega) = \text{supp } P_0$. Finally, the Hamiltonian operator is defined as $H_\omega = H_0 + V_\omega$. For $\mathbb{P}$-almost all ω, it is shown that $\sigma(H)$ is $[0, 4d] + \text{supp } P_0$.

With the proper setup, localization is defined and discussed. When a material has sufficiently high disorder, or is at a sufficiently low energy, localization can be shown to occur. Then, some weaker results are developed related to bounded states of a particle, which are implied by localization. The Green's functions are introduced as a powerful tool to describe how eigenfunctions decay. Some weak results are provided which help gain familiarity with using Green's functions.

A proof of localization is given which relies on strong multiscale analysis, a result which looks at eigenfunction decay on pairs of cubes in space. After this proof is given, the actual statement of multiscale analysis is considered. The heuristics of a proof are provided, with a discussion about large disorder implying localization at the end.

Finally, further discussion is provided about the strength of the assumptions in this paper. Results which are related to localization, as well as current research interests, are also looked at.

Contents

Introduction

This report develops the mathematical and physical preliminaries to understand the Anderson Model and the effect of localization. In some ways this is a condensed version of selected chapters in [13]. Topics are more carefully selected to maintain focus on the goal of proving localization. Some results near the end are not proven with full rigor. However, more proofs are presented in total, and those which follow closely with [13] provide more detail to aid the reader.

The author began this project with a background in undergraduate real analysis, linear algebra, probability and some physics. This thesis is meant to document the author's exploration of the topic in such a way that a student with a similar background and motivation can comfortably understand the topic. By working in the language of finite matrices for intuition, it is quite possible for readers with a similar background to understand the results given here. However, there is material hiding in the background. In particular, the mathematics developed requires some familiarity with measure theory. To fully appreciate the physical phenomenon being represented, it is good to have knowledge of introductory quantum mechanics (see [6], for example.)

A typical introduction to quantum mechanics starts with the time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi, \quad H = -\frac{\hbar^2}{2m}\nabla^2 + V \ . \tag{1}$$

Separation of variables, $\Psi(\mathbf{x},t) = \psi(\mathbf{x})\phi(t)$, yields the equation

$$i\hbar\,\psi(\mathbf{x})\frac{d\phi}{dt} = \phi(t)H\psi(\mathbf{x}) \ .$$

When performing separation of variables, we wish to get all functions of t on the left, and all functions of $\mathbf{x}$ on the right. Then, since these are independent, they must equal some separation constant E:

$$i\hbar\,\frac{1}{\phi(t)}\frac{d\phi}{dt} = \frac{1}{\psi(\mathbf{x})}H\psi(\mathbf{x}) = E \ .$$

The equation in time is an ordinary differential equation with solution $\phi(t) = \exp(-iEt/\hbar)$, and the equation in $\mathbf{x}$ is called the time-independent Schrödinger equation:

$$H\psi(\mathbf{x}) = E\psi(\mathbf{x}), \quad \Psi(\mathbf{x}, t) = \psi(\mathbf{x})e^{-iEt/\hbar} \, . \tag{2}$$

We can thus view E as an eigenvalue for the operator H, and ψ a corresponding eigenvector when viewed in the appropriate function space. More typically, ψ is called an eigenfunction.

This paper looks at a different space, namely $\ell^2(\mathbb{Z}^d)$. This treats physical space as a grid, which helps represent the crystalline structure of many regular solids. In particular, this paper concerns itself with a random form of H in this space, where the potential operator V is a random variable at each point in space. Then, we are interested in the exponential decay of the eigenfunctions ψ. This occurs when localization happens in a solid.

Anderson localization has been an important phenomenon since its discovery in 1958, giving rise to interesting physical systems and providing an impetus to discover new mathematics to describe it. It is used to learn more about conductivity in alloy materials, and is also an important component in the development of LED light bulbs. Mathematicians have used many methods to describe this effect, with it first being proposed in a discrete system on $\mathbb{Z}^d$, as is discussed in this paper; more recently some very unique methods have been developed to study it in continuous space (see [7]).

To understand why Anderson localization is of such interest, imagine the following situation: You are standing in a room, full of boxes and jagged objects across various surfaces, and across the room you see your friend clap. But you never hear the sound which should follow. Somewhere along the way, the sound waves of the clap bounced around in such a way as to destructively interfere, and the wave never reaches you. This, in a sense, is localization.

Localization was first discovered in the context of electron waves moving through a metal alloy. Current moves through an object via electrons, so if the energy of the electrons as they propagate through the material is arranged randomly in a particular way, there becomes a high chance they will never reach the other end of the object. Specifically, electron waves may get trapped, or *localized*, with high probability in a particular part of the material. Hence the expected propagation would not occur.

The study of localization has moved beyond the realm of its discovery with electrons. As waves are such a fundamental object to modern physics, it has been natural for studies to consider how localization occurs in sound and light waves as well [11]. For example, development of LEDs has benefited from mathematical understanding of localization.

LEDs are created by the control of electrons jumping orbitals at specific energy levels, causing emittance of a particular color. When localization can be controlled further, LEDs can become far more efficient.

When discussing the movement of electrons, we use energy as the central piece of information. Their kinetic energy is well-known, but the potential energy is a dynamic quantity, changing with the movement of the electrons themselves. As electrons move through orbitals, their charges interact with other particles. In addition, when an electron leaves an orbital a potential "hole" is left there, which strongly affects electrons in the area. These dynamical interactions make it very difficult to predict where electrons may be localized in a material, but mathematics has come a long way in describing the circumstances under which it will occur. While slightly out of date, [16] has a good description of the state of discovery in localization.

What makes localization so interesting to many is its random nature. Due to the way alloys and other materials are made, their structures are effectively a random mix of multiple elements. This randomness can be mathematicall realized as an electron at point x being affected by a potential $V_\omega(x)$ that is caused by its neighbors. A deterministic potential may have terms of the form $V(x) = \sum_i qf(x-i)$, where q is some coupling constant or charge, and $i \in \mathbb{Z}^d$ are the neighbors of the point x. However, there may be a random component where each $i \in \mathbb{Z}^d$ is displaced some amount from its expected position in the structure, which may be described by $f(x-i-\delta_i(\omega))$. It is also possible that the charges or coupling constants are randomly changing, forcing the terms to become $q_i(\omega)f(x-i)$.

It will likely aid the reader to keep these examples in mind, and continually refer to the physical phenomenon that localization represents. While the math can be heavy when describing the exponential decay of eigenfunctions, recall this means the probability an electron wave is moving outside of some finite space is vanishing to 0. The wave will not propagate, and thus current (or sound, or light) will not travel as expected.

Localization continues to provide interesting mathematical and physical problems to be solved. This is a rich area for discovery, with many deep implications for how the physical world works. As we continue to move forward in understanding this phenomenon, many surprises likely lie in wait.

Chapter 1

Preliminaries

This chapter develops the relevant measure theory and functional analysis required to understand how localization is mathematically described. The proofs presented are mostly adapted from [13] and [17], though arguments are expanded upon to provide greater clarification and context.

Throughout this chapter, we will work on a separable complex Hilbert space $\mathcal{H}$, and use a self-adjoint linear operator T. This means that $T = T^*$, where T^* is the adjoint operator. This implies $\langle v, Tw\rangle = \langle Tv, w\rangle$. There is a well-defined norm

$$\|T\| := \sup_{f\in\mathcal{H}} \frac{\|Tf\|}{\|f\|}$$

which will be used throughout.

1.1 Spectral Calculus

The basic object of interest for much of this paper is the spectrum of an operator. We will see the spectrum has a physical meaning as an energy interval, and is the object through which we define localization. First, we must define its complement, the resolvent set.

Definition 1.1 (Resolvent set)**.** The resolvent set of T, denoted $\rho(T)$, is the set of all complex numbers μ such that $T-\mu := T-\mu I$ is a bijective mapping between the domain of T and $\mathcal{H}$.

Since the mapping is bijective, the operator $T - \mu$ is invertible. We call $(T - \mu)^{-1}$ the resolvent of T and μ.

Definition 1.2 (Spectrum)**.** The spectrum of T, denoted $\sigma(T)$, is the complement of the resolvent, $\mathbb{C}-\rho(T)$. In particular, it consists of all λ such that $T-\lambda$ is not invertible.

The spectrum is an extension of the idea of an eigenvalue. If λ is an eigenvalue, so that $Tv = \lambda v$ for some non-zero $v \in \mathcal{H}$, then $(T-\lambda)v = 0$ for a non-zero v, and hence $T-\lambda$ is not injective. Thus, it can be helpful for the reader to think about matrices and their eigenvalues when working through much of what is to follow. However, it is also necessary to note that not every element of the spectrum is an eigenvalue.

Now we turn to the technical details of why we would want our T to be self-adjoint. It turns out $\sigma(T)$ has very nice properties in this case, as summarized by the following theorem.

Theorem 1.3. *For self-adjoint T, the spectrum $\sigma(T)$ is a closed subset of $\mathbb{R}$. If in addition T is bounded, then $\sigma(T)$ is also bounded, and hence compact.*

Proof. Because T is self-adjoint, $T = T^*$, so $\langle Tx, x\rangle = \langle x, Tx\rangle$ for all x. Assume $\lambda \in \sigma(T)$. Then $Tx = \lambda x$ for some x, so

$$\begin{aligned}\langle Tx, x\rangle &= \langle x, Tx\rangle \\ \langle \lambda x, x\rangle &= \langle x, \lambda x\rangle \\ \langle x, \overline{\lambda} x\rangle &= \langle x, \lambda x\rangle \ .\end{aligned}$$

The last step on the left is moving λ through the inner product, requiring us to take the conjugate. The final equality implies that $\lambda = \overline{\lambda}$, hence $\lambda \in \mathbb{R}$, so $\sigma(T) \subset \mathbb{R}$.
To show $\sigma(T)$ is closed in $\mathbb{R}$, we show $\rho(T)$ is open in $\mathbb{C}$. We first show a calculation with a formal power series, then consider its convergence for a particular case. Choose $\lambda_0 \in \rho(T)$. Then we can write

$$\frac{1}{\lambda - T} = \frac{1}{\lambda - T + \lambda_0 - \lambda_0} = \frac{1}{\lambda_0 - T}\left(\frac{1}{1 - \frac{\lambda_0 - \lambda}{\lambda_0 - T}}\right) \ .$$

We can expand the expression in parentheses as a geometric series:

$$\frac{1}{\lambda - T} = \frac{1}{\lambda_0 - T}\sum_{n=0}^{\infty}\left(\frac{\lambda_0 - \lambda}{\lambda_0 - T}\right)^n \ .$$

This series converges precisely when $|\lambda_0 - \lambda|/|\lambda_0 - T| < 1$. We consider the inverse operator $R_{\lambda_0}(T) = (\lambda_0 I - T)^{-1}$, so we can write

$$\hat{R} := \frac{1}{\lambda I - T} = R_{\lambda_0}(T)\sum_{n=0}^{\infty}(\lambda_0 - \lambda)^n \left[R_{\lambda_0}(T)\right]^n, \quad |\lambda_0 - \lambda| < \frac{1}{\|R_{\lambda_0}(T)\|}$$

based on our analysis of convergence. By definition $\hat{R} = (\lambda I - T)^{-1}$, so we see immediately that $(\lambda I - T)\hat{R} = \hat{R}(\lambda I - T) = I$, which implies $\hat{R} = R_\lambda(T)$, so $\lambda \in \rho(T)$. In particular, since λ is in the $\|R_{\lambda_0}(T)\|^{-1}$ ball around $\lambda_0 \in \rho(T)$, for *any* $\lambda_0 \in \rho(T)$ there is some open neighborhood B such that $\lambda \in B$ implies $\lambda \in \rho(T)$. By definition, $\rho(T)$ is indeed open. Hence $\sigma(T) = \mathbb{C} - \rho(T)$ is closed.
Finally, assume that T is bounded. We will show that $\sigma(T)$ is bounded by $\|T\|$. Indeed, let $|\lambda| > \|T\|$ and define the sequence of operators

$$R_{\lambda,N} = -\frac{1}{\lambda}\sum_{n=0}^{N}\frac{T^n}{\lambda^n}\,. \tag{1.1}$$

Since $|\lambda| > \|T\|$ by assumption, this sequence converges as $N \to \infty$; assume it converges to some operator S_λ. We aim to show that $S_\lambda = R_\lambda(T) = (\lambda I - T)^{-1}$, and thus we compare $S_\lambda(T - \lambda I)$ and $(T - \lambda I)S_\lambda$ with I. We will only show the formal computation for the first case, since the second case is exactly analogous, only switching the order in which we multiply. For any $N \in \mathbb{N}$ we have

$$\begin{aligned}\|S_\lambda(T-\lambda I) - I\| &= \|S(T-\lambda I) - R_{\lambda,N}(T-\lambda I) + R_{\lambda,N}(T-\lambda I) - I\| \\ &\le \|(S - R_{\lambda,N})(T-\lambda I)\| + \|R_{\lambda,N}(T-\lambda I) - I\|\end{aligned}$$

where the second line comes from factoring out the $(T-\lambda I)$ on the right, and the triangle inequality. Now, we can simplify the first term with the Cauchy-Schwarz inequality, and then we expand out the second term with the definition of $R_{\lambda,N}$, obtaining

$$\begin{aligned}\|S_\lambda(T-\lambda I) - I\| &\le \|S_\lambda - R_{\lambda,N}\|\,\|T-\lambda I\| + \left\|-\frac{T}{\lambda}\sum_{n=0}^{N}\frac{T^n}{\lambda^n} + \sum_{n=0}^{N}\frac{T^n}{\lambda^n} - I\right\| \\ &\le \|S_\lambda - R_{\lambda,N}\|\,\|T-\lambda I\| + \left\|\sum_{n=0}^{N}\frac{T^n}{\lambda^n} - \sum_{n=1}^{N+1}\frac{T^n}{\lambda^n} - I\right\| \\ &\le \|S_\lambda - R_{\lambda,N}\|\,\|T-\lambda I\| + \left\|\frac{T^{N+1}}{\lambda^{N+1}}\right\|\,.\end{aligned}$$

We take the limit as $N \to \infty$. By assumption, $R_{\lambda,N} \to S_\lambda$, so the first term goes to 0. Also by assumption, $\|T\|^n/|\lambda|^n \to 0$, hence the entire right side goes to 0. Thus $\|S_\lambda(T-\lambda I) - I\| \to 0$, so $S_\lambda(T-\lambda I) = I$. As stated before, we can do a similar calculation to find $(T-\lambda I)S_\lambda = I$, which implies that $S_\lambda = R_\lambda(T)$. In particular, if $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$. Thus $\sigma(T)$ is bounded by $\|T\|$. By the Heine-Borel theorem, $\sigma(T)$ is also compact. □

These are very useful results when discussing the spectrum. Later, the fact that $\sigma(T)$ is closed will allow us to determine a useful way to approximate the spectrum by the set of

generalized eigenvalues. Its compactness will be used in defining $f(T)$ in the next part of this section. Also, knowing that $\sigma(T) \subset \mathbb{R}$ for self-adjoint T is of significant physical importance. The Hamiltonian operator H we will work with is indeed self-adjoint, thus the eigenvalues have a physical interpretation as the energy levels attainable by a particle.

One property of the spectrum which is familiar in the finite case is *unitary invariance.* When dealing with matrices, we can consider diagonalizing a matrix A as $A = P^{-1}DP$. Then the eigenvalues of A and D are both the same. This concept extends to operators in general. If we have a self-adjoint linear operator A on a Hilbert space, then for any unitary operator U (which means $UU^* = U^*U = I$), we know that the spectrum of A and the spectrum of UAU^{-1} are the same. There is one other useful property, whose proof requires a bit more measure theory than is warranted for this paper. The interested reader can find a proof online in [1].

Theorem 1.4. *Suppose M_f is the multiplication operator defined by $M_f\varphi = f\varphi$, where f is a continuous complex-valued function. Suppose the domain of f is open and the range of f is closed. Then $\sigma(M_f)$ is precisely the range of f.*

More broadly, the theorem says that $\sigma(M_f)$ is the *essential range* of f, which is what the cited proof deals with. For our purposes, it is much easier to note that when mapping from an open set, the essential range of f is just the closure of the image, $\overline{\text{im } f}$. Thus, if the range of f is closed the essential range is the range of f. With this knowledge, it is often useful to find a unitary operator that transforms the operator of interest into a multiplication operator. Then the spectrum is simple to find. This will be the approach taken with the discrete Laplacian in Section 2.1.

The next task is developing $f(T)$ for certain functions f. This will in turn allow us to formulate a calculus based on operators and their spectra. It is simple to define $P(T)$ for any polynomial P. It is understood T^n is T composed with itself n times, so

$$T^n\varphi = \underbrace{T(T(\cdots(T(\varphi))\cdots))}_{n \text{ times}} . \tag{1.2}$$

With this definition, there are a few properties of the spectrum we can make use of. First, a lemma whose proof is is beyond the scope of this paper. (The interested reader can find a proof in [17].)

Lemma 1.5. *Let T be a bounded, self-adjoint operator. Then $\|T\| = \sup_{\lambda\in\sigma(T)} |\lambda|$.*

Another useful result with a rather simple proof tells us the spectrum can "move inside" the polynomial for calculation:

Lemma 1.6. *For a polynomial P and a bounded, self-adjoint operator T,*

$$\sigma(P(T)) = \{P(\lambda) \mid \lambda \in \sigma(T)\} \ . \tag{1.3}$$

Proof. Let P be a polynomial and $\lambda \in \sigma(T)$. Then the polynomial $P(x) - P(\lambda)$ has a root at $x = \lambda$, so it factors as

$$P(x) - P(\lambda) = (x - \lambda)Q(x)$$

for some polynomial Q. Thus, we have $P(T) - P(\lambda) = (T - \lambda)Q(T)$. Since $T - \lambda$ is not invertible by assumption, we must have $P(T) - P(\lambda)$ is also not invertible, thus $P(\lambda)$ is in $\sigma(P(T))$, and $\{P(\lambda) \mid \lambda \in \sigma(T)\} \subset \sigma(P(T))$.

For the converse, suppose $\mu \in \sigma(P(T))$. Factor $P(x)-\mu = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$. Then $P(T) - \mu = (T - \lambda_1)\cdots(T - \lambda_n)$. By assumption, $P(T) - \mu$ is not invertible, hence $T - \lambda_k$ is not invertible for some k. We see that this λ_k is a root of $P(x) - \mu$, so $P(\lambda_k) = \mu$. Hence any $\mu \in \sigma(P(T))$ can be written as $P(\lambda)$ for some $\lambda \in \sigma(T)$, so $\sigma(P(T)) \subset \{P(\lambda) \mid \lambda \in \sigma(T)\}$. By double inclusion, these sets are equal. □

These two results help us easily find the norm of $P(T)$.

Theorem 1.7. *For a bounded, self-adjoint operator T and any polynomial P, we have*

$$\|P(T)\| = \sup_{\lambda \in \sigma(T)} |P(\lambda)| \ . \tag{1.4}$$

Proof. From Lemma 1.5 we know $\|P(T)\| = \sup_{\lambda \in \sigma(P(T))} |\lambda|$. From Lemma 1.6 we know each $\lambda \in \sigma(P(T))$ can be written as $P(\lambda)$ for some $\lambda \in \sigma(T)$. Thus the $\lambda \in \sigma(P(T))$ moves out to taking the supremum of all $P(\lambda)$, where $\lambda \in \sigma(T)$ as shown. □

Now that we fully understand how polynomials interact with our operator T, we can move forward. To reiterate, we are assuming that T is bounded and self-adjoint, so $\sigma(T)$ is a compact set of real numbers. Let $C(\sigma(T))$ be the set of continuous complex-valued functions on $\sigma(T)$. The Weierstrass approximation theorem tells us any $f \in C(\sigma(T))$ can be uniformly approximated by polynomials. In particular, there is a sequence of polynomials $\{P_n\}$ such that $\lim_{n\to\infty} P_n = f$ uniformly on the compact $\sigma(T)$. This implies the natural definition

$$f(T) = \lim_{n\to\infty} P_n(T), \quad f \in C(\sigma(T)) \ . \tag{1.5}$$

Without delving into the technicalities one would cover in a real analysis course, since $f(T)$ is a uniform limit of polynomials it immediately satisfies

$$(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T) \tag{1.6}$$

$$(f \cdot g)(T) = f(T)g(T) \tag{1.7}$$

$$\overline{f}(T) = f(T)^* \, . \tag{1.8}$$

Also, if $f \geq 0$ we can write $f(T) = g(T)^2 = g(T)g(T)$ for $g = \sqrt{f}$, and we see that

$$\langle \varphi, f(T)\varphi \rangle = \langle g(T)\varphi, g(T)\varphi \rangle \geq 0 \tag{1.9}$$

since T is self-adjoint.

We will now develop a formulation of the inner product in terms of an integral with respect to a certain measure. One can see this outlined in [13] or, for more details, see [17]. Notice the map $f \mapsto \langle \varphi, f(T)\varphi \rangle$ is a linear functional. The Riesz-representation theorem says for any $\varphi \in \mathcal{H}$, there is a positive, bounded measure $\mu_{\varphi,\varphi}$ defined on $\sigma(T)$ such that for all $f \in C(\sigma(T))$ we have

$$\langle \varphi, f(T)\varphi \rangle = \int_{\sigma(T)} f(\lambda) d\mu_{\varphi,\varphi}(\lambda) \, . \tag{1.10}$$

In particular, take f to be the identity map, so that $\langle \varphi, \varphi \rangle = \mu_{\varphi,\varphi}$. Then, the polarization identity allows us to extend this relation to any $\varphi, \psi \in \mathcal{H}$. Specifically,

$$\begin{aligned} 4\mu_{\varphi,\psi} &= 4\langle \varphi, \psi \rangle \\ &= \langle \varphi + \psi, \varphi + \psi \rangle + \langle \varphi - \psi, \varphi - \psi \rangle + i \left(\langle \varphi - i\psi, \varphi - i\psi \rangle - \langle \varphi + i\psi, \varphi + i\psi \rangle \right) \\ &= \mu_{\varphi+\psi,\varphi+\psi} + \mu_{\varphi-\psi,\varphi-\psi} + i \left(\mu_{\varphi-i\psi,\varphi-i\psi} - \mu_{\varphi+i\psi,\varphi+i\psi} \right) \end{aligned}$$

Thus for each pair of functions $\varphi, \psi \in \mathcal{H}$ there is a complex-valued measure $\mu_{\varphi,\psi}$ such that

$$\langle \varphi, f(T)\psi \rangle = \int_{\sigma(T)} f(\lambda) d\mu_{\varphi,\psi}(\lambda) \, . \tag{1.11}$$

These relations give us an alternative way to describe bounded, measurable functions, in lieu of the limit of polynomials used for the continuous functions. Indeed, since we are using integrals these operators will similarly satisfy the relations (1.6)-(1.9) above. In addition, we have the inequality

$$\|f(T)\| \leq \sup_{\lambda \in \sigma(T)} |f(\lambda)| \, . \tag{1.12}$$

If f is continuous it can be defined as a polynomial limit, giving equality above.

A topic of vital importance are the spectral measures of an operator. These will show up repeatedly in this paper, more than the definitions of $f(A)$, which are included as a necessary and interesting technicality.

Consider a Borel set $B \subset \mathbb{R}$ (i.e. those sets formed from open sets in $\mathbb{R}$ from complement and countable unions and intersections). We wish to find some operators that work as "building-blocks" for other operators in a natural way. One may imagine a δ function. This is a bit too limiting, so we look at the characteristic function.

Definition 1.8 (Characteristic Function)**.** For any Borel set $B \subset \mathbb{R}$, the characteristic function χ_B is defined as

$$\chi_B(\lambda) = \begin{cases} 1 & \text{if } \lambda \in B \\ 0 & \text{else.} \end{cases} \tag{1.13}$$

Since χ_B is measurable, we can define its operator form $\chi_B(T)$. If we fix an operator T, then we denote $\chi_B(T)$ by $\mu(B)$. Each operator $\mu(B)$ is an orthogonal projection, meaning that $\mu^2(A) = \mu(A)$. In addition, we have

$$\mu(\emptyset) = 0 \tag{1.14}$$

$$\mu(\sigma(T)) = I \tag{1.15}$$

$$\mu(M \cap N) = \mu(M)\mu(N)\ . \tag{1.16}$$

Finally, if there is a family of pairwise disjoint Borel sets B_n, then for all $\varphi \in \mathcal{H}$

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right)\varphi = \sum_{n=1}^{\infty} \mu(B_n)\varphi\ . \tag{1.17}$$

This is very reminiscent of a measure (or if the reader has not seen a formal measure before, perhaps the definition of a probability density may also come to mind). This implies the following definition.

Definition 1.9 (Projection-valued spectral measure)**.** For a fixed self-adjoint operator T, the family of operators $\chi_B(T) = \mu(B)$ is called the projection-valued spectral measure of T.

These operators interact with the previously-defined integral definitions through the relation

$$\langle \varphi, \mu(B)\psi \rangle = \mu_{\varphi,\psi}(B)\ . \tag{1.18}$$

Notice how this relates to the previous definition. We can realize this in integral form as

$$\langle \varphi, \mu(B)\psi \rangle = \int_{\sigma(T)} \chi_B(\lambda)\, d\mu_{\varphi,\psi}(\lambda) \; .$$

Hence these projection-valued spectral measures have a way of measuring to what extent the set B intersects with the spectrum.

Now that this calculus has been well-defined, it allows us to work with a very large class of functions. In particular the operator e^{-itH} can be used, which is a formal solution to the time-independent Schrödinger equation.

There are a few more terms related to the spectrum which will be useful when proving results about localization. It is quite common to classify sets (and domains of functions) by their Lebesgue measure. Now that we have developed a concept of spectral measures, we can do similarly here. We say two Borel sets A and B agree up to a set of spectral measure 0 if $\mu(A - B) = \mu(B - A) = 0$.

1.2 Spectrum Example

It is good to look at an example which directly calculates the spectrum of an operator, using some of the tools developed in the previous section. Consider the left shift operator in $\ell^2(\mathbb{Z})$, call it S. This acts on a sequence $(a_1, a_2, \cdots)$, $a_i \in \mathbb{C}$ by the rule

$$S(a_1, a_2, \cdots) = (a_2, a_3, a_4, \cdots) \; . \tag{1.19}$$

We wish to find the spectrum of the shift operator. We'll first find all eigenvalues. Hence we want all λ such that

$$S(a_1, a_2, \cdots) = \lambda(a_1, a_2, \cdots) \tag{1.20}$$

$$(a_2, a_3, \cdots) = \lambda(a_1, a_2, \cdots) \; . \tag{1.21}$$

In this case, it is easier to first determine what these eigenvectors must be. They satisfy the relationship $a_i = \lambda a_{i-1}$ for $i \geq 2$, with a_1 being arbitrary. Thus any eigenvector v_λ is of the form

$$v_\lambda = a_1(1, \lambda, \lambda^2, \cdots) \; . \tag{1.22}$$

In order to determine which λ are actually eigenvalues, we must remember we are in the space $\ell^2(\mathbb{Z})$. A vector $(a_1, a_2, \cdots) \in \ell^2(\mathbb{Z})$ whenever $\sum |a_i|^2 < \infty$. Thus we want

$$|a_1|^2 \sum_{i=0}^{\infty} |\lambda|^{2i} < \infty \; . \tag{1.23}$$

We know that a geometric series converges if and only if $|\lambda| < 1$. Thus, the set of eigenvalues $\varepsilon(S)$ of S is just

$$\varepsilon(S) = \{\lambda \in \mathbb{C} \mid |\lambda| < 1\} . \tag{1.24}$$

Notice that S is a bounded operator. In particular

$$\|Sa\| = \sum_{i=2}^{\infty} |a_i|^2 \leq \sum_{i=1}^{\infty} |a_i|^2 = \|a\| .$$

Thus, taking the supremum over all $a \in \ell^2(\mathbb{Z}^d)$, we get $\|S\| \leq 1$. Taking the sequence $a_1 = 0$, $a_2 = 1$ and $a_i = 0$ for $i \geq 3$, we have $\|Sa\| = \|a\| = 1$. Since S is bounded its spectrum must be closed. Therefore

$$\sigma(S) = \overline{\varepsilon(S)} = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\} . \tag{1.25}$$

1.3 Measures and Subspaces

As a set, the spectrum does not actually provide much information, particularly when considering the physical phenomena it is supposed to represent. It is the measure-theoretic structure beneath the spectrum which determines everything. This section describes how to partition a Hilbert space $\mathcal{H}$ into three distinct subspaces by considering a measure induced by its elements.

We define a *Bounded Borel measure* to be a complex σ-additive function ν on the Borel sets $B(\mathbb{R})$ such that $\|\nu\|$, defined by

$$\|\nu\| = \sup \left\{ \sum_{i}^{N} |\nu(A_i)| \;\middle|\; A_i \in B(\mathbb{R}) \text{ are pairwise disjoint} \right\}$$

is finite. We call $\|\nu\|$ the total variation. A *positive Borel measure* is a non-negative σ-additive function m on the Borel sets such that $m(A)$ is bounded when A is bounded.

These two definitions are now used for the three types of measures that will concern us for the rest of the paper.

Definition 1.10 (Pure point measure)**.** A bounded Borel measure ν on $\mathbb{R}$ that is concentrated on a countable set A, so that $\nu(\mathbb{R} - A) = 0$.

Any $x \in \mathbb{R}$ such that $\nu(\{x\}) \neq 0$ is called an atom of ν. If $\alpha_i = \nu(\{x_i\})$ and $\delta_{x_i}(y) = 1$ when $y = x_i$ and 0 otherwise, any pure point measure can be explicitly written as

$$\nu(y) = \sum_{i} \alpha_i \delta_{x_i}(y) .$$

In general, we call a measure *continuous* if it has no atoms. We consider two cases of continuous measures.

Definition 1.11 (Absolutely continuous measure)**.** A continuous, bounded Borel measure ν is absolutely continuous (w.r.t. Lebesgue measure L) if there is a measurable function $\varphi \in L$ such that

$$\nu(A) = \int_A \varphi(x)\ L(dx)\ .$$

Hence, if $L(A) = 0$ for a Borel set A, then $\nu(A) = 0$.

Definition 1.12 (Singular continuous measure)**.** A continuous measure ν concentrated on a set N of Lebesgue measure 0. Thus $\nu(\{x\}) = 0$ for all $x \in \mathbb{R}$, and $\nu(\mathbb{R} - N) = 0$.

A typical (non-trivial) example is the measure induced by the Cantor function, where one linearly interpolates between points in the Cantor set.

We then use the following theorem to make use of these definitions.

Theorem 1.13 (Lebesgue's Decomposition Theorem)**.** *Given a bounded Borel measure ν, there exist measures ν_{ac}, ν' such that $\nu = \nu_{ac} + \nu'$, ν_{ac} is absolutely continuous with respect to the Lebesgue measure, and ν' is defined on a set of Lebesgue measure 0. Furthermore, ν' can be decomposed into $\nu' = \nu_{sc} + \nu_{pp}$, where ν_{sc} is singular continuous and ν_{pp} is pure point.*

We now look back at our spectral calculus to determine how to make use of these definitions. Let T be a self-adjoint operator on $\mathcal{H}$, with associated spectral measure μ. Then for a Borel set A we know $\mu(A)$ is a projection operator whose inner product is defined by

$$\langle \varphi, \mu(A)\psi \rangle = \mu_{\varphi,\psi}(A)\ .$$

In addition, we also have the relation

$$\langle \varphi, T\psi \rangle = \int_{\sigma(T)} \lambda d\mu_{\varphi,\psi}(\lambda)$$

with $\mu_{\varphi,\psi}$ a complex-valued measure, and $\mu_\varphi = \mu_{\varphi,\varphi}$ a positive measure. Since $\mu(A)$ is self-adjoint, we show

$$\begin{aligned} |\mu_{\varphi,\psi}(A)| &= |\langle \varphi, \mu(A)\psi \rangle| \\ &= |\langle \mu(A)\varphi, \mu(A)\psi \rangle| \\ &\leq (\langle \mu(A)\varphi, \mu(A)\varphi \rangle)^{\frac{1}{2}} (\langle \mu(A)\psi, \mu(A)\psi \rangle)^{\frac{1}{2}} \\ &= \mu_\varphi(A)^{\frac{1}{2}} \mu_\psi(A)^{\frac{1}{2}}\ . \end{aligned} \tag{1.26}$$

More details about the following discussion can be found in [17], as these results are more technical in nature. First, we define the sets $\mathcal{H}_{pp}, \mathcal{H}_{sc}, \mathcal{H}_{ac}$ as

$$\begin{aligned}\mathcal{H}_{pp} &= \{\varphi \in \mathcal{H} \mid \mu_\varphi \text{ is pure point}\} \\ \mathcal{H}_{sc} &= \{\varphi \in \mathcal{H} \mid \mu_\varphi \text{ is singular continuous}\} \\ \mathcal{H}_{ac} &= \{\varphi \in \mathcal{H} \mid \mu_\varphi \text{ is absolutely continuous}\} \ .\end{aligned}$$

It is interesting to note that each of these subspaces are mutually orthogonal and closed in $\mathcal{H}$. In addition, $\mathcal{H}$ can be written as the direct sum of these three spaces. Given our operator T, we define the restrictions to the spaces naturally. If $D(T)$ is the domain of T, then

$$T_{pp} = T\,\big|_{\mathcal{H}_{pp}\cap D(T)}, \quad T_{sc} = T\,\big|_{\mathcal{H}_{sc}\cap D(T)}, \quad T_{ac} = T\,\big|_{\mathcal{H}_{ac}\cap D(T)} \ .$$

These operators map the subspaces to themselves, so if $\varphi \in \mathcal{H}_{pp}$, then $T_{pp}\varphi \in \mathcal{H}_{pp}$. Finally, we can define the spectra with which we were originally interested.

Definition 1.14 (Pure point spectrum)**.** The pure point spectrum of T is $\sigma(T_{pp})$, denoted by $\sigma_{pp}(T)$.

Similarly, we define the *singular continuous spectrum* and *absolutely continuous spectrum* to be $\sigma(T_{sc})$ and $\sigma(T_{ac})$, denoted by $\sigma_{sc}(T)$ and $\sigma_{ac}(T)$ respectively.

As we consider some specific operators in the next chapter, there are two technical details to work with. First, our main operator is defined as the sum of two other operators. To ensure this operator is self-adjoint, we need the following theorem.

Theorem 1.15 (Kato-Rellich)**.** *Suppose A is self-adjoint on $\mathcal{H}$ and B is bounded with respect to A, so that*

$$\|Bf\| \leq a\|Af\| + b\|f\|, \quad f \in \mathcal{H}, \quad 0 < a < 1,\ b \geq 0 \ .$$

Then $A + B$ is self-adjoint on $\mathcal{H}$.

Second, an operator might not be self-adjoint on the entire Hilbert space $\mathcal{H}$ we are considering. Instead, they may be self-adjoint on a much smaller subspace $\mathcal{V}$. We are then interested in extending this operator to a larger space $\mathcal{F}$ (which may be all of $\mathcal{H}$, or still a proper subspace.)

Definition 1.16 (Essentially self-adjoint)**.** Suppose an operator T is self-adjoint on a subspace $\mathcal{V} \subset \mathcal{H}$. We say T is essentially self-adjoint on $\mathcal{F} \supset \mathcal{V}$ if there exists a unique extension E which is self-adjoint on $\mathcal{F}$, and coincides with T when restricted to $\mathcal{V}$.

It is important to note that the uniqueness of the extension allows the Kato-Rellich theorem above to be true if A is essentially self-adjoint.

Now we can state a final result. It gives a very useful way to determine if a particular λ is an element of the spectrum. More specifically, it tells us that any $\lambda \in \sigma(T)$ is approximately an eigenvalue of T as well. The proof given follows [9].

Theorem 1.17 (Weyl Criterion)**.** *Let T be essentially self-adjoint on $\mathcal{F}$. Then $\lambda \in \sigma(T)$ if and only if there is a sequence $\{\varphi_n\} \in \mathcal{F}$ with $\|\varphi_n\| = 1$ such that $\|(T-\lambda)\varphi_n\| \to 0$ as $n \to \infty$.*

Note, we call the sequence $\{\varphi_n\}$ a Weyl sequence for λ and T.

Proof. The result is trivial if λ is an eigenvalue. So, suppose $\lambda \in \sigma(T)$ is not an eigenvalue. In particular, this means $\ker(T-\lambda) = \{0\}$, and hence $T-\lambda$ is injective. So, for restricted domain we can define an unbounded $(T-\lambda)^{-1}$. In particular, there is a sequence $\{\psi_n\}$ such that $\|\psi_n\| = 1$ and $\|(T-\lambda)^{-1}\psi_n\| \to \infty$. Then, we can define

$$\varphi_n = [(T-\lambda)^{-1}\psi_n]\|(T-\lambda)^{-1}\psi_n\|^{-1} \ .$$

Then $\|\varphi_n\| = 1$ and each $\|\varphi_n\|$ is in the domain of $(T-\lambda)$ by construction. We have

$$\begin{aligned}\|(T-\lambda)\varphi_n\| &= \frac{\|(T-\lambda)(T-\lambda)^{-1}\psi_n\|}{\|(T-\lambda)^{-1}\psi_n\|} \\ &= \frac{\|\psi_n\|}{\|(T-\lambda)^{-1}\psi_n\|} \to 0\end{aligned}$$

since the denominator goes to ∞, as shown above.

For the converse, let $\mu \in \rho(T)$. We know $(T-\mu)^{-1}$ is bounded, so there exists some $M > 0$ such that for all $\varphi \in \mathcal{H}$ we have

$$\|(T-\mu)^{-1}\varphi\| \leq M\|\varphi\| \ .$$

Define $f_\varphi = (T-\mu)\varphi$ for all φ in the domain of T. Then for all such φ we have $(T-\mu)^{-1}f_\varphi = \varphi$. Therefore

$$\begin{aligned}\|(T-\mu)^{-1}f_\varphi\| &\leq M\|f_\varphi\| \\ \|\varphi\| &\leq M\|(T-\mu)\varphi\| \ .\end{aligned}$$

Hence any sequence $\|(T-\mu)\varphi_n\|$ with $\|\varphi_n\| = 1$ will not converge to 0. □

Chapter 2

Schrödinger Operators

The first chapter developed the necessary fundamentals to mathematically describe localization. In this chapter, we specify the operators we are using, and over what space. Specific calculations of the spectrum of our operators are included, as well as a key result about how one can approximate the spectrum from the generalized eigenvalues.

For the rest of the paper, we will be working over the Hilbert space

$$\ell^2(\mathbb{Z}^d) = \left\{ u : \mathbb{Z}^d \to \mathbb{C} \;\middle|\; \sum_{n\in\mathbb{Z}^d} |u(n)|^2 < \infty \right\} \tag{2.1}$$

with the inner product defined as

$$\langle u, v\rangle = \sum_{n\in\mathbb{Z}^d} \overline{u(n)} v(n) \; . \tag{2.2}$$

Then the resulting norm $\|u\|$ is given by

$$\langle u, u\rangle^{\frac{1}{2}} = \left(\sum_{n\in\mathbb{Z}^d} |u(n)|^2 \right)^{1/2} \; . \tag{2.3}$$

We will also make use of two norms on $\mathbb{Z}^d$, namely

$$\|n\|_\infty := \sup_{k=1,\ldots,d} |n_k|, \quad \|n\|_1 := \sum_{k=1}^{d} |n_k|$$

where n_k is the k^{th} coordinate point of n.

This setting is a deviation from what one would normally see during an introductory course in quantum mechanics. The space of interest is typically the square-integrable

functions $L^2(\mathbb{R})$. However, the space $\ell^2(\mathbb{Z}^d)$ does not differ drastically from the continuous case, and allows for many explicit computations. In addition, modeling a material as a discrete grid is quite reasonable, as many solids have some crystalline structure to them. Then, each atom or molecule lies in the vicinity of a point on $\mathbb{Z}^d$.

2.1 Discrete Laplacian

Typically, the (time-independent) Schrödinger equation is written in continuous space as

$$H\psi = E\psi, \quad H = -\nabla^2 + V \ . \tag{2.4}$$

This is of course after changing units appropriately so that the typical coefficient $\frac{\hbar^2}{2m} = 1$. The Laplacian $-\nabla^2$ is the kinetic energy operator, and V is a potential. In this section, we adapt the kinetic energy operator to our discrete space and explore some of its properties.

Definition 2.1 (Discrete Laplacian)**.** We define our kinetic energy operator H_0 to be the discrete Laplacian

$$(H_0\, u)(n) = -\sum_{\|m-n\|_1=1} (u(m) - u(n)) \ . \tag{2.5}$$

From the definition of the inner product over $\ell^2(\mathbb{Z}^d)$, we have

$$\langle u, H_0\, v\rangle = \frac{1}{2}\sum_{n\in\mathbb{Z}^d}\sum_{\|m-n\|_1=1} \overline{(u(n)-u(m))}(v(n)-v(m)) \ . \tag{2.6}$$

Due to the symmetry in this definition, we see H_0 is self-adjoint. In addition, we can show H_0 is bounded, allowing us to make use of the developed material in the previous chapter.

Lemma 2.2. *The discrete Laplacian H_0 is bounded.*

Proof. First, note that a point $n \in \mathbb{Z}^d$ has $2d$ neighbors. Then by definition

$$\|H_0 u\| = \left(\sum_{n\in\mathbb{Z}^d}\left(\sum_{\|m\|_1=1}(u(n+m)-u(n))\right)^2\right)^{\frac{1}{2}} \ . \tag{2.7}$$

Using the triangle inequality on the functions $\sum f_m(n) = \sum[u(n+m) - u(n)]$ gives

$$\begin{aligned}
\|H_0 u\| &\leq \sum_{\|m\|_1=1} \left(\sum_{n\in\mathbb{Z}^d} |u(n+m) - u(n)|^2 \right)^{\frac{1}{2}} \\
&\leq \sum_{\|m\|_1=1} \left(\sum_{n\in\mathbb{Z}^d} |u(n+m)| + |u(n)|^2 \right)^{\frac{1}{2}} \\
&\leq \sum_{\|m\|_1=1} \left[\left(\sum_{n\in\mathbb{Z}^d} |u(n+m)|^2 \right)^{\frac{1}{2}} + \left(\sum_{n\in\mathbb{Z}^d} |u(n)|^2 \right)^{\frac{1}{2}} \right] \\
&\leq 4d\|u\| \ .
\end{aligned}$$

Hence, H_0 is bounded by $4d$. □

We will now find $\sigma(H_0)$ in two parts. First, we will rigorously show the derivation in $\ell^2(\mathbb{Z})$, then briefly cover what small changes and technicalities occur in $\ell^2(\mathbb{Z}^d)$, $d > 1$. As was mentioned in the beginning of this chapter, it is often useful to find unitary operators that transform the operator of interest (H_0) into a multiplication operator. Then it is fairly simple to find the spectrum of this multiplication operator, which exactly coincides with the original operator. Our choice is to introduce an operator that works as a Fourier transform from $\ell^2(\mathbb{Z}^d)$ to $L^2([0,2\pi]^d)$. In arbitrary dimensions, this operator is defined as

$$\mathcal{F}u(k) = \hat{u}(k) = \frac{1}{(2\pi)^{d/2}} \sum_{n\in\mathbb{Z}^d} u(n) e^{-in\cdot k} \ . \tag{2.8}$$

Notice the dot product in the exponent, allowing $\hat{u}$ to be a well-defined function. Also, the factor out in front is omitted in [13], but due to common practice it is included here. The adjoint of $\mathcal{F}$ can be similarly defined as

$$\mathcal{F}^*\hat{u}(n) = u(n) = \frac{1}{(2\pi)^{d/2}} \int_0^{2\pi} \hat{u}(k) e^{ik\cdot n} dk \ . \tag{2.9}$$

Recalling the discussion at the beginning of this chapter, we ultimately want to be able to find the spectrum of $\mathcal{F}H_0\mathcal{F}^*$, and show it is the same as $\sigma(H_0)$. In order to do this, we must show $\mathcal{F}$ is a unitary operator.

Lemma 2.3. *$\mathcal{F}$ is unitary. That is, $\mathcal{F}\mathcal{F}^* = I_L$, $\mathcal{F}^*\mathcal{F} = I_\ell$.*

Proof. We can consider how $\mathcal{F}$ and $\mathcal{F}^*$ act on the basis vectors in their respective spaces. In particular, the domain of $\mathcal{F}$ is $\ell^2(\mathbb{Z}^d)$ and the domain of $\mathcal{F}^*$ is $L^2([0,2\pi]^d)$. We will show they map onto each other's domains.

Consider the standard orthonormal basis $[e_n]$ of $\ell^2(\mathbb{Z})$, where $e_n(m) = 1$ whenever $n = m$, and is 0 otherwise. Then

$$\mathcal{F}e_m(k) = \frac{1}{(\sqrt{2\pi})^d} \sum_{n\in\mathbb{Z}} e_m(n)e^{-in\cdot k} = \frac{1}{(\sqrt{2\pi})^d} e^{-im\cdot k} .$$

since every term is 0 other than $e_m(m)$. Then we have $\hat{u}(k) = e^{-im\cdot k}$, so

$$\mathcal{F}^*\hat{u}(n) = \mathcal{F}^*e^{-im\cdot n} = \frac{1}{(2\pi)^d} \int_0^{2\pi} e^{-im\cdot k} e^{ik\cdot n} dk .$$

We can either check the value of this integral in general, or recognize the two terms as mutually orthogonal for $n \neq m$, so the value of the integral is 1 if and only if $m = n$. Thus, $(F^*F)e_m = e_m$. Hence $F^*F = I_\ell$.
The proof is completely analogous in the $\mathcal{F}\mathcal{F}^*$ case, since the functions $e^{-im\cdot n}$ form a complete, orthonormal basis of $L^2([0, 2\pi]^d)$. □

Now that we know this transform is a unitary operator, we also know the spectrum of H_0 is the same as the spectrum of $\mathcal{F}H_0\mathcal{F}^*$. We can show this latter operator is actually quite useful and simple for calculating the spectrum.

Theorem 2.4. *For any function $\hat{u}(k) \in L^2([0, 2\pi]^d)$ we have $\mathcal{F}H_0\mathcal{F}^*\,\hat{u}(k) = h_0(k)\hat{u}(k)$ where*

$$h_0(k) = 2\sum_{j=1}^{d}(1 - \cos(k_j)) .$$

Proof. We prove this only for $d = 1$, since the proof idea extends without issue: the main part of the argument is most clearly illustrated when just dealing with $\mathbb{Z}$.
Consider the standard orthonormal basis $[e_n]$ of $\ell^2(\mathbb{Z})$, such that $e_n(m) = 1$ if an only if $n = m$, otherwise it is 0. For any $k \in \mathbb{Z}$, the only elements that are 1 away from k are $k - 1$ and $k + 1$. Thus we have

$$\begin{aligned} H_0e_n(k) &= -(e_n(k + 1) - e_n(k) + e_n(k - 1) - e_n(k)) \\ &= 2e_n(k) - e_n(k + 1) - e_n(k - 1) . \end{aligned}$$

Also, due to how this standard basis is set up, we see that $e_n(m) = \delta_{n,m}$. Thus, applying our transform we have

$$\mathcal{F}e_n(k) = \hat{e}_n(k) = \sum_{m\in\mathbb{Z}} \delta_{n,m}e^{-imk} = e^{-ink} .$$

Notice that the dot product becomes just a regular product, as we are in one dimension. Also, the $\delta_{n,m}$ is 0 for each part of the sum except for $m = n$. Now, we have

$$\begin{aligned}(\mathcal{F}H_0\mathcal{F}^*)\mathcal{F}e_n(k) &= \mathcal{F}H_0\mathcal{F}^*e^{-ink} = \mathcal{F}H_0e_n(k)\\ &= \mathcal{F}[2e_n(k) - e_n(k+1) - e_n(k-1)]\\ &= 2e^{-ink} - e^{-in(k+1)} - e^{-in(k-1)}\\ &= e^{-ink}\left(2 - (e^{-ik} + e^{ik})\right)\\ &= e^{-ink}(2 - 2\cos(k))\ .\end{aligned}$$

Thus, we see that applying $\mathcal{F}H_0\mathcal{F}^*$ to some function in $L^2([0, 2\pi])$ is equivalent to multiplying by $h_0(k) = 2 - 2\cos(k)$, as $\left[e^{-ink}\right]$ is an orthonormal basis for all such functions. □

Since the domain of $h_0(k)$ is open, and its range is just $[0, 4d]$, by Theorem 1.4 we have that

$$\sigma(H_0) = [0, 4d], \quad \text{in } \ell^2(\mathbb{Z}^d)\ . \tag{2.10}$$

2.2 Random Potential

The second operator of interest is a random potential operator, V_ω. For a given ω in the sample space we have an instantiation of the deterministic operator V, which acts as a multiplication operator by the random variable $V_\omega(n)$.

The following development of material in probability will be rather brief, as the reader is assumed to be familiar with the content. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the set of outcomes, $\mathcal{F}$ is the set of events, and $\mathbb{P}$ assigns elements of $\mathcal{F}$ to probabilities. Define a *random variable* X to be a real-valued measurable function on this space (i.e. $X^{-1}(A) \in \mathcal{F}$ for any Borel set A). We can define the *distribution* of X to be

$$P(A) = \mathbb{P}(X \in A), \quad A \text{ is a Borel set.} \tag{2.11}$$

We can define the *support* of P to be the set

$$\text{supp } P = \{x \in \mathbb{R} \mid P_0\big((x - \varepsilon, x + \varepsilon)\big) > 0 \text{ for all } \varepsilon > 0\}\ . \tag{2.12}$$

If there is a function g such that for any Borel set A

$$P_0(A) = \int_A g(\lambda)d\lambda$$

we call g the *density* of the random variable X.

Two random variables X and Y with distributions P_X, P_Y are called *identically distributed* if $P_X(A) = P_Y(A)$ for all A. A family $\{X_i\}_{i\in I}$ of random variables is *independent* if for any finite subset $\{i_1, i_2, \ldots, i_k\} \subset I$ we have

$$\begin{aligned}&\mathbb{P}(X_{i_1} \in [a_1, b_1], X_{i_2} \in [a_2, b_2], \ldots, X_{i_k} \in [a_k, b_k])\\ &= \mathbb{P}(X_{i_1} \in [a_1, b_1])\mathbb{P}(X_{i_2} \in [a_2, b_2]) \cdots \mathbb{P}(X_{i_k} \in [a_k, b_k]) \ .\end{aligned}$$

A corollary of this definition is that the distribution also multiplies through disjoint intervals, thus

$$\mathbb{P}(X_{i_1} \in [a_1, b_1], X_{i_2} \in [a_1, b_1], \ldots, X_{i_k} \in [a_k, b_k]) = P_0([a_1, b_1]) \cdots P_0([a_k, b_k]) \ .$$

There are quite a few typical results used throughout any probabilistic arguments. The main result needed in this paper is the Borel-Cantelli lemma, which will be used as below.

Lemma 2.5 (Borel-Cantelli)**.** *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of sets $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$, let A_∞ be the set*

$$A_\infty = \{\omega \in \Omega \mid \omega \in A_n \text{ for infinitely many } n\} \ . \tag{2.13}$$

Then the following are true:

1. *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A_\infty) = 0$.*
2. *If the A_n are all independent and $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(A_\infty) = 1$.*

Finally, we recall the notion of independence of events. Given some sequence of events $\{A_n\}$ in $\mathcal{F}$, we say they are *independent* if for any finite subsequence $\{A_{n_j}\}, j = 1, \ldots, k$

$$\mathbb{P}\left(\bigcap_{j=1}^{k} A_{n_j}\right) = \prod_{j=1}^{M} \mathbb{P}(A_{n_j}) \ . \tag{2.14}$$

There is one more technical result in probability which will be required to find the spectrum of the Hamiltonian. It says, as stated in [13], events which can happen will happen, and infinitely often. It requires the assumption supp P_0 is compact.

Lemma 2.6. *There exists a set Ω_0 with probability one such that for any $\omega \in \Omega_0$, any finite $\Lambda \subset \mathbb{Z}^d$, any sequence $\{q_i\}_{i\in\Lambda}$ with $q_i \in$ supp P_0 and any $\varepsilon > 0$, there exists a sequence $\{j_n\}$ in $\mathbb{Z}^d$, $\|j_n\|_\infty \to \infty$ such that*

$$\sup_{i\in\Lambda} |q_i - V_\omega(i + j_n)| < \varepsilon \ .$$

Proof. We must construct our Ω_0. Thus we fix a finite $\Lambda \subset \mathbb{Z}^d$, a sequence $\{q_i\}_{i\in\Lambda}$ with $q_i \in \text{supp } P_0$ and some $\varepsilon > 0$, as dictated by the hypothesis of the theorem. By the definition of supp P_0, and the fact that the q_i are independent we can say

$$\mathbb{P}\left(\text{Events } \omega \text{ such that } \sup_{i\in\Lambda} |V_\omega(i) - q_i| < \varepsilon\right) > 0\,. \tag{2.15}$$

Now we wish to find a sequence of events which are independent that allow us to employ this fact. Notice we are taking the supremum over Λ, so if we have points ℓ_n which are pairwise far enough away (more than twice the diameter of Λ in fact) we can make sure these points are always in disjoint "copies" of Λ, allowing the above event for $i + \ell_n$ to be independent. Thus we choose such a sequence $\ell_n \in \mathbb{Z}^d$ so that the distance between any ℓ_n, ℓ_m (with $n \neq m$) is more than twice the diameter of Λ. If we let A_n be the set of events ω such that $\sup_{i\in\Lambda} |V_\omega(i + \ell_n) - q_i| < \varepsilon$, then we see the A_n are independent with $\mathbb{P}(A_n) > 0$, since this is the same probability as (2.15).

We employ the Borel-Cantelli lemma (2.5). Since each A_n is independent with non-zero probability, their probabilities do not converge to 0 as $n \to \infty$, hence $\sum A_n$ must diverge to ∞. Thus, the set

$$\Omega_{\Lambda,\{q_i\},\varepsilon} = \{\omega \mid \omega \in A_n \text{ for infinitely many } n\} \tag{2.16}$$

has probability 1. Notice we can vary our finite set Λ, our sequence $\{q_i\}$ and our $\varepsilon > 0$ and still have this Ω be of probability one. In addition, a countable intersection of sets of probability one also has probability one. With this, we are prepared to define Ω_0. Since supp P_0 is compact, it has a countable dense set R_0. In addition the set Γ of all finite subsets of $\mathbb{Z}^d$ is countable. Thus we can define a countable intersection of the above Ω over this countable set R_0, the finite sets Γ, and $\varepsilon = 1/n$ for $n \in \mathbb{N}$. Each of these components is countable. In particular we set

$$\Omega_0 := \bigcap_{\substack{\Lambda \in \Gamma \\ \{q_i\} \in R_0,\, n \in \mathbb{N}}} \Omega_{\Lambda,\{q_i\},\frac{1}{n}}\,. \tag{2.17}$$

As discussed above, for each $\omega \in \Omega_0$, given any finite Λ, a sequence $\{q_i\}$ and a ε, we constructed our sequence ℓ_n. Thus Ω_0 satisfies the conditions of the lemma. □

To fully understand the intuition behind this lemma, consider a cube Λ in $\mathbb{Z}^d$, and suppose at each point n, the IID random variables $V_\omega(n) \in \{0, 1\}$. Thus our Λ has a sequence $\{q_i\}$ of 0s and 1s. Now, consider splitting $\mathbb{Z}^d$ into disjoint cubes of the same size. By the pigeonhole principle, there exists some Λ_k which has the exact same configuration of 0s and 1s. In fact, there are infinitely many. Take the center points of

these to be our j_n, and we have constructed our sequence. If this setup is true for all ω, we have easily found our set Ω_0 with probability one.

Now we are prepared to discuss our random potential operators.

Definition 2.7 (Random potential)**.** V_ω is a random operator, defined by

$$(V_\omega \varphi)(n) = V_\omega(n)\varphi(n), \quad n \in \mathbb{Z}^d \tag{2.18}$$

where the $V_\omega(n)$ are independent and identically distributed (IID) random variables with common distribution P_0.

Notice for a particular $\omega \in \Omega$, our V_ω just becomes an instantiation of the deterministic multiplication operator V.

There are a few properties of V_ω that are good to state here for technical completeness, though they are not illuminating to prove. More details can be found in [17]. First, if P_0 has compact support, then V_ω is a bounded operator. The specific bound given is that if supp $P_0 \subset [-M, M]$, then $\mathbb{P}$-almost surely

$$\sup_{j \in \mathbb{Z}^d} |V_\omega(j)| \leq M \ . \tag{2.19}$$

In addition, if supp P_0 is not compact, V_ω is self-adjoint on the space

$$\mathcal{V} = \{\varphi \in \ell^2(\mathbb{Z}^d) \mid V_\omega \varphi \in \ell^2(\mathbb{Z}^d)\} \ .$$

V_ω is also essentially self-adjoint on the space

$$\ell_0^2(\mathbb{Z}^d) = \{\varphi \in \ell^2(\mathbb{Z}^d) \mid \varphi(i) \neq 0 \text{ for finitely many points } i\} \ . \tag{2.20}$$

This fact will be important, as it will help guarantee our Hamiltonian operator is essentially self-adjoint on a useful space.

Finally, it is easy to determine the spectrum of V_ω.

Lemma 2.8. *Let V_ω by a multiplication operator with function $V_\omega(n)$, which has distribution P_0. Then $\mathbb{P}$-almost surely $\sigma(V_\omega) =$ supp P_0.*

Proof. By Theorem 1.4, we know $\sigma(V_\omega)$ is the closure of im $V_\omega(n)$. However, we can consider the definition of the support. $V_\omega(n)$ maps to elements x in the domain of P_0, which do not map to 0 probability, so that the image of $V_\omega(n)$ is essentially $\mathbb{R} - \ker P_0$. The closure is then all x which have a non-zero neighborhood around them. This is precisely the support. Hence $\sigma(V_\omega) =$ supp P_0. □

2.3 Hamiltonian

We can now describe the discrete Hamiltonian and explore its properties.

Definition 2.9 (Discrete Hamiltonian)**.** The discrete Hamiltonian is $H = H_0 + V$, where H_0 is the discrete Laplacian (2.5) and V is the potential multiplication operator (2.18). The random Hamiltonian is then $H_\omega = H_0 + V_\omega$.

Using the results of Sections 2.1 and 2.2, we can show the spectrum of H_ω is $\mathbb{P}$-almost surely a fixed set, namely $[0, 4d] + \text{supp } P_0$. While the measure-theoretic structure of the spectrum will ultimately be of interest, it is still a good exercise to determine the spectrum as a set.

Theorem 2.10. *For $\mathbb{P}$-almost all ω we have $\sigma(H_\omega) = [0, 4d] + \text{supp } P_0$.*

Proof. We have already seen $\sigma(V_\omega) = \text{supp } P_0$, for $\mathbb{P}$-almost all ω. We also have

$$0 \leq \text{dist}[\sigma(H_\omega), \sigma(V_\omega)]] \leq \|H_0\|$$

since whenever μ is such that $\text{dist}(\mu, \sigma(V_\omega)) > \|H_0\|$ we have $\mu \in \rho(H_\omega)$ (see [12] for more details.) This implies

$$\begin{aligned} \sigma(H_0 + V_\omega) &\subset [0, \|H_0\|] + \sigma(V_\omega) \\ &= [0, 4d] + \text{supp } P_0 \, . \end{aligned}$$

To show the reverse inclusion we use the Weyl Criterion (1.17). Let $\lambda \in [0, 4d] + \text{supp } P_0$. In particular set $\lambda = \lambda_0 + \lambda_1$ where $\lambda_0 \in [0, 4d]$ and $\lambda_1 \in \text{supp } P_0$. By definition there exists a Weyl sequence φ_n for λ_0 such that

$$\|(H_0 - \lambda_0)\varphi_n\| \to 0, \quad \|\varphi_n\| = 1 \, .$$

Since H_0 is essentially self-adjoint on $\ell_0^2(\mathbb{Z}^d)$ and is in fact bounded, we may suppose $\varphi_n \in \ell_0^2(\mathbb{Z}^d)$. (Note this space is taking the place of $\mathcal{F}$ in the Weyl criterion.) In particular, this tells us $|\text{supp } \varphi_n| < \infty$ for each n. If we set $\varphi^{(j)}(i) = \varphi(i - j)$ for any fixed j, we can see

$$\begin{aligned} (H_0\varphi)^{(j)}(i) &= (H_0\varphi)(i - j) \\ &= - \sum_{\|k-(i-j)\|_1 = 1} [\varphi(i - j) - \varphi(k - j)] \\ &= H_0\varphi^{(j)}(i) \end{aligned}$$

So $H_0\varphi^{(j)} = (H_0\varphi)^{(j)}$. Since we have our finite sets supp φ_n, a (constant) sequence $\{\lambda_1\}$ in supp P_0, and setting $\varepsilon = 1/n$, we can employ Lemma 2.6 to show there is a sequence $\{j_n\}, j_n \to \infty$ with probability one such that

$$\sup_{i \in \text{supp } \varphi_n} |V_\omega(i + j_n) - \lambda_1| < \frac{1}{n} \,. \tag{2.21}$$

We claim $\psi_n = \varphi_n^{(j_n)}$, so that $\psi_n(i) = \varphi_n(i - j_n)$, is a Weyl sequence for H_ω and λ. We can calculate

$$\begin{aligned}
\|(H_\omega - \lambda)\psi_n\| &= \|(H_0 - \lambda_0 + V_\omega - \lambda_1)\varphi_n^{(j_n)}\| \\
&\leq \|(H_0 - \lambda_0)\varphi_n^{(j_n)}\| + \|(V_\omega - \lambda_1)\varphi_n^{(j_n)}\| \\
&= \|(H_0 - \lambda_0)\varphi_n\| + \sup_{i \in \text{supp } \varphi_n^{(j_n)}} |V_\omega \varphi_n^{(j_n)}(i) - \lambda_1 \varphi_n^{(j_n)}(i)| \\
&\leq \|(H_0 - \lambda_0)\varphi_n\| + \sup_{i \in \text{supp } \varphi_n^{(j_n)}} |V_\omega(i - j_n) - \lambda_1|
\end{aligned}$$

where the final inequality comes from the fact that $\|\varphi_n\| = 1$. However, φ_n is a Weyl sequence for H_0 and λ_0, so the first term goes to 0. Similarly, by our assumption on the sequence j_n the second term also goes to 0. Hence $\|(H_\omega - \lambda)\psi_n\| \to 0$ as $n \to \infty$. Thus we have constructed a Weyl sequence for H_ω and λ, implying that $\lambda \in \sigma(H_\omega)$. In particular this shows $[0, 4d] + \text{supp } P_0 \subset \sigma(H_\omega)$. In total, $\sigma(H_\omega) = [0, 4d] + \text{supp } P_0$, $\mathbb{P}$-almost surely. □

Finally, we introduce the idea of a generalized eigenvector, which will help us approximate the spectrum in a very useful way. These have generalized eigenfunctions which are bounded, but not necessarily strictly enough to exist in $\ell^2(\mathbb{Z}^d)$.

Definition 2.11 (Polynomially bounded)**.** A function φ is said to be polynomially bounded if there exist constants $C, r > 0$ such that $\varphi(n) \leq C\|n\|_\infty^r$.

Definition 2.12 (Generalized eigenvalue)**.** If there is a polynomially bounded φ such that $H\varphi = \lambda\varphi$, we say λ is a generalized eigenvalue of H. The set of generalized eigenvalues is denoted $\varepsilon_g(H)$.

The following theorem is our approximation of the spectrum $\sigma(H)$ by the generalized eigenvalues. It will be an important final feature in our proof of localization later on.

Theorem 2.13. *The spectrum of the discrete Hamiltonian H agrees up to a set of spectral measure zero with $\varepsilon_g(H)$. In particular, $\sigma(H) = \overline{\varepsilon_g(H)}$, the closure of the set of generalized eigenvalues.*

Proof. See the Appendix for a full proof. □

Chapter 3

Localization

With the mathematical tools developed, and the operators of interest at hand, we can now discuss localization properly. Some simple results related to localization are shown first. Then the important Green's functions, which serve as a proxy for eigenfunction decay, are introduced. Finally, a few results are shown using the Green's functions which are necessary to ultimately prove the existence of localization.

3.1 Statement of Results

We are now in a position to state the main results of interest. These will help focus the discussion moving forward. We will follow the assumptions listed in [13], section 8.3. Namely:

1. H_0 is the discrete Laplacian on $\ell^2(\mathbb{Z}^d)$, defined in (2.5).
2. $V_\omega(i)$, $i \in \mathbb{Z}^d$ are IID random variables with a common distribution P_0.
3. P_0 has a bounded density g, so that $P_0(A) = \int_A g(\lambda)\, d\lambda$ and $\|g\|_\infty < \infty$.
4. supp P_0 is compact.

Under these assumptions, localization can be defined by the measure-theoretic structure of the spectrum.

Definition 3.1 (Spectral Localization)**.** Let H_ω be the random Hamiltonian, and I an energy interval with $I \cap \sigma(H_\omega) \neq \emptyset$. Then H_ω exhibits spectral localization if for $\mathbb{P}$-almost all ω

$$\sigma_c(H_\omega) \cap I = \emptyset \ .$$

This describes the importance of our previous partitioning of the spectrum into its various components. The "RAGE-theorem" (as described in [13] section 7.3) tells us more specifically how the components of the spectrum interplay with localization. These results are summarized below, although the proofs are well beyond the scope of this report.

Theorem 3.2. *Let $H\varphi = E\varphi$ for $\varphi \in \ell^2(\mathbb{Z}^d)$. Then the spectral measure μ induced by H has an atom at E, and μ_φ is a pure point measure with a concentration at E.*

Theorem 3.3. *Any eigenfunction of E belongs to the pure point space $\ell^2(\mathbb{Z}^d)_{pp}$. Moreover, $\ell^2(\mathbb{Z}^d)$ is the closure of the linear span of all eigenfunctions φ.*

In fact, some books use "eigenfunctions span the space" to be the definition of pure point spectrum. Hence we see that the set of all eigenvalues, denoted $\varepsilon(H)$, is a dense subset of $\sigma_{pp}(H)$. In addition, $\varepsilon(H)$ is a countable set. This is due to $\ell^2(\mathbb{Z}^d)$ being a separable Hilbert space. Thus as we move forward, showing the spectrum consists only of eigenvalues is sufficient to prove localization.

The two main results describe how large disorder in a system, or low energy, are both sufficient to cause localization. Note that the specifics of the low energy case are not dealt with in this paper. The interested reader should consult [13] for a rigorous treatment. Now, we must be precise with how we define disorder in a system.

Definition 3.4 (Disorder)**.** The disorder of a density g is $\delta(g) := \|g\|_\infty^{-1}$.

Intuitively, disorder is the spread of the density. If the density is very spread out, then its norm will be small, causing high disorder.

Theorem 3.5. *There exists $E_1 > E_0 := \inf(\sigma(H_\omega))$ such that the spectrum of H_ω exhibits spectral localization in the energy interval $I = [E_0, E_1]$. Specifically, $\sigma(H_\omega) \cap I$ is pure point almost surely and the corresponding eigenfunctions decay exponentially.*

Theorem 3.6. *For any interval $I \neq \emptyset$, there is $C > 0$ such that for any disorder $\delta(g) \geq C$ the operator H_ω exhibits spectral localization in I. Specifically, $\sigma(H_\omega) \cap I$ is pure point almost surely and the corresponding eigenfunctions decay exponentially.*

When describing localization in $\ell^2(\mathbb{Z}^d)$, the natural object to consider is a discrete cube. Thus, we make clear our notation before moving forward.

Definition 3.7 (Cube)**.** We denote the discrete cube centered at the origin with side-length $2L + 1$ by Λ_L. The same cube centered at n is denoted $\Lambda_L(n)$.

3.2 Bounded and Scattering States

Localization describes how a wave stops propagating in space. A weaker version of this is a bounded state, which only describes the probability a particle remains in some large finite area, without the associated exponential decay of eigenfunctions. We consider this, and related concepts, before delving into localization.

As was mentioned in section 1.1, the spectral calculus developed allows us to look at operators such as e^{-itH}. Consider $\psi \in \mathcal{H}_{pp}$, so that $e^{-itH}\psi(x)$ is periodic in t. If the analogy we have been making is to hold true, there should be some decay or boundedness on this function. Indeed, the following theorem tells us that $\mathcal{H}_{pp}$ contains functions which correspond to bounded states of particles. Particularly, for any $\varepsilon > 0$ there is cube such that for any time, the particle will be inside Λ_L with probability $1-\varepsilon$, hence the particle escapes to infinity with vanishing probability.

Theorem 3.8. *Let H be a self-adjoint operator on $\ell^2(\mathbb{Z}^d)$, with $\psi \in \mathcal{H}_{pp}$ and Λ_L the cube centered at the origin with side length $2L+1$. Then*

$$\lim_{L\to\infty} \sup_{t\geq 0} \left(\sum_{x\in\Lambda_L} |e^{-itH}\psi(x)|^2 \right) = \|\psi\|^2$$

$$\lim_{L\to\infty} \sup_{t\geq 0} \left(\sum_{x\notin\Lambda_L} |e^{-itH}\psi(x)|^2 \right) = 0\,.$$

Proof. Note that if ψ is an eigenfunction then, as discussed before, $e^{-itH}\psi$ is periodic in time, yet it is a standing wave physically. In particular, $|e^{-itH}\psi|^2$ is independent of time. To be explicit, let $\Psi(n,t) = e^{-itH}\psi(n)$. Then

$$\begin{aligned} |\Psi(n,t)|^2 &= |e^{-itH}\psi(n)|^2 \\ &= |e^{-itE}\psi(n)|^2 \\ &= |\psi(n)|^2 \end{aligned}$$

since $|e^{-itE}| = 1$. Hence $|\Psi(n,t)|^2$ is independent of time, as claimed. Thus a particle starting in an eigenstate remains there for all time. So the theorem is true for eigenfunctions.

We know e^{-itH} is unitary, meaning that $e^{-itH}e^{itH} = e^{itH}e^{-itH} = I$. This implies that for all t and any $\Lambda \subset \mathbb{Z}^d$

$$\begin{aligned} \|\psi\|^2 &= \|e^{-itH}\psi\|^2 \\ &= \sum_{x\in\Lambda} |e^{-itH}\psi(x)|^2 + \sum_{x\notin\Lambda} |e^{-itH}\psi(x)|^2\,. \end{aligned}$$

In particular, if we can show one of the equations above, the other follows. We will prove the second equation.

We let P_L be the projection onto $\Lambda_L^{\complement}$, the complement of Λ_L. Then we wish to prove $\|P_L e^{-itH}\psi\| \to 0$ uniformly in t as $L \to \infty$. For any finite linear combination of eigenfunctions

$$\psi = \sum_{k=1}^{M} \alpha_k \psi_k, \quad H\psi_k = E_k \psi_k$$

we can apply the Cauchy-Schwarz inequality and the definition of an eigenfunction to show

$$\begin{aligned}
\|P_L e^{-itH}\psi\| &= \left\| \sum_{k=1}^{M} \alpha_k P_L e^{-itH} \psi_k \right\| \\
&\leq \sum_{k=1}^{M} |\alpha_k| \ \|P_L e^{-itH}\psi_k\| \\
&= \sum_{k=1}^{M} |\alpha_k| \ \|P_L e^{-itE_k}\psi_k\|] \\
&= \sum_{k=1}^{M} |\alpha_k| \ \|P_L \psi_k\|
\end{aligned}$$

since $|e^{-itE_k}| = 1$. As L increases we are evaluating $\|P_L\psi_k\|$ on a smaller set. Hence for any $\varepsilon > 0$ we can take L large enough so that each term in the sum is smaller than

$$\varepsilon \left(\sum_{k=1}^{M} \alpha_k \right)^{-1}$$

so that the entire sum is less than ε. Let $\psi \in \mathcal{H}_{pp}$ be arbitrary. From our prior discussion of how eigenfunctions are dense in $\mathcal{H}_{pp}$ we can find a linear combination of eigenfunctions $\psi^{(M)} = \sum_{k=1}^{M} \alpha_k \psi_k$ such that $\|\psi - \psi^{(M)}\| < \varepsilon$. Then we can write

$$\begin{aligned}
\|P_L e^{-itH}\psi\| &= \|P_L e^{-itH}(\psi + \psi^{(M)} - \psi^{(M)})\| \\
&\leq \|P_L e^{-itH}\psi^{(M)}\| + \|P_L e^{-itH}(\psi - \psi^{(M)})\| \\
&\leq \|P_L e^{-itH}\psi^{(M)}\| + \|\psi - \psi^{(M)}\| \ .
\end{aligned}$$

By assumption we can take M large enough so the right term is arbitrarily small. As discussed in our manipulation of the sums above, we can also take large enough L to make the left term arbitrarily small. Hence as $L \to \infty$ we indeed get $\|P_L e^{-itH}\psi\| \to 0$. $\square$

There is an opposite result for absolutely continuous spectra, which correspond to the scattering states of a particle.

Theorem 3.9. *Let H be a self-adjoint operator on $\ell^2(\mathbb{Z}^d)$, with $\psi \in \mathcal{H}_{ac}$ and Λ a finite subset of $\mathbb{Z}^d$. Then*

$$\lim_{t\to\infty} \left(\sum_{x\in\Lambda} |e^{-itH}\psi(x)|^2 \right) = 0 \tag{3.1}$$

$$\lim_{t\to\infty} \left(\sum_{x\notin\Lambda} |e^{-itH}\psi(x)|^2 \right) = \|\psi\|^2 \,. \tag{3.2}$$

Proof. Let $\psi \in \mathcal{H}_{ac}$. Then μ_ψ is absolutely continuous by definition. From the previous bound 1.26, we find that $\mu_{\varphi,\psi}$ is absolutely continuous as well for any $\varphi \in \mathcal{H}$. Then, there is some density h with respect to Lebesgue measure compatible with $\mu_{\varphi,\psi}$. Thus we can write

$$\begin{aligned} \langle \varphi, e^{-itH}\psi \rangle &= \int e^{-it\lambda}\, d\mu_{\varphi,\psi}(\lambda) \\ &= \int e^{-it\lambda} h(\lambda)\, d\lambda \,. \end{aligned}$$

Notice that the second expression is just the L^1 Fourier transform of h, so it converges to 0 as t goes to infinity. Now, consider $\mathcal{H} = \ell^2(\mathbb{Z}^d)$ with $\varphi = \delta_x, x \in \mathbb{Z}^d$. Then for each $x \in \Lambda$, which is a finite set, we can write

$$\begin{aligned} e^{-itH}\psi(x) &= \langle \delta_x, e^{-itH}\psi \rangle \\ &= \int e^{-itH} h(\lambda)\, d\lambda \to 0 \,. \end{aligned}$$

Thus the first sum over Λ is a finite sum of terms which converge to zero as $t \to \infty$. Hence, the sum itself converges to 0 as $t \to \infty$. □

3.3 Green's Functions

A tool which will be used frequently to show localization are the Green's functions. First, we must discuss the sets and domains on which the functions will be defined. Let $\Lambda_L(n)$ be the cube centered at $n \in \mathbb{Z}^d$ of side length $2L+1$. We define the inner and outer boundaries of $\Lambda_L(n)$ in the natural way:

$$\partial^- \Lambda_L(n) = \{m \in \mathbb{Z}^d \mid \|m-n\|_\infty = L\} \tag{3.3}$$

$$\partial^+ \Lambda_L(n) = \{m \in \mathbb{Z}^d \mid \|m-n\|_\infty = L+1\} \,. \tag{3.4}$$

Hence $\partial^- \Lambda_L(n) \cap \Lambda_L(n) \neq \emptyset$ and $\partial^+ \lambda_L(n) \cap \Lambda_L(n) = \emptyset$. We also define a boundary as the edges between the inner and outer boundaries:

$$\partial \Lambda_L(n) = \{(n, m) \mid n \in \partial^- \Lambda_L(n),\ m \in \partial^+ \Lambda_L(n)\} .$$

Let $|A|$ be the number of lattice points inside some $A \subset \mathbb{Z}^d$. Also define $A \nearrow \mathbb{Z}^d$ to mean an ascending chain $A_m \subset A_{m+1} \subset \mathbb{Z}^d$ such that $\bigcup A_m = \mathbb{Z}^d$.

We require a notion of restricting our operator to some subset of $\mathbb{Z}^d$. We follow the approach given in [13], section 5.2.

Definition 3.10 (Restricted Laplacian)**.** Given some $\Lambda \subset \mathbb{Z}^d$, we define the operator $(H_0)_\Lambda$ on the space $\ell^2(\Lambda)$ by

$$(H_0)_\Lambda(n, m) = \langle \delta_n, H_0 \delta_m \rangle \tag{3.5}$$

whenever $n, m \in \Lambda$. We do not consider $n, m \notin \Lambda$.

Then we define $H_\Lambda = (H_0)_\Lambda + V$. Thus, if $E \in \sigma(H_\Lambda)$, we have a finite-dimensional matrix so that E is an eigenvalue, thus there is an eigenfunction ψ which need only solve the equation $H\psi = E\psi$ over Λ.

We can now define the Green's functions.

Definition 3.11 (Green's Function)**.** Given $\Lambda \subset \mathbb{Z}^d$ and an energy $E \notin \sigma(H_\Lambda)$, define the Green's function to be the kernel of the resolvent $(H_\Lambda - E)^{-1}$, namely

$$G_E^\Lambda(n, m) = (H_\Lambda - E)^{-1}(n, m) = \langle \delta_n, (H_\Lambda - E)^{-1} \delta_m \rangle .$$

The idea of the Green's functions is to use it as a proxy for the decay of eigenfunctions, and thus to help show localization. The main effort in [13] and [4] goes to showing this equivalence under certain conditions. We will use the definitions from [13] to provide a vocabulary for the necessary tools. Assume V is a fixed potential operator.

Definition 3.12 (Exponential Decay)**.** A Green's function $G_E^{\Lambda_L(n_0)}$ decays exponentially on $\Lambda_L(n_0)$ with rate $\gamma > 0$ if E is not an eigenvalue of $H_{\Lambda_L(n_0)} = (H_0 + V)_{\Lambda_L(n_0)}$ and

$$\left| G_E^{\Lambda_L(n_0)}(n, m) \right| = |(H_{\Lambda_L(n_0)} - E)^{-1}(n, m)| \leq e^{-\gamma L} \tag{3.6}$$

for all $n \in \Lambda_{\sqrt{L}}(n_0)$ and $m \in \partial^- \Lambda_L(n_0)$.

Definition 3.13 ((γ, E)-good Cube)**.** A cube $\Lambda_L(n_0)$ is (γ, E)-good for V if the Green's function $G_E^{\Lambda_L(n_0)}$ decays exponentially with rate $\gamma > 0$.

Definition 3.14 (γ-good Energy)**.** An energy E is γ-good for V if there is a sequence of cubes $\Lambda_{\ell_m} \nearrow \mathbb{Z}^d$ such that all Λ_{ℓ_i} are (γ, E)-good for a fixed γ.

Let us consider how these Green's functions help. Suppose ψ solves $H\psi = E\psi$. Define

$$\Gamma_\Lambda(n,m) = \begin{cases} -1 & \text{if } (n,m) \in \partial\Lambda, \\ 0 & \text{otherwise .} \end{cases}$$

If we let $\Lambda^\complement$ be the complement $\mathbb{Z}^d - \Lambda$, we can then write

$$\langle \delta_n, H\delta_m \rangle = H_\Lambda \oplus H_{\Lambda^\complement} + \Gamma_\Lambda \tag{3.7}$$

where $H_\Lambda \oplus H_{\Lambda^\complement}$ is 0 for $(n,m) \in \partial\Lambda = \partial\Lambda^\complement$, and evaluates accordingly based on whether $(n,m) \in \Lambda$ or $(n,m) \in \Lambda^\complement$. So, subtracting E from both sides of (3.7), then taking $n_0 \in \Lambda$ so that $H_{\Lambda^\complement}$ disappears, we have

$$(H_\Lambda - E)\psi(n_0) = (-\Gamma_\Lambda \psi)(n_0) \ . \tag{3.8}$$

(See [13] section 9.1 and 5.2 for a few more details.) If we assume that E is not an eigenvalue, we can invert $H_\Lambda - E$ to obtain

$$\psi(n_0) = -[(H_\Lambda - E)^{-1}\Gamma_\Lambda \psi](n_0) \ . \tag{3.9}$$

By our definition of Γ_Λ, we need only look at the boundary to evaluate this. Given the inner-product definition of the Green's function, we can explicitly write $\psi(n_0)$ as a sum (much in the way you can multiply a function by $\delta(n_0)$ and integrate to obtain the function term):

$$\psi(n_0) = -\sum_{\substack{k \in \partial^- \Lambda \\ m \in \partial^+ \Lambda}} G_E^\Lambda(n_0, k)\psi(m) \ . \tag{3.10}$$

The main result that comes from this is that the Green's functions interplay with the generalized eigenvalues in a useful way.

Theorem 3.15. *If E is γ-good for V, then E is not a generalized eigenvalue of $H = H_0 + V$.*

Proof. Suppose ψ is a polynomially bounded solution to $H\psi = E\psi$, so that

$$|\psi(m)| \leq c|m|^r \quad \text{for } m \neq 0 \ .$$

Let $n \in \mathbb{Z}^d$ and $\Lambda_{L_k}(n_0)$ be the sequence of cubes on which E is γ-good. Choose k large enough so that $n \in \Lambda_{\sqrt{L_k}}(n_0)$. Then, we use our equality in (3.10) to write

$$|\psi(n)| \leq \left| \sum_{\substack{m' \in \partial^- \Lambda_k \\ m \in \partial^+ \Lambda_k}} G_E^{\Lambda_{L_k}}(n, m')\psi(m) \right| \tag{3.11}$$

$$\leq \sup_{m \in \partial\Lambda} |\psi(m)| \sum_{m' \in \partial^- \Lambda_k} \left| G_E^{\Lambda_{L_k}}(n, m') \right| . \tag{3.12}$$

By assumption $|\psi(m)|$ is polynomially bounded and $G_E^{\Lambda_{L_k}}$ is decaying exponentially, while the sum has terms on the order of L_k^{d-1}, since that is the order of the size of $\partial\Lambda$. Also, $L_k^{d-1} \geq |m| = L_k$. Hence there exists some constants c_1, c_2 such that

$$|\psi(n)| \leq c_1 L_k^{d-1} \, e^{-\gamma L_k} \cdot c|m|^r \tag{3.13}$$

$$\leq c_2 L_k^{d-1+r} \, e^{-\gamma L_k} . \tag{3.14}$$

This decays exponentially to 0 as $k \to \infty$, since $e^{-\gamma L_k}$ exceeds the polynomial growth of L_k^{d-1+r}. Hence $|\psi(n)| = 0$, so $\psi \equiv 0$ if it is a solution to $H\psi = E\psi$. Thus E is not a generalized eigenvalue. □

The proof of Green's function decay implying localization is one by induction, where the key step is iterating the estimate (3.13).

There are two quick corollaries, whose proofs come from the fact that $\sigma(H) = \overline{\varepsilon_g(H)}$.

Lemma 3.16. *If every $E \in [E_1, E_2]$ is γ-good for V then $\sigma(H_0 + V) \cap (E_1, E_2) = \emptyset$.*

Lemma 3.17. *If all $E \in [E_1, E_2]$ are γ-good for V, except a subset of Lebesgue measure 0, then $\sigma_{ac}(H_0 + V) \cap (E_1, E_2) = \emptyset$.*

Thus the Green's functions, and the related concepts of a (γ, E)-good cube are powerful in showing localization occurs.

3.4 Weak Results

There are two results which closely hint at localization and the exponential decay of eigenfunctions. The first proof shows how one can iterate the estimate (3.13), which will be key in the next chapter.

Theorem 3.15 implies a relationship between $\sigma(H)$ and whether (γ, E)-good cubes exist throughout $\mathbb{Z}^d$. The next two results heavily use aspects of this theorem. We can

iterate the estimate (3.13) to obtain a bound on the growth of approximate eigenfunctions. Then we begin a probabilistic analysis of $\sigma(H_\omega)$. First, let $\Lambda \subset \mathbb{Z}^d$ be some set. We define the *collection of cubes well-inside* Λ by

$$C_L(\Lambda) = \{\Lambda_L(n) \mid \Lambda_L(n) \subset \Lambda \text{ and } \Lambda_L(n) \cap \partial\Lambda = \emptyset\} \,. \tag{3.15}$$

In addition, we define an inner boundary of width L by

$$\partial_L^-(\Lambda) = \{m \in \Lambda \mid \operatorname{dist}(m, \partial^-\Lambda) \leq L\}, \quad \operatorname{dist}(m, A) = \inf_{k \in A} \|m - k\|_\infty \,. \tag{3.16}$$

The following theorem roughly says that regions filled with (γ, E)-good cubes are also regions where eigenfunctions exponentially decay. Physically, this reduces the probability of quantum tunneling to be exponentially small.

Theorem 3.18. *Let $A \subset \mathbb{Z}^d$ be finite. Suppose each cube in $C_M(A)$ is (γ, E)-good and M is large enough. For any integer $k > 0$, if ψ is a solution of $H\psi = E\psi$ in A, and $n_0 \in A$ such that*

$$\operatorname{dist}(n_0, \partial^- A) \geq k(M+1) \tag{3.17}$$

then

$$|\psi(n_0)| \leq e^{-\gamma' kM} \sup_{m \in \partial_M^- A} |\psi(M)| \,. \tag{3.18}$$

Proof. First, for notational ease we set the constants

$$r = 2d(2M+1)^{d-1} e^{-\gamma M} \tag{3.19}$$

$$\gamma' = \gamma - \frac{1}{M} \ln\left(2d(2M+1)^{d-1}\right) \,. \tag{3.20}$$

Then we have $r = e^{-\gamma' M}$. In addition, it is helpful to note $|\partial\Lambda_M| = 2d(2M+1)^{d-1}$.

By assumption, $k > 0$. We have $n_0 \in A$ and $\operatorname{dist}(n_0, \partial^- A) \geq (M+1)$. This implies that $\Lambda_M(n_0) \in C_M(A)$. Hence we apply estimate (3.10) to find

$$\begin{aligned}
|\psi(n_0)| &\leq \sum_{\substack{q \,\in\, \partial^+\Lambda_M(n_0) \\ q' \,\in\, \partial\Lambda_M(n_0)}} \left| G_E^{\Lambda_M(n_0)}(n_0, q) \right| |\psi(q')| \\
&\leq |\partial\Lambda_M(n_0)| e^{-\gamma M} \sup_{q' \in \partial^+\Lambda_M(n_0)} |\psi(q')| \\
&\leq 2d(2M+1)^{d-1} e^{-\gamma M} |\psi(n_1)| \\
&= r|\psi(n_1)|
\end{aligned}$$

for some $n_1 \in \partial^+\Lambda_M(n_0)$. Now, if $n_1 \in \partial_M^- A$, that is if n_1 is within M of the boundary of A, then n_1 does not satisfy the hypotheses of the theorem. This would imply we have

our bound 3.18 from the theorem statement for $k = 1$. In fact, by the triangle equality we have

$$\operatorname{dist}(n_1, \partial_M^- A) \geq \operatorname{dist}(n_0, \partial_M^- A) - (M+1) \, . \tag{3.21}$$

Thus $n_1 \in \partial_M^- A$ only if $k = 1$. Otherwise, we have $n_1 \notin \partial_M^- A$ and we can iterate again as $\Lambda_M(n_1) \in C_M(A)$. We find $n_2 \in \partial^+ \Lambda_M(n_1)$ such that

$$|\psi(n_1)| \leq r|\psi(n_2)| \tag{3.22}$$

$$|\psi(n_0)| \leq r^2|\psi(n_2)| \, . \tag{3.23}$$

The image below illustrates this procedure. The solid lines represent inner boundaries, while the dashed lines represent outer boundaries.

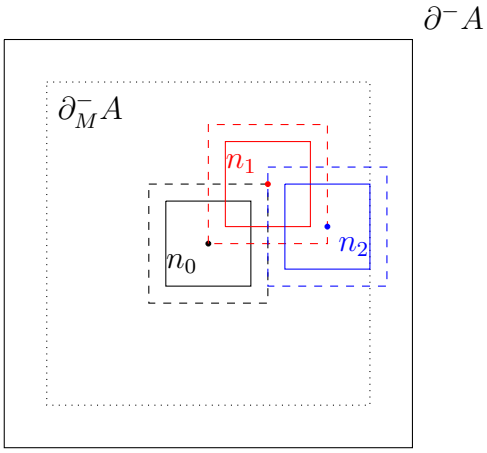

Figure 3.1: The iteration procedure. Each new point n_i is chosen on the outer boundary of the cube $\Lambda_M(n_{i-1})$, denoted by a dashed line. We choose points until a cube intersects $\partial_M^- A$, denoted by the dotted line.

Again using the triangle inequality, we get

$$\begin{aligned} \operatorname{dist}(n_2, \partial_M^- A) &\geq \operatorname{dist}(n_1, \partial_M^- A) - (M+1) \\ &\geq \operatorname{dist}(n_0, \partial_M^- A) - 2(M+1) \, . \end{aligned}$$

So $n_2 \in \partial_M^- A$ only if $k \geq 2$. If not, we can iterate yet again, obtaining a chain of estimates $|\psi(n_0)| \leq r^\ell |\psi(n_\ell)|$, which hold as long as the next point in the sequence n_ℓ does not lie in $\partial_M^- A$. By our assumption that n_0 is at least $k(M+1)$ from $\partial^- A$, we can iterate at least k times. Since we can iterate at least k times, we get

$$|\psi(n_0)| \leq r^k \sup_{q \in \partial_M^- A} |\psi(q)| = e^{-\gamma' k M} \sup_{q \in \partial_M^- A} |\psi(q)| \, . \tag{3.24}$$

This is precisely the bound we wanted to show. □

Notice if all $n_i \notin \partial_M^- A$ then $|\psi(n_0)| \leq r^i \sup_{q \in A} |\psi(q)|$ for any $i \in \mathbb{N}$. Since $|r| < 1$, this implies $\psi \equiv 0$, which is a trivial case for the theorem. Hence the iteration procedure does terminate for $\psi \neq 0$.

We move onto our first probabilistic analysis, which hints at a deeper theory which will be explored in the next chapter.

Theorem 3.19. *Let $R_k \to \infty$ be a sequence of integers such that for every k, every energy $E \in I = [E_1, E_2]$, and a constant $\gamma > 0$ we have $\mathbb{P}(\Lambda_{R_k}$ is not (γ, E)-good$) \to 0$ exponentially fast. Then with probability one*

$$\sigma_{ac}(H_\omega) \cap (E_1, E_2) = \emptyset \, . \tag{3.25}$$

Proof. We closely follow the proof given in section 9.2 of [13], with some extra explanation. Define $p_k = \mathbb{P}(\Lambda_{R_k}$ is not (γ, E)-good$)$. Since $R_k \to \infty$, there exists some subsequence of R_k that is increasing. By assumption, each $p_k \to 0$ exponentially, so $\sum p_k < \infty$. The Borel-Cantelli lemma (2.5) tells us the Λ_{R_k} are *not* (γ, E)-good for at most finitely many k, with probability one. Hence with probability one there exists a k_0 such that Λ_{R_k} is (γ, E)-good for all $k \geq k_0$. In particular, we have all $E \in [E_1, E_2]$ are (γ, E)-good for $\mathbb{P}$-almost all ω, by taking the sequence R_k with $k \geq k_0$.

We define three sets which we can relate through various probabilistic means. The first set consists of (E, ω) pairs which are not γ-good for V_ω, where $E \in [E_1, E_2]$ and ω is some outcome defining V_ω.

$$\mathcal{N} = \{(E, \omega) \mid E \text{ is not } \gamma\text{-good for } V_\omega\} \, . \tag{3.26}$$

The other two sets are the E and ω "components" of $\mathcal{N}$. Thus

$$\mathcal{N}_E = \{\omega \mid E \text{ is not } \gamma\text{-good for } V_\omega\} \tag{3.27}$$

$$\mathcal{N}_\omega = \{E \mid E \text{ is not } \gamma\text{-good for } V_\omega\} \, . \tag{3.28}$$

Thus, our first part of the proof above was showing $\mathbb{P}(\mathcal{N}_E) = 0$ for any $E \in [E_1, E_2]$. Let L be the Lebesgue measure on $\mathbb{R}$. Since $\mathcal{N} = \bigcup_\omega \mathcal{N}_\omega = \bigcup_E \mathcal{N}_\mathcal{E}$, Fubini's theorem tells us we can integrate over $E \in [E_1, E_2]$ or $\omega \in \Omega$. Specifically, we can write

$$\int_{E_1}^{E_2} \mathbb{P}(\mathcal{N}_E) \, dL(E) = \int_\Omega L(\mathcal{N}_\omega) \, d\mathbb{P}(\omega) \, . \tag{3.29}$$

However, we already showed $\mathbb{P}(\mathcal{N}_E) = 0$ for all $E \in [E_1, E_2]$, hence

$$0 = \int_\Omega L(\mathcal{N}_\omega) \, d\mathbb{P}(\omega) \, . \tag{3.30}$$

Thus $L(\mathcal{N}_\omega) = 0$. This tells us that $\mathcal{N}_\omega$, which is a set of energies E, has Lebesgue-measure 0. Hence Lebesgue-almost all E are γ-good for V_ω. By Lemma 3.17 we know there is no absolutely continuous spectrum in (E_1, E_2), as desired. $\square$

Chapter 4

Multiscale Analysis

We have shown the absence of absolutely continuous spectrum under strong assumptions, namely the existence of a sequence of cubes which have an increasingly high probability of being (γ, E)-good. It remains to be shown under what conditions this sequence actually exists. In addition, the absence of absolutely continuous spectrum is not enough to ensure pure point spectrum. We still have the pathological, yet physically realizable, singular continuous spectrum that could exist. The result needed is known as multiscale analysis. For a very detailed discussion of the various aspects of multiscale analysis and its development, see sections 9.2 and 9.3 in [13]. For now, we will state the result of multiscale analysis and prove what its consequences are. Recall we are working in $\ell^2(\mathbb{Z}^d)$, and are interested in an energy interval $I = [E_1, E_2]$.

Theorem 4.1 (Strong multiscale)**.** *There exist $p > 2d$, some α with $1 < \alpha < \frac{2p}{p+2d}$ and a $\gamma > 0$ such that for any disjoint cubes $\Lambda_1 = \Lambda_{L_k}(m)$ and $\Lambda_2 = \Lambda_{L_k}(n)$*

$$\mathbb{P}(\text{There is } E \in I \text{ such that both } \Lambda_1 \text{ and } \Lambda_2 \text{ are not } (\gamma, E)\text{-good}) \leq L_k^{-2p}.$$

As is discussed in [13], the proof of this result is more technically difficult than a weaker analogue, although it does not require significantly different methods. The upside to strong multiscale analysis is the relative simplicity of proving pure point spectrum as a consequence. The weak multiscale analysis immediately implies absence of absolutely continuous spectrum, and more machinery is required to prove there is only pure point spectrum. As we will see in the next section, assuming the strong multiscale analysis allows for a more transparent proof localization. After the given proof localization, a detailed summary of the proof for strong multiscale analysis is provided.

4.1 Proof of Localization

This section proves localization under the assumption of strong multiscale analysis.

Theorem 4.2. *Assume Theorem 4.1 holds for an energy interval $I = [E_1, E_2]$, so a p and α have been found. Then with probability one,*

$$\sigma_c(H_\omega) \cap (E_1, E_2) = \emptyset \ .$$

In particular, the spectrum of H_ω inside (E_1, E_2) is strictly pure point, and the corresponding eigenfunctions decay exponentially.

Proof. There are three main parts of this proof. We develop properties of necessary geometric objects in $\mathbb{Z}^d$, then bound the probability of cubes inside these objects not being (γ, E)-good. Finally, we use this bound to directly prove localization.

The geometric objects of interest are cubes Λ_{L_k} which are centered at the origin, and annuli A_k which do not contain these cubes centered at the origin, but grow linearly with them. To be specific, we choose our sequence L_k such that $L_k = L_{k-1}^\alpha$, where $\alpha > 1$, and L_0 will not be defined until the proof of Theorem 4.1. Two annuli will be defined:

$$A_k = \Lambda_{6L_{k+1}} - \Lambda_{3L_k}, \quad A_k^+ = \Lambda_{8L_{k+1}} - \Lambda_{2L_k} \ . \tag{4.1}$$

Notice that A_k is *not* disjoint from A_{k+1}, and the union of all A_k is $\mathbb{Z}^d - \Lambda_{3L_0}$. These annuli provide sufficient "room " to consider cubes in space, and compare them to the generic center cube we have constructed. Also, since we have assumed *any* disjoint cubes have low probability of simultaneously being not (γ, E)-good, this construction is sufficiently general.

Now, $A_k \subset A_k^+$ for all k, and any $n \in A_k$ has a lower bound on its distance from ∂A_k^+, as described by the following lemma.

Lemma 4.3. *For each $n \in A_k$,* $\operatorname{dist}(n, \partial A_k^+) \geq \frac{1}{3}\|n\|_\infty$.

Proof. For concreteness consider the figure below, where A_k is denoted by the blue dashed lines, and A_k^+ by the solid lines. Specifically, the lines give the *inner* boundary of each annulus. We pick some $n \in A_k$. Notice that $\operatorname{dist}(n, \partial\Lambda_{2L_k}) = \|n\|_\infty - 2L_k$. Since $\|n\|_\infty \geq 3L_k$, we have $\frac{2}{3}\|n\|_\infty \geq 2L_k$. Hence

$$\operatorname{dist}(n, \partial\Lambda_{2L_k}) \geq \|n\|_\infty - \frac{2}{3}\|n\|_\infty = \frac{1}{3}\|n\|_\infty \ .$$

Similarly $\operatorname{dist}(n, \partial\Lambda_{8L_{k+1}}) = 8L_{k+1} - \|n\|_\infty$. Since $\|n\|_\infty \leq 6L_{k+1}$, we also get the bound $\frac{8}{6}\|n\|_\infty \leq 8L_{k+1}$. Hence

$$\operatorname{dist}(n, \partial\Lambda_{8L_{k+1}}) \geq \frac{8}{6}\|n\|_\infty - \|n\|_\infty = \frac{1}{3}\|n\|_\infty \ .$$

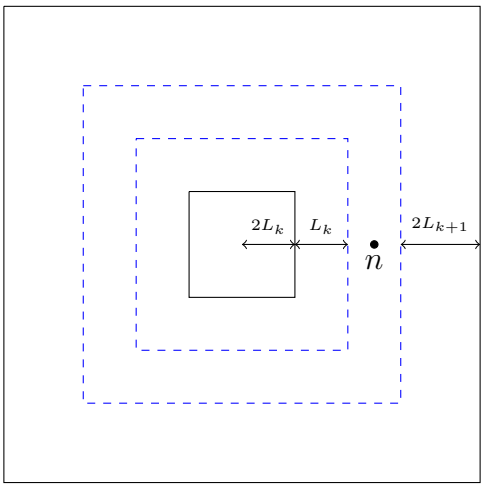

Figure 4.1: The blue dashed lines are the annulus A_k, while the solid black lines are A_k^+. The relevant distances are given.

By definition ∂A_{k+1} consists of $\partial\Lambda_{2L_k}$ and $\partial\Lambda_{8L_{k+1}}$, so n is at least $\frac{1}{3}\|n\|_\infty$ from ∂A_k^+. $\square$

We will use this bound in the final step of the proof. Next, we consider the probability that Λ_{L_k} is not (γ, E)-good while at least one of the size L_k cubes in A_k^+ is simultaneously not (γ, E)-good. We set

$$C_k^+ = C_{L_k}(A_k^+) = \{\Lambda_{L_k}(m) \mid \Lambda_{L_k}(m) \subset A_k^+ \text{ and } \Lambda_{L_k}(m) \cap \partial A_k^+ = \emptyset\} \,. \tag{4.2}$$

Thus C_k^+ is the collection of all cubes of size L_k which are far enough inside A_k^+ such that they do not intersect the boundary. As stated before, we are interested in the following probability for each k:

$$p_k = \mathbb{P}(\text{For some } E \in [E_1, E_2], \text{ both } \Lambda_{L_k} \text{ and at least one } \Lambda \in C_k^+ \text{ are } \textit{not } (\gamma, E)\text{-good}) \tag{4.3}$$

If we can show $p_k > 0$ for at most finitely many k ($\mathbb{P}$-almost surely), this would imply for all $k \geq k_0$ either Λ_{L_k} is (γ, E)-good or all $\Lambda \in C_k^+$ are (γ, E)-good. In order to do this, we prove there is some constant C such that for every k

$$p_k \leq \frac{C}{L_k^{2p-\alpha d}} \,. \tag{4.4}$$

Indeed, for any fixed $\Lambda_{L_k}(m) \in C_k^+$, our assumption of Theorem 4.1 tells us

$$\mathbb{P}(\text{For some } E \in [E_1, E_2] \text{ both } \Lambda_{L_k} \text{ and } \Lambda_{L_k}(m) \text{ are not } (\gamma, E)\text{-good}) \leq \frac{1}{L_k^{2p}} \,. \tag{4.5}$$

Considering every cube in C_k^+, we can bound p_k using the fact that $|C_k^+|$ grows as L_{k+1}^d, which itself is equal to $L_k^{\alpha d}$.

$$p_k \leq |C_k^+| \frac{1}{L_k^{2p}} \leq C(L_k^\alpha)^d \frac{1}{L_k^{2p}} = \frac{C}{L_k^{2p-\alpha d}} \tag{4.6}$$

By our assumption in Theorem 4.1, we know $\alpha < 2p/d$. This implies that $2p - \alpha d > 0$, so by comparison we easily have

$$\sum_k p_k \leq \sum_k C L_k^{-(2p-\alpha d)} < \infty \, . \tag{4.7}$$

The Borel-Cantelli lemma (2.5) tells us

$$\mathbb{P}(p_k > 0 \text{ for infinitely many } k) = 0 \, . \tag{4.8}$$

Then, as stated above, we have shown the following:

Lemma 4.4. *Assume Theorem 4.1 holds for $[E_1, E_2]$. For $\mathbb{P}$-almost all ω there exists $k_0 = k_0(\omega)$ such that for every $k \geq k_0$: For any $E \in [E_1, E_2]$ either Λ_{L_k} is (γ, E)-good or* all cubes $\Lambda_{L_k}(m) \in C_k^+$ *are (γ, E)-good.*

We are now prepared to directly prove localization. Take ω such that Lemma 4.4 is true and suppose $E \in [E_1, E_2]$ is a generalized eigenvalue. Then Theorem 3.15 says there does not exist a sequence $D_k \to \infty$ such that Λ_{D_k} is (γ, E)-good for all k. Thus Lemma 4.4 tells us there is a k_1 such that for all $k \geq k_1$, as none of the Λ_{L_k} are (γ, E)-good, it must be that *every* cube $\Lambda_{L_k}(m) \in C_k^+$ is (γ, E)-good.

Let ψ be a generalized eigenfunction for E. By definition ψ is polynomially bounded. Take $\|n\|_\infty$ large enough and find $k_2 \geq k_1$ such that $n \in A_{k_2}$. As we showed earlier, we have $\operatorname{dist}(n, \partial A_{k_2}^+) \geq \frac{1}{3}\|n\|_\infty$. Since $A_{k_2}^+$ is finite, by replacing $k(M+1)$ with this bound in Theorem 3.18, we can say

$$|\psi(n)| \leq e^{-\gamma''\|n\|_\infty} \sup_{m \in A_{k_2}^+} |\psi(m)| \, . \tag{4.9}$$

Note that for any $m \in A_{k_2}^+$ we have $\|m\|_\infty \leq 8L_{k_2+1}$, and since $n \in A_{k_2}^+$ we have $\|n\|_\infty > L_{k_2}$. Since ψ is polynomially bounded there is some power s such that

$$\begin{aligned} |\psi(m)| &\leq C_0 \|m\|_\infty^s \\ &\leq C_1 (8L_{k_2+1})^s \\ &\leq C_2 L_{k_2}^{\alpha s} \\ &\leq C_3 \|n\|_\infty^{\alpha s} \, . \end{aligned}$$

Using this bound we have

$$\begin{aligned}|\psi(n)| &\leq C_3 e^{-\gamma''\|n\|_\infty}\|n\|_\infty^{\alpha s} \\ &\leq e^{-\tilde{\gamma}\|n\|_\infty}\end{aligned}$$

since the polynomial growth is dominated by the exponential decay. So, if E is a generalized eigenvalue of H_ω then its generalized eigenfunction ψ decays exponentially fast. But any exponentially decaying function is definitely in $\ell^2(\mathbb{Z}^d)$. Hence ψ is an actual eigenfunction, so E is a true eigenvalue. Thus any generalized eigenvalue is actually an eigenvalue. From our discussion of Theorem 3.2 we know eigenvalues imply a dense pure point spectrum. Since Theorem 2.13 tells us the spectrum is approximated by the set of generalized eigenvalues, it follows that $\sigma(H_\omega)$ is pure point in (E_1, E_2). □

4.2 Analytic Estimates

In the next two sections we summarize the approach given in [13], chapters 10 and 11, to prove the multiscale analysis (Theorem 4.1). Chapter 12 also provides a brief roadmap to guide the reader through the proof. The goal of these two sections is to provide more details than chapter 12, yet not be totally rigorous in our approach (the full proof in [13] is 25 pages). There are a number of concepts and results necessary for the proof which go beyond the intended level of this paper. Hence, the broad arguments and results will be given here, with intuitive explanations intended to make multiscale analysis both a believable result, and more accessible for those interested.

The first part of the procedure requires analytic estimates to bound the growth of Green's functions. We use the familiar iteration of (3.18) to obtain these. Once these bounds are made, a probabilistic estimate is used to bring the likelihood of bad cubes to the threshold required for multiscale analysis and the resulting proof of localization in the previous section.

We begin by proving our induction step. That is, we assume multiscale analysis holds for some $\ell = L_k$. We want to show it holds for $L = L_{k+1}$. We also assume we have our scaling factor $1 < \alpha < 2$ so that $\ell^\alpha = L$ (or more specifically, the smallest integer which is at least ℓ^α.) Thus we are presuming the probability any Λ_ℓ is (γ, E)-good is very high, and the intermediary step is to show we have an exponential bound on the Green's function inside our Λ_L. The idea is to show that if a certain amount of cubes Λ_ℓ which are well-inside Λ_L are (γ, E)-good, then so is Λ_L. It so happens that assuming all but one cube of size ℓ is (γ, E)-good is sufficient. However, to illustrate the procedure we begin by assuming every cube of size ℓ inside Λ_L is (γ, E)-good.

We know from our iteration procedure (3.13) we can obtain the bound

$$|G_E^{\Lambda_L}(n,m)| \leq e^{-\tilde{\gamma}k\ell}|G_E^{\Lambda_L}(n_k,m)|$$

where n_k is sufficiently far away from $\partial\Lambda_L$. In particular, if we are assuming *every* cube of size ℓ inside Λ_L is good, then starting at a given point we can iterate up to $k = L/\ell$ times. Hence we get a bound of the form

$$|G_E^{\Lambda_L}(n,m)| \leq e^{-\tilde{\gamma}L}|G_E^{\Lambda_L}(n',m)| \ .$$

This is a good first start, but we need a bound on the second term if we are to obtain exponential decay. There is a useful property of self-adjoint operators T

$$\|(T-\mu)^{-1}\| = \frac{1}{\operatorname{dist}(\mu,\sigma(T))}, \quad \mu \in \rho(T) \ . \tag{4.10}$$

Hence to bound our Green's functions, which scale with the resolvent $(H-E)^{-1}$, we would expect decay when E is relatively far away from the spectrum. So we introduce the concept of a *resonant* cube, which puts a bound on how bad a cube can be.

Definition 4.5 (Resonant cube)**.** A cube is resonant if $\operatorname{dist}(E,\sigma(H_{\Lambda_L})) < e^{-\sqrt{L}}$.

Then our bound on the resolvent above tells us if a cube $\Lambda_L(n)$ is not resonant, then we have some weak bound on the Green's functions, namely

$$|G_E^{\Lambda_L}(n,m)| \leq e^{\sqrt{L}} \ .$$

Hence we can replace our iterative bound to finally obtain $|G_E^{\Lambda_L}(n,m)| \leq e^{-\gamma' L}$, for appropriate γ'. Notice this strongly depended on our assumption that every cube $\Lambda_\ell(m) \in C_\ell(\Lambda_L)$ is (γ,E)-good. As it turns out, this assumption is slightly too strong for our proof to hold. The probability this occurs is simply too low to maintain the bound necessary to prove multiscale analysis. This will briefly be covered in the next section.

Having failed to make use of the simplest possible case, we move to a slightly more complicated situation. We allow a single cube to not be good (but still not be resonant.) This will necessarily complicate our estimates above, yet make the probabilistic estimate sufficiently small.

The key aspect of this situation is being careful near the point $m_0 \in \Lambda_L$ such that $\Lambda_\ell(m_0)$ is *not* (γ,E)-good. In particular, any point $m \in \Lambda_{2\ell}(m_0)$ is such that

$$\Lambda_\ell(m_0) \cap \Lambda_\ell(m) \neq \emptyset \ .$$

The figure below illustrates this critical region, where $m \in \Lambda_{2\ell}(m_0)$ and thus intersects our bad cube, while $m' \notin \Lambda_{2\ell}(m_0)$ and stays clear.

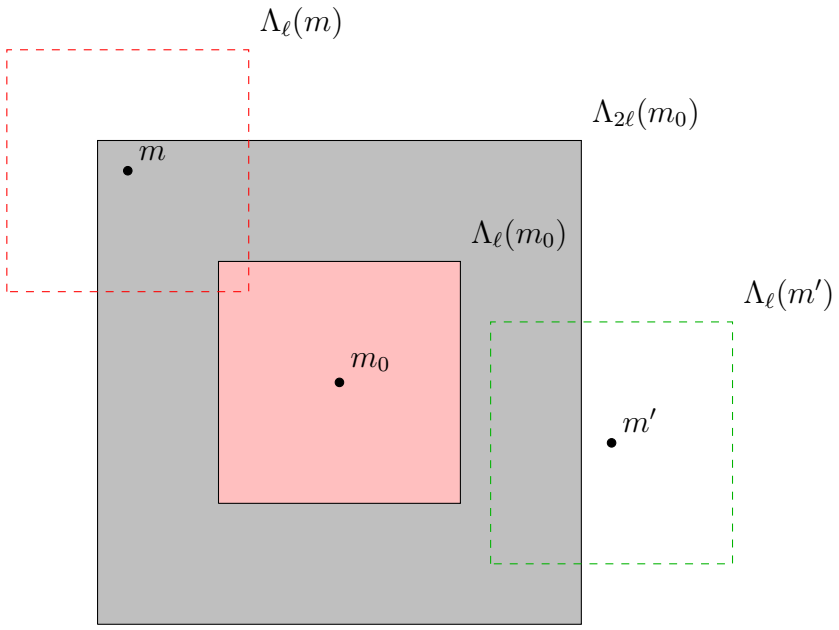

Figure 4.2: The bad cube, colored in red, and the "critical region" outside in grey. A cube of size ℓ centered at m in the critical region intersects the bad cube, while a cube of the same size at m' outside the critical region does not.

Now, we once again start at a good cube and iterate our estimate (3.13) to obtain

$$|G_E^{\Lambda_L}(m,n)| \leq e^{-\tilde{\gamma}\ell r}|G_E^{\Lambda_L}(n_r,n)| \,.$$

This estimate is iterated for as large of a value r as possible. However, we cannot iterate approximately L/ℓ times without worry as before, since we could hit our critical region $\Lambda_{2\ell}(m_0)$ at some point. So we investigate what bounds we can obtain if we have some r such that $n_r = x \in \Lambda_{2\ell}(m_0)$. Applying the same estimate as before, we can obtain some bound given that $\Lambda_\ell(m_0)$ is not resonant, which scales with the size of the cube (since the estimate uses terms across the boundary $\partial\Lambda_\ell(m_0)$). We get

$$|G_E^{\Lambda_L}(x,n)| \leq 2d(4\ell+1)^{d-1}e^{\sqrt{2\ell}}|G_E^{\Lambda_L}(x',n)|$$

where $x' \in \Lambda_L - \Lambda_{2\ell}(m_0)$. Since x' is not within our critical region, the cube $\Lambda_\ell(x')$ is (γ_r, E)-good for some γ_r. Hence we can iterate the estimate again to obtain

$$|G_E^{\Lambda_L}(x,n)| \leq (2d)^2(4\ell+1)^{d-1}(2\ell+1)^{d-1}e^{\sqrt{2\ell}}e^{-\gamma_r\ell}|G_E^{\Lambda_L}(n_{r+1},n)| \tag{4.11}$$

where we pick up the extra boundary terms from the second iteration, and the exponential decay terms from our (γ_r, E)-good cube $\Lambda_\ell(x')$. We need the full term

$$(2d)^2(4\ell+1)^{d-1}(2\ell+1)^{d-1}e^{\sqrt{2\ell}}e^{-\gamma_r \ell}$$

to be less than 1 if we want decay to occur.

It is good to notice that as we iterate through these steps, our γ changes. We start with γ_0 in our first step, and move to some $\gamma_k < \gamma_0$, so our decay does not necessarily increase. Our ultimate goal is to guarantee $\gamma_k \to \gamma^* > 0$. As it turns out, this is indeed possible. The following criteria is given in [13], which is consistent with the constant above being less than 1.

Lemma 4.6. *Suppose $L_0 \geq M(\alpha, d)$ and*

$$\gamma_{k+1} \geq \gamma_k - \gamma_k \frac{4}{L_k^{\alpha-1}} - \frac{2}{L_k^{\alpha/2}} \, .$$

If $\gamma_0 \geq 2L_0^{-1/2}$ then $\gamma_k \geq 2L_k^{-1/2}$ for all k.

Hence we can iterate our estimate even within our critical region (via the "double-step" (4.11) shown above), and also maintain a true exponential decay rate. Thus we can safely iterate until we are near $\partial\Lambda_L$, as before. The result is summarized as

Theorem 4.7. *Suppose L_0 is large enough and $L_{k+1} = L_k^\alpha$ with $1 < \alpha < 2$. Also suppose for a certain k, with $\ell = L_k$ and $L = L_{k+1}$ as before, we have the following:*

1. *There do not exist two disjoint cubes in $C_\ell(\Lambda_L)$ which are not (γ_k, E)-good with rate $\gamma_k \geq 2\ell^{-1/2}$.*
2. *No cube $\Lambda_{2\ell(m)}$ in Λ_L is E-resonant.*
3. *The cube Λ_L is not E-resonant.*

Then Λ_L is (γ_{k+1}, E)-good such that $\gamma_{k+1} \geq 2L^{-1/2}$, chosen so that

$$\gamma_{k+1} \geq \gamma_k - \gamma_k \frac{C}{L_k^{\alpha-1}} - \frac{C}{L_k^{\alpha/2}} \, .$$

As was discussed above, this final part guarantees our exponential decay.

Note again these are the heuristics for the weak multiscale. For the strong multiscale we utilized, it is necessary to allow up to 3 disjoint cubes which are simultaneously not (γ, E)-good. Of course, the work above becomes more technically involved, but the

procedure overall is quite similar. We now consider not just one critical region, but 3 critical regions depending on how many cubes are not (γ, E)-good and how close they are to each other. We look at the cubes of size 2ℓ as before, but also the cubes of size $6\ell+1$ and $10\ell+2$, all inside Λ_L. If we can guarantee these are not resonant, we can get a similar bound as before, summarized below.

Theorem 4.8. *Suppose L_0 is large enough with $L_{k+1} = L_k^\alpha$ with $1 < \alpha < 2$. Also suppose for a certain k, with $\ell = L_k$ and $L = L_{k+1}$ as before, we have the following:*

1. *There do not exist four disjoint cubes in $C_\ell(\Lambda_L)$ which are not (γ_k, E)-good with rate $\gamma_k \geq 12\ell^{-1/2}$.*
2. *No cube in*
$$C_{2\ell}(\Lambda_L) \cup C_{6\ell+1}(\Lambda_L) \cup C_{10\ell+1}(\Lambda_L)$$
is E-resonant.
3. *The cube Λ_L is not E-resonant.*

Then Λ_L is (γ_{k+1}, E)-good such that $\gamma_{k+1} \geq 12L^{-1/2}$, chosen so that

$$\gamma_{k+1} \geq \gamma_k - \gamma_k \frac{C}{L_k^{\alpha-1}} - \frac{C}{L_k^{\alpha/2}} \, .$$

Again, this allows for proper exponential decay at each step, and guarantees that $\gamma_k \not\to 0$ as $k \to \infty$. With the analytic parts out of the way, we turn our attention to the probabilistic approaches that followed, and the initial scale estimate for the basis of induction.

4.3 Probabilistic and Initial Scale Estimates

First, let us discuss where the initial analytic estimate went awry. Recall we were assuming every cube in $C_\ell(\Lambda_L)$ was (γ, E)-good, and that Λ_L was not E-resonant. This allowed us to prove Λ_L was (γ, E)-good. Then, the ultimate goal for a probabilistic bound was

$$\mathbb{P}(\Lambda_L \text{ is not } (\gamma, E)\text{-good}) \leq L^{-p}$$

given the the same is true for Λ_ℓ, bounded by ℓ^{-p}. However, our assumptions give us an estimate on the probability Λ_L is not (γ, E)-good. Indeed, it must be less than

$$\mathbb{P}(\Lambda_L \text{ is not resonant}) + \mathbb{P}(\text{One or more } \Lambda \in C_\ell(\Lambda_L) \text{ is not } (\gamma, E)\text{-good}) \, .$$

While it is possible to bound the first term, the only bound we have access to for the second term is ℓ^{-p}, by the induction hypothesis on Λ_ℓ. Hence the best possible case is $\mathbb{P}(\Lambda_L \text{ is not } (\gamma, E)\text{-good})$ becomes bounded by $\ell^{-p} = L^{-p/\alpha}$. However, doing this at each step gives us no actual progress, and is significantly worse than the bound we are working for. Thus, assuming all cubes are good is too strong.

We then worked with a single bad cube. With slightly more work, we were able to get the decay we wanted. The resulting probabilistic estimates are much better as a result. Recall again that the distribution P_0 of the random variables $V_\omega(n)$ has a bounded density g. Without worrying about a base case L_0 for now, we take the assumptions and result from Theorem 4.7.

Theorem 4.9. *For any k, let $\ell = L_k$ and $L = L_{k+1} = \ell^\alpha$. If*

$$\mathbb{P}(\Lambda_\ell \text{ is not } (\gamma, E)\text{-good}) \leq \ell^{-p},$$

then

$$\mathbb{P}(\Lambda_L \text{ is not } (\gamma, E)\text{-good}) \leq L^{-p} .$$

The method is a typical analysis-type argument. We split the probability we want into a bound containing three parts, then show each part is bounded by $\frac{1}{3}L^{-p}$. The three parts are the opposites of our assumptions. Namely, the probability Λ_L is E-resonant, the probability one of the cubes $\Lambda_{2\ell}(m)$ inside Λ_L is E-resonant, and the probability there are two disjoint cubes in Λ_L which are not (γ, E)-good.

There is an estimate that is repeatedly used throughout the analytic and probabilistic arguments, which we will make explicit here. It is called the *Wegner estimate*. We will introduce the forms most-used in these arguments.

Theorem 4.10 (Wegner Estimate)**.** *Suppose P_0 has a bounded density g. Then*

$$\mathbb{P}(\operatorname{dist}(E, \sigma(H_\Lambda)) < \varepsilon) \leq C\|g\|_\infty \, \varepsilon \, |\Lambda| .$$

If Λ_1 and Λ_2 are disjoint, finite subsets of $\mathbb{Z}^d$, then

$$\mathbb{P}(\text{There is } E \text{ such that } \operatorname{dist}(E, \sigma(H_{\Lambda_1})), \operatorname{dist}(E, \sigma(H_{\Lambda_2}) < \varepsilon) \leq 2C\|g\|_\infty \, \varepsilon \, |\Lambda_1| \, |\Lambda_2| .$$

So, to bound the probability Λ_L is E-resonant, we set $\varepsilon = e^{-\sqrt{L}}$, and we have $|\Lambda_L| \propto (2L+1)^d$. Hence this probability is bounded by $(2L+1)^d e^{-\sqrt{L}}$. Due to this exponential decay, for L large enough this will fall below the polynomial-type decay $\frac{1}{3}L^{-p}$.

We can similarly bound the probability one of the cubes in $C_{2\ell}(\Lambda_L)$ is E-resonant by just taking the probability $\Lambda_{2\ell}$ is E-resonant, and multiply it by the order $(2L+1)^d$ other points which could have cubes of size 2ℓ which are E-resonant. The probability this particular cube is E-resonant is again bounded by approximately $(4\ell+1)^d e^{-\sqrt{2\ell}}$ (since we are changing ℓ to 2ℓ everywhere). We can put this in terms of L by changing ℓ to $L^{1/\alpha}$, and by the same argument the exponential decay falls below $\frac{1}{3}L^{-p}$ for large enough L.

Finally, we consider the probability there are two disjoint cubes which are not (γ, E)-good. We can take a sum of products of probabilities for each individual cube. Specifically, assuming all $\Lambda_\ell(i)$ and $\Lambda_\ell(j)$ are within Λ_L, we have

$$\begin{aligned}
&\sum_{\Lambda_\ell(i)\cap\Lambda_\ell(j)=\emptyset} \mathbb{P}(\Lambda_\ell(i) \text{ and } \Lambda_\ell(j) \text{ are both not } (\gamma, E)\text{-good}) \\
&\leq \sum_{i,j\in\Lambda_L} \mathbb{P}(\Lambda_\ell(i) \text{ is not } (\gamma, E)\text{-good})\, \mathbb{P}(\Lambda_\ell(j) \text{ is not } (\gamma, E)\text{-good})
\end{aligned}$$

By the induction hypothesis, each of these probabilities are individually less than ℓ^{-p}. Since we are summing twice over Λ_L, this sum is bounded by $(2L+1)^{2d}\ell^{-2p}$. Multiplying the power on ℓ by α, we end up being able to limit this as well to $\frac{1}{3}L^{-p}$. Summing up these three situations, we get our bound of L^{-p} that we desired.

Kirsch gives an explanation of why two "bad" cubes are not sufficient to prove the strong multiscale analysis. It is quite analogous to the situation we saw above, when assuming all cubes were good in the weak case.

We are already familiar with the statement of the strong multiscale analysis, Theorem 4.1. The bound here is of the order L_k^{-2p} for each k. We are eventually able to obtain an estimate in terms of the probabilities of just two well-defined events, each of which can be bounded by $\frac{1}{4}L^{-2p}$.

Hopefully the reader can find the rigorous arguments presented in the proof of multiscale analysis to be more readable and attainable after this brief overview of the past two sections. In particular, the analytic estimates can be somewhat tricky, so having a good geometric image in mind can help significantly. The probabilities often feel like a game of various sets, offering a sort of puzzle to obtain the desired bound.

When reading such proofs, it is natural to wonder if a bound could be tighter than given, and if there would be any additional benefit to the proof, or the result, in such a case. It is good to keep these questions in mind. Due to the assumptions made on the

distribution P_0, there were easier jumps which could be made. As will be briefly discussed in the next chapter, weaker assumptions require even more technical arguments requiring yet more ingenuity.

One final remark is in place for the initial scale estimate, proven in chapter 11 of [13]. For small energies, the proof is somewhat involved and will not be reproduced in any manner here. However, it is more intuitive to consider how high disorder immediately implies localization.

First, high disorder means the norm of the density $\|g\|_\infty$ is quite small, as we defined the disorder to be $\delta := \|g\|_\infty^{-1}$. We have already shown that if we assume some L_0 exists with particular properties (the exponential decay there is sufficient, for example), we get all other cubes in the sequence outside it are (γ, E)-good. For high disorder, it is simply a matter of obtaining the required initial bound on the probability two cubes of size L_0 are not both (γ, E)-good. This will be shown in the proof of the theorem.

Theorem 4.11. *Suppose P_0 has bounded density g. Then for any L_0 and $\gamma > 0$, there is a $\rho > 0$ such that if $\|g\|_\infty < \rho$ and $\Lambda_{L_0}(m) \cap \Lambda_{L_0}(n) = \Lambda_1 \cap \Lambda_2 = \emptyset$, then*

$$\mathbb{P}(\text{There is } E \text{ such that both } \Lambda_1,\ \Lambda_2 \text{ are not } (\gamma, E)\text{-good}) \leq \frac{1}{L_0^{2p}} .$$

Proof. We know $|G_E^\Lambda(m,n)| \leq \|(H_\Lambda - E)^{-1}\|$ in general. Hence, if both cubes are not good then their resolvents $(H_\Lambda - E)^{-1}$ are at least $e^{-\gamma L_0}$, by definition 3.12. But we also know $\|(H_\lambda - E^{-1})\| = [\text{dist}(E, \sigma(H_\Lambda))]^{-1}$. Hence we can bound

$$\begin{aligned} & \mathbb{P}(\text{There is } E \text{ such that both } \Lambda_{L_0}(m),\ \Lambda_{L_0}(n) \text{ are not } (\gamma, E)\text{-good}) \\ \leq\ & \mathbb{P}(\text{dist}(E, \sigma(H_{\Lambda_1})) \leq e^{\gamma L_0} \text{ and } \text{dist}(E, \sigma(H_{\Lambda_2})) \leq e^{\gamma L_0} \text{ for some } E) \\ \leq\ & 2C\|g\|_\infty e^{\gamma L_0}(2L_0 + 1)^{2d} . \end{aligned}$$

In the final step, we used the Wegner estimate (4.10) across both cubes. Notice at this point that we do not actually have a good bound on the probability at all. However, for any γ and L_0, which are fixed at this point, we can indeed find a ρ small enough such that if $\|g\|_\infty < \rho$ we can bring this bound just below the threshold L_0^{-2p}. □

The end of this proof shows the disorder may have to be extremely high to force localization at some arbitrary point in the energy interval. Yet this gives intuition about some of the core aspects of localization. While it does naturally occur under some specific circumstances and energy levels, given a high enough random nature to a system, localization becomes quite probable.

Chapter 5

Discussion

As localization has been treated mathematically, experimentalists are also working hard to create the conditions under which localization is predicted. This chapter serves to highlight a few of the mathematical developments beyond the scope of this paper, as well as the future direction of localization research in a physical setting.

5.1 Mathematical Theory

This section describes further mathematical results which, while surveyed by the author, became beyond the scope of this paper. It is first good to note that the assumptions used throughout this paper, the same as those in [13], were strictly stronger than necessary for some of the results given. For example, the following result (due to Carmona et. al. [4]) which solved the issue of Bernoulli potentials has weaker assumptions for dimension $d = 1$.

Theorem 5.1. *Given dimension $d = 1$ and the potential $V_\omega(n)$ has density P_0, suppose* supp P_0 *is not concentrated at a single point and P_0 has a finite moment. That is, there exists $n > 0$ such that*

$$\int |\lambda|^n \, dP_0(v) < \infty \, .$$

Then the spectrum of H_ω is pure point with probability one and the corresponding eigenfunctions are exponentially localized.

The general approach for this proof is reminiscent of our own. Notice for the Green's functions G_E to decay as necessary for a given E, we need this E to be relatively far away from the spectrum. Thus the above theorem ends up being precisely equivalent to the following:

Theorem 5.2. *Suppose we have dimension $d = 1$, supp P_0 is not concentrated at a single point and P_0 has a finite moment. Let I be a compact interval. For any $\beta \in (0,1)$ and $\delta > 0$ there exists ℓ_0 and $\alpha > 0$ such that for any $E \in I$ and $\ell \geq \ell_0$*

$$\mathbb{P}\left(\text{dist}(E, \sigma(H_{\Lambda_\ell})) \leq e^{-\delta \ell^\beta}\right) \leq e^{-\alpha \ell^\beta}$$

When localization was first being explained and mathematically proven, it was mostly concerned with the broad conditions under which it will occur. As we have seen throughout this chapter, the methods are concerned with showing there will be cubes in space where eigenfunctions are exponentially localized. Due to the randomness inherent to the phenomenon, the arguments tell us such things occur almost certainly. Some delicate experiments back this up. Thus, given some medium through which a wave will travel, mathematicians and physicists are increasingly confident whether localization will occur somewhere within the medium. This theory's predictive power ends at this point. Thus, current research is moving away from the question of if localization will occur, and transitioning into asking where in the medium itself waves will be localized.

An informative article [7] describes work in the continuous time domain by Dr. Mayboroda. She and her collaborators have been involved in the question of precisely where waves will be localized given a certain medium. While their work exists firmly in the continuous space (as opposed to our discrete grid $\mathbb{Z}^d$), the ramifications of what they are working on will be of great interest to anyone who has gone through this paper. They are working with a *landscape function*, called such as it creates a map of the potential wells in a medium, predicting a few small subsets at a time where electrons will likely be localized [5]. In particular, the (numerically evaluated) solution to a PDE reveals a small network of subregions which define the possible areas of localization, together with a judgment of how strong localization would be in each region.

5.2 Future Experimental Directions

Prediction and Control of Localization. The mathematical theory to predict localization coupled with modern computing has allowed significant physical research into how it works with different waves and various media (e.g. see [3]). As was mentioned in the introduction, LED development has been greatly affected by recent discoveries [8]. The ability to further analyze what materials will have strong localization properties will likely allow for the creation of efficient LED lighting solutions across the world. Furthermore, it will increase our understanding of conductivity in alloys, which has been a very important field in solid state physics for a significant period of time.

Diffusion Boundary of Light. Diffusion is what occurs when an area is close to localization but does not obtain the sufficient decay of eigenfunctions, causing the tails of the functions to "interact" with each other. Imagine a few wave packets localized to various regions, yet there is some non-zero, noticeable probability their tail-ends interact at some intersecting region. Then localization will not occur, and instead one experiences diffusion. A wave still propagates somewhat, yet it becomes "spread out". The threshold between diffusion and localization in light waves specifically has become of increasing interest among physicists [2]. The reason for this is while, as we saw in our prior discussions, electrons can interact with each other and are greatly affected by random potential wells, photons do not share these properties.

Weak Localization and Destructive Interference. In some sense, weak localization is the physical precursor, or a sign of potential localization in an area. When calculating the probabilities of electron wave movement through a solid, for example, when a solid is disorganized it becomes necessary to consider the probability an electron will complete a path back to where it began. This is the essence of weak localization. It is easy to conceptually understand how this would begin to imply localization. If enough electrons are being deflected in such a way to move back to where they started along various paths, and if this continually occurs with high probability, the wave would indeed stop propagating. If the electron waves stay within some small fixed area, they become localized as a result. Hence, this weak localization was used as a signal for Anderson localization. Throughout various physical experiments to understand how localization plays into the transition between conductors and insulators, it was nearly impossible to determine whether electrons were localized in a particular instance due to some potential, or due to weak localization and the resulting destructive interference of waves. Light offers the opportunity to investigate the latter due to their lack of interaction with potential wells. Experiments can slowly move from diffusion to localization, in an attempt to understand this change more fully.

Notation

H_0	p. 20	Discrete Laplacian.
V	p. 25	Potential operator, works as a multiplication operator.
V_ω	p. 25	Random potential operator.
H	p. 26	Discrete Hamiltonian, defined as $H_0 + V$.
H_ω	p. 26	Discrete random Hamiltonian, defined as $H_0 + V_\omega$.
$\ell^2(\mathbb{Z}^d)$	p. 19	Space of square-summable complex sequences.
$\|n\|_\infty$	p. 19	Supremum norm on $\mathbb{Z}^d$, defined as $\sup_{k=1,\dots,d} \|n_k\|$.
$\|n\|_1$	p. 19	Sum norm on $\mathbb{Z}^d$, defined as $\sum_{k=1}^d \|n_k\|$.
$\mathcal{H}$	p. 7	Generic Hilbert space.
$\mathcal{H}_\mathbf{x}$	p. 17	The subspace $\{\varphi \mid \varphi \text{ has property } \mathbf{x}\}$.
$T_\mathbf{x}$	p. 17	Restriction of T to the space $\mathcal{H}_\mathbf{x}$.
H_A	p. 34	Restriction of H to the space $\ell^2(A)$, $A \subset \mathbb{Z}^d$.
$\rho(T)$	p. 7	Resolvent set of operator T.
$\sigma(T)$	p. 8	Spectrum of operator T.
$\sigma_\mathbf{x}(T)$	p. 17	Spectrum of $T_\mathbf{x}$.
$\mu(B)$	p. 13	Projection valued spectral measure for T, defined as $\chi_B(T)$.
$\varepsilon_g(T)$	p. 14	Set of generalize eigenvalues of T.
P	p. 23	Distribution of random variable (typically V_ω).
supp P	p. 23	Support of P.
g	p. 23	Density function of P.
$\delta(g)$	p. 30	Disorder, defined as $\|g\|_\infty^{-1}$.
$\mathbb{P}(\omega)$	p. 23	Probability event ω occurs.
$\Lambda_L(m)$	p. 30	Cube centered at $m \in \mathbb{Z}^d$ with side length $2L+1$.
$\partial^- \Lambda_L$	p. 34	Inner boundary of Λ_L.
$\partial^+ \Lambda_L$	p. 34	Outer boundary of $\Lambda_L(n)$.
$\partial \Lambda_L$	p. 25	Boundary of Λ_L, the union of the inner and outer boundaries.
$\partial_L^- \Lambda$	p. 36	Width L inner boundary of Λ.
25 $C_L(A)$	p. 36	Set of $\Lambda_L(m) \subset A$ with $\Lambda_L(m) \cap \partial A = \emptyset$.
$A \nearrow \mathbb{Z}^d$	p. 34	Ascending chain $A_m \subset A_{m+1} \subset \mathbb{Z}^d$ whose union is $\mathbb{Z}^d$.
$G_E^\Lambda(n,m)$	p. 34	Green's function of H_Λ.

Appendix: Generalized Eigenvalues

Here we provide a proof of Theorem 2.13 with a few more details than those in section 7.1 of [13]. We restate the theorem here for convenience.

Theorem. *The spectrum of the discrete Hamiltonian H agrees up to a set of spectral measure zero with $\varepsilon_g(H)$. In particular, $\sigma(H) = \overline{\varepsilon_g(H)}$.*

We first provide a few reminders and some additional machinery. We remind the reader of the definition of a projection valued spectral measure found in section 1.1. Given our operator H we let $\mu(B) = \chi_B(H)$, where B is some Borel set of $\mathbb{R}$. Then we can write

$$\langle \varphi, H\psi \rangle = \int_{\sigma(H)} \lambda \, d\mu_{\varphi,\psi}(\lambda), \quad \mu_{\varphi,\psi}(B) = \langle \varphi, \mu(B)\psi \rangle \; . \tag{5.1}$$

In our particular space $\ell^2(\mathbb{Z}^d)$ we also set $\mu_{n,m}(B) = \langle \delta_n, \mu(B)\delta_m \rangle$. Next, given some sequence of positive real numbers $\{\alpha_n\}_{n\in\mathbb{Z}^d}$ such that $\sum \alpha_n = 1$, we can define

$$\rho(B) = \sum_{n\in\mathbb{Z}^d} \alpha_n \mu_{n,n}(B) \; . \tag{5.2}$$

Then $\rho(\mathbb{R}) = 1$ and ρ is a finite positive Borel measure. We call ρ a *real spectral measure*. Suppose $\rho(B) = 0$. Since $\alpha_n \neq 0$ for all n it follows that $\mu_{n,n}(B) = \langle \delta_n, \mu(B)\delta_n \rangle = 0$ for all n. This implies that $\mu(B) = 0$. Conversely, if $\mu(B) = 0$ it immediately follows that $\rho(B) = 0$. Hence

$$\rho(B) = 0 \iff \mu(B) = 0 \; . \tag{5.3}$$

Since $\mu(\sigma(H)) = I$ this also implies supp $\rho = \sigma(H)$. Recall two sets A and B agree up to a set of spectral measure zero if $\mu(A - B) = \mu(B - A) = 0$. The above condition tells us it is also sufficient to prove $\rho(A - B) = \rho(B - A) = 0$.

We define $\Lambda_L = \{|n| \leq L\}$ to be the cube of side length $2L + 1$, and $\|\psi\|_S$ to be the ℓ^2 norm of ψ over $S \subset \mathbb{Z}^d$. Finally, we need the following simple lemma.

Lemma 5.3. *If $\psi \neq 0$ is polynomially bounded and ℓ is a positive integer, then there is a sequence $L_n \to \infty$ such that*

$$\frac{\|\psi\|_{\Lambda_{L_n+\ell}}}{\|\psi\|_{\Lambda_{L_n}}} \to 1 \; . \tag{5.4}$$

Proof. Suppose there is no such sequence L_n. Then there exists some $r > 1$ and L_0 such that for $L \geq L_0$ we have $\|\psi\|_{\Lambda_{L+\ell}} \geq r\|\psi\|_{\Lambda_L}$. Specifically, the ratio (at best) converges to some constant larger than 1. Applying this bound k times past L_0 we have

$$\|\psi\|_{\Lambda_{L_0+\ell k}} \geq r^k \|\psi\|_{\Lambda_{L_0}} \tag{5.5}$$

However, ψ is polynomially bounded so

$$\|\psi\|_{\Lambda_{L_0+\ell k}} \leq C(L_0 + \ell k)^M \leq C' k^M \tag{5.6}$$

for appropriate $C', M > 0$. This contradicts 5.5, as a function cannot simultaneously exponentially grow while being polynomially bounded. □

With this, we are ready to prove our theorem. We break it into two parts. The first lemma says that approximately all $\lambda \in \sigma(H)$ are generalized eigenvalues. The second lemma says that every generalized eigenvalue is indeed in the spectrum.

Lemma 5.4. *Let ρ be a real spectral measure for $H = H_0 + V$. Then for ρ-almost all $\lambda \in \sigma(H)$ there exists a polynomially bounded solution of $H\psi = \lambda\psi$.*

Proof. Using the Cauchy-Schwarz inequality on inner products we have

$$|\mu_{n,m}(B)| \leq \mu_{n,n}(B)^{\frac{1}{2}} \mu_{m,m}(B)^{\frac{1}{2}} \; .$$

So, assume $\rho(B) = 0$. Then we know $\mu(B) = 0$, but this only occurs when $\mu_{n,n}(B) = 0$ for all n, which also implies $\mu_{n,m}(B) = 0$. Thus $\rho(B) = 0$ implies $\mu_{n,m}(B) = 0$ for all n, m. We say the $\mu_{n,m}$ are absolutely continuous with respect to ρ. Since ρ is a probability measure, the Radon-Nikodym theorem says there exist measurable densities $F_{n,m}$ such that

$$\mu_{n,m}(B) = \int_B F_{n,m}(\lambda) \, d\rho(\lambda) \; . \tag{5.7}$$

Since these $F_{n,m}$ are defined through an integral with respect to ρ, they are defined up to sets of ρ-measure zero. In addition note that $\mu_{n,n} \geq 0$ and thus $F_{n,n} \geq 0$, ρ-almost

surely. If we consider our definition of ρ we see

$$\begin{aligned}\rho(B) &= \sum \alpha_n \mu_{n,n}(B) \\ \int_B d\rho(\lambda) &= \sum \alpha_n \int_B F_{n,n}(\lambda) d\rho(\lambda) \\ \int_B d\rho(\lambda) &= \int_B \left(\sum \alpha_n F_{n,n}(\lambda)\right) d\rho(\lambda) \ .\end{aligned}$$

This tells us ρ-almost surely that $\sum \alpha_n F_{n,n}(\lambda) = 1$. In particular for any n we have $\alpha_n F_{n,n}(\lambda) \leq 1$, so

$$F_{n,n}(\lambda) \leq \frac{1}{\alpha_n} \ . \tag{5.8}$$

Using this bound it is easy to show

$$\begin{aligned}\left| \int_B F_{n,m}(\lambda)\ d\rho(\lambda) \right| &= |\mu_{n,m}(B)| \\ &\leq \mu_{n,n}(B)^{\frac{1}{2}} \mu_{m,m}(B)^{\frac{1}{2}} \\ &= \left(\int_B F_{n,n}(\lambda)\ d\rho(\lambda) \right)^{\frac{1}{2}} \left(\int_B F_{m,m}(\lambda)\ d\rho(\lambda) \right)^{\frac{1}{2}} \\ &\leq \alpha_n^{-\frac{1}{2}} \alpha_m^{-\frac{1}{2}} \rho(B) \ .\end{aligned}$$

Differentiating with respect to ρ, we see

$$|F_{n,m}| \leq \alpha_n^{-\frac{1}{2}} \alpha_m^{-\frac{1}{2}} \ . \tag{5.9}$$

Now, using our original definition we have

$$\mu_{n,m}(B) = \langle \delta_n, \mu(B)\delta_m \rangle = \int_B F_{n,m}(\lambda)\ d\rho(\lambda) \ .$$

Using our knowledge of the spectral calculus (section 1.1), we can replace our $\mu(B)$ with a bounded, measurable function f so that

$$\langle \delta_n, f(H)\delta_m \rangle = \int_B f(\lambda) F_{n,m}(\lambda)\ d\rho(\lambda) \ . \tag{5.10}$$

We take $f(\lambda) = \lambda g(\lambda)$ where g has compact support. Then,

$$\begin{aligned}&\int \lambda g(\lambda) F_{n,m}(\lambda) d\rho(\lambda) \\ &= \langle \delta_n, Hg(H)\delta_m \rangle \quad \text{by definition given above} \\ &= \langle H\delta_n, g(H)\delta_m \rangle \quad \text{by self-adjointness of } H.\end{aligned}$$

Recall the definition of $H(n) = H_0 + V(n)$, where

$$H_0 u(n) = -\sum_{\|e\|=1} (u(n+e) - u(n)) = 2du(n) - \sum_{\|e\|=1} u(m+e)$$

since in d dimensions, any point m has $2d$ neighbors of distance 1, so the $u(n)$ terms in the sum can be combined $2d$ times. Also notice that $V\delta_n = V(n)$. Using this, we get

$$\begin{aligned}
&\int \lambda g(\lambda) F_{n,m}(\lambda) d\rho(\lambda) \\
&= \langle H_0\delta_n, g(H)\delta_m\rangle + \langle V\delta_n, g(H)\delta_m\rangle \\
&= \sum_{\|e\|=1} (-\langle \delta_{n+e}, g(H)\delta_m\rangle) + 2d\langle\delta_n, g(H)\delta_m\rangle + V(n)\langle\delta_n, g(H)\delta_m\rangle \\
&= \sum_{\|e\|=1} (-\langle \delta_{n+e}, g(H)\delta_m\rangle) + (V(n)+2d)\langle\delta_n, g(H)\delta_m\rangle \\
&= \sum_{\|e\|=1} \left(-\int g(\lambda) F_{n+e,m}(\lambda) d\rho(\lambda)\right) + \int g(\lambda)(V(n)+2d) F_{n,m}(\lambda) d\rho(\lambda) \quad \text{by (5.10)} \\
&= -\sum_{\|e\|=1} \left(\int g(\lambda)[F_{n+e,m}(\lambda) - F_{n,m}(\lambda)] d\rho(\lambda)\right) + \int g(\lambda) V(n) F_{n,m}(\lambda) d\rho(\lambda) \\
&= \int g(\lambda)[H_0 F_{n,m}(\lambda)](n) d\rho(\lambda) + \int g(\lambda) V(n) F_{n,m}(\lambda) d\rho(\lambda) \\
&= \int g(\lambda)[H F_{n,m}(\lambda)](n) d\rho(\lambda) \ .
\end{aligned}$$

The text uses the notation $H^{(n)} F_{n,m}(\lambda)$, which is equivalent to our notation; however ours emphasizes that we are evaluating the function $HF_{n,m}(\lambda)$ at the point n, i.e. we have fixed m and λ throughout. To summarize this string of equations, we have

$$\int_B g(\lambda)\lambda F_{n,m}(\lambda)\, d\rho(\lambda) = \int_B g(\lambda)[HF_{n,m}(\lambda)](n)\, d\rho(\lambda) \ . \tag{5.11}$$

for any bounded measurable function g which has compact support. Thus for ρ-almost all λ and fixed $m \in \mathbb{Z}^d$, we have

$$HF_{n,m}(\lambda) = \lambda F_{n,m}(\lambda) \ .$$

In particular, if we set $\psi(n) = F_{n,m}(\lambda)$ then $H\psi = \lambda\psi$. From our bound on $F_{n,m}$ we know

$$|\psi(n)| \le C_0 \alpha_n^{-\frac{1}{2}} \ .$$

Although we have not chosen our sequence α_n yet, it need only fulfill $\alpha_n > 0$ with $\sum \alpha_n = 1$. Hence we can choose $\alpha_n = c(1 + \|n\|_\infty)^{-\beta}$ for appropriate c and $\beta > d$. These are positive and will sum to 1. In particular we have

$$|\psi(n)| \leq C(1 + \|n\|_\infty)^{\frac{d}{2}+\varepsilon}$$

for some $\varepsilon > 0$, since $\beta > d$. Thus ψ is polynomially bounded. Hence we see that for ρ-almost all $\lambda \in \sigma(H)$, we can explicitly construct a generalized eigenfunction ψ, proving the lemma. □

Thus we have shown that nearly all elements of the spectrum are generalized eigenvalues as well. Next, we will show the set of generalized eigenvalues is entirely within the spectrum.

Lemma 5.5. *If $H\psi = \lambda\psi$ has a polynomially bounded solution ψ, then $\lambda \in \sigma(H)$.*

Proof. Let ψ be our polynomially bounded solution. We will create a sequence of functions which will allow us to use Lemma 5.3. Set

$$\psi_L(n) = \begin{cases} \psi(n) & \text{for } |n| \leq L, \\ 0 & \text{otherwise .} \end{cases}$$

We also normalize these functions as $\varphi_L = \|\psi_L\|^{-1}\psi_L$. We can see that $(H-\lambda)\psi_L(n) = 0$ as long as

$$n \notin S_L := \{m \mid L - 1 \leq |m| \leq L + 1\} .$$

This is because S_L represents a "transition region" between ψ and 0, so behavior may be unexpected. However, since ψ is polynomially bounded we have

$$\begin{aligned} \|(H - \lambda)\psi_L\|^2 &\leq \|\psi\|^2_{S_L} \\ &= \sum_{m \in S_L} |\psi(m)|^2 \\ &= \|\psi\|^2_{\Lambda_{L+1}} - \|\psi\|^2_{\Lambda_{L-2}} \end{aligned}$$

where the final bound is just a matter of summing over two cubes. The difference is precisely the set S_L. We can apply Lemma 5.3 with $\ell = 3$ to find a sequence $L_n \to \infty$ such that

$$\frac{\|\psi\|^2_{\Lambda_{L_n+1}}}{\|\psi\|^2_{\Lambda_{L_n-2}}} \to 1 \tag{5.12}$$

and use our previous bound to say

$$\begin{aligned}\|(H-\lambda)\varphi_{L_n}\|^2 &\leq \|(H-\lambda)\psi_{L_n}\|^2\|\psi_{L_n}\|^{-2} \\ &\leq \frac{\|\psi\|^2_{S_L}}{\|\psi\|^2_{\Lambda_{L_n-2}}} \\ &\leq \frac{\|\psi\|{\Lambda_{L_n+1}}^2 - \|\psi\|^2_{\Lambda_{L_n-2}}}{\|\psi\|^2_{\Lambda_{L_n-2}}} \to 0\ .\end{aligned}$$

This tells us that φ_{L_n} is a Weyl sequence for H and λ. Using the Weyl Criterion (1.17) we see that $\lambda \in \sigma(H)$. □

We complete the proof of the theorem with a final lemma.

Lemma 5.6. *$\sigma(H)$ is the closure of $\varepsilon_g(H)$.*

Proof. We have just shown in Lemma 5.5 that $\varepsilon_g(H) \subset \sigma(H)$. Now, we know that $\sigma(H)$ is closed (Theorem 1.3) hence $\overline{\varepsilon_g(H)} \subset \sigma(H)$. If we denote $\sigma(H) - \varepsilon_g(H) = \varepsilon_g(H)^{\complement}$, then the first statement of the theorem we just proved specifically says

$$\rho(\varepsilon_g(H)^{\complement}) = 0 \implies \rho\left(\overline{\varepsilon_g(H)}^{\complement}\right) = 0\ .$$

Since supp $\rho = \sigma(H)$, we can finally say

$$\overline{\varepsilon_g(H)}^{\complement} \cap \sigma(H) = \overline{\varepsilon_g(H)}^{\complement} \cap \text{supp } \rho = \emptyset$$

Thus there are no elements $\lambda \in \sigma(H)$ such that $\lambda \notin \overline{\varepsilon_g(H)}$. Hence $\sigma(H) \subset \overline{\varepsilon_g(H)}$. By double inclusion this shows $\sigma(H) = \overline{\varepsilon_g(H)}$. □

Bibliography

[1] Decomposition of spectrum (functional analysis). *Wikipedia*, Feb 2017.

[2] Christof M. Aegerter, Martin Strzer, Peter Gross, and Georg Maret. Observation of the Critical Regime Near Anderson Localization of Light. *Photonic Metamaterials: From Random to Periodic*, 2006.

[3] D. Arnold, G. David, M. Filoche, D. Jerison, and S. Mayboroda. Computing spectra without solving eigenvalue problems. *ArXiv e-prints*, November 2017.

[4] Rene Carmona, Abel Klein, and Fabio Martinelli. Anderson Localization for Bernoulli and Other Singular Potentials. *Communications in Mathematical Physics*, 108(1):4166, 1987.

[5] Marcel Filoche and Svitlana Mayboroda. The landscape of anderson localization in a disordered medium. *Contemporary Mathematics*, page 113121, 2013.

[6] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, 2 edition, 2005.

[7] Kevin Hartnett. Mathematicians Tame Rogue Waves, Lighting Up Future of LEDs. *Quanta magazine*, Aug 2017.

[8] Kevin Hartnett. The Math That Promises to Make the World Brighter. *Quanta magazine*, Sep 2017.

[9] P. D. Hislop and Israel Michael. Sigal. *Introduction to spectral theory: with applications to Schrodinger operators*. Springer, 2012.

[10] D. Hundertmark. A short introduction to Anderson Localization. September 2007.

[11] Sajeev John. Localization of light. *Physics Today*, 44(5):3240, 1991.

[12] Tosio Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, 1980.

[13] W. Kirsch. An Invitation to Random Schrödinger operators. *ArXiv e-prints*, September 2007.

[14] Erwin Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, 1978.

[15] Hervé Kunz and Bernard Souillard. Sur le spectre des opérateurs aux différences finies aléatoires. *Communications in Mathematical Physcs*, 78:201–246, 1980.

[16] Ad Lagendijk, Bart Van Tiggelen, and Diederik S. Wiersma. Fifty years of Anderson localization. *Physics Today*, 62(8):2429, 2009.

[17] Michael Reed and Barry Simon. *Methods of modern mathematical physics I: Functional analysis*. Academic Press, 1980.

[18] Akito Suzuki. Spectrum of Discrete Schrödinger Operator and Discrete Laplacian, Sep 2012.