Ergodic Properties of Random Schrödinger Operators

by

Irina Y. Zhecheva

A Thesis
Submitted in partial fulfillment
of the requirements for the Degree of Bachelor of Arts with Honors
in Mathematics

WILLIAMS COLLEGE

Williamstown, Massachusetts

May 05, 2008

Acknowledgements

I would like to thank my thesis advisor, Professor Stoiciu, for his enormous patience and understanding, and for his incredible encouragement and support. I would also like to thank professor Silva, my second reader, for giving me valuable suggestions.

Contents

Chapter 1

Introduction

The Schrödinger equation plays a central role in quantum mechanics and mathematical physics in general. In this thesis we attempt to give an introduction to some of the basic theory of random Schrödinger operators. In particular, we look at some major results concerning the spectral theory of the one-dimensional Schrödinger operator with a random potential. The random Schrödinger operator has been intensively studied in the last few decades, in both its discrete and its continuous formulations. Variations of this operator arise very naturally in physics and mathematics, in the theory of electron conduction and in Hamiltonian mechanics. In the theory of electron conduction, random Schrödinger operators model disordered solids. This exposition focuses on the one-dimensional Anderson model, namely the model with random operators on the Hilbert space $l^2(\mathbb{Z})$. The theory underlying the study of random Schrödinger operators is fascinating in its own right, even abstracting from its applications to physics. What is so interesting about it is that it combines several areas of mathematics and uses tools that have applications in applied probability, stochastic processes and other fields.

This thesis introduces many of the tools that are needed to understand the theory of random Schrödinger operators. We begin by introducing fundamental results in ergodic theory, functional analysis, probability, and measure theory that we will need to be able to understand the theory of random Schrödinger operators.

The focus of this thesis is to examine the spectrum of the random Schrödinger operator and to show that, contrary to what we would expect, it is deterministic. More precisely, we show that it is non-random with probability 1. This result was proved by Pastur. A sketch of the proof can be found in [1], but a lot of the details are omitted. The goal of this thesis is to provide the missing details in the proofs of the theorems leading up to this important result by Pastur. What enables us to prove that the spectrum is deterministic is the fact that the random Schrödinger operator pertains to a class of operators called ergodic operators. The framework of ergodic theory turns out to be particularly suitable for investigating the properties of the random Schrödinger operator. We start by giving a careful proof of the Mean Ergodic theorem, which is a very important result in ergodic theory.

Chapter 2

Background from Ergodic Theory

2.1 Basic Notions from Ergodic Theory

The theory of dynamical systems studies the long-term behavior of states in a system given a transformation which drives the system's evolution. A dynamical system is a transformation which maps a space into itself. Some dynamical systems are measure-preserving in the sense that the measure of a set is always the same as the measure of the set of points which map to it. More formally, a transformation is measure preserving if $\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A)$ for any measurable set A. A set is called invariant if it is the same as its image. An ergodic dynamical system is one in which, with respect to some probability distribution, all invariant sets have measure either 0 or 1.

Here are the formal definitions of the concepts from ergodic theory that we are going to need. Throughout we assume that we are given a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is a probability space, $\mathcal{F}$ is a σ-algebra of sets, and $\mathbb{P}$ is a probability measure.

Definition 1. *A measurable mapping $T : \Omega \to \Omega$ is a measure-preserving transformation if $\mathbb{P}(T^{-1}(A) = \mathbb{P}(A)$ for all $A \in \mathcal{F}$.*

Definition 2. *Let $\{T_i\}_{i\in\mathbb{Z}}$ be a family of measure-preserving transformations. A set A is invariant under $\{T_i\}$ if $T_i^{-1}(A) = A$ for all $i \in \mathbb{Z}$.*

Definition 3. *A triple $(\Omega, \mathcal{F}, T_i)$ of an event space, probability measure and a transformation is called stationary if $\mathbb{P}(T_i^{-1}A) = \mathbb{P}(A)$ for any $A \in \mathcal{F}$.*

Definition 4. *A family $\{T_i\}$ of measure-preserving transformations is called ergodic with respect to the probability measure $\mathbb{P}$ if any invariant set $A \in \mathcal{F}$ has probability 0 or 1.*

Definition 5. *A family $\{X_i\}_{i\in\mathbb{Z}}$ of random variables is called a stochastic process (with index set $\mathbb{Z}$). This means that there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ such that the X_i are real-valued, measurable functions on $(\Omega, \mathcal{F})$.*

Definition 6. *A stochastic process $\{X_i\}_{i\in\mathbb{Z}}$ is called ergodic if there exists an ergodic family of measure-preserving transformations $\{T_i\}_{i\in\mathbb{Z}}$ such that $X_i(T_j\omega) = X_{i-j}(\omega)$.*

In our case we are going to take T_i to be the shifting transformations. Thus, we also need the following definition.

Definition 7. *A set A is called shift invariant if $T_i^{-1}A = A$ for all $i \in \mathbb{Z}$.*

Definition 8. *A stationary probability measure $\mathbb{P}$ is called ergodic if any shift-invariant set A has probability 0 or 1.*

In the next section we introduce the mean ergodic theorem due to Birkhoff, which is one of the most important results in Ergodic Theory.

2.2 The Mean Ergodic Theorem

The mean ergodic theorem intuitively says the following. For any integrable function of our space, if we randomly pick a point according to the ergodic distribution and calculate the time average of the function along the point's orbit, then as time goes to infinity, the time average converges to the space average, which is equal to the weighted average of the value of the function at all points in the space. More formally, consider a point x in the state space, a function f that we are averaging, a map T and a probability measure $\mathbb{P}$. The space average is $\bar{f} = \int f(x)d\mathbb{P}(x)$. The time-average starting from x is $\langle f\rangle_x = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n} f(T^i(x))$. The ergodic theorem states that if f is integrable and T is ergodic with respect to $\mathbb{P}$, then $\langle f\rangle_x$ exists, and $\mathbb{P}\left(\{x : \langle f\rangle_x = \bar{f}\}\right) = 1$.

Here we state a common formulation of the mean ergodic theorem due to Birkhoff.

Theorem 2.2.1. *Let (X, S, μ) be a probability measure space, and $T : X \to X$ be a measure-preserving transformation. Let $I = \{g \in L^2(X) : g \circ T = g \text{ a.e.}\}$ be the subspace of T-invariant functions and let P be the projection from $L^2(X)$ to I. Then for any $f \in L^2$, the sequence*

$$\frac{1}{n}\sum_{i=0}^{n-1} f\left(T^i x\right) \tag{2.1}$$

converges to $P(f)$ in $L^2(X)$.

Proof. We prove the theorem for each of two orthogonal subspaces that span L^2. Then since any element in L^2 can be written as a linear combination of elements in the two subspaces, the theorem holds for the general case as well. When $f \in I$, $f \circ T^n = f$ a.e., so the statement of the theorem holds:

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i = \frac{1}{n}nf = f = P(f) \tag{2.2}$$

Here $f = P(f)$ because $f \in I$, so the projection of f is also in I. Consider a subspace B of functions,

$$B = \{f : f = g \circ T - g, g \in L^2\} \tag{2.3}$$

Then if $f = g \circ T - g$,

$$\sum_{i=0}^{n-1} f \circ T^i = \sum_{i=0}^{n-1} (g \circ T - g) \circ T^i = \sum_{i=0}^{n-1} \{g \circ T^{i+1} - g \circ T^i\} = g \circ T^n - g \tag{2.4}$$

Thus, we can see that

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i = \frac{1}{n}(g \circ T^n - g) \tag{2.5}$$

converges to 0 as $n \to \infty$.
Thus, the ergodic theorem holds for $f \in P(f)$. Now we are going to show that it also holds for $f \in \bar{B}$, where $\bar{B}$ denotes the norm closure of B. Let f_j be a sequence of functions converging in norm to a function $f \in L^2$. Then

$$\begin{aligned} ||\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i||_2 =& ||\frac{1}{n} \sum_{i=0}^{n-1} (f - f_j + f_j) \circ T^i||_2 \\ &\leq ||\frac{1}{n} \sum_{i=0}^{n-1} (f - f_j) \circ T^i||_2 + ||\frac{1}{n} \sum_{i=0}^{n-1} f_j \circ T^i||_2 \\ &\leq \frac{1}{n} \sum_{i=0}^{n-1} ||(f - f_j) \circ T^i||_2 + \frac{1}{n} \sum_{i=0}^{n-1} ||f_j \circ T^i||_2 \\ &\frac{1}{n} \sum_{i=0}^{n-1} ||(f - f_j)||_2 + \frac{1}{n} \sum_{i=0}^{n-1} ||f_j \circ T^i||_2 \end{aligned}$$

Therefore, for functions in the norm closure of $B, \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i \to 0$ as $n \to \infty$. Consider $f \in I$ and $h \in B$. We can write h as $h = g \circ T - g$ for $g \in L^2$. Then,

$$\langle f, h\rangle = \langle f, g \circ T - g\rangle = \langle f, g \circ T\rangle - \langle f, g\rangle = \langle f \circ T, g \circ T\rangle - \langle f, g\rangle = \langle f, g\rangle - \langle f, g\rangle = 0 \tag{2.6}$$

where the last equality follows because T is measure-preserving. Therefore, $f \in B^\perp$, and so $I \subset B^\perp$. Consider a sequence $h_k \in B$ converging to h in norm. Since $||h_k - h|| \to 0$, and $\langle f, h_k\rangle \to 0$, for all $k > 0$, it follows that

$$\langle f, h\rangle \to 0. \tag{2.7}$$

Therefore, we have shown that f is in the orthogonal complement of the norm closure of B, so $I \subset \bar{B}^\perp$. Now we are going to show that $\bar{B}^\perp \subset I$. Let $f \in \bar{B}^\perp$. Then $\langle f, h\rangle = 0$ for any $h \in \bar{B}$. Since $B = \{h : h = g \circ T - g, g \in L^2\}$, it follows that $\langle f, g \circ T - g\rangle = 0$ for any $g \in L^2$. Therefore, $\langle f, g\rangle = \langle f, g \circ T\rangle$, and

$$\langle f, g\rangle = \langle f, g \circ T\rangle = \int f\bar{g} \circ T d\mu = \int f \circ T^{-1}\bar{g} d\mu = \langle f \circ T^{-1}, g\rangle. \tag{2.8}$$

Thus, since $\langle f, g\rangle = \langle f \circ T^{-1}, g\rangle$, it follows that $\langle f - f \circ T^{-1}, g\rangle = 0$ for all $g \in L^2$. Therefore, $f = f \circ T^{-1} = f \circ T$, so $f \in I$. Thus, we have shown that $\bar{B}^\perp \in I$, and we have completed the proof that $\bar{B}^\perp = I$.

We have shown that the ergodic theorem holds for $\bar{B}$ and I, 2 orthogonal subspaces of L^2. Since any element of L^2 can be written as the sum of an element in $\bar{B}$ and an element in I, it follows that we have proved the ergodic theorem for L^2. □

In physics, the ergodic hypothesis says that over long periods of time, the time spent by a particle in some region of the phase space of microstates with the same energy is proportional to the volume of this region, i.e. all accessible microstates are equally likely over a long period of time.

In ergodic theory, an ergodic measure is a measure that satisfies the ergodic hypothesis for a given map of a measurable space into itself. Intuitively, an ergodic measure is one with respect to which the points of the space are "well mixed up" by the map.

Definition 9. *(Ergodic measure) Let (X, Ω) be a measurable space, let $T : X \to X$ be a measurable function and let $f : X \to X$ be a measurable and integrable function. A finite measure $\mu : \Sigma \to [0, \infty]$ is called an ergodic measure for the transformation T and the function f if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(T^k(x)\right) = \frac{1}{\mu(X)} \int_X f(y)d\mu(y) \tag{2.9}$$

for μ-almost all points $x \in X$.

The left-hand side term can be interpreted as the long-term average of f, since it is the average value of f over the first n iterates of x under T in the limit when $n \to \infty$. The right-hand side term can be interpreted as the space average of f, since it is the average value of f over the phase space X with respect to the measure μ. Thus, for an ergodic measure, the space average of f equals the time average. The measure $\mu(X)$ is finite and so it can be normalized so that $\mu(X) = 1$. Thus, μ can be interpreted as a probability measure.

An alternative version of the mean ergodic theorem more clearly express the fact that under certain conditions the time average of a function is equal to the space average. This version is the one most commonly used in statistical mechanics.

Theorem 2.2.2. *(Birkhoff Ergodic Theorem) Let T be a measure-preserving transformation on a measure space (Ω, μ). If μ is ergodic and $\mu(\Omega) = 1$, then for any $f \in L^1(\Omega, \mu)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f\left(T^k x\right) = \int_\Omega f(y)d\mu(y) \tag{2.10}$$

for almost all x.

The mean ergodic theorem is a very important result in statistical mechanics. It guarantees that starting from a random point in the state space, orbits will almost certainly pass arbitrarily close to every point in the phase space. To see this more clearly, consider the indicator function I_A of a set A. The time-average of I is the fraction of time that the orbit spends in the set A. The space average of I is the probability that a randomly chosen point is in A. Since the two trajectories are almost always equal, almost all trajectories cover almost all of the state space.

Another way to think about the mean ergodic theorem is that it generalizes the strong law of large numbers - a basic result about independent, identically distributed random variables - to ergodic processes. Birkhoff's theorem tells us intuitively that a sufficiently large sample (i.e. an average over all time) is representative of the whole population (i.e. the space average). We can restate the ergodic theorem in a form that more clearly reveals the connection with the strong law of large numbers.

Theorem 2.2.3. *If $\{X_i\}_{i\in\mathbb{Z}}$ is an ergodic process and $\mathbb{E}(|X_0|) < \infty$ then*

$$\lim_{L\to\infty} \frac{1}{(2L+1)} \sum_{i\in\wedge_L} X_i \to \mathbb{E}(X_0) \tag{2.11}$$

for $\mathbb{P}$-almost ω.

Here Λ_L denotes the cube of side length $2L+1$:

$$\Lambda_L(n_0) := \{n \in \mathbb{Z}; \|n - n_0\|_\infty \leq L\} \tag{2.12}$$

It contains $\|\Lambda_L(n_0)| := (2L+1)$ points. The mean ergodic theorem expressed in this form is very useful for proving claims related to the density of states, a very important quantity in condensed matter physics. This concept will be explored in the last chapter of this thesis.
Next we state the most commonly used condition for ergodicity. Below, we denote by $\mathcal{S}(\mathcal{H})$ the space of self-adjoint operators.

Definition 10. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $H : \Omega \to \mathcal{S}(\mathcal{H})$ be measurable. $H : \Omega \to \mathcal{S}(\mathcal{H})$ is called ergodic if there exists an ergodic family of measure-preserving transformations $(T_i)_{i\in I}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ and a family of unitary operators $(U_i)_{i\in I}$ on $\mathcal{H}$ such that the following condition holds:*

$$H\big(T_i(\omega)\big) = U_i^* H(\omega) U_i. \tag{2.13}$$

This condition turns out to be crucial in proving properties about the spectrum of the random Schrödinger operator. It will enable us to prove that its spectrum is deterministic.

Chapter 3

The Schödinger Operator

In quantum mechanics, a particle moving in one-dimensional space is described by a vector ψ in the Hilbert space $L^2(\mathbb{R})$. The time evolution of the particle's state is modeled by the Schrödinger operator

$$H = H_0 + V \tag{3.1}$$

acting on $L^2(\mathbb{R})$. The operator H_0 is called the free operator and represents the kinetic energy of the particle. In the absence of magnetic fields, the free operator is the Laplacian

$$H_0 = \frac{\hbar^2}{2m}\Delta = -\frac{\hbar^2}{2m}\sum_{\nu=1}^{d}\frac{\partial^2}{\partial x_\nu^2} \tag{3.2}$$

In physics the need arises to add a potential V, in which case we get the following operators:

$$\big(-\Delta + V\big)f = -f'' + vf \tag{3.3}$$

In two dimensions, we get

$$\big(-\Delta + V\big)f(x,y) = -\frac{\partial^2 f^2}{\partial x^2}(x,y) - \frac{\partial^2 f^2}{\partial y^2}(x,y) + V(x,y)f(x,y) \tag{3.4}$$

The physical properties of the system are encoded in the potential V which is the multiplication operator with the function $V(x)$ in the Hilbert space $L^2(\mathbb{R})$. In this exposition we study the one-dimensional Anderson model, which replaces the Hilbert space $L^2(\mathbb{R})$ with the sequence space $l^2(\mathbb{Z})$. This is the discretized version of the Schrödinger operator in 1 dimension acting on the sequence space $l^2(\mathbb{Z})$. We start by looking at the discrete Laplacian, but first we are going to define the concept of the spectrum of an operator.

3.1 The Spectrum

In functional analysis, the spectrum of an operator is a generalization of the concept of eigenvalues in the case of matrices. Operators on infinite-dimensional spaces may have no eigenvalues. For example, on the Hilbert space l^2 the unilateral shift operator,

$$(x_1, x_2, \cdots) \to (0, x_1, x_2, \cdots) \tag{3.5}$$

has no eigenvalues, but its spectrum is non-empty. In fact, any bounded linear operator on a complex Banach space has a non-empty spectrum. We first define the resolvent set of a bounded linear operator, which has some nice properties because it is an open set.

Definition 11. *The resolvent set of an operator T is $\rho(T) = \mathbb{C} \setminus \sigma(T) = \{\lambda, T - \lambda I$ is invertible$\}$.*

The spectrum of an operator is the complement of the resolvent.

Definition 12. *The spectrum of an operator T is $\sigma(T) = \{\lambda, T - \lambda I$ is not invertible$\}$.*

We are also going to need the definition of a unitary operator and unitary equivalence.

Definition 13. *In an inner product space $\mathbb{H}$, the adjoint U^* of an operator U is defined by the requirement that $\langle Uu, v\rangle = \langle u, U^*v\rangle$.*

Definition 14. *A unitary operator is a bounded linear operator $U : \mathbb{H} \to \mathbb{H}$ on a Hilbert space $\mathbb{H}$ satisfying*

$$U^*U = UU^* = I \tag{3.6}$$

where U^* is the adjoint of U, and $I : \mathbb{H} \to \mathbb{H}$ is the identity operator. A unitary operator preserves the inner product $\langle\rangle$ on the Hilbert space, i.e. for all vectors x and y in the Hilbert space,

$$\langle Ux, Uy\rangle = \langle x, y\rangle. \tag{3.7}$$

Definition 15. *Two operators A and B are unitarily equivalent if there exists a unitary operator U such that $B = UAU^*$.*

Theorem 3.1.1. *If T and S are unitarily equivalent, then $\sigma(T) = \sigma(S)$.*

Proof. Since S and T are unitarily equivalent, there exists a unitary matrix U such that $S = UTU^{-1}$. Thus, we need to show that $\sigma(T) = \sigma(UTU^{-1})$. It suffices to show that $\mathbb{C} \setminus \sigma(T) = \mathbb{C} \setminus \sigma(UTU^{-1})$.

$\lambda \in \mathbb{C} \setminus \sigma(T) \iff (T - \lambda I)$ is invertible
$\iff U(T - \lambda I)U^{-1}$ is invertible
$\iff UTU^{-1} - U(\lambda I)U^{-1}$ is invertible
$\iff UTU^{-1} - \lambda I$ is invertble
$\iff \lambda \in \mathbb{C} \setminus \sigma(UTU^{-1})$ □

This theorem shows that the spectrum of an operator is invariant under unitary equivalence. The following theorem can be proved analogously.

Theorem 3.1.2. *If $S = T + c \cdot I$ then $\sigma(S) = \sigma(T) + c$.*

Having discussed the concept of the spectrum of an operator, we are going to give a detailed exposition of the formal construction of the random Schrödinger operator. The details of this construction can be found in [1].

3.2 The Discrete Laplace Operator

We start by introducing the discrete Laplacian. The discrete Laplace operator is essentially a discrete analog of the continuous Laplace operator adapted to the case when we have a discretized grid. The discrete Laplace operator has applications in a variety of fields. For example, in numerical analysis it is used as an approximation to the continuous Laplacian. The discrete Laplacian Δ acting on u is defined by

Definition 16. $(\Delta_d u)(i) = \sum_{j:|j-i|=1}[u(j) - u(i)]$

where the sum is over the nearest neighbors of the ith element of the sequence u. Note that we can rewrite Δ_d as

$$\Delta_d u = \sum_{j:|j-i|=1} \{u(j)\} - 2vu(i) \tag{3.8}$$

It is easy to see why this relationship holds in the 1-dimensional case:

$$\begin{aligned}
&(\Delta_d u)(i) = \sum_{j:|j-i|=1} \{u(j) - u(i)\} = \\
&u(i+1) - u(i) + u(i-1) - u(i) \\
&= u(i-1) + u(i+1) - 2u(i) = \sum_{j:|j-i|=1} \{u(j)\} - 2u(i)
\end{aligned}$$

We can think of Δ_d as a discrete analog of the Laplacian $\frac{\partial^2 f}{\partial x^2}$ because

$$\begin{aligned}
\Delta_d u(i) &= u(i+1) + u(i-1) - 2u(i) \\
&= \frac{u(i+1) - u(i)}{(i+1) - i} - \frac{u(i) - u(i-1)}{i - (i-1)} \\
&= \frac{u'(i+1) - u'(i)}{(i+1) - i} \\
&= u''(i+1)
\end{aligned}$$

It is interesting to note that Δ_d satisfies an identity similar to the identity $< f(x), -\frac{\partial^2 f(x)}{\partial x^2} >= ||\frac{\partial f}{\partial x}||_2^2$ for the continuous case.

Proposition 3.2.1. $2 < u, -\Delta_d u >= \sum_{j:|i-j|=1} |u(i) - u(j)|^2$

Proof. The left hand side equals

$$\begin{aligned}
&2 < u, -\Delta_d u >= \\
&= 2\sum \bar{u}_i(-u_{i-1} - u_{i+1} + 2u_i) = \\
&= -2\sum \bar{u}_i u_{i-1} - 2\sum \bar{u}_i u_{i+1} + 4\sum |u_i|^2 = \\
&= -2\sum \bar{u}_{i+1} u_i - 2\sum \bar{u}_i u_{i+1} + 4\sum |u_i|^2
\end{aligned}$$

The right hand side equals

$$
\begin{aligned}
&\sum_{j:|i-j|=1} |u_i - u_j|^2 \\
&= \sum |u_i - u_{i-1}|^2 + \sum |u_i - u_{i+1}|^2 \\
&= \sum (\bar{u}_i - \bar{u}_{i-1})(u_i - u_{i-1}) + \sum (\bar{u}_i - \bar{u}_{i+1})(u_i - u_{i+1}) \\
&= \sum |u_i|^2 + \sum |u_{i-1}|^2 - \sum \bar{u}_{i-1} u_i - \sum \bar{u}_i u_{i-1} + \sum |u_i|^2 + \sum |u_{i+1}|^2 - \\
&- \sum \bar{u}_{i+1} u_i - \sum \bar{u}_i u_{i+1} = \\
&= \sum |u_i|^2 + \sum |u_i|^2 - \sum \bar{u}_i u_{i+1} - \sum \bar{u}_{i+1} u_i + \sum |u_i|^2 + \sum |u_i|^2 - \\
&- \sum \bar{u}_{i+1} u_i - \sum \bar{u}_i u_{i+1} \\
&= 4 \sum |u_i|^2 - 2 \sum \bar{u}_i u_{i+1} - 2 \sum \bar{u}_{i+1} u_i
\end{aligned}
$$

Thus, we see that the left-hand side and the right-hand side are equal. This completes the proof. □

It is useful to consider the representation of the Laplacian as an infinite matrix. Notice that

$$
\begin{pmatrix}
\ddots & & & & & & \\
\cdots & -2 & 1 & 0 & 0 & 0 & \cdots \\
\cdots & 1 & -2 & 1 & 0 & 0 & \cdots \\
\cdots & 0 & 1 & -2 & 1 & 0 & \cdots \\
\cdots & 0 & 0 & 1 & -2 & 1 & \cdots \\
\cdots & 0 & 0 & 0 & 1 & -2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}
\begin{pmatrix}
\vdots \\ u_{-2} \\ u_{-1} \\ u_0 \\ u_1 \\ u_2 \\ \vdots
\end{pmatrix}
=
\begin{pmatrix}
\vdots \\ u_{-3} + u_{-1} - 2u_{-2} \\ u_{-2} + u_0 - 2u_{-1} \\ u_{-1} + u_1 - 2u_0 \\ u_0 + u_2 - 2u_1 \\ u_1 + u_3 - 2u_2 \\ \vdots
\end{pmatrix}
$$

Therefore, we can see that $\Delta_d(i)$ is a tridiagonal matrix with -2's on the diagonal and 1's on the two diagonals to the left and to the right of the main diagonal.

The spectrum of the discrete Laplacian can be found relatively easily. We can show that $\sigma(\Delta_d) = [-4v, 0]$ by Fourier transformation. The Discrete Fourier transform is defined by $\mathcal{F}(\{u_n\})(t) = \sum_{n=-\infty}^{\infty} u_n e^{-int}$.

In order to find the spectrum, we are first going to look at how the Discrete Laplacian acts on the level of functions.

$$
\begin{aligned}
F\{\Delta_d u\} &= \mathcal{F}(\{u_{n-1} + u_{n+1} - 2u_n\}) \\
&= \sum_{n=-\infty}^{\infty} (u_{n-1} + u_{n+1} - 2u_n) e^{-int} \\
&= \sum_{n=-\infty}^{\infty} u_{n-1} e^{-int} + \sum_{n=-\infty}^{\infty} u_{n+1} e^{-int} - 2 \sum_{n=-\infty}^{\infty} u_n e^{-int} \\
&= \sum_{n=-\infty}^{\infty} u_n e^{-i(n+1)t} + \sum_{n=-\infty}^{\infty} u_n e^{-i(n-1)t} - 2 \sum_{n=-\infty}^{\infty} u_n e^{-int}
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-it}\sum_{n=-\infty}^{\infty} u_n e^{-int} + e^{it}\sum_{n=-\infty}^{\infty} u_n e^{-int} - 2\sum_{n=-\infty}^{\infty} u_n e^{-int} \\
&= e^{-it}F\{u_n\} + e^{it}F\{u_n\} - 2F\{u_n\} \\
&= F\{u_n\}[e^{-it} + e^{it} - 2] \\
&= (2\cos t - 2)F\{u_n\}
\end{aligned}
$$

Thus, we can see that $f(t) = 2\cos t - 2$ sends $F\{u\}$ to $F\{\Delta_d u\}$
Therefore, the spectrum of $f(t)$ is $\sigma(\delta_d) = [-4; 0]$, since that is the range of $f(t)$. This is consistent with the result that $\sigma(\Delta_d) = [-4v; 0]$ in the case when the dimension is $v = 1$.

Now we proceed to formally define the Schrödinger operator. The details of this construction are taken from [1]. Now we proceed to the formal construction of the Schrödinger operator. First we define the operator $(-1)^N$ by $[(-1)^N u](n) = (-1)^{|n|+}u(n)$ Consider how this operator acts on a sequence $u(n)$:

$$
(\ (-1)^n\ )\begin{pmatrix} \vdots \\ u_{-2} \\ u_{-1} \\ u_0 \\ u_1 \\ u_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ u_{-2} \\ -u_{-1} \\ u_0 \\ -u_1 \\ u_2 \\ \vdots \end{pmatrix}.
$$

Notice that
$$
\begin{pmatrix} \ddots & & & & & & \\ & 1 & 0 & 0 & 0 & 0 & \\ & 0 & -1 & 0 & 0 & 0 & \\ & 0 & 0 & 1 & 0 & 0 & \\ & 0 & 0 & 0 & -1 & 0 & \\ & 0 & 0 & 0 & 0 & 1 & \\ & & & & & & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ u_{-2} \\ u_{-1} \\ u_0 \\ u_1 \\ u_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ u_{-2} \\ -u_{-1} \\ u_0 \\ -u_1 \\ u_2 \\ \vdots \end{pmatrix}.
$$

Therefore, we get the following matrix representation of $[(-1)^n]$:

$$
[(-1)^n] = \begin{pmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & -1 & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & 1 & \\ & & & & & & \ddots \end{pmatrix}
$$

It is easy to show that $[(-1)^n]$ is unitary. We also observe that $[(-1)^n]^{-1} = (-1)^n$.
Now we proceed by adding a potential to the Laplacian. Define $\bar{H} = -\Delta_d + V$ where

$$V = \begin{pmatrix} \ddots & & & & & & \\ & v_{-2} & & & & & \\ & & v_{-1} & & & & \\ & & & v_0 & & & \\ & & & & v_1 & & \\ & & & & & v_2 & \\ & & & & & & \ddots \end{pmatrix}$$

$\bar{H}$ can be regarded as the discrete version of the Schrödinger operator. Here the Hilbert space is the sequence space $l^2(\mathbb{Z}^d)$ instead of $L^2(\mathbb{R}^d)$. We call this setting the discrete case in contrast to Schrödinger operators on $L^2(\mathbb{R}^d)$ which we refer to as the continuous case.

Remark 3.2.1. *$\bar{H}$ is unitarily equivalent to $4 + \Delta_d + V$.*

Proof.

$$\begin{aligned}
&(-1)^N \bar{H}[(-1)^N]^{-1} = (-1)^N \bar{H}(-1)^N \\
&= \begin{pmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & -1 & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & 1 & \\ & & & & & & \ddots \end{pmatrix} \\
&\times \begin{pmatrix} \ddots & & & & & & & \\ -1 & v_{-2}+2 & -1 & & & & & \\ & -1 & v_{-1}+2 & -1 & & & & \\ & & -1 & v_0+2 & -1 & & & \\ & & & -1 & v_1+2 & -1 & & \\ & & & & -1 & v_2+2 & -1 & \\ & & & & & & \ddots & \end{pmatrix} \\
&\begin{pmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & -1 & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & 1 & \\ & & & & & & \ddots \end{pmatrix} =
\end{aligned}$$

$$= \begin{pmatrix} \ddots & & & & & & & \\ 1 & v_{-2}+2 & 1 & & & & & \\ & 1 & v_{-1}+2 & 1 & & & & \\ & & 1 & v_0+2 & 1 & & & \\ & & & 1 & v_1+2 & 1 & & \\ & & & & 1 & v_2+2 & 1 & \\ & & & & & & \ddots & \end{pmatrix}$$

Therefore, $\bar{H}$ is unitarily equivalent to $4 + \Delta_d + V$ □

Consider the operators

$$(H_0 u)(n) = \sum_{j:|j-n|_+=1} u(j)$$

and

$$(Hu)(n) = (H_0 u)(n) + V(n)u(n) \tag{3.9}$$

Note that $H_0(n) = \sum_{j:|j-n|_+=1} u(j) = u(n+1)+u(n-1)$, so $H_0(n)$ can be represented by the matrix $\begin{pmatrix} \ddots & & & & & & \\ 1 & 0 & 1 & & & & \\ & 1 & 0 & 1 & & & \\ & & 1 & 0 & 1 & & \\ & & & 1 & 0 & 1 & \\ & & & & 1 & 0 & 1 \\ & & & & & & \ddots \end{pmatrix}$

Subsuming the diagonal terms of Δ_d into the potential V, we can see that $\bar{H}$ is unitarily equivalent to H up to a constant. Since theorem 3.1.1 guarantees that unitary equivalence preserves the spectral properties of the operators, we can use $\bar{H}$ instead of H to represent the Schrödinger operator.

3.3 Randomizing the Potential

So far we have completed the construction of the Schrödinger operator. In this thesis, we are interested in random Schrödinger operators, i.e. operators where the potential is random. In quantum mechanics, these operators model disordered solids. Most solids do not constitute ideal crystals. The positions of the atoms may deviate from the ideal lattice positions in a non-regular way due to imperfections in the crystallization process. Or the positions of the atoms may be completely disordered as is the case in amorphous or glassy materials. The solid may also be a mixture of various materials which is the case for example for alloys or doped semiconductors. In all these cases it seems reasonable to think of the potential as a random quantity. This description of the motivation for considering random Schrödinger operators is based on [7].

With this in mind, we can model the potential as a sequence of independent, identically distributed random variables $\{V_n(\omega)\}_n$ on a probability space Ω. For any $n \in \mathbb{Z}$, the potential $V(n)$ evaluated at n is a random variable, i.e. a measurable function on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. To make this concept precise, next we are going to discuss some fundamental concepts from the theory of probability.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A probability or probability measure is a measure whose total mass is one. Because the origins of probability lie in statistics rather than analysis, some of the terminology is different. For example, instead of denoting a measure space by $(X, \mathcal{A}, \mu)$, in probability the notation is $(\Omega, \mathcal{F}, \mathbb{P})$. Here Ω is a set, $\mathcal{F}$ is called a σ-field (i.e. σ-algebra), and $\mathbb{P}$ is a measure with $\mathbb{P}(\Omega) = 1$. Elements of $\mathcal{F}$ are called events, and elements of Ω are denoted ω. Events such as $(\omega : X(\omega) > a)$ are abbreviated by $(X > a)$. Similarly, for a set $A \subset \mathbb{R}$, $(\omega : X(\omega) \in A$ is abbreviated by $(X \in A)$ Real-valued measurable functions from Ω to $\mathbb{R}$ are called random variables and are usually denoted by X or Y or other capital letters. Integration (in the sense of Lebesgue) is called expectation or expected value, and we write $\int X d\mathbb{P}$. The random variable 1_A is the function that is one if $\omega \in A$ and zero otherwise. It is called the indicator function of A. This is the analog of the term characteristic function used in analysis.

For the convenience of the reader, here we give the rigorous definitions of the concepts we just discussed. These definitions can be found for example in [10].

Definition 17. *An ($\mathcal{F}$)-measurable function is a map $f : \Omega \to \mathbb{R}$ from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $f^{-1}(A) \equiv \{\omega \in \Omega : f(\omega) \in A\} \in \mathcal{F}$ for every $A \in \mathcal{B}(\mathbb{R})$.*

Definition 18. *A random variable is a real-valued measurable function $f : \Omega \to \mathbb{R}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

Definition 19. *If X is a random variable, we can define a probability measure F_X on $\mathbb{R}$ called the distribution of X:*

$$F_X(A) = \mathbb{P}\big(\{\omega \big| X(\omega) \in A\}\big) \tag{3.10}$$

for any Borel set A.

Alternatively, we can define a probability on $\mathbb{R}$ by

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A), A \subset \mathbb{R} \tag{3.11}$$

where the probability $\mathbb{P}_X$ is called the law of X or the distribution of X. We define $F_X : \mathbb{R} \to [0, 1]$ by

$$F_X(x) = \mathbb{P}_X((-\infty, x]) = \mathbb{P}(X \le x) \tag{3.12}$$

The function F_X is called the distribution function of X.

Theorem 3.3.1. *The distribution function F_X of a random variable X satisfies:*
a) F_X is non-decreasing;
b) F_X is right continuous with left limits;
c) $\lim_{x\to\infty} F_X(x) = 1$ and $\lim_{x\to-\infty} F_X(x) = 0$.

If a sequence of random variables $\{X_i\}_{i\in I}$ have the same distributions, we say that they are identically distributed. More formally,

Definition 20. *The random variables X and Y are identically distributed if for every set $A \in S$, $\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$. Equivalently, X and Y are identically distributed if $F_X(x) = F_Y(x)$ for every x where F_X denotes the cumulative distribution function of X.*

Now we formally define the notion of independence for events, σ-algebras, and random variables.

Definition 21. *A sequence of events $A_1, \cdots, A_n$ are independent if*

$$\mathbb{P}\big(A_{i_1} \cap A_{i_2} \cap \cdots A_{i_j}\big) = \mathbb{P}\big(A_{i_1}\big)\mathbb{P}\big(A_{i_2}\big)\cdots\mathbb{P}\big(A_{i_j}\big) \tag{3.13}$$

for every subset $\{i_1, \cdots, i_j\}$ of $\{1, 2, \cdots, n\}$.

Definition 22. *A countable set $\{\mathcal{G}_n\}_{n\in I}$ of σ-algebras $G_n \subset \mathcal{F}$ are independent if any finite set of events $A_1, \cdots, A_n$ from distinct $\mathcal{G}_k$ are independent.*

Definition 23. *A countable set $\{X_n\}_{n\in I}$ of random variables are independent if the σ-algebras $\sigma(X_n)$ generated by these random variables are independent. Alternatively, a sequence of random variables $\{X_i\}_{i\in I}$ is called independent if for any finite subset $\{i_1, \cdots, i_n\}$ of I,*

$$\begin{aligned}&\mathbb{P}\big(\{\omega \big| X_{i_1}(\omega) \in [a_1, b_1], X_{i_2}(\omega) \in [a_2, b_2], \cdots, X_{i_n}(\omega) \in [a_n, b_n]\}\big) = \\ &\mathbb{P}\big(\{\omega \big| X_{i_1}(\omega) \in [a_1, b_1]\}\big) \cdot \ldots \cdot \mathbb{P}\big(\{\omega \big| X_{i_n}(\omega) \in [a_n, b_n]\}\big).\end{aligned}$$

It is important to remark that if X_i are independent and identically distributed (i.i.d.) with common distribution F_X, then

$$\begin{aligned}&\mathbb{P}\big(\{\omega \big| X_{i_1}(\omega) \in [a_1, b_1], X_{i_2}(\omega) \in [a_2, b_2], \cdots, X_{i_n}(\omega) \in [a_n, b_n]\}\big) = \\ &F\big([a_1, b_1]\big) \cdot F\big([a_2, b_2]\big) \cdot \ldots \cdot F\big([a_n, b_n]\big).\end{aligned}$$

Here we are going to include another definition from probability that we are going to need in the next chapter.

Definition 24. *(Product space) Given two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, we define the product space $\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$. The product σ-algebra on $\Omega_1 \times \Omega_2$ is defined as $\mathcal{F}_1 \times \mathcal{F}_2 = \sigma\{\rho_1, \rho_2\}$, where $\rho_1 : \Omega_1 \times \Omega_2 \to \Omega_i$ are the projection maps $\rho_1 : (\omega_1, \omega_2) \to \omega_1$ and $\rho_2 : (\omega_1, \omega_2) \to \omega_2$.*

Definition 25. *(Product measure) Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be two probability spaces , and denote by $\rho_1 : \Omega_1 \times \Omega_2 \to \Omega_1$ and $\rho_2 : \Omega_1 \times \Omega_2 \to \Omega_2$ the projection maps. There exists a unique probability measure $\mathbb{P}_1 \times \mathbb{P}_2$ on $\mathcal{F}_1 \times \mathcal{F}_2$ called the product measure, under which the distribution of ρ_1 is $\mathbb{P}_1$, the distribution of ρ_2 is $\mathbb{P}_2$, and ρ_1 and ρ_2 are independent Ω_1- and Ω_2- valued random variables, respectively.*

Another difference in terminology between analysis and probability is the set-theoretic notion of 'almost everywhere.' In probability, instead of saying that a property occurs almost everywhere, we talk about properties occurring almost surely, written *a.s.*

With these definitions in mind, we add a subscript ω to the potential V to make it clear that V_ω depends on (unknown) random parameters.

The special way in which Ω is constructed is irrelevant. Without loss of generality, we may use the probability space defined by

$$\Omega = S^{\mathbb{Z}} \tag{3.14}$$

where S is a (Borel-) subset of $\mathbb{R}$. The corresponding σ-algebra $\mathcal{F}$ is generated by cylinder sets of the form

$$\{\omega | \omega_{i_1} \in A_1, \cdots, \omega_{i_n} \in A_n\} \tag{3.15}$$

where $A_1, \cdots, A_n$ are Borel subsets of $\mathbb{R}$. On Ω the random variables X_i can be realized by $X_i(\omega) = \omega_i$.

Even without randomization, the potential is already a quite complicated function. At first glance it is not clear why we need to randomize it. The reason is that with the introduction of random variables we change our point of view. Now we are no longer interested in properties of the Schrödinger operator for a single given ω, but instead we are interested in its properties for 'almost all' ω, or, in other words, for 'typical' ω. In mathematical terms, we are interested in results of the form: the set of all ω such that H_ω has the property $\mathcal{P}$ as probability 1. In short, $\mathcal{P}$ holds for $\mathbb{P}$-almost all ω (or $\mathbb{P}$-almost surely). Here $\mathbb{P}$ is the probability measure on the underlying probability space. There are a number of such properties that hold for random Schrödinger operators. In this thesis, we are going to focus on giving the details of the proof of the property that the spectrum of the random Schrödinger operator is deterministic with probability 1. In other words, we are going to show that there is a closed, nonrandom subset Σ of the real line such that $\Sigma = \sigma(H_\omega)$, the spectrum of the operator H_ω, $\mathbb{P}$-a.s. The proof of this property is due to Pastur and is well understood among experts in the field of mathematical physics and random Schrödinger operators in particular. An outline of the proof is given in [1], but a lot of details are omitted. The result that the spectrum of the random Schrödinger operator is deterministic is truly unexpected in light of the fact that the potential of H_ω is non-random. Therefore, the main goal of this thesis is to provide the full details of this important proof so that it is accessible to the non-specialist.

Recall the way we defined the operator H in (3.9). Now that we have modeled the potential as a sequence of i.i.d. random variables, we add it to the (negative) Laplacian to get the random Schrödinger operator. We denote the Schrödinger operator with randomized potential by H_{V_ω}. It can be represented as a matrix in the following way:

$$\begin{pmatrix} \ddots & & & & & & & \\ 1 & v_{-2}(\omega) & 1 & & & & & \\ & 1 & v_{-1}(\omega) & 1 & & & & \\ & & 1 & v_0(\omega) & 1 & & & \\ & & & 1 & v_1(\omega) & 1 & & \\ & & & & 1 & v_1(\omega) & 1 & \\ & & & & & & \ddots & \end{pmatrix} \tag{3.16}$$

where $v_0(\omega)$ is at position (0,0).

Now that we have completed the construction of the random Schödinger operator, we are going to devote the next chapter to introducing some background from ergodic theory, and then using ideas from this field to gain insight into the spectrum of the random Schrödiger operator.

Chapter 4

Connecting Spectral Theory with Ergodicity

In this section we are going to show that the random Schrödinger operator is ergodic. Recall that a family of measure-preserving transformations is called ergodic if any invariant set has probability 0 or 1. Now we define what it means for an operator to be ergodic.

4.1 Ergodicity of the Random Schrödinger Operator

Definition 26. *An operator A_ω is called ergodic if it satisfies the identity*

$$A_{T_i\omega} = U_i A_\omega U_i^* \tag{4.1}$$

where $\{T_i\}$ are ergodic transformations, U_i is a unitary operator, and U_i^* is its adjoint.

It turns out that the random Schrödinger operator is ergodic with respect to the shift operators $\{T_i\}_{i\in\mathbb{Z}}$ which are defined as

$$T_i w(j) = w(j-i) \tag{4.2}$$

where $\omega \in \Omega$.

Recall that we defined the potential V_ω so that for any $n \in \mathbb{Z}$, the potential $V(n)$ evaluated at n is a random variable, and the set $\{V_n(\omega)\}_n$ is a sequence of i.i.d. random variables. In general, any sequence of i.i.d. random variables is an ergodic stochastic process. In our case, we know that the sequence of i.i.d. random variables $X_n(\omega) = V_\omega(n)$ is an ergodic process. Recall the definition of product measure introduced in the previous chapter. Because of the independence of the random variables, the probability measure $\mathbb{P}$ on our probability space $\Omega = S^{\mathbb{Z}}$ is the infinite product measure of the probability measure $\mathbb{P}_0(A) = \mathbb{P}(V_\omega(0) \in A)$, where $\mathbb{P}_0$ is the distribution of $V_\omega(0)$. Since any set must have probability strictly between

0 and 1, when we take the infinite product of the probabilities, in the limit we get either 0 or 1. This is why the process $\{V_\omega(n)\}$ is ergodic.

Since the random potential $\{V_\omega(n)\}_{n\in\mathbb{Z}}$ is an ergodic process, by definition there exists a family of ergodic measure-preserving transformations $\{T_i\}$ on Ω such that $V_\omega(n)$ satisfies

$$V_{T_i\omega}(n) = V_\omega(n-i) \tag{4.3}$$

and any measurable subset of Ω which is invariant under the $\{T_i\}$ has probability 0 or 1.

Because of the way that we constructed our probability space as a product of infinitely many copies of S where S is a Borel subset of $\mathbb{R}$, we can infer that the equation

$$V_{T_i\omega}(n) = V_\omega(n-i) \tag{4.4}$$

holds if we take the $\{T_i\}$ to be the shift operators that we defined above. Because of ergodicity, it also follows that any measurable subset of Ω which is invariant under the $\{T_i\}$ has trivial probability (i.e. has probability 0 or 1).

Now we are going to define the unitary operators U_i and U_i^* with respect to which H_ω is ergodic as in definition (26) above.

Define the translation operator $\{U_i\}_{i\in\mathbb{Z}}$ on $l^2(\mathbb{Z})$ in the following way:

Definition 27. $(U_i u)(n) = u(n-i)$

We can represent this operator as a matrix which is essentially the identity matrix with its diagonal shifted down by i positions. Thus, for example, for $i = 2$ we can write U_i as

$$\begin{pmatrix} \ddots & & & & & & \\ 0 & 0 & 0 & & & & \\ 1 & 0 & 0 & 0 & & & \\ & 1 & 0 & 0 & 0 & & \\ & & 1 & 0 & 0 & 0 & \\ & & & 1 & 0 & 0 & 0 \\ & & & & 1 & & \ddots \end{pmatrix} \tag{4.5}$$

Clearly, applying U_i to a vector u shifts the vector down by i positions.
Then the adjoint of U_i, U_i^*, can be derived from the definition of the adjoint of an operator which we presented earlier. We find that

Theorem 4.1.1. *The adjoint of U_i is defined by $(U_i^* v)(n) = v(n+i)$.*

Proof. It suffices to verify that $\langle u(n-i), v\rangle = \langle u, v(n+i)\rangle$ which is a simple computation. □

Applying U_i^* to a vector v shifts it up by i positions. We can also think of U_i^* as the identity matrix with its diagonal shifted up by 2 positions. For example, for $i = 2$ we can write U_i^* as

$$\begin{pmatrix} \ddots & & & & & & \\ 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & 1 & & & \\ & 0 & 0 & 0 & 1 & & \\ & & 0 & 0 & 0 & 1 & \\ & & & 0 & 0 & 0 & 0 \\ & & & & 0 & & \ddots \end{pmatrix} \tag{4.6}$$

Recall that we defined $T_i w(j) = w(j - i)$. Now we are going to show that

$$H_{T_i\omega} = U_i H_\omega U_i^* \tag{4.7}$$

holds if we take $T_i\omega$, U_i, and U_i^* are as defined above. Rather than giving the proof in its full generality, we are going to consider what happens in the case when $i = 2$. Checking that the identity holds for any i is analogous.
T_i is a matrix that for $i = 2$ looks like this:

$$\begin{pmatrix} \ddots & & & & & & \\ 0 & 0 & 0 & & & & \\ 1 & 0 & 0 & 0 & & & \\ & 1 & 0 & 0 & 0 & & \\ & & 1 & 0 & 0 & 0 & \\ & & & 1 & 0 & 0 & 0 \\ & & & & 1 & & \ddots \end{pmatrix} \tag{4.8}$$

and $T_2 w$ can be represented as

$$\begin{pmatrix} \vdots \\ \omega(-4) \\ \omega(-3) \\ \omega(-2) \\ \omega(-1) \\ \omega(0) \\ \vdots \end{pmatrix} \tag{4.9}$$

where $\omega(-2)$ is at position (0,0).
Consider $H_{T_i\omega}$. For $i = 2$, we can represent it as a matrix as follows:

$$\begin{pmatrix} \ddots & & & & & & & \\ 1 & v_{-4}(\omega) & 1 & & & & & \\ & 1 & v_{-3}(\omega) & 1 & & & & \\ & & 1 & v_{-2}(\omega) & 1 & & & \\ & & & 1 & v_{-1}(\omega) & 1 & & \\ & & & & 1 & v_0(\omega) & 1 & \\ & & & & & & \ddots & \end{pmatrix} \tag{4.10}$$

where $v_{-2}(\omega)$ is at position (0,0).

In order to show that $H_{T_i\omega} = U_i H_\omega U_i^*$ holds, it is sufficient to show that $H_{T_i\omega}$ and $U_i H_\omega U_i^*$ act in the same way on each of the basis vectors of $\mathbb{R}^{\ltimes}$. For instance, consider the vector e_0 with 1 in position 0 and 0's everywhere else, and let us consider the case when we are applying first U_2^*, then H_ω, and then U_2. Applying U_2^* on e_0 yields e_{-2}, i.e. the basis vector with 1 in position $j = -2$ and 0's everywhere else, since U_2^* causes an upward shift of the vector by 2 positions. Next, calculating $H_\omega U_2^* e_0$ yields a vector with 1 in position $j = -3$ and $j = -1$, and $v_{-2}(\omega)$ in position $j = -2$.
Note: recall the way H_ω acts on a vector u_i:

$$H_\omega \begin{pmatrix} \vdots \\ u_{-2} \\ u_{-1} \\ u_0 \\ u_1 \\ u_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ u_{-2}v_{-2}(\omega) + u_{-3} + u_{-1} \\ u_{-1}v_{-1}(\omega) + u_{-2} + u_0 \\ u_0 v_0(\omega) + u_{-1} + u_1 \\ u_1 v_1(\omega) + u_0 + u_2 \\ u_2 v_2(\omega) + u_1 + u_3 \\ \vdots \end{pmatrix} \tag{4.11}$$

Finally, computing $U_2 \cdot H_\omega \cdot U_2^* e_0$ just shifts $H_\omega \cdot U_2^* \cdot e_0$ down by 2 positions, resulting in a vector with $v_{-2}(\omega)$ in position $j = 0$ and $1's$ in positions $j = -1$ and $j = 1$. We can summarize applying $U_2 H_\omega U_2^* e_0$ as follows:

$$\begin{pmatrix} \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots \\ 1 \\ v_{-2}(\omega) \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ v_{-2}(\omega) \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{4.12}$$

Now consider applying $H_{T_2\omega}e_0$.

$$\begin{pmatrix} \ddots & & & & & & & \\ 1 & v_{-4}(\omega) & 1 & & & & & \\ & 1 & v_{-3}(\omega) & 1 & & & & \\ & & 1 & v_{-2}(\omega) & 1 & & & \\ & & & 1 & v_{-1}(\omega) & 1 & & \\ & & & & 1 & v_0(\omega) & 1 & \\ & & & & & & & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ 0 \\ 0 \\ 0 \\ 1 \\ v_{-2}(\omega) \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{4.13}$$

Thus, we have shown that

$$H_{T_2\omega}e_0 = U_2 H_\omega U_2^* e_0 \tag{4.14}$$

The same approach can be used to show that the more general result

$$H_{T_i\omega}e_j = U_i H_\omega U_i^* e_j \tag{4.15}$$

holds for any index i and any basis vector e_j.
Thus we have shown that the random Schrödinger operator satisfies the condition

$$H_{T_i\omega} = U_i H_\omega U_i^* \tag{4.16}$$

where $\{T_i\}$ are the shift operators and U_i are the translation operators defined earlier. Therefore, the random Schrödinger operator is ergodic.

4.2 Connections with Ergodic Theory

As a next step towards proving that the spectrum of the random Schrödinger operator is deterministic, we are going to need the following theorem.

Theorem 4.2.1. *Consider an ergodic family of measure-preserving transformations $\{T_i\}$. Let $f : \Omega \to \mathbb{R}$ be a random variable which is invariant under $\{T_i\}$. Then f is constant P-a.s.*

Proof. It suffices to show that $\mathbb{P}(\{w|f(w) = M_0\}) = 1$ for some constant M_0. Define $\Omega_M = \{w|f(w) \leq M\} = f^{-1}\{(-\infty, M]\}.f$ is invariant under $\{T_i\}$, so $f(T_i w) = f(w)$, i.e. $f \circ T_i = f$. We are going to show that Ω_M is invariant under T_i. Take $x \in T_i^{-1}\Omega_M$. Therefore, $T_i x \in \Omega_M$. Thus, by definition, $f(T_i x) \leq M$, or $f \circ T_i(x) \leq M$. We know that $f \circ T_i = f$ since f is invariant under T_i. Therefore, $f(x) \leq M$ so $x \in \Omega_M$. Thus, we have shown that if $x \in T_i^{-1}\Omega_M$ then $x \in \Omega_M$, i.e. $T_i^{-1}\Omega_M \subset \Omega_M$. Now consider $y \in \Omega M$. Then $f(y) \leq M$ by definition of Ω_M. Since $f \circ T_i = f$, it follows that $f \circ T_i(y) \leq M$, which means that $T_i y \in \Omega_M$. Therefore, $y \in T_i^{-1}\Omega_M$. Thus, we have

shown that if $y \in \Omega_M$ then $y \in T_i^{-1}\Omega_M$, which means that $\Omega_M \subset T_i^{-1}\Omega_M$. We have shown double inclusion, so it follows that $T_i^{-1}\Omega_M = \Omega_M$, i.e. that Ω_M is invariant under T_i.
Since $\{T_i\}$ are ergodic and Ω_M is invariant, $P(\Omega_M) = 1$ or $P(\Omega_M) = 0$.
Consider $M_1 < M_2$. Then $\{w|f(w) \leq M_1\} \subset \{w|f(w) \leq M_2\}$, which means that $\Omega_{M_1} \subset \Omega_{M_2}$.
Also notice that

Lemma 4.2.2. $\bigcup_{M\in\mathbb{R}} \Omega_M = \bigcup_{M\in\mathbb{Z}} \Omega_M$.

Proof. We prove this by double inclusion. Clearly $\bigcup_{M\in\mathbb{Z}} \Omega_M \subset \bigcup_{M\in\mathbb{R}} \Omega_M$ because $\mathbb{Z} \subset \mathbb{R}$. Now we show that $\bigcup_{M\in\mathbb{R}} \Omega_M \subset \bigcup_{M\in\mathbb{Z}} \Omega_M$. Let $x \in \bigcup_{M\in\mathbb{R}} \Omega_M$. Therefore, $x \in \Omega_{M'}$ for some $M' \in R$. But there exists an integer M'' that is greater than M', so $\Omega_{M''} \in \bigcup_{M\in\mathbb{Z}} \Omega_M$. However, $\Omega_{M'} \subset \Omega_{M''}$, so $\Omega_{M'} \in \bigcup_{M\in\mathbb{Z}} \Omega_M$. Since $x \in \Omega_{M'}, x \in \bigcup_{M\in\mathbb{Z}} \Omega_M$. Thus, it follows that $\bigcup_{M\in R} \Omega_M = \bigcup_{M\in Z} \Omega_M$. □

Similarly, it can be shown that $\bigcap\limits_{M\in\mathbb{R}} \Omega_M = \bigcap\limits_{M\in\mathbb{Z}} \Omega_M$.

Clearly, $\mathbb{P}(\bigcap_{M\in\mathbb{R}} \Omega_M) = 0$. Suppose that $\mathbb{P}(\bigcap_{M\in\mathbb{R}} \Omega_M) = \alpha \neq 0$. Since $\Omega_M \supset \bigcap_{M\in\mathbb{R}} \Omega_M$ for any M, $\mathbb{P}(\Omega_M) > \mathbb{P}(\bigcap_{M\in\mathbb{R}} \Omega_M)$, i.e. $\mathbb{P}(\Omega_M) > \alpha > 0$ for any M. But we have shown earlier that $\mathbb{P}(\Omega_M)$ is either 0 or 1, which implies that $\mathbb{P}(\Omega_M) = 1$ for any M. However, this contradicts the fact that $f(w)$ is finite. Suppose that $\mathbb{P}(\Omega_M) = 1$. Then $\mathbb{P}(\{w|f(w) \leq M\}) = 1$ for all M, which implies that $f(w) \leq M$ for all M. This in turn implies that $f(w) = -\infty$ a.s. Thus, we have found a contradiction and shown that $\mathbb{P}(\bigcap_{M\in\mathbb{R}} \Omega_M) = 0$.

Define

$$M_0 = \inf_{\mathbb{P}(\Omega_M)=1} M. \tag{4.17}$$

Lemma 4.2.3. M_0 *is finite.*

Proof. Suppose that M_0 is infinite. Then either $M_0 = \infty$ or $M_0 = -\infty$. Suppose that $M_0 = \infty$. Then $\mathbb{P}(\Omega_M) \neq 1$ for all M. Since $\mathbb{P}(\Omega_M)$ is either 0 or 1, this implies that $\mathbb{P}(\Omega_M) = 0$ for all M. Thus, $\{\omega|f(\omega) \leq M)\}$ has measure 0 for any M, which implies that $f = \infty$ a.s. - a contradiction. Therefore, $M_0 \neq \infty$.
Now suppose that $M_0 = -\infty$. Then $\mathbb{P}(\Omega_M) = 1$ for all M, and the proof continues the same way as the proof that $\mathbb{P}(\bigcap_{M\in\mathbb{R}} \Omega_M) = 0$. Thus, we have established that M_0 is finite. □

Notice that

$$\Omega_{M_0} = \bigcap_{n\in\mathbb{N}} \Omega_{M_0+(\frac{1}{n})} \tag{4.18}$$

Define

$$\tilde{\Omega}_{M_0} = \{\omega|f(\omega) < M_0\}. \tag{4.19}$$

Notice also that

$$\tilde{\Omega}_{M_0} = \bigcup_{n\in\mathbb{N}} \Omega_{M_0-\frac{1}{n}}. \tag{4.20}$$

Clearly, $\mathbb{P}(\Omega_{M_0+\frac{1}{n}}) = 1$ for each $n \in \mathbb{N}$ because $M_0 + \frac{1}{n} > M_0$ and $M_0 = \inf_{\mathbb{P}(\Omega_M)=1} M$. Moreover, $\mathbb{P}(\Omega_{M_0} - \frac{1}{n}) = 0$ because $\mathbb{P}(\Omega_{M_0} - \frac{1}{n}) \neq 1$ and each set Ω_M has probability 0 or 1. Therefore,

$$\mathbb{P}(\Omega_{M_0}) = \mathbb{P}(\bigcap_{n\in\mathbb{N}} \Omega_{M_0+\frac{1}{n}}) = 1 \tag{4.21}$$

and

$$\mathbb{P}(\tilde{\Omega}_{M_0}) = \mathbb{P}(\bigcup_{n\in\mathbb{N}} \Omega_{M_0-\frac{1}{n}}) = 0. \tag{4.22}$$

Because $\Omega_{M_0} = \{\omega | f(\omega) \leq M_0\}$ and $\tilde{\Omega}_{M_0} = \{\omega | f(\omega) < M_0\}$, we see that $\{\omega | f(\omega) = M_0\} = \Omega_{M0} \backslash \tilde{\Omega}_{M_0}$. Thus,

$$\mathbb{P}(\{\omega | f(\omega) = M_0\}) = \mathbb{P}(\Omega_{M_0} \backslash \tilde{\Omega}_{M_0}) = \mathbb{P}(\Omega_{M_0}) - \mathbb{P}(\tilde{\Omega}_{M_0}) = 1 - 0 = 1 \tag{4.23}$$

since $\Omega_{M_0} \supset \tilde{\Omega}_{M_0}$. Thus, we have shown that f is constant $\mathbb{P}$-a.s. □

In order to prove that the spectrum of the random Schrödinger operator is deterministic, we are going to need another result concerning the dimension of the range of orthogonal projections P_ω satisfying the condition for ergodicity. The next chapter introduces some more tools from functional analysis that we need in order to prove the desired lemma.

Chapter 5

More Results from Functional Analysis

Before we prove our next lemma, we are going to introduce some definitions and theorems about projections.

5.1 More Definitions

First we define what it means for an operator to be a projection.

Definition 28. *If P is a bounded linear operator on a Hilbert space $\mathbb{H}$ and $P^2 = P$, then P is called a projection. If also $P = P^*$, then P is called an orthogonal projection.*

The projection is a positive operator as defined below.

Definition 29. *An operator B in a Hilbert space H is called positive if*

$$\langle Bx, x \rangle \geq 0 \tag{5.1}$$

for all $x \in H$. Any positive operator is self-adjoint.

Lemma 5.1.1. *(Square root lemma) Let A be a positive operator. Then there is a unique positive operator B such that $B^2 = A$. We can also write $B = A^{\frac{1}{2}}$. B commutes with every bounded operator which commutes with A.*

We are also going to need Parseval's identity, which is a very commonly used tool in functional analysis.

Theorem 5.1.2. *(Parseval's identity) Let $\mathbb{H}$ be a Hilbert space and $\{\phi_n\}$ be an orthonormal basis. Then for each $y \in \mathbb{H}$,*

$$y = \sum_n |\langle \phi_n, y \rangle| \phi_n \tag{5.2}$$

and

$$\|y\|^2 = \sum_n |\langle \phi_n, y \rangle|^2 \tag{5.3}$$

Another important concept is the trace of an operator.

Definition 30. *Let $\mathbb{H}$ be a Hilbert space, and let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal basis. Then for any positive operator A we define the trace of A:*

$$\mathrm{tr}A = \sum_{n=1}^{\infty} \langle \phi_n, A\phi_n \rangle. \tag{5.4}$$

The trace is a linear map $A \to \mathrm{tr}A$ in the sense that $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(cA) = c \cdot \mathrm{tr}(A)$ where c is a constant. It has some nice properties such as commutativity: if A and B are two positive operators, then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. We are going to need this property later in proving a claim about the density of states.

A fundamental result regarding the trace of an operator is that it does not depend on the choice of orthonormal basis. The proof of this result is standard and can be found for example in [4]. It is included below because it illustrates several important techniques that are commonly employed in manipulations of expressions we will encounter later.

Lemma 5.1.3. *The trace of an operator A does not depend on the choice of orthonormal basis.*

Proof. Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal basis. We can define $\mathrm{tr}_\phi(A) = \sum_{n=1}^{\infty} \langle \phi_n, A\phi_n \rangle$. Let $\{\psi_m\}_{m=1}^{\infty}$ be another orthonormal basis. Then, since A is positive, we can write

$$\sum_{n=1}^{\infty} \langle \phi_n, A\phi_n \rangle = \sum_{n=1}^{\infty} \langle \phi_n, B^2\phi_n \rangle$$

using the property of the adjoint

$$= \sum_{n=1}^{\infty} \langle B^*\phi_n, B\phi_n \rangle$$

since B is a positive operator and is therefore self-adjoint

$$= \sum_{n=1}^{\infty} \langle B\phi_n, B\phi_n \rangle$$

$$= \sum_{n=1}^{\infty} \|B\phi_n\|^2$$

using Parseval's identity

$$= \sum_{n=1}^{\infty} \Big(\sum_{m=1}^{\infty} |\langle \psi_m, B\phi_n \rangle|^2 \Big)$$

interchanging the order of the summations since all terms are positive

$$= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |\langle \psi_m, B\phi_n\rangle|^2$$

again using the property of the adjoint and the fact that B is self-adjoint

$$= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |\langle B\psi_m, \phi_n\rangle|^2$$

again using Parseval's identity

$$= \sum_{m=1}^{\infty} \|B\psi_m\|^2$$

$$= \sum_{m=1}^{\infty} \langle B\psi_m, B\psi_m\rangle$$

again using the property of the adjoint

$$= \sum_{m=1}^{\infty} \langle \psi_m, B^*B\psi_m\rangle$$

using the fact that $B^* = B$ since B is self-adjoint

$$= \sum_{m=1}^{\infty} \langle \psi_m, B^2\psi_m\rangle$$

$$= \sum_{m=1}^{\infty} \langle \psi_m, A\psi_m\rangle$$

□

Another important result is the fact that if two operators are unitarily equivalent, they have the same trace.

Lemma 5.1.4. *If A is a positive operator and U is a unitary operator,*

$$\mathrm{tr}(UAU^{-1}) = \mathrm{tr}(A) \tag{5.5}$$

Proof. Note that if $\{\phi_n\}$ is an orthonormal basis then so is $\{U\phi_n\}$. Since the trace is independent of the choice of basis,

$$\begin{aligned}
\mathrm{tr}(UAU^{-1}) &= \mathrm{tr}_{(U\phi)}(UAU^{-1}) \\
&= \sum \langle U\phi, UAU^{-1}U\phi\rangle \\
&= \sum \langle U\phi, UA\phi\rangle \\
&= \sum \langle U^*U\phi, A\phi\rangle \quad \text{(using the adjoint property)} \\
&= \sum \langle \phi, A\phi\rangle \qquad \text{(since } U \text{ is unitary)} \\
&= \mathrm{tr}_{(\phi)}(A) = \mathrm{tr}(A)
\end{aligned}$$

□

Lemma 5.1.5. *(Projection theorem) Let $\mathbb{H}$ be a Hilbert space, $\mathbb{M}$ a closed subspace. Then every $x \in \mathbb{H}$ can be uniquely written $x = z + w$ where $z \in \mathbb{M}$ and $w \in \mathbb{M}^{\perp}$.*

Note that if x is decomposed as in the above definition, then $Px = z, Pz = z$, and $Pw = 0$.

Now we have all the tools we need in order to prove the next theorem, which will bring us one step closer to characterizing the spectrum of the random Schrödinger operator.

Lemma 5.1.6. *Suppose that $\{P_\omega\}_{\omega\in\Omega}$ is a weakly measurable family of orthogonal projections satisfying*

$$P_{T_i\omega} = U_i P_\omega U_i^* \tag{5.6}$$

where U_i is a unitary operator. Then $\dim \operatorname{Ran}(P_\omega)$ is either zero or infinite $\mathbb{P}$-a.s.

Proof. We fix ω and choose an orthonormal basis $\{a_1(\omega), a_2(\omega), \ldots\}$ of $\operatorname{Ran}(P_\omega)$ and an orthonormal basis $\{b_1(\omega), b_2(\omega), \ldots\}$ of $\operatorname{Ran}(P_\omega)^{\perp}$. Then by the projection theorem the orthonormal basis for $\operatorname{Ran}(P_\omega)$ and the orthonormal basis for $\operatorname{Ran}(P_\omega)^{\perp}$ span the entire space $\mathbb{H}$ and thus form an orthonormal basis for it. Since $\{b_i(\omega)\}$ is in $\operatorname{Ran}(P_\omega)^{\perp}$, we know that $P_\omega b_i(\omega) = 0$. Also, since $\{a_i(\omega)\}$ is in $\operatorname{Ran}(P_\omega)$, we know that $P_\omega a_i(\omega) = a_i(\omega)$. Thus, we get

$$\mathrm{tr} P_\omega \;=\; \sum \langle a_i(\omega), P_\omega a_i(\omega)\rangle + \sum \langle b_i(\omega), P_\omega b_i(\omega)\rangle = \sum \langle a_i(\omega), a_i(\omega)\rangle = \dim \operatorname{Ran}(P_\omega)$$

The last equality follows from the fact that $\{a_i(\omega)\}$ is an orthonormal basis for $\operatorname{Ran}(P_\omega)$.

Now let $\{e_i, i\in Z\}$ be the standard orthonormal basis in $l^2(Z)$, i.e. $e_i(n) = \delta_{in}$.Note that $\mathrm{tr} P_\omega = \sum \langle e_i, P_\omega e_i\rangle$is a random variable and is therefore measurable.
Recall that the projections P_ω satisfy

$$P_{T_i\omega} = U_i P_\omega U_i^* \tag{5.7}$$

where U_i is a unitary operator defined as

$$(U_i u)(n) = u(n - i) \tag{5.8}$$

and U_i^* is its adjoint defined as

$$(U_i^* u)(n) = u(n + i) \tag{5.9}$$

Using that fact, we can write

$$\mathrm{tr}(P_{T_j\omega}) = \mathrm{tr}(U_j P_\omega U_j^*) = \mathrm{tr}(U_j P_\omega U_j^{-1}) = \mathrm{tr}(P_\omega)$$

where for the last equality we used Lemma 2.1.11. Thus, we have shown that the random variable $\mathrm{tr}(P_\omega)$ is invariant under T_i, and is therefore constant $\mathbb{P}$-a.s. by Theorem 2.1.5. Since we also showed that $\mathrm{tr}(P_\omega) = \dim \mathrm{Ran}(P_\omega)$, it follows that

$$\dim \mathrm{Ran}(P_\omega) = \mathrm{tr}(P_\omega) \tag{5.10}$$

is constant $\mathbb{P}$-a.s. Hence $\mathbb{P}$-a.s.:

$$\begin{aligned}
\mathrm{tr}P_\omega = \mathbb{E}(\mathrm{tr}P_\omega) = \mathbb{E}\left(\sum_{i=1}^{\infty}\langle e_i, P_\omega e_i\rangle\right) &= \sum_{i=1}^{\infty}\mathbb{E}(\langle e_i, P_\omega e_i\rangle) \\
&\geq \sum_{|i|\leq N}\mathbb{E}(\langle e_i, P_\omega e_i\rangle) = \sum_{|i|\leq N}\mathbb{E}(\langle e_i, U_i^* P_{T_i\omega} U_i e_i\rangle) \\
&= \sum_{|i|\leq N}\mathbb{E}(\langle U_i e_i, P_{T_i\omega} U_i e_i\rangle) \qquad (5.11) \\
&= \sum_{|i|\leq N}\mathbb{E}(\langle e_0, P_{T_i\omega} e_0\rangle) \\
&= \sum_{|i|\leq N}\mathbb{E}(\langle e_0, P_\omega e_0\rangle) = (2N+1)\,\mathbb{E}(\langle e_0, P_\omega e_0\rangle) \qquad (5.12)
\end{aligned}$$

Since N was arbitrary, letting $N \to \infty$ yields $\mathrm{tr}P_\omega = 0$ if $\mathbb{E}(\langle e_0, P_\omega e_0\rangle) = 0$ or $\mathrm{tr}P_\omega = \infty$ if $\mathbb{E}(\langle e_0, P_\omega e_0\rangle) > 0$. $\square$

An outline of this proof is given in [1]. Part of my contribution has been to provide the details of the proof.

In the next chapter we present the functional calculus and the spectral theorem - two very powerful results that are commonly used in various areas of mathematics.

Chapter 6

The Functional Calculus and the Spectral Theorem

6.1 The Functional Calculus

The functional calculus allows us to define functions of self-adjoint operators, and thus to give meaning to expressions such as $f(A)$ where A is an operator. There are different versions of the functional calculus, reaching different degrees of generality regarding what the function f can be. The one that suffices for us is the continuous functional calculus.

Definition 31. *(Continuous Functional Calculus) Let A be a self-adjoint operator on a Hilbert space $\mathbb{H}$. Then there is a unique map $\phi : C(\sigma(A)) \to \mathcal{L}(\mathbb{H}), \phi \mapsto \phi(A)$ with the following properties:*

(i) ϕ is a homomorphism, i.e. $\phi(fg) = \phi(f)\phi(g), \phi(\lambda f) = \lambda\phi(f), \phi(1) = I, \phi(\bar{f}) = \phi(f)^*$

(ii) ϕ is continuous, i.e. $||\phi(f)||_{\mathcal{L}(\mathbb{H})} \leq C||f||_\infty$

(iii) Let f be the function $f(x) = x$; then $\phi(f) = A$.

Moreover, ϕ has the following properties:

(iv) If $A\psi = \lambda\psi$, then $\phi(f)\psi = f(\lambda)\psi$

(iv) (Spectral Mapping Theorem) $\sigma[\phi(f)] = \{f(\lambda) \big| \lambda \in \sigma(A)\}$

Note that one should interpret $\phi(f)$ as defining a function of the operator A, i.e. as $f(A)$. Therefore, we sometimes write $f(A)$ for $\phi(f)$. Thus, property (i) should be interpreted as saying that

$$
\begin{aligned}
&(\alpha f + \beta g)(A) = \alpha f(A) + \beta g(A) \\
&f \cdot g(A) = f(A)g(A) \\
&\bar{f}(A) = f(A)^*
\end{aligned}
$$

Moreover, if $f \geq 0$ then $\langle \phi, f(A)\phi \rangle \geq 0$ for all $\phi \in \mathcal{H}$. $\langle \phi, f(A)\phi \rangle$ is called a positive linear functional. The definition of this concept is given below.

Definition 32. *Let V be a vector space over a field $\mathbb{F}$. A linear functional on V is a mapping $T : V \to \mathbb{F}$ satisfying the following conditions:*
1) For every $x, y \in V$, $T(x + y) = T(x) + T(y)$;
2) For every $c \in \mathbb{F}$ and $x \in V$ we have $T(cx) = cT(x)$.

Theorem 6.1.1. *(Riesz Representation Theorem) Let X be a compact Hausdorff space, and let $C(X)$ denote the space of continuous real-valued functions on X. For any positive linear functional ϕ on $C(X)$ there is a unique measure μ on X with*

$$\phi(f) = \int_X f(x)d\mu(x) \tag{6.1}$$

The Riesz representation theorem allows one to construct a measure from any positive linear functional. Thus, the Riesz representation theorem guarantees that for each $\varphi \in \mathcal{H}$ there is a positive and bounded measure μ_φ on $\sigma(A)$ such that for all $f \in \mathcal{C}(\sigma(A))$,

$$\langle \varphi, f(A)\varphi \rangle = \int f(\lambda)d\mu_\varphi(\lambda). \tag{6.2}$$

Definition 33. *The measure μ_ψ is called the spectral measure associated with the vector ψ.*

One can use the polarization identity to find complex-valued measures $\mu_{\phi,\psi}$ for $\phi, \psi \in \mathcal{H}$ such that

$$\langle \phi, f(A)\psi \rangle = \int_{\sigma(A)} f(\lambda)d\mu_{\phi,\psi} \tag{6.3}$$

This equation can be used to define the operator $f(A)$ for bounded measurable functions. The properties of the functional calculus allow us to use the same rules that apply to ordinary functions on the real line in order to perform calculations involving a self-adjoint operator. In this exposition we are going to use property (iv), the spectral mapping theorem, especially often. We can express it compactly in the following more intuitive way:

$$\sigma\big(\varphi(A)\big) = \varphi\big(\sigma(A)\big) \tag{6.4}$$

For any Borel set $M \subset \mathbb{R}$ we denote by χ_M the characteristic function of M, $\chi_M(\lambda)$, where $\chi_M(\lambda) = 1$ if $\lambda \in M$, and is equal to 0 otherwise.
The operators $\mu(A) = \chi_M(A)$ are very important and are called spectral projections. One can easily check that they are projections. In particular, they satisfy the following conditions

$$\begin{gathered} \mu(A) \text{ is an orthogonal projection} \\ \mu(\emptyset) = 0 \text{ and } \mu\left(\sigma(A)\right) = 1 \\ \mu(M \cap N) = \mu(M)\mu(N) \end{gathered}$$

If the Borel sets M_n are pairwise disjoint, then for each $\phi \in \mathcal{H}$,

$$\mu\left(\cup_{n=1}^{\infty} M_n\right)\varphi = \sum_{n=1}^{\infty} \mu(M_n)\varphi \tag{6.5}$$

$\mu(A) = \chi_M(A)$ is called the projection-valued measure associated with the operator A because it satisfies the above conditions which are analogous to those for a measure. $\chi_M(A)$ is also called the projection valued spectral measure of A.

Definition 34. *Let A be a bounded self-adjoint operator and Ω a Borel set of $\mathbb{R}$. The spectral projection P_Ω of A is defined as $\chi_\Omega(A)$ where $\chi_\Omega(A)$ is the characteristic function of $\sigma(A)$ on the set Ω.*

Taking the dimension of the spectral projection of a self-adjoint operator on an interval Δ is equivalent to counting the eigenvalues in that interval. This can be done if the operator has discrete spectrum.

Definition 35. *The discrete spectrum σ_{disc} of an operator A consists of all $\lambda \in \sigma(A)$ such that $P_{(\lambda-\epsilon,\lambda+\epsilon)}(A)$ is finite-dimensional for some $\epsilon > 0$, where $P(A)$ denotes the spectral projection of A.*

Theorem 6.1.2. $\lambda \in \sigma_{disc}(A)$ *if and only if both λ is an isolated point of $\sigma(A)$ (i.e. for some $\epsilon, (\lambda - \epsilon, \lambda + \epsilon) \cap \sigma(A) = \{\lambda\}$, and λ is an eigenvalue of finite multiplicity.*

In the next section, we introduce the spectral theorem which is also a very fundamental result. We also elaborate on the spectral calculus.

6.2 The Spectral Theorem

The spectral theorem is extremely important in various areas of mathematics and has many alternative formulations. In a first linear algebra course it comes up as the fact that every symmetric matrix can be diagonalized. The version that we are going to use is a generalization to bounded self-adjoint operators. We state it below

Theorem 6.2.1. *For any self-adjoint operator $A : \mathcal{H} \to \mathcal{H}$ with cyclic vector ψ, there exists a measure μ on $\mathbb{R}$ and a unitary operator $U : \mathcal{H} \to L^2(\mathbb{R}, d\mu)$ such that $UAU^{-1} = M$ where $M : L^2(\mathbb{R}, d\mu) \to L^2(\mathbb{R}, d\mu)$ is multiplication by x, i.e. $Mf(x) = xf(x)$*

Definition 36. *A spectral resolution is a mapping $E : \mathbb{R} \to \mathcal{L}(\mathcal{H})$ with the following properties:*
(i) $E(t)$ *is an orthogonal projection for every* $t \in \mathbb{R}$.
(ii) $E(s) \leq E(t)$ *for* $s \leq t$ *(monotonicity)*
(iii) $\lim_{\epsilon\to 0} E(t + \epsilon)x = E(t)x$ *for all* $t \in \mathbb{R}, x \in \mathcal{H}$ *(right continuity)*
(iv) $\lim_{t\to-\infty} E(t)x = 0$ *and* $\lim_{t\to\infty} E(t)x = x$ *for all* $x \in \mathcal{H}$.

One of the many equivalent versions of the spectral theorem says that every self-adjoint operator H is of the form $H = \int_{-\infty}^{\infty} tdE(t)$ for some spectral resolution E.

In the proof of Pastur's theorem, we are going to need the following result which follows from the functional calculus.

Lemma 6.2.2. *Let A be a self-adjoint operator and U a unitary operator. Then for any bounded measurable function f we have*

$$f(UAU^*) = Uf(A)U^* \tag{6.6}$$

Also note that if μ and ν are projection valued measures for A and $B = UAU^*$ respectively, then we have

$$\int f(\lambda)d\nu_{\phi,\psi}(\lambda) = \langle \phi, f(B)\psi \rangle = \langle \phi, Uf(A)U^*\psi \rangle = \langle U^*\phi, f(A)U^*\psi \rangle$$
$$= \int f(\lambda)d\mu_{U^*\phi,U^*\psi}(\lambda)$$

Thus, the measures $\mu_{\phi,\psi}$ and $\nu_{U^*\phi,U^*\psi}$ agree.

Chapter 7

The Spectrum of the Random Schrödinger Operator

Now we have all the tools we need to give a complete proof of Pastur's theorem about the spectrum of the random Schrödinger operator.

7.1 Proving that the spectrum is deterministic

Theorem 7.1.1. *Let V_ω be an ergodic potential. Then*

$$\begin{aligned} \sigma(H_\omega) \ \textit{is constant}\ \mathbb{P} - a.s. \\ \sigma_{disc}(H_\omega) = \varnothing\ \mathbb{P} - a.s. \end{aligned}$$

Proof. Let us denote the spectral projection of H_ω by $E_\Delta(\omega)$ where Δ is an interval. Note that the spectral projections satisfy the following identity

$$E_\Delta(T_i\omega) = U_i E_\Delta(\omega) U_i^* \tag{7.1}$$

where U_i is a unitary operator. This identity follows from the functional calculus, which allows us to apply a function f to an operator. Recall that

$$H_{T_i\omega} = U_i H_\omega U_i^* \tag{7.2}$$

Applying the characteristic function χ_Δ on both sides of this equation, we get

$$\chi_\Delta(H_{T_i\omega}) = \chi_\Delta(U_i H_\omega U_i^*) = \chi_\Delta(U_i)\chi_\Delta(H_\omega)\chi_\Delta(U_i^*) \tag{7.3}$$

where $\chi_\Delta(A)$ denotes the characteristic function of $\sigma(A)$ on the set Δ. By definition, χ_Δ is a spectral projection. Thus, we can rewrite this as

$$E_\Delta(T_i\omega) = U_i E_\Delta(\omega) U_i^* \tag{7.4}$$

where the last equality follows from the functional calculus because $\chi_\Delta(U_i) = P_\Delta(U_i) = U_i$, and similarly for U_i^* since U_i and U_i^* are unitary operators.

Note that H_ω^n is weakly measurable since products of bounded, weakly measurable functions are weakly measurable. Therefore, $\{E_\Delta(\omega)\}_{\omega\in\Omega}$ is a weakly measurable family of orthogonal projections, since they can be approximated by ω-independent polynomials in H_ω. Thus, $\{E_\Delta(\omega)\}_{\omega\in\Omega}$ satisfy the conditions of lemma 2.1.7, and therefore,

$$\dim \operatorname{Ran}\left(E_\Delta(\omega)\right) = 0 \;\; \mathbb{P}\text{-a.s. or} \dim \operatorname{Ran}\left(E_\Delta(\omega)\right) = \infty \;\; \mathbb{P}\text{-a.s.} \tag{7.5}$$

We also define a function η for any $p, q \in \mathbb{Q}$:

$$\eta(p,q) = 0 \quad \text{if} \dim \operatorname{Ran}\left(E_{(p,q)}(\omega)\right) = 0 \;\mathbb{P}\text{-a.s.}$$
$$\eta(p,q) = \infty \quad \text{if} \dim \operatorname{Ran}\left(E_{(p,q)}(\omega)\right) = \infty \;\mathbb{P}\text{-a.s.}$$

(2.35) guarantees that $\eta(p,q)$ is well-defined.

Define $\Omega_{(p,q)} = \{\omega | \dim \operatorname{Ran}\left(E_{(p,q)}(\omega)\right) = \eta(p,q)\}$. From the definition of η we can see that $\eta(p,q) = \dim \operatorname{Ran}\left(E_{(p,q)}(\omega)\right)$ $\mathbb{P}$-a.s. Therefore, $\mathbb{P}\left(\Omega_{(p,q)}\right) = 1$ for any $p, q \in \mathbb{Q}$. Set $\Omega_0 = \cap_{p,q\in\mathbb{Q}}\Omega_{p,q}$. Since $\mathbb{Q}$ is countable, Ω_0 is a countable intersection of sets with probability 1. Therefore, $\mathbb{P}\left(\Omega_0\right) = 1$ We claim that

Lemma 7.1.2. *If ω_1 and $\omega_2 \in \Omega_0$, then $\sigma(H_{\omega_1}) = \sigma(H_{\omega_2})$.*

Proof. We give a proof by contrapositive. First, suppose that $\lambda \notin \sigma(H_{\omega_1})$. By the definition of spectral projection, $E_{(\lambda_1,\lambda_2)}$ is a characteristic function on (λ_1, λ_2), taking the value 1 if (λ_1, λ_2) contains a value in $\sigma(H_\omega)$ and 0 if it doesn't.

Then, since we assume $\lambda \notin \sigma(H_\omega)$, for all λ_1, λ_2 with $\lambda_1 < \lambda < \lambda_2$ sufficiently close to λ,

$$E_{(\lambda_1,\lambda_2)}(\omega_1) = 0 \Rightarrow \dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega_1) = 0 \tag{7.6}$$

By assumption, $\omega_1, \omega_2 \in \Omega_0$. Therefore, $\omega_1, \omega_2 \in \cap_{p,q\in\mathbb{Q}}\Omega_{p,q}$, i.e. $\omega_1, \omega_2 \in \Omega_{p,q}$ for every $p, q \in \mathbb{Q}$. Thus, for each pair $(p,q) \in \mathbb{Q}$, $\dim \operatorname{Ran} E_{(p,q)}(\omega_1) = \eta(p,q)$ and $\dim \operatorname{Ran} E_{(p,q)}(\omega_2) = \eta(p,q)$, so $\dim \operatorname{Ran} E_{(p,q)}(\omega_1) = \dim \operatorname{Ran} E_{(p,q)}(\omega_2)$ for every pair $(p,q) \in \mathbb{Q}$. In particular, for $\lambda_1, \lambda_2 \in \mathbb{Q}$ sufficiently close to λ $(\lambda_1 < \lambda < \lambda_2)$, $\dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega_1) = \dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega_2)$. Using (2.34), we get

$$\dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega_1) = \dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega_2) = 0 \tag{7.7}$$

for λ_1, λ_2 sufficiently close to λ. It follows that

$$E_{(\lambda_1,\lambda_2)}(\omega_2) = 0 \tag{7.8}$$

Since by the definition of spectral projection, $E_{(\lambda_1,\lambda_2)}$ is the characteristic function of $\sigma(H_\omega)$ on the interval (λ_1, λ_2), it follows that $\lambda \notin \sigma(H_{\omega_2})$. Thus, we have shown that if $\lambda \notin \sigma(H_{\omega_1})$, then $\lambda \notin \sigma(H_{\omega_2})$. Taking the contrapositive, we get that $\lambda \in \sigma(H_{\omega_2})$ implies $\lambda \in \sigma(H_{\omega_1})$, i.e. $\sigma(H_{\omega_2}) \subset \sigma(H_{\omega_1})$. Interchanging ω_1 and ω_2, we similarly get that $\sigma(H_{\omega_1}) \subset \sigma(H_{\omega_2})$. Thus, we have shown that $\sigma(H_{\omega_1}) = \sigma(H_{\omega_2})$. □

Since ω_1 and ω_2 are arbitrary elements of Ω_0 which has probability 1, it follows that

$$\sigma(H_\omega) \text{ is constant } \mathbb{P} - a.s.$$

This finishes the proof of the first part of Pastur's theorem.

Now we prove that $\sigma_{disc}(A)$ is empty. Suppose that $\lambda \in \sigma_{disc}(H_\omega)$ for some $\omega \in \Omega_0$. Therefore, by theorem 2.1.13, λ is an eigenvalue of finite multiplicity. Then if we take $\lambda_1 < \lambda < \lambda_2$ sufficiently close to λ, the dimension of the projection of H_ω on the interval (λ_1, λ_2) should equal the multiplicity of λ:

$$0 < \dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega) < \infty \tag{7.9}$$

But we showed that for $\omega \in \Omega_0$, $\dim \operatorname{Ran} E_{(p,q)}(\omega) = 0$ for every $p, q \in \mathbb{Q}$. Therefore, we must have $\dim \operatorname{Ran} E_{(\lambda_1,\lambda_2)}(\omega) = 0$, which contradicts equation 2.37. This proves that $\sigma_{disc}(H_\omega) = \varnothing$.

□

Thus, we have proved

Theorem (Pastur) Let V_ω be an ergodic potential. Then there exists a set $\Sigma \subset \mathbb{R}$ such that $\sigma(H_\omega) = \Sigma$ $\mathbb{P} - a.s.$ Moreover, $\sigma_{dis}(H_\omega) = \varnothing$ $\mathbb{P} - a.s.$

Kunz and Souillard obtained with probability one, in various situations, the exact location of the spectrum, and criteria for a given part in the spectrum to be pure point, absolutely continuous, or singularly-continuous. We would also like to study and give a detailed proof of this result. Thus, in the next section we are going to look at the spectral decomposition of the random Schrödinger operator.

7.2 Decomposing the Spectrum

Consider the spectral measure μ_h corresponding to an element $h \in H$, and defined on $\sigma(T)$. Suppose μ is a Borel measure on $\mathbb{R}$. Lebesgue's decomposition theorem says that this measure can be decomposed into several distinct measures:

Theorem 7.2.1. *(Lebesgue Decomposition Theorem) Let μ be a measure on $\mathbb{R}$. Then μ can be uniquely written as $\mu = \mu_{ac} + \mu_{sing} + \mu_{pp}$, i.e. μ can be decomposed as a sum of a measure which is absolutely continuous w.r.t Lebesgue, a measure which is singular w.r.t Lebesgue, and a pure point measure.*

Definition 37. *A Borel measure μ on $\mathbb{R}$ is called continuous if it has no pure points. μ is called a pure point measure if $\mu(X) = \sum_{x \in X} \mu(x)$ for any Borel set X.*

Definition 38. *A measure μ is said to be absolutely continuous with respect to a measure λ, denoted as $\mu \ll \lambda$, if $\mu(S) = 0$ for all events (i.e. for all sets) S such that $\lambda(S) = 0$.*

Definition 39. *A measure μ is absolutely continuous with respect to Lebesgue measure if there exists an L^1 function f so that*

$$\int g d\mu = \int g f dx \tag{7.10}$$

for any Borel function $g \in L^1(\mathbb{R}, d\mu)$. Then we can write $d\mu = f dx$.

Definition 40. *A measure μ is singular relative to Lebesgue measure if and only if $\mu(S) = 0$ for some set S such that $\mu(\mathbb{R} \backslash S) = 0$.*

Since the three measures are mutually singular, we can write

$$L^2(\mathbb{R}, d\mu) = L^2(\mathbb{R}, d\mu_{pp}) \bigoplus L^2(\mathbb{R}, d\mu_{ac}) \bigoplus L^2(\mathbb{R}, d\mu_{sing}) \tag{7.11}$$

It can be shown that any $\psi \in L^2(\mathbb{R}, d\mu)$ has an absolutely continuous spectral measure $d\mu_\psi$ if and only if $\psi \in L^2(\mathbb{R}, d\mu_{ac})$. The analogous result holds for pure point and singular measures.

μ_{ac}, μ_{sing}, and μ_{pp} are invariant under linear operations.

Theorem 7.2.2. *If μ is absolutely continuous then $f(\lambda)\mu$ is also absolutely continuous.*

Proof. Since μ is absolutely continuous w.r.t. the Lebesgue measure λ, we know that if there exists a set S with $\lambda(S) = 0$, then we also have that $\mu(S) = 0$.

$$f(\lambda)\mu(S) = \int_{\sigma(T)} \chi_S(\lambda) d\left(f(\lambda)\mu\right) = \int_{\sigma(T)} \chi_S(\lambda) \cdot f(\lambda) d\mu = \int_S f(\lambda) d\mu = 0.$$

Since S has Lebesgue measure 0, the integral above equals 0. □

Let $\mathcal{H}_{ac}$ be the subspace of the Hilbert space $\mathbb{H}$ consisting of vectors whose spectral measures are absolutely continuous with respect to Lebesgue measure. Define $\mathcal{H}_{sc}$ and $\mathcal{H}_{pp}$ in the same way. Then we have the following theorem

Theorem 7.2.3. $\mathbb{H} = \mathcal{H}_{ac} \bigoplus \mathcal{H}_{sc} \bigoplus \mathcal{H}_{pp}$ *where $\mathcal{H}_{ac} = \{\psi | \mu_\psi$ is absolutely continuous, $\mathcal{H}_{sing} = \{\psi | \mu_\psi$ is singularly-continuous$\}$, and $\mathcal{H}_{pp} = \{\psi | \mu_\psi$ is pure point.$\}$*

Theorem 7.2.4. *The subspaces $\mathcal{H}_{ac}, \mathcal{H}_{pp}$, and $\mathcal{H}_{sing}$ are invariant under A where A is a self-adjoint operator.*

Proof. Let $h \in \mathcal{H}_{ac}$, and let $k = Th$. Let χ be the characteristic function so that

$$\langle k, \chi(T) k \rangle = \int_{\sigma(T)} \chi(\lambda) d\mu_k(\lambda) \tag{7.12}$$

We can use the functional calculus to show that the above expression equals

$$\int_{\sigma(T)} \chi(\lambda) \lambda^2 d\mu_h(\lambda) \tag{7.13}$$

so that $\lambda^2 d\mu_h = d\mu_k$. This would imply that μ_k is also absolutely continuous, and thus that $k \in \mathcal{H}_{ac}$. This would give us the desired result that $\mathcal{H}_{ac}$ is invariant under A where A is a self-adjoint operator.

$$\begin{aligned}\langle k, \chi(T)k\rangle &= \langle Th, \chi(T)\cdot Th\rangle = \langle h, T\chi(T)\cdot Th\rangle = \langle h, Id(T)\cdot \chi_B(T)\cdot Id(T)h\rangle \\ &= \langle h, (Id\cdot\chi_B\cdot Id)(T)h\rangle = \int_{\sigma(T)} p(\lambda)d\mu_h(\lambda)\end{aligned}$$

where $p(\lambda) = \lambda^2$ if $\lambda \in B$, and 0 otherwise, and $Id(T)$ is the identity function. □

The spectrum of a self-adjoint operator can be decomposed into several components. We introduce a family of projection operators uniquely associated to the self-adjoint operator H_ω.

Define $E^{ac}_\Delta(\omega) := E_\Delta(\omega)P_{ac}(\omega)$ where $P_{ac}(\omega)$ is the projector onto the continuous subspace w.r.t H_ω. Similarly, define $E^{sc}_\Delta(\omega) := E_\Delta(\omega)P_{sc}(\omega)$ and $E^{pp}_\Delta(\omega)$.

Definition 41. *The spectrum of an operator H restricted to $\mathcal{H}_{ac}$ is called the absolutely continuous spectrum of H, $\sigma_{ac}(H)$.*

Definition 42. *The spectrum of an operator H restricted to $\mathcal{H}_{sc}$ is called the singular spectrum of H, $\sigma_{sc}(H)$.*

Definition 43. *The set of eigenvalues of H is called the pure point spectrum of H, $\sigma_{pp}(H)$.*

Now we have everything we need in order to be able to prove a result analogous to Pastur's theorem regarding the components of the spectrum of H_ω.

Theorem 7.2.5. *(Kunz-Souillard). Let V_ω be an ergodic potential. Then there exist sets $\Sigma_{ac}, \Sigma_{sc}, \Sigma_{pp} \subset \mathbb{R}$ such that*
$\sigma_{ac}(H_\omega) = \Sigma_{ac} \quad \mathbb{P} - a.s.$
$\sigma_{sc}(H_\omega) = \Sigma_{sc} \quad \mathbb{P} - a.s.$
$\sigma_{pp}(H_\omega) = \Sigma_{pp} \quad \mathbb{P} - a.s.$

This theorem says that for $\mathbb{P}$-almost all ω, the absolutely continuous spectrum, the singular spectrum, and the pure point spectrum of H_ω are nonrandom sets.

Proof. If we show that the projections $E^{ac}_\Delta, E^{sc}_\Delta$, and E^{pp}_Δ are weakly measurable, we can just apply the proof of Pastur's theorem 3 times with $E_\Delta(\omega)$ replaced by $E^{ac}_\Delta(\omega), E^{sc}_\Delta(\omega)$, and $E^{pp}_\Delta(\omega)$. We can do that because each of the spectral projections would satisfy the conditions of lemma 5.1.6. This would complete the proof Kunz-Souillard's theorem. Verifying the weak measurability of these spectral projections requires the RAGE theorem. We omit this part of the proof and refer the reader to [1] for the details.
For the remainder of this proof we assume the fact that the spectral projections are weakly measurable. Define

$$\begin{aligned}T_1 &= H_\omega\big|\mathcal{H}_{ac} : \mathcal{H}_{ac} \to \mathcal{H}_{ac}\\ T_2 &= H_\omega\big|\mathcal{H}_{sc} : \mathcal{H}_{sc} \to \mathcal{H}_{sc}\\ T_3 &= H_\omega\big|\mathcal{H}_{pp} : \mathcal{H}_{pp} \to \mathcal{H}_{pp}\end{aligned}$$

By definition, $\sigma(T_1) = \sigma_{ac}(H_\omega), \sigma(T_2) = \sigma_{sc}(H_\omega), \sigma(T_3) = \sigma_{pp}(H_\omega)$
Then we can apply Pastur's theorem using each of $E_\Delta^{ac}, E_\Delta^{sc}, E_\Delta^{pp}$ to obtain the results

$$\sigma(T_1) = \sigma_{ac}(H_\omega) = \Sigma_{ac} \tag{7.14}$$

$$\sigma(T_2) = \sigma_{sc}(H_\omega) = \Sigma_{sc} \tag{7.15}$$

$$\sigma(T_3) = \sigma_{pp}(H_\omega) = \Sigma_{pp} \tag{7.16}$$

with probability 1 for some fixed non-random sets Σ_{ac}, Σ_{sc},Σ_{pp}. This completes the proof of Kunz-Souillard's theorem. □

In the next chapter we briefly discuss the density of states which is an important concept for disordered systems.

Chapter 8

The Density of States

The density of states is a very concept in statistical and condensed matter physics. Intuitively, it measures the number of available states at each energy level. Recall that we defined the finite cube Λ_L in equation (2.12). Let χ_Λ denote the characteristic function of the set Λ, i.e. $\chi_\Lambda(x) = 1$ if $x \in \Lambda$ and $\chi_\Lambda(x) = 0$ otherwise. For any bounded measurable function ϕ on $\mathbb{R}$ we define

$$\xi_L(\phi) := \frac{1}{|\Lambda_L|}\operatorname{tr}\left(\chi_{\Lambda_L}\phi(H_\omega)\chi_{\Lambda_L}\right) \tag{8.1}$$

Note that this equals

$$\frac{1}{|\Lambda_L|}\operatorname{tr}\left(\phi(H_\omega)\chi^2_{\Lambda_L}\right) = \frac{1}{\Lambda_L}\operatorname{tr}\left(\phi(H_\omega)\chi_{\Lambda_L}\right) \tag{8.2}$$

where we used commutativity of the trace ($\operatorname{tr}(AB) = \operatorname{tr}(BA)$) and the fact that $\chi^2_{\Lambda_L} = \chi_{\Lambda_L}$. Clearly, $\phi \to \xi_L(\phi)$ is a positive linear functional on the bounded continuous functions because of the linearity of the trace operator.
The functional calculus we introduced earlier allows us to define $\phi(H_\omega)$. The Riesz representation theorem (6.1.1) tells us that there exists a unique measure ν_L such that

$$\xi_L(\phi) = \int_{\mathbb{R}} \phi(\lambda) d\nu_L(\lambda) \tag{8.3}$$

Theorem 8.0.6. *If ϕ is a bounded measurable function, then for $\mathbb{P}$-almost all ω,*

$$\lim_{L\to\infty} \frac{1}{|\Lambda_L|}\operatorname{tr}\left(\phi(H_\omega)\chi_{\Lambda_L}\right) = \mathbb{E}\left(\langle\delta_0, \phi(H_\omega)\delta_0\rangle\right). \tag{8.4}$$

Proof. Let $\{\delta_i\}_{i=1}^\infty$ be an orthonormal basis. Then by definition (30), we have

$$\frac{1}{|\Lambda_L|}\operatorname{tr}\left(\phi(H_\omega)\chi_{\Lambda_L}\right) = \frac{1}{|\Lambda_L|}\sum_{i=0}^{\infty}\langle\delta_i, \phi(H_\omega)\chi_{\Lambda_L}\delta_i\rangle = \frac{1}{(2L+1)}\sum_{i\in\Lambda_L}^{\infty}\langle\delta_i, \phi(H_\omega)\delta_i\rangle$$

To prove this claim, we take the limit as $L \to \infty$ and then use the mean ergodic theorem (2.2.3). However, first we need to verify that the random variables $X_i(\omega) := \langle\delta_i, \phi(H_\omega)\delta_i\rangle$ form a stochastic ergodic process as defined in (6).

Lemma 8.0.7. *The sequence* $\{X_i(\omega)\} := \{\langle \delta_i, \phi(H_\omega)\delta_i\rangle\}$ *is a stochastic ergodic process.*

Proof.

$$\begin{aligned} X_i(T_j\omega) &= \langle \delta_i, \phi(H_{T_j\omega})\delta_i\rangle \\ &= \langle \delta_i, U_j\phi(H_\omega)U_j^*\delta_i\rangle \\ &= \langle U_j^*\delta_i, \phi(H_\omega)U_j^*\delta_i\rangle \\ &= \langle \delta_{i-j}, \phi(H_\omega)\delta_{i-j}\rangle \\ &= X_{i-j}(\omega) \end{aligned}$$

where we used the fact that $U_j^*\delta_i(n) = \delta_i(n+j) = \delta_{i-j}(n)$. □

Applying the mean ergodic theorem (2.2.3), we get

$$\lim_{L\to\infty} \frac{1}{(2L+1)} \sum_{i\in\Lambda_L} \langle \delta_i, \phi(H_\omega)\delta_i\rangle = \mathbb{E}\left(\langle \delta_0, \phi(H_\omega)\delta_0\rangle\right). \tag{8.5}$$

This finishes the proof. This result and its proof are given in [7]. □

Notice that $\mathbb{E}\left(\langle \delta_0, \phi(H_\omega)\delta_0\rangle\right)$ is a positive linear functional. Therefore, by the Riesz representation theorem there exists a unique positive measure ν defined by

$$\int \varphi(\lambda)d\nu(\lambda) = \mathbb{E}\left(\langle \delta_0, \varphi(H_\omega)\delta_0\rangle\right). \tag{8.6}$$

Definition 44. *The measure* ν *defined by*

$$\nu(A) = \mathbb{E}\left(\langle \delta_0, \chi_A(H_\omega)\delta_0\rangle\right) \tag{8.7}$$

is called the density of states measure.

The distribution function F of ν, $F(E) = \nu((-\infty, E])$, is called the integrated density of states.

Since $\nu(\mathbb{R}) = 1$, ν is a probability measure. We can show that the sequence of measures $\{\nu_L\} \to \nu$ as $L \to \infty$ in the vague sense for $\mathbb{P}$-almost all ω. Vague convergence is defined below.

Definition 45. *(Vague convergence) A series* ν_n *of Borel measures on* $\mathbb{R}$ *converges vaguely to a Borel measure* ν *if*

$$\lim_{n\to\infty} \int \phi(x)d\nu_n(x) = \int \phi(x)d\nu(x) \tag{8.8}$$

for all functions $\phi \in \mathcal{C}_0(\mathbb{R})$ *where* $\mathcal{C}_0(\mathbb{R})$ *is the set of continuous functions with compact support.*

The following claim and its proof can be found in [7].

Theorem 8.0.8. *The measures $\{\nu_L\}$ converge vaguely to the measure ν $\mathbb{P}$-almost surely, i.e. there is a set Ω_0 of probability one such that*

$$\lim_{L\to\infty}\int\varphi(\lambda)d\nu_L(\lambda)=\int\varphi(\lambda)d\nu(\lambda) \tag{8.9}$$

for all $\varphi\in\mathcal{C}_0(\mathbb{R})$ and all $\omega\in\Omega_0$.

Proof. Equations (8.1), (8.2), and (8.3) and theorem (8.0.6) imply that

$$\lim_{L\to\infty}\int\varphi(\lambda)d\nu_L(\lambda)=\mathbb{E}(\langle\delta_0,\varphi(H_\omega)\delta_0\rangle). \tag{8.10}$$

But equation (8.6) implies

$$\lim_{L\to\infty}\int\varphi(\lambda)d\nu_L(\lambda)=\int\varphi(\lambda)d\nu(x) \tag{8.11}$$

Thus, by theorem (8.0.6) we have established the claim for a fixed φ for $\mathbb{P}$-almost all ω. Thus, the claim holds true on a set Ω_φ of probability one if φ is fixed. We need to prove that it holds for all functions $\varphi\in\mathcal{C}_0(\mathbb{R})$. Note that the claim holds for all ϕ for $\omega\in\cap_\varphi\Omega_\varphi$ where $\mathbb{P}(\Omega_\varphi)=1$ for each φ. However, this is an uncountable intersection of sets with probability one, and it is not clear whether $\mathbb{P}(\cap_\varphi\Omega_\varphi)=1$. Since by definition $\mathcal{C}_0(\mathbb{R})$ is the set of continuous functions with compact support, it must contain a dense set D. Let us define

$$\Omega_0:=\cap_{\varphi\in D}\Omega_\varphi.\infty \tag{8.12}$$

Ω_0 is a countable intersection of sets with probability one, so $\mathbb{P}(\Omega_0)=1$. Thus, we have established the claim for $\omega\in\Omega_0$ for all $\varphi\in D$. It remains to establish the claim for $C_0(\mathbb{R})\backslash D$. Since D is dense, every $\varphi\in\mathcal{C}_0(\mathbb{R})$ is the uniform limit of a sequence $\{\varphi_n\}\in D$. Thus, if $\varphi\in\mathcal{C}_0(\mathbb{R})$, there exists a sequence $\{\varphi_n\}\in D$ such that $\varphi_n\to\varphi$ uniformly. Then, by the triangle inequality,

$$\begin{aligned}
&|\int\varphi(\lambda)d\nu(\lambda)-\int\varphi(\lambda)d\nu_L(\lambda)|\\
&\le|\int\varphi(\lambda)d\nu(\lambda)-\int\varphi_n(\lambda)d\nu(\lambda)|+|\int\varphi_n(\lambda)d\nu(\lambda)-\int\varphi_n(\lambda)d\nu_L(\lambda)|+\\
&|\int\varphi_n(\lambda)d\nu_L(\lambda)-\int\varphi(\lambda)d\nu_L(\lambda)|\le\|\varphi-\varphi_n\|_\infty\cdot\int_{\mathbb{R}}d\nu+\\
&|\int\varphi_n(\lambda)d\nu(\lambda)-\int\varphi_n(\lambda)d\nu_L(\lambda)|+\|\varphi-\varphi_n\|_\infty\cdot\int_{\mathbb{R}}d\nu_L=\|\varphi-\varphi_n\|_\infty\cdot\nu(\mathbb{R})+\\
&|\int\varphi_n(\lambda)d\nu(\lambda)-\int\varphi_n(\lambda)d\nu_L(\lambda)|+\|\varphi-\varphi_n\|_\infty\cdot\nu_L(\mathbb{R})
\end{aligned}$$

Note that $\nu_L(\mathbb{R})$ and $\nu(\mathbb{R})$ are equal to 1 because they are probability measures. Since the claim holds for all $\varphi\in D$ and since $\{\varphi_n\}\in D$, we know that

$$\lim_{L\to\infty}\int\varphi_n(\lambda)d\nu_L(\lambda)=\int\varphi_n(\lambda)d\nu(\lambda) \tag{8.13}$$

Therefore, there exists N_1 such that when $L > N$,

$$\Big| \int \varphi_n(\lambda) d\nu(\lambda) - \int \varphi_n(\lambda) d\nu_L(\lambda) \Big| < \frac{\epsilon}{3} \tag{8.14}$$

Since $\varphi_n \to \varphi$ uniformly, we can also find N_2 such that when $n > N_2$, $\|\varphi - \varphi_n\|_\infty < \frac{\epsilon}{3}$. Therefore,

$$\Big| \int \varphi(\lambda) d\nu(\lambda) - \int \varphi(\lambda) d\nu_L(\lambda) \Big| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \tag{8.15}$$

This completes the proof. □

Chapter 9

Further Work

In this thesis we have started exploring the connection between ergodicity and the spectral properties of the random Schrödinger operator. We have introduced many of the relevant tools from ergodic theory and functional analysis, and have seen how and where they are applied in proving important results about the spectrum of the random Schrödinger operator. Much more can be done to achieve a more complete understanding of the connection between ergodicity and the spectral properties of this operator. For example, it would be helpful to consider other properties of this operator that hold with probability 1, and to think about exactly where ergodicity is necessary in order to be able to prove the desired results.

Bibliography

[1] Cycon, H. L.; Froese, R. G.; Kirsch, W.; Simon, B. *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*. Texts and Monographs in Physics. Springer Study Edition. Springer-Verlag, Berlin, 1987.

[2] Jitomirskaya, S., *Ergodic Schrödinger operators (on one foot)*. Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon's 60th birthday, 613–647, Proc. Sympos. Pure Math., 76, Part 2, Amer. Math. Soc., Providence, RI, 2007.

[3] Silva, C., *Invitation to Ergodic Theory*. Student Mathematical Library, 42. American Mathematical Society, Providence, RI, 2007.

[4] Reed, M.; Simon, B. *Methods of modern mathematical physics. I. Functional analysis*. Second edition. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1980.

[5] Demuth, M. *Determining Spectra in Quantum Theory*. Progress in Mathematical Physics, 44. Birkhuser Boston, Inc., Boston, MA

[6] Stollman, P. *Caught by Disorder*. Bound states in random media. Progress in Mathematical Physics, 20. Birkhuser Boston, Inc., Boston, MA, 2001.

[7] Kirsch, W., *An Invitation to Random Schrodinger Operators*, lecture notes available at *http://arxiv.org/pdf/0709.3707.pdf*

[8] Shalizi, C., *Notes* available at *http://www.cscs.umich.edu/ crshalizi/notabene/*

[9] Bass, R., *Lecture Notes on Probability* available at *www.math.uconn.edu/ bass/prob.pdf*

[10] Van Handel, R., *Lecture Notes for Stochastic Calculus and Stochastic Control* available at *http://www.its.caltech.edu/ ramon/acm217/ACM217.pdf*