Introduction to Quantum Theory

William G. Faris

October 14, 1998

Contents

Abstract

These lectures give a brief and elementary introduction to quantum mechanics that starts from zero and leads to ideas from current research. The topics are:

- The uncertainty principle. This is an account of the general physical significance of quantum mechanics. The uncertainty principle gives a picturesque way of thinking about what quantum mechanics is about. However it can also be given a precise mathematical formulation. In this account the uncertainty principle for quantum mechanics is deduced from an uncertainty principle for diffusion.
- The geometry of quantum mechanics. Throughout these lectures the account will be of quantum mechanics formulated in the finite dimensional case. The fundamental probability interpretation of quantum mechanics is related to the geometry of inner product spaces, in particular to the theorem of Pythagorus. There is a different version of the uncertainty principle that may be formulated in the context of such elementary geometry.
- Commutation relations. There is more to quantum mechanics than abstract geometry; the geometry must be related to physical quantities like position and momentum. This lecture presents a characterization of a quantum mechanical system describing a particle present in a finite one-dimensional crystal consisting of N points (arranged in a circle). A simple commutation relation has an essentially unique solution that characterizes the physical system.
- Extended and localized states. The ultimate goal of quantum mechanics is to compute quantities associated to the energy. It is easy to compute functions of the kinetic energy associated to the particle moving in the crystal. However the total energy, consisting of kinetic energy and potential energy, is more complicated. One question considered in current research is whether the states of fixed energy are extended over the crystal, or localized near particular places. The usual picture of conduction electrons in a metal is via extended states. However in one dimension the smallest amount of randomness is enough to destroy this and produce localization. What happens in higher dimensions is largely unknown, at least as far as rigorous mathematical results are concerned.

Chapter 1

The uncertainty principle

1.1 Planck's constant

The characteristic sign of quantum mechanics is the rationalized Planck's constant

$$\hbar = 1.054 \times 10^{-27} \text{ erg-sec.} \tag{1.1}$$

(To say that it is a rationalized constant means in this case that it incorporates a factor of $1/(2\pi)$.) The units of this quantity are energy times time, or momentum times distance. In other words, the units are those of area in phase space. When considering an area in phase space smaller than this scale, the world must be regarded in a radically different way. This is the quantum theory.

The most famous expression of this new way of looking at the world is the Heisenberg uncertainty principle. This says that the product of the uncertainties of the momentum p and the position x satisfy

$$\Delta p \, \Delta x \geq \frac{\hbar}{2}. \tag{1.2}$$

Let us be more precise about what we mean by uncertainty. Suppose that the probability density of position is given by a density $\rho(x)$. This is a positive function with integral over the x axis equal to one. Then the average (mean, expected value) of x is

$$\langle x \rangle = \int_{-\infty}^{\infty} x\rho(x) \, dx. \tag{1.3}$$

The average squared deviation (variance) of x is

$$(\Delta x)^2 = \langle (x - \langle x \rangle)^2 \rangle = \int_{-\infty}^{\infty} (x - \langle x \rangle)^2 \rho(x) \, dx. \tag{1.4}$$

The square root Δx of this is the uncertainty (standard deviation). Sometimes this is called the root mean square (RMS) deviation.

Suppose there is some other probability density $\tilde{\rho}(p)$ for momentum. In the same way we define the mean value of p as

$$\langle p \rangle = \int_{-\infty}^{\infty} p\tilde{\rho}(p)\, dp. \tag{1.5}$$

The mean square of p is

$$(\Delta p)^2 = \langle (p - \langle p \rangle)^2 \rangle = \int_{-\infty}^{\infty} (p - \langle p \rangle)^2 \tilde{\rho}(p)\, dp. \tag{1.6}$$

The root mean square is Δp.

The uncertainty principle says that the two probability densities are such that the product of the standard deviations is bounded below by a small but strictly positive number, $\hbar/2$. If one density is strongly peaked, then the other must be somewhat spread out. Note that the principle deals with the two probability distributions for momentum and position separately. In quantum mechanics there is no joint distribution for these two quantities. This is interpreted as saying that measurement of one quantity precludes in principle the measurement of the other quantity.

1.2 The size of atoms

Now we give a physical argument that is very non-rigorous. However there is a rigorous version of it. This can be found, for instance, in my article "Inequalities and uncertainty principles," *J. Math. Phys.* **19** (1978), 461–466.

The non-rigorous argument is the following. Consider a system consisting of a nucleus and an electron. The nucleus is regarded as fixed. The kinetic energy of the electron is

$$T = \frac{p^2}{2m}, \tag{1.7}$$

where m is the mass of the electron. The kinetic energy is always positive, and it is zero when the momentum p is zero.

The potential energy of the electron is

$$V = -\frac{e^2}{r} \tag{1.8}$$

where e is proportional to the charge of the electron. Here r is the distance of the electron from the fixed nucleus. The potential energy is always negative. It is close to zero when the electron is far away from the nucleus, but it becomes extremely negative when the electron is very close to the nucleus.

The total energy is the sum

$$H = T + V = \frac{p^2}{2m} - \frac{e^2}{r} \tag{1.9}$$

of the kinetic and potential energies. In classical mechanics the state of lowest energy would be when $p = 0$, $r = 0$, and $H = -\infty$. This is a rather singular situation. However quantum mechanics, in the form of the uncertainty principle, comes to the rescue.

Let us assume (this is the non-rigorous part) that the lowest energy situation in quantum mechanics is when we take $\langle p \rangle = 0$ and that the remaining kinetic energy has an average value

$$T = \frac{(\Delta p)^2}{2m}. \tag{1.10}$$

Let us also assume that the distance r of the electron from the nucleus may be regarded as comparable to $r = 2\Delta x$. (The factor of two is for convenience; it makes the final formulas look nicer.) Thus we write

$$V = -\frac{e^2}{2\Delta x}. \tag{1.11}$$

Finally, we assume that we have the least possible uncertainty: in the form $\Delta p \Delta x = \hbar/2$. From these equations we get that the total energy is

$$H = \frac{(\Delta p)^2}{2m} - \frac{e^2 \Delta p}{\hbar}. \tag{1.12}$$

This has its minimum value when

$$\frac{\Delta p}{m} = \frac{e^2}{\hbar}. \tag{1.13}$$

The corresponding radius is

$$r = 2\Delta x = \frac{\hbar^2}{me^2}. \tag{1.14}$$

How big are these numbers? The quantity $e^2/\hbar$ is a velocity, very roughly equal to

$$\frac{e^2}{\hbar} = 2 \times 10^8 \,\mathrm{cm/sec}. \tag{1.15}$$

The quantity $\hbar/m$ is a diffusion constant, which for the case of an electron is roughly equal to

$$\frac{\hbar}{m} = 1 \,\mathrm{cm}^2/\mathrm{sec}. \tag{1.16}$$

Thus the radius is about

$$r = \frac{1}{2} \times 10^{-8} \,\mathrm{cm}. \tag{1.17}$$

Let us say that this is about half of the diameter of a typical atom, which is 10^{-8} cm. Then the number of atoms in a cubic centimeter would have to be about 10^{24}. This is in fact rather close to the number used in chemical calculations.

1.3 The uncertainty principle in diffusion

Is the uncertainty principle unique to quantum mechanics? Not really. Let us look at how such an uncertainty principle might occur in a probability problem.

Let us make the postulate that particles are diffusing with diffusion constant $\sigma^2/2$. This means that the flow of particles is given by $-(\sigma^2/2)\, d\rho(x)/dx$, where $\rho(x)$ is the density of particles. The particles diffuse from a region of higher density to a region of lower density.

Let us also assume that equilibrium is maintained by a tendency of the particles to drift at a velocity $u(x)$. This should be thought of as a terminal velocity that results from an external force and opposing frictional force. Then the flow of particle due to this drift is $u(x)\rho(x)$.

The detailed balance between these two effects that produces equilibrium is

$$u(x)\rho(x) - \frac{\sigma^2}{2}\frac{d\rho(x)}{dx} = 0. \tag{1.18}$$

As an example, one can think of the case when x represents height and $u(x)$ represents the terminal velocity of a falling particle. Then $u(x)$ would be negative. This would compensate the positive flow of particles due to diffusion. The density would be decreasing with height in such a way as to maintain the balance given by this equation.

We have the following uncertainty principle for diffusion.

Theorem 1.1 *Consider a smooth probability density* $\rho(x)$ *that goes to zero sufficiently fast at infinity. Define* $u(x)$ *by the detailed balance equation. Then* $\langle u \rangle = 0$ *and*

$$\Delta u\, \Delta x \geq \frac{\sigma^2}{2}. \tag{1.19}$$

This says that it is impossible to concentrate the probability near its average value (Δx very small) unless the drift that maintains the equilibrium has values that are far from constant (Δu very large).

Proof: First we compute

$$\langle u \rangle = \int_{-\infty}^{\infty} u(x)\rho(x)\, dx = \frac{\sigma^2}{2}\int_{-\infty}^{\infty} \frac{d\rho(x)}{dx}\, dx = 0. \tag{1.20}$$

Then we use the identity

$$\frac{d}{dx}(x - \langle x \rangle)\rho(x) = \rho(x) + (x - \langle x \rangle)\frac{d}{dx}\rho(x). \tag{1.21}$$

When we integrate this, the integral of the left hand side is zero. Therefore we get

$$1 = -\int_{-\infty}^{\infty} (x - \langle x \rangle)\frac{d}{dx}\rho(x)\, dx. \tag{1.22}$$

This is the same as

$$\frac{\sigma^2}{2} = -\int_{-\infty}^{\infty} (x - \langle x \rangle)\, u(x)\rho(x)\, dx. \tag{1.23}$$

This says that the deviation has a non-zero negative correlation with the drift that maintains equilibrium. The proof of the uncertainty principle is concluded by an application of the Schwarz inequality, so

$$\frac{\sigma^2}{2} \leq \sqrt{\int_{-\infty}^{\infty} u(x)^2 \rho(x)\, dx} \sqrt{\int_{-\infty}^{\infty} (x - \langle x \rangle)^2 \rho(x)\, dx}. \tag{1.24}$$

1.4 The uncertainty principle in quantum mechanics

In quantum mechanics the physics is much more mysterious. The fundamental quantity is the wave function $\psi(x)$, which is a complex valued function. The relation of the wave function to the position probability density is

$$\rho(x) = |\psi(x)|^2. \tag{1.25}$$

Define the osmotic velocity $u(x)$ and the current velocity $v(x)$ by

$$u(x) + iv(x) = \frac{\hbar}{m} \frac{1}{\psi(x)} \frac{d\psi(x)}{dx}. \tag{1.26}$$

It may be shown in quantum mechanics that $\langle u \rangle = 0$ and that

$$\langle p \rangle = m \langle v \rangle. \tag{1.27}$$

This identity says that the expected momentum is the mass times the expected current velocity. It is more subtle than it looks. The expectation on the left is taken with respect to the quantum mechanical momentum distribution $\tilde{\rho}(p)$, while the expectation on the right is taken with respect to the position distribution $\rho(x)$.

It may be shown in quantum mechanics that

$$\Delta p = m \sqrt{(\Delta u)^2 + (\Delta v)^2}. \tag{1.28}$$

This identity relates the uncertainty in the momentum to the uncertainties in the osmotic and current velocities. Again the left hand side is computed with respect to the quantum mechanical momentum density, while the right hand side is computed with respect to the position density.

The conclusion is that if we define the quantum mechanical diffusion constant by

$$\sigma^2 = \frac{\hbar}{m}, \tag{1.29}$$

then the diffusion uncertainty principle implies the quantum mechanical uncertainty principle

$$\Delta p \, \Delta x \geq \frac{\hbar}{2}. \tag{1.30}$$

Chapter 2

Geometry of quantum mechanics

2.1 Complex inner product space

This section introduces the notion of complex inner product space. For simplicity, we shall consider only the cases when the space is finite-dimensional.

We consider a complex vector space **H**. Thus if ψ and ϕ are vectors in **H** and z and w are complex numbers, the linear combination $z\psi + w\phi$ is defined and is in **H**. The usual vector space axioms are satisfied.

The vector space has a given inner product. For every pair of vectors the inner product $\langle \phi, \psi \rangle$ is a complex number. Furthermore, reversing the order of the vectors is equivalent to complex conjugation, so

$$\langle \psi, \phi \rangle = \langle \phi, \psi \rangle^*, \tag{2.1}$$

where the star denotes complex conjugation. We also make the convention that the inner product is linear in the second variable. Thus

$$\langle \chi, z\psi + w\phi \rangle = z\langle \chi, \psi \rangle + w\langle \chi, \phi \rangle. \tag{2.2}$$

It follows that it is conjugate linear in the first variable:

$$\langle z\psi + w\phi, \chi \rangle = z^*\langle \psi, \chi \rangle + w^*\langle \phi, \chi \rangle. \tag{2.3}$$

(This is the most common convention in physics. In mathematics the opposite convention is prevalent.) Finally, we assume that $\langle \psi, \psi \rangle \geq 0$ and that it is zero only for the zero vector.

The norm of a vector ψ is

$$\|\psi\| = \sqrt{\langle \psi, \psi \rangle}. \tag{2.4}$$

Thus it is always true that $\|\psi\| \geq 0$, and this is zero only for the zero vector.

The distance between two vectors is defined to be the norm of their difference. With this notion of distance, the vector space **H** is a metric space.

Two vectors ϕ and χ are said to be orthogonal or perpendicular if $\langle \phi, \chi \rangle = 0$.

The following result is the theorem of Pythagoras:

Theorem 2.1 *If ϕ and χ are orthogonal, then*

$$\|\phi\|^2 + \|\chi\|^2 = \|\phi + \chi\|^2. \tag{2.5}$$

If M is a subspace of the space **H**, then $M^{\perp}$ is defined to be the set of all vectors χ perpendicular to every vector in M. It is also a subspace of **H**.

The following theorem is fundamental. It is known as the projection theorem.

Theorem 2.2 *Consider a finite dimensional complex inner product space* **H**. *Let M be a linear subspace of* **H**. *Let ψ be a vector in* **H**. *Then there are unique vectors ϕ in M and χ in $M^{\perp}$ such that*

$$\psi = \phi + \chi. \tag{2.6}$$

We call ϕ the orthogonal projection of ψ onto M.

2.2 Quantum mechanics

We consider quantum mechanics in the finite-dimensional case. Let **H** be a complex inner product space. The collection of all the one-dimensional subspaces of **H** is called a complex projective space. In the following we choose to represent a one-dimensional subspace by a unit vector ψ in the space. Two unit vectors determine the same one-dimensional subspace if and only if one is a multiple of the other by a complex number of absolute value one.

1. There is a complex inner product space **H** such that the states of the system correspond to the one-dimensional subspaces of **H**. Such a state is represented by a unit vector ψ in the one-dimensional subspace.

2. The quantum events are in one-to-one correspondence with the linear subspaces M of the Hilbert space **H**.

3. The probability of an event M when the state is given by the unit vector ψ is
$$P_{\psi}[M] = \|\phi\|^2, \tag{2.7}$$
where ϕ is the orthogonal projection of ψ onto M.

2.3 Negation

For each quantum event M there is a complementary event $M^{\perp}$, the orthogonal complement. We can use notation from mathematical logic and write this as $\neg M$, the negation of M. We have the following fundamental result:

$$P_{\psi}[M] + P_{\psi}[\neg M] = 1. \tag{2.8}$$

This of course is just the theorem of Pythagoras in quantum mechanical language. The interpretation of this is that the event $\neg M$ happens precisely when the event M does not happen. So it is reasonable that the probabilities add to one. Sometimes we say that the zero subspace represents the impossible event and the entire space $\mathbf{H}$ represents the sure event. Whatever the state, the probability of the impossible event is zero, while the probability of the sure event is one.

2.4 Conjunction

Given two closed subspaces M and N, $M \cap N$ is also a closed subspace. We say that two quantum events M and N are compatible if every vector in $\mathbf{H}$ may be written as a sum of its projections on the four subspaces $M \cap N$, $M \cap N^{\perp}$, $M^{\perp} \cap N$, and $M^{\perp} \cap N^{\perp}$.

If M and N are compatible, then define the conjunction of the two events to be

$$M \& N = M \cap N. \tag{2.9}$$

This makes sense, because for compatible events we have the identity

$$P_{\psi}[M \& N] + P_{\psi}[M \& \neg N] + P_{\psi}[\neg M \& N] + P_{\psi}[\neg M \& \neg N] = 1 \tag{2.10}$$

for every vector ψ. In particular, we have the identity

$$P_{\psi}[M] = P_{\psi}[M \& N] + P_{\psi}[M \& \neg N] \tag{2.11}$$

for every vector ψ.

If M and N are not compatible, then we do not try to define the conjunction. Subspaces in general position tend not to be compatible. In ordinary probability theory there is no obstacle to forming the conjunction of two events. This is the main distinction between quantum mechanics and probability.

The philosophical justification for incompatible quantum events is often taken to be the following. An experiment that measures whether one event occurs or not may preclude an experiment that measures whether the other event occurs or not. So in principle it is meaningless to speak of the joint occurrence of the events.

2.5 An uncertainty principle

Even when the conjunction of two events M and N is not defined, it is possible to compute the probabilities of the two events separately. We want to give an uncertainty principle relating these probabilities.

Given two subspaces, we want to define the minimum angle between these two subspaces. It will be an angle θ between 0 and $\pi/2$. The definition is that $\cos^2(\theta)$ is the maximum value of $|\langle\psi, \phi\rangle|^2$ for unit vectors ψ in M and ϕ in N. If $\cos^2(\theta)$ is near zero, then θ is near $\pi/2$, and the subspaces are nearly orthogonal.

If $\cos^2(\theta)$ is close to one, then θ is near zero, and the subspaces can be rather close to each other.

The uncertainty principle is stated in the following theorem.

Theorem 2.3 *Let M and N be two subspaces corresponding to quantum events. Let θ be the minimum angle (between 0 and $\pi/2$) between the two subspaces. Let ψ be a unit vector representing a quantum state. Then*

$$P_\psi[M] + P_\psi[N] \leq 1 + \cos(\theta). \tag{2.12}$$

Thus if the angle θ is close to $\pi/2$, the subspaces are nearly orthogonal, $\cos(\theta)$ is close to zero, and the right hand side is close to one. This says that the two probabilities cannot both be close to one. This is the sense in which this is an uncertainty principle.

Consider two intervals, an interval I of position values and and interval $\tilde{I}$ of momentum values. Suppose that the length of I is δx and the length of $\tilde{I}$ is Δp. (Here Δx and Δp are arbitrary positive numbers, not necessarily arising as standard deviations.) It will be shown in the next lecture that if M is the event that the position x is in the interval I and N is the event that the momentum p is in the interval $\tilde{I}$, then

$$\cos(\theta) \leq \sqrt{\frac{\Delta x \Delta p}{2\pi\hbar}}. \tag{2.13}$$

Thus when the product of the two lengths is small with respect to Planck's constant, the cosine is close to zero, and the uncertainty principle constrains the probabilities. For example, if the product of the numbers Δx and Δp happens to be $\frac{1}{100}$ of $2\pi\hbar$, the sum of the probabilities that x is in the space interval I and p is in the momentum interval $\tilde{I}$ can be at most $\frac{1}{10}$. This is quite far from the maximum value of 2 that one would get classically by constraining the position and momentum to be in these intervals with probability one. For example, if the probability that x is in I is close to one, then the probablity that p is in $\tilde{I}$ is close to zero.

Technical note: Hilbert space

These lectures for the main part only consider quantum mechanics in which the complex inner product space is finite dimensional. However the geometric picture of states and quantum events also works in the infinite dimensional case. There is an additional condition on the complex inner product space $\mathbf{H}$. It must be a Hilbert space, which means that it is a complete metric space with respect to the norm. The quantum events are then identified with those linear subspaces that are themselves Hilbert spaces. Since a subset of a complete metric space is complete if and only if it is closed, these are the same as the closed linear subspaces. With this restriction, the projection theorem remains true, and the theory proceeds as above.

Chapter 3

Commutation relations

3.1 Unitary operators

As before we consider a finite dimensional complex inner product space $\mathbf{H}$.

We have seen that the fundamental quantities in quantum mechanics are the states (given by unit vectors) and the quantum events (given by subspaces).

Let A be a linear transformation from $\mathbf{H}$ to itself. We shall call such a linear transformation an operator. We need the concept of adjoint operator A^*. This is the operator that satisfies the identity

$$\langle A^*\phi, \psi\rangle = \langle \phi, A\psi\rangle \tag{3.1}$$

for all pairs of vectors ϕ and ψ.

We also need the concepts of self-adjoint operator and unitary operator. A self-adjoint operator is equal to its own adjoint: $A^* = A$. A unitary operator is an operator whose adjoint is equal to its inverse: $U^* = U^{-1}$. In particular,

$$\langle U^{-1}\phi, \psi\rangle = \langle \phi, U\psi\rangle \tag{3.2}$$

which means that

$$\langle \chi, \psi\rangle = \langle U\chi, U\psi\rangle. \tag{3.3}$$

A unitary operator preserves the inner product.

3.2 Projections

We can identify the subspaces by the corresponding projection operators. If M is a closed subspace of $\mathbf{H}$, then there is an operator E of orthogonal projection onto M. This is a linear transformation E from $\mathbf{H}$ to itself such that E is self-adjoint and idempotent, that is, such that $E = E^*$ and $E^2 = E$. Conversely, given such a projection operator E, its range is a subspace M. So there is a one-to-one correspondence between projection operators and quantum events.

The logical operators on projections follow from the corresponding logical operations on subspaces. The projection corresponding to the negation of E is $I-E$. It is also a projection, since it is self-adjoint and $(I-E)^2 = I-2E+E^2 = I-2E+E = I-E$.

The logical operation of conjunction is multiplication, but it is only defined when the projections commute. Thus if E and F are projections, the adjoint is $(EF)^* = F^*E^* = FE$, and the square is $EFEF$. If $EF = FE$, then $(EF)^* = EF$ and $EFEF = E^2F^2 = EF$. So when E and F commute, then EF is also a projection.

Finally, the logical operation of disjunction is defined for commuting projections E and F by $I-(I-E)(I-F) = E+F-EF$. From now on we identify quantum events with projections.

We can write the probability of a quantum event E in the state ψ to be

$$P_\psi[E] = \|E\psi\|^2 = \langle E\psi, E\psi\rangle = \langle \psi, E\psi\rangle. \tag{3.4}$$

The equivalence between the last two ways of writing this uses the relations $E = E^*$ and $E^2 = E$.

If U is a unitary operator and M is a subspace, then the image UM is another subspace. If E is the projection on M, then the projection on UM is UEU^{-1}. In fact, it is easy to verify that $UEU^{-1} = UEU^*$ is a projection and has range UM. The quantum mechanical probabilities are invariant under such unitary transformations. This is expressed by the equation

$$P_{U\psi}[UEU^{-1}] = P_\psi[E]. \tag{3.5}$$

Consequently, such an overall unitary transformation of both states and events should not affect the physical predictions. This property is called the unitary invariance of quantum mechanics.

3.3 Discrete position and momentum

As an example of a quantum mechanical system we consider a one-dimensional periodic system with N points arranged in a circle. We think of the N points mathematically as the set $0, 1, 2, \ldots, N-1$ with addition defined modulo N. Thus this is a finite commutative group. Physically, we think of the points as the sites in a one-dimensional crystal. The crystal is taken to be periodic for mathematical convenience, but it turns out that many physical predictions do not depend heavily on this assumption.

It turns out that in this model the momentum will also assume a finite set of N values. This seems curious, but it is actually quite common in the description of matter at the atomic level, in the context of solid state physics. Of course the value of N will be extremely large, perhaps on the order of 10^8. (Even more common are two or three dimensional systems, where the number of crystal sites would be 10^{16} or 10^{24}.)

The initial description of this system will be one in which the inner product space **H** has a specific form as a space of functions (actually finite sequences).

Thus it is not in the spirit of unitary invariance. We shall see later in this lecture how to give a description in the spirit of unitary invariance.

The complex vector space is $\mathbf{C}^N$ with the usual inner product. We think of this space as the space of all complex functions defined on the set of crystal sites $0, 1, 2, \ldots, N-1$. The inner product of two functions is

$$\langle f, g\rangle = \sum_{n=0}^{N-1} f(n)^* g(n). \tag{3.6}$$

The superscript $*$ denotes complex conjugate.

Next we need some projections. The event E_m that the position of the particle is at m is the orthogonal projection onto the indicator function δ_m whose value at n is given by the identity matrix δ_{mn}. Thus

$$E_m f(n) = \delta_{mn} f(m). \tag{3.7}$$

Thus in the representation E_m is a diagonal matrix with a 1 on the mth place on the diagonal.

The event F_k that the momentum of the particle is k is the orthogonal projection onto the normalized exponential function ϕ_k whose value at n is

$$\phi_k(n) = \frac{1}{\sqrt{N}} \exp(\frac{2\pi i k n}{N}). \tag{3.8}$$

Thus $F_k f = \phi_k \langle \phi_k, f\rangle$, or more explicitly,

$$F_k f(n) = \phi_k(n) \sum_{r=0}^{N-1} \phi_k(r)^* f(r) = \frac{1}{N} \exp(\frac{2\pi i k n}{N}) \sum_{r=0}^{N-1} \exp(-\frac{2\pi i k r}{N}) f(r). \tag{3.9}$$

This is a rank one matrix.

Furthermore, we need unitary operators. The unitary operator that translates in space is

$$V f(n) = f(n-1) \tag{3.10}$$

where the subtraction is taken modulo N. This is a matrix most of whose entries are ones just below the diagonal. It is a position shift because it takes the indicator function δ_m to the indicator function δ_{m+1}. The effect of this on the position projection operators is thus $V E_m V^{-1} = E_{m+1}$. Furthermore, it is possible to calculate that the position shift V may be expressed in terms of the momentum operators F_k by

$$V = \sum_{k=0}^{N-1} \exp(\frac{-2\pi i k}{N}) F_k. \tag{3.11}$$

This is because shifting the exponential function is the same as multiplying it by the phase factor.

The unitary operator that translates in momentum is

$$Uf(n) = \exp(\frac{2\pi i n}{N})f(n). \tag{3.12}$$

This is a diagonal matrix. It is a momentum shift because it takes the function ϕ_k with momentum k to the function ϕ_{k+1} with momentum $k+1$. The effect on the momentum projection operators is thus $UF_kU^{-1} = F_{k+1}$. The momentum shift may be expressed in terms of the position operators E_m by

$$U = \sum_{m=0}^{N-1} \exp(\frac{2\pi i m}{N})E_m. \tag{3.13}$$

We can summarize these results in the following list.

- The position projection E_m projects onto the subspace spanned by the indicator function δ_m.
- The momentum projection F_k projects onto the subspace spanned by the exponential function ϕ_k.
- The position shift operator V shifts position according to $V\delta_m = \delta_{m+1}$, or
$$VE_mV^{-1} = E_{m+1}. \tag{3.14}$$
It is a function of momentum.
- The momentum shift operator U shifts momentum according to $U\phi_k = \phi_{k+1}$, or
$$UF_kU^{-1} = F_{k+1}. \tag{3.15}$$
It is a function of position.

There is a particularly elegant representation of these facts in terms of the unitary operators alone. This is expressed in the following relations. The interesting thing is that these relations are just different ways of writing the same thing.

- The position shift operator V shifts U according to
$$VUV^{-1} = \exp(-\frac{2\pi i}{N})U. \tag{3.16}$$
- The momentum shift operator U shifts V according to
$$UVU^{-1} = \exp(\frac{2\pi i}{N})V. \tag{3.17}$$

Note: For mathematical convenience we have been considering the position n and quantum mechanical momentum k as having integer values. In actual quantum mechanical calculations the position x is $n\,\delta x$, where δx is a distance unit. Similarly, the momentum p is $k\,\delta p$, where δp is a momentum unit equal to Planck's constant $\hbar$ times $2\pi/(N\delta x)$.

3.4 A discrete uncertainty principle

Let us look at the probability that a particle is simultaneously localized in space and in momentum. Say that it is localized in a set A in space and a set B in momentum. The corresponding projections are

$$E(A) = \sum_{n \in A} E_n \tag{3.18}$$

and

$$F(B) = \sum_{k \in B} F_k. \tag{3.19}$$

We know that

$$\langle \psi, E(A)\psi \rangle + \langle \psi, F(B)\psi \rangle \leq 1 + \cos(\theta), \tag{3.20}$$

where θ is the minimum angle between the range of $E(A)$ and the range of $E(B)$. Thus if $\cos(\theta)$ is close to zero, that is, θ is close to $\pi/2$, then it is impossible that the position is localized in A and the momentum in B.

Now it is not hard to see that

$$\cos^2(\theta) \leq \operatorname{tr}(E(A)F(B)E(A)). \tag{3.21}$$

It suffices to take ψ to be a unit vector in the range of $E(A)$ and ϕ to be a unit vector in the range of $F(B)$, with $|\langle \psi, \phi \rangle|^2$ maximal. This maximal value is $\cos^2(\theta)$. Then the projection of ψ onto the range of $F(B)$ is $\langle \phi, \psi \rangle \phi$. Finally,

$$\langle \psi, E(A)F(B)E(A)\psi \rangle = \langle \psi, F(B)\psi \rangle = |\langle \psi, \phi \rangle|^2 \tag{3.22}$$

is a diagonal matrix element and so is a lower bound for the trace.

Let a be the number of points in A and b be the number of points in B. Then

$$\operatorname{tr}(E(A)F(B)E(A)) = \frac{ab}{N}. \tag{3.23}$$

Thus if ab is considerably less than N, then the probabilities of simultaneous localization in position and momentum cannot both be near one.

It is interesting to look at the continuum limit. Think of the points in space as being spaced by δx. Then the points in momentum space are spaced by δp, where

$$\frac{\delta p \, \delta x}{2\pi\hbar} = \frac{1}{N}. \tag{3.24}$$

So the condition for the uncertainty principle to apply is that the area in phase space divided by Planck's constant be considerably less than one. The situation is summarized in the following theorem.

Theorem 3.1 *Consider intervals of length* $\Delta x = a\delta x$ *and* $\Delta p = b\delta p$, *where* a *and* b *are the number of points, and where* δx *and* δp *are the position and momentum spacings. Then the minimum angle* θ *between the corresponding position and momentum subspaces is constrained by*

$$\cos^2(\theta) \leq \frac{\Delta x \, \Delta p}{2\pi\hbar}. \tag{3.25}$$

3.5 A discrete commutation relation

Up to now we have considered the events by specifying them in a particular representation of the vector space as a space $\mathbf{C}^N$ of functions. It would be nice to have a more intrinsic characterization of these physical quantities, one compatible with the principle of unitary invariance. In this section we will see that the structure of the position and momentum events and of the position and momentum shift operators may be determined from abstract conditions that make no assumption about the particular form of the vector space.

Thus we assume that we have a finite-dimensional complex inner product space $\mathbf{H}$. We have a number N and two unitary operators U and V. The operator V will describe translation in space by one step, while the U will describe translation in momentum by one step. In our case both space and momentum are periodic with period N. Thus we require that $U^N = I$ and $V^N = I$. Finally we require that the operators satisfy the commutation relation

$$UV = \exp(\frac{2\pi i}{N})VU. \tag{3.26}$$

The commutation relation relates shifts in space (according to V) and in momentum (according to U). If we write it as $VUV^{-1} = \exp(-2\pi i/N)U$, then it says that a shift in space given by V makes a phase change in the momentum shift operator U. We shall see that the commutation relation not only relates the operators to each other, but also restricts the form of the individual operators.

We already know examples of operators U and V satisfying these relations. These are the concrete examples that we saw before, which for the moment we will denote by $\hat{U}$ and $\hat{V}$. In this standard representation the inner product space is $\mathbf{C}^N$ and the operators $\hat{U}$ and $\hat{V}$ are given by

$$\hat{U}f(n) = \exp(\frac{2\pi i n}{N})f(n) \tag{3.27}$$

and

$$\hat{V}f(n) = f(n-1). \tag{3.28}$$

The goal is to show that every solution satisfying the commutation relations together with a multiplicity condition is isomorphic to the standard representation. Thus one would like to show that there is a unitary operator W from $\mathbf{C}^N$ to $\mathbf{H}$ such that the operators U and V take the form

$$U = W\hat{U}W^{-1} \tag{3.29}$$

and

$$V = W\hat{V}W^{-1}. \tag{3.30}$$

Theorem 3.2 *Let $\mathbf{H}$ be a complex inner product space of finite dimension. Suppose that U and V are unitary operators with $U^N = I$ and $V^N = I$. Suppose that they satisfy the commutation relation. Finally, suppose that the eigenvalues of U and V have multiplicity one. Then U and V are isomorphic to the momentum and position translation operators of the standard representation.*

Proof: First we construct the projection operator corresponding to the mth point in space by

$$E_m = \frac{1}{N} \sum_{j=0}^{N-1} \exp(-\frac{2\pi imj}{N}) U^j. \tag{3.31}$$

We can recover the unitary operator U from these projection operators. In fact the jth power of U is given by

$$U^j = \sum_{m=0}^{N-1} \exp(\frac{2\pi imj}{N}) E_m. \tag{3.32}$$

In particular from the case $j = 0$ we see that the projections E_m sum to the identity operator. Furthermore, if the eigenvalues of U have multiplicity one, then the projections E_m have one-dimensional ranges.

Then the commutation relation implies the relation

$$VE_m = E_{m+1}V. \tag{3.33}$$

This relates shift in space by one unit to location in space. If we write it as $VE_mV^{-1} = E_{m+1}$, then it says that the position subspaces are shifted by the position shift operator in the expected way.

Take a unit vector ψ in the range of E_0. Then the relation above implies that the vector $V^n\psi$ is in the range of E_n. Since $\sum_n E_n = I$, every vector can be expanded in terms of these unit vectors. Thus the vectors $V^n\psi$ form an orthonormal basis for the vector space $\mathbf{H}$.

Consider the space of all complex functions f in $\mathbf{C}^N$. Let

$$Wf = \sum_n f(n)V^n\psi. \tag{3.34}$$

Then W is unitary from $\mathbf{C}^N$ to $\mathbf{H}$. Furthermore,

$$VWf = \sum_n f(n)V^{n+1}\psi = \sum_n f(n-1)V^n\psi = W\hat{V}f. \tag{3.35}$$

Also

$$E_mWf = f(m)V^m\psi = \sum_n \delta_{nm}f(m)V^n\phi = W\hat{E}_mf. \tag{3.36}$$

Here

$$\hat{E}_mf(n) = \delta_{nm}f(m) \tag{3.37}$$

is the position projection in the standard representation. This shows that $VW = W\hat{V}$ and $E_mW = W\hat{E}_m$. In other words, the operators V and E_m are isomorphic to the position shift and position localization operators in the standard representation. It is then easy to see that $UW = W\hat{U}$, so U is isomorphic to the momentum shift operator in the standard representation.

Remark: We have assumed in the above theorem that the multiplicity of the eigenvalues is one. If instead the multiplicity of the eigenvalues is q, then there is a unitary operator W from $\mathbf{H}$ to $\mathbf{C}^{Nq}$ so that under W the operators U and V act like q copies of the standard example. Some applications require higher multiplicities. For example, if we have a three dimensional crystal with N^3 points, then there will be translations in three directions. The translations in one specified direction will have multiplicity $q = N^2$. This number corresponds the number of points in a two dimensional layer that is carried along in the translation.

Chapter 4

Extended and localized states

4.1 Energy: the index of extension

The commutation relations determine the structure of quantum events. Which events are important? By far the most important are the ones corresponding to values of the energy. The energy H is a self-adjoint operator, and its eigenspaces are the relevant subspaces. Thus for each λ that is an eigenvalue of H, the corresponding subspace where $H - \lambda I$ is zero is the event that the energy has value λ. In the following we shall write λ_α with α in some index set for the possible energy eigenvalues. The corresponding projections onto the eigenspaces are written F_α. If ψ is the state of the quantum system, then the probability that the energy is λ_α is given by $\langle \psi, F_\alpha \psi \rangle$.

Each eigenspace corresponding to the energy λ_α consists of eigenvectors. Denote a typical eigenvector by ϕ_α. Then $F_\alpha \phi_\alpha = \phi_\alpha$, so the probability $\langle \phi_\alpha, F_\alpha \phi_\alpha \rangle = 1$. Thus ϕ_α represents a state in which the energy is sure to be λ_α. These energy states are important in many physical problems. For this reason it is customary to talk almost interchangeably about energy events and energy states.

The energy H is the sum of two terms:

$$H = H_0 + V. \tag{4.1}$$

The first term represents kinetic energy; the second term represents potential energy. The kinetic energy and the potential energy are each easy to understand. The challenge is to understand the properties of the total energy. In particular, we want to understand the properties of the eigenfunctions of H. In the following we deal with the situation when the operator H acts in the space $\mathbf{C}^N$ corresponding to N sites in a periodic one-dimensional crystal.

Take a real number λ that is not an eigenvalue of H. Then the operator $(H - \lambda I)^{-1}$ exists. Consider a site m, and let δ_m be the function that is one at m and zero elsewhere. It follows that $(H-\lambda I)^{-1}\delta_m$ is also a well-defined function. The square of its norm is

$$\|(H-\lambda I)^{-1}\delta_m\|^2 = \sum_\alpha \frac{1}{(\lambda_\alpha-\lambda)^2}\langle \delta_m, F_\alpha \delta_m\rangle. \tag{4.2}$$

This can also be written as

$$\|(H-\lambda I)^{-1}\delta_m\|^2 = \sum_\alpha \frac{1}{(\lambda_\alpha-\lambda)^2}|\phi_\alpha(m)|^2, \tag{4.3}$$

where ϕ_α is the eigenfunction corresponding to eigenvalue λ_α. This comes from expanding the operator in eigenvectors of H. It is a measure of how many energy events near λ have significant probability when the state is located at site m. We shall use this quantity as an index of how much the individual energy events are extended in space. If the energy events are each extended in space, then they will all involve the site m, and the index $\|(H-\lambda I)^{-1}\|$ will be large. If the energy events are each localized in space, then only a few of these events will involve the site m; the rest will be localized somewhere else. In this latter case the index will be small.

In summary, the index of extension $\|(H-\lambda I)^{-1}\delta_m\|^2$ measures the extent to which eigenvalues close to λ have eigenfunctions that overlap significantly with site m. When the number of sites N is large, there can be many eigenvalues close to λ. The index of extension can still be bounded independently of N, provided that most of the eigenfunctions do not overlap very much with the site m. When this happens for each site m, the eigenfunctions are said to be localized.

4.2 Kinetic energy: extended energy states

We continue with the example of a one-dimensional periodic system with N points arranged in a circle. The complex vector space is $\mathbf{C}^N$ with the usual inner product.

The kinetic energy operator is a function of momentum. Recall that the unitary operator that translates in space is

$$Vf(n) = f(n-1) \tag{4.4}$$

where the subtraction is taken modulo N. This is a function of momentum. The adjoint operator is

$$V^*f(n) = f(n+1). \tag{4.5}$$

The kinetic energy operator is

$$H_0 = -[V^* + V - 2I]. \tag{4.6}$$

Thus it has the explicit form

$$H_0 f(n) = -[f(n+1) - 2f(n) + f(n-1)]. \tag{4.7}$$

The expression in brackets is the second difference operator that is the discrete analog of the second derivative.

The projection onto the subspace with the kth momentum value is the orthogonal projection onto the function whose value at n is

$$\phi_k(n) = \frac{1}{\sqrt{N}} \exp(\frac{2\pi i k n}{N}). \tag{4.8}$$

These are oscillatory functions that run over the entire crystal. Therefore they are called extended states.

If we apply H_0 to this function, we obtain

$$H_0 \phi_k = 2[1 - \cos(\frac{2\pi k}{N})]\phi_k. \tag{4.9}$$

This can also be written as

$$H_0 \phi_k = 4 \sin^2(\frac{\pi k}{N})\phi_k. \tag{4.10}$$

From this we see that the energy eigenvalues are all of the form

$$\lambda_k = 4 \sin^2(\frac{\pi k}{N}). \tag{4.11}$$

The values of k run from 0 up to $N - 1$. There are N energy eigenvalues, and they are all in the interval from 0 to 4. When N is large they are rather closely spaced, in fact distance between adjacent eigenvalues is always bounded by $4\pi/N$.

This picture of a band of closely spaced energies is the usual picture of a conduction band in the physics of solids. It describes those electrons in a metallic crystal that are free to move from site to site. The state of each such conduction electron is described by one of the energy states of H_0. These are extended states. So the probability for finding an electron is spread evenly over the entire crystal.

One mathematical representation of this fact is as follows. As usual, let δ_m be the function that is equal to 1 at site m and zero elsewhere. Then the position projection E_m is the projection onto δ_m. The probability that the position is at m when the state is ϕ_k is

$$\langle \phi_k, E_m \phi_k \rangle = \frac{1}{N}. \tag{4.12}$$

This is the same probability for all sites.

There is a complementary result in which the state has fixed position and one is interested in the value of the momentum. Recall that the projection onto the eigenspace for the kth momentum value is

$$F_k f(n) = \frac{1}{N} \sum_{r=1}^{N-1} \exp(\frac{2\pi i k(n-r)}{N}) f(r). \tag{4.13}$$

Conssider again the function δ_m that is 1 at site m and zero elsewhere. This represents a state at which the particle is sure to be at site m. It is easy to see that the probability that the momentum is k in this state is

$$\langle \delta_m, F_k \delta_m \rangle = \frac{1}{N}. \tag{4.14}$$

For the kinetic energy the index of extension is

$$\|(H_0 - \lambda I)^{-1} \delta_m\|^2 = \sum_k \frac{1}{(\lambda_k - \lambda)^2} \frac{1}{N}. \tag{4.15}$$

If λ is in the range from 0 to 4, then it is within a distance $2\pi/N$ from some λ_k. This shows that

$$\|(H_0 - \lambda I)^{-1} \delta_m\|^2 \geq \left(\frac{N}{2\pi}\right)^2 \frac{1}{N} = \frac{N}{4\pi^2}. \tag{4.16}$$

When N is large, the index of extension is huge. This is because each momentum eigenvector is extended over the whole crystal, and so all of them contribute to energy associated with the state at site m.

4.3 Potential energy: localized energy states

The potential energy is a function of position. It is specified by giving a sequence of real numbers v_n for $n = 0, 1, 2, \ldots, N-1$. The potential energy operator is

$$V = \sum_n v_n E_n. \tag{4.17}$$

This acts on a function by the formula

$$V f(n) = v_n f(n). \tag{4.18}$$

Thus V is given by a diagonal matrix.

For the potential energy the index of extension at site m is just

$$\|(V - \lambda I)^{-1} \delta_m\|^2 = \frac{1}{(v_m - \lambda)^2}. \tag{4.19}$$

This will be small unless by accident v_m happens to be close to λ. In any case, it is independent of N. A state located at m will only experience the one value of potential energy associated with m. The other potential energy eigenvectors do not overlap with m at all. This is localization at its most extreme manifestation.

4.4 Random environments

The Anderson model is an attempt to model a disordered system. This is given by taking each coefficient v_n to be random. Perhaps the simplest scheme is to take the v_n to be independent and to have uniform distribution on the interval $[-\epsilon, \epsilon]$. Each list of values of the v_n coefficients is represented by a point in an N dimensional cube of side 2ϵ. The probability associated with a subset S of the cube is just

$$\mathrm{Prob}[S] = \frac{\mathrm{volume}(S)}{\mathrm{volume}(\mathrm{cube})} = \frac{\mathrm{volume}(S)}{(2\epsilon)^N}, \tag{4.20}$$

where the volume is the volume in the N dimensional space. Since this is not the quantum mechanical probability, but rather is imposed from the outside, we shall refer to it as the environmental probability.

To think of these probabilities in a more intuitive manner, imagine a random number generator that generates random numbers uniformly in the interval $-\epsilon$ to ϵ. The values of the potential function are obtained by running the random number generator and recording the value, once for each site.

In the Anderson model we can compute environmental probabilities associated with the index of extension for the potential energy. Thus for $-\epsilon < \lambda < \epsilon$ we have

$$\mathrm{Prob}[\|(V - \lambda I)^{-1}\delta_m\|^2 \geq K] = \mathrm{Prob}[|v_m - \lambda| \leq \frac{1}{\sqrt{K}}] = \frac{1}{\epsilon\sqrt{K}} \tag{4.21}$$

for large K. Thus for fixed $\epsilon > 0$ the probability that the index is very large is small.

4.5 Total energy: localized energy states

The one-dimensional Anderson model describes a quantum mechanical particle with a random potential energy function. It would seem natural that if the randomness is rather small, then the eigenfuctions would somewhat resemble the eigenfunctions for the kinetic energy, and they would thus be extended in space. However, it turns out that the total energy in the one-dimensional Anderson model has only localized eigenfunctions.

The total energy is given by the operator

$$H = H_0 + V. \tag{4.22}$$

Again this operator will have eigenvalues λ_α and corresponding projections F_α onto the eigenspaces. But now they will be much more difficult to compute.

For the random potential energy function of the Anderson model, the corresponding H is a random operator. This means that it is a function of the v_n coefficients, and one computes probabilities with respect to the environmental probability measure defined above.

The following is the fundamental result about the index of extension for the Anderson model. It says that the index of extension is not likely to be large, no matter how large the crystal. If the crystal is large, there are more energy eigenfunctions, but the number that have significant overlap with the site m always remains about the same.

Theorem 4.1 *Let H be the energy operator in the crystal with N points. It is the sum of the kinetic energy H_0 with the potential energy V defined by a random potential function with independent values $v(n)$, each uniformly distributed on $[-\epsilon, \epsilon]$. For every $\epsilon > 0$ there are constants $C < \infty$ and $s > 0$ such that for all N and for all energy values λ and index bounds K we have*

$$\mathrm{Prob}[\|(H - \lambda I)^{-1}\delta_m\|^2 \geq K] \leq \frac{C}{K^s}. \tag{4.23}$$

The theorem says that by taking K extremely large, the environmental probability that $\|(H - \lambda I)^{-1}\delta_m\|$ exceeds K can be made so small that this will never be observed in practice. The remarkable thing is that this is independent of the size N of the crystal.

How can this be, if the eigenvalues λ_α are close to λ? They are going to be very close for large N, since the λ_α will be very closely spaced. The answer is that although λ_α is likely to be close to λ, this is more than compensated by the fact that the eigenfunction associated with λ_α is likely to be very close to zero except in an interval very far from m. This is the phenomenon of localization.

In physical terms, the localization of the eigenfunctions is saying that in a disordered one-dimensional system the electrons do not travel freely through the crystal, but each one hangs around near one spot. Even a slight randomness destroys their capability to propagate.

What is the situation in two and three dimensions? Do we have extended states, or localized states? It is known that if the randomness is very strong, then one has localized states. This famous result now has a relatively elementary proof; see the Aizenman and Molchanov paper mentioned below. Aside from this, almost all problems are open. It is not even known rigorously if it is possible to have extended states for a random potential, though it is universally believed that this is the case in three dimensions. For a discussion, see my article on random waves and localization (Notices of the American Mathematical Society 42 (1995), 848–853).

4.6 Estimating fractional powers

The proof of the theorem given above is not elementary, but it may be useful to sketch some of the ideas. These are ideas that are useful in various problems involving wave propagation in random environments.

We can compute the environmental expectation of functions X of the v_n coefficients by the formula

$$\mathrm{Expect}[X] = \frac{1}{(2\epsilon)^N} \int_{-\epsilon}^{\epsilon} \cdots \int_{-\epsilon}^{\epsilon} X(v_0, \ldots, v_{N-1})\, dv_0 \cdots dv_{N-1}. \tag{4.24}$$

As an example, we can try to take the environmental expectation of the index of extension for the potential energy. When λ is in the interval fro $-\epsilon$ to ϵ this is

$$\frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} \frac{1}{(v_m - \lambda)^2}\, dv_m = \infty. \tag{4.25}$$

The best we can do is to have some fractional power that has finite expectation. Thus if $s < 1/2$, we have

$$\frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} \frac{1}{(v_m - \lambda)^{2s}}\, dv_m < \infty. \tag{4.26}$$

A recent idea in this field is that the same should apply for the total energy: the expectation of a small fractional power of the index of extension should be finite. The most spectacular application of this idea was given in a paper of Aizenman and Molchanov (Communications in Mathematical Physics 157 (1993), 245–278). They dealt with the multidimensional case with large randomness parameter ϵ. There is a similar result in the one dimensional case with any value of $\epsilon > 0$.

Theorem 4.2 *Let H be the energy operator in the crystal with N points defined by the sum of the kinetic energy H_0 with the potential energy V given by a random potential function with independent values $v(n)$, each uniformly distributed on $[-\epsilon, \epsilon]$. Then for every $\epsilon > 0$ there are constants $C < \infty$ and $s > 0$ such that for all N and all λ*

$$\mathrm{Expect}[\|(H - \lambda I)^{-1}\delta_m\|^{2s}] \leq C. \tag{4.27}$$

The proof of this theorem is not elementary, but we can sketch some of the ingredients. The vector $g = (H - \lambda I)^{-1}\delta_m$ is the periodic function that solves the difference equation

$$-[g(n+1) - 2g(n) + g(n-1)] + v_n g(n) - \lambda g(n) = \delta_{mn} g(n). \tag{4.28}$$

This is a random difference equation, since the coefficients v_n are random. One needs properties of the solutions of random difference equations.

One useful device is to reduce this to a result about products of random matrices. Thus we could write the equation for $n \neq m$ by introducing a second coordinate $h(n) = g(n-1)$ so that

$$\begin{pmatrix} g(n+1) \\ h(n+1) \end{pmatrix} = \begin{pmatrix} (v_n + 2 - \lambda) & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} g(n) \\ h(n) \end{pmatrix}. \tag{4.29}$$

Thus the solution obtained by iterating this equation is given by a product of 2 by 2 matrices, each with determinant one.

The book by Bougerel and Lacroix (Products of Random Matrices with Applications to Schrödinger Operators, Birkhäuser, Boston, 1985) describes the theory of such matrix products. There is a kind of instability: The solutions of such random equations tend either to grow exponentially or to decay exponentially. In the case of the solution above, the function $g(n)$ should decay exponentially from the source point m, at least up to the maximum distance from m around the circle. The rate of exponential decay should depend on the randomness parameter ϵ but not on the size of the system N. The theorem is a way of making this idea quantitative.

The proof that the theorem about probabilities follows from the theorem on expectations is elementary. In fact, it follows from the Chebyshev inequality. The Chebyshev is a very general and simple probability result. It says that if ϕ is a positive increasing function on the range of X, then

$$\mathrm{Prob}[X \geq a] \leq \frac{\mathrm{Expect}[\phi(X)]}{\phi(a)}. \tag{4.30}$$

The proof is simple. If $X \geq a$, then since ϕ is increasing, it follows that $\phi(X) \geq \phi(a)$. Hence $\mathrm{Prob}[X \geq a] \leq \mathrm{Prob}[\phi(X) \geq \phi(a)]$. Define $Y = \phi(a)$ if $\phi(X) \geq \phi(a)$ and $Y = 0$ otherwise. Then $Y \leq \phi(X)$, so $\mathrm{Expect}[Y] \leq \mathrm{Expect}[\phi(X)]$. This says $\phi(a)\,\mathrm{Prob}[\phi(X) \geq \phi(a)] \leq \mathrm{Expect}[\phi(X)]$. Hence $\phi(a)\,\mathrm{Prob}[X \geq a] \leq \mathrm{Expect}[\phi(X)]$, which is Chebyshev's inequality.

For the problem at hand we get the estimate by taking the increasing function to be $\phi(x) = x^s$ for $x \geq 0$. Then the Chebyshev inequality gives

$$\mathrm{Prob}[\|(H - \lambda I)^{-1}\delta_m\|^2 \geq K] \leq \frac{\mathrm{Expect}[\|(H - \lambda I)^{-1}\delta_m\|^{2s}]}{K^s} \leq \frac{C}{K^s}. \tag{4.31}$$