

# Formalism of quantum mechanics

August 22, 2014

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Classical physics</b>	<b>1</b>
<b>3</b>	<b>Quantum physics</b>	<b>3</b>
<b>4</b>	<b>Harmonic oscillator</b>	<b>10</b>

## 1 Introduction

Quite a bit of the serious mathematical theory of self-adjoint operators was created to serve the needs of quantum mechanics. These notes are a quick-and-dirty outline of the simplest mathematical setting of quantum mechanics. None of it should be taken too seriously: real physics is hard, and requires more than a few nice mathematical ideas.

## 2 Classical physics

In order to describe a toy picture of quantum mechanics, I'll start with a toy picture of classical mechanics. In the toy picture, the state of a physical system is described by a point  $x$  belonging to a *state space*  $X$ . For example, if we want to study the earth going around the sun classically, we might think of the universe as  $\mathbb{R}^3$ , and the sun as fixed at the center of the universe (and therefore at the point  $0 \in \mathbb{R}^3$ ). The state of the earth is then given by two vectors in  $\mathbb{R}^3$ : its *position*  $\mathbf{p}$  and its *velocity*  $\mathbf{v}$ . The *state* of the earth is the pair of vectors

$$x = (\mathbf{p}, \mathbf{v}) \in \mathbf{R}^3 \oplus \mathbf{R}^3 \simeq \mathbf{R}^6 = X. \quad (2.1)$$

Everything we want to know about the earth (assuming we take a sufficiently superior attitude toward television and so on) is recorded in the pair of vectors  $(\mathbf{p}, \mathbf{v}) \in \mathbf{R}^6$ .

Of course the state of the earth changes with time: it moves, with a changing velocity. What we are really interested in is the entire history of the earth, which is a function assigning to each time  $t \in \mathbb{R}$  a state  $x(t)$ . For example, if we choose coordinates so that the earth's orbit is in the plane of the first two coordinates, and take the orbit to be a perfect circle, the function describing the earth now is something like

$$x(t) = ((R \cos(\phi + 2\pi t/T), R \sin(\phi + 2\pi t/T), 0), \\ (-2\pi R/T) \sin(\phi + 2\pi t/T), (2\pi R/T) \cos(\phi + 2\pi t/T), 0). \quad (2.2)$$

Here

$$R = \text{distance to the sun} \approx 1.496 \times 10^{11} \text{ meters,}$$

$$T = \text{length of the year} \approx 3.15569 \times 10^7 \text{ seconds,}$$

and  $\phi$  is the angular position of the earth in its orbit at time zero.

The job of classical physics is to tell you how to find the future  $x(t)$  from knowledge of the state  $x(t_0)$  at one time  $t_0$ . Typically the answer is in the form of a differential equation

$$\frac{dx}{dt} = F(x, t), \quad (2.3)$$

which says that how the state of the system changes depends on its present state and the present time. The right side  $F(x, t)$  is a "direction" at  $x$ . If  $X$  is inside a vector space, then  $F(x, t)$  is always a vector in that same space; so  $F(x, t)$  is a vector field on  $X$  (maybe changing with time.) If you know what a manifold is, and  $X$  is a manifold, then  $F(x, t)$  belongs to the tangent space  $T_x(X)$  to  $X$  at  $x$ .

For the motion of the earth around the sun, Newton's laws give the differential equation as

$$\frac{d(\mathbf{p}, \mathbf{v})}{dt} = (\mathbf{v}(t), \frac{-GM_s \mathbf{p}(t)}{\|\mathbf{p}(t)\|^3}).$$

(The first three coordinates of the equation are just the definition of velocity. What's interesting is how the velocity is changing.) Here  $G$  is Newton's gravitational constant, and  $M_s$  is the mass of the sun. (By comparing this equation with the formula for  $x(t)$  above, you can get a value for  $GM_s$ .)

A nice feature of this equation is that it depends only on  $x(t)$  and not otherwise on  $t$ : the laws of physics are not changing with time.

In addition to predicting the future, physicists like to look at things. A *classical observable* is a real-valued function on the state space  $X$ :

$$A: X \rightarrow \mathbb{R}.$$

If the system is in the state  $x \in X$ , making the observation yields the number  $A(x)$ . The equation of motion (2.3) and the chain rule tell you how the observable  $A$  changes as the system evolves:

$$\frac{dA(x(t))}{dt} = (F(*, t) \cdot A)(x(t)). \quad (2.4)$$

The rate of change of  $A$  is gotten by taking the directional derivative of  $A$  in the direction  $F$  (in abstract mathematical language, applying the vector field  $F$  to  $A$ ) then evaluating at  $x(t)$ .

**Definition 2.5.** The observable  $A$  is *conserved* by the classical physical system (2.3) if  $F(*, t) \cdot A = 0$ ; that is, if  $A$  is constant in the directions of the vector fields  $F(*, t)$ . In this case, the value of  $A$  is constant in  $t$  for any possible history  $x(t)$ .

For the motion of the earth around the sun, a typical observable is the distance to the sun:

$$d(\mathbf{p}, \mathbf{v}) = \|\mathbf{p}\|.$$

Even though this observable happens to be constant in the circular orbit solution (2.2), it is not conserved: there are other possible histories (like elliptical orbits) in which it is not constant.

### 3 Quantum physics

So here is the corresponding toy picture of quantum mechanics. The underlying mathematical structure of a state space  $X$  is replaced by a *complex inner product space*, often denoted  $\mathcal{H}$ . (This is what we have called  $V$  in class.) A *state* of the physical system is a line (a one-dimensional complex subspace) of  $\mathcal{H}$ . Another way to say this is that a state is a nonzero vector  $\psi \in \mathcal{H}$ , and that  $\psi$  and  $z\psi$  define the same state whenever  $z$  is a nonzero complex number. (The vector  $\psi$  is just like the vectors we've been calling  $v$ ; I switched to the Greek letter only to follow common physics notation.)

A lot of sources say that a state is a vector  $\psi$  with  $\|\psi\| = 1$  (called a *unit vector*) and that the unit vectors  $\psi$  and  $e^{i\theta}\psi$  define the same state. (In

the Copenhagen Interpretation of quantum mechanics, this corresponds to the idea that no experiment can be designed that will distinguish between the state  $\psi$  and the state  $e^{i\theta}\psi$ .)

I'll stick with the idea that a state is a line  $\mathbb{C}\psi$ , and that the chosen basis vector  $\psi$  for the line need not be a unit vector.

Just as in classical mechanics, the state of the system changes in time. In general, this evolution is not described as a changing line, but as a changing basis vector for the line. That is, there is supposed to be a function

$$\psi: \mathbb{R} \rightarrow \mathcal{H} - 0, \quad \mathbb{C} \cdot \psi(t) = \text{state at time } t. \quad (3.1)$$

The laws of physics are supposed to be summarized by a self-adjoint operator

$$H \in \mathcal{L}(\mathcal{H}), \quad H^* = H, \quad (3.2)$$

called the *Hamiltonian* of the physical system. I'll say more in a moment about the physical interpretation of this linear transformation. The first point is that the quantum-mechanical version of (2.3) is the *Schrödinger equation*

$$\frac{d\psi(t)}{dt} = \frac{1}{i\hbar} H\psi(t). \quad (3.3)$$

The first big difference from (2.3) is that the Schrödinger equation is required to be *linear*. This seems to make quantum mechanics simpler than classical mechanics. One reason that quantum mechanics can stay frightening is that the vector space  $\mathcal{H}$  is most often infinite-dimensional.

Suppose  $\psi(t)$  is a solution of the Schrödinger equation. Let us see how the length of the vector  $\psi(t) \in \mathcal{H}$  changes in time. We calculate

$$\begin{aligned} \frac{d\langle \psi(t), \psi(t) \rangle}{dt} &= \left\langle \frac{d\psi(t)}{dt}, \psi(t) \right\rangle + \left\langle \psi(t), \frac{d\psi(t)}{dt} \right\rangle && \text{(product rule)} \\ &= \left\langle \frac{1}{i\hbar} H\psi(t), \psi(t) \right\rangle + \left\langle \psi(t), \frac{1}{i\hbar} H\psi(t) \right\rangle && \text{(Schrödinger)} \\ &= \left\langle \frac{1}{i\hbar} H\psi(t), \psi(t) \right\rangle - \left\langle \frac{1}{i\hbar} \psi(t), H\psi(t) \right\rangle && \text{(Hermitian)} \\ &= \left\langle \frac{1}{i\hbar} H\psi(t), \psi(t) \right\rangle - \left\langle \frac{1}{i\hbar} H\psi(t), \psi(t) \right\rangle && \text{(selfadjoint)} \\ &= 0. \end{aligned} \quad (3.4)$$

Therefore

$$\text{a solution } \psi(t) \text{ to Schrödinger's equation (3.3) has constant length.} \quad (3.5)$$

A *quantum observable* is a self-adjoint linear transformation on

$$A \in \mathcal{L}(\mathcal{H}), \quad A^* = A. \quad (3.6a)$$

If  $\mathcal{H}$  is finite-dimensional, then the spectral theorem says

$$\mathcal{H} = \bigoplus_{\lambda \in \mathbb{R}} \mathcal{H}_\lambda, \quad (3.6b)$$

with  $\mathcal{H}_\lambda$  the  $\lambda$ -eigenspace of the observable  $A$ . (If  $\mathcal{H}$  is infinite-dimensional, there are still versions of the spectral theorem available. The number of eigenvalues may be infinite, and proving the spectral theorem requires more work; but the finite-dimensional case still gives a reasonable picture of what is going on.)

The possible values of the observable  $A$  are the real numbers that are eigenvalues of  $A$ . If a state  $\mathbb{C}\psi$  is contained in the eigenspace  $\mathcal{H}_\lambda$ , then we say that the *value of the observable is  $\lambda$* .

The central idea of quantum mechanics is this: most vectors in  $\mathcal{H}$  (and therefore most states of the corresponding physical system) do not belong to *any* one eigenspace. Instead the vectors (and so the states) are *linear combinations* (the physics word is *superpositions*) of many eigenvectors. What this means is that the observable is *simultaneously taking on many different values*. This isn't how classical observables work, but quantum mechanics is just different.

To be more concrete, suppose that the observable  $A$  has the spectral decomposition written in (3.6b). If  $\mathbb{C}\psi$  is a state, then the eigenspace decomposition of  $\psi$  is

$$\psi = \sum_{\lambda \in \mathbb{R}} \psi_\lambda, \quad A\psi_\lambda = \lambda\psi_\lambda. \quad (3.6c)$$

Because the decomposition is orthogonal, the Pythagorean Theorem says

$$\langle \psi, \psi \rangle = \sum_{\lambda \in \mathbb{R}} \langle \psi_\lambda, \psi_\lambda \rangle, \quad (3.6d)$$

or equivalently

$$1 = \sum_{\lambda \in \mathbb{R}} \frac{\langle \psi_\lambda, \psi_\lambda \rangle}{\langle \psi, \psi \rangle}. \quad (3.6e)$$

On the right we have a bunch of nonnegative numbers adding up to one. Whenever you see such a thing, you should think *probability*. In this case, one of the standard ways of thinking about quantum mechanics is

$$\frac{\langle \psi_\lambda, \psi_\lambda \rangle}{\langle \psi, \psi \rangle} = \text{probability that observable } A \text{ has value } \lambda. \quad (3.6f)$$

The Copenhagen Interpretation raises this idea a bit higher: it says that if the quantum system is in the state  $\mathbb{C}\psi$ , and you perform an experiment to measure the observable  $A$ , then the probability that you will measure  $\lambda$  is equal to  $\frac{\langle \psi_\lambda, \psi_\lambda \rangle}{\langle \psi, \psi \rangle}$ . Quantum mechanics doesn't tell you what *will* happen: it offers a library of possible outcomes (the eigenvalues of  $A$ ) and tells you the probability of seeing each of them.

One of the basic ideas in probability is the idea of *expected value*. This is the average of all possible outcomes of some trial, with the outcomes weighted by their probability. If I have five test papers in a hat, with the scores 58, 63, 79, 87, and 98, and I draw one paper at random from the hat, then the expected value of the score is

$$\frac{1}{5} \cdot 58 + \frac{1}{5} \cdot 63 + \frac{1}{5} \cdot 79 + \frac{1}{5} \cdot 87 + \frac{1}{5} \cdot 98 = 77,$$

the ordinary average of the scores. If I roll a fair die, the expected number appearing is

$$\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3.5.$$

(Notice that this “expected value” cannot actually occur.)

In the Copenhagen Interpretation, the expected value of the experiment to measure the value of the observable  $A$  is

$$\begin{aligned} E(A) &= \sum_{\lambda \in \mathbb{R}} \frac{\langle \psi_\lambda, \psi_\lambda \rangle}{\langle \psi, \psi \rangle} \cdot \lambda && \text{(weighted average of outcomes)} \\ &= \sum_{\lambda \in \mathbb{R}} \frac{\langle \lambda \psi_\lambda, \psi_\lambda \rangle}{\langle \psi, \psi \rangle} \\ &= \sum_{\lambda \in \mathbb{R}} \frac{\langle \lambda \psi_\lambda, \psi \rangle}{\langle \psi, \psi \rangle} && \text{(orthogonality of eigenspaces)} \\ &= \frac{\langle \sum_{\lambda \in \mathbb{R}} \lambda \psi_\lambda, \psi \rangle}{\langle \psi, \psi \rangle} && \text{(linearity of inner product)} \\ &= \frac{\langle \sum_{\lambda \in \mathbb{R}} A \psi_\lambda, \psi \rangle}{\langle \psi, \psi \rangle} \\ E(A) &= \frac{\langle A\psi, \psi \rangle}{\langle \psi, \psi \rangle}. \end{aligned} \tag{3.6g}$$

This expected value varies continuously with the quantum state, from a maximum of the largest eigenvalue of  $A$  (if the state belongs to the corresponding eigenspace) to a minimum of the smallest eigenvalue of  $A$ .

It's natural to try to concentrate on observables that you can reliably observe; that is, to look only at states  $\psi \in \mathcal{H}_\lambda$  that are eigenvalues for  $A$ . The essential difficulties of quantum mechanics arise when you are interested in two different observables  $A$  and  $A'$ , so that there are two different eigenspace decompositions

$$\mathcal{H} = \bigoplus_{\lambda \in \mathbb{R}} \mathcal{H}_{\lambda,A}, \quad \mathcal{H} = \bigoplus_{\lambda' \in \mathbb{R}} \mathcal{H}_{\lambda',A'}. \quad (3.6h)$$

In this situation you might like to concentrate on states in which you can reliably observe both  $A$  and  $A'$ ; that is, states that are in a “simultaneous eigenspace”

$$\mathcal{H}_{(\lambda,\lambda')} = \mathcal{H}_{\lambda,A} \cap \mathcal{H}_{\lambda',A'}. \quad (3.6i)$$

The difficulty is that all these spaces can be zero: there may be *no* simultaneous eigenspaces of  $A$  and  $A'$ . This is not something pathological, but rather entirely typical.

*Example 3.7.* Suppose  $\mathcal{H} = \mathbb{C}^2$ , so that self-adjoint operators are  $2 \times 2$  matrices

$$\begin{pmatrix} a & z \\ \bar{z} & b \end{pmatrix} \quad (a \in \mathbb{R}, z \in \mathbb{C}).$$

Two natural observables are

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Each of these observables has eigenvalues  $+1$  and  $-1$ ; each can take just those two values “classically.” (In honest physics, observables like this arise for example in describing the polarization of light.) The spectral decompositions are

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_{1,A} \oplus \mathcal{H}_{-1,A} \\ &= \mathbb{C} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus \mathbb{C} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_{1,A'} \oplus \mathcal{H}_{-1,A'} \\ &= \mathbb{C} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \oplus \mathbb{C} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

The eigenspaces for  $A$  are the two coordinate axes; and the eigenspaces for  $A'$  are the two diagonal lines. There are no simultaneous eigenvectors except  $0$ , which does not represent a state (since a quantum state is a one-dimensional subspace).

Problem set 9 says that if the observables  $A$  and  $A'$  commute, then there is a simultaneous eigenspace decomposition. (It's easy to see that commuting is a necessary condition for the simultaneous eigenspace decomposition: do you see why?) So *quantum-mechanical difficulties are attached to observables that do not commute*. Example 3.7 says that this is mathematically very common. One of the fundamental physical principles of quantum mechanics is that *the observables position and momentum never commute*. That is, the most basic physical observables—position and velocity—can never be reliably observed at the same time. That's what the *Heisenberg uncertainty principle* is about. Some details may be found in Section 4 (see (4.5b)).

The Hamiltonian  $H$  appearing in the Schrödinger equation (3.3) is a self-adjoint operator, so it is also an observable. It's the most important observable of all: the *energy* of the physical system. The eigenvalues of  $H$  are the possible energies that the system can have. If  $\psi_0 \in \mathcal{H}_E$  is in the  $E$ -eigenspace—that is, if  $\psi_0$  has energy equal to  $E$ —then  $H\psi_0 = E\psi_0$ , and we get an easy solution to the Schrödinger equation:

$$\psi(t) = e^{Et/i\hbar}\psi_0. \quad (3.8)$$

That is, a state purely in the energy level  $E$  does not change; only its phase oscillates, with a frequency of

$$\nu = 2\pi E/\hbar = E/h. \quad (3.9)$$

This is Planck's relation between frequency and energy: originally found for photons, but now making some kind of sense for any quantum-mechanical system.

**Definition 3.10.** The observable  $A$  is *conserved* by the quantum physical system (3.3) if the expected value of the observable

$$\frac{\langle A\psi(t), \psi(t) \rangle}{\langle \psi(t), \psi(t) \rangle}$$

is constant in time.

**Proposition 3.11.** *Suppose*

$$\psi: \mathbb{R} \rightarrow \mathcal{H}$$

*is a solution of the Schrödinger equation (3.3), and that  $A \in \mathcal{L}(\mathcal{H})$  is an observable. Then the expected value of  $A$*

$$E(A)(t) = \frac{\langle A\psi(t), \psi(t) \rangle}{\langle \psi(t), \psi(t) \rangle}$$



satisfies the differential equation

$$\frac{dE(A)}{dt} = E\left(\frac{1}{i\hbar}(AH - HA)\right).$$

Here  $\frac{1}{i\hbar}(AH - HA) \in \mathcal{L}(\mathcal{H})$  is a self-adjoint linear transformation (and so a new observable).

In particular,  $A$  is conserved if and only if  $AH = HA$ ; that is, if and only if the linear maps  $A$  and  $H$  commute with each other.

To a mathematician, the proposition above suggests that one should replace each physical observable

$$A = \text{selfadjoint operator} \quad (A^* = A) \quad (3.12a)$$

by a “mathematical observable”

$$S_A = \text{skew-adjoint operator} \quad (S^* = -S), \quad (3.12b)$$

according to the rule

$$S_A = \frac{1}{i\hbar}A, \quad A = i\hbar S. \quad (3.12c)$$

Given two mathematical observables  $S$  and  $S'$ , one can form a new mathematical observable

$$[S, S'] = SS' - S'S, \quad (3.12d)$$

called the *commutator* of  $S$  and  $S'$ . This operation, called *commutator* or *Lie bracket* makes the mathematical observables into a *Lie algebra*. The Schrödinger equation takes the form

$$\frac{d\psi(t)}{dt} = S_H\psi(t). \quad (3.12e)$$

The differential equation in the proposition above is

$$\frac{dE(A)}{dt} = E(i\hbar[S_A, S_H]). \quad (3.12f)$$

So the commutator describes exactly the how observables change in time.

## 4 Harmonic oscillator

This possible future section will write down the classical and quantum models for a one-dimensional harmonic oscillator, and compare their solutions. For now I'll just write a little. The inner product space is

$$\mathcal{H} = L^2(\mathbb{R}) = \{\psi: \mathbb{R} \rightarrow \mathbb{C} \mid \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty\}. \quad (4.1)$$

We'll be vague about exactly which functions  $\psi$  are allowed; this is one of the places where the details are difficult.

The inner product is like the one used on functions on  $[0, 1]$  in the text and problem sets:

$$\langle \psi_1, \psi_2 \rangle = \int_{-\infty}^{\infty} \psi_1(x) \overline{\psi_2(x)} dx. \quad (4.2)$$

Recall that a state corresponds to a line  $\mathbb{C}\psi \subset \mathcal{H}$ . We think of it as corresponding to a (quantum) particle living at some indeterminate place on the real line; the size  $|\psi(x)|^2$  of the function represents the probability that the particle is at  $x$ . (A little more precisely, the integral

$$\frac{1}{\langle \psi, \psi \rangle} \int_a^b |\psi(x)|^2 dx$$

is the probability that the particle is between  $a$  and  $b$ .)

The classical observable *position* corresponds to the (selfadjoint) quantum observable

$$(M\psi)(x) = x\psi(x). \quad (4.3a)$$

That is, we multiply the function  $\psi$  by the function  $x$ . (The letter  $M$  is chosen to be a reminder of “multiply.” The eigenspace for a real number  $\lambda$  is

$$\begin{aligned} \mathcal{H}_{\lambda, M} = \mathcal{H}_{\lambda} &= \{\psi \mid x\psi(x) = \lambda\psi(x)\} \\ &= \{\psi \mid \psi(x') = 0, x' \neq \lambda\}. \end{aligned} \quad (4.3b)$$

That is, the eigenspace consists of multiples of the Dirac delta “function”

$$\delta_{\lambda}, \quad \delta_{\lambda}(x) = \begin{cases} \infty & x = \lambda, \\ 0 & x \neq \lambda. \end{cases} \quad (4.3c)$$

The “function”  $\delta_{\lambda}$  is not in  $L^2(\mathbb{R})$ , because it is too concentrated at the point  $\lambda$ . (I can't make this precise, because I haven't said exactly what functions are allowed in  $L^2(\mathbb{R})$ ). The idea is that a *nice* function on the real line that is

zero except at one point must be zero everywhere.) Nevertheless, one should think of every real number  $\lambda$  as an eigenvalue of  $M$ , with eigenvector  $\delta_\lambda$ .

Even though it is not an actual function, certain integrals involving  $\delta_\lambda$  make sense. What is required (and what becomes the mathematical definition of  $\delta_\lambda$ ) is that for any nice function  $\phi$ ,

$$\int_{-\infty}^{\infty} \phi(x) \delta_\lambda(x) dx = \phi(\lambda) \quad (4.3d)$$

Using the formal “definition” (4.3c), it is clear that

$$\delta_\lambda(x) = \delta_x(\lambda). \quad (4.3e)$$

Using the mathematical definition hinted at in (4.3d), it is possible to make this statement meaningful and precise; but I won’t worry about it.

In the case of a self-adjoint operator  $A$  on a finite-dimensional vector space  $V$ , we can choose a basis of eigenvectors

$$(e_{\lambda_1}, \dots, e_{\lambda_n}), \quad Ae_{\lambda_i} = \lambda_i e_{\lambda_i}$$

Then any vector in  $v$  has a finite sum expansion  $v = \sum_i a_i e_{\lambda_i}$ ; the magnitude  $|a_i|^2$  tells how much of  $v$  is in the direction of the eigenvalue  $\lambda_i$ .

In this infinite-dimensional, setting, the finite sum is replaced by an integral

$$\phi = \int_{-\infty}^{\infty} a(\lambda) \delta_\lambda d\lambda \quad (4.3f)$$

In order to see what the coefficients  $a(\lambda)$  of the various eigenvectors  $\delta_\lambda$  should be, first plug in  $x$ ; then use the formula (4.3e); and finally use (4.3d):

$$\begin{aligned} \phi(x) &= \int_{-\infty}^{\infty} a(\lambda) \delta_\lambda(x) d\lambda \\ &= \int_{-\infty}^{\infty} a(\lambda) \delta_x(\lambda) d\lambda = a(x). \end{aligned} \quad (4.3g)$$

The conclusion is that

$$a(x) = \phi(x), \quad \phi = \int_{-\infty}^{\infty} \phi(\lambda) \delta_\lambda d\lambda. \quad (4.3h)$$

That is, the size of  $|\phi(\lambda)|^2$  measures how much of the state  $\phi$  is at the position  $\lambda$ .

The classical observable *momentum* corresponds to the (selfadjoint) quantum observable

$$(D\psi)(x) = -i\hbar \frac{d\psi}{dx}, \quad (4.4a)$$

a multiple of the derivative of  $\psi$ . (The letter  $D$  is chosen to be a reminder of “derivative.” The eigenspace for a real number  $\mu$  is

$$\begin{aligned} \mathcal{H}_{\mu,D} = \mathcal{H}_\mu &= \{\psi \mid D\psi(x) = \mu\psi(x)\} \\ &= \{\psi(x) = Ae^{i\mu x/\hbar}\}. \end{aligned} \quad (4.4b)$$

That is, the  $\mu$  eigenspace consists of multiples of the exponential function

$$e_{\mu/\hbar}, \quad e_\xi(x) =_{\text{def}} e^{2\pi i \xi x}. \quad (4.4c)$$

These exponential functions are not in  $L^2(\mathbb{R})$ , because they are too spread out over  $\mathbb{R}$ : the absolute value squared is one everywhere, so cannot have finite integral. Nevertheless, one should think of every real number  $\mu$  as an eigenvalue of  $D$ , with eigenvector  $e_{\mu/\hbar}$ .

Just as in the case of position, we want to replace the finite sum in the finite-dimensional spectral theorem by an integral, and write any vector  $\phi$  as an integral of eigenvectors:

$$\phi = \int_{-\infty}^{\infty} b(\mu) e_{\mu/\hbar} d\mu \quad (4.4d)$$

In order to see what the coefficients  $b(\mu)$  of the various momentum eigenvectors  $e_{\mu/\hbar}$  should be, first plug in  $x$ , then perform the change of variable  $\xi = \mu/\hbar$ :

$$\begin{aligned} \phi(x) &= \int_{-\infty}^{\infty} b(\mu) e_{\mu/\hbar} d\mu \\ &= \frac{1}{\hbar} \int_{-\infty}^{\infty} b(\hbar\xi) e_\xi d\xi \end{aligned} \quad (4.4e)$$

The integral on the left is an *inverse Fourier transform*. Here are the basic definitions and facts. If  $\phi$  and  $\psi$  are nice complex-valued functions on  $\mathbb{R}$ , then the *Fourier transform of  $\phi$*  is a new (nice) complex-valued function on  $\mathbb{R}$ , defined by

$$\hat{\psi}(\xi) = \int_{-\infty}^{\infty} \psi(x) e_{-\xi}(x) dx = \int_{-\infty}^{\infty} \psi(x) e^{-2\pi i x \xi} dx. \quad (4.4f)$$

The *inverse Fourier transform of  $\phi$*  is the new (nice) complex-valued function on  $\mathbb{R}$  defined by

$$\check{\phi}(x) = \int_{-\infty}^{\infty} \phi(\xi) e_x(\xi) d\xi = \int_{-\infty}^{\infty} \phi(\xi) e^{2\pi i x \xi} d\xi. \quad (4.4g)$$

The main theorem (Fourier inversion) is that the Fourier transform is an *isometry* (length-preserving invertible linear map) on  $L^2(\mathbb{R})$ , with inverse equal to the inverse Fourier transform:

$$(\hat{\phi})^\sim = \phi, \quad \|\hat{\phi}\| = \|\phi\| \quad (\phi \in L^2(\mathbb{R})). \quad (4.4h)$$

Comparing (4.4e) with (4.4g), we conclude that

$$b(\mu) = \hat{\phi}(\mu/h), \quad \phi = \frac{1}{h} \int_{-\infty}^{\infty} \hat{\phi}(\mu/h) e_{\mu/h} d\mu. \quad (4.4i)$$

That is, the size of  $|\hat{\phi}(\mu/h)|^2$  measures how much of the state  $\phi$  has momentum  $\mu$ .

The main point of all this is that the states corresponding to a precise position  $\lambda$ —the “Dirac delta functions”  $\delta_\lambda$ —are completely different from the states corresponding to a precise momentum  $\mu$ —the complex exponentials  $e_{\mu/h}$ . We saw in a pset that commuting selfadjoint operators could be simultaneously diagonalized: that there would in lots of simultaneous eigenstates, where both observables are precisely known. The position and momentum operators  $P$  and  $D$  do *not* commute; in fact

$$\begin{aligned} (MD - DM)\phi &= -ix\hbar \frac{d\phi}{dx} + i\hbar \frac{d}{dx}(x\phi) \\ &= -ix\hbar \frac{d\phi}{dx} + i\hbar \left(x \frac{d\phi}{dx} + \phi\right) \\ &= i\hbar\phi. \end{aligned} \quad (4.5a)$$

That is,

$$[M, D] = i\hbar, \quad (4.5b)$$

the *canonical commutation relation* of Heisenberg. If we use the language of skew-adjoint “mathematical observables” discussed in (3.12), then the skew-adjoint operators corresponding to position and momentum are

$$S_M = x/i\hbar, \quad S_D = -\frac{d}{dx}, \quad [S_M, S_D] = \frac{1}{i\hbar}. \quad (4.5c)$$

As explained in the text, no such commutation relations can be satisfied by operators on a finite-dimensional inner product space: the left side has trace equal to zero, and the right side has trace equal to  $(i\hbar)^{\pm 1}$  times the dimension of the space. So any quantum mechanics involving positions and momenta has to live in an infinite-dimensional inner product space.

So you need to take more math classes after this one!

MIT OpenCourseWare  
<http://ocw.mit.edu>

18.700 Linear Algebra  
Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.