# Lecture 9 – Quantum mechanical description of physical systems
MATH-GA 2710.001 Mechanics

## 1   Limitations of classical mechanics

There are several ways of motivating the study of quantum mechanics. A common way is a historical presentation of the limitations of classical mechanics in explaining key experimental observations at the turn of the 19th and 20th centuries. The "ultraviolet catastrophe" is one of the most famous examples: classical electromagnetism fails to correctly predict the radiated power per unit frequency of a black body at thermal equilibrium, and predicts that such a black body will emit radiation with infinite power. O. Bühler has a clear discussion of this in the first pages of the chapter on quantum mechanics in his book *A Brief Introduction to Classical, Statistical and Quantum mechanics.*

Another way of motivating the study of quantum mechanics is to describe a fairly simple experiment whose results are very confusing for someone used to classical mechanics, and can only be explained within the quantum framework. Young's double slits experiment with electrons is a canonical example in that respect. In my opinion, the most brilliant description and interpretation of that experiment has been given by Richard Feynman, and can be seen in the following video: https://youtu.be/aAgcqgDc-YM. Given the brilliance of Feynman and this video, we will not add any more words on this topic, and recommend everyone to listen to this great, great mind.

## 2   Mathematical formulation of quantum mechanics

### 2.1   Representing quantum states

We saw in Lectures 1 and 2 that in classical mechanics the knowledge of the position $\mathbf{r}(t_0)$ and of the velocity $\dot{\mathbf{r}}(t_0)$ (or equivalently the momentum $\mathbf{p}(t_0)$ was necessary and sufficient to uniquely determine the physical evolution of a system for later times and/or earlier times. In other words, the classical state at a given instant in time $t_0$ is uniquely determined by $(\mathbf{r}(t_0), \mathbf{p}(t_0))$.

In quantum mechanics, the situation is quite different. The quantum state of a particle is represented

- **Either** by a function $\Psi(\mathbf{r}, t)$ usually called the *wave function*

- **Or** (exclusive or) by the wave function in *momentum space* $\Phi(\mathbf{p}, t)$

The function $\Psi(\mathbf{r}, t)$ (or $\Phi(\mathbf{p}, t)$) describes the maximal amount of available knowledge about the particle before a measurement is made. If a measurement is made, the probability $P(\mathbf{r}, t)d\mathbf{r}$ that the position of the particle at time $t$ is in the infinitesimal volume $d\mathbf{r}$ is given by

$$|\Psi(\mathbf{r}, t)|^2 d\mathbf{r} = P(\mathbf{r}, t)d\mathbf{r}$$

In other words, $\Psi(\mathbf{r}, t)$ allows us to predict the probability density function $P(\mathbf{r}, t)$ of $\mathbf{r}$ associated with a measurement of position.

Likewise, $\Phi(\mathbf{p}, t)$ allows us to predict the probability density function $Q(\mathbf{p}, t)$ of $\mathbf{p}$ associated with a measurement of momentum, with

$$|\Phi(\mathbf{p}, t)|^2 d\mathbf{r} = Q(\mathbf{p}, t)d\mathbf{p}$$

$\Psi$ and $\Phi$ are complex valued functions, and $|\cdot|^2$ in the formulae above represents their modulus squared.

A quantum state can thus be described in several ways, two of which we have just seen: an "$\mathbf{r}$ representation", through $\Psi$, and a "$\mathbf{p}$ representation" through $\Phi$. There are actually infinitely many ways of representing a quantum state, each highlighting particular variables. We will soon learn how to go from one representation to the other.

### 2.2   Hilbert space

The functions $\Psi$ and $\Phi$ are in $L^2$ Hilbert spaces with the following inner products

$$\langle \chi, \phi \rangle = \int_V \chi^* \phi \, d\mathbf{r}$$

for wavefunctions in $\mathbf{r}$ space, where $V$ is the volume of the configuration space, and

$$\langle \chi, \phi \rangle = \int_{V_p} \chi^* \phi \, d\mathbf{p}$$

for wavefunctions in momentum space, where $V_p$ is the momentum space volume.

The $\mathbf{r}$ wave function $\Psi$ and the $\mathbf{p}$ wave function $\Phi$ can be expressed in terms of one another through the Fourier transform relations

$$\Psi(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{V_p} \Phi(\mathbf{p}, t) \exp\left(\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}\right) d\mathbf{p} \tag{1}$$

$$\Phi(\mathbf{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{V} \Psi(\mathbf{x}, t) \exp\left(-\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}\right) d\mathbf{r} \tag{2}$$

From Parseval's theorem and the fact that one is certain to find the particle somewhere in the volume $V$, we can write

$$\langle \Psi, \Psi \rangle = \int_{V} |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} = \int_{V_p} |\Phi(\mathbf{p}, t)|^2 d\mathbf{p} = \langle \Phi, \Phi \rangle = 1$$

Finally, note that by differentiating under the integral sign, one can write

$$\nabla \Psi(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{V_p} \frac{i\mathbf{p}}{\hbar} \Phi(\mathbf{p}, t) \exp\left(\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}\right) d\mathbf{p}$$

and

$$\nabla^2 \Psi(\mathbf{r}, t) = -\frac{1}{(2\pi\hbar)^{3/2}} \int_{V_p} \frac{\mathbf{p}^2}{\hbar^2} \Phi(\mathbf{p}, t) \exp\left(\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}\right) d\mathbf{p}$$

In other words, the Fourier transform of $i\mathbf{p}/\hbar\Phi(\mathbf{p}, t)$ is $\nabla\Psi(\mathbf{r}, t)$, and the Fourier transform of $-\mathbf{p}^2/\hbar^2\Phi(\mathbf{p}, t)$ is $\nabla^2\Psi$. We will use these facts very soon.

## 2.3   Representing physical quantities

We have just seen how to describe the probability density functions $P(\mathbf{r}, t)$ for $\mathbf{r}$ and $Q(\mathbf{p}, t)$ for $\mathbf{p}$. Since the dynamical quantities of interest depend on $\mathbf{r}$ and $\mathbf{p}$, we hope to be able to build on this to construct the probability density of any physical quantity. This is the goal of this subsection. We start with a general discussion on probability distributions.

### *Characterising probability density functions*

Let $u$ be a random variable defined on an interval $\Omega$ with an associated probability density function $P(u)$. $P(u_0)du$ is then the probability of finding $u$ in $[u_0, u_0+du]$. The $n-th$ moment of the probability distribution is the quantity defined by

$$M_n = \overline{u^n} \equiv \int_{\Omega} u^n P(u) du$$

We are usually well acquainted with the first few moments. By definition of the probability density function, we must have

$$M_0 = \int_{\Omega} P(u) du = 1$$

The mean $\bar{u}$ of $u$ is given by $M_1$:

$$\bar{u} = \int_{\Omega} u P(u) du$$

The variance $\Delta u^2$ can be expressed in terms of $M_2$ and $M_1$ as follows:

$$\Delta u^2 = \overline{u^2} - \bar{u}^2 = M_2 - M_1^2$$

Now, one may naturally ask oneself the following question: if we do not know $P(u)$ itself but know all its moments $M_n$, can we use this information to compute $P(u)$? The answer is yes, and here is a simple way to see this. Consider the function

$$Q(v) = \overline{\exp(iuv)} = \int_{\Omega} P(u) \exp(iuv) du \ , \ v \in \mathbb{R} \tag{3}$$

$Q$ is sometimes called the *characteristic function* for the probability density $P$. Note that for all $u$ in $\Omega$, $P(u)$ can be calculated from $Q$ using the inverse Fourier transform:

$$P(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Q(v) \exp(-iuv) dv$$

So the question becomes: can one compute $Q(v)$ if we know all the moments $M_n$? The answer is clearly yes, since we can write the exponential in the integrand in (3) as an infinite series:

$$Q(v) = \int_{\Omega} P(u) \sum_{n=0}^{\infty} \frac{(iuv)^n}{n!} du = \sum_{n=0}^{\infty} \frac{(iv)^n}{n!} \int_{\Omega} u^n P(u) du = \sum_{n=0}^{\infty} \frac{(iv)^n}{n!} M_n$$

The consequence of this result is as follows. Suppose we know how to evaluate the mean of some physical quantity $A$. We can then use the same method to calculate the mean of $A^n$ for any $n$, from which we can evaluate the characteristic function $Q$ and hence the probability density function $P$. We will show how this works next. Before we do so, we give the expression for $Q$ for the case in which $u$ is a discrete random variable with probabilities $P_1, P_2, \ldots, P_k$ (the set of the $P_k$ is possibly infinite, provided $\sum_k P_k = 1$). The discrete version of Eq.(3) is

$$Q(v) = \overline{\exp(iuv)} = \sum_k P_k \exp(ivu_k) \tag{4}$$

### Computing the mean value of physical quantities

Here we will work at fixed $t$, so for the simplicity of the notation, we will omit the time dependence and write $\Psi(\mathbf{r})$ or $\Phi(\mathbf{p})$. Let us start by finding the expression for the mean value of $x$ and $p_x$ in $\mathbf{r}$ representation. The formula for $x$ is quite simple

$$\overline{x} = \int_V x P(\mathbf{r}) d\mathbf{r} = \int_V \Psi(\mathbf{r})^* x \Psi(\mathbf{r}) d\mathbf{r} = \langle \Psi(\mathbf{r}), x\Psi(\mathbf{r}) \rangle \tag{5}$$

Similar expressions can be written for $\overline{y}$ and $\overline{z}$.

To obtain $\overline{p_x}$ in the $\mathbf{r}$ representation, we start with the expression of $\overline{p_x}$ in the $\mathbf{p}$ representation:

$$\overline{p_x} = \int_{V_p} p_x Q(\mathbf{p}) d\mathbf{p} = \int_{V_p} \Phi^*(\mathbf{p}) p_x \Phi(\mathbf{p}) d\mathbf{p} = \langle \Phi(\mathbf{p}), p_x \Phi(\mathbf{p}) \rangle \tag{6}$$

Now, the Fourier transform preserves the inner product, and we have proved that $\nabla\Psi(\mathbf{r}, t)$ is the Fourier transform of $ip_x/\hbar \Phi(\mathbf{p}, t)$, so

$$\langle \Phi(\mathbf{p}), \frac{ip_x}{\hbar} \Phi(\mathbf{p}) \rangle = \langle \Psi(\mathbf{r}), \frac{\partial \Psi(\mathbf{r})}{\partial x} \rangle$$

We can now use this identity in (6) to write

$$\overline{p_x} = \langle \Psi(\mathbf{r}), \frac{\hbar}{i} \frac{\partial \Psi(\mathbf{r})}{\partial x} \rangle \tag{7}$$

Identical expressions can be written for $\overline{p}_y$ and $\overline{p}_z$.

Both Equations (5) and (7) have the form

$$\langle \Psi(\mathbf{r}), \hat{O}\Psi(\mathbf{r}) \rangle$$

where $\hat{O}$ is a linear operator acting on $\Psi$. In the case of Eq. (5), the operator is a multiplication by $x$, in the case of Eq. (7), the operator is the differential operator $\hbar/i \partial/\partial x$. From now on, we will associate to any physical quantity an operator acting on the wavefunctions; this operator will be represented by the symbol of the physical variable with a hat on top. So far, we showed that

$$\hat{x}\Psi(\mathbf{r}) = x\Psi(\mathbf{r}) \qquad \hat{p}_x \Psi = \frac{\hbar}{i} \frac{\partial \Psi}{\partial x}$$

and equivalent expressions for $y$, $z$, $p_y$ and $p_z$. Let us emphasize once more the difference here between the operators associated with the coordinates – simple multiplications – and the operators associated with the

3

momenta. This asymmetry is not surprising since we are working with the "**r** representation" here. You can convince yourself that if we had worked with the "**p** representation", the situation would be inverted. The operators associated with the momenta would then be simple multiplications, and the operators associated with the coordinates would be differential operators.

Let us now turn to a key physical quantity that in classical mechanics is a function of the position of the particle and of its momentum: the Hamiltonian. Let us consider the case in which the forces acting on the particle can be written in terms of the gradient of a potential $U(\mathbf{r})$ that does not depend on time so that the Hamiltonian is

$$H(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + U(\mathbf{r}) = E$$

where $m$ is the mass of the particle, and $E$ the energy of the particle.

The mean value of the energy of the particle is

$$\overline{E} = \frac{\overline{\mathbf{p}^2}}{2m} + \overline{U(\mathbf{r})}$$

In the **r** representation, the mean value of $U(\mathbf{r})$ has a simple expression:

$$\overline{U}(\mathbf{r}) = \int_V U(\mathbf{r})|\Psi(\mathbf{r})|^2 d\mathbf{r} = \langle \Psi(\mathbf{r}), U(\mathbf{r})\Psi(\mathbf{r}) \rangle$$

In other words,

$$\hat{U}\Psi(\mathbf{r}) = U(\mathbf{r})\Psi(\mathbf{r})$$

For $\overline{\mathbf{p}^2}$, we use the fact that the Fourier transform of $\mathbf{p}^2/\hbar^2 \Phi(\mathbf{p}, t)$ is $\nabla^2 \Psi(\mathbf{r}, t)$ and the fact that the Fourier transform preserves the inner product to write

$$\overline{\mathbf{p}^2} = \langle \Phi(\mathbf{p}), \mathbf{p}^2 \Phi(\mathbf{p}) \rangle = \langle \Psi(\mathbf{r}), -\hbar^2 \nabla^2 \Psi \rangle$$

We can thus write $\widehat{\mathbf{p}^2} = -\hbar^2 \nabla^2$. Strictly speaking, this is the product of two operators

$$\widehat{\mathbf{p}^2} = \frac{\hbar}{i}\nabla \cdot \frac{\hbar}{i}\nabla = -\hbar^2 \nabla^2$$

i.e. $\widehat{\mathbf{p}^2}$ is a product of operators of the form $\hat{A}\hat{B}$, so that

$$\hat{A}\hat{B}\Psi(\mathbf{r}) = \hat{A}(\hat{B}\Psi(\mathbf{r}))$$

We apply $\hat{B}$ first, and then $\hat{A}$. In general, the order of the operators matters, as we will soon see.

So we have constructed the linear operator associated with the Hamiltonian:

$$\hat{H} = \frac{\widehat{\mathbf{p}^2}}{2m} + \hat{U}(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r})$$

### Correspondence principle

The first three operators we have constructed here illustrate a general principle often known as the *correspondence principle*, first formulated by Niels Bohr and which can be stated as follows. To compute the mean value of any physical quantity $A(\mathbf{r}, \mathbf{p})$, one should take the expression for $A(\mathbf{r}, \mathbf{p})$ given by classical mechanics and replace the variables $x$, $y$, $z$, $p_x$, $p_y$, $p_z$ with their associated linear operator to form the linear operator $\hat{A}$ that can be used to calculate the mean value of $A$ according to

$$\overline{A} = \langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle \tag{8}$$

Using the correspondence principle, the mean values

$$\overline{A^n} = \langle \Psi(\mathbf{r}), \hat{A}^n \Psi(\mathbf{r}) \rangle$$

can also be calculated, and following our discussion on moments of a probability density function, this means that we can then reconstruct the probability density function associated with the random variable $A$.

### Non-commutative operators

The principle of correspondence is somewhat incomplete in that it does not say in which orders the operators should be applied. We did not run into this problem when constructing the operator associated with the Hamiltonian, because the two successive operators where momentum operators. Commutativity matters for mixed operators of the type $\hat{x}\hat{p}_x$ and $\hat{p}_x\hat{x}$. Indeed, for any $L^2$ function $\Psi$, we can write

$$\hat{x}\hat{p}_x = x\left(\frac{\hbar}{i}\frac{\partial\Psi}{\partial x}\right) = \frac{\hbar}{i}x\frac{\partial\Psi}{\partial x}$$

$$\hat{p}_x\hat{x} = \frac{\hbar}{i}\frac{\partial}{\partial x}\left(x\Psi(\mathbf{r})\right) = \frac{\hbar}{i}\left(\Psi(\mathbf{r}) + x\frac{\partial\Psi}{\partial x}\right)$$

In other words,

$$(\hat{x}\hat{p}_x - \hat{p}_x\hat{x})\Psi(\mathbf{r}) = i\hbar\Psi(\mathbf{r})$$

The operator $(\hat{x}\hat{p}_x - \hat{p}_x\hat{x})$ is called the *commutator* of $\hat{x}$ and $\hat{p}_x$, written $[\hat{x}, \hat{p}_x] \equiv (\hat{x}\hat{p}_x - \hat{p}_x\hat{x})$. We have

$$[\hat{x}, \hat{p}_x] = i\hbar\hat{I} \qquad [\hat{y}, \hat{p}_y] = i\hbar\hat{I} \qquad [\hat{z}, \hat{p}_z] = i\hbar\hat{I}$$

where $\hat{I}$ is the identity operator, and where the second and third equality can be easily proved in the same way as we have done for $\hat{x}$ and $\hat{p}_x$.

It is clear that all the coordinate operators commute with one another, as do the momentum operators. It can also be seen that $\hat{x}$ commutes with $\hat{p}_y$ and $\hat{p}_z$, and so on. We will show later that the only nonzero commutators are obtained for variables that are not compatible and have an associated Heisenberg uncertainty relation.

Now that we saw that some operators do not commute, the question is: how do we choose the order of the operators? This is what we address next.

### Constructing Hermitian operators

The mean value $\overline{A}$ of a physical quantity $A$ is a real quantity. From Eq. (8), we can thus write

$$\langle\Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r})\rangle = \left(\Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r})\right)^* = \langle\hat{A}\Psi(\mathbf{r}), \Psi(\mathbf{r})\rangle$$

In other word, the operator $\hat{A}$ associated with a physical quantity must be self-adjoint (also called Hermitian).

The operator $\hat{x}$ is clearly self-adjoint, by commutativity of the multiplication and the fact that $x$ is a real quantity:

$$\langle\Psi(\mathbf{r}), \hat{x}\Psi(\mathbf{r})\rangle = \int_V \Psi^*(\mathbf{r})x\Psi(\mathbf{r})d\mathbf{r} = \int_V (x\Psi(\mathbf{r}))^*\Psi(\mathbf{r})d\mathbf{r} = \langle\hat{x}\Psi(\mathbf{r}), \Psi(\mathbf{r})\rangle$$

Likewise, $\hat{p}_x$ is self-adjoint. Indeed,

$$\langle\Psi(\mathbf{r}), \hat{p}_x\Psi(\mathbf{r})\rangle = \int_V \Psi(\mathbf{r})^*\frac{\hbar}{i}\frac{\partial\Psi}{\partial\mathbf{r}}d\mathbf{r} = -\int_V \left(\frac{\hbar}{i}\Psi(\mathbf{r})\right)^*\frac{\partial\Psi}{\partial\mathbf{r}}d\mathbf{r} = \int_V \left(\frac{\hbar}{i}\frac{\partial\Psi}{\partial\mathbf{r}}\right)^*\Psi(\mathbf{r})d\mathbf{r} = \langle\hat{p}_x\Psi(\mathbf{r}), \Psi(\mathbf{r})\rangle$$

where we have used integration by parts and the fact that $\Psi$ vanishes on the boundary of $V$.

The operator $\hat{x}\hat{p}_x$, on the other hand, is not Hermitian. As an exercise, you may verify that

$$\langle\Psi(\mathbf{r}), \hat{x}\hat{p}_x\Psi(\mathbf{r})\rangle = \langle\hat{p}_x\hat{x}\Psi(\mathbf{r}), \Psi(\mathbf{r})\rangle$$

To make the operator associated with $\hat{x}\hat{p}_x$ self-adjoint, we thus replace $\hat{x}\hat{p}_x$ and $\hat{p}_x\hat{x}$ with the symmetric half-sum

$$\frac{1}{2}\left(\hat{x}\hat{p}_x + \hat{p}_x\hat{x}\right)$$

This is the general rule that has to be followed for non-Hermitian operators.

### Certainty

We now know how to compute the mean value of any physical quantity $A$ and all the moments of its probability density function. At that point, one may ask what happens if $A$ can be predicted with certainty. This is the case if and only if the variance of $A$ is zero:

$$\overline{A^2} = \overline{A}^2 \qquad \Leftrightarrow \qquad \langle \Psi(\mathbf{r}), \hat{A}^2 \Psi(\mathbf{r}) \rangle = \left( \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \right)^2 \tag{9}$$

Now, since $\hat{A}$ is Hermitian, we can write

$$\langle \Psi(\mathbf{r}), \hat{A}^2 \Psi(\mathbf{r}) \rangle = \langle \Psi(\mathbf{r}), \hat{A}(\hat{A}\Psi(\mathbf{r})) \rangle = \langle \hat{A}\Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle$$

Using the fact that $\langle \Psi(\mathbf{r}), \Psi(\mathbf{r}) \rangle = 1$, Eq. (9) can be rewritten as

$$\langle \Psi(\mathbf{r}), \Psi(\mathbf{r}) \rangle \langle \hat{A}\Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle = \left( \langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle \right)^2 = \left| \langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle \right|^2 \tag{10}$$

where the last equality holds because $\langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle$ is a real quantity.

From the Cauchy-Schwarz inequality, we know that

$$\langle \Psi(\mathbf{r}), \Psi(\mathbf{r}) \rangle \langle \hat{A}\Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle \geq \left| \langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle \right|^2$$

and equality holds if $\Psi(\mathbf{r})$ and $\hat{A}\Psi(\mathbf{r})$ are colinear. So Eq.(10) tells us that there exists a real constant $a$ such that

$$\hat{A}\Psi(\mathbf{r}) = a\Psi(\mathbf{r}) \tag{11}$$

and clearly $\overline{A} = a$, i.e. the result of any measurement of $A$ is $a$.

We just showed that certainty of the measurement is obtained when $\Psi(\mathbf{r})$ is an eigenfunction of the Hermitian operator $\hat{A}$. The result of any measurement then is the associated eigenvalue.

### Observables with discrete spectrum

A physical quantity whose associated Hermitian operator has eigenfunctions which form a basis of the Hilbert space are called *observable*. It can be shown that the physical quantities we will deal with in this class are observable. Now, consider the set of all the eigenvalues of an operator, called the *spectrum* of the operator. If the set is discrete, we say that the spectrum is discrete. Otherwise, the spectrum is said to be continuous. In general, the full spectrum of an operator can be made in part by a discrete spectrum and in part by a continuous spectrum. This is the case for the Hamiltonian operator associated with the hydrogen atom, for example: bound states of the electron correspond to the discrete spectrum, and the ionized atom has a continuous spectrum.

We first look at the case of an observable quantity $A$ with a discrete spectrum, and consider the basis of the Hilbert space made of the eigenfunctions of $\hat{A}$, which we normalize so they have norm 1. It can be the case that several eigenfunctions have the same eigenvalue. This eigenvalue is then said to be degenerate. In our notation, we thus use two indices: a first index $n$ numbering a given eigenvalue $a_n$, and a second index $(k)$ numbering the eigenfunction $\phi_{n,(k)}(\mathbf{r})$ of the degenerate eigenvalue $a_n$.

If two eigenfunctions have two distinct eigenvalues $a_n$ and $a_m$, they are orthogonal. Using the Gram-Schmidt process, we can also construct eigenfunctions of a given degenerate eigenvalue $a_n$ to be orthogonal to one another. The whole basis then is orthonormal:

$$\langle \phi_{n,(k)}(\mathbf{r}), \phi_{m,(l)}(\mathbf{r}) \rangle = \delta_{mn}\delta_{kl}$$

where $\delta_{ij}$ is the Kronecker delta symbol: $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ if $i \neq j$.

Suppose now that the physical system is described by the wave function $\Psi(\mathbf{r})$. Writing $\Psi(\mathbf{r})$ in the orthonormal basis we just constructed, we have

$$\Psi(\mathbf{r}) = \sum_{n,(k)} c_{n,(k)}\phi_{n,(k)}(\mathbf{r}) \tag{12}$$

where the complex coefficients $c_{n,(k)}$ are given by

$$c_{n,(k)} = \langle \phi_{n,(k)}(\mathbf{r}), \Psi(\mathbf{r}) \rangle$$

If we apply $\hat{A}$ to $\Psi(\mathbf{r})$, we have

$$\hat{A}\Psi(\mathbf{r}) = \sum_{n,(k)} c_{n,(k)} a_n \phi_{n,(k)}(\mathbf{r})$$

The average value of $A$ thus is

$$\overline{A} = \langle \Psi(\mathbf{r}), \hat{A}\Psi(\mathbf{r}) \rangle = \langle \sum_{m,(l)} c_{m,(l)} \phi_{m,(l)}, \sum_{n,(k)} c_{n,(k)} a_n \phi_{n,(k)} \rangle = \sum_{m,(l)} \sum_{n,(k)} c_{m,(l)} c_{n,(k)} a_n \langle \phi_{m,(l)}(\mathbf{r}), \phi_{n,(k)}(\mathbf{r}) \rangle$$

$$= \sum_{n,(k)} a_n |c_{n,(k)}|^2 = \sum_n a_n \sum_{(k)} |c_{n,(k)}|^2$$

Likewise, we have

$$\overline{A^p} = \sum_n a_n^p \sum_{(k)} |c_{n,(k)}|^2 = \sum_n a_n^p P_n$$

where

$$P_n = \sum_{(k)} |c_{n,(k)}|^2$$

The characteristic function $Q$ for the probability density $P(A)$ is

$$Q(v) = \overline{\exp(iAv)} = \sum_{p=0}^{\infty} \frac{(iv)^p}{p!} \overline{A^p} = \sum_n P_n \left( \sum_{p=0}^{\infty} \frac{(iva_n)^p}{p!} \right) = \sum_n P_n \exp(ia_n v) \tag{13}$$

Comparing Eq.(13) with Eq.(4), we see that $Q(v)$ is the characteristic function of a discrete random variable which can take the values $a_n$ with associated probability $P_n$. The physical interpretation is that the only possible values one can obtain from a measurement of $A$ are the eigenvalues $a_n$ of the operator $\hat{A}$. The probability $P_m$ of measuring the particular value $a_m$ is

$$P_m = \sum_{(k)} |c_{m,(k)}|^2$$

In a representation associated with the basis $\phi_{n,(k)}$, the state of the system is described by the column vector of the $c_{n,(k)}$. We say that the $c_{n,(k)}$ are the *discrete probability amplitude* associated with the operator $A$, whose squared moduli allow us to compute the probabilities $P_n$. If one works in the "$\mathbf{r}$ representation", the $P_n$ are computed according to

$$P_n = \sum_{(k)} |\langle \phi_{n,(k)}(\mathbf{r}), \Psi(\mathbf{r}) \rangle|^2$$

### Observables with continuous spectrum

If the spectrum of the operator associated with a physical quantity is continuous, then the wave function $\Psi$ can still be written in terms of a basis $\phi_a(\mathbf{r})$ of eigenfunctions, in the following integral form

$$\Psi(\mathbf{r}) = \int_{\Omega} \gamma(a) \phi_a(\mathbf{r}) da \tag{14}$$

where $\Omega$ is the interval corresponding to the continuous spectrum.

Let us look at two concrete examples that we already know and illustrate this situation. We will quickly see that it is much subtler from a mathematical point of view than the discrete spectrum.

- Eigenfunctions of the momentum operator $\hat{\mathbf{p}}$

  The eigenfunctions of the operator $\hat{\mathbf{p}} = \hbar/i \nabla$ satisfy

$$\frac{\hbar}{i} \nabla \phi = \mathbf{p} \phi$$

where we have called $\mathbf{p}$ the eigenvalues of the operator $\hat{\mathbf{p}}$. Clearly, the plane waves

$$\phi(\mathbf{r}) = K \exp\left(i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}\right)$$

are the eigenfunctions of the momentum operator. This is not surprising if one remembers De Broglie's insight that to any plane wave with wave vector $\mathbf{k}$ one can associate the momentum $\mathbf{p} = \hbar\mathbf{k}$. In that case, the expansion (14) is just Eq. (1), where the probability amplitudes $\Phi(\mathbf{p}, t)$ play the role of the amplitude $\gamma(a)$, and we saw that $|\Phi(\mathbf{p}, t)|^2$ is the probability density of $\mathbf{p}$.

It is important to note a mathematical difficulty here: the plane waves, eigenfunctions of $\hat{\mathbf{p}}$, are a basis for functions in $L^2$ but do not belong to that Hilbert space themselves!

- Eigenfunctions of the operator $\hat{\mathbf{r}}$

  The case of the operator $\hat{\mathbf{r}}$ is even stranger. Let us call $\mathbf{a}$ the eigenvalues of $\hat{\mathbf{r}}$. The eigenfunctions of $\hat{\mathbf{r}} = \mathbf{r}$ satisfy

  $$(\mathbf{r} - \mathbf{a})\Psi(\mathbf{r}) = \mathbf{0}$$

  In other words, $\phi(\mathbf{r})$ is zero for all $\mathbf{r} \neq \mathbf{a}$, but must have unit norm. We recognize here that $\phi$ is the Dirac delta "function" $\delta(\mathbf{r} - \mathbf{a})$ such that

  $$\langle\delta(\mathbf{r} - \mathbf{a}), \phi(\mathbf{r})\rangle = \int_V \delta(\mathbf{r} - \mathbf{a})\phi(\mathbf{r})d\mathbf{r} = \phi(\mathbf{a})$$

  The delta "function" does not belong to $L^2$ either.

We have highlighted here a general property of eigenfunctions for a continuous spectrum: their inner product and norm cannot be defined in the usual sense. However, distribution theory resolves these issues. Loosely speaking, the functions seen above have to be seen as the continuous limit (in the $L^2$ sense) of a sequence of functions in $L^2$. Distribution theory allows us to generalize the concept of orthonormality for eigenfunctions $\phi_a(\mathbf{r})$ and $\phi_b(\mathbf{r})$ according to the relation

$$\langle\phi_a(\mathbf{r}), \phi_b(\mathbf{r})\rangle = \delta(a - b)$$

where the right-hand side is the one-dimensional Dirac delta "function". This is how $\gamma(a)$ in (14) can be considered a probability amplitude, and $|\gamma(a)|^2$ is the probability density of the variable $a$.

### Eigenvalue equation for the Hamiltonian

In the $\mathbf{r}$ representation, the eigenvalue equation for any operator associated with a physical quantity is a partial differential equation. In the remainder of this class, the physical quantity we will most often consider is the energy, whose operator is the Hamiltonian. For a particle subject to forces that can be written in terms of the potential $U(\mathbf{r})$, the eigenvalue equation for the eigenvalues $E$ of the Hamiltonian is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r}) + U(\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$

This equation implies that the second derivatives of $\Psi$ exist. They can be discontinuous if the potential $U$ is discontinuous. However, $\Psi$ as well as its first derivatives must be continuous.

## 2.4 Temporal evolution

### The role of time in non-relativistic quantum mechanics

Thus far, we have hidden the time dependence of all the quantities, for the simplicity of the notation. In reality, we have $\Psi(\mathbf{r}, t)$, $\Phi(\mathbf{r}, t)$, $\gamma(a, t)$, $c_{n,(k)}(t)$. $t$ should be seen as a parameter indicating the time at which a measurement is potentially made. If no measurement is made, it characterizes the evolution of the probability amplitudes. In non-relativistic quantum mechanics, time is not considered as a physical quantity, and it does not have an associated operator. Moreover, measurements are considered instantaneous.

We will now see how to describe the evolution of the probability amplitudes as a function of time.

### *Two types of time evolution*

One of the major peculiarities of quantum mechanics is the distinction between two types of time evolution: 1) the evolution of the probability amplitudes in the absence of measurements; 2) the modification of the probability amplitudes when a measurement is made.

1. **Evolution in the absence of measurement**

   A postulate of quantum mechanics is that the evolution of the probability amplitudes in the absence of measurements is *deterministic*: the knowledge of the quantum state at one instant in time is sufficient to compute the quantum state at any later stage, as long as a measurement is not made. The partial differential equation must be linear because the phenomena satisfy the principle of linear superposition: if $\Psi_1$ and $\Psi_2$ are solutions and $\lambda_1$ and $\lambda_2$ are two constants, then $\lambda_1\Psi_1 + \lambda_2\Psi_2$ also is a solution. The partial differential equation must also be first order in time, since we have seen that unlike classical mechanics, the knowledge of $\Psi(\mathbf{r}, t_0)$ alone is sufficient to specify the quantum state. A postulate of quantum mechanics is that the time evolution in the absence of measurement is given by the *Schrödinger equation*, which satisfies all the properties mentioned above. We will present the *Schrödinger equation* shortly.

2. **Evolution during a measurement**

   During an instantaneous measurement, another type of evolution takes place, in which the measurement apparatus is involved. Suppose that a measurement has been made at the time $t_1$. Just before the measurement, at time $t_1 - \epsilon$, the amplitude $\Psi(\mathbf{r}, t_1 - \epsilon)$ gives the probability density for the measurement of a given value of the physical quantity $A$. However, just after the measurement, at time $t_1 + \epsilon$, our information on the system has changed since a given value $a$ was measured. Experiments show that if one repeats the same measurement immediately after, one finds the same value $a$ as long as $\epsilon$ is short enough: nature does not contradict itself! One must therefore admit that after the measurement the *a posteriori probability amplitude $\Psi(\mathbf{r}, t_1 + \epsilon)$ is reduced to the projection of the a priori amplitude $\Psi(\mathbf{r}, t_1 - \epsilon)$ on the subspace corresponding to the eigenvalue $a$.*

   Consider for example an operator with a discrete spectrum, and the expansion (12) on the basis of eigenfunctions for $\Psi(\mathbf{r}, t_1 - \epsilon)$. If the measurement gives the value $a_n$, then $\Psi(\mathbf{r}, t_1 + \epsilon)$ will be reduced to

   $$\Psi(\mathbf{r}, t_1 + \epsilon) = K \sum_{(k)} c_{n,(k)}(t_1) \phi_{n,(k)}(\mathbf{r})$$

   where $K$ is a coefficient that renormalizes the probabily properly:

   $$K = \frac{1}{\sqrt{\sum_{(k)} |c_{n,(k)}(t_1)|^2}}$$

   In the case of a continuous spectrum, a measurement does not yield an exact value $a$ because there always is some uncertainty $\Delta a$ associated with the measurement, as we will discuss later. To the interval with size $\Delta a$ around the value $a$ one associates the probability $P(a)\Delta a = |\gamma(a)|^2 \Delta a$. After the measurement, the wave function is given by an integral of the form

   $$\Psi(\mathbf{r}, t_1 + \epsilon) = K \int_{\Gamma} \gamma(a)\phi_a(\mathbf{r}) da$$

   where $\Gamma$ is the interval of size $\Delta a$ around $a$ and $K$ is the renormalizing constant.

   This *wave function collapse* has to be seen as another postulate of quantum mechanics, describing the irreversible evolution of the wave function following a measurement. This postulate summarizes and models the complex chain of processes taking place during a measurement. Exciting experimental and theoretical research is currently being done to better understand the origins of this postulate and the mechanisms at play in a phenomenon also known as *decoherence*. Nobel prizes were awarded in 2005 and 2012 to scientists working on these questions.

### *The Schrödinger equation*

A central postulate of quantum mechanics is that in the absence of measurement, the time evolution of a quantum system with Hamiltonian $H$ and associated operator $\hat{H}$ is given by the Schrödinger equation

$$\hat{H}\Psi(\mathbf{r},t) = i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} \tag{15}$$

The Schrödinger equation has two key properties that we will prove: 1) it conserves the norm; 2) it agrees with the classical evolution in the limit in which the dimensions of the wave packet (and therefore all the wavelengths) are small compared to the characteristic dimensions of the problem.

1. Conservation of the norm

   Let $\Psi(\mathbf{r},t)$ be a solution of the Schrödinger equation. Let us compute

   $$\begin{aligned}\frac{d}{dt}\left(\langle\Psi(\mathbf{r},t),\Psi(\mathbf{r},t)\rangle\right) &= \langle\frac{\partial\Psi(\mathbf{r},t)}{\partial t},\Psi(\mathbf{r},t)\rangle + \langle\Psi(\mathbf{r},t),\frac{\partial\Psi(\mathbf{r},t)}{\partial t}\rangle\\ &= \frac{i}{\hbar}\langle\hat{H}\Psi(\mathbf{r},t),\Psi(\mathbf{r},t)\rangle - \frac{i}{\hbar}\langle\Psi(\mathbf{r},t),\hat{H}\Psi(\mathbf{r},t)\rangle\end{aligned}$$

   where we have used the fact that $\Psi$ satisfies Eq.(15), the antilinearity of the inner product with respect to the left entries, and the linearity of the inner product with respect to the right entries. Now, since $\hat{H}$ is a Hermitian operator, the last expression is zero:

   $$\frac{d}{dt}\left(\langle\Psi(\mathbf{r},t),\Psi(\mathbf{r},t)\rangle\right) = 0$$

2. Classical limit

   We want to show that in the limit in which the probability distributions have a negligible spread and in which we can see the mean values of the distributions as the deterministic quantities of classical mechanics the Schödinger equation agrees with the laws for the evolution of classical systems we have encountered in previous lectures.

   To do this, we look at the time derivative of the space variable $x$ and the momentum variable $p_x$. Since the calculation is similar in both cases, let us call $\hat{O}$ the associated operator to either physical quantity. We have

   $$\begin{aligned}\frac{d\overline{O}}{dt} &= \frac{d}{dt}\left(\langle\Psi(\mathbf{r},t),\hat{O}\Psi(\mathbf{r},t)\rangle\right)\\ &= \langle\frac{\partial\Psi(\mathbf{r},t)}{\partial t},\hat{O}\Psi(\mathbf{r},t)\rangle + \langle\Psi(\mathbf{r},t),\hat{O}\frac{\partial\Psi(\mathbf{r},t)}{\partial t}\rangle\end{aligned}$$

   Using the Schrödinger equation, we can rewrite this as

   $$\frac{d\overline{O}}{dt} = \frac{i}{\hbar}\left(\langle\hat{H}\Psi(\mathbf{r},t),\hat{O}\Psi(\mathbf{r},t)\rangle - \langle\Psi(\mathbf{r},t),\hat{O}\hat{H}\Psi(\mathbf{r},t)\rangle\right)$$

   and since $\hat{H}$ is Hermitian,

   $$\frac{d\overline{O}}{dt} = \frac{i}{\hbar}\left(\langle\Psi(\mathbf{r},t),\hat{H}\hat{O}\Psi(\mathbf{r},t)\rangle - \langle\Psi(\mathbf{r},t),\hat{O}\hat{H}\Psi(\mathbf{r},t)\rangle\right) = \frac{i}{\hbar}\langle\Psi(\mathbf{r},t),\left[\hat{H},\hat{O}\right]\Psi(\mathbf{r},t)\rangle = \frac{i}{\hbar}\overline{\left[\hat{H},\hat{O}\right]} \tag{16}$$

   Equation (16) is important. Note first that if $\hat{O}$ is the identity operator $\hat{I}$, then $\overline{I}$ is just the norm of $\Psi(\mathbf{r},t)$. Since $\hat{I}$ commutes with all operators, Eq.(16) is a proof, in that case, of the norm conserving property we proved earlier.

   A second observation one can make is that Eq. (16) proves that if the operator associated with a physical quantity commutes with the Hamiltonian operator, the mean value of this physical quantity is a conserved quantity: its time derivative is zero.

Let us now assign the operators $\hat{x}$ and $\hat{p_x}$ to $\hat{O}$. Thus far, we have proved that

$$\begin{cases} \frac{d\overline{x}}{dt} = \frac{1}{i\hbar} \overline{\left[\hat{x}, \hat{H}\right]} \\ \frac{d\overline{p_x}}{dt} = \frac{1}{i\hbar} \overline{\left[\hat{p_x}, \hat{H}\right]} \end{cases}$$

Let us see how to evaluate $\left[\hat{x}, \hat{H}\right]$. The same technique applies to the evaluation of $\left[\hat{p_x}, \hat{H}\right]$.

The first idea is to expand the classical expression of the Hamiltonian $H(x, y, z, p_x, p_y, p_z)$ in a power series in $p_x$, which is the only variable whose operator does not commute with $\hat{x}$. The coefficients in the series expansion then only depend on variables whose operators commute with $\hat{x}$.

We then compute the commutator $[\hat{x}, \hat{p_x^n}]$ for any $n$ by induction. We start with

$$[\hat{x}, \hat{p_x^2}] = (\hat{x}\hat{p_x})\hat{p_x} - \hat{p_x}(\hat{p_x}\hat{x}) = (\hat{p_x}\hat{x} + i\hbar)\hat{p_x} - \hat{p_x}(\hat{x}\hat{p_x} - i\hbar) = 2i\hbar\hat{p_x}$$

It is then straighforward, using the same method, to prove by induction that for any $n \in \mathbb{N}$,

$$[\hat{x}, \hat{p_x^n}] = i\hbar n \widehat{p_x^{n-1}}$$

Therefore, the commutator of $\hat{x}$ and $\hat{H}$ is

$$[\hat{x}, \hat{H}] = i\hbar \widehat{\frac{\partial H}{\partial p_x}}$$

where $\widehat{\frac{\partial H}{\partial p_x}}$ is the operator corresponding to the partial derivative $\partial H / \partial p_x$ of the Hamiltonian.

In a similar way, we would show that

$$[\hat{p_x}, \hat{H}] = -i\hbar \widehat{\frac{\partial H}{\partial x}}$$

We finally obtain the desired equations, known as Ehrenfest's equations:

$$\begin{cases} \frac{d\overline{x}}{dt} = \overline{\widehat{\frac{\partial H}{\partial p_x}}} \\ \frac{d\overline{p_x}}{dt} = -\overline{\widehat{\frac{\partial H}{\partial x}}} \end{cases}$$

Now, the classical limit is understood as the limit in which the probability distributions have negligible spread compared with the accuracy of the measurement, and the measured values and the mean values can be seen as the same. In this limit, it is clear that Ehrenfest's equations agree with the canonical equations of Hamiltonian mechanics – Equation 10 in Lecture 6.