

Fundamentals and formalism of quantum mechanics

Asaf Pe'er¹

March 28, 2018

This part of the course is based on Refs. [1] – [4].

We saw how quantum mechanics deals with some simple problems. This is very different than classical physics. We now want to generalize the treatment, in order to enable us to solve more complicated problems. For that, we need to introduce two things: first, a more in depth discussion in some of the mathematical concepts of quantum mechanics. Second, some basic principles - or postulations about the behavior of physical systems. The verification of these postulations is experimental.

1. Operators (part 1)

Let us look again at the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r},t) \right] \psi(\vec{r},t) \quad (1)$$

The function $\psi(\vec{r})$ is a **wave function**: namely, for every point in space (and time), it obtains a (complex) value.

We further know that in order for ψ to be physically meaningful, it must fulfill the condition of being **square integrable** and normalized, namely

$$\int_V |\psi(\vec{r},t)|^2 d^3r = \int_V \psi^*(\vec{r},t)\psi(\vec{r},t) d^3r = 1. \quad (2)$$

Schrödinger equation contains terms such as $\frac{\partial\psi}{\partial t}$. Such a term takes the function ψ and turns it into another function of space (and time): namely, we can write

$$\tilde{\psi}(\vec{r},t) = \frac{\partial}{\partial t}\psi(\vec{r},t) \quad (3)$$

¹Physics Dep., University College Cork

Thus, by applying the derivative $\partial/\partial t$, we changed the function ψ into $\tilde{\psi}$. Generally speaking, we made an *operation* that mapped ψ into $\tilde{\psi}$. That is, $\tilde{\psi}$ is obtained through the action of an **operator** on the function ψ :

$$\tilde{\psi} = A\psi \quad (4)$$

It is typical to denote operators by capital letters. By definition, an operator A is **linear**, if the result of it acting on the function $\lambda_1\psi_1 + \lambda_2\psi_2$ is given by

$$A(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1(A\psi_1) + \lambda_2(A\psi_2) \quad (5)$$

where λ_1 and λ_2 are arbitrary complex numbers.

Examples of linear operators.

1. The differential operators: $\partial/\partial x$, $\partial/\partial t$, etc.
2. Operators of the form $f(\vec{r}, t)$ whose operation is to multiply the wavefunction ψ by the function $f(\vec{r}, t)$.
3. Sum of two linear operators, $S = A + B$.
4. Multiplication of any linear operator by constant, c : $(cA)\psi = c(A\psi)$.
5. Product of two operators, $P = AB$.

We must pay careful attention to the following point. There is a fundamental difference between “ordinary algebra” and the algebra of linear operators. **The product of two operators is not necessarily commutative.** The difference $AB - BA$ is called the **commutator**, and is denoted by $[A, B]$:

$$[A, B] \equiv AB - BA \quad (6)$$

If $[A, B] = 0$, we say that **A and B commute**.

Example. Consider the two operators $A = \partial/\partial x$ and $B = f(x)$ (the operator that multiplies ψ by $f(x)$). For any wave function ψ we have:

$$AB\psi \equiv \frac{\partial}{\partial x}(f(x)\psi) = \frac{\partial f(x)}{\partial x}\psi + f(x)\frac{\partial\psi}{\partial x} = \left(\frac{\partial f(x)}{\partial x} + f(x)\frac{\partial}{\partial x}\right)\psi \quad (7)$$

From which it follows that

$$\left[\frac{\partial}{\partial x}, f(x)\right] = \frac{\partial f(x)}{\partial x}. \quad (8)$$

In particular, if $f(x) = x$, we have

$$\left[\frac{\partial}{\partial x}, x\right] = 1.$$

1.1. The Hamiltonian operator

Let us look at the (time-independent) Schrödinger equation (Equation 1). As we will see during this course, the operator on the right hand side of this equation plays a very important role in quantum mechanics. It has a special name: **Hamiltonian operator**, denoted by H (sometimes: \hat{H}), honoring **William Hamilton**:

$$\hat{H} \equiv -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}, t) \quad (9)$$

1.2. Eigenvalues and eigenfunctions

Let us look again at the *time-independent* Schrödinger equation:

$$\hat{H}\phi(\vec{r}) \equiv \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) \right] \phi(\vec{r}) = E\phi(\vec{r}) \quad (10)$$

In the right hand side, we have a constant: E , multiplying a wavefunction, ϕ . On the left hand side, we have an operator, \hat{H} multiplying the wave function ϕ ; As we saw, this results in a different wavefunction,

$$\tilde{\phi} \equiv \hat{H}\phi.$$

If ϕ is a solution of the time-independent Schrödinger equation, we see that the new wave function, $\tilde{\phi}$ is equal to the original wavefunction ϕ - up to a constant, E . In such cases we say that ϕ is an **eigenfunction** of the operator \hat{H} , and E is its corresponding **eigenvalue**.

1.3. Hermitian operators (self-adjoint)

The interpretation of the wave function in terms of probability of finding the particle,

$$\int_{\text{all space}} P(\vec{r}, t) dV = \int_{\text{all space}} \psi^* \psi d^3r = 1 \quad (11)$$

led us to conclude that

$$\frac{\partial}{\partial t} \int_{\text{all space}} P(\vec{r}, t) dV = \int_{\text{all space}} \left[\psi^* \left(\frac{\partial \psi}{\partial t} \right) + \left(\frac{\partial \psi^*}{\partial t} \right) \psi \right] d^3r = 0. \quad (12)$$

(see “waves”, equations 58, 61).

In order to proceed, we can write the (time-dependent) Schrödinger Equation (1) in the form

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \hat{H} \psi(\vec{r}, t), \quad (13)$$

and its complex conjugate

$$-i\hbar \frac{\partial}{\partial t} \psi^*(\vec{r}, t) = \left[\hat{H} \psi(\vec{r}, t) \right]^*. \quad (14)$$

Using these equations in Equation 12, we find

$$\begin{aligned} \frac{\partial}{\partial t} \int_{all \ space} P(\vec{r}, t) dV &= \int_{all \ space} \left[\psi^* \left(\frac{\partial \psi}{\partial t} \right) + \left(\frac{\partial \psi^*}{\partial t} \right) \psi \right] d^3r \\ &= \frac{1}{i\hbar} \int_{all \ space} \left[\psi^* \left(\hat{H} \psi \right) - \left(\hat{H} \psi \right)^* \psi \right] d^3r = 0. \end{aligned} \quad (15)$$

We therefore find that the condition that the total probability of finding the particle somewhere is independent on time can be written as a restriction on the Hamiltonian operator:

$$\int \psi^* \left(\hat{H} \psi \right) d^3r = \int \left(\hat{H} \psi \right)^* \psi d^3r. \quad (16)$$

Operators which fulfill the condition set in Equation 16 for all functions ψ which they act on, are called **Hermitian operators**. As we will show below, Hermitian operators play a particularly important role in Quantum Mechanics.

In particular, we just showed that the Hamiltonian operator that appears in Schrödinger equation is Hermitian.

Hermitian operators, their eigenfunctions and eigenvectors have some interesting mathematical properties.

1.3.1. Eigenvalues of a Hermitian operators are real

Let us assume that A is a Hermitian operator. Let ψ be an eigenfunction of A , with an associated eigenvalue λ . We have:

$$A\psi = \lambda\psi. \quad (17)$$

Taking the the complex conjugate,

$$(A\psi)^* = \lambda^* \psi^* \quad (18)$$

Multiplying Equation 17 by ψ^* and integrating over all space, we find

$$\int \psi^*(A\psi) d^3r = \int \lambda \psi^* \psi d^3r = \lambda \int \psi^* \psi d^3r. \quad (19)$$

Similarly, multiplying Equation 18 by ψ and integrating over all space, we have

$$\int (A\psi)^*\psi d^3r = \int \lambda^*\psi^*\psi d^3r = \lambda^* \int \psi^*\psi d^3r. \quad (20)$$

We assumed that A is Hermitian, therefore it satisfies Equation 16. Thus, Equations 19 and 20 are equal, which leads us to conclude that

$$\lambda \int \psi^*\psi d^3r = \lambda^* \int \psi^*\psi d^3r \quad (21)$$

from which we find that $\lambda = \lambda^*$. Thus, the eigenvalue λ must be a real number.

1.4. Orthogonality of wavefunctions of a Hermitian operator

Let us assume that $\psi_i(\vec{r}, t)$ and $\psi_j(\vec{r}, t)$ are two wavefunctions that are both eigenfunctions of a Hermitian operator A , but with **different eigenvalues**, $\lambda_i \neq \lambda_j$.

Each wavefunction is normalized:

$$\int_{all\ space} \psi_i^*\psi_i d^3r = \int_{all\ space} \psi_j^*\psi_j d^3r = 1. \quad (22)$$

We now show that these functions are **orthogonal**, namely

$$\int_{all\ space} \psi_i^*\psi_j d^3r = 0 \quad \text{for } i \neq j. \quad (23)$$

Proof. We take

$$A\psi_j = \lambda_j\psi_j,$$

multiply by ψ_i^* and integrate over all space,

$$\int \psi_i^*(A\psi_j) d^3r = \int \psi_i^*\lambda_j\psi_j d^3r = \lambda_j \int \psi_i^*\psi_j d^3r \quad (24)$$

(since λ_j is just a constant, we can take it outside of the integration).

Similarly, we have

$$(A\psi_i)^* = \lambda_i^*\psi_i^* = \lambda_i\psi_i^*$$

(since λ_i is real), from which it follows that

$$\int (A\psi_i)^* \psi_j d^3r = \lambda_i \int \psi_i^*\psi_j d^3r \quad (25)$$

We now subtract equation 24 from Equation 25, to get

$$(\lambda_i - \lambda_j) \int \psi_i^* \psi_j d^3r = \int [(A\psi_i)^* \psi_j - \psi_i^* (A\psi_j)] d^3r = 0 \quad (26)$$

This is equal to zero using the definition of a Hermitian operator (Equation 16), and our assumption that A is Hermitian. Since, by assumption $\lambda_i \neq \lambda_j$, we are left with

$$\int \psi_i^* \psi_j d^3r = 0.$$

In particular, since the Hamiltonian, \hat{H} is a Hermitian operator, two wavefunctions that are eigenfunctions of the Hamiltonian with different eigenvalues $E_i \neq E_j$ are orthogonal to each other.

Example. We showed (“waves”, §8.2) that the wavefunctions representing a particle in an infinite square potential well of length L are $\phi_i(x) = \sqrt{(2/L)} \sin(i\pi x/L)$. These are real functions, that fulfill

$$\int_0^L \phi_i^* \phi_j dx = \frac{2}{L} \int_0^L \sin\left(\frac{i\pi x}{L}\right) \sin\left(\frac{j\pi x}{L}\right) = \frac{\sin\left(\frac{(i-j)\pi x}{L}\right)}{(i-j)\pi x} - \frac{\sin\left(\frac{(i+i)\pi x}{L}\right)}{(i+j)\pi x} \Bigg|_0^L = 0$$

(unless $i = j$).

2. Expectation values and operators

I have stressed that in quantum mechanics, the **wave function** plays a fundamental role: it contains all possible information on a particle. It replaces the classical concept of *trajectory*, as a particle does not have a well defined trajectory. The wave function contains all possible information on a particle. **We interpret it in a statistical way:** the square of a wave function tells us what is the probability of finding the particle at a given position *if we conduct a measurement*,

$$dP(\vec{r}, t) = \psi^*(\vec{r}, t)\psi(\vec{r}, t)d^3r. \quad (27)$$

Operators are equally important in quantum mechanics. This is due to the fact that **operators represent quantities that can be measured**, such as energy, position, momentum, etc. In fact,

Every measurable physical quantity is represented by a Hermitian operator.

Before getting deeper into the explanation of this, let us consider a particle that is described by a wavefunction ψ . Let us show how one can use ψ to extract information about the particle.

Consider an arbitrary function $F = F(\vec{r})$ of the coordinates of the particle. We can define the mean value of F as

$$\langle F(\vec{r}) \rangle = \int P(\vec{r}, t) F(\vec{r}) d\vec{r} = \int \psi^*(\vec{r}, t) F(\vec{r}) \psi(\vec{r}, t) d\vec{r} \quad (28)$$

The mean value $\langle F \rangle$ is also called the **expectation value of F** .

It has the following physical meaning: $\langle F(\vec{r}) \rangle$ is the average value that we obtain if we measure $F(\vec{r})$ many ($N \gg 1$) times, in N different systems that are all identical and independent on each other, and are all represented by the same wavefunction ψ .

As a concrete example, if we set $F(\vec{r}) = \vec{r}$ (the position of the particle), then we find

$$\langle \vec{r} \rangle = \int P(\vec{r}, t) \vec{r} d\vec{r} = \int \psi^*(\vec{r}, t) \vec{r} \psi(\vec{r}, t) d\vec{r} \quad (29)$$

Equation 29 is equivalent to the three equations

$$\begin{aligned} \langle x \rangle &= \int \psi^*(\vec{r}, t) x \psi(\vec{r}, t) d\vec{r}, \\ \langle y \rangle &= \int \psi^*(\vec{r}, t) y \psi(\vec{r}, t) d\vec{r}, \\ \langle z \rangle &= \int \psi^*(\vec{r}, t) z \psi(\vec{r}, t) d\vec{r}. \end{aligned} \quad (30)$$

We can think of the result in Equation 28 in the following way. Inside the integral (the right hand side of the equation), $F(\vec{r})$ can be thought of as an **operator** that acts on the wavefunction $\psi(\vec{r}, t)$. Outside the integral (the left hand side), $\langle F(\vec{r}) \rangle$ is the **expectation value** of the physical quantity F (which, for example, could be the position, momentum, energy,...): this is the average value we obtain when measuring F multiple times.

With this interpretation, we can state the results in Equation 29, 30 as follows: **the operator \vec{r} is associated with the position vector \vec{r}** .

There is one important restriction on operators that are associated with physical quantities. When we measure a physical quantity (such as position, momentum, etc.) obviously the result is always some **real** quantity. This means that if F is an operator associated with a measurable physical quantity, then it must satisfy

$$\int \psi^*(\vec{r}, t) [F(\vec{r})\psi(\vec{r}, t)] d\vec{r} = \int [F(\vec{r})\psi(\vec{r}, t)]^* \psi(\vec{r}, t) d\vec{r} \quad (31)$$

In other words, **the operator $F(\vec{r})$ must be Hermitian**. Such operators are also called **observables**.

2.1. Functions of momentum

Often, we encounter functions not of the position of a particle, but of its momentum - such as the momentum itself. We would like to calculate the expectation value of such a function, $G = G(\vec{p})$.

In order to do this, recall that we can use the superposition principle, to write every wave function as a combination of plane waves (see “waves”, equations 47, 48; here we generalize to 3-d):

$$\psi(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \phi(\vec{k}) e^{i(\vec{k}\cdot\vec{r} - \omega(|k|)t)} d^3k = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \phi(\vec{p}, t = 0) e^{i(\vec{p}\cdot\vec{r} - E(|p|)t)/\hbar} d^3p, \quad (32)$$

where in the last equality we used de-Broglie relations, $p = \hbar k$ and $E = \hbar\omega$.

The function $\phi(\vec{p}, t = 0)$ is the **Fourier transform** of $\psi(\vec{r}, t = 0)$:

$$\phi(\vec{p}, t = 0) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{r}, t = 0) e^{-i\vec{p}\cdot\vec{r}/\hbar} d^3r. \quad (33)$$

This can be generalized to any time, t : we can always write

$$\psi(\vec{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \phi(\vec{p}, t) e^{i(\vec{p}\cdot\vec{r})/\hbar} d^3p, \quad (34)$$

and

$$\phi(\vec{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{r}, t) e^{-i\vec{p}\cdot\vec{r}/\hbar} d^3r. \quad (35)$$

The function $\phi(\vec{p}, t)$ is the wave function of the particle in momentum space. Thus, in analogy to Equation 28, if we have a function $G = G(\vec{p})$, we can write its expectation value as

$$\langle G(\vec{p}) \rangle = \int \phi^*(\vec{p}, t) G(\vec{p}) \phi(\vec{p}, t) d\vec{p}. \quad (36)$$

For example, the expectation value of the kinetic energy is

$$\left\langle \frac{|\vec{p}|^2}{2m} \right\rangle = \int \phi^*(\vec{p}, t) \frac{\vec{p}^2}{2m} \phi(\vec{p}, t) d\vec{p}. \quad (37)$$

We now write Equation 36 in a different form. We start by writing

$$\phi^*(\vec{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \psi^*(\vec{r}', t) e^{+i\vec{p}\cdot\vec{r}'/\hbar} d^3r'. \quad (38)$$

Inserting into Equation 36, we get

$$\langle p_x \rangle = \frac{1}{(2\pi\hbar)^3} \int d\vec{p} \int d\vec{r} \int d\vec{r}' e^{+i\vec{p}\cdot\vec{r}'/\hbar} \psi^*(\vec{r}', t) p_x e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) \quad (39)$$

Now, look at:

$$p_x e^{-i\vec{p}\cdot\vec{r}/\hbar} = p_x e^{-i(p_x x + p_y y + p_z z)/\hbar} = i\hbar \frac{\partial}{\partial x} e^{-i\vec{p}\cdot\vec{r}/\hbar}. \quad (40)$$

Equation 39 thus becomes:

$$\langle p_x \rangle = \frac{1}{(2\pi\hbar)^3} \int d\vec{p} \int d\vec{r} \int d\vec{r}' e^{+i\vec{p}\cdot\vec{r}'/\hbar} \psi^*(\vec{r}', t) \left(i\hbar \frac{\partial}{\partial x} e^{-i\vec{p}\cdot\vec{r}/\hbar} \right) \psi(\vec{r}, t) \quad (41)$$

We now integrate by parts with respect to x . Recall that at $x \rightarrow \infty$, $\psi(\vec{r}, t) = 0$, since $\psi(\vec{r}, t)$ must be normalized. We thus have:

$$\langle p_x \rangle = \frac{1}{(2\pi\hbar)^3} \int d\vec{p} \int d\vec{r} \int d\vec{r}' \psi^*(\vec{r}', t) e^{+i\vec{p}\cdot(\vec{r}'-\vec{r})/\hbar} \left[-i\hbar \frac{\partial}{\partial x} \psi(\vec{r}, t) \right] \quad (42)$$

In order to proceed, we use a property of the δ -function,

$$\delta(\vec{r}' - \vec{r}) = \frac{1}{(2\pi\hbar)^3} \int d\vec{p} e^{+i\vec{p}\cdot(\vec{r}'-\vec{r})/\hbar}, \quad (43)$$

to perform the integration over $d\vec{p}$:

$$\langle p_x \rangle = \int d\vec{r} \int d\vec{r}' \psi^*(\vec{r}', t) \delta(\vec{r}' - \vec{r}) \left[-i\hbar \frac{\partial}{\partial x} \psi(\vec{r}, t) \right] = \int \psi^*(\vec{r}, t) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(\vec{r}, t) d\vec{r}. \quad (44)$$

In a similar way,

$$\begin{aligned} \langle p_y \rangle &= \int \psi^*(\vec{r}, t) \left(-i\hbar \frac{\partial}{\partial y} \right) \psi(\vec{r}, t) d\vec{r}, \\ \langle p_z \rangle &= \int \psi^*(\vec{r}, t) \left(-i\hbar \frac{\partial}{\partial z} \right) \psi(\vec{r}, t) d\vec{r}. \end{aligned} \quad (45)$$

This result can be obviously generalized to

$$\langle \vec{p} \rangle = \int \psi^*(\vec{r}, t) \left(-i\hbar \vec{\nabla} \right) \psi(\vec{r}, t) d\vec{r}, \quad (46)$$

and to any function of the momentum, $G = G(\vec{p}, t)$:

$$\langle G(\vec{p}, t) \rangle = \int \psi^*(\vec{r}, t) G \left(-i\hbar \vec{\nabla}, t \right) \psi(\vec{r}, t) d\vec{r}. \quad (47)$$

This result implies that **the operator associated with the momentum, \vec{p} is $-i\hbar \vec{\nabla}$.**

Example. Suppose that we want to calculate the expectation value of the total energy of a particle that moves in a potential $V(r, t)$. In Newtonian physics, we have $E = |p|^2/2m + V(r, t)$. Using Equation 46, we derive the quantum mechanical analogue:

$$\langle E \rangle = \int \psi^*(\vec{r}, t) \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}, t) \right] \psi(\vec{r}, t) d\vec{r} = \langle \hat{H} \rangle \quad (48)$$

This leads to an important conclusion: **the operator associated with the total energy E of a particle is the Hamiltonian, \hat{H} .**

2.2. Functions of both position and momentum

We saw that the expectation values of function of the position: $F = F(\vec{r}, t)$ or the momentum: $G = G(\vec{p}, t)$ are always real. Hence, the operators associated with these functions are Hermitian.

However, if we have a function of *both* \vec{r} and \vec{p} , this is not necessarily the case. For example, consider a particle moving in 1-d, characterized by wavefunction $\psi(x, t)$. What is the expectation value of xp_x ? We have:

$$\langle xp_x \rangle = \int_{-\infty}^{+\infty} \psi^*(x, t) x \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x, t) dx \quad (49)$$

We can integrate by parts:

$$\begin{aligned} \langle xp_x \rangle &= -i\hbar x \psi^* \psi \Big|_{-\infty}^{+\infty} + i\hbar \int_{-\infty}^{+\infty} \psi(x, t) \frac{\partial}{\partial x} (x \psi^*(x, t)) dx \\ &= 0 + i\hbar \int_{-\infty}^{+\infty} \psi x \frac{\partial \psi^*}{\partial x} dx + i\hbar \int_{-\infty}^{+\infty} \psi \psi^* dx \end{aligned} \quad (50)$$

(the first term is 0, since $|\psi| \rightarrow 0$ when $x \rightarrow \pm\infty$).

We thus find:

$$\langle xp_x \rangle = \langle xp_x \rangle^* + i\hbar. \quad (51)$$

The existence of the second term, $i\hbar$ in Equation 51, forces us to conclude that the expectation value of $\langle xp_x \rangle$ cannot be real. This means that the associated operator, $x \left(-i\hbar \frac{\partial}{\partial x} \right)$ is not Hermitian.

Mathematically, we just showed that the operators representing the physical quantities of position (x) and momentum (p_x) do not commute. Physically, this means that **we cannot measure simultaneously both the position and momentum of a particle to infinite precision.**

3. Ehrenfest theorem: transition between quantum mechanics and classical mechanics

Obviously, the above result has no classical analogue. However, we know that on a large scale, the laws of classical mechanics do hold. Thus, there must exist conditions under which classical mechanics provides good description of a physical system. The general rule was formulated by Niels Bohr in 1923, and is known as the **correspondence principle**. A common formulation of the principle states that **the results of a quantum theory must approach those of classical theory asymptotically in the limit of large quantum numbers**.

In order to investigate this point quantitatively and gain an insight into the nature of quantum mechanics, we will prove the following theorem, known as **Ehrenfest's theorem**. Consider Newton's laws, written in the following form, known as **Hamilton-Jacobi equations** (yes, the same Hamilton from the Hamiltonian):

$$\begin{aligned}\frac{d\vec{r}}{dt} &= \frac{\vec{p}}{m}, \\ \frac{d\vec{p}}{dt} &= -\vec{\nabla}V.\end{aligned}\tag{52}$$

Clearly, these are classical equations of motion.

Ehrenfest's theorem states that **the same equations are satisfied by the expectation values (=average values) of the corresponding quantum mechanics operators**. Note an important point. In quantum mechanics, there are no well-define trajectories, as in classical mechanics, so in this sense terms such as $d\vec{r}/dt$ which gives the change in direction along a classical path does not exist. However, the expectation values of operators such as \vec{r} do change with time, and so it does make sense to look at their temporal evolution.

In the mathematical proof, we will use Green's identities from vector calculus. Green's first identity states that if u and v are two scalar functions of the position ($u = u(x, y, z)$, $v = v(x, y, z)$), both have continuous first and second derivatives, then

$$\int_V [u\nabla^2v + (\nabla u) \cdot (\nabla v)] d^3r = \int_S u(\nabla v) \cdot dS,\tag{53}$$

where S is the surface enclosing the volume V . Green's second identity states that for these functions,

$$\int_V [u(\nabla^2v) - v(\nabla^2u)] d^3r = \int_S [u(\nabla v) - v(\nabla u)] \cdot dS.\tag{54}$$

(These theorems can be viewed as special cases of the divergence theorem, $\int_V (\vec{\nabla} \cdot \vec{F})dV = \oint \vec{F} \cdot d\vec{S}$).

proof. Using Equation 30, we have

$$\begin{aligned} \frac{d}{dt}\langle x \rangle &= \frac{d}{dt} \int \psi^*(\vec{r}, t) x \psi(\vec{r}, t) d\vec{r} \\ &= \int \psi^*(\vec{r}, t) x \frac{d\psi(\vec{r}, t)}{dt} d\vec{r} + \int \frac{d\psi^*(\vec{r}, t)}{dt} x \psi(\vec{r}, t) d\vec{r} \end{aligned} \quad (55)$$

Using Schrödinger equation and its complex conjugate (Equations 13, 14), we have

$$\begin{aligned} \frac{d}{dt}\langle x \rangle &= \frac{1}{i\hbar} \left[\int \psi^* x (\hat{H}\psi) d\vec{r} - \int (\hat{H}\psi)^* x \psi d\vec{r} \right] \\ &= \frac{i\hbar}{2m} \int [\psi^* x (\nabla^2 \psi) - (\nabla^2 \psi^*) x \psi] d\vec{r}. \end{aligned} \quad (56)$$

(The terms involve the potential V cancel out).

We now use Green's first identity, setting $u = x\psi$ and $v = \psi^*$ to write the second integral:

$$\int (\nabla^2 \psi^*) x \psi d\vec{r} = \int_S x \psi (\nabla \psi^*) \cdot dS - \int (\nabla \psi^*) \cdot (\nabla x \psi) d\vec{r} = - \int (\nabla \psi^*) \cdot (\nabla x \psi) d\vec{r} \quad (57)$$

(the first integral over dS must be equal to zero, since the surface encloses all space, and $\psi \rightarrow 0$ at $|\vec{r}| \rightarrow \infty$).

We use again Green's first identity, to write

$$- \int (\nabla \psi^*) \cdot (\nabla x \psi) d\vec{r} = - \int_S \psi^* \nabla(x\psi) \cdot dS + \int \psi^* \nabla^2(x\psi) d\vec{r} = \int \psi^* \nabla^2(x\psi) d\vec{r} \quad (58)$$

(from the same reasoning).

Putting this in Equation 56, we find

$$\frac{d}{dt}\langle x \rangle = \frac{i\hbar}{2m} \int \psi^* [x(\nabla^2 \psi) - \nabla^2(x\psi)] d\vec{r} = -\frac{i\hbar}{m} \int \psi^* \frac{\partial \psi}{\partial x} d\vec{r} \quad (59)$$

where the last equality comes from expansion of the second term.

Comparing this result to Equation 44, one finds

$$\frac{d}{dt}\langle x \rangle = \frac{\langle p_x \rangle}{m}, \quad (60)$$

which is the quantum mechanical counterpart of the first of the classical Hamilton-Jacobi Equations (Equation 52).

The second part of Ehrenfest's theorem is proven in a similar way. From the expectation value of the momentum (Equation 44) we have

$$\frac{d}{dt}\langle p_x \rangle = -i\hbar \frac{d}{dt} \int \psi^* \frac{\partial \psi}{\partial x} d\vec{r} = -i\hbar \left[\int \psi^* \frac{\partial}{\partial x} \frac{\partial \psi}{\partial t} d\vec{r} + \int \frac{\partial \psi^*}{\partial t} \frac{\partial \psi}{\partial x} d\vec{r} \right] \quad (61)$$

Using Schrödinger equation and its complex conjugate to re-write $\partial\psi/\partial t$ and $\partial\psi^*/\partial t$, one obtains

$$\begin{aligned} \frac{d}{dt}\langle p_x \rangle &= -\int \psi^* \frac{\partial}{\partial x} \left(-\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi \right) d\vec{r} + \int \left(-\frac{\hbar^2}{2m} \nabla^2 \psi^* + V\psi^* \right) \frac{\partial \psi}{\partial x} d\vec{r} \\ &= \frac{\hbar^2}{2m} \int [\psi^* (\nabla^2 \frac{\partial \psi}{\partial x}) - (\nabla^2 \psi^*) \frac{\partial \psi}{\partial x}] d\vec{r} - \int \psi^* [\frac{\partial}{\partial x}(V\psi) - V \frac{\partial \psi}{\partial x}] d\vec{r} \end{aligned} \quad (62)$$

Assuming that ψ and $\partial\psi/\partial x$ vanish at large x , the first integral vanishes by Green's second identity (using $u = \psi^*$ and $v = \partial\psi/\partial x$). The second integral is

$$-\int \psi^* \left[\frac{\partial}{\partial x}(V\psi) - V \frac{\partial \psi}{\partial x} \right] d\vec{r} = -\int \psi^* \frac{\partial V}{\partial x} \psi d\vec{r} = -\left\langle \frac{\partial V}{\partial x} \right\rangle. \quad (63)$$

We thus found that

$$\frac{d}{dt}\langle p_x \rangle = -\left\langle \frac{\partial V}{\partial x} \right\rangle \quad (64)$$

which completes the proof.

Physical meaning. Let us assume that the wave function ψ that describes the particle takes the form of a wavepacket. We can therefore consider the average value (=expectation value) $\langle \vec{r} \rangle(t)$ to be the center of the wavepacket at time t . Thus, the set of all points $\langle \vec{r} \rangle(t)$ corresponding to different times form a trajectory: this is the **trajectory followed by the center of the wave packet**.

The conclusion is, that although we cannot speak of a trajectory of the particle itself (since ψ extends over all space), we can speak of a trajectory of the center of the wave packet. If all the distances involved in the problem are much bigger than the extension of the wave packet, we can approximate the wave packet by its center, and recover the classical description of trajectory, and the classical results.

4. Dirac's bracket notation

We saw that in quantum mechanics, the quantum state of a particle is fully specified by a wave function, $\psi(\vec{r}, t)$. On the other hand, not every wave function can describe a physical particle: a wave function must be square integrable and normalized, in order for it to describe a physical particle:

$$\int |\psi(\vec{r}, t)|^2 d^3r = 1. \quad (65)$$

From a physical point of view, we add a few requirements: these functions must be everywhere defined, continuous, and infinitely differentiable.

Technically, the set of functions ψ for which the integral converges, form what is called by mathematicians L^2 space. This space has all the properties of a vector space, which we

will not go into in full details here. However, I do point into two important aspects of it: (i) the **superposition principle**, according to which if ψ_1 and ψ_2 belong to this space, so will

$$\psi(\vec{r}) = \lambda_1\psi_1(\vec{r}) + \lambda_2\psi_2(\vec{r}),$$

where λ_1 and λ_2 are two arbitrary complex numbers; and (ii) the ability to define a **scalar product** between two functions, $\phi(\vec{r})$ and $\psi(\vec{r}, t)$ using

$$\int \phi^*(\vec{r}, t)\psi(\vec{r}, t) d^3r.$$

For those who are familiar with the mathematical definitions, we say that L^2 is an example of an infinite-dimensions **Hilbert space**.

This association of wave functions as elements in a vector space, enable us to use a convenient notation, according to **Paul Dirac**. To each wave function, $\psi(\vec{r})$, we associate a *ket vector*, or simply **ket**, which we denote by $|\psi\rangle$. Thus:

$$\boxed{\psi(\vec{r}) \iff |\psi\rangle} \tag{66}$$

The set of ket vectors, also known as **state vectors** of a quantum system, define a space. It is known as the “state space” of the system. (Technically, this space is isomorphic to the space defined by the functions $\psi(\vec{r})$, but we do not need to get into the mathematical details of this). Note that we generalize $|\psi\rangle$ to describe the state of a quantum system, which may include more than a single particle.

We omit the \vec{r} dependence from $|\psi\rangle$. In fact (we will not prove it here), we could think of $\psi(\vec{r})$ as the components of the ket $|\psi\rangle$ in a particular basis, \vec{r} . However, this goes somewhat deeper than currently needed.

One point which is important to remember, is that the use of Dirac’s notation enables not only to simplify the calculations, but also to generalize them. There exist physical systems whose quantum description cannot be fully given by a wave function. A classical example is the spin degree of freedom (which we may or may not have time to cover this year). Nonetheless, such systems can still be described by specifying their state vector.

4.1. Bras and Scalar product

To each ket vector, $|\psi\rangle$, there is an associated *bra vector*, or simply **bra**, denoted by $\langle\psi|$. Technically, the bra can be thought of a **linear functional**. A linear functional is a linear operation that associates with every ket $|\psi\rangle$ a complex number. This is done in a linear way.

Note the difference between a linear functional that associates with every ket a number, and a linear operator discussed above, that associates with every ket another ket. For those who are familiar with the term, we say that $\langle\psi|$ belongs to the *dual space* of the vector space defined by the kets.

With each pair of bra $\langle\phi|$ and ket $|\psi\rangle$, **taken in this order**, we can therefore associate a complex number - which is their **scalar product**, defined by

$$\langle\phi|\psi\rangle \equiv \int \phi^*(\vec{r})\psi(\vec{r}) d^3r. \quad (67)$$

It is obvious from this definition that

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*. \quad (68)$$

Furthermore, if λ is a complex number, we obviously have

$$\begin{aligned} \langle\phi|\lambda\psi\rangle &= \lambda\langle\phi|\psi\rangle \\ \langle\lambda\phi|\psi\rangle &= \lambda^*\langle\phi|\psi\rangle \\ \langle\phi|\psi_1 + \psi_2\rangle &= \langle\phi|\psi_1\rangle + \langle\phi|\psi_2\rangle \end{aligned} \quad (69)$$

Using this Dirac notation, the normalization condition of the wavefunction (Equation 65) can be written in a compact form,

$$\langle\psi|\psi\rangle = 1. \quad (70)$$

Furthermore, the functions ψ and ϕ are said to be orthogonal, if their scalar product vanishes:

$$\langle\phi|\psi\rangle = 0. \quad (71)$$

5. Linear operators (part II)

We saw that a linear operator A associates with every wavefunction ψ another wavefunction $\bar{\psi} = A\psi$, in a linear way, namely $A(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1(A\psi_1) + \lambda_2(A\psi_2)$ (see Equations 4, 5). Using Dirac's notation, we write

$$\begin{aligned} |\bar{\psi}\rangle &= A|\psi\rangle, \\ A(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle) &= \lambda_1A|\psi_1\rangle + \lambda_2A|\psi_2\rangle. \end{aligned} \quad (72)$$

Using Equations 67 and 72, we know that $\langle\phi|\bar{\psi}\rangle = \langle\phi|(A|\psi\rangle)$ is a complex number. Since both the scalar product and the operator A are linear, this complex number depends linearly

on $|\psi\rangle$. We can therefore think of both $\langle\phi|$ and A as defining a new bra, that associates with each ket $|\psi\rangle$ a complex number; in other words, we can write

$$(\langle\phi|A)|\psi\rangle = \langle\phi|(A|\psi\rangle) \equiv \langle\phi|A|\psi\rangle \quad (73)$$

where we omitted the parenthesis, as their place is unimportant. This is known as the **matrix element** between the kets $|\phi\rangle$ and $|\psi\rangle$ (obviously, $|\phi\rangle$ is the ket associated with the bra $\langle\phi|$).

Example: the projection operator.

Let $|\psi\rangle$ be a ket which is normalized to one:

$$\langle\psi|\psi\rangle = 1.$$

Consider the operator P_ψ defined by:

$$P_\psi = |\psi\rangle\langle\psi| \quad (74)$$

apply it to arbitrary ket, $|\phi\rangle$:

$$P_\psi|\phi\rangle = |\psi\rangle\langle\psi|\phi\rangle \quad (75)$$

We see that we obtain a ket which is proportional to $|\psi\rangle$; the coefficient of proportionality, $\langle\psi|\phi\rangle$ is a complex number, which is the scalar product of $\langle\psi|$ and $|\phi\rangle$. This has an obvious “geometrical” interpretation.

Furthermore, we have

$$P_\psi^2 = P_\psi P_\psi = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P_\psi, \quad (76)$$

where we used the fact that $\langle\psi|\psi\rangle = 1$.

5.1. Adjoint operators

The association between kets and bras enable to associate with each linear operator A , another linear operator, called **the adjoint operator**, that is denoted by A^\dagger . Another name for adjoint operators is **Hermitian conjugates** (not to be confused with Hermitian operators!).

Let $|\psi\rangle$ be an arbitrary ket. The operator A associates with it another ket, $|\bar{\psi}\rangle$. Now, let $\langle\psi|$ be the bra associated with $|\psi\rangle$, and $\langle\bar{\psi}|$ the bra associated with the ket $|\bar{\psi}\rangle$. The bra-ket correspondence, enable us to define the operator A^\dagger **that acts on the bras**:

$$\langle\bar{\psi}| = \langle\psi|A^\dagger \iff A|\psi\rangle = |\bar{\psi}\rangle. \quad (77)$$

Using the property of the scalar product (equation 68), $\langle \bar{\psi} | \phi \rangle = \langle \phi | \bar{\psi} \rangle^*$, we can write

$$\langle \psi | A^\dagger | \phi \rangle = \langle \phi | A | \psi \rangle^* \quad (78)$$

(note that both sides of Equation 78 are c-numbers !). This is valid for all $\langle \phi |$ and $| \psi \rangle$. The operator A^\dagger is known as the **Hermitian conjugate** of the operator A .

The concept of adjoint operators can be understood as operator generalization of the the complex conjugate of a complex number.

It is straight forward to show that

$$\begin{aligned} (A^\dagger)^\dagger &= A \\ (\lambda A)^\dagger &= \lambda^* A^\dagger \\ (A + B)^\dagger &= A^\dagger + B^\dagger \end{aligned} \quad (79)$$

By applying two operators once at a time, one can also show that

$$(AB)^\dagger = B^\dagger A^\dagger \quad (80)$$

5.2. Hermitian operators (again)

If a linear operator A satisfies

$$A^\dagger = A \quad (81)$$

then it is called **self adjoint**. In this case, Equation 78 becomes

$$\langle \psi | A | \phi \rangle = \langle \phi | A | \psi \rangle^* \quad (82)$$

Comparing this to Equation 16, we see that this is similar to the definition of a Hermitian operator (identical if $|\phi\rangle = |\psi\rangle$). In fact, it is not difficult to show that the two definitions are equivalent. Thus, self-adjoint is another name to Hermitian operator.

Alternatively, we could define Hermitian operator A by

$$\langle \phi | (A\psi) \rangle = \langle (A\phi) | \psi \rangle \quad (83)$$

where we used the notation $|(A\psi)\rangle \equiv A|\psi\rangle$ and $\langle (A\psi)| \equiv \langle \psi|A^\dagger$.

We can now repeat the discussion above about the properties of the eigenfunctions and eigenvalues of Hermitian operators, using Dirac's notation. For example, if ψ_n is an eigenfunction of the operator A with eigenvalue a_n , then from $A\psi_n = a_n\psi_n$ we have

$$\langle \psi_n | A | \psi_n \rangle = a_n \langle \psi_n | \psi_n \rangle \quad (84)$$

and since $(A\psi_n)^* = a_n^* \psi_n^*$, we have

$$\langle (A\psi_n) | \psi_n \rangle = a_n^* \langle \psi_n | \psi_n \rangle. \quad (85)$$

If A is Hermitian, the left hand sides of Equations 84, 85 are equal, hence $a_n = a_n^*$, namely the eigenvalues of A are real.

Similarly, we can prove the orthogonality properties of the eigenfunctions of A . If ψ_i and ψ_j are eigenfunctions of A with two different eigenvalues, a_i and a_j , we have $A\psi_i = a_i\psi_i$ and $A\psi_j = a_j\psi_j$. Thus,

$$(a_i - a_j) \langle \psi_i | \psi_j \rangle = \langle a_i \psi_i | \psi_j \rangle - \langle \psi_i | a_j \psi_j \rangle = \langle (A\psi_i) | \psi_j \rangle - \langle \psi_i | (A\psi_j) \rangle = 0 \quad (86)$$

where we have used the fact that A is Hermitian. Since $a_i \neq a_j$, it follows immediately that $\langle \psi_i | \psi_j \rangle = 0$ ($i \neq j$).

Furthermore, if ψ_n is an eigenfunction of the operator A , its corresponding ket, $|\psi_n\rangle$ is called **eigenvector** of A .

5.3. Physical interpretation: observables

A basic postulate of quantum mechanics is that

Every measurable physical quantity (position, energy, momentum, etc.) is described by a Hermitian operator, A . This operator is an observable.

Now, if the wavefunction of a system happens to be an eigenfunction ψ_n of the operator A (corresponding eigenvalue a_n), then (second postulate): **a measurement of the dynamical variable A will certainly produce the result a_n** . We say that **the system is in an eigenstate of A characterized by eigenvalue a_n** . The set of all eigenvalues of A are called the **spectrum** of A .

If, on the other hand, the wave function ψ describing the system is not an eigenfunction of A , then, when measuring A , any one of the results a_1, a_2, \dots can be obtained. We don't know in advanced which result will be obtained; but we can calculate the probability of obtaining a certain result (we will show shortly how).

If the system is described by wavefunction ψ that is normalized, namely $\langle \psi | \psi \rangle = 1$,

then in analogy to the discussion above, we define the **expectation value** of A as

$$\langle A \rangle \equiv \langle \psi | A | \psi \rangle.^1 \quad (87)$$

5.4. Functions of operators

Consider an arbitrary operator A . One can easily define the operator A^n : this is the operator which corresponds to n successive applications of the operator A : $A^n|\psi\rangle = A \cdot A \cdot \dots \cdot A|\psi\rangle$. But what about a general function of operator, $f(A)$?

In analogy to power series expansion of functions of a variable, z : $f(z) = \sum_{i=0}^{\infty} c_i z^i$, we can define the operator function $f(A)$ as

$$f(A) \equiv \sum_{i=0}^{\infty} c_i A^i \quad (88)$$

Consequently, if $|\psi_n\rangle$ is one of the eigenfunctions of A corresponding to the eigenvalue a_n , we have: $A^i|\psi_n\rangle = (a_n)^i|\psi_n\rangle$. Thus,

$$f(A)|\psi_n\rangle = f(a_n)|\psi_n\rangle. \quad (89)$$

In words: we just proved that if $|\psi_n\rangle$ is an eigenvector of A with the eigenvalue a_n , then $|\psi_n\rangle$ is also an eigenvector of $f(A)$ with the eigenvalue $f(a_n)$.

This leads to another insight. As we will immediately show, operators are best represented by matrices, once we specify a set of basis vectors. If we find a basis such that the matrix representing an operator A is diagonal (namely, zero everywhere except from the diagonal) - which is something that we can always do for an observable - then the matrix representing $f(A)$, in the same basis, is also diagonal, whose elements are $f(a)$.

Example. The operator e^A is defined be

$$e^A \equiv \sum_{n=0}^{\infty} \frac{A^n}{n!} = I + A + \frac{A^2}{2!} + \dots + \frac{A^n}{n!} + \dots \quad (90)$$

Note that as opposed to numbers, in general, $f(A)f(B) \neq f(B)f(A)$, unless the operators A and B commute. In fact, one can show that $e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]}$. This is known as *Glauber's formula*.

¹If the wavefunction ψ is not normalized, we define the expectation value of A as $\langle A \rangle \equiv \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}$.

This example is particularly useful when solving the Schrödinger equation, $i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle$. For a time-independent Hamiltonian, the solution to this equation is

$$|\psi(t)\rangle = e^{\frac{-iHt}{\hbar}}|\psi(t=0)\rangle,$$

and so the state vector at arbitrary time $|\psi(t)\rangle$ can be immediately calculated if $|\psi(t=0)\rangle$ (the state of the system at $t=0$) is known and the Hamiltonian can be written in the form of a diagonal matrix, as is frequently the case.

Example - 2. The operator σ_z (known as one of **Pauli's spin matrices**) is represented by the matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (91)$$

It follows directly that

$$e^{\sigma_z} = \begin{pmatrix} e & 0 \\ 0 & 1/e \end{pmatrix}. \quad (92)$$

6. Representations in state space

We stated that for each wavefunction ψ there corresponds a ket vector, $|\psi\rangle$ in the “state space” of the system. In order to calculate the expectation values of the different observables (as well as other calculations), we need first to define an orthonormal basis for this space (namely, all the basis vectors are normalized to unity, and are orthogonal to each other). This basis can be either *discrete*, or *continuous*, depending on the spectrum (=all possible results of measurements) of the particular observable.

Once we specify a basis, vectors and operators are represented by **numbers**: for vectors, these are their components in the given basis (exactly the same way as we represent vectors by their x , y and z components in Cartesian coordinate system). Operators will be represented by their matrix elements. In theory, the choice of representation is arbitrary. Practically, we obviously choose a representation which simplifies the calculations as much as possible.

Let us discriminate between the two cases: discrete and continuous basis. (To avoid confusion, I am omitting discussion on degeneracy).

6.1. Discrete basis

A set of discrete kets, $\{|\psi_i\rangle\}$ is said to be **orthonormal**, if the kets of this set satisfy the orthonormalization condition:

$$\langle\psi_i|\psi_j\rangle = \delta_{ij} \quad (93)$$

where δ_{ij} is the Kronecker delta symbol,

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases} \quad (94)$$

In order for the set $\{|\psi_i\rangle\}$ to form a basis, **every** wavefunction must have a unique expansion on the $|\psi_i\rangle$, namely every wavefunction can be expressed as a linear combination of $|\psi_i\rangle$:

$$|\Psi\rangle = \sum_i c_i |\psi_i\rangle \quad (95)$$

Since this must be true for *all* wave functions, the set $|\psi_i\rangle$ is said to be **complete**.

We can multiply both sides of Equation 95 by $\langle\psi_j|$, to get

$$\langle\psi_j|\Psi\rangle = c_j \quad (96)$$

Using this in Equation 95, one obtains:

$$|\Psi\rangle = \sum_i (\langle\psi_i|\Psi\rangle) |\psi_i\rangle = \sum_i |\psi_i\rangle \langle\psi_i|\Psi\rangle \quad (97)$$

We can think of $\sum_i |\psi_i\rangle \langle\psi_i|$ as an operator, that when acting on any arbitrary ket $|\Psi\rangle$ gives the same ket; thus,

$$\sum_i |\psi_i\rangle \langle\psi_i| = I \quad (98)$$

where I denotes the identity operator. Equation 98 is also known as the **closure relation**, as it expresses the completeness of the set of functions $\{|\psi_i\rangle\}$.

Probability amplitude.

Consider a system that is described by a wavefunction Ψ , which is normalized to unity. Let $\{\psi_n\}$ be a complete set of eigenfunctions of A , *whose corresponding kets $\{|\psi_n\rangle\}$ form an orthonormal basis.*

The expectation value of an (arbitrary) observable A is given by Equation 87,

$$\begin{aligned} \langle A \rangle &\equiv \langle \Psi | A | \Psi \rangle \\ &= \sum_m \sum_n c_m^* c_n \langle \psi_m | A | \psi_n \rangle \\ &= \sum_m \sum_n c_m^* c_n a_n \langle \psi_m | \psi_n \rangle \\ &= \sum_n |c_n|^2 a_n \end{aligned} \quad (99)$$

Since Ψ is normalized to unity, $\langle \Psi | \Psi \rangle = 1$ namely $\sum_n |c_n|^2 = 1$.

Now, recall the following: when we measure A , the only possible results that we can obtain are a_n . When we measure A multiple times (in a large number of identical systems), the average value we obtain is $\langle A \rangle$. Thus, we interpret the quantity

$$P_n = |c_n|^2 = |\langle \psi_n | \Psi \rangle|^2 \quad (100)$$

as *the probability of obtaining the particular value a_n in a given measurement*. The coefficients $c_n = \langle \psi_n | \Psi \rangle$ are called **probability amplitudes**.

Let us consider a scenario in which the state of a system, $|\Psi\rangle$ is known, and we want to measure an observable A . Before the measurement, we don't know what value we will get; we only know the probability amplitude of obtaining the various possible values. Suppose now we actually measure A , and obtained a value a_n . **Immediately after the measurement**, we obviously cannot speak of “probability of obtaining a value”; the value is known. Thus, clearly, the state of the system is no longer $|\Psi\rangle$. Instead,

The state of the system immediately after the measurement is the eigenvector $|\psi_n\rangle$ associated with a_n .

6.2. Continuous spectrum

The spectrum of some observables, e.g., the momentum operator p_x is continuous, while the spectrum of others, e.g., the Hamiltonian, possess both a discrete and continuous part. We thus need to generalize the treatment in the previous section to these cases.

The treatment is very similar to the discrete case. Let us denote the set of continuous eigenvalues by a , and eigenfunctions by ψ_a , namely $A\psi_a = a\psi_a$. We can write any arbitrary wave function Ψ as

$$\Psi = \int c(a)\psi_a da, \quad (101)$$

where the integral runs over all values of a . Alternatively, we could write $|\Psi\rangle = \int da c(a)|\psi_a\rangle$.

The set of eigenfunctions must be orthonormal, namely

$$\langle \psi_{a'} | \psi_a \rangle = \delta(a - a'), \quad (102)$$

(compare to Equation 93), enabling us to write the coefficients $c(a)$ as

$$c(a) = \langle \psi_a | \Psi \rangle \quad (103)$$

(compare to Equation 96).

The closure relation takes the form

$$\int da |\psi_a\rangle \langle \psi_a| = I. \quad (104)$$

The expectation value of an operator is written as

$$\begin{aligned} \langle A \rangle &\equiv \langle \Psi | A | \Psi \rangle \\ &= \int da \int da' c^*(a') c(a) \langle \psi_{a'} | A | \psi_a \rangle \\ &= \int da \int da' c^*(a') c(a) a \langle \psi_{a'} | \psi_a \rangle \\ &= \int da |c(a)|^2 a \end{aligned} \quad (105)$$

6.3. Matrix representation of wave functions and operators

6.3.1. Representation of kets

Let us assume that $\{|\psi_i\rangle\}$ form a complete set of orthonormal basis (for simplicity, I consider only the discrete case here). As we saw, this means that every wave function can be written as

$$|\Psi\rangle = \sum_i c_i |\psi_i\rangle$$

(see Equation 95). Therefore, once we know the set $\{|\psi_i\rangle\}$, the set of coefficients $c_i = \langle \psi_i | \Psi \rangle$ fully determines $|\Psi\rangle$. It is analogous to the way we describe a vector (in 3-d space) by its \hat{x} , \hat{y} and \hat{z} components, which we can specify after we define the coordinate system.

Thus, we can think of the set of functions $\{|\psi_i\rangle\}$ as analogue to the set of orthogonal axes, and c_i as analogue to the components of the 'vector' Ψ along each of these axes. We can arrange the c_i vertically, in a column vector = one column matrix (it can, in principle, be of infinite length!)

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle \psi_1 | \Psi \rangle \\ \langle \psi_2 | \Psi \rangle \\ \vdots \\ \langle \psi_n | \Psi \rangle \\ \vdots \end{pmatrix} \quad (106)$$

6.3.2. Representation of bras and scalar product

Similar to kets, every bra $\langle\Psi|$ has a unique representation in the orthonormal (bra) basis $\{|\psi_i\rangle\}$. Using the closure relation (Equation 98), we can write

$$\langle\Psi| = \langle\Psi|I = \sum_i \langle\Psi|\psi_i\rangle\langle\psi_i| \quad (107)$$

Obviously, the components of $\langle\Psi|$, namely $\langle\Psi|\psi_i\rangle$ are the complex conjugates of the components of the ket $|\Psi\rangle$ associated with the bra $\langle\Psi|$ (see Equation 67), and are therefore equal to c_i^* .

We thus arrange the components of the bra $\langle\Psi|$ horizontally, forming a row vector,

$$(\langle\Psi|\psi_1\rangle, \langle\Psi|\psi_2\rangle, \dots, \langle\Psi|\psi_n\rangle) \quad (108)$$

We can now look again at the scalar product (Equation 67). We can use the closure relation (Equation 98) to write

$$\langle\phi|\Psi\rangle = \langle\phi|I|\Psi\rangle = \sum_i \langle\phi|\psi_i\rangle\langle\psi_i|\Psi\rangle = \sum_i b_i^* c_i \quad (109)$$

where we wrote $b_i^* \equiv \langle\phi|\psi_i\rangle$.

Not surprising at all, this is exactly what we get when we multiply the row vector representing $\langle\phi|$ (Equation 108) with the column vector representing $|\Psi\rangle$ (Equation 106)!

6.3.3. Matrix representation of operators

Given a linear operator A , we can, in the $\{|\psi_i\rangle\}$ basis associate with it a series of numbers defined by

$$A_{ij} = \langle\psi_i|A|\psi_j\rangle. \quad (110)$$

These numbers are called the **matrix elements** of the operator A in the basis $\{|\psi_i\rangle\}$.

As these numbers depend on two indices, we can therefore arrange them in a matrix form:

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1j} & \dots \\ A_{21} & A_{22} & \dots & A_{2j} & \dots \\ \vdots & \vdots & & \vdots & \\ A_{i1} & A_{i2} & \dots & A_{ij} & \dots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \quad (111)$$

The operator A , when acting on the ket $|\Psi\rangle$ produces a new ket,

$$|\Phi\rangle = A|\Psi\rangle.$$

If we know the representation of the ket $|\Psi\rangle$ (namely, its components in the basis $\{|\psi_i\rangle\}$) and the matrix representation of A in the the same basis, we can calculate the components of $|\Phi\rangle$ using

$$b_i = \langle\psi_i|\Phi\rangle = \langle\psi_i|A|\Psi\rangle = \sum_j \langle\psi_i|A|\psi_j\rangle\langle\psi_j|\Psi\rangle = \sum_j A_{ij}c_j \quad (112)$$

where we made use of the closure relation (Equation 98).

This equation can also be written as a matrix equation, using the standard formula for matrix multiplication:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1j} & \dots \\ A_{21} & A_{22} & \dots & A_{2j} & \dots \\ \vdots & \vdots & & \vdots & \\ A_{i1} & A_{i2} & \dots & A_{ij} & \dots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \\ c_j \end{pmatrix} \quad (113)$$

Using this representation the matrix representation of the operator C which is obtained by the product of two operators, $C = AB$ is simply the matrix product of the two matrices representing the operators A and B . This can be verified by using again the closure relation (equation 98):

$$C_{ij} = \langle\psi_i|C|\psi_j\rangle = \langle\psi_i|AB|\psi_j\rangle = \sum_k \langle\psi_i|A|\psi_k\rangle\langle\psi_k|B|\psi_j\rangle = \sum_k A_{ik}B_{kj} \quad (114)$$

Matrix representation of adjoint operators.

Using Equation 78 ($\langle\psi|A^\dagger|\phi\rangle = \langle\phi|A|\psi\rangle^*$), we find:

$$(A^\dagger)_{ij} = \langle\psi_i|A^\dagger|\psi_j\rangle = \langle\psi_j|A|\psi_i\rangle^* = A_{ji}^* \quad (115)$$

Therefore, **the matrix elements of A^\dagger are obtained from the matrix elements of A by interchanging rows and columns and then taking the complex conjugate of each element.** (And vice versa, of course).

If A is **Hermitian**, then we have $A^\dagger = A$; therefore, $(A^\dagger)_{ij} = A_{ij}$ and therefore

$$A_{ij} = A_{ji}^* \quad (116)$$

Such a matrix is called **Hermitian matrix**. Namely, a Hermitian operator is represented by a matrix in which **any two elements which are symmetric with respect to the principle diagonal are complex conjugates of each other**.

In particular, all the elements on the diagonal, A_{ii} must have $A_{ii} = A_{ii}^*$, namely, **the diagonal elements of a Hermitian matrix are always real numbers**.

6.4. Finding eigenvalues and eigenvectors of an operator

Clearly, we say that $|\Psi\rangle$ is an eigenvector (or eigenket) of the operator A with eigenvalue λ if

$$A|\Psi\rangle = \lambda|\Psi\rangle. \quad (117)$$

We would like to answer the following question: given an operator A , how do we find all its eigenvalues and the corresponding eigenvectors ?

Let us assume that $\{|\psi_i\rangle\}$ forms a basis in some particular representation. We can then project the vector Equation 117 onto the various orthonormal basis vectors,

$$\langle\psi_i|A|\Psi\rangle = \lambda\langle\psi_i|\Psi\rangle \quad (118)$$

Inserting the closure relation between A and $|\Psi\rangle$ gives

$$\sum_j \langle\psi_i|A|\psi_j\rangle\langle\psi_j|\Psi\rangle = \lambda\langle\psi_i|\Psi\rangle \quad (119)$$

Using Equations 96 ($\langle\psi_j|\Psi\rangle = c_j$) and 110 ($A_{ij} = \langle\psi_i|A|\psi_j\rangle$), this becomes

$$\sum_j A_{ij}c_j = \lambda c_i, \quad (120)$$

or

$$\sum_j [A_{ij} - \lambda\delta_{ij}] c_j = 0. \quad (121)$$

This equation can be thought of a linear and homogeneous system of N equations, in N unknowns (c_j). From linear algebra, we know that non-trivial solution (namely, a solution in which not all $c_j = 0$) exists if and only if the determinant

$$\text{Det}[A - \lambda I] = 0, \quad (122)$$

where A is the $(N \times N)$ matrix of the A_{ij} and I is the unit matrix. Equation 122 is known as the **characteristic equation** (also: secular equation). The eigenvalues of the operator are the roots of its characteristic equation.

Once we solve the equation and find the eigenvalues, we can find the eigenvectors. Let us assume that λ_0 is an eigenvalue that solves the set of N equations. When we put $\lambda = \lambda_0$ in equation 120, we are left with $N - 1$ independent equations, but with N unknowns (c_j). Thus, we need to fix, say c_1 , and the rest will depend on this choice, namely $c_j = \alpha_j c_1$ for $j \neq 1$. (Obviously, for, $j = 1$, $\alpha_1 = 1$). We obtain as the eigenket, $|\psi(c_1)\rangle = \sum_j \alpha_j c_1 |\psi_j\rangle$.

7. Commuting observables, compatibility and the Heisenberg uncertainty relation

Consider two operators A and B . We defined earlier (see Equation 6) their commutator by

$$[A, B] \equiv AB - BA.$$

If the commutator vanishes when acting on **any** wave function ψ , then the two operators A and B are said to commute, namely $AB = BA$.

Assume now that A and B are two observables. If there exists a complete set of wavefunctions ψ_n such that each function is simultaneously an eigenfunction of both A and B , the observables A and B are said to be **compatible**. Denote by a_n and b_n the eigenvalues of A and B corresponding to the eigenkets $|\psi_n\rangle$:

$$A|\psi_n\rangle = a_n|\psi_n\rangle, \quad B|\psi_n\rangle = b_n|\psi_n\rangle. \quad (123)$$

If a state is described by $|\psi_n\rangle$, a measurement of A produces an exact result a_n , and a measurement of B produces an exact result b_n - both accurate to any precision.

Examples of compatible observers are the Cartesian components, x y and z of the position vector \vec{r} of a particle. Another example are the Cartesian components of its momentum, p_x , p_y and p_z . However, recall that x and p_x are not compatible (see Equation 51).

Two compatible observables commute. If $|\psi_n\rangle$ is a common eigenket, we have

$$AB|\psi_n\rangle = a_n b_n |\psi_n\rangle = b_n a_n |\psi_n\rangle = BA|\psi_n\rangle \quad (124)$$

Since any ket can be expanded as $|\Psi\rangle = \sum_n c_n |\psi_n\rangle$ (see Equation 95), we have

$$(AB - BA)|\Psi\rangle = \sum_n c_n (AB - BA)|\psi_n\rangle = 0, \quad (125)$$

from which it is obvious that $[A, B] = 0$.

The opposite statement, namely that two observables which commute possess a complete set of common eigenfunctions is also true.² Assume that $|\psi_n\rangle$ is an eigenket of A with eigenvalue a_n . If A and B commute, we have

$$A(B|\psi_n\rangle) = BA|\psi_n\rangle = a_n(B|\psi_n\rangle) \quad (126)$$

Thus, the ket $(B|\psi_n\rangle)$ is an eigenket of A with eigenvalue a_n . By assumption, a_n is non-degenerate, and so $(B|\psi_n\rangle)$ can only differ from $|\psi_n\rangle$ by a constant, which we call b_n , namely

$$B|\psi_n\rangle = b_n|\psi_n\rangle \quad (127)$$

We thus see that $|\psi_n\rangle$ is simultaneously an eigenket of the operators A and B , with the corresponding eigenvalues a_n and b_n .

7.1. Heisenberg uncertainty relation

We are now at a stage to obtain a precise form of Heisenberg's uncertainty relation. Consider two observables, A and B , whose expectation values are $\langle A \rangle \equiv \langle \Psi|A|\Psi \rangle$ and $\langle B \rangle \equiv \langle \Psi|B|\Psi \rangle$ when the system is in state $|\Psi \rangle$ (see Equation 87).

We **define** the uncertainty ΔA to be

$$(\Delta A)^2 \equiv \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \quad (128)$$

and a similar definition of ΔB .

We shall now prove the mathematical formula:

$$\Delta A \cdot \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle| \quad (129)$$

Proof. Consider the operators

$$\bar{A} \equiv A - \langle A \rangle, \quad \bar{B} \equiv B - \langle B \rangle.$$

The expectation values of both these operators vanish, $\langle \bar{A} \rangle = \langle \bar{B} \rangle = 0$. We further have

$$\langle \bar{A}^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 = (\Delta A)^2 \quad (130)$$

²I prove this statement here only for the non-degenerate case.

Furthermore,

$$[\bar{A}, \bar{B}] = [A - \langle A \rangle, B - \langle B \rangle] = [A, B] \quad (131)$$

since the expectation values are just numbers.

Consider now the linear (non-Hermitian) operator

$$C = \bar{A} + i\lambda\bar{B}, \quad (132)$$

where λ is a *real* constant. The adjoint operator is $C^\dagger = \bar{A} - i\lambda\bar{B}$ (recall that both \bar{A} and \bar{B} are Hermitian), and thus the expectation value of CC^\dagger is real and non-negative, since

$$\langle CC^\dagger \rangle = \langle \Psi | CC^\dagger | \Psi \rangle = \langle C^\dagger \Psi | C^\dagger \Psi \rangle \geq 0. \quad (133)$$

We thus deduce that the expectation value of

$$\langle (\bar{A} + i\lambda\bar{B})(\bar{A} - i\lambda\bar{B}) \rangle = \langle \bar{A}^2 + \lambda^2\bar{B}^2 - i\lambda[\bar{A}, \bar{B}] \rangle \quad (134)$$

is both real and non-negative. Using Equations 134, 130 and 131, this means that the function

$$f(\lambda) = \langle \bar{A}^2 \rangle + \lambda^2 \langle \bar{B}^2 \rangle - i\lambda \langle [\bar{A}, \bar{B}] \rangle = (\Delta A)^2 + \lambda^2 (\Delta B)^2 - i\lambda \langle [A, B] \rangle \quad (135)$$

is also real and non-negative. Therefore, $\langle [A, B] \rangle$ must be purely imaginary.

By differentiating with respect to λ , one finds the minimum of $f(\lambda)$ to be given for

$$\lambda = \lambda_0 = \frac{i \langle [A, B] \rangle}{2 (\Delta B)^2}. \quad (136)$$

At $\lambda = \lambda_0$, the value of $f(\lambda)$ is

$$f(\lambda_0) = (\Delta A)^2 + \frac{1}{4} \frac{(\langle [A, B] \rangle)^2}{(\Delta B)^2}. \quad (137)$$

Since we know that $f(\lambda_0) \geq 0$, we find that

$$(\Delta A)^2 (\Delta B)^2 \geq -\frac{1}{4} (\langle [A, B] \rangle)^2. \quad (138)$$

Finally, since $\langle [A, B] \rangle$ is purely imaginary, Equation 129

$$\Delta A \cdot \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|$$

follows, which completes the proof.

Example. For two observables for which $[A, B] = i\hbar$, we have $\langle [A, B] \rangle = i\hbar$ and we therefore deduce

$$\Delta A \Delta B \geq \frac{\hbar}{2} \quad (139)$$

Specifically, we have

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \quad \Delta y \Delta p_y \geq \frac{\hbar}{2}, \quad \Delta z \Delta p_z \geq \frac{\hbar}{2}. \quad (140)$$

where $\Delta x = [(\langle (x - \langle x \rangle)^2 \rangle)]^{1/2}$, $\Delta p_x = [(\langle (p_x - \langle p_x \rangle)^2 \rangle)]^{1/2}$, etc.

8. Summary: basic postulates of quantum mechanics

Let us summarize the basic principles and postulates of quantum mechanics.

1. The state of a system is completely specified by a wave function ψ , or state vector, $|\psi\rangle$, which contain all the information that can be known about the system. This function is, in general, complex. The interpretation of this wave function is probabilistic.
2. Every measurable physical quantity is described by a Hermitian operator A . The operator is called an observable.
3. The only possible result of a precise measurement of a physical quantity is one of the eigenvalues of the corresponding observable A .
4. In a series of measurements of a variable A made in an ensemble of systems all having normalized state function $|\Psi\rangle$, the expectation value of this variable is

$$\langle A \rangle = \langle \Psi | A | \Psi \rangle.$$

The probability of obtaining the (non-degenerated) eigenvalue a_n is

$$P(a_n) = |\langle \psi_n | \Psi \rangle|^2,$$

where $|\psi_n\rangle$ is the normalized eigenvector of A associated with the eigenvalue a_n .

5. (Completeness): Any wavefunction can be expressed as a linear combination of the eigenfunctions of A , where A is an operator associated with a dynamical variable (i.e., an observable).

6. The temporal evolution of a state vector $|\psi\rangle$ is governed by the Schrödinger equation,

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi\rangle$$

where \hat{H} is the Hamiltonian operator, which is the observable associated with the total energy of the system.

Using Schrödinger equation, we can calculate the temporal evolution of the expectation value of an operator. We have

$$\begin{aligned} \frac{d}{dt} \langle A \rangle &= \frac{d}{dt} \langle \Psi | A | \Psi \rangle = \left\langle \frac{\partial \Psi}{\partial t} | A | \Psi \right\rangle + \left\langle \Psi | \frac{\partial A}{\partial t} | \Psi \right\rangle + \left\langle \Psi | A | \frac{\partial \Psi}{\partial t} \right\rangle \\ &= -\frac{1}{i\hbar} \langle \hat{H} \Psi | A | \Psi \rangle + \left\langle \Psi | \frac{\partial A}{\partial t} | \Psi \right\rangle + \frac{1}{i\hbar} \langle \Psi | A | \hat{H} \Psi \rangle \\ &= \frac{1}{i\hbar} \langle [A, \hat{H}] \rangle + \left\langle \frac{\partial A}{\partial t} \right\rangle \end{aligned} \quad (141)$$

In particular, if the operator A does not depend explicitly on time, namely $\partial A / \partial t = 0$, we have

$$\frac{d}{dt} \langle A \rangle = \frac{1}{i\hbar} \langle [A, \hat{H}] \rangle. \quad (142)$$

This means that if the operator A commutes with \hat{H} (and $\partial A / \partial t = 0$), its expectation value does not vary in time. We say that **the observable A is a constant of motion**.

Example. Consider a system in which the Hamiltonian is time-independent, $\partial \hat{H} / \partial t = 0$. Using Equation 142 with $A = \hat{H}$, we see that $d\langle \hat{H} \rangle / dt = 0$. This implies that *the total energy is constant of the motion*. This is the quantum mechanical analogue of energy conservation in classical mechanics.

A. Appendix: important examples of representations.

We argued that for every wavefunction ψ there corresponds a ket $|\psi\rangle$ in the “state space” of the system. In order to calculate the expectation values of the different observables, we need first to define an orthonormal basis for this space. Here, I will demonstrate two important examples of such bases: the $\{|r\rangle\}$ and $\{|p\rangle\}$ representations.

A.1. The $\{|r\rangle\}$ representation

Let us, for the moment, return to the description of wavefunction $\psi(\vec{r}, t)$ as a function of space and time (rather than ket vectors). Consider a set of functions $\{\xi_{\vec{r}_0}(\vec{r})\}$, which are

labels by a **continuous** index \vec{r}_0 . These are defined via

$$\xi_{\vec{r}_0}(\vec{r}) = \delta(\vec{r} - \vec{r}_0). \quad (\text{A1})$$

Thus, $\{\xi_{\vec{r}_0}(\vec{r})\}$ represent a set of delta functions, each centered at a point \vec{r}_0 of space. Obviously, each individual function $\xi_{\vec{r}_0}(\vec{r})$ by itself is **not** square integrable, hence these functions are not legitimate wave functions.³

On the other hand, every (legitimate) wave function $\psi(\vec{r})$ can be written as

$$\psi(\vec{r}) = \int d^3r_0 \psi(\vec{r}_0) \delta(\vec{r} - \vec{r}_0) = \int d^3r_0 \psi(\vec{r}_0) \xi_{\vec{r}_0}(\vec{r}) \quad (\text{A2})$$

Moreover, this expansion of the (arbitrary) wavefunction in terms of $\{\xi_{\vec{r}_0}(\vec{r})\}$ is unique. For this reason, despite the fact that $\xi_{\vec{r}_0}(\vec{r})$ is not a legitimate wavefunction, we can look at the set of delta functions $\{\xi_{\vec{r}_0}(\vec{r})\}$ as forming a basis for the wavefunction space.

Let us return now to Dirac notation. We denote the ket associated with $\xi_{\vec{r}_0}(\vec{r})$ simply by $|\vec{r}_0\rangle$,

$$\xi_{\vec{r}_0}(\vec{r}) \iff |\vec{r}_0\rangle. \quad (\text{A3})$$

The set of ket vectors $|\vec{r}_0\rangle$ form a (continuous) basis; it is known as **the $\{|r\rangle\}$ representation**.

In analogy to Equation A2, we have

$$|\psi\rangle = \int d^3r_0 c(\vec{r}_0) |\vec{r}_0\rangle \quad (\text{A4})$$

where

$$c(\vec{r}_0) = \langle \vec{r}_0 | \psi \rangle = \int d^3r \xi_{\vec{r}_0}^*(\vec{r}) \psi(\vec{r}) = \psi(\vec{r}_0) \quad (\text{A5})$$

where we used the fact that $\xi_{\vec{r}_0}^*(\vec{r}) = \delta(\vec{r} - \vec{r}_0) = \xi_{\vec{r}_0}(\vec{r})$.

³This statement follows directly from the properties of the δ -function.

$$\int \delta(\vec{r} - \vec{r}_0) f(\vec{r}) d^3r = f(\vec{r}_0)$$

Use $f(\vec{r}) = \delta(\vec{r} - \vec{r}_0)$, to get

$$\int \delta(\vec{r} - \vec{r}_0) \delta(\vec{r} - \vec{r}_0) d^3r = \delta(\vec{r}_0 - \vec{r}_0) = \delta(0) = \infty$$

which is not square integrable.

Orthonormalization. Using the definition of the scalar product, let us calculate:

$$\langle \vec{r}_0 | \vec{r}'_0 \rangle = \int d^3r \xi_{\vec{r}_0}^*(\vec{r}) \xi_{\vec{r}'_0}(\vec{r}) = \int d^3r \delta(\vec{r} - \vec{r}_0) \delta(\vec{r} - \vec{r}'_0) = \delta(\vec{r}_0 - \vec{r}'_0) \quad (\text{A6})$$

The $\{|r_0\rangle\}$ basis is therefore orthonormal.

Closure. From Equations A4 and A5 above, it follows immediately that

$$\int d^3r_0 |\vec{r}_0\rangle \langle \vec{r}_0| = I \quad (\text{A7})$$

These results prove that indeed the set $\{|r_0\rangle\}$ form an orthonormal basis.

From Equation A5, it follows that the value $\psi(\vec{r}_0)$ of any (arbitrary) wavefunction ψ at the point \vec{r}_0 is simply the component of the ket $|\psi\rangle$ on the basis vector $|\vec{r}_0\rangle$ of the $\{|r_0\rangle\}$ representation (in short, it is often denoted simply as the $\{|r\rangle\}$ representation).

A.2. The $\{|p\rangle\}$ representation

We previously discussed functions of momentum. In the derivation, we used the Fourier transform to write

$$\psi(\vec{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \phi(\vec{p}, t) e^{i(\vec{p}\cdot\vec{r})/\hbar} d^3p, \quad (\text{A8})$$

and

$$\phi(\vec{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{r}, t) e^{-i\vec{p}\cdot\vec{r}/\hbar} d^3r. \quad (\text{A9})$$

This naturally leads to the following idea: one can define a new function

$$v_{\vec{p}}(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i(\vec{p}\cdot\vec{r})/\hbar}. \quad (\text{A10})$$

The function $v_{\vec{p}}(\vec{r})$ is a plane wave, with wave vector p/\hbar . Similar to the function $\xi_{\vec{r}_0}(\vec{r})$, each function $v_{\vec{p}}(\vec{r})$ is **not** square integrable, as $|v_{\vec{p}}(\vec{r})|^2$ diverges. Therefore, it cannot represent a legitimate wave function.

However, similar to the set $\{\xi_{\vec{r}_0}(\vec{r})\}$, each legitimate wave function can be completely written in terms of the set $\{v_{\vec{p}}(\vec{r})\}$. This follows directly from Equation A8, which can be written as

$$\psi(\vec{r}, t) = \int_{-\infty}^{+\infty} \phi(\vec{p}, t) v_{\vec{p}}(\vec{r}) d^3p. \quad (\text{A11})$$

In Dirac notation, we write Equation A10 using

$$v_{\vec{p}}(\vec{r}) \iff |\vec{p}\rangle, \quad (\text{A12})$$

as

$$|\psi\rangle = \int d^3p c(\vec{p}) |\vec{p}\rangle \quad (\text{A13})$$

where

$$c(\vec{p}) = \langle \vec{p} | \psi \rangle = \int d^3r v_{\vec{p}}^*(\vec{r}) \psi(\vec{r}) = \phi(\vec{p}) \quad (\text{A14})$$

(see Equation A9).

Physically, Equation A11 proves that each wavefunction can be expanded in one (and only one) way in terms of the plane waves $v_{\vec{p}}(\vec{r})$.

Furthermore, it is easy to show that the plane waves are orthonormal:

$$\langle \vec{p} | \vec{p}' \rangle = \int d^3r v_{\vec{p}}^*(\vec{r}) v_{\vec{p}'}(\vec{r}) = \frac{1}{(2\pi\hbar)^3} \int d^3r e^{i(\vec{p}-\vec{p}')\cdot\vec{r}/\hbar} = \delta(\vec{p}-\vec{p}') \quad (\text{A15})$$

Combining Equations A13, A14 and A15 further implies the closure of the set $\{v_{\vec{p}}(\vec{r})\}$.

A.3. Changing between the $\{|r\rangle\}$ and $\{|p\rangle\}$ representations

While this is part of a bigger question, let us use the example given above to demonstrate how one can change between different bases. We know:

$$\begin{aligned} \langle \vec{r}_0 | \vec{p} \rangle &= \langle \vec{p} | \vec{r}_0 \rangle^* = \int d^3r \xi_{\vec{r}_0}^*(\vec{r}) v_{\vec{p}}(\vec{r}) \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r \delta(\vec{r} - \vec{r}_0) e^{i\vec{p}\cdot\vec{r}/\hbar} \\ &= \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{p}\cdot\vec{r}_0/\hbar} \end{aligned} \quad (\text{A16})$$

A given ket $|\psi\rangle$ is represented by $\langle \vec{r}_0 | \psi \rangle = \psi(\vec{r}_0)$ in the $\{|r_0\rangle\}$ representation, and by $\langle \vec{p} | \psi \rangle = \phi(\vec{p})$ in the $\{|p\rangle\}$ representation (see Equations A5 and A14).

Given a representation of $|\psi\rangle$ in the $\{|\vec{p}\rangle\}$ representation, we can use the closure relation of the $\{|p\rangle\}$ representation to write it in the $\{|\vec{r}_0\rangle\}$ representation:

$$\langle \vec{r}_0 | \psi \rangle = \int d^3p \langle \vec{r}_0 | \vec{p} \rangle \langle \vec{p} | \psi \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p e^{i\vec{p}\cdot\vec{r}_0/\hbar} \phi(\vec{p}) \quad (\text{A17})$$

where we made use of Equations A16 and A14. But this is of course identical to the Fourier transform in Equation A8!

Obviously, the inverse transform is also true.

REFERENCES

- [1] B.H. Bransden & C.J. Joachain, *Quantum Mechanics* (Second edition), Chapter 3, 5.
- [2] C. Cohen-Tannoudji, B. Diu & F. Laloe, *Quantum Mechanics*, Vol. 1., Chapters 2,3.
- [3] A. Messiah, *Quantum Mechanics*, Vol. 1., Chapters 4, 5.
- [4] D. Griffiths, *Introduction to Quensutm Mechanics*, Chapter 3.