2. Introduction to Quantum Mechanics

2.1 Laws of Quantum Mechanics

Every physical theory is formulated in terms of mathematical objects. It is thus necessary to establish a set of rules to map physical concepts and objects into mathematical objects that we use to represent them. Sometimes this mapping is evident, as in classical mechanics, while for other theories, such as quantum mechanics, the mathematical objects are not intuitive.

In the same way as classical mechanics is founded on Newton's laws or electrodynamics on the Maxwell-Boltzmann equations, quantum mechanics is also based on some fundamental laws, which are called the postulates or axioms of quantum mechanics.

We want in particular to develop a mathematical model for the dynamics of closed quantum systems [1]: therefore we are interested in defining

$$\text{states} \quad - \quad \text{observables} \quad - \quad \text{measurements} \quad - \quad \text{evolution}$$

Some subtleties will arise since we are trying to define measurement in a closed system, when the measuring person is instead outside the system itself. A more complete picture, that can explain some of the confusion arising from the measurement process, is possible, but we will not study it in this course.

We are interested in giving a description of physical phenomena and in particular in how they emerge during an experiment.

Experiments – A physical experiment can be divided into two steps: preparation and measurement.

In classical mechanics (CM):

- the first step determines the possible outcomes of the experiment,

- while the measurement retrieves the value of the outcome.

In quantum mechanics (QM) the situation is slightly different:

- the first step (preparation) determines the *probabilities* of the various possible outcomes,

- the second step (measurement) retrieve the *value* of a particular outcome, in a statistic manner.

This separation of the experiment in two steps is reflected into the two types of operators that we find in QM.

- The first step corresponds to the concept of a **state** of the system,

---

[1] We define a *closed* system any system that is isolated, thus not exchanging any input or output and not interacting with any other system. An *open* system instead interacts e.g., with an external environment.

- while the second step corresponds to **observables**.

In CM the state of a system is described by a set of properties. For example, if we consider a ball, we can define its state by giving its position, momentum, energy, angular momentum (if for example the ball is spinning), its temperature etc. We can then perform a measurement on this ball, for example measuring its position. This will give us one value for one possible observable (the position).

We can express this process in mathematical terms. The state of the system is defined by a set of values: $\{\vec{r}, \vec{p}, E, \vec{L}, T, \dots\}$. All of these values (and there might be of course more that I haven't written down) are needed to fully describe the state of the ball. Performing a measurement of the position, will retrieve the values $\{r_x, r_y, r_z\} = \vec{r}$ (the same values that describe the state).

If we now consider a nucleus, we can as well give a description of its state. In quantum mechanics, a complete description of the state of a quantum object (or system) is given mathematically by the state vector $|\psi\rangle$ (or wavefunction $\psi(\vec{r})$). The situation is however different than in classical mechanics.

The state vector is no longer a collection of values for different properties of the system. The state gives instead a complete description of the set of *probabilities* for all the physical properties (or observables). All the information is contained in the state, irrespectively on how I got the state, of its previous history.

On the other hand, the observables are all the physical properties that in principle can be measured, in the same way as it was in classical mechanics. Since however the state only gives probabilities for all observables, the result of measurement will be a statistical variable.

All of these considerations are made more formal in the axioms of quantum mechanics that also indicate the mathematical formalism to be used.

1. The properties of a quantum system are completely defined by specification of its state vector $|\psi\rangle$. The state vector is an element of a complex Hilbert space H called the space of states.
2. With every physical property $\mathcal{A}$ (energy, position, momentum, angular momentum, ...) there exists an associated linear, Hermitian operator A (usually called observable), which acts in the space of states H. The eigenvalues of the operator are the possible values of the physical properties.
3.a If $|\psi\rangle$ is the vector representing the state of a system and if $|\varphi\rangle$ represents another physical state, there exists a probability $p(|\psi\rangle, |\varphi\rangle)$ of finding $|\psi\rangle$ in state $|\varphi\rangle$, which is given by the squared modulus of the inner product on $\mathcal{H}$: $p(|\psi\rangle, |\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$ (Born Rule).
3.b If A is an observable with eigenvalues a_n and eigenvectors $|n\rangle$ [such that the eigenvalue equation is $A|n\rangle = a_n |n\rangle$], given a system in the state $|\psi\rangle$, the probability of obtaining a_n as the outcome of the measurement of A is $p(a_n) = |\langle n|\psi\rangle|^2$. After the measurement the system is left in the state projected on the subspace of the eigenvalue a_n (Wave function collapse).
4. The evolution of a closed system is unitary (reversible). The evolution is given by the time-dependent Schrödinger equation

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = \mathcal{H}|\psi\rangle$$

where $\mathcal{H}$ is the Hamiltonian of the system (the energy operator) and $\hbar$ is the reduced Planck constant $h/2\pi$ (with h the Planck constant, allowing conversion from energy to frequency units).

2.2 States, observables and eigenvalues

$\mathcal{D}$:   From the first postulate we see that the state of a quantum system is given by the state vector $|\psi(t)\rangle$ (or the wavefunction $\psi(\vec{x}, t)$). The state vector contains all possible information about the system. The state vector is a vector in the Hilbert space. A Hilbert space H is a complex vector space that possess an inner product.

An example of Hilbert space is the usual Euclidean space of geometric vectors. This is a particularly simple case since the space in this case is real. In general as we will see, Hilbert space vectors can be complex (that is, some of their components can be complex numbers). In the 3D Euclidean space we can define vectors, with a representation such as $\vec{v} = \{v_x, v_y, v_z\}$ or :

$$\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

This representation corresponds to choose a particular basis for the vector (in this case, the usual $\{x, y, z\}$ coordinates). We can also define the inner product between two vectors, $\vec{v}$ and $\vec{u}$ (which is just the usual scalar product):

$$\vec{v} \cdot \vec{u} = [v_x \; v_y \; v_z] \cdot \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = v_x u_x + v_y u_x + v_z u_z$$

Notice that we have taken the transpose of the vector $\vec{v}$, $\vec{v}^T$ in order to calculate the inner product. In general, in a Hilbert space, we can define the dual of any vector. The Dirac notation makes this more clear.

The notation $|\psi\rangle$ is called the Dirac notation and the symbol $|\cdot\rangle$ is called *ket*. This is useful in calculating inner products of state vectors using the *bra* $\langle\cdot|$ (which is the dual of the ket), for example $\langle\varphi|$. An inner product is then written as $\langle\varphi|\psi\rangle$ (this is a *bracket*, hence the names).

We will often describe states by their wavefunction instead of state vector. The wavefunction is just a particular way of writing down the state vector, where we express the state vector in a *basis* linked to the position of a particle itself (this is called the position representation). This particular case is however the one we are mostly interested in this course. Mathematically, the wavefunction is a complex function of space and time. In the position representation (that is, the position basis) the state is expressed by the wavefunction via the inner product $\psi(x) = \langle x|\psi\rangle$.

The properties of Hilbert spaces, kets and bras and of the wavefunction can be expressed in a more rigorous mathematical way. In this course as said we are mostly interested in systems that are nicely described by the wavefunction. Thus we will just use this mathematical tool, without delving into the mathematical details. We will see some more properties of the wavefunction once we have defined observables and measurement.

$\mathcal{D}$: All physical observables (defined by the prescription of experiment or measurement) are represented by a linear operator that operates in the Hilbert space H (a linear, complex, inner product vector space).

In mathematics, an operator is a type of function that acts on functions to produce other functions. Formally, an operator is a mapping between two function spaces[2] $A : g(I) \to f(I)$ that assigns to each function $g \in g(I)$ a function $f = A(g) \in f(I)$.

Examples of observables are what we already mentioned, e.g. position, momentum, energy, angular momentum. These operators are associated to classical variables. To distinguish them from their classical variable counterpart, we will thus put a *hat* on the operator name. For example, the position operators will be $\hat{x}$, $\hat{y}$, $\hat{z}$. The momentum operators $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ and the angular momentum operators $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$. The energy operator is called Hamiltonian (this is also true in classical mechanics) and is usually denoted by the symbol $\mathcal{H}$.

There are also some operators that do not have a classical counterpart (remember that quantum-mechanics is more general than classical mechanics). This is the case of the spin operator, an observable that is associated to each particle (electron, nucleon, atom etc.). For example, the spin of an electron is usually denoted by S; this is also a vector variable (i.e. we can define S_x, S_y, S_z). I am omitting here the *hat* since there is no classical variable we can confuse the spin with. While the position, momentum etc. observable are continuous operator, the spin is a discrete operator.

The second postulate states that the possible values of the physical properties are given by the eigenvalues of the operators.

$\mathcal{D}$: **Eigenvalues and eigenfunctions** of an operator are defined as the solutions of the eigenvalue problem:

$$A[u_n(\vec{x})] = a_n u_n(\vec{x})$$

where $n = 1, 2, \dots$ indexes the possible solutions. The a_n are the eigenvalues of A (they are scalars) and $u_n(\vec{x})$ are the eigenfunctions.

The eigenvalue problem consists in finding the functions such that when the operator A is applied to them, the result is the function itself multiplied by a scalar. (Notice that we indicate the action of an operator on a function by $A[f(\cdot)]$).

You should have seen the eigenvalue problem in linear algebra, where you studied eigenvectors and eigenvalues of matrices. Consider for example the spin operator for the electron S. The spin operator can be represented by the following matrices (this is called a matrix representation of the operator; it's not unique and depends on the basis chosen):

$$S_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We can calculate what are the eigenvalues and eigenvectors of this operators with some simple algebra. In class we considered the eigenvalue equations for S_x and S_z. The eigenvalue problem can be solved by setting the determinant of the matrix $S_\alpha - s\mathbb{1}$ equal to zero. We find that the eigenvalues are $\pm\frac{1}{2}$ for both operators. The eigenvectors are different:

$$v_1^z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad v_2^z = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$v_1^x = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad v_2^x = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

---

[2] A function space $f(I)$ is a collection of functions satisfying certain properties.

We proved also that $v_1 \cdot v_2 = 0$ (that is, the eigenvectors are orthogonal) and that they form a complete basis (we can write any other vector, describing the state of the electron spin, as a linear combination of either the eigenvectors of S_z or of S_x).

The eigenvalue problem can be solved in a similar way for continuous operators. Consider for example the differential operator, $\frac{d\,[\cdot]}{d\,x}$. The eigenvalue equation for this operator reads:

$$\frac{d\,f(x)}{d\,x} = af(x)$$

where a is the eigenvalue and $f(x)$ is the eigenfunction.

? Question: what is $f(x)$? What are all the possible eigenvalues (and their corresponding eigenfunctions)?

Examples – The eigenvalue equation for the operator is $x\frac{d\,[\cdot]}{d\,x}$ is:

$$x\frac{d\,f(x)}{d\,x} = af(x)$$

which is solved by $f(x) = x^n$, $a = n$.

The "standard" Gaussian function $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is the eigenfunction of the Fourier transform. The Fourier transform is an operation that transforms one complex-valued function of a real variable into another one (thus it is an operator):

$$\mathcal{F}_x : f(x) \to \tilde{f}(k), \qquad \text{with} \qquad \tilde{f}(k) = \mathcal{F}_x[f(x)](k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx$$

Notice that sometimes different normalizations are used. With this definition, we also find that the inverse Fourier transform is given by:

$$\mathcal{F}_k^{-1} : \tilde{f}(k) \to f(x), \qquad f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{f}(k)e^{ikx}dk$$

Let's now turn to quantum mechanical operators.

$\mathcal{D}$: **Position operator** – The position operator for a single particle $\hat{\vec{x}}$ is simply given by the scalar $\vec{x}$. This means that the operator $\hat{\vec{x}}$ acting on the wavefunction $\psi(\vec{x})$ simply multiplies the wavefunction by $\vec{x}$. We can write

$$\boxed{\hat{\vec{x}}[\psi(\vec{x})] = \vec{x}\,\psi(\vec{x}).}$$

We can now consider the eigenvalue problem for the position operator. For example, for the x-component of $\vec{x}$ this is written as:

$$\hat{x}[u_n(x)] = x_n u_n(x) \quad \to \quad x\,u_n(x) = x_n u_n(x)$$

where we used the definition of the position operator. Here x_n is the eigenvalue and $u_n(x)$ the eigenfunction. The solution to this equation is not a proper function, but a distribution (a generalized function): the Dirac delta function: $u_n(x) = \delta(x - x_n)$

$\mathcal{D}$: **Dirac Delta function** $\delta(x - x_0)$ is equal to zero everywhere except at x_0 where it is infinite. The Dirac Delta function also has the property that $\int_{-\infty}^{\infty} \delta(x)dx = 1$ and of course $x\delta(x - x_0) = x_0\delta(x - x_0)$ (which corresponds to the eigenvalue problem above). We also have:

$$\int dx\delta(x - x_0)f(x) = f(x_0)$$

That is, the integral of any function multiplied by the delta function gives back the function itself evaluated at the point x_0. [See any textbook (and recitations) for other properties.]

How many solutions are there to the eigenvalue problem defined above for the position operator? One per each possible position, that is an infinite number of solutions. Conversely, all possible positions are allowed values for the measurement of the position (a continuum of solutions in this case).

$\mathcal{D}$: **Momentum operator** – The momentum operator is defined (in analogy with classical mechanics) as the generator of translations. This means that the momentum modifies the position of a particle from $\vec{x}$ to $\vec{x} + d\vec{x}$. It is possible to show that this definition gives the following form of the position operator (in the position representation, or position basis)

$$\boxed{\hat{p}_x = -i\hbar\frac{\partial}{\partial x}, \ \hat{p}_y = -i\hbar\frac{\partial}{\partial y}, \ \hat{p}_z = -i\hbar\frac{\partial}{\partial z}}$$
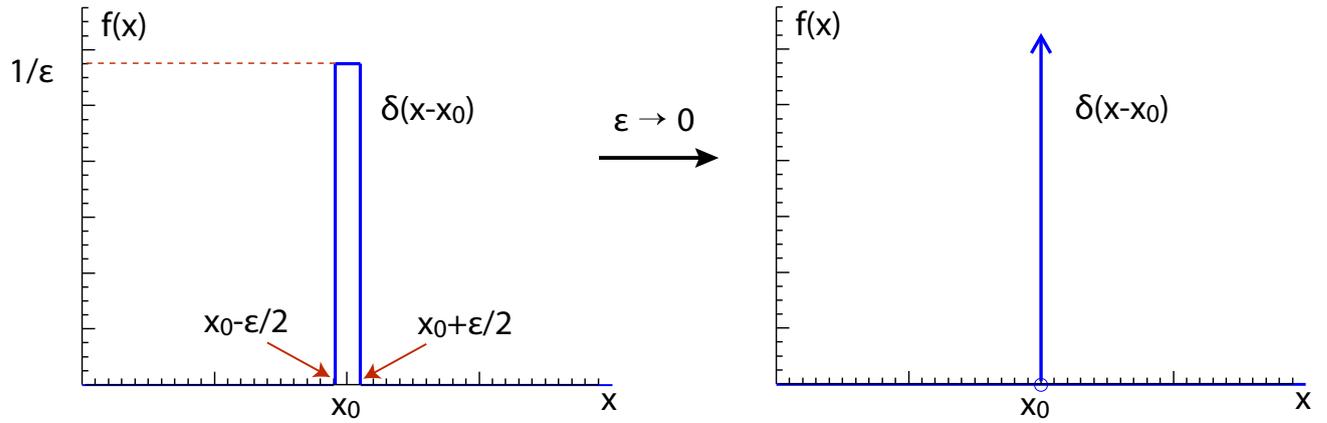
Fig. 9: Schematics of Dirac's delta function. Left: the rectangular function of base ϵ and height ϵ becomes the delta-function (right) in the limit of $\epsilon \to 0$.

or in vector notation $\hat{\mathbf{p}} = -i\hbar\nabla$. Here $\hbar$ is the reduced Planck constant $h/2\pi$ (with h the Planck constant) with value

$$\hbar = 1.054 \times 10^{-34} \text{J s}.$$

Planck's constant is introduced in order to make the values of quantum observables consistent with the corresponding classical values.

We now study the momentum operator eigenvalue problem in 1D. The problem's statement is

$$\hat{p}_x[u_n(x)] = p_n u_n(x) \quad \to \quad -i\hbar\frac{\partial u_n(x)}{\partial x} = p_n u_n(x)$$

This is a differential equation that we can solve quite easily. We set $k = p/\hbar$ and call k the wavenumber (for reasons clear in a moment). The differential equation is then

$$\frac{\partial u_n(x)}{\partial x} = ik_n u_n(x)$$

which has as solution the complex function:

$$u_n(x) = Ae^{ik_n x} = Ae^{i\frac{p_n}{\hbar}x}$$

The momentum eigenfunctions and eigenvalues are thus $u_n = Ae^{ik_n x}$ and k_n.

Now remember the meaning of the eigenvalues. By the second postulate, the eigenvalues of an operator are the possible values that one can obtain in a measurement.

Obs. 1 There are no restrictions on the possible values obtained from a momentum measurements. All values $p = \hbar k$ are possible.

Obs. 2 The eigenfunction $u_n(x)$ corresponds to a wave traveling to the right with momentum $p_n = \hbar k_n$. This was also expressed by De Broglie when he postulated the existence of matter waves.

Louis de Broglie (1892-1987) was a French physicist. In his Ph.D thesis he postulated a relationship between the momentum of a particle and the wavelength of the wave associated with the particle (1922). In de Broglie's equation a particle wavelength is the Planck's constant divided by the particle momentum. We can see this behavior in the electron interferometer video[3]. For classical objects the momentum is very large (since the mass is large), then the wavelength is very small and the object loose its wave behavior. De Broglie equation was experimentally confirmed in 1927 when physicists Lester Germer and Clinton Davisson fired electrons at a crystalline nickel target and the resulting diffraction pattern was found to match the predicted values.

2.2.1 Properties of eigenfunctions

From these examples we can notice two properties of eigenfunctions which are valid for any operator:

---

[3] A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki and H. Ezawa, Am. J. of Phys. 57, 117 (1989)

1. The eigenfunctions of an operator are orthogonal functions. We will as well assume that they are normalized. Consider two eigenfunctions u_n, u_m of an operator A and the inner product defined by $\langle f|g \rangle = \int d^3x \, f^*(\mathbf{x})g(\mathbf{x})$. Then we have

$$\int d^3x \, u_m^*(\mathbf{x})u_n(\mathbf{x}) = \delta_{nm}$$

2. The set of eigenfunctions forms a **complete basis**. This means that any other function can be written in terms of the set of eigenfunctions $\{u_n(\mathbf{x})\}$ of an operator A:

$$f(\mathbf{x}) = \sum_n c_n u_n(\mathbf{x}), \qquad \text{with} \qquad c_n = \int d^3x \, u_n^*(\mathbf{x})f(\mathbf{x})$$

[Note that the last equality is valid iif the eigenfunctions are normalized, which is exactly the reason for normalizing them].
If the eigenvalues are a continuous parameter, we have a continuum of eigenfunctions, and we will have to replace the sum over n with an integral.

Consider the two examples we saw. From the property of the Dirac Delta function we know that we can write any function as:

$$f(x) = \int dx' \delta(x' - x) f(x')$$

We can interpret this equation as to say that any function can be written in terms of the position eigenfunction $\delta(x' - x)$ (notice that we are in the continuous case mentioned before, since the x-eigenvalue is a continuous function). In this case the coefficient c_n becomes also a continuous function

$$c_n \to c(x_n) = \int dx \, \delta(x - x_n) f(x) = f(x_n).$$

This is not surprising as we are already expressing everything in the position basis.
If we want instead to express the function $f(x)$ using the basis given by the momentum operator eigenfunctions we have: (consider 1D case)

$$f(x) = \int dk \, u_k(x) c(k) = \int dk \, e^{ikx} c(k)$$

where again we need an integral since there is a continuum of possible eigenvalues. The coefficient $c(k)$ can be calculated from

$$c(k) = \int dx \, u_k^*(x) f(x) = \int dx \, e^{-ikx} f(x)$$

We then have that $c(k)$ is just the Fourier transform of the function $f(x)$ (up to a multiplier).
The Fourier transform is an operation that transforms one complex-valued function of a real variable into another:

$$\mathcal{F}_x : f(x) \to \tilde{f}(k), \qquad \text{with} \qquad \tilde{f}(k) = \mathcal{F}_x[f(x)](k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$$

Notice that sometimes different normalizations are used. With this definition, we also find that the inverse Fourier transform is given by:

$$\mathcal{F}_k^{-1} : \tilde{f}(k) \to f(x), \qquad f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk$$

2.2.2 Review of linear Algebra

This is a very concise review of concepts in linear algebra, reintroducing some of the ideas we saw in the previous paragraphs in a slightly more formal way.

A. Vectors and vector spaces

Quantum mechanics is a linear theory, thus it is well described by vectors and vector spaces. Vectors are mathematical objects (distinct from scalars) that can be added one to another and multiplied by a scalar. In QM we denote vectors by the Dirac notation: $|\psi\rangle$, $|\varphi\rangle$, ... Then, these have the properties:

− If $|\psi_1\rangle$ and $|\psi_2\rangle$ are vectors, then $|\psi_3\rangle = |\psi_1\rangle + |\psi_2\rangle$ is also a vector.

– Given a scalar s, $|\psi_4\rangle = s |\psi_1\rangle$ is also a vector.

A vector space is a collection of vectors. For example, for vectors of finite dimensions, we can define a vector space of dimensions N over the complex numbers as the collection of all complex-valued N-dimensional vectors.

Example 1 – A familiar example of vectors and vector space are the Euclidean vectors and the real 3D space.

Example 2 – Another example of a vector space is the space of polynomials of order n. Its elements, the polynomials $P_n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ can be proved to be vectors since they can be summed to obtain another polynomial and multiplied by a scalar. The dimension of this vector space is $n + 1$.

Example 3 – In general, functions can be considered vectors of a vector space with infinite dimension (of course, if we restrict the set of functions that belong to a given space, we must ensure that this is still a well-defined vector space. For example, the collection of all function $f(x)$ bounded by 3 $[f(x) < 3, \forall x]$ is not a well defined vector-space, since $sf(x)$ (with s a scalar > 1) is not a vector in the space.

B. Inner product

We denote by $\langle \psi | \varphi \rangle$ the scalar product between the two vectors $|\psi\rangle$ and $|\varphi\rangle$. The inner product or scalar product is a mapping from two vectors to a complex scalar, with the following properties:

– It is linear in the second argument: $\langle \psi | a_1 \varphi_1 + a_2 \varphi_2 \rangle = a_1 \langle \psi | \varphi_1 \rangle + a_2 \langle \psi | \varphi_2 \rangle$.

– It has the property of complex conjugation: $\langle \psi | \varphi \rangle = \langle \varphi | \psi \rangle^*$.

– It is positive-definite: $\langle \psi | \psi \rangle = 0 \Leftrightarrow |\psi\rangle = 0$.

Example 1 – For Euclidean vectors the inner product is the usual scalar product $\vec{v}_1 \cdot \vec{v}_2 = |\vec{v}_1||\vec{v}_2| \cos \vartheta$.

Example 2 – For functions, the inner product is defined as:

$$\langle f | g \rangle = \int_{-\infty}^{\infty} f(x)^* g(x) dx$$

C. Linearly independent vectors (and functions)

We can define linear combinations of vectors as $|\psi\rangle = a_1 |\varphi_1\rangle + a_2 |\varphi_2\rangle + \dots$. If a vector cannot be expressed as a linear superposition of a set of vectors, than it is said to be linearly independent from these vectors. In mathematical terms, if

$$|\xi\rangle \neq \sum_i a_i |\varphi_i\rangle, \quad \forall a_i$$

then $|\xi\rangle$ is linearly independent of the vectors $\{|\varphi_i\rangle\}$.

D. Basis

A basis is a linearly independent set of vectors that spans the space. The number of vectors in the basis is the vector space dimension. Any other vector can be expressed as a linear combination of the basis vectors. The basis is not unique, and we will usually choose an orthonormal basis.

Examples – For the Polynomial vector space, a basis are the monomials $\{x^k\}$, $k = 0, \dots, n$. For Euclidean vectors the vectors along the 3 coordinate axes form a basis.

We have seen in class that eigenvectors of operators form a basis.

E. Unitary and Hermitian operators

An important class of operators are self adjoint or Hermitian operators, as observables are described by them. We need first to define the adjoint of an operator A. This is denoted $A^\dagger$ and it is defined by the relation:

$$\langle (A^\dagger \psi) | \varphi \rangle = \langle \phi | (A\psi) \rangle \qquad \forall \{ |\psi\rangle, |\varphi\rangle \}.$$

This condition can also be written (by using the second property of the inner product) as:

$$\langle \psi | A^\dagger | \varphi \rangle = \langle \varphi | A | \psi \rangle^*$$

If the operator is represented by a matrix, the adjoint of an operator is the conjugate transpose of that operator: $A^\dagger_{k,j} = \langle k | A^\dagger | j \rangle = \langle j | A | k \rangle^* = A^*_{j,k}$.

$\mathcal{D}$: **Self-adjoint** . A self adjoint operator is an operator such that $A^\dagger = A$, or more precisely

$$\langle \psi | A | \varphi \rangle = \langle \varphi | A | \psi \rangle^*$$

For matrix operators, $A_{ki} = A_{ik}^*$.

An important properties of Hermitian operators is that their eigenvalues are always real (even if the operators are defined on the complex numbers). Then, all the observables must be represented by hermitian operators, since we want their eigenvalues to be real, as the eigenvalues are nothing else than possible outcomes of experiments (and we wouldn't want the position of a particle, for example, to be a complex number).

Then, for example, the Hamiltonian of any system is an hermitian operator. For a particle in a potential, it's easy to check that the operator is real, thus it is also hermitian.

$\mathcal{D}$: **Unitary operators** U are such that their inverse is equal to their adjoint: $U^{-1} = U^\dagger$, or

$$UU^\dagger = U^\dagger U = \mathbb{1}.$$

We will see that the evolution of a system is given by a unitary operator, which implies that the evolution is time-reversible.

2.3 Measurement and probability

From the second postulate we have seen that the possible outcomes of a measurement are the eigenvalues of the operator corresponding to the measured observable. The question of which one of the eigenvalue we will obtain is still open. This question is resolved by the third postulate, which gives a recipe on how to predict which outcome will be observed. However, this mathematical recipe does not tell us (in general) with absolute certainty which one of the eigenvalue will be the outcome. It only provides us with the probability of obtaining each eigenvalue.

In postulate 3.b we consider an observable $\mathcal{A}$ and a system in the state $|\psi\rangle$. The eigenvalue equation for the operator A (corresponding to $\mathcal{A}$) can be expressed as

$$A |n\rangle = a_n |n\rangle$$

where a_n are the eigenvalues and $|n\rangle$ the eigenvectors . The postulate states that the probability of obtaining a_n as the outcome of the measurement of A is $p(a_n) = |\langle n|\psi\rangle|^2$.

We want to re-express the postulate in terms of the wavefunction $\psi(\vec{x})$. To do so, we need to define the inner product in the Hilbert space of the wavefunctions. Given two wave functions $\psi(\vec{x})$ and $\varphi(\vec{x})$, the inner product $\langle \varphi|\psi\rangle$ is given by:

$$\langle \varphi|\psi\rangle = \int d^3\vec{x}\, \varphi(x)^* \psi(x)$$

(where $*$ indicates the complex conjugate).

We first rewrite the eigenvalue problem for the operator A in terms of the eigenfunctions $u_n(\vec{x})$ and the associated eigenvalues a_n:

$$A[u_n(\vec{x})] = a_n u_n(\vec{x})$$

Then, we have seen that any function can be expressed in terms of the eigenfunctions $u_n(\vec{x})$. We can as well express the wavefunction in terms of these eigenfunctions:

$$\psi(\vec{x}) = \sum_n c_n u_n(\vec{x}), \quad \text{with} \quad c_n = \int d^3\vec{x}\, u_n^*(\vec{x})\psi(\vec{x})$$

Finally, according to postulate 3.b the probability of obtaining the outcome a_n if the system is in the state $\psi(\vec{x})$ is given by the inner product :

$$\boxed{p(a_n) = \left| \int d^3\vec{x}\, u_n^*(\vec{x})\psi(\vec{x}) \right|^2 = |c_n|^2}$$

where the last equality follows from the orthogonality of the eigenfunctions $\int d^3\vec{x}\, u_n^*(\vec{x})u_m(\vec{x}) = 0$, for $m \neq n$.

Since $|c_n|^2 = p(a_n)$ is a probability, the coefficients c_n of the wavefunction expansion are called *probability amplitudes*. We now confirm that the wavefunction contain all information about the state of the system, since given the wavefunction we can calculate all the probabilities of each outcome for each possible observable with the following procedure:

1) Find the eigenfunctions of the observable's operator. E.g., given an operator O, we will calculate its eigenfunctions $w_n(x)$, such that $O[w_n(x)] = o_n w_n(x)$.

2) Find the probability amplitude of the wavefunction with respect to the eigenfunction of the desired eigenvalue outcome.
E.g., if the outcome is o_m, such that $O[w_m(x)] = o_m w_m(x)$ we will calculate $c_m = \int d^3\vec{x}\, w_m^*(\vec{x})\psi(\vec{x})$.

3) The probability of obtaining the given eigenvalue in the measurement is the probability amplitude modulus square. E.g. $p(o_m) = |c_m|^2$.

2.3.1 Wavefunction collapse

The third postulate states also that after the measurement the system is left in the eigenstate corresponding to the eigenvalue found (more generally, if more than one eigenstate is associated to the same eigenvalue, the state is projected on the subspace of the eigenvalue a_n, that is, the subspace spanned by all the eigenstates associated with a_n).
This is the *wavefunction collapse*, a concept that is usually quite puzzling in quantum mechanics. We can make this statement at least a bit less puzzling by taking into account the following two considerations.
The wavefunction collapse is puzzling because it predicts an instantaneous evolution of the system from its pre-measurement state $\psi(x)$ to its post-measurement state $u_n(x)$ (when we measure a_n). This type of evolution is very different than the usual evolution predicted by the fourth postulate (that we will see in a later lecture). However, this weird behavior arises from considering the measurement apparatus (and hence the measurement) as a classical system, outside the realm of quantum mechanics. Although this view gives most of the time a correct answer – and thus we will use it in this class – it is a quite imprecise description. More rigorous descriptions of the measurement process, invoking for example decoherence[4], can give a better picture of what actually happens (e.g. the wave-function collapse can take a finite time and be measured experimentally in some cases).
More pragmatically, the wavefunction collapse is needed in order to make experiment consistent. What the collapse entails is that if I make a measurement and I obtain as an outcome the eigenvalue a_n, I can check that result again, by repeating the measurement just after the first one (with no time for any change in the system between the two measurement). If I could not make this second check, I could never be able to be confident that I got the correct answer the first time (e.g. my detector could be wrong) and so I could never gain any knowledge at all on my system.
Obs.: I want to clarify the meaning of "subspace of the eigenvalue a_n". If there is a set of eigenstates associated with the eigenvalue a_n, $|n_j\rangle$, then the state $|\psi\rangle$ is projected onto a superposition of these eigenstates $|\psi\rangle \to |n\rangle = \sum_j c_j |n_j\rangle$.

2.3.2 Position measurement

We have already calculated the eigenvalues and eigenfunctions of the position operator. The eigenfunctions were $u_n(x) = \delta(x - x_n)$ with eigenvalues x_n. We also calculated the expansion of a function in terms of the position eigenfunctions. Repeating the calculation for the wavefunction we find:

$$c(x_n) = \int dx\, \delta(x - x_n)\psi(x) = \psi(x_n),$$

from which we obtain that the probability of finding a particle in the position x_n is given by:

$$p(x_n) = |\psi(x_n)|^2$$

More generally, since x is continuous, we can drop the subscript n, as any value of x is an eigenvalue. Then, generalizing to the 3D case, we can say that the probability of finding a particle described by the wavefunction $\psi(\vec{x})$ at the position $\vec{x}$ is given by the modulus square of the wavefunction itself:

$$\boxed{p(\vec{x}) = |\psi(\vec{x})|^2}$$

We can also say that the wavefunction is the probability amplitude for the position measurement. More precisely, we should say that the probability of finding a particle between x and $x + dx$ is $p(x)dx = |\psi(x)|^2 dx$ while $|\psi(x)|^2$ is a probability density (per unit length). In 3D, $|\psi(\vec{x})|^2$ is the probability density per unit volume and the probability is given by $|\psi(\vec{x})|^2 d^3x$.

---

[4] Decoherence is the phenomenon by which an *open* quantum system, interacting with the environment, undergoes an irreversible evolution that often leaves it in a state best described by the rules of classical mechanics.

Given this interpretation of the wavefunction, it becomes natural to require that the wavefunction be normalized. We require that integrating the probability of a particular position over all possible position we obtain 1 (i.e. certainty, the particle **has** to be somewhere!). Then

$$\int d^3\vec{x}\, p(\vec{x}) = 1 \quad \rightarrow \quad \int d^3\vec{x}\, |\psi(\vec{x})|^2 = 1$$

From being a very abstract notion, the wavefunction has thus assumed a very physical meaning. We can still say that the wavefunction describes the state of a system and contains all the information about this system. In addition, and more concretely, the absolute value of the wavefunction tells us where it is more probable to find the system.

2.3.3 Momentum measurement

We calculated the eigenvalues and eigenfunctions of the momentum operator to be

$$p = \hbar k \qquad \text{and} \qquad u_k(x) = \Upsilon e^{ikx}$$

(Notice that we could have labeled the wave-numbers as k_n —and the momentum p_n— to have the eigenvalue equation: $\hat{p}u_n = p_n u_n$, but we omitted the subscript n since momentum is a continuous variable; then we also simply label the eigenfunctions by k instead of n).
As usual, we would like the eigenfunctions to be normalized. However notice that $\int u_k^*(x)u_k(x)dx = \int |\Upsilon|^2 dx = \infty$, so we cannot fix Υ such that the result is normalized as usual. For convention we set $\Upsilon = \frac{1}{\sqrt{2\pi}}$: $u_k(x) = \frac{1}{\sqrt{2\pi}}e^{ikx}$
(we are considering the 1D case). Now we can calculate the probability amplitude for the momentum measurement, by calculating the coefficients of the expansion of the wavefunction in terms of the momentum eigenfunctions basis. Here we rename the coefficients $c(k)$ of the expansion $\varphi(k)$. This is given by:

$$c(k) \equiv \varphi(k) = \int u_k(x)^*\psi(x) \quad \rightarrow \quad \varphi(k) = \frac{1}{\sqrt{2\pi}}\int e^{-ikx}\psi(x)$$

Notice that this last equation is simply stating that the probability amplitude for the momentum measurement is the Fourier transform of the wavefunction, $\varphi(k) = \mathcal{F}[\psi(x)]$. Then

$$p(k \rightarrow k + dk) = |\varphi(k)|^2 dk$$

is the probability of finding that the particle has a momentum between $\hbar k$ and $\hbar(k + dk)$ when measuring the momentum.

A. Flux of particles

The choice of the coefficient Υ implies that:

$$\int u_k^*(x)u_{k'}(x)dx = \delta(k - k')$$

which is not the usual normalization for the wavefunction.
Why is it not possible to normalize the momentum eigenstates?
We saw that for the wavefunction, the normalization was related to its interpretation as the probability amplitude for the position. If a wavefunction is in a momentum eigenstate $\psi(x) = u_k(x)$, then we cannot really talk about a particle, but rather the system is better described by a wave. In fact, the probability of finding the system at any position x is constant and equal to Υ. Thus the coefficient Υ can be better linked to a flux of particles rather than a particle density.
We can set $v|\psi|^2 = \Gamma$ where $v = \frac{p}{m} = \frac{\hbar k}{m}$ is the velocity and Γ correspond to a flux of particle, as described by the plane wave e^{ikx}. Then $\frac{\hbar k}{m}|\Upsilon|^2 = \Gamma$ fixes the value of Υ to $\Upsilon = \sqrt{\frac{m\Gamma}{\hbar k}}$.

2.3.4 Expectation values

We have just seen that the outcome of a measurement is a random quantity (although it is of course limited to a given set of values –the eigenvalues– from which we can choose from). In order to know more about the state of a system we need then to repeat the measurement several times, in order to build a statistics of the observable.

For example, we could be interested in knowing what is the average of the measurements of a particular observable. This quantity is usually called in QM the **expectation value** of an observable.

How do we usually calculate the average of a given quantity? Consider for example the average number obtained by throwing a dice. In an experiment, I would have to repeatedly throw the dice, record the number that comes out (say n_i) and then calculate the sum: $\langle n \rangle = \frac{1}{N} \sum_{i=1}^{N} n_i$. Equivalently, I could count the number of times t_n that each number n appears and calculate the average number from the frequencies $\nu_n = t_n/N$: $\langle n \rangle = \sum_{n=1}^{6} \nu_n n$. In the limit of $N \to \infty$, the frequencies ν_n approach the probabilities $\nu_n \to p_n$. Then for the dice we have $p_n = 1/6$ ($\forall n$) and the average is just calculates from the sum $\frac{1}{6}(1 + 2 + \ldots 6) = 3.5$.

The procedure for calculating the average (or expectation value) of a quantity is very general. We have for discrete and continuous probability distribution functions respectively

$$\langle x \rangle = \sum_i p_i x_i \qquad \langle x \rangle = \int dx p(x) x$$

In QM we just need to replace the probability p by its value as given by the third postulate. For example, the expectation value of the position can be expressed in terms of the probability density function given by the modulus square of the wavefunction:

$$\boxed{\langle x \rangle = \int_{-\infty}^{\infty} x |\psi(x,t)|^2 dx}$$

How can we in practice obtain this expectation value? (that is, by performing real experiments). If we make a first measurement on a single particle and then repeat the measurement over and over again this is not what we measure. In fact, in that case we know (from the postulates) that after the first measurement we expect always to get the same result. In fact, these repeated successive measurement are only a way to check that yes, we got the first answer correct. (otherwise we could never be certain of anything, since we would not even know that our experimental apparatus works).

Instead what we can do is adopt one of two strategies. Either we can repeat the same *experiment* (not measurement) many times on the same system. This implies first preparing the system in a given state with a reproducible procedure and then performing a measurement. The second option is to make the experiment on a set (an *ensemble*) of identical systems. In order to obtain the exact expectation value we would need an infinite number (of repetitions or systems), however a large enough sample is usually practically enough.

Obs.: Notice that we can rewrite the expression above for the expectation value of the position as

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi(x,t)^* x \, \psi(x,t) dx = \int_{-\infty}^{\infty} \psi(x,t)^* \hat{x} [\psi(x,t)] dx$$

This last form is a more general one that is valid for any operator.

$\mathcal{D}$: **Expectation value** The expectation value of an observable $\hat{\mathcal{O}}$ is

$$\boxed{\left\langle \hat{\mathcal{O}} \right\rangle = \langle \psi | \hat{\mathcal{O}} | \psi \rangle = \int_{-\infty}^{\infty} \psi(x,t)^* \hat{\mathcal{O}} [\psi(x,t)] dx}$$

where we first used the Dirac notation to express the expectation value, which is an even more general expression.

Example – The expectation value for the momentum operator is given by:

$$\int_{-\infty}^{\infty} \psi(x,t)^* \hat{p} [\psi(x,t)] dx = \int_{-\infty}^{\infty} \psi(x,t)^* \left(-i\hbar \frac{\partial \psi(x,t)}{\partial x} \right) dx$$

2.4 Energy eigenvalue problem

The energy operator is called Hamiltonian. The first postulate stated that the time dependence of the wavefunction is dictated by the Schrödinger equation:

$$i\hbar \frac{\partial \psi(\vec{x}, t)}{\partial t} = \mathcal{H}\psi(\vec{x}, t)$$

If we assume that $\psi(\vec{x}, t)$ is the product of a time-dependent part $T(t)$ and a time-independent one $\varphi(\vec{x})$, we can attempt to solve the equation by separation of variables. From $\psi(\vec{x}, t) = T(t)\varphi(\vec{x})$, we can rewrite the Schrödinger equation (using the fact that $\mathcal{H}$ does not change $T(t)$):

$$i\hbar \frac{\partial T(t)\varphi(\vec{x})}{\partial t} = \mathcal{H}\left[T(t)\varphi(\vec{x}, t)\right] \quad \rightarrow \quad \varphi(\vec{x}) \cdot i\hbar \frac{\partial T(t)}{\partial t} = T(t) \cdot \mathcal{H}\left[\varphi(\vec{x}, t)\right]$$

and we rearrange terms based on their dependence on t or $\vec{x}$

$$\frac{1}{T(t)} i\hbar \frac{\partial T(t)}{\partial t} = \frac{1}{\varphi(\vec{x})} \mathcal{H}\left[\varphi(\vec{x}, t)\right]$$

Each side has to be equal to a constant, in order for the equality to hold. Then the time-independent wavefunction obeys the time-independent Schrödinger equation:

$$\boxed{\mathcal{H}\varphi(\vec{x}) = E\varphi(\vec{x})}$$

where E is identified as the energy of the system. If the wavefunction is given by just its time-independent part, $\psi(\vec{x}, t) = \varphi(\vec{x})$, the state is *stationary*. Thus, the time-independent Schrödinger equation allows us to find stationary states of the system, given a certain Hamiltonian.

Notice that the time-independent Schrödinger equation is nothing else than the eigenvalue equation for the Hamiltonian operator. It is thus particularly interesting to study eigenvalues and eigenfunctions of this operator (which, as said, correspond to the energies and stationary states of the system) [5].

In general, the wavefunction describing the state of a quantum system is not

The energy of a particle has contributions from the kinetic energy as well as the potential energy:

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z)$$

In quantum mechanics we can find the equivalent operator by substituting the quantum operators for the position and momentum in the above expression:

$$\mathcal{H} = \frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V(\hat{x}, \hat{y}, \hat{z})$$

or more explicitly:

$$\mathcal{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x, y, z)$$

and in a compact form

$$\boxed{\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(x, y, z)}$$

(Notice that $V(x, y, z)$ is just a multiplicative operator, in the same way as the position is).

---

[5] I want to clarify the distinction between eigenfunctions and wavefunctions. In this class we are interested in both. The eigenfunctions are related to a given operator, and they are the solutions to the eigenvalue equation for that operator. They are important since they form a basis and they allow us to calculate the probability of obtaining a given measurement outcome. The wavefunction describes the state of the quantum system. In general, it is not an eigenfunction. However, *if* we are considering a *stationary* state, the wavefunction that represents it must be an eigenfunction of the Hamiltonian (energy) operator. Thus in that particular case only (which is a quite common case!) the wavefunction is also an eigenfunction.

2.4.1 Free particle

In 1D, for a free particle there is no potential energy, but only kinetic energy that we can rewrite as:

$$\mathcal{H} = \frac{1}{2m}p^2 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

The eigenvalue problem $\mathcal{H}w_n(x) = E_n w_n(x)$ is then the differential equation

$$\mathcal{H}w_n(x) = E_n w_n(x) \quad \rightarrow \quad -\frac{\hbar^2}{2m}\frac{\partial^2 w_n(x)}{\partial x^2} = E_n w_n(x)$$

we rewrite it as:

$$\frac{\partial^2 w_n(x)}{\partial x^2} + \frac{2mE_n}{\hbar^2}w_n = 0 \quad \rightarrow \quad \frac{\partial^2 w_n(x)}{\partial x^2} + k_n^2 w_n = 0$$

where we used the identity

$$\boxed{\frac{\hbar^2 k_n^2}{2m} = E_n}$$

between the kinetic energy eigenvalue E_n and the wavenumber k_n (and the momentum $p_n = \hbar k_n$).
For a free particle there is no restriction on the possible energies, E_n can be any positive number. The solution to the eigenvalue problem is then the eigenfunction:

$$w_n(x) = A\sin(k_n x) + B\cos(k_n x) = A'e^{ik_n x} + B'e^{-ik_n x}$$

We see that there are two independent functions for each eigenvalue E_n. Also there are two distinct momentum eigenvalues $\pm k_n$ for each energy eigenvalue, which correspond to two different directions of propagation of the wave function $e^{\pm ik_n x}$.

2.5 Operators, Commutators and Uncertainty Principle

2.5.1 Commutator

$\mathcal{D}$: **Commutator**   The Commutator of two operators A, B is the operator $C = [A, B]$ such that $C = AB - BA$.
Example 1 – If the operators A and B are scalar operators (such as the position operators) then $AB = BA$ and the commutator is always zero.
Example 2 – If the operators A and B are matrices, then in general $AB \neq BA$. Consider for example:

$$A = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then

$$AB = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad BA = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

Then $[A, B] = 2AB$.
Example 3 – A is *Turn to your right*. B is *Take 3 steps to your left*.
? Question: do these two operators commute?

Fig. 10: Two rotations A, B along the x-axis. Left: we apply AB (first the $3\pi/4$ rotation), right we apply BA. Since the two operators commute, the result is the same.

Example 4 – Let A and B be two rotations. First assume that A is a $\pi/4$ rotation around the x direction and B a $3\pi/4$ rotation in the same direction. Now assume that the vector to be rotated is initially around z. Then, if we apply AB (that means, first a $3\pi/4$ rotation around x and then a $\pi/4$ rotation), the vector ends up in the negative z direction. The same happen if we apply BA (first A and then B).

Fig. 11: Two rotations A, B along the x- and z-axis. Left: we apply AB (first the $\pi/2$ rotation along z), right: we apply BA. Since the two operators do not commute, the result is not the same.

Now assume that A is a $\pi/2$ rotation around the x direction and B around the z direction. When we apply AB, the vector ends up (from the z direction) along the y-axis (since the first rotation does not do anything to it), if instead we apply BA the vector is aligned along the x direction. In this case the two rotations along different axes do not commute.
These examples show that commutators are not specific of quantum mechanics but can be found in everyday life. We now want an example for QM operators.

The most famous commutation relationship is between the position and momentum operators. Consider first the 1D case. We want to know what is $[\hat{x}, \hat{p}_x]$ (I'll omit the subscript on the momentum). We said this is an operator, so in order to know what it is, we apply it to a function (a wavefunction). Let's call this operator C_{xp}, $C_{xp} = [\hat{x}, \hat{p}_x]$.

$$[\hat{x}, \hat{p}]\psi(x) = C_{xp}[\psi(x)] = \hat{x}\left[\hat{p}[\psi(x)]\right] - \hat{p}\left[\hat{x}[\psi(x)]\right] = -i\hbar\left(x\frac{d}{dx} - \frac{d}{dx}x\right)\psi(x)$$

$$-i\hbar\left(x\frac{d\psi(x)}{dx} - \frac{d}{dx}(x\psi(x))\right) = -i\hbar\left(x\frac{d\psi(x)}{dx} - \psi(x) - x\frac{d\psi(x)}{dx}\right) = i\hbar\psi(x)$$

From $[\hat{x}, \hat{p}]\psi(x) = i\hbar\psi(x)$ which is valid for all $\psi(x)$ we can write

$$\boxed{[\hat{x}, \hat{p}] = i\hbar}$$

Considering now the 3D case, we write the position components as $\{r_x, r_y r_z\}$. Then we have the commutator relationships:

$$\boxed{[\hat{r}_a, \hat{p}_b] = i\hbar\delta_{a,b}}$$

that is, vector components in different directions commute (the commutator is zero).

A. Properties of commutators

- Any operator commutes with scalars $[A, a] = 0$
- $[A, BC] = [A, B]C + B[A, C]$ and $[AB, C] = A[B, C] + [A, C]B$
- Any operator commutes with itself $[A, A] = 0$, with any power of itself $[A, A^n] = 0$ and with any function of itself $[A, f(A)] = 0$ (from previous property and with power expansion of any function).

From these properties, we have that the Hamiltonian of the free particle commutes with the momentum: $[p, \mathcal{H}] = 0$ since for the free particle $\mathcal{H} = p^2/2m$. Also, $[x, p^2] = [x, p]p + p[x, p] = 2i\hbar p$.

We now prove an important theorem that will have consequences on how we can describe states of a systems, by measuring different observables, as well as how much information we can extract about the expectation values of different observables.

■ Theorem: If A and B commute, then they have a set of non-trivial common eigenfunctions.
Proof: Let φ_a be an eigenfunction of A with eigenvalue a:

$$A\varphi_a = a\varphi_a$$

Then

$$BA\varphi_a = aB\varphi_a$$

But since $[A, B] = 0$ we have $BA = AB$. Let's substitute in the LHS:

$$A(B\varphi_a) = a(B\varphi_a)$$

This means that $(B\varphi_a)$ is also an eigenfunction of A with the same eigenvalue a. If φ_a is the only linearly independent eigenfunction of A for the eigenvalue a, then $B\varphi_a$ is equal to φ_a at most up to a multiplicative constant: $B\varphi_a \propto \varphi_a$. That is, we can write

$$B\varphi_a = b_a\varphi_a$$

But this equation is nothing else than an eigenvalue equation for B. Then φ_a is also an eigenfunction of B with eigenvalue b_a. We thus proved that φ_a is a common eigenfunction for the two operators A and B. $\square$

Example – We have just seen that the momentum operator commutes with the Hamiltonian of a free particle. Then the two operators should share common eigenfunctions.
This is indeed the case, as we can verify. Consider the eigenfunctions for the momentum operator:

$$\hat{p}[\psi_k] = \hbar k\psi_k \quad \rightarrow \quad -i\hbar\frac{d\psi_k}{dx} = \hbar k\psi_k \quad \rightarrow \quad \psi_k = Ae^{-ikx}$$

What is the Hamiltonian applied to ψ_k?

$$\mathcal{H}[\psi_k] = -\frac{\hbar^2}{2m}\frac{d^2(Ae^{-ikx})}{dx^2} = \frac{\hbar^2 k^2}{2m}Ae^{-ikx} = E_k\psi_k$$

thus we found that ψ_k is also a solution of the eigenvalue equation for the Hamiltonian, which is to say that it is also an eigenfunction for the Hamiltonian.

2.5.2 Commuting observables

A. Degeneracy

In the proof of the theorem about commuting observables and common eigenfunctions we took a special case, in which we assume that the eigenvalue a was *non-degenerate*. That is, we stated that φ_a was the only linearly independent eigenfunction of A for the eigenvalue a (functions such as $4\varphi_a$, $\alpha\varphi_a$ don't count, since they are not linearly independent from φ_a).

$\mathcal{D}$: **Degeneracy** In general, an eigenvalue is degenerate if there is more than one eigenfunction that has the same eigenvalue. The degeneracy of an eigenvalue is the number of eigenfunctions that share that eigenvalue.
For example a is n-degenerate if there are n eigenfunction $\{\varphi_j^a\}$, $j = 1, 2, \ldots, n$, such that $A\varphi_j^a = a\varphi_j^a$.
What happens if we relax the assumption that the eigenvalue a is not degenerate in the theorem above? Consider for example that there are two eigenfunctions associated with the same eigenvalue:

$$A\varphi_1^a = a\varphi_1^a \quad \text{and} \quad A\varphi_2^a = a\varphi_2^a$$

then any linear combination $\varphi^a = c_1\varphi_1^a + c_2\varphi_2^a$ is also an eigenfunction with the same eigenvalue (there's an infinity of such eigenfunctions). From the equality $A(B\varphi^a) = a(B\varphi^a)$ we can still state that $(B\varphi^a)$ is an eigenfunction of A but we don't know which one. Most generally, there exist $\tilde{c}_1$ and $\tilde{c}_2$ such that

$$B\varphi_1^a = \tilde{c}_1\varphi_1^a + \tilde{c}_2\varphi_2^a$$

but in general $B\varphi_1^a \not\propto \varphi_1^a$, or φ_1^a is not an eigenfunction of B too.
Example – Consider again the energy eigenfunctions of the free particle. To each energy $E = \frac{\hbar^2 k^2}{2m}$ are associated two linearly-independent eigenfunctions (the eigenvalue is doubly degenerate). We can choose for example $\varphi_E = e^{ikx}$ and $\varphi_E = e^{-ikx}$. Notice that these are also eigenfunctions of the momentum operator (with eigenvalues $\pm k$). If we had chosen instead as the eigenfunctions $\cos(kx)$ and $\sin(kx)$ these are not eigenfunctions of $\hat{p}$.

■ Theorem: In general, it is always possible to choose a set of (linearly independent) eigenfunctions of A for the eigenvalue a such that they are also eigenfunctions of B.
For the momentum/Hamiltonian for example we have to choose the exponential functions instead of the trigonometric functions. Also, if the eigenvalue of A is degenerate, it is possible to label its corresponding eigenfunctions by the eigenvalue of B, thus lifting the degeneracy. For example, there are two eigenfunctions associated with the energy E: $\varphi_E = e^{\pm ikx}$. We can distinguish between them by labeling them with their momentum eigenvalue $\pm k$: $\varphi_{E,+k} = e^{ikx}$ and $\varphi_{E,-k} = e^{-ikx}$.

Proof: Assume now we have an eigenvalue a with an n-fold degeneracy such that there exists n independent eigenfunctions φ_k^a, $k = 1, \ldots, n$. Any linear combination of these functions is also an eigenfunction $\tilde{\varphi}^a = \sum_{k=1}^{n} \tilde{c}_k \varphi_k^a$. For any of these eigenfunctions (let's take the hth one) we can write:

$$B[A[\varphi_h^a]] = A[B[\varphi_h^a]] = aB[\varphi_h^a]$$

so that $\bar{\varphi}_h^a = B[\varphi_h^a]$ is an eigenfunction of A with eigenvalue a. Then this function can be written in terms of the $\{\varphi_k^a\}$:

$$B[\varphi_h^a] = \bar{\varphi}_h^a = \sum_k \bar{c}_{h,k} \varphi_k^a$$

This notation makes it clear that $\bar{c}_{h,k}$ is a tensor (an $n \times n$ matrix) operating a transformation from a set of eigenfunctions of A (chosen arbitrarily) to another set of eigenfunctions. We can write an eigenvalue equation also for this tensor,

$$\bar{c}v^j = b^j v^j \quad \rightarrow \quad \sum_h \bar{c}_{h,k} v_h^j = b^j v^j$$

where the eigenvectors v^j are vectors of length n.
If we now define the functions $\psi_j^a = \sum_h v_h^j \varphi_h^a$, we have that ψ_j^a are of course eigenfunctions of A with eigenvalue a. Also

$$B[\psi_j^a] = \sum_h v_h^j B[\varphi_h^a] = \sum_h v_h^j \sum_{k=1}^{n} \bar{c}_{h,k} \varphi_k^a$$

$$= \sum_k \varphi_k^a \sum_h \bar{c}_{h,k} v_h^j = \sum_k \varphi_k^a b^j v_k^j = b^j \sum_k v_k^j \varphi_k^a = b^j \psi_j^a$$

We have thus proved that ψ_j^a are eigenfunctions of B with eigenvalues b^j. The ψ_j^a are simultaneous eigenfunctions of both A and B.

32

Consider the set of functions $\{\psi_j^a\}$. From the point of view of A they are not distinguishable, they all have the same eigenvalue so they are degenerate. Taking into account a second operator B, we can lift their degeneracy by labeling them with the index j corresponding to the eigenvalue of B (b^j).

Example – Assume that we choose $\varphi_1 = \sin(kx)$ and $\varphi_2 = \cos(kx)$ as the degenerate eigenfunctions of $\mathcal{H}$ with the same eigenvalue $E_k = \frac{\hbar^2 k^2}{2m}$. We now want to find with this method the common eigenfunctions of $\hat{p}$. We first need to find the matrix $\bar{c}$ (here a 2×2 matrix), by applying $\hat{p}$ to the eigenfunctions.

$$\hat{p}\varphi_1 = -i\hbar \frac{d\varphi_1}{dx} = i\hbar k \cos(kx) = -i\hbar k \varphi_2$$

and $\hat{p}\varphi_2 = i\hbar k \varphi_1$. Then the matrix $\bar{c}$ is:

$$\bar{c} = \begin{pmatrix} 0 & i\hbar k \\ -i\hbar k & 0 \end{pmatrix}$$

with eigenvalues $b^j = \pm\hbar k$, and eigenvectors (not normalized)

$$v^1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}, \quad v^2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

We then write the ψ eigenfunctions:

$$\psi^1 = v_1^1 \varphi_1 + v_2^1 \varphi_2 = -i\sin(kx) + \cos(kx) \propto e^{-ikx}, \quad \psi^2 = v_1^2 \varphi_1 + v_2^2 \varphi_2 = i\sin(kx) + \cos(kx) \propto e^{ikx}$$

B. Complete set of commuting observables

We have seen that if an eigenvalue is degenerate, more than one eigenfunction is associated with it. Then, if we measure the observable A obtaining a we still do not know what the state of the system after the measurement is. If we take another observable B that commutes with A we can measure it and obtain b. We have thus acquired some extra information about the state, since we know that it is now in a common eigenstate of both A and B with the eigenvalues a and b. Still, this could be not enough to fully define the state, if there is more than one state φ_{ab}. We can then look for another observable C, that commutes with both A and B and so on, until we find a set of observables such that upon measuring them and obtaining the eigenvalues a, b, c, d, ... the function $\varphi_{abcd...}$ is uniquely defined. Then the set of operators $\{A, B, C, D, ...\}$ is called a complete set of commuting observables. The eigenvalues $a, b, c, d, ...$ that specify the state are called *good quantum numbers* and the state is written in Dirac notation as $|a\,b\,c\,d...\rangle$.
Obs. The set of commuting observable is not unique.

2.5.3 Uncertainty principle

A. Uncertainty for waves

The uncertainty principle, which you probably already heard of, is not found just in QM. Consider for example the propagation of a wave. If you shake a rope rhythmically, you generate a stationary wave, which is not localized (where is the wave??) but it has a well defined wavelength (and thus a momentum).
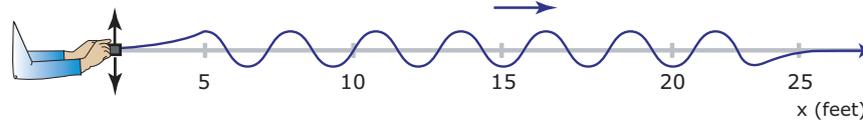
Image by MIT OpenCourseWare.

Fig. 12: A wave with a well defined wavelength but no well-defined position.

If instead you give a sudden jerk, you create a well localized wavepacket. Now however the wavelength is not well defined (since we have a superposition of waves with many wavelengths). The position and wavelength cannot thus be well defined at the same time. In QM we express this fact with an inequality involving position and momentum $p = \frac{2\pi\hbar}{\lambda}$. Then we have $\sigma_x \sigma_p \geq \frac{\hbar}{2}$. We are now going to express these ideas in a more rigorous way.

Fig. 13: A wave packet with a well defined position but no well-defined wavelength. (From Griffith)

B. Repeated measurements

Recall that the third postulate states that after a measurement the wavefunction *collapses* to the eigenfunction of the eigenvalue observed.

Let us assume that I make two measurements of the same operator A one after the other (no evolution, or time to modify the system in between measurements). In the first measurement I obtain the outcome a_k (an eigenvalue of A). Then for QM to be consistent, it must hold that the second measurement also gives me the same answer a_k. How is this possible? We know that if the system is in the state $\psi = \sum_k c_k \varphi_k$, with φ_k the eigenfunction corresponding to the eigenvalue a_k (assume no degeneracy for simplicity), the probability of obtaining a_k is $|c_k|^2$. If I want to impose that $|c_k|^2 = 1$, I must set the wavefunction after the measurement to be $\psi = \varphi_k$ (as all the other c_h, $h \neq k$ are zero). This is the so-called *collapse of the wavefunction*. It is not a mysterious accident, but it is a prescription that ensures that QM (and experimental outcomes) are consistent (thus it's included in one of the postulates).

Now consider the case in which we make two successive measurements of two different operators, A and B. First we measure A and obtain a_k. We now know that the state of the system after the measurement must be φ_k. We now have two possibilities.

If $[A, B] = 0$ (the two operator commute, and again for simplicity we assume no degeneracy) then φ_k is also an eigenfunction of B. Then, when we measure B we obtain the outcome b_k with certainty. There is no uncertainty in the measurement. If I measure A again, I would still obtain a_k. If I inverted the order of the measurements, I would have obtained the same kind of results (the first measurement outcome is always unknown, unless the system is already in an eigenstate of the operators).

This is not so surprising if we consider the classical point of view, where measurements are not probabilistic in nature.

The second scenario is if $[A, B] \neq 0$. Then, φ_k is not an eigenfunction of B but instead can be written in terms of eigenfunctions of B, $\varphi_k = \sum_h c_h^k \psi_h$ (where ψ_h are eigenfunctions of B with eigenvalue b_h). A measurement of B does not have a certain outcome. We would obtain b_h with probability $|c_h^k|^2$.

There is then an intrinsic uncertainty in the successive measurement of two non-commuting observables. Also, the results of successive measurements of A, B and A again, are different if I change the order B, A and B.

It means that if I try to know with certainty the outcome of the first observable (e.g. by preparing it in an eigenfunction) I have an uncertainty in the other observable. We saw that this uncertainty is linked to the commutator of the two observables. This statement can be made more precise.

■ **Theorem**: Define $C = [A, B]$ and ΔA and ΔB the uncertainty in the measurement outcomes of A and B: $\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2$, where $\langle \hat{O} \rangle$ is the expectation value of the operator $\hat{O}$ (that is, the average over the possible outcomes, for a given state: $\langle \hat{O} \rangle = \langle \psi | \hat{O} | \psi \rangle = \sum_k O_k |c_k|^2$).

Then:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle C \rangle|$$

This is **Heisenberg Uncertainty Principle**.

Example – The most important example is the uncertainty relation between position and momentum. We know that these two operators do not commute and their commutator is $[\hat{x}, \hat{p}] = i\hbar$. Then

$$\Delta \hat{x} \Delta \hat{p} \geq \frac{\hbar}{2}$$

34

22.02 Introduction to Applied Nuclear Physics
Spring 2012