

2.2: States, Observables and Eigenvalues

Definition: State vector

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):

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

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.

Definition: Observable

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 spaces² $A : g(I) \rightarrow f(I)$ 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.

Note

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

Definition: Eigenvalues and eigenfunctions

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 $S_\alpha - s1$ equal to zero. We find that the eigenvalues are $\pm \frac{1}{2}$ for both operators. The eigenvectors are different:

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

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

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, $\frac{d[\cdot]}{dx}$. The eigenvalue equation for this operator reads:

$$\frac{df(x)}{dx} = a f(x)$$

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 $x \frac{d[\cdot]}{dx}$ is:

$$x \frac{df(x)}{dx} = a f(x)$$

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):

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

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

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

Let's now turn to quantum mechanical operators.

Definition: Position operator

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

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

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

$$\hat{x}[u_n(x)] = x_n u_n(x) \rightarrow x u_n(x) = x_n u_n(x)$$

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: $u_n(x) = \delta(x - x_n)$

Definition: Dirac Delta function

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).

Definition: 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)

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \hat{p}_z = -i\hbar \frac{\partial}{\partial z}$$

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

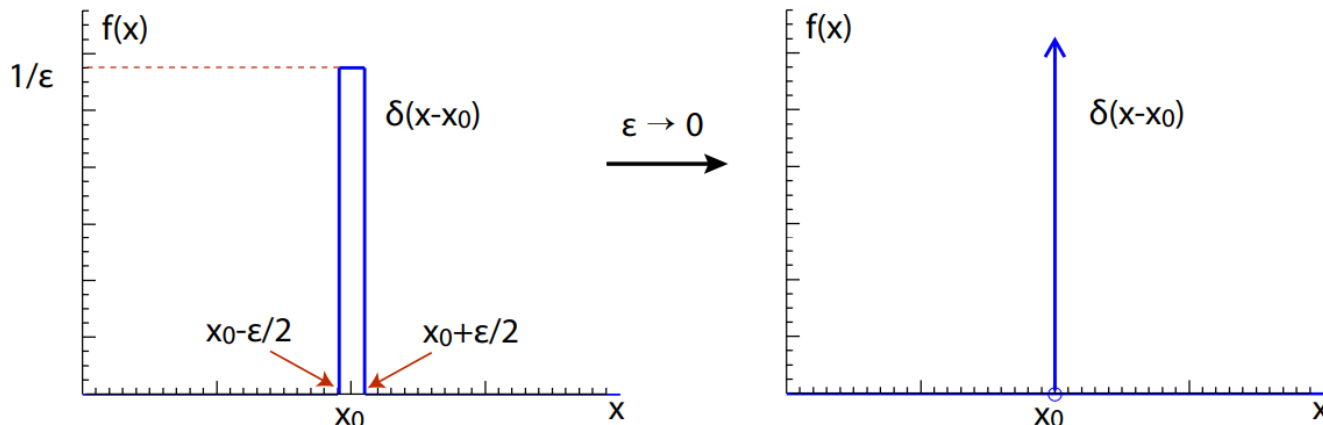


Figure 2.2.1: Schematics of Dirac's delta function. Left: the rectangular function of base ϵ and height $1/\epsilon$ becomes the delta-function (right) in the limit of $\epsilon \rightarrow 0$. (CC BY-NC-ND; Paola Cappellaro)

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

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

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 video³. For classical objects the momentum is very large (since the mass is large), then the wavelength is very small and the object loses 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.

Note

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

Properties of eigenfunctions

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

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 $\langle f | g \rangle = \int d^3x f^*(x)g(x)$. Then we have

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

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 $\{u_n(x)\}$ of an operator A :

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

[Note that the last equality is valid iff 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

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

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:

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

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

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

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.

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:

$|\psi\rangle, |\varphi\rangle, \dots$. 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 A.1

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

✓ Example A.2

Another example of a vector space is the space of polynomials of order n . Its elements, the polynomials $P_n = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ 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 A.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.

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 B.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 B.2

For functions, the inner product is defined as:

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

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, then 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 $\{|\varphi_i\rangle\}$.

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.

✓ Example D.1

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.

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:

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

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:

$$A_{k,j}^\dagger = \langle k | A^\dagger | j \rangle = \langle j | A | k \rangle^* = A_{j,k}^* .$$

Definition: 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.

Definition: Unitary operators

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

$$UU^\dagger = U^\dagger U = 1.$$

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

This page titled [2.2: States, Observables and Eigenvalues](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Paola Cappellaro \(MIT OpenCourseWare\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.