Overview

1   Linear algebra

A **Hilbert space** is a linear space over $\mathbb{C}$ with a (non-degenerate) inner product, and that is complete under the metric induced by this inner product. We do not go into the details of this induced metric or of the completeness.

Vectors in the Hilbert space are usually denoted by $|v\rangle$. The formal way of denoting the inner product between two vectors $|v\rangle$ and $|w\rangle$ is $(|v\rangle, |w\rangle)$. By linearity, $a|v\rangle + b|w\rangle$ is also a vector. The inner product is **anti-linear** on the first vector, and **linear** on the second:

$$(a|v\rangle + b|w\rangle, |u\rangle) = a^*(|v\rangle, |u\rangle) + b^*(|w\rangle, |u\rangle), \quad (|u\rangle, a|v\rangle + b|w\rangle) = a(|u\rangle, |v\rangle) + b(|u\rangle, |w\rangle).$$

The inner product also satisfies

$$(|v\rangle, |w\rangle)^* = (|w\rangle, |v\rangle).$$

Dual vectors are linear maps from the Hilbert space to $\mathbb{C}$. For any vector $|v\rangle$, the dual vector associated to it is denoted $\langle v|$. Its action on a vector $|w\rangle$ is denoted by $\langle v|w\rangle$: this gives a complex number in general. By definition, the action of $\langle v|$ on $|w\rangle$ is given by the inner product between the vectors $|v\rangle$ and $|w\rangle$, that is,

$$\langle v|w\rangle \equiv (|v\rangle, |w\rangle).$$

Dual vectors also form a linear space, so that $a\langle v| + b\langle w|$ is also a dual vector, whose action on $|u\rangle$ is obtained by linearity:

$$(a\langle v| + b\langle w|)|u\rangle = a\langle v|u\rangle + b\langle w|u\rangle.$$

By anti-linearity on the first factor of the inner product, we have that the dual vector associated to the vector $a|v\rangle$ is $a^*\langle v|$.

A **linear operator** on the Hilbert space (or simply an operator) is a linear map from the Hilbert space to itself. The application of the linear operator $\hat{A}$ on the vector $|v\rangle$ is denoted $\hat{A}|v\rangle$: this gives a new vector. The product of two linear operators, $\hat{A}\hat{B}$, means that we first apply $\hat{B}$, then we apply $\hat{A}$. That is, $(\hat{A}\hat{B})|v\rangle = \hat{A}(\hat{B}|v\rangle)$.

An **orthonormal basis** for the Hilbert space is a set of orthogonal vectors $|v_j\rangle$ of length 1, $\langle v_j|v_{j'}\rangle = \delta_{j,j'}$, such that any other vector can be written as a linear combination of these. If $|w\rangle = \sum_j c_j|v_j\rangle$, then by orthonormality, $c_j = \langle v_j|w\rangle$. This can be written in a different form, in the form of the completeness relation, using simply $|w\rangle = \mathbf{1}|w\rangle$ with the identity operator written as

$$\mathbf{1} = \sum_j |v_j\rangle\langle v_j|$$

(both sides are seen as linear maps). If the basis forms a continuum, $|v_s\rangle$ for $s \in \mathbb{R}$, then the normalisation is rather that of a delta-function: $\langle v_s|v_{s'}\rangle = \delta(s - s')$, and the expansion of other vectors is through integrals: $|w\rangle = \int ds\, c_s|v_s\rangle$. The completeness relation is then

$$\mathbf{1} = \int ds\, |v_s\rangle\langle v_s|,$$

and we have $c_s = \langle v_s|w\rangle$.

An **eigenvector** of a linear operator $\hat{A}$ is a vector $|v\rangle$ such that $\hat{A}|v\rangle = \lambda|v\rangle$ for some number λ. This number is the **eigenvalue** associated to the vector $|v\rangle$. For a given eigenvalue λ, there may be many eigenvectors, $|v_{\lambda,i}\rangle$ for i in some set. Any linear combinations of these is still an eigenvector with eigenvalue λ, so these eigenvectors form a linear space by themselves, the eigenspace associated to λ. We may choose i to parametrise an orthonormal basis in this space.

The **Hermitian conjugate** of a linear map $\hat{A}$ is a linear map $\hat{A}^\dagger$ such that

$$(|v\rangle, \hat{A}|w\rangle) = (\hat{A}^\dagger|v\rangle, |w\rangle)$$

for all $|v\rangle$ and $|w\rangle$ in the Hilbert space (this is its definition). Its has the properties that $(\hat{A}^\dagger)^\dagger = \hat{A}$ and $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$. We also talk of the dual vector to $|v\rangle$ as its Hermitian conjugate: $\langle v| = (|v\rangle)^\dagger$. Then, by definition of the Hermitian conjugation, $(\hat{A}|v\rangle)^\dagger = \langle v|\hat{A}^\dagger$.

A **Hermitian operator** is a linear operator $\hat{A}$ such that $\hat{A}^\dagger = \hat{A}$. It has the properties that its eigenvalues are real, that eigenvectors with different eigenvalues are orthonormal, and if it is complete, that its normalised eigenvectors form an orthonormal basis for the Hilbert space (if we choose in each eigenspace an orthonormal basis).

Two very much used types Hilbert spaces are:

- The vector spaces $\mathbb{R}^n$ for $n = 1, 2, 3, \ldots$. Vectors can be simply represented by **column vectors** with n components,

$$|v\rangle = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

and linear maps by **matrices** n by n. Hermitian conjugation is a **complex conjugation** of the matrix followed by a **transpose** (or the other way around, it's the same thing):

$$\hat{A}^\dagger = (\hat{A}^*)^t.$$

Likewise, the dual vector to $|v\rangle$ is

$$\langle v| = ((|v\rangle)^*)^t = \begin{pmatrix} v_1^* & \cdots & v_n^* \end{pmatrix},$$

from which follows the inner product:

$$\langle v|w\rangle = \sum_{j=1}^n v_j^* w_j.$$

2

- The vector spaces of functions on $\mathbb{R}^n$. Vectors are functions $\mathbb{R}^n \mapsto \mathbb{C}$, linear maps are for instance the derivatives, or the multiplication by a given function. For $|f\rangle$ represented by the function $f : \mathbf{x} \mapsto f(\mathbf{x})$, we sometimes write $|f\rangle = f$. The standard inner product is

$$(f, g) = \langle f|g \rangle = \int_{-\infty}^{\infty} dx_1 \cdots dx_n \, (f(\mathbf{x}))^* g(\mathbf{x}).$$

2 Physical concepts

A **physical state** is a **ray** in a Hilbert space. This means that a given physical state may be represented by a **normalised vector**, a vector $|v\rangle$ whose length is $\langle v|v \rangle = 1$, or by any other vector obtained from $|v\rangle$ by multiplying by a phase, $e^{i\theta}|v\rangle$, $\theta \in \mathbb{R}$.

An **observable** is a linear operator on the Hilbert space, with the properties that it is Hermitian and complete. For matrices, all we have to verify is that $(\hat{A}^*)^t = \hat{A}$. For operators on a space of functions, we have to go back to the basic definition of Hermitian conjugation: if $\hat{A}$ is Hermitian, then $(f, \hat{A}g) = (\hat{A}f, g)$. For differential operators, this can be verified by using integration by parts. In general we don't have to worry about completeness (it is usually there essentially by definition of our Hilbert space as a representation space of the observables).

The **spectrum** of an observable is the possible values that can be obtained when we make a measurement associated to this observable. The spectrum is given by the **set of all possible eigenvalues** of the observable. In order to find the eigenvalues of a matrix, we have to solve the characteristic equation $\det(\hat{A} - \lambda \mathbf{1}) = 0$ (the set of solutions is the set of eigenvalues). In order to find the eigenvectors, we have to solve $\hat{A}|v_\lambda\rangle = \lambda|v_\lambda\rangle$ for λ in the set of eigenvalues. For operators on the space of functions, in order to find the eigenvectors and eigenvalues we have to solve the (usually differential) equation $\hat{A}|f_\lambda\rangle = \lambda|f_\lambda\rangle$. The set of allowed eigenvalues is obtained by imposing constraints on the **asymptotic behaviour** of the function at large $|\mathbf{x}|$ (for instance, $\lim_{|\mathbf{x}|\to\infty} f_\lambda(\mathbf{x}) = 0$), or by imposing **boundary conditions** (for instance, $f_\lambda(L) = f_\lambda(0)$, or $f_\lambda(0) = 0$, $f_\lambda(L) = 0$). Which asymptotic condition or boundary condition we must consider depend on the problem at hand.

The **probability of observing λ upon a measurement of $\hat{A}$ in a state** $|w\rangle$ is given by

$$|\langle v_\lambda|w\rangle|^2$$

if there is a unique eigenvector $|v_\lambda\rangle$ associated to λ (all vectors must be normalised). If there are many eigenvectors associated to λ, then they form a linear space, and we have to choose a basis of orthonormal eigenvectors in this space, $|v_{\lambda,i}\rangle$. Then, the probability is

$$\sum_i \langle v_{\lambda,i}|w\rangle|^2.$$

A similar formula holds if the index i is in fact a continuous index, where the sum is replaced by an integral.

Just after the measurement, the state has collapsed to the eigenstate (or the eigenspace) corresponding to the value measured:

$$|w\rangle \mapsto |v_\lambda\rangle \quad \text{or} \quad |w\rangle \mapsto \frac{\sum_i |v_{\lambda,i}\rangle\langle v_{\lambda,i}|w\rangle}{\sqrt{\sum_i \langle w|v_{\lambda,i}\rangle\langle v_{\lambda,i}|w\rangle}},$$

if the value λ was actually obtained in the laboratory.

For a quantum mechanical particle, the **position observable** (here in one dimension) is denoted $\hat{X}$. Its eigenvectors are denoted by $|x\rangle$ for $x \in \mathbb{R}$ corresponding to positions on the real line (in principle x has appropriate dimensions, like centimeters, etc.), and the associated eigenvalues are x. That is,

$$\hat{X}|x\rangle = x|x\rangle.$$

The orthonormality relation is $\langle x|x'\rangle = \delta(x-x')$ and the completeness relation is $\mathbf{1} = \int_{-\infty}^{\infty} dx\, |x\rangle\langle x|$.

The **wave function** associated to a vector $|f\rangle$ is the function $f : \mathbf{x} \mapsto f(\mathbf{x}) = \langle \mathbf{x}|f\rangle$. The space of wave functions forms a natural Hilbert space with the inner product mentioned above, because, for instance in one dimension,

$$\langle f|g\rangle = \int dx\, \langle f|x\rangle\langle x|g\rangle = \int dx\, (f(x))^* g(x).$$

The position observable $\hat{X}$ acts by multiplication by the argument x: $(\hat{X}f)(\mathbf{x}) = x f(\mathbf{x})$.

The **momentum observable** (again, in one dimension) is denoted $\hat{P}$. It satisfies the **canonical commutation relation** with the position observable,

$$[\hat{X}, \hat{P}] = i\hbar.$$

The momentum operator acts by differentiation on the space of wave functions: $(\hat{P}_1 f)(\mathbf{x}) = -i\hbar\frac{d}{dx}f(\mathbf{x})$ (here with $\hat{P}_1$ representing the momentum in the x direction; in one dimension, it is just denoted $\hat{P}$). Its eigenfunctions are denoted by $|p\rangle$ for $p \in \mathbb{R}$ (with appropriate momentum dimensions), with corresponding eigenvalues p, that is,

$$\hat{P}|p\rangle = p|p\rangle.$$

The orthonormality relation is $\langle p|p'\rangle = \delta(p-p')$ and the completeness relation is $\mathbf{1} = \int_{-\infty}^{\infty} dp\, |p\rangle\langle p|$. The overlap between position and momentum eigenfunctions are

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ixp/\hbar},$$

this gives the wave function of $|p\rangle$, $f_p : x \mapsto \langle x|p\rangle$, which indeed satisfies $(\hat{P}f_p)(x) = -i\hbar df_p(x)/dx = p f_p(x)$.

The **Dirac quantisation condition** (or correspondence principle) is the general statement from which the canonical commutation relation is derived. Given a classival variable A (in general, for a particle in one dimension for instance, this a function of X and P), there is a map

ˆ : $A \mapsto \hat{A}$ to a quantum observable, such that for any two classical variables A and B and their corresponding quantum observables $\hat{A}$ and $\hat{B}$, we have the Dirac quantisation condition:

$$[\hat{A}, \hat{B}] = i\hbar \widehat{\{A, B\}}$$

where $\{A, B\}$ is the Poisson bracket of classical mechanics,

$$\{A, B\} = \frac{\partial A}{\partial X}\frac{\partial B}{\partial P} - \frac{\partial B}{\partial X}\frac{\partial A}{\partial P}.$$

In the case $A = X$ and $B = P$, this give the canonical commutation relation. In general, this allows us to identify the classical meaning of a given quantum observable.

The **Probability density** is

$$\rho(\mathbf{x}) = |\psi(\mathbf{x})|^2$$

for a wave function $\psi(\mathbf{x})$. This means that the probability of finding the particle in a volume V is $\int_V dx dy dz \, |\psi(\mathbf{x})|^2$. The **Probability current** is

$$\mathbf{j}(\mathbf{x}) = \text{Re}\left((\psi(\mathbf{x}))^* \frac{\hat{\mathbf{P}}}{m} \psi(\mathbf{x})\right) = \text{Re}\left((\psi(\mathbf{x}))^* \frac{(-i\hbar)}{m} \nabla \psi(\mathbf{x})\right)$$

The continuity relation, which has to do with time evolution, is

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{x}, t) = 0.$$

Additional topics: Heisenberg uncertainty relation, the harmonic oscillator, the angular momentum, ...

3 Time evolution

The observable that allows us to describe time evolution is the **Hamiltonian** $\hat{H}$, the quantum observable associated to the Hamiltonian of classical mechanics. When the Hamiltonian is explicitly time-independent (which is the only case that we actually considered), then it is the total energy observable. For one particle of mass m in a potential $V(\mathbf{x})$, the quantum Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2m} + V(\hat{\mathbf{X}}).$$

For other quantum systems, the Hamiltonian may look different; for instance, for systems with just finitely many basis states, the Hamiltonian is just some Hermitian matrix.

In the **Schrödinger picture**, what changes in time is the **physical state**, and the observables stay the same. The state vector changes in time as:

$$|\psi\rangle(t) = e^{-i\hat{H}t/\hbar}|\psi\rangle(0).$$

This means that if we want to evolve a given initial state vector $|\psi\rangle$ in time, we may first write it in the basis of Hamiltonian eigenstates $|E, i\rangle$ (with eigenvalue E),

$$|\psi\rangle = \sum_{E,i} |E, i\rangle\langle E, i|\psi\rangle$$

then evolve each eigenstate in this expansion:

$$|\psi\rangle(t) = \sum_{E,i} e^{-i\hat{H}t/\hbar}|E,i\rangle\langle E,i|\psi\rangle = \sum_{E,i} e^{-iEt/\hbar}|E,i\rangle\langle E,i|\psi\rangle.$$

A **stationary state** is a physical state that does not change in time. The state vector of a stationary state can only acquire a phase when evolved in time (so that it corresponds to the same physical state). Hence, the state vector of a stationary state is an **eigenstate of the Hamiltonian**, $\sum_i c_i|E,i\rangle$. All physical properties of a stationary state are invariant with time, no matter what observable we are looking at.

The evolution equation can also be written in infinitesimal form,

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

This, in terms of wave functions for one particle, gives us the **time-dependent Schrödinger equation**:

$$i\frac{d}{dt}\psi(\mathbf{x},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) + V(\mathbf{x})\psi(\mathbf{x},t)$$

or in one dimension

$$i\frac{d}{dt}\psi(x,t) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x,t) + V(x)\psi(x,t).$$

Clearly, knowing the Hamiltonian eigenstates (the stationary states) is a good way of obtaining the time evolution of states. Hence, we often have to find the eigenstates and eigenvalues of the Hamiltonian. For wave functions, the Hamiltonian eigenvalue equation $\hat{H}|\psi\rangle = E|\psi\rangle$ is the **time-independent Schrödinger equation**:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) + V(\mathbf{x})\psi(\mathbf{x},t) = E\psi(\mathbf{x},t).$$

In the **Heisenberg picture**, as time passes all states stay the same, but the observables change in time. The observables evolve as

$$\hat{A}(t) = e^{i\hat{H}t/\hbar}\hat{A}(0)e^{-i\hat{H}t/\hbar}.$$

Likewise, then, all eigenvectors of the observables change in time, but the eigenvalues always stay the same: if $\hat{A}(0)|a\rangle = a|a\rangle$, then $\hat{A}(t)|a\rangle(t) = a|a\rangle(t)$ with $|a\rangle(t) = e^{i\hat{H}t/\hbar}|a\rangle(0)$.

A **conserved quantity** is an observable $\hat{A}$ that does not change with time, $\hat{A}(t) = \hat{A}(0)$. The only condition we need is that

$$[\hat{A}, \hat{H}] = 0.$$

All averages and probabilities associated to a conserved quantity are invariant in time, no matter what the initial state is.

The equivalence between the two pictures is obtained by realising, besides the fact that the eigenvalues of observables are the same, that all overlaps between observable eigenvectors and the state vector are the same in both pictures:

$$(|a\rangle(t))^*|\psi\rangle = \langle a|e^{-i\hat{H}t/\hbar}|\psi\rangle = \langle a|\psi\rangle(t).$$

This means that all probabilities are the same, and that the collapsed vector after the measurement is also the same.

Additional topics: Ehrenfest's theorem, confined/scattering states, reflection/transmission coefficients, three-dimensional problems, conserved quantities and separation of variables, the hydrogen atom, ...

4   Commutators and related formulas

The **commutator** is $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$. We immediately have $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$. The Jacobi identity tells us that

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0.$$

The commutator is useful if we want to try to exchange the order of two operators: $\hat{A}\hat{B} = \hat{B}\hat{A} + [\hat{A}, \hat{B}]$. For instance, $\hat{X}\hat{P} = \hat{P}\hat{X} + i\hbar$.

The basic commutator formula for product of operators is

$$[\hat{A}\hat{B}, \hat{C}] = [\hat{A}, \hat{C}]\hat{B} + \hat{A}[\hat{B}, \hat{C}].$$

We have to be careful with the order of operators everywhere. The formula also holds for the other factor

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}].$$

It also generalises:

$$[\hat{A}, \hat{B}\hat{C}\hat{D}] = [\hat{A}, \hat{B}]\hat{C}\hat{D} + \hat{B}[\hat{A}, \hat{C}]\hat{D} + \hat{B}\hat{C}[\hat{A}, \hat{D}]$$

and

$$[\hat{A}\hat{B}\hat{C}, \hat{D}] = [\hat{A}, \hat{D}]\hat{B}\hat{C} + \hat{A}[\hat{B}, \hat{D}]\hat{C} + \hat{A}\hat{B}[\hat{C}, \hat{D}],$$

etc.

In order to evaluate commutators between operators on a space of functions, like $[x^2, xd/dx]$, there are two possible ways. We may just recall that $[d/dx, x] = 1$ and use the formulas above for product of operators. Or, we may apply both operators on arbitrary functions:

$$[x^2, xd/dx]f(x) = x^2\,x(d/dx)f(x) - x(d/dx)(x^2f(x)) = x^3 df(x)/dx - 2x^2 f(x) - x^3 df(x)/dx = -2x^2 f(x)$$

from which we conclude that $[x^2, xd/dx] = -2x^2$.

Another important formula is that related to the exponential. We may define in general

$$e^{\hat{A}} = 1 + \hat{A} + \frac{1}{2}\hat{A}^2 + \frac{1}{3!}\hat{A}^3 + \dots.$$

Then, putting terms together, it turns out that there is a nice formula for the following object:

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots.$$

We often have something like $e^{\hat{A}}\hat{B}e^{-\hat{A}}$ in quantum mechanics (for instance for the time evolution in the Heisenberg picture), so this is very useful.