2. Mathematical Formalism of Quantum Mechanics

Quantum mechanics is a linear theory, and so it is natural that vector spaces play an important role in it. A physical state is represented mathematically by a vector in a Hilbert space (that is, vector spaces on which a positive-definite scalar product is defined); this is called the space of states. Physical properties like momentum, position, energy, and so on will be represented by operators acting in the space of states. We will introduce the essential properties of Hilbert spaces, mainly in the case of finite dimension, as the mathematical theory of Hilbert spaces of infinite dimension is much more complicated than that of spaces of finite dimension

2.1 Linear vectors and Hilbert space

$\mathcal{D}$: **Linear vector space** A linear vector space is a set of elements, called vectors, which is closed under addition and multiplication by scalars.
Using Dirac notation, the vectors are denoted by *kets*: $|k\rangle$. We can associate to each ket a vector in the dual space called *bra*: $\langle\psi|$.
If two vectors $|\psi\rangle$ and $|\varphi\rangle$ are part of a vector space, then $|\psi\rangle + |\varphi\rangle$ also belongs to the space. If a vector $|\psi\rangle$ is in the space, then $\alpha\,|\psi\rangle$ is also in the space (where α is a complex scalar).
A set of *linearly independent* vectors $\{|\varphi_i\rangle\}$ is such that $\sum_k c_k\,|\varphi_k\rangle = 0$ if and only if $c_k = 0\ \forall k$ (no trivial combination of them sums to zero).
The *dimension* of the space N is the maximum number of linearly independent vectors (which is also the smallest number of vectors that span the space).

$\mathcal{D}$: **Basis** A maximal set of linearly independent vectors in the space is called a basis. (e.g. $\{|\phi_k\rangle\}$, $k = 1, \ldots,\ N$).
Any vector in the space can be written as a linear superposition of the basis vectors:

$$|\psi\rangle = \sum_k a_k\,|\phi_k\rangle \tag{1}$$

To any vector we can thus associate a column vector of N complex numbers $(a_1, a_2 \ldots a_n)^T$. Here we are going to restrict ourselves to bounded, finite dimension spaces (even if many physical spaces are not: for example energy spaces can be unbounded and position has infinite dimension).

$\mathcal{D}$: **Hilbert space** The Hilbert space is a linear vector space over complex numbers with an *inner product*.

$\mathcal{D}$: **Inner product** An inner product is an ordered mapping from two vectors to a complex number (for a Hilbert space a mapping from a ket and a bra to a complex number $c = \langle\psi|\varphi\rangle$) with the following properties:

– positivity: $\langle\psi|\psi\rangle \geq 0$. The equality holds only for the zero vector $|\psi\rangle = 0$.
– linearity in the second function: $\langle\psi|(c_1\varphi_1\rangle + c_2|\varphi_2\rangle) = c_1\langle\psi|\varphi_1\rangle + c_2\langle\psi|\varphi_2\rangle$.

– anti-linearity in the first function:$(\langle c_1\varphi_1 + \langle c_2|\varphi_2)|\psi\langle = c_1^*\langle\varphi_1|\psi\rangle + c_2^*\langle\varphi_2|\psi\rangle$.
– skew symmetry: $\langle\psi|\varphi\rangle = \langle\varphi|\psi\rangle^*$

$\mathcal{D}$: **Norm** The norm of a vector is $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$.
Since the Hilbert space is characterized by its inner product, vectors are defined up to a global phase, that is, $|\psi\rangle = e^{i\vartheta}|\psi\rangle$. Relative phase is instead very important: $|\psi\rangle + e^{i\vartheta}|\phi\rangle \neq |\psi\rangle + |\phi\rangle$.
The inner product properties allow us to define two geometric inequalities:

– Schwartz inequality: $|\langle\psi|\varphi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\varphi|\varphi\rangle$.
– Triangular inequality: $\|(\psi+\varphi)\| \leq \|\varphi\| + \|\psi\|$.

The equality holds only if the two vectors are in the same direction: $|\psi\rangle = c|\varphi\rangle$.
There is also an antilinear correspondence between the dual vectors ket and bra:

$$c_1|\psi_1\rangle + c_2|\psi_2\rangle \rightarrow c_1^*\langle\psi_1| + c_2^*\langle\psi_2|$$

$\mathcal{D}$: **Orthonormal set** A set of vectors $\{|\varphi_k\rangle\}$ is orthonormal if for each pair the inner product $\langle\varphi_k|\,\varphi_j\rangle = \delta_{k,j}$.

2.2 Operators

We can define a set of operators that acting on the vectors return vectors:

$\mathcal{D}$: **Operator** An operator A on a vector space is a mapping between two vectors in that space: $A|\psi\rangle = |\phi\rangle$.
A *linear* operator satisfies:

$$A(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1A|\psi_1\rangle + c_2A|\psi_2\rangle$$

To characterize and parametrize A we look at its action on each vector in the space. Because of linearity, it is however enough to characterize A with its action on the N basis vectors $\{|\phi\rangle_k\}$. In this way we can associate a matrix representation to any operator, in the same way we associated arrays of complex numbers with the vectors. In particular, given an orthonormal basis $\{|v\rangle_k\}$, the matrix representation of the operator A is an $N \times N$ square matrix A whose elements are given by $A_{k,j} = \langle v_k|\,A\,|v_j\rangle$.
Let us consider an orthonormal basis $\{v_i\}$, then as seen any vector can be written as: $|\psi\rangle = \sum_{i=1}^N a_i|v_i\rangle$. The action of an operator A becomes:

$$A|\psi\rangle = |\varphi\rangle \rightarrow \sum_{i=1}^N Aa_i|v_i\rangle = \sum_{i=1}^N b_i|v_i\rangle$$

To extract one of the coefficients, say b_k we multiply by the bra $\langle v_k|$, obtaining:

$$\sum_{i=1}^N \langle v_k|Aa_i|v_i\rangle = b_k \rightarrow \sum_i A_{ki}a_i = b_k$$

The action of an operator can be thus seen as a matrix multiplication (again, here we are restricting to bounded, finite dimension spaces that support finite operators, hence this simple matrix representation).

? Question: Perform a simple matrix multiplication.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

This is equivalent to $R_x \cdot \vec{v}_z = \vec{v}_y$. □

The *domain* of an operator is the subspace on which it acts non-trivially (spanned by $k \leq N$ vectors).

Two operators A and B are equal if their domains are the same and their action is equal $\forall|\psi\rangle$ in their domains. The sum and product of operators are then defined as

$$(A+B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle)$$

$$(AB)|\psi\rangle = A(B|\psi\rangle)$$

The operators are associative:

$$A(BC)|\psi\rangle = (AB)C|\psi\rangle$$

But they are not in general commutative:

$$AB|\psi\rangle \neq BA|\psi\rangle$$

$\mathcal{D}$: **Commutator** . The commutator of two operators is $[A, B] = AB - BA$. Two operators commute/are commutable if $[A, B] = 0$.

2.2.1 Hermitian operators

An important class of operators are self adjoint operators, as observables are described by them.

$\mathcal{D}$: **Adjoint** The adjoint of an operator $A^\dagger$ is an operator acting on the dual space with the property: $\left\langle (A^\dagger\psi) \middle| \varphi\right\rangle = \langle\psi|(A\varphi)\rangle$, $\forall\{|\psi\rangle, |\varphi\rangle\}$. We can also have other notations. From $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$ (where $*$ indicates the complex conjugate) we have $\left\langle (A^\dagger\psi) \middle| \varphi\right\rangle = \langle\psi|(A\varphi)\rangle = \langle\varphi|A^\dagger\psi\rangle^*$. Also, we can write the inner product as $\langle\varphi|(A\psi)\rangle = \langle\varphi|A|\psi\rangle$ and $\langle(A\varphi)|\psi\rangle = \langle\varphi|A^\dagger|\psi\rangle$. In matrix representation, this means that 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 and $A^\dagger$ operate on the same domain and with the property

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

or shortly, $A^\dagger = A$. In matrix representation we have then: $A_{ki} = A^*_{ik}$.

? Question: Prove that $(cA)^\dagger = c^* A^\dagger$
We want to prove that $(cA)^\dagger = c^* A^\dagger$. We can take two strategies:
1) From the adjoint operator definition in the form:

$$\langle B^\dagger\phi|\psi\rangle = \langle\phi|B\psi\rangle,$$

with $B = cA$ we obtain:

$$\langle(cA)^\dagger\phi|\psi\rangle = \langle\phi|cA\psi\rangle = c\langle\phi|A\psi\rangle = c\langle A^\dagger\phi|\psi\rangle = \langle c^* A^\dagger\phi|\psi\rangle$$

2) Alternatively, we can use the adjoint definition in Dirac's notation:

$$\langle\varphi|B^\dagger|\psi\rangle = \langle\psi|B|\varphi\rangle^*,$$

to get:

$$\langle\varphi|(cA)^\dagger|\psi\rangle = \langle\psi|cA|\varphi\rangle^* = c^*\langle\psi|A|\varphi\rangle^* = c^*\langle\varphi|A^\dagger|\psi\rangle = \langle\varphi|c^* A^\dagger|\psi\rangle$$

□

Note that we can write

$$\langle B^\dagger\phi|\psi\rangle = \langle\phi|B\psi\rangle = \langle\varphi|B|\psi\rangle = \langle\psi|B^\dagger|\varphi\rangle^*.$$

The second notation (based on Dirac's notation) could be seen as implying $(|\varphi\rangle)^\dagger = \langle\varphi|$ (and thus $(A|\varphi\rangle)^\dagger = \left\langle A^\dagger\phi\right|$. However, this applies the adjoint operation to a vector, while the adjoint is only properly defined for operators. For discrete dimensional spaces, which allow a matrix representation, there is no ambiguity since we have the equivalence of the adjoint with the complex-transpose of an operator (which can be defined also for vectors)[4].

[4] See also quant-ph/9907069 page 12, for a subtle difference between Hermitian and self-adjoint infinite-dimensional operators

? Question: Prove that $(AB)^\dagger = B^\dagger A^\dagger$
$\forall|\psi\rangle$ we have $|\varphi\rangle = (AB)|\psi\rangle \to \langle\phi| = \langle\psi|\,(AB)^\dagger$. Define $|\phi\rangle = B|\psi\rangle$, then $|\varphi\rangle = A|\phi\rangle$, $\langle\varphi| = \langle\psi|\,B^\dagger$ and $\langle\phi| = \langle\varphi|\,A^\dagger$, so that $\langle\phi| = \langle\psi|\,B^\dagger A^\dagger$. □

A self-adjoint operator is also Hermitian in bounded, finite space, therefore we will use either term. Hermitian operators have some properties:

1. if A, B are both Hermitian, then $A + B$ is Hermitian (but notice that AB is a priori not, unless the two operators commute, too.).
2. if A, B are both Hermitian but do not commute, then at least $AB + BA$ is Hermitian.

? Question: Prove property # 2.
$(AB + BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB$. □

Before describing other properties we need the following definition.

$\mathcal{D}$: **Eigenvector** We define a right eigenvector as a column vector $|\psi\rangle_R$ satisfying $A|\psi\rangle_R = \lambda_R|\psi\rangle_R$, so $(A - \lambda_R\mathbb{1})|\psi\rangle_R = 0$, which means the right eigenvalues λ_R must have zero determinant, i.e., $\det(A - \lambda_R\mathbb{1}) = 0$. Similarly, a left eigenvector is such that $\langle\psi|_L A = \lambda_L\langle\psi|_L$.

The following properties will be very important in QM:

3. if A is Hermitian its eigenvalues are real (eigenvalues: scalar a such that $A|\psi\rangle = a|\psi\rangle$). It is easy to show this properties from $\langle\psi|A|\psi\rangle = a = a^*$.
4. distinct eigenvectors of an Hermitian operator are orthogonal: $A|\psi_1\rangle = a_1|\psi_1\rangle,\ A|\psi_2\rangle = a_2|\psi_2\rangle \to \langle\psi_1|\psi_2\rangle = 0$ unless $a_1 = a_2$.
5. distinct eigenvalues correspond to orthogonal eigenvectors:
Given $A\,|\psi_1\rangle = c_1\,|\psi_1\rangle$ and $A\,|\psi_2\rangle = c_2\,|\psi_2\rangle$, if $c_1 \neq c_2 \ \to \langle\psi_1|\,\psi_2\rangle = 0$.

As observables are given by Hermitian operators, the first properties will imply that the values that an observable can take on are only real values (as needed for the observable to have a physical meaning). On the domain of the operator, the eigenvectors form a complete orthogonal basis set.

? Question: Prove property # 5.
$\langle\psi_2|\,A\psi_1\rangle = \langle\psi_2|\,c_1\psi_1\rangle = \langle c_2^*\psi_2|\,\psi_1\rangle$. For Hermitian operators then $c_1\,\langle\psi_2|\,\psi_1\rangle = c_2\,\langle\psi_2|\,\psi_1\rangle$, which is satisfied only if $c_1 = c_2$ or if $\langle\psi_1|\,\psi_2\rangle = 0$. □

? Question: Prove property # 4.
Consider two eigenstates of A $|a_1\rangle$ and $|a_2\rangle$. We have $\langle a_2|\,A\,|a_1\rangle = \langle a_1|\,A\,|a_2\rangle^*$ since A is Hermitian. Now $\langle a_2|\,A\,|a_1\rangle = a_1\langle a_2|a_1\rangle$ and $\langle a_1|\,A\,|a_2\rangle^* = (a_2\langle a_1|a_2\rangle)^* = a_2(\langle a_1|a_2\rangle)^*$ since a_2 is real (being an eigenvector of A. We thus have $a_1\langle a_2|a_1\rangle = a_2\langle a_2|a_1\rangle$ which is satisfied iif $a_1 = a_2$ (contrary to the hypothesis) or if $\langle a_2|a_1\rangle = 0$. □

2.2.2 Operators and their properties

$\mathcal{D}$: **The Outer Product** $|\psi\rangle\langle\varphi|$ is an operator, since acting on a vector returns a vector: $(|\psi\rangle\langle\varphi|)\,|\phi\rangle = \langle\varphi|\phi\rangle|\psi\rangle$.
It defines a projector operator $P_i = |v_i\rangle\langle v_i|$. The sum over all projectors on the space is the identity, therefore, for any basis set we have: $\sum_i |v_i\rangle\langle v_i| = \mathbb{1}$ (closure relation). The product of two projectors is $P_jP_k = \delta_{ik}P_j$. Projectors derive their name from the property that they project out a vector component of the related basis vector: given $P_j = |v_j\rangle\,\langle v_j|$, $P_j|\psi\rangle = P_j\sum_k c_k\,|v_k\rangle = c_j\,|v_j\rangle$.

$\mathcal{D}$: **Trace** - The trace of an operator is the sum of the diagonal elements of an operator $\text{Tr}\,\{A\} = \sum_{j=1}^N A_{jj} = \sum_j \langle v_j|\,A\,|v_j\rangle$. It is independent of the choice of basis.

$\mathcal{D}$: **Spectral Decomposition** - The spectral theorem states that given a self-adjoint operator A on a linear space $\mathcal{H}$, there exists an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of A. Equivalently, we can state that A can be written as a linear combination of pairwise orthogonal projections (which are formed from its eigenvectors). This representation of A is called its spectral decomposition: $A = \sum_j a_j\,|v_j\rangle\,\langle v_j|$, where $A\,|v_j\rangle = a_j\,|v_j\rangle$. In this basis, the matrix representation of A is diagonal.

■ Theorem: If two hermitian operators commute, they share a common set of eigenvectors.
If $[A, B] = 0$ then $AB = BA$. Given two eigenvectors of A, we have $\langle a'|\,(AB - BA)\,|a''\rangle = a'\,\langle a'|\,B\,|a''\rangle - a''\,\langle a'|\,B\,|a''\rangle$. This is zero if $a'' = a'$ (and $\langle a'|\,B\,|a'\rangle$ is a diagonal term of B and it can be anything) or if $\langle a'|\,B\,|a''\rangle = 0$ (off-diagonal, with $a' \neq a''$). Thus B is diagonal in the basis of A's eigenvectors, hence A's eigenvectors are also eigenvectors of B. □
A simultaneous eigenvector of A and B $|a, b\rangle$ has the property: $A\,|a, b\rangle = a\,|a, b\rangle$ and $B\,|a, b\rangle = b\,|a, b\rangle$. The notation $|a, b\rangle$ is useful when the eigenvector is **degenerate**, that is, there exist more than one eigenvector with the same eigenvalue: $A\,\left|a^{(i)}\right\rangle = a\,\left|a^{(i)}\right\rangle$, $i = 1, \ldots n$, where n is the degeneracy. Then the label b serves to distinguish different eigenvectors.

$\mathcal{D}$: **Unitary operator** An operator fulfilling the conditions $U^\dagger U = \mathbb{1}$ and $UU^\dagger = \mathbb{1}$ is called unitary.

■ Theorem: Given two sets of basis kets $\{|\psi\rangle_i\}$ and $\{|\phi\rangle_i\}$ there exist a unitary operator such that $|\phi\rangle_i = U|\psi\rangle_i$, $\forall i$. (The unitary operator is $U = \sum_k |\varphi_k\rangle\,\langle\psi_k|$).

2.2.3 Functions of operators

Functions of operators are defined by the corresponding Taylor expansion of the function (if that exists). If $f(x) = f(0) + f'(0)x + \cdots + \frac{1}{n!}f^{(n)}(0)x^n + \ldots$, then $f(A) = f(0)\mathbb{1} + f'(0)A + \cdots + \frac{1}{n!}f^{(n)}(0)A^n + \ldots$, where the matrix power is defined recursively in terms of products $A^n = A(A^{n-1})$.

? Question: Show that given the spectral decomposition of the operator $A = \sum_a \lambda_a |a\rangle\langle a|$ we have $f(A) = \sum_a f(\lambda_a)|a\rangle\langle a|$.
We can first prove that $A^2 = (\sum_a \lambda_a |a\rangle\langle a|)(\sum_a \lambda_a |a\rangle\langle a|) = \sum_{a,b} \lambda_a\lambda_b(|a\rangle\langle a|)(|b\rangle\langle b|) = \sum_a \lambda_a^2 |a\rangle\langle a|$. Then show that if the theorem is valid for $n-1$ it is also valid for n. Finally, use the Taylor expansion to show it's true. □

? Question: Consider in particular the exponential function: $\exp(\xi A) = \sum_n \frac{1}{n!}(\xi A)^n = \sum_k \exp(\xi a_k)|a_k\rangle\langle a_k|$. Prove that $f(ABA^{-1}) = Af(B)A^{-1}$
It's easy to show that $(ABA^{-1})^n = AB^nA^{-1}$ by expanding the product and using $AA^{-1} = \mathbb{1}$. In particular for unitary matrices $U^{-1} = U^\dagger \to f(UAU^\dagger) = Uf(A)U^\dagger$. □

22.51 Quantum Theory of Radiation Interactions
Fall 2012