Chapter 2

Mathematical formalism of quantum mechanics

Quantum mechanics is the best theory that we have to explain the physical phenomena (if we exclude gravity). The elaboration of the theory has been guided by real experiments as well as thought experiments and conceptual ideas of a great generation of young physicist [milestones from 1900 to 1930 are: Planck on black body spectrum (1900), Einstein on the photon (1905), Bohr on the atom (1913), De Broglie on the wave function (1924), Schroedinger on the wave function evolution (1926), Born on the interpretation of the wave function (1926), Heisenberg on matrix mechanics (1925), Dirac on relativistic QM (1930)]. Some never completely accepted their own ideas, although these still form the best theory that we have today. The mathematical form of the theory that we find in textbooks has been put forward by Dirac and von Neumann in the 30's Since then the quantum laws of physics have been used unchanged[1] to successfully describe an impressive range of phenomena ranging from macroscopic solid state, molecular to atomic, nuclear, sub-nuclear and particle physics scales.

The arena of QM is Hilbert space so we begin with some mathematical reminders on linear algebra in such spaces. Our goal is also to carefully introduce the reader to Dirac's bra and ket notation. Then we introduce 5 basic principles that define QM. We also discuss two genuine quantum notions, namely, entangled states and the no-cloning theorem.

[1] combined with special relativity when needed

2.1 Linear algebra in Dirac notation

A *Hilbert space* $\mathcal{H}$ is a vector space over the field of complex numbers $\mathbf{C}$, with an inner product. For a finite dimensional Hilbert space that is all. For an infinite dimensional Hilbert space we require that it is complete and separable[2]. In quantum information theory we will almost always deal with Hilbert spaces of quantum bits which are discrete by nature, hence our Hilbert spaces are finite dimensional and we do not have to worry about completeness and separability.

The vectors will be denoted $|\psi\rangle$ (pronounced ket psi). The hermitian conjugate (transpose and complex conjugate) is denoted by $\langle\psi|$ (pronounced bra psi). The inner product is denoted $\langle\phi|\psi\rangle$. This is the inner product of the vectors $|\phi\rangle$ and $|\psi\rangle$ and is called a *bracket* (for bra-ket). The inner product must satisfy:

1. *Positivity*: $\langle\phi|\phi\rangle \geq 0$ with equality if and only if $|\phi\rangle = 0$.
2. *Linearity*: $\langle\phi|(\alpha|\psi_1\rangle + \beta|\psi_2\rangle) = \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle$, $\alpha, \beta \in \mathbf{C}$
3. *Skew symmetry*: $\langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}$ where the bar denotes complex conjugation.

A ray is an equivalence class of vectors of the form $\lambda|\psi\rangle$ where $\lambda \in \mathbf{C}$ and $|\psi\rangle$ is a specified vector. This specified vector is a representative of the ray.

Example 1: Qbit or two level system. $\mathcal{H} = \mathbf{C}^2 = \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{ with } \alpha, \beta \in \mathbf{C} \right\}$. The inner product is $(\overline{\gamma}, \overline{\delta}) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \overline{\gamma}\alpha + \overline{\delta}\beta$. In Dirac notation we have

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha|0\rangle + \beta|1\rangle$$

where $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Moreover

$$(\bar{\gamma}, \bar{\delta}) = \bar{\gamma}\langle 0| + \bar{\delta}\langle 1|$$

and

$$(\bar{\gamma}\langle 0| + \bar{\delta}\langle 1|)(\alpha|0\rangle + \beta|1\rangle) = \overline{\gamma}\alpha\langle 0|0\rangle + \overline{\gamma}\beta\langle 0|1\rangle + \overline{\delta}\alpha\langle 1|0\rangle + \overline{\delta}\beta\langle 1|1\rangle = \overline{\gamma}\alpha + \overline{\delta}\beta$$

[2]Complete means that all Cauchy sequences converge in the norm induced by the inner product and separable that there is a contable orthonormal basis.

.

Example 2: particle in three dimensional space. $\mathcal{H} = L^2(\mathbf{R}^3) = \{f : \mathbf{R}^3 \to \mathbf{C},\ \int d^3x|f(x)|^2 < \infty\}$. The inner product is $\langle f|g\rangle = \int d^3x\overline{f(x)}g(x)$ and the induced norm $||f||_2 = \langle f|f\rangle^{1/2} = \int d^3x|f(x)|^2$. This space plays a fundamental role in quantum mechanics but we will not need it in this course, since we deal only with discrete degrees of freedom.

We will need the notion of *tensor product*. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces with two finite basis. Let the basis of the first space be $|i\rangle_1$, $i = 1, ..., n_1$, $\dim \mathcal{H}_1 = n_1$ and that of the second space $|j\rangle_2$, $j = 1, ..., n_2$, $\dim \mathcal{H}_2 = n_2$. We can form the tensor product space

$$\mathcal{H}_1 \otimes \mathcal{H}_2$$

which is simply the new Hilbert space spanned by the basis vectors

$$|i\rangle_1 \otimes |j\rangle_2$$

(also denoted $|i, j\rangle$ or $|i\rangle_1|j\rangle_2$). There are n_1n_2 such vectors so

$$\dim \mathcal{H}_1 \otimes \mathcal{H}_2 = n_1n_2$$

A general element of the tensor product space is of the form

$$|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}|i, j\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}|i\rangle_1 \otimes |j\rangle_2$$

Lastly we have to say what is the inner product in the product space:

$$\langle i', j'|i, j\rangle = (\langle i'|_1 \otimes \langle j'|_2)(|i\rangle_1 \otimes |j\rangle_2) = \langle i'|i\rangle_1\langle j'|j\rangle_2$$

.

Example 3. For one Qbit the Hilbert space is $\mathbf{C}^2$. We will see that the Hilbert space of two Qbits is $\mathbf{C}^2 \otimes \mathbf{C}^2$. The basis vectors of $\mathbf{C}^2 \otimes \mathbf{C}^2$ are $\{|0\rangle \otimes |0\rangle,\ |0\rangle \otimes |1\rangle,\ |1\rangle \otimes |0\rangle,\ |1\rangle \otimes |1\rangle\}$ or $\{|0, 0\rangle,\ |0, 1\rangle,\ |1, 0\rangle,\ |1, 1\rangle\}$. A general state is

$$|\psi\rangle = \alpha_{00}|0, 0\rangle + \alpha_{01}|0, 0\rangle + \alpha_{10}|0, 0\rangle + \alpha_{11}|1, 1\rangle$$

We have $\dim \mathbf{C}^2 \otimes \mathbf{C}^2 = 4$ and of course $\mathbf{C}^2 \otimes \mathbf{C}^2$ is isomorphic to $\mathbf{C}^4$: However it is important to stress that in QM the meaning of the first representation is really that *two Qbits are involved*: in general it is too difficult (meaningless in some sense) to do physics in a bad representation. Here a few inner products

are $\langle 0,0|0,0\rangle = \langle 0|0\rangle\langle 0|0\rangle = 1$, $\langle 0,1|0,1\rangle = \langle 0|0\rangle\langle 1|1\rangle = 1$, $\langle 0,1|1,1\rangle = \langle 0|1\rangle\langle 1|1\rangle = 0$ etc... From these one can compute the inner product of $|\psi\rangle$ and $|\phi\rangle = \beta_{00}|0,0\rangle + \beta_{01}|0,0\rangle + \beta_{10}|0,0\rangle + \beta_{11}|1,1\rangle$. We find the natural product of $\mathbf{C}^4$, $\langle\phi|\psi\rangle = \overline{\beta}_{00}\alpha_{00} + \overline{\beta}_{01}\alpha_{01} + \overline{\beta}_{10}\alpha_{10} + \overline{\beta}_{11}\alpha_{11}$. It is often useful to work in the canonical basis of $\mathbf{C}^4$

$$\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix} = |0,0\rangle \quad \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} = |0,1\rangle \quad \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} = |1,0\rangle \quad \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix} = |1,1\rangle$$

Once this (conventional) correspondence is fixed we can infer the rules for tensoring vectors in their coordinate representation

$$\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$$

$$\begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

You can see that in this course the convention is that you multiply the first set of coordinates by the second vector . All these rules generalize to $\mathbf{C}^2\otimes\mathbf{C}^2\otimes\mathbf{C}^2$ etc...

Cauchy-Schwarz inequality. As usual:

$$|\langle\phi|\psi\rangle| \leq \langle\phi|\phi\rangle^{1/2}\langle\psi|\psi\rangle^{1/2}$$

Closure relation. Let $|i\rangle$, $i = 1, ..., n$ be an orthonormal basis of the n-dimensional Hilbert space. Any vector $|\phi\rangle$ can be expanded as

$$|\phi\rangle = \sum_{i=1}^{n} c_i|i\rangle, \qquad c_i = \langle i|\phi\rangle$$

where the components c_i are obtained by projecting $|\phi\rangle$ over the basis vectors. The above expansion can be rewritten as

$$|\phi\rangle = \sum_{i=1}^{n} |i\rangle\langle i|\phi\rangle$$

Note that $|i\rangle\langle i|$ is the projection operator on vector $|i\rangle$. We can view $\sum_{i=1}^{n} |i\rangle\langle i|$ as the identity operator acting on $|\phi\rangle$, thus we have the *closure relation*

$$\sum_{i=1}^{n} |i\rangle\langle i| = I$$

This turns out to be a very useful identity for doing practical calculations in Dirac notation. Note that this identity is simply the spectral decomposition of the identity.

Observables. In QM observable quantities are represented by linear operators (matrices) on $\mathcal{H}$. Let us briefly review a few important facts. The map $A : \mathcal{H} \to \mathcal{H}$, $|\psi\rangle \to A|\psi\rangle$ is linear if

$$A(\alpha|\phi_1\rangle + \beta|\phi_2\rangle) = \alpha(A|\phi_1\rangle) + \beta(A|\phi_2\rangle)$$

The matrix elements of A in a basis $\{|i\rangle, i = 1, ..., n\}$ of $\mathcal{H}$ are denoted by $\langle i|A|j\rangle$ or A_{ij}. Given A, the *adjoint* of A is denoted $A^\dagger$ and defined by

$$\langle\phi|A^\dagger|\psi\rangle = \overline{\langle\psi|A|\phi\rangle}$$

So the adjoint (or hermitian conjugate) is the operator which has transposed and conjugate matrix elements. We say that A is self-adjoint (or hermitian) if $A = A^\dagger$. The later type of operators play a very central role in QM because observable quantities are represented by self-adjoint operators: the reader can guess that this must be so because any physical measurement is expressed by a real number (why ?) and self-adjoint operators have real eigenvalues. The reader can check that $(A + B)^\dagger = A^\dagger + B^\dagger$ and $(AB)^\dagger = B^\dagger A^\dagger$.

We will also need the following notations for the *commutator*

$$[A, B] = AB - BA$$

and the anticommutator

$$\{A, B\} = AB + BA$$

Projectors in Dirac notation. The linear operator $|i\rangle\langle i| = P_i$ is the projector on the basis vector $|i\rangle$. To check that P_i is a projector we need to verify that $P_i^\dagger = P_i$ and $P_i^2 = P_i$. Here is how one does it in Dirac notation

$$P_i^\dagger = (|i\rangle\langle i|)^\dagger = (\langle i|)^\dagger(|i\rangle)^\dagger = |i\rangle\langle i| = P_i$$

$$P_i^2 = (|i\rangle\langle i|)(|i\rangle\langle i|) = |i\rangle\langle i|i\rangle\langle i| = |i\rangle\langle i| = P_i$$

Since $|i\rangle$ and $|j\rangle$ are orthogonal for $i \neq j$ we have $P_i P_j = P_j P_i = 0$. Indeed

$$P_i P_j (|i\rangle\langle i|)(|j\rangle)(\langle j|) = |i\rangle\langle i|j\rangle\langle j| = 0$$

$$P_j P_i (|j\rangle\langle j|)(|i\rangle\langle i|) = |j\rangle\langle j|i\rangle\langle i| = 0$$

Note that if $|\phi\rangle$ is any vector of the Hilbert space, then $P_\phi = |\phi\rangle\langle\phi|$ is the projector on $|\phi\rangle$.

Spectral decomposition. Hermitian operators (matrices) on a Hilbert space have a *spectral decomposition* or *spectral representation*,

$$A = \sum_n a_n P_n$$

where $a_n \in \mathbf{R}$ are the eigenvalues and P_n the eigenprojectors of A. The eigenspaces of A are spanned by the orthonormal eigenvectors $|\phi_{nj}\rangle$ associated to the eigenvalue a_n:

$$A|\phi_{nj}\rangle = a_n|\phi_{nj}\rangle, \;\; P_n = \sum_j |\phi_{nj}\rangle\langle\phi_{nj}\rangle$$

The index j takes into account the possible degeneracy of a_n. From the orthonormality of the eigenvectors one sees that $P_n P_m = P_m P_n = 0$ for $n \neq m$. Note that for given n one always has the liberty to rotate the basis $\{|\phi_{nj}\rangle\}$ in the subspace of P_n. Moreover we have the closure relation

$$I = \sum_n P_n = \sum_{n,j} |\phi_{n,j}\rangle\langle\phi_{nj}\rangle$$

We will often write the spectral decomposition as

$$A = \sum_{n,j} a_n |\phi_{nj}\rangle\langle\phi_{nj}\rangle$$

In the non-degenerate case this becomes simply $A = \sum_n a_n |\phi_n\rangle\langle\phi_n\rangle$.

2.2 Principles of quantum mechanics

In this paragraph we explain the 5 basic principles of QM. In a nutshell:

- isolated systems are described by *states of a Hilbert space*,
- they *evolve unitarily with time*,

- *observable quantities are described by hermitian matrices*,
- measurement is a *distinct process* from time evolution: it is a *random projection*,
- systems can be brought together and *composed*: their Hilbert space is a tensor product space.

Their meaning, interpretation and soundness has been debated over the first half of the 20-th century by the founding fathers of QM and by their followers, specially the measurement postulate.

Principle 1: states. The state of a quantum system - that is isolated from the rest of the universe - is *completely* described by a ray in a Hilbert space. We require that the representative vector $|\psi\rangle \in \mathcal{H}$ is normalized to one, $\langle\psi|\psi\rangle = 1$.

Example 4.

- To describe the polarization of the photon we take $\mathcal{H} = \mathbf{C}^2$. States are vectors in $\mathbf{C}^2$, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$. For a linearly polarized state $|\theta\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$, for a circularly polarized state $|\tilde{\theta}\rangle = \cos\theta|0\rangle + i\sin\theta|1\rangle$, and for elliptic polarization $\cos\theta|0\rangle + e^{i\delta}\sin\theta|1\rangle$.
- The spin $\frac{1}{2}$ of an electron (say) is described by the same Hilbert space.
- For a Benzene molecule the Hilbert space is again the same and is spanned by the two valence bond states (see chapter 1):

$$|\psi\rangle = \alpha|1\rangle + \beta|2\rangle$$

- For a particle in $\mathbf{R}^3$ we have $\mathcal{H} = L^2(\mathbf{R}^3)$ as explained before. These are called wave functions and are normalized $\int d^3x|\psi(x)|^2 = 1$.

Remark. If $|\psi\rangle$ is a description of a system then $e^{i\lambda}|\psi\rangle$ is an equally good description. The *global* phase $\lambda \in \mathbf{R}$ is not an observable quantity and can be fixed arbitrarily. This is why QM states should really be defined as rays. However the *relative* phase of states is observable through interference effects. You might also wonder what is the difference between spin one-half and photon polarization. In fact photon polarization states and spin one-half states behave very differently under spatial rotations of the coordinate system (or the lab). Under a rotation of the reference frame the state of polarization of a photon behaves like a vector. In particular under a 2π rotation we recover the same state. On the other hand for spin one-half behaves as a

spinor (sometimes called half-vector) under a rotation of the reference frame. In particular, under a 2π rotation we recover the opposite state. In QM the representations of the rotation group (and any other group) on the Hilbert space does not have to satisfy $\mathbf{R}(2\pi) = 1$, precisely because states are rays. Therefore a phase is allowed for $\mathbf{R}(2\pi)|\psi\rangle = e^{i\lambda}|\psi\rangle$. All these aspects of QM will not mater too much in this course so we omit more explanations on what "spin" and "photon polarization" really are. A more profound discussion of these aspects would require to explain the representation theory of the Lorentz group of special relativity.

Principle 2: time evolution. An isolated quantum system evolves with time in a unitary fashion. This means that if $|\psi\rangle$ is the state at time 0, the state at time t is of the form $U_t|\psi\rangle$ where U_t is a unitary operator from $\mathcal{H} \to \mathcal{H}$. Here unitary means that $U_t^\dagger U_t = U_t U_t^\dagger = 1$ or equivalently $U_t^{-1} = U_t^\dagger$.

Unitary time evolution forms a group (it is a representation of translations along the time axis) in the sense that

$$U_{t=0} = I, \qquad U_{t_1} U_{t_2} = U_{t_1+t_2}$$

QM tells us how to compute U_t for a given system: one has to solve the *Schroedinger equation* or the *Heisenberg equations of motion*. These are equivalent in fact. The first one is the quantum mechanical version of the Hamilton-Jacobi equation of classical mechanics while the second is the quantum version of the Hamilton equations of motion. In quantum computation (at least in theory) we do not bother too much about these equations : we optimistically assume that if we need a specified U_t then somebody (a physicist, an engineer) will be able to construct a device (an electronic or optical device for example) which realizes the time evolution U. For us a specified time evolution is a *gate* that will ultimately be part of a quantum circuit.

It is very important to realize that *time evolution is linear*: this is quite surprising because in the classical regime one should get back the classical equations of motion which are generally non-linear[3].

Example 5. A semi-transparent mirror decomposes an incident ray into a reflected and a transmitted part (see chapter 1). Let $\mathcal{H} = \mathbf{C}^2$ the Hilbert space with basis $|T\rangle$, $|R\rangle$. The semi-transparent mirror acts in a unitary way

$$|T\rangle \to \boxed{\text{H}} \to H|T\rangle = \frac{1}{\sqrt{2}}(|T\rangle + |R\rangle)$$

[3]The study of this sort of reduction has led to a whole discipline called quantum chaos. Let us also point out that non-linear versions of the Schroedinger equation may arise when some degrees of freedom are integrated out, in other words for non-isolated systems.

$$|R\rangle \rightarrow \boxed{\text{H}} \rightarrow H|R\rangle = \frac{1}{\sqrt{2}}(|T\rangle - |R\rangle)$$

The unitary matrix H is called a Hadamard gate

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

One checks that $HH^\dagger = H^\dagger H = 1$. If we put two semi-transparent mirrors in series (see exercises)

$$|\psi\rangle \rightarrow \boxed{\text{H}} \rightarrow \boxed{\text{H}} \rightarrow H^2|\psi\rangle = |\psi\rangle$$

the output is equal to the input because $H^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In other words if the input state is $|T\rangle$ then the output is also $|T\rangle$. If we wish to take more seriously into account the effect of the perfect mirrors in-between the semi-transparent mirrors, we insert between the two Hadamard matrices the gate $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$$|\psi\rangle = \alpha|T\rangle + \beta|R\rangle \rightarrow \boxed{\text{H}} \rightarrow \boxed{\text{X}} \rightarrow \boxed{\text{H}} \rightarrow HXH|\psi\rangle = \alpha|T\rangle - \beta|R\rangle$$

Principle 3: observable quantities. In quantum mechanics an observable quantity (energy, magnetic moment, position, momentum,...) is represented by a linear self-adjoint operator[4] on $\mathcal{H}$. For us this just means a hermitian matrix.

Examples 6.

- Position x, momentum $p = \frac{\hbar}{i}\frac{\partial}{\partial x}$, energy or Hamiltonian $\frac{p^2}{2m} + V(x)$. We will not need these.
- However we will need things like the polarization of a photon. Suppose we send a photon in a polarized beam-splitter (see chapter 1). If D_y clicks we record a -1 while if D_x clicks we record a $+1$. Our observations can be described by the observable

  $$\mathcal{P} = (+1)|x\rangle\langle x| + (-1)|y\rangle\langle y|$$

  This is the self-adjoint matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (in the $|x\rangle$, $|y\rangle$ basis).

[4]There is a "correspondence principle" which is a rule of thumb on how to construct the appropriate self-adjoint operator from the classical one; in fact this procedure may sometimes be a bit ambiguous due to non-commutativity of operators

- General observables in $\mathcal{H} = \mathbf{C}^2$ can always be represented by 2×2 hermitian matrices

$$A = \begin{pmatrix} \alpha & \beta \\ \overline{\beta} & \gamma \end{pmatrix}$$

or in Dirac notation

$$A = \alpha|0\rangle\langle 0| + \beta|0\rangle\langle 1| + \overline{\beta}|1\rangle\langle 0| + \gamma|1\rangle\langle 1|$$

All such matrices can be written as linear combinations of

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

The observables (hermitian matrices !) X, Y, Z are called Pauli matrices. One of their uses is the description of the spin observable for spin $\frac{1}{2}$ particles: this is a "vector" with 3 components $\Sigma = (X, Y, Z)$. In the physics literature the notation is $\Sigma = (\sigma_x, \sigma_y, \sigma_z)$. Important properties of these matrices are

$$X^2 = Y^2 = Z^2 = I,\ XY = -YX,\ XZ = -ZX,\ YZ = -ZY$$

and

$$[X, Y] = Z,\ [Y, Z] = X,\ [Z, X] = Y$$

This algebra is a special example of spin or Clifford algebras which play an important role in QM.

Principle 4: measurement postulate. This is the most disturbing postulate: it requires a rather big leap of intuition (or stroke of genius) which goes back to Max Born (one also speaks of the Born interpretation of the wave function). Let a system be prepared in a state $|\psi\rangle$. The system is to be measured with an apparatus. The apparatus is modeled by a set of orthonormal projectors $\{P_n\}$ satisfying $\sum_n P_n = I$. A single measurement *reduces*[5] the state ψ of the system to

$$|\phi_n\rangle = \frac{P_n|\psi\rangle}{||P_n|\psi\rangle||} = \frac{P_n|\psi\rangle}{\langle\psi|P_n|\psi\rangle^{1/2}}$$

For a single measurement *there is no way to predict* what will be the specific outcome n: it is random. If the experiment is repeated many times (assuming this is a reproducible experiment) one finds that the probability (in a frequentist interpretation of the term) of the outcome n is

$$\text{Prob(outcome } n) = |\langle\phi_n|\psi\rangle|^2 = \langle\psi|P_n|\psi\rangle$$

[5]physicist are used to say that "the wave function collapses"

Remark 1. Since $\sum_j P_j = I$ and $|\psi\rangle$ are normalized we have $\sum_j \text{Prob(outcome } j) = 1$.

Remark 2. When the eigenprojectors are not degenerate these formulas are slightly simpler. If $P_j = |j\rangle\langle j|$ the probability of the outcome j is

$$\text{Prob(outcome } j) = \langle\psi|P_j|\psi\rangle = |\langle j|\psi\rangle|^2$$

and the state just after the measurement is $|j\rangle$.

Consequences for the measurement of observables. This is a very important point because ultimately one really measures physical quantities. The above measurement apparatus $\{P_n\}$ gives the value of any observable of the form $A = \sum_j a_j P_j$. The measurement makes $|\psi\rangle \to |\phi_n\rangle$ for some n. Since $A|\phi_n\rangle = a_n|\phi_n\rangle$ the value of A given by the measurement is precisely a_n when the outcome is n. In particular we can know simultaneously the value of many observables, by measuring them with the same appratus, as long as they have the same eigenspaces. Such observables commute and are sometimes said to be compatible.

The average value that the measurement, on the state $|\psi\rangle$, will yield can be calculated from the probability distribution above. One finds

$$\sum_j a_j \langle\psi|P_j|\psi\rangle = \langle\psi|A|\psi\rangle$$

and the variance is

$$\sum_j a_j^2 \langle\psi|P_j|\psi\rangle - (\sum_j a_j \langle\psi|P_j|\psi\rangle)^2 = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2$$

In practice one uses the right hand side of these two formulas. That is basically all that a theorist can predict.

After a measurement the state vector is reduced $|\psi\rangle \to |\phi_n\rangle$, for some n, and thus the expectation value in the new state (i.e $|\phi_n\rangle$) becomes a_n and the variance 0. This means that if we repeat the same measurement on the same state we will get precisely the value a_n again and again.

We will return to this point when we will consider the Heisenberg uncertainty principle.

Example 7: measurement of photon polarization. Suppose we want to measure the observable $\mathcal{P} = |x\rangle\langle x| - |y\rangle\langle y|$ For this we use the apparatus constituted of an analyzer oriented along x and a detector. This apparatus is the physical realization of the measurement basis. If a photon is detected the state just after the measurement is $|x\rangle$ and if a photon is not detected

(it has been absorbed by the analyzer) the state just after the measurement is $|y\rangle$. The probabilities of these outcomes are

$$\text{Prob(outcome } +1) = |\langle x|\psi\rangle|^2, \;\; \text{Prob(outcome } -1) = |\langle y|\psi\rangle|^2$$

If the initial preparation of the beam is $|\psi\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$ these probabilities are simply $\cos^2\theta$ and $\sin^2\theta$. Suppose that now we rotate the analyzer by an angle γ. This means that we wish to measure the observable $\mathcal{P} = |\gamma\rangle\langle\gamma| - |\gamma_\perp\rangle\langle\gamma_\perp|$. then we can compute again the probabilities of the outcomes

$$\text{Prob(outcome } +1) = |\langle\gamma|\psi\rangle|^2 = \cos^2(\theta-\gamma)$$

$$\text{Prob(outcome } -1) = |\langle\gamma_\perp|\psi\rangle|^2 = \sin^2(\theta-\gamma)$$

Finally let us note that in the first case the measured observable in matrix form is

$$\mathcal{P} = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and in the second

$$\mathcal{P} = \begin{pmatrix} \cos 2\gamma & \sin 2\gamma \\ \sin 2\gamma & -\cos 2\gamma \end{pmatrix} = (\cos 2\gamma)Z + (\sin 2\gamma)X$$

.

Uncertainty principle Suppose that we have a system in a state ψ and we consider two observables A and B. We assume that these have spectral representations

$$A = \sum_j a_j P_j, \quad B = \sum_j b_j Q_j$$

As discussed previously in a general state $|\psi\rangle$ each of these is not fixed but has an average value $\langle\psi|A|\psi\rangle$, $\langle\psi|B|\psi\rangle$ and a standard deviation $\Delta A = \sqrt{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2}$, $\Delta B = \sqrt{\langle\psi|B^2|\psi\rangle - \langle\psi|B|\psi\rangle^2}$. The Heisenberg uncertainty relation states that

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

The interpretation of this inequality as first discussed by Heisenberg is that when $[A,B] \neq 0$ it is not possible to measure A and B simultaneously with infinite precision. If we manage to make $\Delta A = 0$ then we will have $\Delta B = \infty$. The prototypical and most striking example is $A = x$ (position) and $B = p = \frac{\hbar}{i}\frac{\partial}{\partial x}$ (momentum). In this case $\Delta x \Delta p \geq \frac{h}{4\pi}$ and we cannot measure simultaneously with infinite precision the position and the momentum of a

particle: this is not a technological limitation but ultimately a "God given" limitation.

Note that if $[A, B] = 0$ then there exist a common basis of the Hilbert space in which A and B are both diagonal. Then by measuring in this basis, the measurement postulate tells us that both observables can be determined with infinite precision. There is no clash with the uncertainty relation because the right hand side of the inequality vanishes.

There is a related principle called the "entropic uncertainty principle" which we now state. Suppose A and B have non degenerate eigenvalues

$$A = \sum_{n_a} a_{n_a} |n_a\rangle\langle n_a|$$

$$B = \sum_{m_b} b_{m_b} |m_b\rangle\langle m_b|$$

Set

$$H(A) = -\sum_{n_a} p(n_a) \ln p(n_a), \qquad H(B) = -\sum_{m_b} p(m_b) \ln p(m_b)$$

where

$$p(n_a) = |\langle n_a|\psi\rangle|^2, \qquad p(m_b) = |\langle m_b|\psi\rangle|^2$$

We have

$$H(A) + H(B) \geq -2 \ln\Big(\frac{1 + \max_{n_a, m_b} |\langle n_a|m_b\rangle|}{2}\Big)$$

Principle 5: composite quantum systems. Suppose we have two systems $\mathcal{A}$ and $\mathcal{B}$ with Hilbert spaces $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$. The Hilbert space of the composite system $\mathcal{AB}$ is given by the tensor product space

$$\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$$

The states of $\mathcal{AB}$ are vectors $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. The previous postulates apply to the composite system.

This is also a highly non trivial postulate as will be seen from its consequences throughout the course. In a famous paper Einstein, Podolsky, Rosen were the first to make a sharp analysis of its consequences. This has ultimately led to Bell inequalities and to important primitive protocols of quantum information such as teleportation and dense coding.

Example 8. Two photons with polarization degrees of freedom have Hilbert space $\mathbf{C}^2 \otimes \mathbf{C}^2$. Examples of states are $|x\rangle_\mathcal{A} \otimes |y\rangle_\mathcal{B}$ or $|x\rangle_\mathcal{A} \otimes |y\rangle_\mathcal{B} + |\theta\rangle_\mathcal{A} \otimes |\theta\rangle_\mathcal{B}$.

N Qbits live in the space

$$\underbrace{\mathbf{C}^2 \otimes \mathbf{C}^2 \otimes \mathbf{C}^2 \otimes ... \otimes \mathbf{C}^2}_{N \text{ copies}}$$

If $|0\rangle, |1\rangle$ is a canonical basis for $\mathbf{C}^2$, a basis for the composite system is given by

$$|b_1\rangle \otimes |b_2\rangle ... \otimes |b_N\rangle = |b_1, ..., b_N\rangle$$

where $b_i = \{0, 1\}$. There are 2^N such states and they are in one to one correspondence with the 2^N classical bit strings of length N. A general N Qbit state is a linear superposition of the basis states:

$$|\psi\rangle = \sum_{b_1,...,b_N} c_{b_1,...,b_N} |b_1, ..., b_N\rangle$$

where the coefficients $c_{b_1...b_N}$ satisfy

$$\sum_{b_1,...,b_N} |c_{b_1,...,b_N}|^2$$

2.3 Tensor product versus entangled states

States of a composite system $\mathcal{AB}$ lie in $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. We say that a state is a *tensor product state* (or is *not entangled*) if it can be written as

$$|\psi\rangle = |\phi\rangle_\mathcal{A} \otimes |\phi\rangle_\mathcal{B}$$

An entangled state $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ is one for which it is impossible to find $|\phi\rangle_\mathcal{A} \in \mathcal{H}_\mathcal{A}$ and $|\phi\rangle_\mathcal{B} \in \mathcal{H}_\mathcal{B}$ such that ψ is of the tensor product form.

Entangled states have very special correlations between their parts $\mathcal{A}$ and $\mathcal{B}$. These are genuine quantum correlations with no classical counterpart and as we will see later in the course they play a very important role (for example in teleportation). These definitions generalize to multipartite systems.

example 9. Two Qbit system with $\mathcal{A} \otimes \mathcal{B} = \mathbf{C}^2 \otimes \mathbf{C}^2$. Some product states are : $|0\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} = |0, 0\rangle$, $|0\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B} = |0, 1\rangle$, $|1\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} = |1, 0\rangle$, $|1\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B} = |1, 1\rangle$. Two less trivial ones are

$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} + |1\rangle_\mathcal{B}) \otimes |0\rangle_\mathcal{B} = \frac{1}{2}(|0, 0\rangle + |1, 0\rangle)$$

and

$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} + |1\rangle_\mathcal{B}) \otimes \frac{1}{\sqrt{2}}(|0\rangle_\mathcal{B} - |1\rangle_\mathcal{B}) = \frac{1}{2}(|0, 0\rangle - |0, 1\rangle + |1, 0\rangle - |1, 1\rangle)$$

In the same space there are also entangled states that simply *cannot* be written as a tensor product form. For example,

$$\frac{1}{\sqrt{2}}(|0\rangle_{\mathcal{A}} \otimes |0\rangle_{\mathcal{B}} + |1\rangle_{\mathcal{A}} \otimes |1\rangle_{\mathcal{B}}) = \frac{1}{\sqrt{2}}(|0,0\rangle + |1,1\rangle)$$

$$\frac{1}{\sqrt{2}}(|0\rangle_{\mathcal{A}} \otimes |0\rangle_{\mathcal{B}} - |1\rangle_{\mathcal{A}} \otimes |1\rangle_{\mathcal{B}}) = \frac{1}{\sqrt{2}}(|0,0\rangle - |1,1\rangle)$$

$$\frac{1}{\sqrt{2}}(|1\rangle_{\mathcal{A}} \otimes |0\rangle_{\mathcal{B}} + |0\rangle_{\mathcal{A}} \otimes |1\rangle_{\mathcal{B}}) = \frac{1}{\sqrt{2}}(|1,0\rangle + |0,1\rangle)$$

$$\frac{1}{\sqrt{2}}(|0\rangle_{\mathcal{A}} \otimes |0\rangle_{\mathcal{B}} - |1\rangle_{\mathcal{A}} \otimes |0\rangle_{\mathcal{B}}) = \frac{1}{\sqrt{2}}(|0,1\rangle + |1,0\rangle)$$

As we will see these four particular states play a special role and are called Bell states. The reader can check that they form a basis of the 2 Qbit space.

Production of entangled states. Suppose we have a composite system in an initial tensor product state $|\phi\rangle_{\mathcal{A}} \otimes |\chi\rangle_{\mathcal{B}}$. These could for example be two electrons in the spin state $|\uparrow\rangle \otimes |\downarrow\rangle$. If we let them evolve separately and without interaction, the unitary operator for the time evolution is of the form $U_{\mathcal{A}} \otimes U_{\mathcal{B}}$ and

$$U_{\mathcal{A}} \otimes U_{\mathcal{B}}(|\uparrow\rangle \otimes |\downarrow\rangle) = U_{\mathcal{A}}|\uparrow\rangle \otimes U_{\mathcal{B}}|\downarrow\rangle$$

so that the system remains in a tensor product state.

Thus to produce entangled states systems $\mathcal{A}$ and $\mathcal{B}$ must interact at some point in time in order to have an evolution $U_{\mathcal{AB}} \neq U_{\mathcal{A}} \otimes U_{\mathcal{B}}$. With an appropriate interaction we might be able to achieve

$$U_{\mathcal{AB}}(|\uparrow\rangle \otimes |\downarrow\rangle)$$

All known physical interactions are local: this means that in order to interact (in a non-negligible way) two systems must be "close in space-time". In particular if we are presented with an entangled state we know that the two parties have interacted in the past, i.e they have been "sufficiently close in the past".

2.4 No cloning theorem

Classical bits can be copied. For example any latex file can be duplicated or any text can be copied with a (universal) Xerox machine.

Suppose we have a set of quantum states $|\psi\rangle \in \mathcal{H}$ and we want to build a (universal) "quantum Xerox machine" to copy $|\psi\rangle$. This machine should be

able to copy any state of $\mathcal{H}$. A quantum Xerox machine should be described by some unitary operator U (this is true for any physical process except measurement). The Hilbert space is composite $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ where $\mathcal{A}$ is the quantum file to be copied and $\mathcal{B}$ the duplicated file. We start from the state

$$|\psi\rangle \otimes |\text{blank}\rangle$$

and we feed it in the Xerox machine

$$|\psi\rangle \otimes |\text{blank}\rangle \rightarrow \boxed{\text{U}} \rightarrow |\psi\rangle \otimes |\psi\rangle$$

In mathematical terms the question is: can one find a unitary operator such that for a reasonably large set of ψ

$$U(|\psi\rangle \otimes |\text{blank}\rangle) = |\psi\rangle \otimes |\psi\rangle$$

The answer is NO and this is sometimes called the "no cloning theorem". However it is possible to copy a set of orthogonal states with an appropriate U depending on the set.

Proof of no-cloning theorem. Suppose there exists U such that $U^\dagger U = UU^\dagger = 1$ with

$$U(|\phi_1\rangle \otimes |\text{blank}\rangle) = |\phi_1\rangle \otimes |\phi_1\rangle$$
$$U(|\phi_2\rangle \otimes |\text{blank}\rangle) = |\phi_2\rangle \otimes |\phi_2\rangle$$

conjugating the second equation

$$(\langle\phi_2| \otimes \langle\text{blank}|)U^\dagger = \langle\phi_2| \otimes \langle\phi_2|$$

Taking the inner product with the first equation

$$\langle\phi_2| \otimes \langle\text{blank}|U^\dagger U|\phi_1\rangle \otimes |\text{blank}\rangle = (\phi_2| \otimes \langle\phi_2|)(|\phi_1\rangle \otimes |\phi_1\rangle)$$

which implies

$$\langle\phi_2|\phi_1\rangle\langle\text{blank}|\text{blank}\rangle = \langle\phi_2|\phi_1\rangle^2$$

so

$$\langle\phi_2|\phi_1\rangle = 0 \text{ or } \langle\phi_2|\phi_1\rangle = 1$$

We conclude that we cannot copy states $|\phi_1\rangle$ and $|\phi_2\rangle$ that are not identical or orthogonal, with the same U. In fact it is possible to copy a given orthogonal basis . To see this the reader has to construct a unitary operation that does the job.

Non orthogonal states cannot be perfectly distinguished. There are many variants and refinements of the no-cloning theorem. let us just show

one such variant. Suppose we have two states $|\psi\rangle$ and $|\phi\rangle$ and we want to build a (unitary) machine to distinguish them. We seek a U such that

$$U|\psi\rangle \otimes |a\rangle = |\psi\rangle \otimes |v\rangle$$

$$U|\phi\rangle \otimes |a\rangle = |\phi\rangle \otimes |v'\rangle$$

where the outputs $|v\rangle$ and $|v'\rangle$ give some information about $|\psi\rangle$ and $|\phi\rangle$. Taking the inner product of these two equations yields

$$\langle\phi| \otimes \langle a|U^\dagger U|\psi\rangle \otimes |a\rangle = (\langle\phi| \otimes \langle v'|)(|\psi\rangle\rangle \otimes |v\rangle)$$

This implies

$$\langle\phi|\psi\rangle\langle a|a\rangle = \langle\phi|\psi\rangle\langle v'|v\rangle$$

If $|\phi\rangle$ is not orthogonal to $|\psi\rangle$ we have $\langle\phi|\psi\rangle \neq 0$ thus

$$\langle v'|v\rangle = \langle a|a\rangle = 1$$

Thus $|v\rangle = |v'\rangle$ so there is no information in $|v\rangle$ and $|v'\rangle$ distinguishing $|\psi\rangle$ and $|\phi\rangle$.