B Basic concepts from quantum theory

B.1 Introduction

B.1.a Bases

In quantum mechanics certain physical quantities are quantized, such as the energy of an electron in an atom. Therefore an atom might be in certain distinct energy states $|\text{ground}\rangle$, $|\text{first excited}\rangle$, $|\text{second excited}\rangle$, ... Other particles might have distinct states such as spin-up $|\uparrow\rangle$ and spin-down $|\downarrow\rangle$. In each case these alternative states correspond to orthonormal vectors:

$\langle\uparrow|\downarrow\rangle = 0,$
$\langle\text{ground} \mid \text{first excited}\rangle = 0,$
$\langle\text{ground} \mid \text{second excited}\rangle = 0,$
$\langle\text{first excited} \mid \text{second excited}\rangle = 0.$

In general we may express the same state with respect to different bases, such as vertical or horizontal polarization $|\rightarrow\rangle$, $|\uparrow\rangle$; or orthogonal diagonal polarizations $|\nearrow\rangle$, $|\searrow\rangle$.

B.1.b Superpositions of Basis States

One of the unique characteristics of quantum mechanics is that a physical system can be in a superposition of basis states, for example,

$$|\psi\rangle = c_0|\text{ground}\rangle + c_1|\text{first excited}\rangle + c_2|\text{second excited}\rangle,$$

where the c_j are complex numbers, called *(probability) amplitudes*. With respect to a given basis, a state $|\psi\rangle$ is interchangeable with its vector of coefficients, $\mathbf{c} = (c_0, c_1, \ldots, c_n)^{\mathrm{T}}$. When the basis is understood, we can use $|\psi\rangle$ as a name for this vector. This ability of a quantum system to be in many states simultaneously is the foundation of *quantum parallelism*.

As we will see, when we measure the quantum state

$$c_0|E_0\rangle + c_1|E_1\rangle + \ldots + c_n|E_n\rangle$$

with respect to the $|E_0\rangle, \ldots, |E_n\rangle$ basis, we will get the result $|E_j\rangle$ with probability $|c_j|^2$ and the state will "collapse" into state $|E_j\rangle$. Since the probabilities must add to 1, $|c_0|^2 + |c_1|^2 + \cdots + |c_n|^2 = 1$, we know $\||\psi\rangle\| = 1$, that is, the vector is normalized.

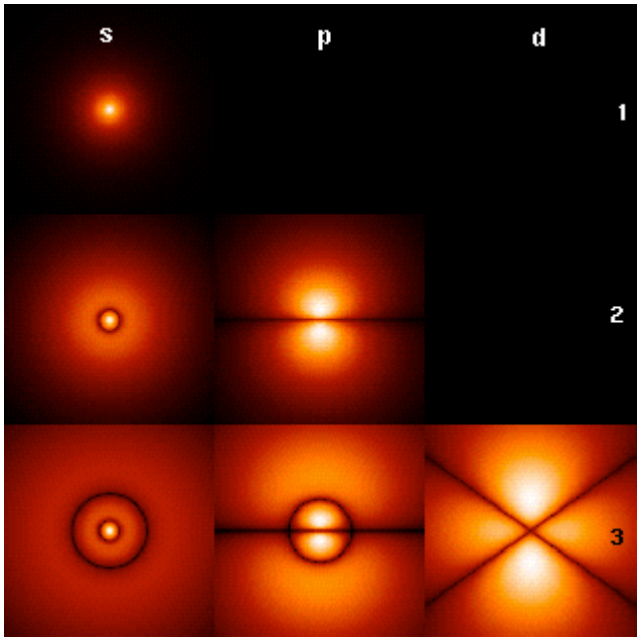

Figure III.1: Probability density of first six hydrogen orbitals. The main quantum number ($n = 1, 2, 3$) and the angular momentum quantum number ($\ell = 0, 1, 2 =$ s, p, d) are shown. (The magnetic quantum number $m = 0$ in these plots.) [fig. from wikipedia commons]

For the purposes of quantum computation, we usually pick two basis states and use them to represent the bits 1 and 0, for example, $|1\rangle = |\text{ground}\rangle$ and $|0\rangle = |\text{excited}\rangle$. We call this the *computational basis*. I've picked the opposite of the "obvious" assignment ($|0\rangle = |\text{ground}\rangle$) just to show that the assignment is arbitrary (just as for classical bits). Note that $|0\rangle \neq \mathbf{0}$, the zero element of the vector space, since $\||0\rangle\| = 1$ but $\|\mathbf{0}\| = 0$. (Thus $\mathbf{0}$ does not represent a physical state, since it is not normalized.)

B.2 Postulates of QM

In this section you will learn the four fundamental postulates of quantum mechanics.[1]

[1] Quotes are from Nielsen & Chuang (2010) unless otherwise specified.

B.2.a Postulate 1: state space

Postulate 1: Associated with any isolated physical system is a *state space*, which is a Hilbert space. The state of the system "is completely defined by its *state vector*, which is a unit vector in the system's state space" (Nielsen & Chuang, 2010). The state vector has to be normalized so that the total probability is 1; it is equivalent to the probability axiom that states that the maximum probability (probability of the whole sample space) = 1.

In previous examples, the state vectors have been finite dimensional, but Hilbert spaces can be infinite dimensional as well. For example, a quantum system might have an unlimited number of energy levels, $|0\rangle, |1\rangle, |2\rangle, \ldots$. If the state of the system is a superposition, $|\psi\rangle = \sum_{k=0}^{\infty} c_k|k\rangle$, then the squared amplitudes must sum to 1, $\sum_{k=0}^{\infty} |c_k|^2 = 1$.

A quantum state $|\psi\rangle$ is often a *wavefunction*, which defines the *probability amplitude* distribution (actually, the probability density function) of some continuous quantity. For example, $|\psi\rangle$ may define the complex amplitude $\psi(\mathbf{r})$ associated with each location $\mathbf{r}$ in space, and $|\Psi\rangle$ may define the complex amplitude of $\Psi(\mathbf{p})$ associated with each momentum $\mathbf{p}$ (see Fig. III.1). Infinite dimensional Hilbert spaces also include spaces of wavefunctions such as these. The inner product of wavefunctions is defined:

$$\langle \phi \mid \psi \rangle = \int_{\mathbb{R}^3} \overline{\phi(\mathbf{r})}\psi(\mathbf{r})\mathrm{d}\mathbf{r}.$$

(For this example we are assuming the domain is 3D space.) Wavefunctions are also normalized, $1 = \||\psi\rangle\|^2 = \int_{\mathbb{R}^3} |\psi(\mathbf{r})|^2 \mathrm{d}\mathbf{r}$. For our purposes, finite dimensional spaces are usually adequate.

In quantum mechanics, global phase has no physical meaning; all that matters is relative phase. In other words, if you consider all the angles around the circle, there is no distinguished 0° (see Fig. III.2). Likewise, in a continuous wave (such as a sine wave), there is no distinguished starting point (see Fig. III.3).

To say all quantum states are normalized is equivalent to saying that their absolute length has no physical meaning. That is, only their *form* (shape) matters, not their absolute size. This is a characteristic of *information*.

Another way of looking at quantum states is as *rays* in a *projective Hilbert space*. A *ray* is an equivalence class of nonzero vectors under the relation, $\phi \cong \psi$ iff $\exists z \neq 0 \in \mathbb{C} : \phi = z\psi$, where $\phi, \psi \neq \mathbf{0}$. That is, global magnitude and phase (r and ϕ in $z = re^{i\phi}$) are irrelevant (i.e., have no physical meaning).

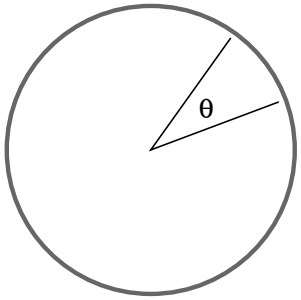

Figure III.2: Relative phase vs. global phase. What matters in quantum mechanics is the relative phase between state vectors (e.g., θ in the figure). Global phase "has no physical meaning"; i.e., we can choose to put the 0° point anywhere we like.

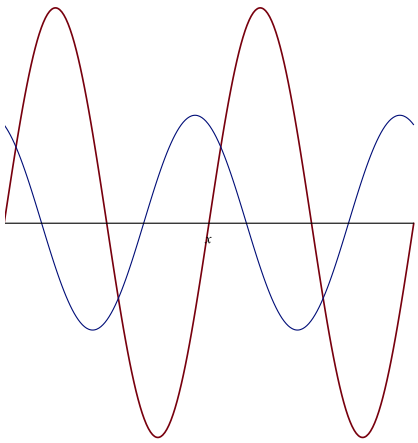

Figure III.3: Relative phase vs. global phase of sine waves. There is no privileged point from which to start measuring absolute phase, but there is a definite relative phase between the two waves.

This is another way of expressing the fact that the *form* is significant, but not the *size*. However, it is more convenient to use normalized vectors in ordinary Hilbert spaces and to ignore global phase.

B.2.b Postulate 2: evolution

Postulate 2: "The evolution of a closed quantum system is described by a unitary transformation" (Nielsen & Chuang, 2010). Therefore a closed quantum system evolves by "complex rotation" of a Hilbert space. More precisely, the state $|\psi\rangle$ of the system at time t is related to the state $|\psi'\rangle$ of the system at time t' by a unitary operator U which depends only on the times t and t',

$$|\psi'\rangle = U(t,t')|\psi\rangle = U|\psi\rangle.$$

This postulate describes the evolution of systems that don't interact with the rest of the world. That is, the quantum system is a dynamical system of relatively low dimension, whereas the environment, including any measurement apparatus, is a thermodynamical system (recall Ch. II, Sec. B).

Dynamics (supplementary) The laws of quantum mechanics, like the laws of classical mechanics, are expressed in differential equations. However, in quantum computation we usually deal with quantum gates operating in discrete time, so it is worth mentioning their relation.

The continuous-time evolution of a closed quantum mechanical system is given by the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle,$$

or more compactly, $i\hbar|\dot{\psi}\rangle = H|\psi\rangle$. H is the Hamiltonian of the system (a fixed Hermitian operator), and $\hbar$ is the reduced Planck constant (often absorbed into H).

Since H is Hermitian, it has a spectral decomposition, $H = \sum_E E|E\rangle\langle E|$, where the normalized $|E\rangle$ are *energy eigenstates* (or *stationary states*) with corresponding energies E. The lowest energy is the *ground state energy*.

In quantum computing, we are generally interested in the discrete-time dynamics of quantum systems. Stone's theorem shows that the solution to the Schrödinger equation is:

$$|\psi(t+s)\rangle = e^{-iHt/\hbar}|\psi(s)\rangle.$$

Therefore define $U(t) \stackrel{\text{def}}{=} \exp(-iHt/\hbar)$; then $|\psi(t+s)\rangle = U(t)|\psi(s)\rangle$. It turns out that U is unitary (Exer. III.5). Hence the evolution of a closed quantum mechanical system from a state $|\psi\rangle$ at time t to a state $|\psi'\rangle$ at time t' can be described by a unitary operator, $|\psi'\rangle = U|\psi\rangle$. Conversely, for any unitary operator U there is a Hermitian K such that $U = \exp(iK)$ (Exer. III.6).

B.2.c Postulate 3: quantum measurement

What happens if the system is no longer closed, that is, if it interacts with the larger environment? In particular, what happens if a quantum system interacts with a much larger measurement apparatus, the purpose of which is to translate a microscopic state into a macroscopic, observable effect? For example, suppose we have a quantum system that can be in two distinct states, for example, an atom that can be in a ground state $|0\rangle$ and an excited state $|1\rangle$. Since they are distinct states, they correspond to orthogonal vectors, $\langle 0 \mid 1\rangle = 0$. Suppose further that we have a measurement apparatus that turns on one light if it measures state $|0\rangle$ and a different light if it measures state $|1\rangle$.

Now consider an atom in a quantum state $|\psi\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$, a superposition of the states $|0\rangle$ and $|1\rangle$. When we measure $|\psi\rangle$ in the computational basis, we will measure $|0\rangle$ with probability $\left|\frac{1}{2}\right|^2 = \frac{1}{4}$, and we will measure $|1\rangle$ with probability $\left|\frac{\sqrt{3}}{2}\right|^2 = \frac{3}{4}$. After measurement, the system is in the state we measured ($|0\rangle$ or $|1\rangle$, respectively); this is the "collapse" of the wavefunction. We depict the possibilities as follows:

$$\begin{aligned} |\psi\rangle &\xrightarrow{1/4} |0\rangle, \\ |\psi\rangle &\xrightarrow{3/4} |1\rangle. \end{aligned}$$

Now consider a more complicated example, a quantum system that can be in three distinct states, say an atom that can be in a ground state $|0\rangle$ or two excited states, $|1\rangle$ and $|2\rangle$. Note that $\langle 0 \mid 1\rangle = \langle 1 \mid 2\rangle = \langle 0 \mid 2\rangle = 0$. Suppose the quantum system is in state $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle$. Further, suppose we have a measurement apparatus that turns on a light if it measures state $|0\rangle$ and does not turn it on otherwise. When we measure $|\psi\rangle$, with probability $\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$ we will measure $|0\rangle$ and after measurement it will collapse to state $|0\rangle$. With probability $\left|\frac{1}{2}\right|^2 + \left|\frac{1}{2}\right|^2 = \frac{1}{2}$ it will not measure state $|0\rangle$ and the

light won't go on. In this case, it will collapse to state $\frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle$, which we get by renormalizing the state measured:

$$\frac{\frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle}{\sqrt{\left|\frac{1}{2}\right|^2 + \left|\frac{1}{2}\right|^2}} = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle.$$

We can depict the possible outcomes as follows:

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|1\rangle + \frac{1}{2}|2\rangle \quad \begin{cases} \xrightarrow{1/2} & |0\rangle \\ \xrightarrow{1/2} & \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle \end{cases}.$$

In other words, we zero out the coefficients of the states we didn't measure and renormalize (because quantum states are always normalized). Now we develop these ideas more formally.

A measurement can be characterized by a set of projectors P_m, for each possible measurement outcome m. In the first example above, the measurement operators are $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1|$. In the second example, the operators are $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1| + |2\rangle\langle 2|$. In the latter case, P_1 projects the quantum state into the subspace spanned by $\{|0\rangle\}$, and P_2 projects the quantum state into the subspace spanned by $\{|1\rangle, |2\rangle\}$. These are *orthogonal subspaces* of the original space (spanned by $\{|0\rangle, |1\rangle, |2\rangle\}$).

Since a measurement must measure some definite state, a projective measurement is a set of projectors $P_1, \ldots, P_N$ satisfying: (1) They project into orthogonal subspaces, so for $m \neq n$ we have $P_m P_n = \mathbf{0}$, the identically zero operator. (2) They are complete, that is, $I = \sum_{m=1}^{N} P_m$, so measurement always produces a result. Projectors are also idempotent, $P_m P_m = P_m$, since if a vector is already projected into the m subspace, projecting it again has no effect. Finally, projectors are Hermitian (self-adjoint), as we can see:

$$P_m^\dagger = \left(\sum_j |\eta_j\rangle\langle\eta_j|\right)^\dagger = \sum_j (|\eta_j\rangle\langle\eta_j|)^\dagger = \sum_j |\eta_j\rangle\langle\eta_j| = P_m.$$

Now we can state Postulate 3.

Postulate 3: Quantum measurements are described by a complete set of orthogonal *projectors*, P_m, for each possible measurement outcome m.

Measurement projects the state into a subspace with a probability given by the squared magnitude of the projection. Therefore, the probability of measurement m of state $|\psi\rangle$ is given by:

$$p(m) = \|P_m|\psi\rangle\|^2 = \langle\psi \mid P_m^\dagger P_m \mid \psi\rangle = \langle\psi \mid P_m P_m \mid \psi\rangle = \langle\psi \mid P_m \mid \psi\rangle. \tag{III.1}$$

This is *Born's Rule*, which gives the probability of a measurement outcome. The measurement probabilities must sum to 1, which we can check:

$$\sum_m p(m) = \sum_m \langle\psi \mid P_m \mid \psi\rangle = \langle\psi| \left(\sum_m P_m\right) |\psi\rangle = \langle\psi \mid I \mid \psi\rangle = \langle\psi \mid \psi\rangle = 1.$$

This follows from the completeness if the projectors, $\sum_m P_m = I$.

For an example, suppose $P_m = |m\rangle\langle m|$, and write the quantum state in the measurement basis: $|\psi\rangle = \sum_m c_m|m\rangle$. Then the probability $p(m)$ of measuring m is:

$$\begin{aligned} p(m) &= \langle\psi \mid P_m \mid \psi\rangle \\ &= \langle\psi|(|m\rangle\langle m|)|\psi\rangle \\ &= \langle\psi \mid m\rangle\langle m \mid \psi\rangle \\ &= \overline{\langle m \mid \psi\rangle}\langle m \mid \psi\rangle \\ &= |\langle m \mid \psi\rangle|^2 \\ &= |c_m|^2. \end{aligned}$$

More generally, the same holds if P_m projects into a subspace, $P_m = \sum_k |k\rangle\langle k|$; the probability is $p(m) = \sum_k |c_k|^2$. Alternatively, we can "zero out" the c_j for the orthogonal subspace, that is, for the $|j\rangle\langle j|$ omitted by P_m. To maintain a total probability of 1, the normalized state vector after measurement is

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}} = \frac{P_m|\psi\rangle}{\|P_m|\psi\rangle\|}.$$

B.2.d Postulate 4: composite systems

Postulate 4: "The state space of a composite physical system is the tensor product of the state spaces of the component physical systems" (Nielsen & Chuang, 2010). If there are n subsystems, and subsystem j is prepared in state $|\psi_j\rangle$, then the composite system is in state

$$|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle = \bigotimes_{j=1}^{n} |\psi_j\rangle.$$

B.3 Wave-particle duality (supplementary)

Some of the capabilities of quantum computation depend on the fact that microscopic objects behave as both particles and waves. To see why, imagine performing the double-slit experiment with three different kinds of objects.

Imagine a stream of classical particles impinging on the two slits and consider the probability of their arriving on a screen. Define $P_j(x)$ to be the probability of a particle arriving at x with just slit j open, and $P_{12}(x)$ to be the probability of a particle arriving at x with both open. We observe $P_{12} = P_1 + P_2$, as expected.

Now consider classical waves, such as water waves, passing through the two slits. The energy I of a water wave depends on the square of its height H, which may be positive or negative. Hence,

$$I_{12} = H_{12}^2 = (H_1 + H_2)^2 = H_1^2 + 2H_1H_2 + H_2^2 = I_1 + 2H_1H_2 + I_2.$$

The $2H_1H_2$ term may be positive or negative, which leads to constructive and destructive interference.

Finally, consider quantum particles. The probability of observing a particle is given by the rule for waves. In particular, the probability P is given by the square of a complex amplitude A:

$$\begin{aligned} P_{12} &= |A_1 + A_2|^2 = \overline{A_1}A_1 + \overline{A_1}A_2 + \overline{A_2}A_1 + \overline{A_2}A_2, \\ &= P_1 + \overline{A_1}A_2 + A_1\overline{A_2} + P_2. \end{aligned}$$

Again, the interference terms $\overline{A_1}A_2 + A_1\overline{A_2}$ can be positive or negative leading to constructive and destructive interference. How does a particle going through one slit "know" whether or not the other slit is open?

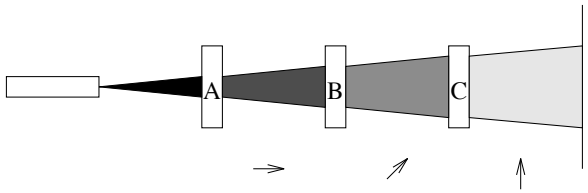

Figure III.4: Fig. from Rieffel & Polak (2000).

B.4 Superposition

A simple experiment demonstrates quantum effects that can not be explained by classical physics (see Fig. III.4). Suppose we have three polarizing filters, A, B, and C, polarized horizontally, 45°, and vertically, respectively. Place the horizontal filter A between a strong light source, such as a laser, and a screen. The light intensity is reduced by one half and the light is horizontally polarized. (Note: Since the light source is unpolarized, i.e., it has all polarizations, the resulting intensity would be much less than one half if the filter allowed only exactly horizontally polarized light through, as would be implied by a sieve model of polarization.) Next insert filter C, polarized vertically, and the intensity drops to zero. This is not surprising, since the filters are cross-polarized. Finally, insert filter B, polarized diagonally, between A and C, and surprisingly some light (about 1/8 intensity) will return! This can't be explained by the sieve model. How can putting in more filters increase the light intensity?

Quantum mechanics provides a simple explanation of the this effect; in fact, it's exactly what we should expect. A photon's polarization state can be represented by a unit vector pointing in appropriate direction. Therefore, arbitrary polarization can be expressed by $a|0\rangle+b|1\rangle$ for any two basis vectors $|0\rangle$ and $|1\rangle$, where $|a|^2 + |b|^2 = 1$.

A polarizing filter measures a state with respect to a basis that includes a vector parallel to its polarization and one orthogonal to it. The effect of filter A is the projector $P_{\mathrm{A}} = |{\rightarrow}\rangle\langle{\rightarrow}|$. To get the probability amplitude, apply it to $|\psi\rangle \stackrel{\text{def}}{=} a|{\rightarrow}\rangle + b|\uparrow\rangle$:

$$p(\mathrm{A}) = |\langle{\rightarrow}|\psi\rangle|^2 = |\langle{\rightarrow}|(a|{\rightarrow}\rangle + b|\uparrow\rangle)|^2 = |a\langle{\rightarrow}|{\rightarrow}\rangle + b\langle{\rightarrow}|\uparrow\rangle|^2 = |a|^2.$$

So with probability $|a|^2$ we get $|{\rightarrow}\rangle$ (recall Eqn. III.1, p. 84). So if the polarizations are randomly distributed from the source, half will get through

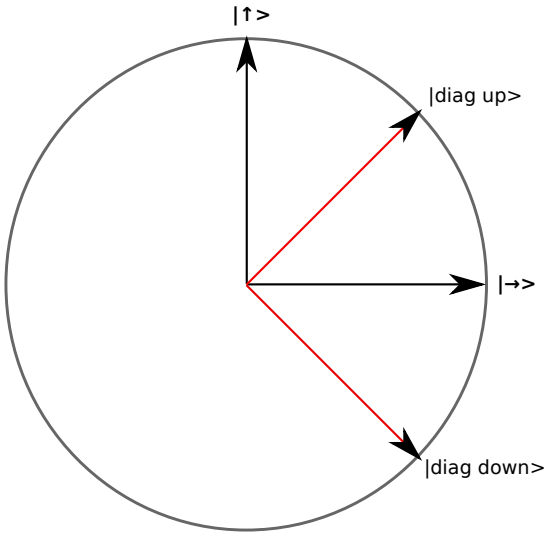

Figure III.5: Alternative polarization bases for measuring photons (black = rectilinear basis, red = diagonal basis). Note $|\nearrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\rightarrow\rangle)$ and $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\nearrow\rangle + |\searrow\rangle)$.

---

and all of them will be in state $|\rightarrow\rangle$. Why one half? Note that $a = \cos\theta$, where θ is the angle between $|\psi\rangle$ and $|\rightarrow\rangle$, and that

$$\langle a^2\rangle = \frac{1}{2\pi}\int_0^{2\pi} \cos^2\theta \, \mathrm{d}\theta = \frac{1}{2}.$$

When we insert filter C we are measuring with the projector $P_{\mathrm{C}} = |\uparrow\rangle\langle\uparrow|$ and the result is 0, as expected:

$$p(\mathrm{AC}) = |\langle\uparrow|\rightarrow\rangle|^2 = 0.$$

Now insert the diagonal filter B between the horizontal and vertical filters A and C. Filter B measures with respect to the projector $\{|\nearrow\rangle, |\searrow\rangle\}$ basis (see Fig. III.5). Transmitted light is given by the projector $P_{\mathrm{B}} = |\nearrow\rangle\langle\nearrow|$. To find the result of applying filter B to the horizontally polarized light emerging from filter A, we must express $|\rightarrow\rangle$ in the diagonal basis:

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\nearrow\rangle + |\searrow\rangle).$$

So if filter B is $|\nearrow\rangle\langle\nearrow|$ we get $|\nearrow\rangle$ photons passing through filter B with probability 1/2:

$$p(\mathrm{B}) = |\langle\nearrow|\rightarrow\rangle|^2 = \left|\langle\nearrow| \left[\frac{1}{\sqrt{2}}(|\nearrow\rangle + |\searrow\rangle\right]\right|^2 = \left|\frac{1}{\sqrt{2}}\langle\nearrow|\nearrow\rangle = \frac{1}{\sqrt{2}}\langle\nearrow|\searrow\rangle\right|^2 = \frac{1}{2}.$$

Hence, the probability of source photons passing though filters A and B is $p(\mathrm{AB}) = p(\mathrm{A})p(\mathrm{B}) = 1/4$.

The effect of filter C, then, is to measure $|\nearrow\rangle$ by projecting against $|\uparrow\rangle$. Note that

$$|\nearrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\rightarrow\rangle).$$

The probability of these photons getting through filter C is

$$|\langle\rightarrow|\nearrow\rangle|^2 = \left|\langle\rightarrow| \left[\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\rightarrow\rangle)\right]\right|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}.$$

Therefore we get $|\rightarrow\rangle$ with another 1/2 decrease in intensity (so 1/8 overall).

B.5 No-cloning theorem

Copying and erasing are two of the fundamental (blackboard-inspired) operations of conventional computing. However, the *No-cloning Theorem* of quantum mechanics states that it is impossible to copy the state of a qubit. To see this, assume on the contrary that we have a unitary transformation U that does the copying, so that $U(|\psi\rangle \otimes |c\rangle) = |\psi\rangle \otimes |\psi\rangle$, where $|c\rangle$ is an arbitrary constant qubit (actually, $|c\rangle$ can be any quantum state). That is, $U|\psi c\rangle = |\psi\psi\rangle$. Next suppose that $|\psi\rangle = a|0\rangle + b|1\rangle$. By the linearity of U:

$$\begin{aligned} U\ |\psi\rangle|c\rangle &= U\ (a|0\rangle + b|1\rangle)|c\rangle \\ &= U(a|0\rangle|c\rangle + b|1\rangle|c\rangle) \quad \text{distrib. of tensor prod.} \\ &= U(a|0c\rangle + b|1c\rangle) \\ &= a(U|0c\rangle) + b(U|1c\rangle) \quad \text{linearity} \\ &= a|00\rangle + b|11\rangle \quad \text{copying property.} \end{aligned}$$

On the other hand, by expanding $|\psi\psi\rangle$ we have:

$$\begin{aligned} U\ |\psi c\rangle &= |\psi\psi\rangle \\ &= (a|0\rangle + b|1\rangle) \otimes (a|0\rangle + b|1\rangle) \\ &= a^2|00\rangle + ba|10\rangle + ab|01\rangle + b^2|11\rangle. \end{aligned}$$

Note that these two expansions cannot be made equal in general, so no such unitary transformation exists. Cloning is possible only in the special cases $a = 0, b = 1$ or $a = 1, b = 0$, that is, only where we know that we are cloning a determinate (classical) basis state. The inability to simply copy a quantum state is one of the characteristics of quantum computation that makes it significantly different from classical computation.

B.6 Entanglement

B.6.a Entangled and decomposable states

The possibility of *entangled quantum states* is one of the most remarkable characteristics distinguishing quantum from classical systems. Suppose that $\mathcal{H}'$ and $\mathcal{H}''$ are the state spaces of two quantum systems. Then $\mathcal{H} = \mathcal{H}' \otimes \mathcal{H}''$ is the state space of the *composite system* (Postulate 4). For simplicity, suppose that both spaces have the basis $\{|0\rangle, |1\rangle\}$. Then $\mathcal{H}' \otimes \mathcal{H}''$ has the

basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. (Recall that $|01\rangle = |0\rangle \otimes |1\rangle$, etc.) Arbitrary elements of $\mathcal{H}' \otimes \mathcal{H}''$ can be written in the form

$$\sum_{j,k=0,1} c_{jk}|jk\rangle = \sum_{j,k=0,1} c_{jk}\ |j'\rangle \otimes |k''\rangle.$$

Sometimes the state of the composite systems can be written as the tensor product of the states of the subsystems, $|\psi\rangle = |\psi'\rangle \otimes |\psi''\rangle$. Such a state is called a *separable*, *decomposable* or *product state*. In other cases the state cannot be decomposed, in which case it is called an *entangled state*

For an example of an entangled state, consider the *Bell state* $|\beta_{01}\rangle$, which might arise from a process that produced two particles with opposite spin (but without determining which is which):

$$|\beta_{01}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \stackrel{\text{def}}{=} |\Phi^+\rangle. \tag{III.2}$$

(The notations $|\beta_{01}\rangle$ and $|\Phi^+\rangle$ are both used.) Note that the states $|01\rangle$ and $|10\rangle$ both have probability 1/2. Such a state might arise, for example, from a process that emits two particles with opposite spin angular momentum in order to preserve conservation of spin angular momentum.

To show that $|\beta_{01}\rangle$ is entangled, we need to show that it cannot be decomposed, that is, that we cannot write $|\beta_{01}\rangle = |\psi'\rangle \otimes |\psi''\rangle$, for two state vectors $|\psi'\rangle = a_0|0\rangle + a_1|1\rangle$ and $|\psi''\rangle = b_0|0\rangle + b_1|1\rangle$. Let's try a separation or decomposition:

$$|\beta_{01}\rangle \stackrel{?}{=} (a_0|0\rangle + a_1|1\rangle) \otimes (b_0|0\rangle + b_1|1\rangle).$$

Multiplying out the RHS yields:

$$a_0b_0|00\rangle + a_0b_1|01\rangle + a_1b_0|10\rangle + a_1b_1|11\rangle.$$

Therefore we must have $a_0b_0 = 0$ and $a_1b_1 = 0$. But this implies that either $a_0b_1 = 0$ or $a_1b_0 = 0$ (as opposed to $1/\sqrt{2}$), so the decomposition is impossible.

For an example of a decomposable state, consider $\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$. Writing out the product $(a_0|0\rangle + a_1|1\rangle) \otimes (b_0|0\rangle + b_1|1\rangle)$ as before, we require $a_0b_0 = a_0b_1 = a_1b_0 = a_1b_1 = \frac{1}{2}$. This is satisfied by $a_0 = a_1 = b_0 = b_1 = \frac{1}{\sqrt{2}}$, therefore the state is decomposable.

In addition to Eq. III.2, the other three Bell states are defined:

$$|\beta_{00}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \stackrel{\text{def}}{=} |\Psi^+\rangle, \tag{III.3}$$

$$|\beta_{10}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \stackrel{\text{def}}{=} |\Psi^-\rangle, \tag{III.4}$$

$$|\beta_{11}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \stackrel{\text{def}}{=} |\Phi^-\rangle. \tag{III.5}$$

The Ψ states have two identical qubits, the Φ states have opposite qubits. The $+$ superscript indicates they are added, the $-$ that they are subtracted. The general definition is:

$$|\beta_{xy}\rangle = \frac{1}{\sqrt{2}}(|0, y\rangle + (-1)^x|1, \neg y\rangle).$$

Remember this useful formula! The Bell states are orthogonal and in fact constitute a basis for $\mathcal{H}' \otimes \mathcal{H}''$ (exercise).

B.6.b EPR PARADOX

The EPR Paradox was proposed by Einstein, Podolsky, and Rosen in 1935 to show problems in quantum mechanics. Our discussion here will be informal.

Suppose a source produces an entangled *EPR pair* (or *Bell state*) $|\Psi^+\rangle = |\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, and the entangled particles are sent to Alice and Bob. If Alice measures her particle and gets $|0\rangle$, then that collapses the state to $|00\rangle$, and so Bob will have to get $|0\rangle$ if he measures his particle. Likewise, if Alice happens to get $|1\rangle$, Bob is also required to get $|1\rangle$ if he measures. This happens instantaneously (but it does not permit faster-than-light communication, as explained below).

One explanation is that there is some internal state in the particles that will determine the result of the measurement. Both particles have the same internal state. Such *hidden-variable theories* of quantum mechanics assume that particles are "really" in some definite state and that superposition reflects our ignorance of its state. However, they cannot explain the results of measurements in different bases. In 1964 John Bell showed that any local hidden variable theory would lead to measurements satisfying a certain inequality (Bell's inequality). Actual experiments, which have been conducted over tens of kilometers, violate Bell's inequality. Thus local hidden variable theories cannot be correct.

Another explanation is that Alice's measurement affects Bob's (or vice versa, if Bob measures first). These are called *causal theories*. According to relativity theory, however, in some frames of reference Alice's measurement comes first, and in other frames, Bob's comes first. Therefore there is no consistent cause-effect relation. This is why Alice and Bob cannot use entangled pairs to communicate.

B.7 Uncertainty principle (supplementary)

You might be surprised that the famous Heisenberg uncertainty principle is not among the postulates of quantum mechanics. That is because it is not a postulate, but a theorem, which can be proved from the postulates. This section is optional, since the uncertainty principle is not required for quantum computation.

B.7.a Informally

The uncertainty principle states a lower bound on the precision with which certain pairs of variables, called *conjugate variables*, can be measured. These are such pairs as position and momentum, and energy and time. For example, the same state can be represented by the wave function $\psi(x)$ as a function of space and by $\phi(p)$ as a function of momentum. The most familiar version of the Heisenberg principle, limits the precision with which location and momentum can be measured simultaneously: $\Delta x\ \Delta p \geq \hbar/2$, where the reduced Plank constant $\hbar = h/2\pi$, where h is Planck's constant.

It is often supposed that the uncertainty principle is a manifestation of the *observer effect*, the inevitable effect that measuring a system has on it, but this is not the case. "While it is true that measurements in quantum mechanics cause disturbance to the system being measured, this is most emphatically *not* the content of the uncertainty principle."(Nielsen & Chuang, 2010, p. 89)

Often the uncertainty principle is a result of the variables representing measurements in two bases that are Fourier transforms of each other. Consider an audio signal $\psi(t)$ and its Fourier transform $\Psi(\omega)$ (its spectrum). Note that ψ is a function of time, with dimension t, and its spectrum Ψ is a function of frequency, with dimension t^{-1}. They are reciprocals of each other, and that is always the case with Fourier transforms. Simultaneous measurement in the time and frequency domains obeys the uncertainty relation $\Delta t \Delta\omega \geq 1/2$. (For more details on this, including an intuitive explanation, see MacLennan (prep, ch. 6).)

Time and energy are also conjugate, as a result of the de Broglie relation, according to which energy is proportional to frequency: $E = h\nu$ (ν in Hertz, or cycles per second) or $E = \hbar\omega$ (ω in radians per second). Therefore simultaneous measurement in the time and energy domains obeys the uncertainty principle $\Delta t \Delta E \geq \hbar/2$.

More generally, the observables are represented by Hermitian operators P, Q that do not commute. That is, to the extent they do not commute, to that extent you cannot measure them both (because you would have to do either PQ or QP, but they do not give the same result). The best interpretation of the uncertainty principle is that if you set up the experiment multiple times, and measure the outcomes, you will find

$$2\ \Delta P\ \Delta Q \geq |\langle [P, Q] \rangle|,$$

where P and Q are conjugate observables. (The commutator $[P, Q]$ is defined below, Def. B.2, p. 96.)

Note that this is a *purely mathematical* result (proved in Sec. B.7.b). Any system obeying the QM postulates will have uncertainty principles for every pair of non-commuting observables.

B.7.b Formally

In this section we'll derive the uncertainty principle more formally. Since it deals with the variances of measurements, we begin with their definition. To understand the motivation for these definitions, suppose we have a quantum system (such as an atom) that can be in three distinct states $|\text{ground}\rangle$, $|\text{first excited}\rangle$, $|\text{second excited}\rangle$ with energies e_0, e_1, e_2, respectively. Then the *energy observable* is the operator

$$\begin{aligned} E \ &= \ e_0 |\text{ground}\rangle\langle\text{ground}| + e_1 |\text{first excited}\rangle\langle\text{first excited}| \\ &\quad + e_2 |\text{second excited}\rangle\langle\text{second excited}|, \end{aligned}$$

or more briefly, $\sum_{m=0}^{2} e_m |m\rangle\langle m|$.

Definition B.1 (observable) *An* observable M *is a Hermitian operator on the state space.*

An observable M has a spectral decomposition (Sec. A.2.g):

$$M = \sum_{m=1}^{N} e_m P_m,$$

where the P_m are *projectors* onto the eigenspaces of M, and the eigenvalues e_m are the corresponding measurement results. The projector P_m projects

into the eigenspace corresponding to eigenvalue e_m. (For projectors, see Sec. A.2.d.) Since an observable is described by a Hermitian operator M, it has a spectral decomposition with real eigenvalues, $M = \sum_{m=1}^{N} e_m |m\rangle\langle m|$, where $|m\rangle$ is the measurement basis. Therefore we can write $M = UEU^\dagger$, where $E = \mathrm{diag}(e_1, e_2, \ldots, e_N)$, $U = (|1\rangle, |2\rangle, \ldots, |N\rangle)$, and

$$U^\dagger = (|1\rangle, |2\rangle, \ldots, |N\rangle)^\dagger = \begin{pmatrix} \langle 1| \\ \langle 2| \\ \vdots \\ \langle N| \end{pmatrix}.$$

$U^\dagger$ expresses the state in the measurement basis and U translates back. In the measurement basis, the matrix for an observable is a diagonal matrix: $E = \mathrm{diag}(e_1, \ldots, e_N)$. The probability of measuring e_m is

$$p(m) = \langle \psi \mid P_m^\dagger P_m \mid \psi \rangle = \langle \psi \mid P_m P_m \mid \psi \rangle = \langle \psi \mid P_m \mid \psi \rangle.$$

We can derive the mean or expectation value of an energy measurement for a given quantum state $|\psi\rangle$:

$$\begin{aligned} \langle E \rangle &\stackrel{\text{def}}{=} \mu_E \stackrel{\text{def}}{=} \mathcal{E}\{E\} \\ &= \sum_m e_m p(m) \\ &= \sum_m e_m \langle \psi \mid m \rangle \langle m \mid \psi \rangle \\ &= \sum_m \langle \psi| \, e_m |m\rangle\langle m| \, |\psi\rangle \\ &= \langle \psi| \left(\sum_m e_m |m\rangle\langle m| \right) |\psi\rangle \\ &= \langle \psi \mid E \mid \psi \rangle. \end{aligned}$$

This formula can be used to derive the standard deviation σ_E and variance σ_E^2, which are important in the uncertainty principle:

$$\begin{aligned} \sigma_E^2 &\stackrel{\text{def}}{=} (\Delta E)^2 \stackrel{\text{def}}{=} \mathrm{Var}\{E\} \\ &= \mathcal{E}\{(E - \langle E \rangle)^2\} \\ &= \langle E^2 \rangle - \langle E \rangle^2 \\ &= \langle \psi \mid E^2 \mid \psi \rangle - (\langle \psi \mid E \mid \psi \rangle)^2. \end{aligned}$$

Note that E^2, the matrix E multiplied by itself, is also the operator that measures the square of the energy, $E^2 = \sum_j e_m^2 |m\rangle\langle m|$. (This is because E is diagonal in this basis; alternately, E^2 can be interpreted as an operator function.)

We now proceed to the derivation of the uncertainty principle.[2]

Definition B.2 (commutator) *If $L, M : \mathcal{H} \to \mathcal{H}$ are linear operators, then their* commutator *is defined:*

$$[L, M] = LM - ML. \tag{III.6}$$

Remark B.1 *In effect, $[L, M]$ distills out the non-commutative part of the product of L and M. If the operators commute, then $[L, M] = \mathbf{0}$, the identically zero operator. Constant-valued operators always commute ($cL = Lc$), and so $[c, L] = \mathbf{0}$.*

Definition B.3 (anti-commutator) *If $L, M : \mathcal{H} \to \mathcal{H}$ are linear operators, then their* anti-commutator *is defined:*

$$\{L, M\} = LM + ML. \tag{III.7}$$

If $\{L, M\} = \mathbf{0}$, we say that L and M anti-commute*, $LM = -ML$.*

See B.2.c (p. 82) for the justification of the following definitions.

Definition B.4 (mean of measurement) *If M is a Hermitian operator representing an observable, then the mean value of the measurement of a state $|\psi\rangle$ is*

$$\langle M \rangle = \langle \psi \mid M \mid \psi \rangle.$$

Definition B.5 (variance and standard deviation of measurement) *If M is a Hermitian operator representing an observable, then the variance in the measurement of a state $|\psi\rangle$ is*

$$\mathrm{Var}\{M\} = \langle (M - \langle M \rangle^2) \rangle = \langle M^2 \rangle - \langle M \rangle^2.$$

As usual, the standard deviation *ΔM of the measurement is defined*

$$\Delta M = \sqrt{\mathrm{Var}\{M\}}.$$

[2]The following derivation is from MacLennan (prep, ch. 5).

Proposition B.1 *If L and M are Hermitian operators on $\mathcal{H}$ and $|\psi\rangle \in \mathcal{H}$, then*

$$4\langle\psi \mid L^2 \mid \psi\rangle \, \langle\psi \mid M^2 \mid \psi\rangle \geq |\langle\psi \mid [L, M] \mid \psi\rangle|^2 + |\langle\psi \mid \{L, M\} \mid \psi\rangle|^2.$$

More briefly, in terms of average measurements,

$$4\langle L^2\rangle\langle M^2\rangle \geq |\langle[L, M]\rangle|^2 + |\langle\{L, M\}\rangle|^2.$$

Proof: Let $x + iy = \langle\psi \mid LM \mid \psi\rangle$. Then,

$$\begin{aligned} 2x &= \langle\psi \mid LM \mid \psi\rangle + (\langle\psi \mid LM \mid \psi\rangle)^* \\ &= \langle\psi \mid LM \mid \psi\rangle + \langle\psi \mid M^\dagger L^\dagger \mid \psi\rangle \\ &= \langle\psi \mid LM \mid \psi\rangle + \langle\psi \mid ML \mid \psi\rangle \quad \text{since } L, M \text{ are Hermitian} \\ &= \langle\psi \mid \{L, M\} \mid \psi\rangle. \end{aligned}$$

Likewise,

$$\begin{aligned} 2iy &= \langle\psi \mid LM \mid \psi\rangle - (\langle\psi \mid LM \mid \psi\rangle)^* \\ &= \langle\psi \mid LM \mid \psi\rangle - \langle\psi \mid ML \mid \psi\rangle \\ &= \langle\psi \mid [L, M] \mid \psi\rangle. \end{aligned}$$

Hence,

$$\begin{aligned} |\langle\psi \mid LM \mid \psi\rangle|^2 &= 4(x^2 + y^2) \\ &= |\langle\psi \mid [L, M] \mid \psi\rangle|^2 + |\langle\psi \mid \{L, M\} \mid \psi\rangle|^2. \end{aligned}$$

Let $|\lambda\rangle = L|\psi\rangle$ and $|\mu\rangle = M|\psi\rangle$. By the Cauchy-Schwarz inequality, $\||\lambda\rangle\| \, \||\mu\rangle\| \geq |\langle\lambda \mid \mu\rangle|$ and so $\langle\lambda \mid \lambda\rangle \, \langle\mu \mid \mu\rangle \geq |\langle\lambda \mid \mu\rangle|^2$. Hence,

$$\langle\psi \mid L^2 \mid \psi\rangle \, \langle\psi \mid M^2 \mid \psi\rangle \geq |\langle\psi \mid LM \mid \psi\rangle|^2.$$

The result follows.

□

Proposition B.2 *Prop. B.1 can be weakened into a more useful form:*

$$4\langle\psi \mid L^2 \mid \psi\rangle \, \langle\psi \mid M^2 \mid \psi\rangle \geq |\langle\psi \mid [L, M] \mid \psi\rangle|^2,$$

or $4\langle L^2\rangle\langle M^2\rangle \geq |\langle[L, M]\rangle|^2$

Proposition B.3 (uncertainty principle) *If Hermitian operators P and Q are measurements (observables), then*

$$\Delta P \, \Delta Q \geq \frac{1}{2} |\langle \psi \mid [P, Q] \mid \psi \rangle|.$$

That is, $\Delta P \, \Delta Q \geq |\langle [P, Q] \rangle|/2$. So the product of the variances is bounded below by the degree to which the operators do not commute.

Proof: Let $L = P - \langle P \rangle$ and $M = Q - \langle Q \rangle$. By Prop. B.2 we have

$$\begin{aligned} 4 \operatorname{Var}\{P\} \operatorname{Var}\{Q\} &= 4 \langle L^2 \rangle \langle M^2 \rangle \\ &\geq |\langle [L, M] \rangle|^2 \\ &= |\, \langle [P - \langle P \rangle, Q - \langle Q \rangle] \rangle \,|^2 \\ &= |\langle [P, Q] \rangle|^2. \end{aligned}$$

Hence,

$$2 \, \Delta P \Delta Q \geq |\langle [P, Q] \rangle|$$

□