Quantum Information Theory: MAT 4953 XLS 5983.
Lesson 03 (Sep 06-08), Fall 2022.

José A. Morales E., UTSA Math and Physics & A. Depts.

Abstract

Communications Theory. Physical qubits: spinning particles and photons.

1. Communications Theory.

The problem is the following: there will be communication between a sender (called Alice) and a receiver (named Bob). However, there might also be noise interfering across the communication. We will have two events:

- Event A: Alice selects a message from the set $\{a_i\}$ of possible messages.
- Event B: Bob detects a message from the set $\{b_j\}$ of possibilities.

1.1. Shannon's work on Communications.

Claude Shannon wrote the paper "A Mathematical Theory of Communication", fundamental to Classical Information Theory. He derived two theorems:

1. Noiseless coding theorem: how much we can compress a message and still can read it.
2. Noisy channel coding theorem: how much redundancy is needed to correct errors.

1.1.1. Noiseless coding theorem

The key idea of the theorem is that the number of bits needed to efficiently encode a message is related to Entropy.

Remark 1. Most messages have some redundancy, therefore they can be compressed. For example, if I write "QNTM NFRMTN THRY RLS", what do you think this message is? Answer: QUANTUM INFORMATION THEORY RULES.

Example 2. Some examples of easy compression in the English language are:

- queen –> qeen (since q is usually followed by u)

Email address: jose.morales4@utsa.edu (José A. Morales E., UTSA Math and Physics & A. Depts.)

- colour −> color (difference between British and American English)
- the −> th (a 3-letter word beginning with th: most certainly the word is the)

Let's consider a message with N bits. Since for each bit we have 2 possible states (0 and 1),

$$\underbrace{010110001\cdots011}_{N \quad \text{bits}}$$

we can express 2^N possible messages with N bits.

We will assume for simplicity that any bit can take the value 0 with a probability p,therefore the value 1 has probability $1-p$. Moreover, we'll assume each bit takes its value independently of all other bits.

If we want to calculate the probability that any particular message has n zeros, we have that

$$P(n) = \frac{N!}{(N-n)!n!} p^n (1-p)^{N-n}$$

When N is large, the most probable scenario is that the number of zeros will be close to Np (under the assumptions by a law of large numbers argument). So we will restrict to consider only "typical messages", for which $n = Np$.

The number of different messages with $n = Np$ zeros is only (by combinatorics arguments)

$$W = \frac{N!}{(N-n)!n!}$$

If we use Stirling's approximation for the factorial (incidentally, it appears commonly on Statistical Mechanics), then

$$N! \approx \sqrt{2\pi} N^{N+1/2} \exp(-N), \quad N \ggg 1$$

and

$$\ln(N!) \approx \ln(\sqrt{2\pi} N^{N+1/2} \exp(-N)) \approx \ln(\sqrt{2\pi}) + (N+1/2)\ln(N) - N \approx N\ln(N) - N$$

Then

$$W = \frac{N!}{(N-n)!n!}$$

so

$$\ln W = \ln N! - \ln\{(N-n)!\} - \ln\{n!\} \approx$$
$$N\ln(N) - N - (N-n)\ln\{(N-n)\} + (N-n) - (n\ln(n) - n) =$$
$$= N\ln(N) - (N-n)\ln(N-n) - n\ln(n) =$$
$$= N\{\ln(N) - \ln(N-n)\} + n\{\ln(N-n) - \ln(n)\} =$$
$$= N\{-\ln(1-\frac{n}{N})\} + n\{-\ln(\frac{1}{N/n-1})\} =$$

$$= N\{-\ln(1-p) + p\ln(1/p - 1)\} =$$

$$= N[-\ln(1-p) + p\ln\{(1-p)/p\}] =$$

$$= N[-\ln(1-p) + p\ln\{(1-p)\} - p\ln p] =$$

$$= N[-(1-p)\ln(1-p) - p\ln p] = NH_e(p)$$

remembering that $p = n/N$. Equivalently, in base 2, we have

$$\log_2 W \approx NH(p)$$

where incidentally Stirling's approximation of the logarithm base 2 of the factorial is

$$\log_2(N!) = N\log_2 N - \frac{N}{\ln 2}$$

We remember that $H(p)$ is the entropy associated with the probabilities p and $1-p$ from our previous studies. We have then that

$$W \approx 2^{NH(p)}$$

estimates approximately the number of different "typical messages".

Therefore, we only need $NH(p)$ bits, rather than N bits for typical messages. We can prove that $H(p) \leq 1$, which implies that we have a "compression factor" of $H(p)$.

This result is precisely Shannon's noiseless coding theorem.

Remark 3. We don't need all the approximations taken at this point (that might have seemed restrictive). Allowing more typical messages than the ones with Np zeros the result still holds, particularly n differing from pN by a small amount, it can be derived that the typical number of messages is

$$W \approx 2^{N(H(p)+\delta)}, \quad \lim_{N\to\infty} \delta = 0.$$

Remark 4. Shannon's theorem is of existence: it indicates that there exists an optimal coding scheme, but it doesn't tell us exactly what this scheme is or how to find it (not a proof by construction).

Example 5. 4 letter "alphabet" problem: We will consider a simple alphabet of 4 letters (which natural phenomena sounds similar btw?), for simplicity, labeled as A, B, C, D.

Each letter abovementioned occurs in a message with an associated probability specified below:

$$P(A) = 1/2, \quad P(B) = 1/4, \quad P(C) = 1/8 = P(D)$$

What is the information (entropy) associated with the set of messages then?

Solution 6. We have that

$$H = -\sum_{i=1}^{4} p_i \log_2(p_i) =$$

$$H = -[\frac{1}{2}\log_2(\frac{1}{2}) + \frac{1}{4}\log_2(\frac{1}{4}) + 2\{\frac{1}{8}\log_2(\frac{1}{8})\}]$$

$$H = [\frac{1}{2} + \frac{2}{4} + 2\{\frac{3}{8}\}] = 1 + 3/4 = 7/4$$

so $H = 7/4$ bits (we are in base 2).

Remark 7. If a message in this alphabet is N characters long, it requires $2N$ bits (since each character needs 2 bits for its encoding, meaning A=00, B=01, C=10, D=11, for example). However, the result above indicates that we can compress the $2N$ bits message to a message of $7N/4$ bits. How can we reduce the length of our messages though? For example, we can replace common elements of a message. The most common letter in our alphabet is A (since it has the highest probability), so we can replace A by short sequences and leave longer sequences for the least likely letters C and D. Intuitively, that is what we do when we text "c u @ utsa l8r".

We can use the following coding to obtain the abovementioned: A=0, B=10, C=110, D=111 (again, the least likely events are the ones that give the most information, remember this?).

BTW, it can be shown that any message coded with the scheme above can be uniquely decoded to recover the original sequence of letters (a bijective correspondence, one-to-one and onto).

Now, if we calculate the average number of bits necessary to code a sequence of N letters in this encoding we have that it's

$$N(\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + 2 \times \frac{1}{8} \times 3) = \frac{7}{4}N = HN$$

exactly the Shannon's bound based on entropy considerations! Since Shannon's calculation is a bound, this means that it's impossible to have a better coding (meaning no shorter messages).

1.1.2. Noisy coding theorem

It's usual to have noise entering communications. This noise will introduce errors. The way we can try to compensate for this is by introducing "redundancy". It is thought that this is why spoken languages have plenty of redundancy. For example, if one writes "WNTM NARMQN THRS SLS", it's very hard to guess what it originally meant, as redundancy was removed before errors were introduced. On the other hand, if we left the vowels in before and then added errors after, we would have "WANTFM INAORMAQION THEORS SULES" (QUANTUM INFORMATION THEORY RULES, with errors). This is easier to read, even if there were errors both in vowels and consonants.

Problem 8. Amount of redundancy needed.

Let's assume each bit in a message can be flipped (with a given probability q), which would give an error.

Theorem 9. *Shannon's noisy coding theorem: on average, we need at least* $N_0/(1-H(q))$ *bits to encode one of* 2^{N_0} *equally probable messages, with*

$$H(q) = -[q \log q + (1-q)\log(1-q)]$$

the associated entropy with the single-bit error probability q.

Proof. Look at Barnett's "Quantum Information" 2009 book for the proof. □

Essentially, if we first take out the redundancy according to the previous "noiseless coding theorem", then we need to reintroduce this redundancy to make up for possible errors.

2. Physical qubits: spinning particles and photons.

2.1. Stern-Gerlach experiment (1922)

Look at the following illustration of this experiment.

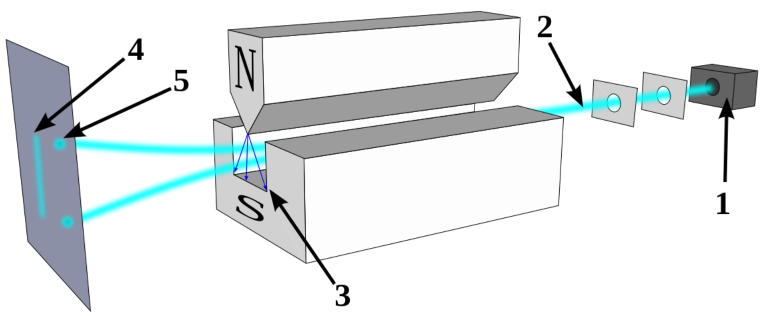


Fig. By Tatoute - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=34095239

We have neutral silver atoms, which have one "unpaired" electron (with respect to spin, the intrinsic angular momentum), whose spin gives the silver atoms a certain angular momentum. Therefore the atoms behave as little magnets. We will apply a nonhomogeneous magnetic field to the atoms, which will be under the potential

$$U = -\vec{\mu} \cdot \vec{B}$$

where $\vec{\mu}$ is the magnetic moment (classically, with magnitude μ =area times current, so we have a current loop "model" for the spin) and therefore will feel the force

$$\vec{F} = -\nabla U$$

and since the force is in the $z-$direction, we have

$$F_z = -\partial_z U = \mu \frac{\partial B_z}{\partial z}$$

since our magnetic field is non-homogeneous in the $z-$direction.

Before performing the experiment, people expected a vertical line, which would stand for a continuous distribution of spin directions. However, the surprising experimental findings were that Stern & Gerlach got two bands as in the figure below, which indicated that spin (intrinsic angular momentum) is quantized. Also, you can orient the magnets at any angle, and you will find that the spin is quantized along that direction too.

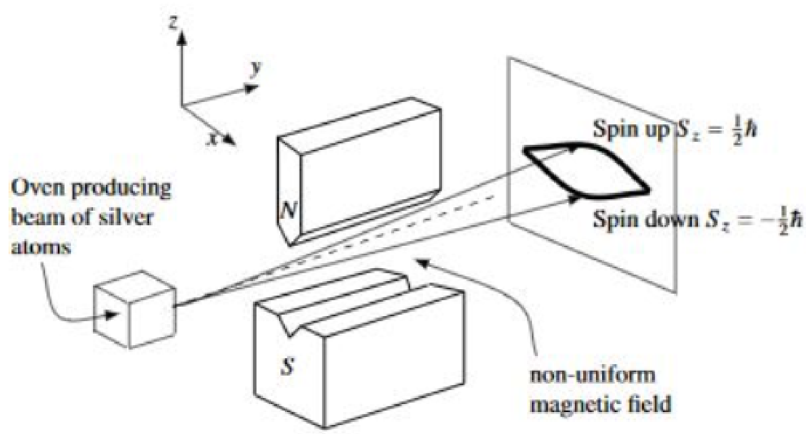


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

2.1.1. Experimental Facts

Fact 10. *If we measure the spin of a "spin-1/2" particle, we either get $+\hbar/2$ or $-\hbar/2$ (not any other in between), with*

$$\hbar = \frac{h}{2\pi} \approx 1.0546 \times 10^{-34} \text{Joules}$$

the "reduced Planck's constant" (called "ħ bar" by notation, whereas h is Planck's constant).

Fact 11. *This experimental result does not depend on the direction along which the measurement is performed.*

2.2. Succesive Measurements of Spin

Let's assume we have the experimental capabilities to perform the measurements of Stern-Gerlach twice (succesively), as in the diagram below.

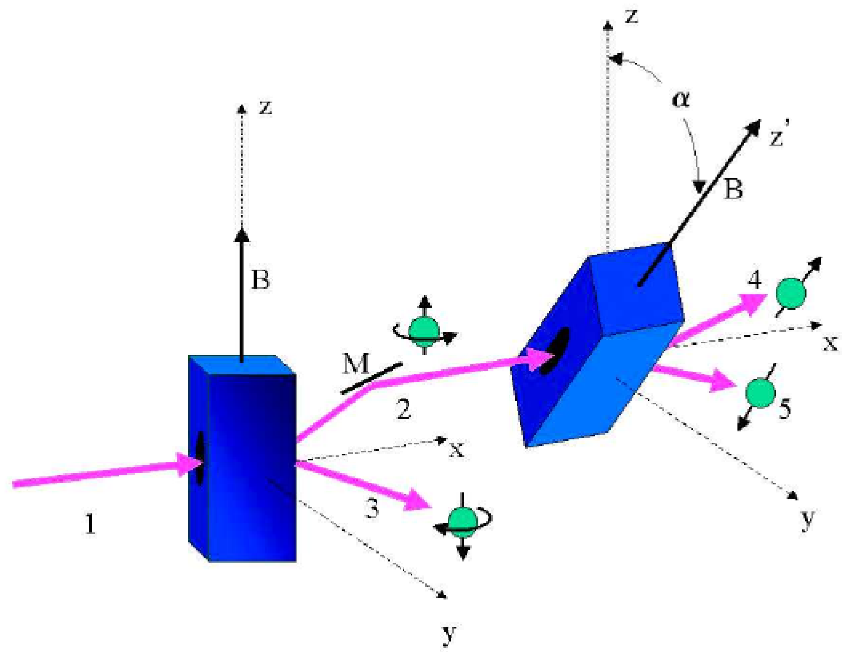


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

We perform the first measurement. After it, we only keep the spin along the $+z$ direction. Notice that the second measurement is defined by the angle α. If $\alpha = 90$ degrees, the probability of spin z' (notice the tilted axis) versus $-z'$ is fifty-fifty percent. However, the individual outcome of each particle is random. For any α in general, the probability of spin along z' is $\cos^2(\alpha/2)$, and the probability of spin along $-z'$ is $\sin^2(\alpha/2)$. Naturally, we have

$$\cos^2(\alpha/2) + \sin^2(\alpha/2) = 1$$

2.3. Light: wave/particle theories

The first studies of light (classical optics) understood it in a "particle" fashion (to make sense of reflection, for example). However, later on it was understood that light has wave properties (Huygens' superposition principle). J. C. Maxwell (1864) developed a mathematical theory of light, explaining it as an electromagnetic wave. That means that a light wave is an oscillatory wave of both electric and magnetic fields, which in the free-space (in the vacuum, absent of air for example, as somewhere in outer space) propagate following the wave equations

$$\partial_t \vec{E} = c^2 \nabla^2 \vec{E}$$
$$\partial_t \vec{B} = c^2 \nabla^2 \vec{B}$$

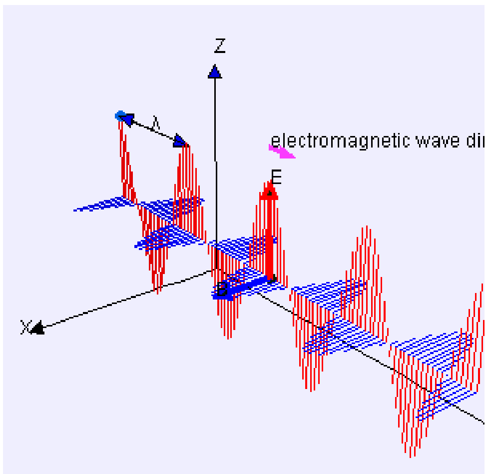


Fig. By Lookang - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16874302

Look at the Wikipedia animation related to the figure above (will be included in Blackboard).

Now, what we know is that light (and matter too actually) has wave-like properties and particle-like properties (as in photons) at the same time. Essentially, the concepts of wave and particle are abstractions of the physical properties we observe objects in the natural phenomena to have. Regarding light in particular, it has wave properties such as interference (as when you see water waves interfere in a pond, the ocean, or in a water tank). One particular experiment that makes evidence these interference properties for light is the so-called "double-slit experiment" (by Thomas Young in 1801). Look at the figure below:

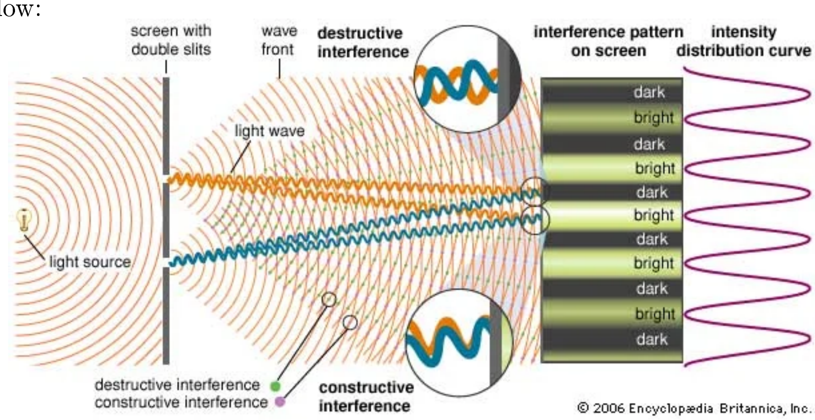


Fig. Young's double-slit experiment (from Encyclopædia Britannica, Inc.)

We have a light source, and we put at a distance a screen with double slits (the slits separated by a distance d). Now, by Huygens' principle ("every point on a wavefront is itself the source of spherical wavelets") the light passing through the two slits become synchronized light sources (since they emanate from the common original light source), but because the slits are separated by a distance d their respective waves will interfere. A screen will show the interference pattern (constructive-destructive-constructive-...-etc.). This experiment verified the wave-properties of light at its time: if light (in the scale of the observation) behaved as a particle, then (assuming a small probability of collisions) no destructive interference would be observed, only constructive (think about it).

Let's study the double-slit experiment in more detail. Look at the following diagram:

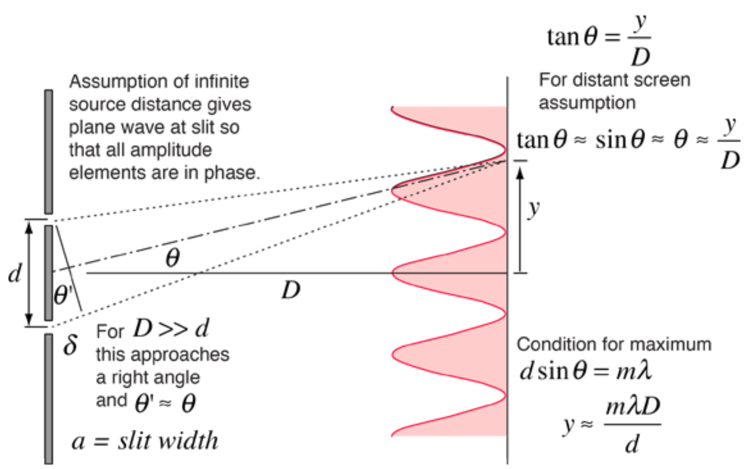


Fig. From http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/slits.html

We assume the slits are very narrow. In that case, they both become "point sources", so each one of them is now a source of a spherical wave of the form

$$A\exp(ikr)/r$$

Now, if we consider the screen, the superposition of the waves from both sources generates the net contribution of

$$A(\exp(ikr_1)/r_1 + \exp(ikr_2)/r_2)$$

with r_1 the distance from the observation point in the screen to slit 1, and r_2 the distance from the same observation point in the screen to slit 2. If we assume that the slits are very far away from the screen, we can approximate

$$\exp(ikr_1)/r_1 + \exp(ikr_2)/r_2 \approx \frac{1}{r_{avg}}(\exp(ikr_1) + \exp(ikr_2))$$

and if we now play with the complex exponentials,

$$\exp(ikr_1) + \exp(ikr_2) = e^{ik(r_1+r_2)/2}[e^{ik(r_1-r_2)/2} + e^{ik(r_2-r_1)/2}] =$$

$$= e^{ik(r_1+r_2)/2}[e^{ik(r_1-r_2)/2} + e^{-ik(r_1-r_2)/2}] = e^{ik(r_1+r_2)/2}2\cos(k(r_2-r_1)/2)$$

This is approximately the net wave amplitude. Therefore, the light intensity (norm squared of the amplitude) is proportional to

$$\propto 4\cos^2(\frac{k(r_2-r_1)}{2})$$

and we now simply estimate, given the figure above, that

$$r_2 - r_1 \approx d\sin(\theta') \approx d\sin\theta \approx d\frac{y}{D}$$

2.3.1. Polarization of electro-magnetic waves

As we mentioned on the go, formally speaking light also behaves like a particle in quantum mechanics. This particle is called a photon (photo is light in Greek). Photons are quantum-mechanical particles with spin-1 (spin is a quantum property that we've mentioned before). So, when one measures the photon angular momentum, we could get either $\pm\hbar$ or 0 but since photons (light) have no mass, the 0 state can't be observed. Somehow then, since only $\pm\hbar$ can be observed, photon seems like a "spin-1/2" particle in some aspects, but not in others, such as behavior under rotation.

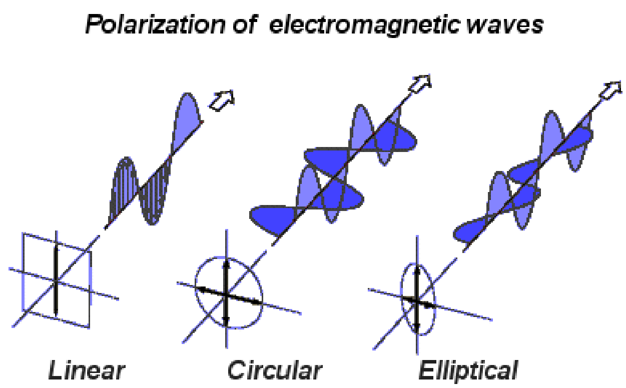


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

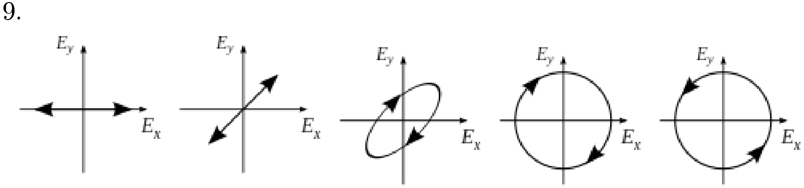


Fig. Linear polarizations, elliptic, and circular ones, respectively. From https://en.wikipedia.org/wiki/Polarization_(waves)

For the moment we'll focuse on the electric field vector.

Let's assume we have a plane EM wave, traveling along a direction $\vec{k}$, with $\widehat{k} = \widehat{z}$. Then the electric field vector is given by

$$\vec{E} = \Re\{\vec{E}_0 e^{i(kz-wt)}\}$$

with the fixed (constant) $\vec{E_0}$ given by

$$\vec{E}_0 = E_{0x}\widehat{i} + E_{0y}\widehat{j}$$

We can write it as a column vector, called the Jones vector:

$$\vec{E}_0 = \begin{pmatrix} E_{0x} \\ E_{0y} \end{pmatrix}$$

There are different cases of Jones vectors. Some examples appear in the following figure (with a normalized Jones vector), together with its respective polarization in the left:

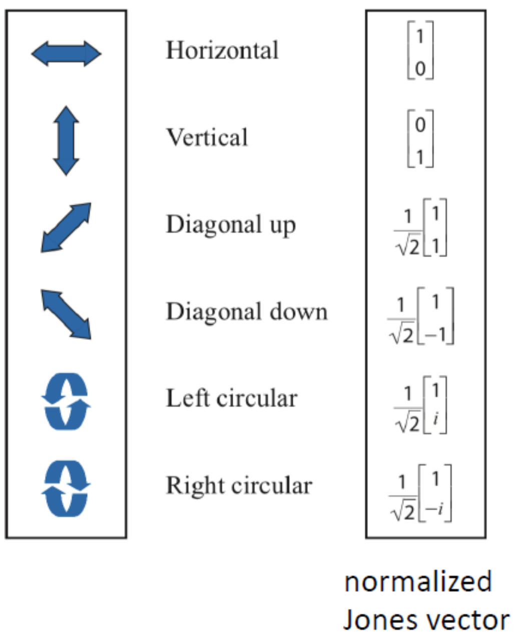


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

Remark 12. The convention above for left and right circular polarization is not unique. Some other sources exchange the two definitions abovementioned.

If we think of "crossed polarizers" (we have two subsequent polarizers whose transmission directions are orthogonal to each other), we would ideally eliminate transmission at the end, as in the figure below.

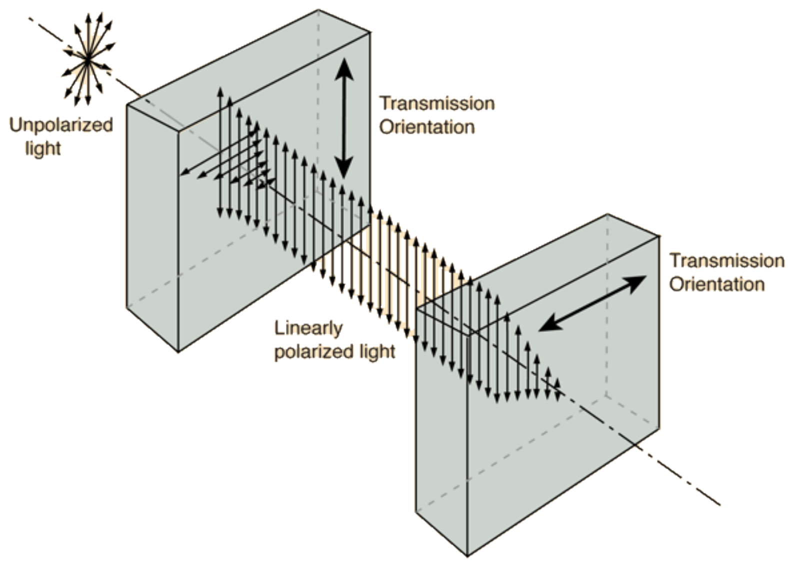


Fig. From http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/polcross.html

Now, in general, if there's an angle θ between the light polarization and the polaroid (transmission direction), the fraction of light intensity transmitted through this single polaroid is $\cos^2(\theta)$.

2.3.2. *Beam splitters*

This idealized physical device appears very frequently in Optics experiments.

A polarizing beam splitter is a device that lets horizontally polarized light to be transmitted, but it reflects the vertically polarized light (btw, part of it is lost by absorption), as in the figure below:

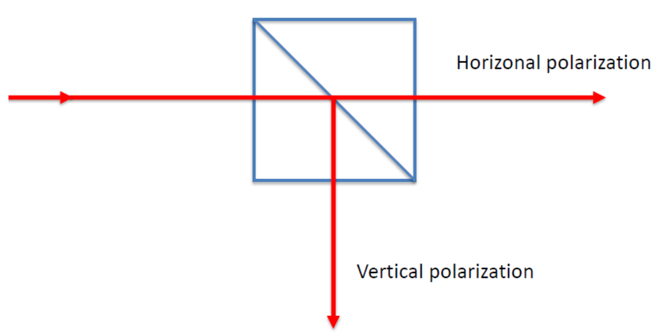


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

2.3.3. Waveplates

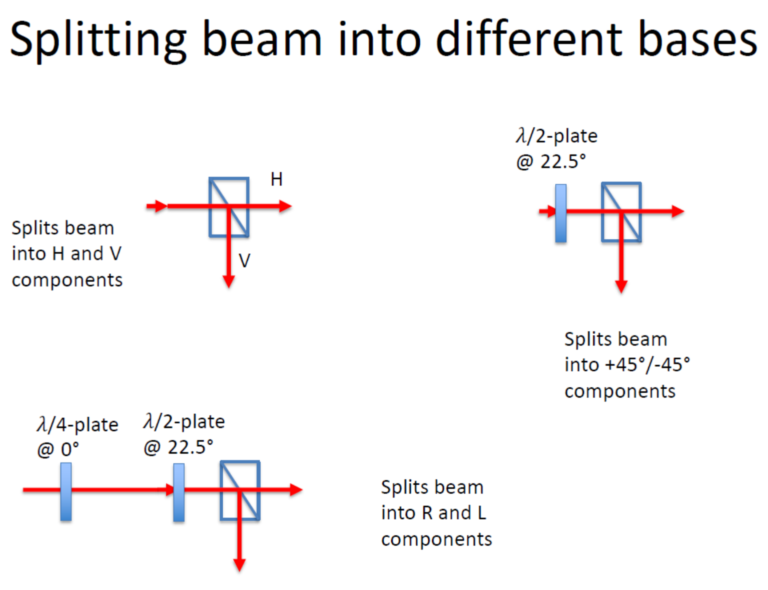
Waveplates are made from materials called birefringent (with different refraction indices for different polarizations, therefore anisotropic), such as quartz or calcite. Light only conserves its polarization if it's linearly polarized along either the crystal's optic axis direction (ordinary direction) or the orthogonal one (extraordinary direction). Because it introduces different phase shifts along these directions, it can rotate the polarization. Two particular cases are:

- $\lambda/2$-plate with the optical axis @ 22.5° to horizontal interconverts between horizontal and +45° polarizations, as well as between the vertical and -45° polarizations

- $\lambda/4$-plate with optical axis phase oriented horizontally or vertically interconverts between the circular and +/- 45° polarizations

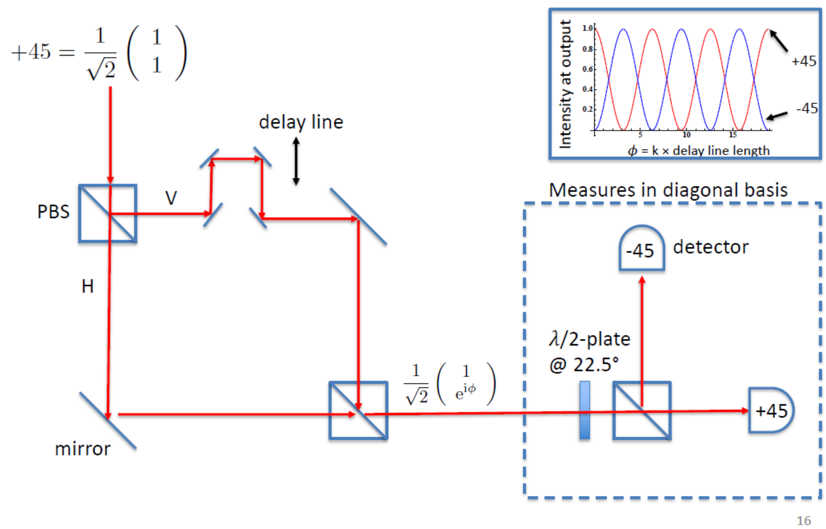
2.3.4. Splitting beam into different bases

Look at the following Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.



2.3.5. Mach-Zehnder Interferometer

Look at the following experimental diagram, where it's assumed the light is polarized in the diagonal-up direction (the interference pattern shows both diagonal up/down cases):



The array of mirrors in the upper part of the experimental setting delays the light in its path w.r.t. the lower part of the setting (this is why we denote it as the "delay line").

The delay line causes a phase difference ϕ between the beams. The wave-plate at the end splits the beam into ± 45 degrees components, still with a phase difference. The outcome of the MZ experiment is then interference fringes (detected by an interferometer), dependant on the phase $\phi \propto$delay line length. If we calculate the probabilities in terms of the ± 45 degrees basis, we have to project onto these vectors, so

$$+\vec{45}^\circ \cdot \vec{E}_0 = \frac{1}{\sqrt{2}}(1,1)\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1\\\exp(i\phi)\end{pmatrix} = \frac{1+\exp(i\phi)}{2} = \exp(i\phi/2)\frac{\exp(-i\phi/2)+\exp(i\phi/2)}{2}$$

$$\implies P(+45^\circ) = \cos^2(\phi/2)$$

$$-\vec{45}^\circ \cdot \vec{E}_0 = \frac{1}{\sqrt{2}}(1,-1)\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1\\\exp(i\phi)\end{pmatrix} = \frac{1-\exp(i\phi)}{2} = \exp(i\phi/2)\frac{\exp(-i\phi/2)-\exp(i\phi/2)}{2}$$

$$\implies P(-45^\circ) = \sin^2(\phi/2)$$

2.3.6. Quantum Theory of Light

Different experimental evidence conducted physicists to conclude that, at certain scale, light behaves as a particle. The only way to explain certain experiments was to conclude that the energy of a light wave is quantized into "light particles" called photons, following Planck's relation

$$E = h\nu = \hbar\omega$$

where ν is the frequency of the light wave, ω is the angular frequency, h is Planck's constant, and the reduced Planck's constant is

$$\hbar = \frac{h}{2\pi}$$

In Quantum Theory, a similar relation between wavelength and momentum is satisfied. It is called De Broglie's relation and it states that the momentum of the quantum particle representing a wave (in this case light) is

$$p = h/\lambda = \hbar k$$

In a nutshell, light has a "wave-particle duality": it behaves like either a wave or a particle, depending on the scale of observation of the related physical phenomena. Moreover, this wave-particle duality is not exclusive of light. It applies also to particles like electrons. This means that, if we are in the appropriate scale, we can try the double-slit experiment with electrons and we'll obtain an interference pattern as in Young's experiment for light.

Now, light of a single frequency is quantized in the sense that

$$E = \hbar\omega$$

is the minimum energy for EM waves of that frequency and they represent the presence of a single particle. Given the wave-particle duality, we can try now to send a single photon (light particle) into an experiment where waves outcomes are expected if the scale is appropriate, such as MZ (a quantum interferometer). Although only one of either detectors measures a photon, the probability of detection is actually given by a "classical" (light wave) intensity pattern. This confirms the wave particle duality. There's an appearance of quantum complementarity as well: we cannot know the particle (path) and the wave (phase) properties at the same time.

2.4. A summary of Quantum Mechanics

2.4.1. Quantum state vectors in Dirac notation

We recall that Dirac's notation is specific for the Hilbert space (spaces with inner product, possibly infinite dimensional) to which the quantum state vectors (called ket vectors belong). There's the following "bracket" notation

$$\langle \text{bra} | \text{ket} \rangle$$

where "bra vectors" $\langle \text{bra}|$ represent the functionals of the Hilbert space, and "ket vectors" $|\text{ket}\rangle$ represent the actual vectors in the space. The bracket notation above denotes the inner product (the correspondance between a ket vector and its respective bra vector representing a functional giving a scalar via the inner product is given by Riesz representation theorem).

If for example, our ket (state vector) is given by

$$|\Psi\rangle = \psi_x |x\rangle + \psi_y |y\rangle \to \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

one understands that $|x\rangle, |y\rangle$ are "basis kets", and $\psi_x, \quad \psi_y$ are "probability amplitudes", as we have mentioned, the latter being complex numbers. In this example, we can represent them as a complex-valued 2D vector.

Because Quantum Mechanics follows Schroedinger equation, which is a linear PDE, it admits the Superposition Principle (together with a normalization condition). For the moment we haven't said what x, y represent (it could be position or not). Let's assume in the example above we have a 2-dimensional orthonormal basis, so

$$\langle x|y\rangle = 0 = \langle y|x\rangle\,, \quad \langle x|x\rangle = 1 = \langle y|y\rangle$$

or in the vector notation,

$$(1,0)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 = (0,1)\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (1,0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 = (0,1)\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and even in a "coordinate direction" notation, we can represent the inner products above as

$$\widehat{i}\cdot\widehat{j} = 0 = \widehat{j}\cdot\widehat{i}, \quad \widehat{i}\cdot\widehat{i} = 1 = \widehat{j}\cdot\widehat{j}.$$

Noticeably, if we want to get the $x-$amplitude in our example, it is given by

$$\langle x|\Psi\rangle = \psi_x\,\langle x|x\rangle + \psi_y\,\langle x|y\rangle = \psi_x$$

given the orthonormality conditions.

2.4.2. Jones vectors and qubits

There's an equivalence between a normalized Jones vector and its respective qubit in dirac notation. Again, recalling that in the case of left/right circular polarization the convention is not universal (as they're sometimes swapped), examples of these equivalences are:

- Horizontal: normalized Jones vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, equivalent qubit in Dirac notation $|0\rangle$
- Vertical: normalized Jones vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, equivalent qubit in Dirac notation $|1\rangle$
- Diagonal up: normalized Jones vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}/\sqrt{2}$, equivalent qubit in Dirac notation $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$
- Diagonal down: normalized Jones vector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}/\sqrt{2}$, equivalent qubit in Dirac notation $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$
- Left circular: normalized Jones vector $\begin{pmatrix} 1 \\ i \end{pmatrix}/\sqrt{2}$, equivalent qubit in Dirac notation $\frac{|0\rangle+i|1\rangle}{\sqrt{2}}$
- Diagonal down: normalized Jones vector $\begin{pmatrix} 1 \\ -i \end{pmatrix}/\sqrt{2}$, equivalent qubit in Dirac notation $\frac{|0\rangle-i|1\rangle}{\sqrt{2}}$

2.4.3. Scalar product of two quantum vectors

More specifically in terms of how to perform the scalar/inner product (bracket, bra-ket) given the components of the state vectors, if

$$|\Psi\rangle = \psi_x |x\rangle + \psi_y |y\rangle \rightarrow \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

and

$$|\Phi\rangle = \phi_x |x\rangle + \phi_y |y\rangle \rightarrow \begin{pmatrix} \phi_x \\ \phi_y \end{pmatrix}$$

then

$$\langle\Phi|\Psi\rangle = (\phi_x^*, \phi_y^*)\begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \phi_x^*\psi_x + \phi_y^*\psi_y$$

As you can see, given the ket (state vector) $|\Phi\rangle$, its respective bra vector (functional) is

$$\langle\Phi| = \phi_x^* \langle x| + \phi_y^* \langle y| \rightarrow (\phi_x^*, \phi_y^*)$$

which is akin to say that the "dual" of $|\Phi\rangle$ is $\langle\Phi|$ and it's obtained by conjugating and transposing $|\Phi\rangle$.

We recall that in quantum mechanics (bounded) state vectors have a normalization condition (as their coefficients represent probability amplitudes):

$$\langle\Psi|\Psi\rangle = (\psi_x^*, \psi_y^*)\begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \psi_x^*\psi_x + \psi_y^*\psi_y = |\psi_x|^2 + |\psi_y|^2 = 1$$

2.4.4. Hilbert space

A Hilbert space $\mathbb{V}$ is a linear vector space (either finite or infinite dimensional) in which a scalar/inner product is defined (incidentally, this lets us define angles between infinite dimensional vectors).

Again, the kets denote vectors living on our Hilbert space. In Quantum Mechanics, the Hilbert space is over the complex numbers and therefore the respective scalar product is complex.

A single qubit belongs to a 2-dimensional Hilbert space. When we have N qubits, they live in a 2^N-dimensional Hilbert space.

2.4.5. Adjoint space

As we mentioned, Dirac introduced as well a "bra" notation for functionals:

$$\langle\Psi| = \psi_x^* \langle x| + \psi_y^* \langle y| \rightarrow (\psi_x^*, \psi_y^*)$$

In principle, bras don't live in the same Hilbert space $\mathbb{V}$ as kets. Bras live in the adjoint space $\mathbb{V}^\dagger$. However, for every ket (vector) $|\Psi\rangle$ there exists a unique bra (functional) $\langle\Psi|$ representing it in the inner product (this is precisely Riesz representation theorem). If you have done some QM, Linear Algebra over the complex numbers, or Functional Analysis, you might know that bras are the "adjoints" (Hermitian conjugates) of the kets: if

$$|\Psi\rangle = \psi_x |x\rangle + \psi_y |y\rangle \rightarrow \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

then

$$\langle\Psi| = |\Psi\rangle^\dagger = \psi_x^* \langle x| + \psi_y^* \langle y| \to (\psi_x^*, \psi_y^*)$$

and so now we are comfortable expressing the "bra-ket"

$$\langle\Phi|\Psi\rangle = (\phi_x^*, \phi_y^*)\begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \phi_x^*\psi_x + \phi_y^*\psi_y = \langle\Psi|\Phi\rangle^*$$

2.4.6. Change of basis

Sometimes we have to introduce a change of basis. For example, if we rotate our coordinate system xy into a new one $x'y'$ by an angle θ, then

$$|x\rangle = \cos(\theta)\,|x'\rangle - \sin(\theta)\,|y'\rangle$$
$$|y\rangle = \sin(\theta)\,|x'\rangle + \cos(\theta)\,|y'\rangle$$

If we want to express our ket in the new basis, we have

$$|\Psi\rangle = \begin{pmatrix} \langle x'|\Psi\rangle \\ \langle y'|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \psi_{x'} \\ \psi_{y'} \end{pmatrix}$$

Example 13. Let's say the rotation angle is 45 degrees. In this particular case, we know that

$$\sin(\theta) = \frac{1}{\sqrt{2}} = \cos(\theta)$$

So (in our "interferometer" notation), we have that

$$|x'\rangle = |{+45^\circ}\rangle$$
$$|y'\rangle = |{-45^\circ}\rangle$$

Therefore, if

$$|\Psi\rangle = \psi_x\,|x\rangle + \psi_y\,|y\rangle \to \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

then, since

$$|x\rangle = \frac{1}{\sqrt{2}}\,|x'\rangle - \frac{1}{\sqrt{2}}\,|y'\rangle$$
$$|y\rangle = \frac{1}{\sqrt{2}}\,|x'\rangle + \frac{1}{\sqrt{2}}\,|y'\rangle$$

we get

$$|\Psi\rangle = \psi_x(\frac{1}{\sqrt{2}}\,|x'\rangle - \frac{1}{\sqrt{2}}\,|y'\rangle) + \psi_y(\frac{1}{\sqrt{2}}\,|x'\rangle + \frac{1}{\sqrt{2}}\,|y'\rangle)$$
$$|\Psi\rangle = \frac{\psi_x + \psi_y}{\sqrt{2}}\,|x'\rangle + \frac{\psi_y - \psi_x}{\sqrt{2}}\,|y'\rangle = \frac{\psi_x + \psi_y}{\sqrt{2}}\,|{+45^\circ}\rangle + \frac{\psi_y - \psi_x}{\sqrt{2}}\,|{-45^\circ}\rangle$$

So, for example, if $\psi_x = 0$ and $\psi_y = 1$, then

$$|\Psi\rangle = |y\rangle = \frac{1}{\sqrt{2}}\,|{+45^\circ}\rangle + \frac{1}{\sqrt{2}}\,|{-45^\circ}\rangle$$

therefore there's a 50:50 probability of vertically polarized photons passing a ± 45 degrees polaroid.

2.4.7. Projector and identity operators

If we want to project into the 1D space generated by $|x\rangle$, we apply to the state vector $|\Psi\rangle$ the following operator,

$$\widehat{P}_x = |x\rangle \langle x|$$

which is called the projector onto $x-$subspace. Likewise, the projector onto $y-$subspace is

$$\widehat{P}_y = |y\rangle \langle y|$$

Both of the projectors above are indeed operators, since when they act over $|\Psi\rangle$ the outcome is the following, for example when projecting onto the $x-$subspace,

$$\widehat{P}_x |\Psi\rangle = |x\rangle \langle x| |\Psi\rangle = \langle x|\Psi\rangle |x\rangle = \psi_x |x\rangle$$

If our Hilbert space $\mathbb{V}$ is 2D and generated by the basis vectors $\{|x\rangle, |y\rangle\}$, then the identity operator can be represented as

$$\widehat{\mathbb{I}} = |x\rangle \langle x| + |y\rangle \langle y|$$

and for any set (2D in this case) that is a basis for our Hilbert space $\mathbb{V}$,such as $\{|L\rangle, |R\rangle\}$ (regarding polarization) or $\{|+45^\circ\rangle, |-45^\circ\rangle\}$, we can also express the identity operator as

$$\widehat{\mathbb{I}} = |L\rangle \langle L| + |R\rangle \langle R|$$

$$\widehat{\mathbb{I}} = |+45^\circ\rangle \langle +45^\circ| + |-45^\circ\rangle \langle -45^\circ|$$

In all cases, it is because the identity operator is expanding over all the 2 elements of the respective basis and getting their components by projection. For example,

$$\widehat{\mathbb{I}} |\Psi\rangle = |x\rangle \langle x| |\Psi\rangle + |y\rangle \langle y| |\Psi\rangle = \langle x|\Psi\rangle |x\rangle + \langle y|\Psi\rangle |y\rangle = \psi_x |x\rangle + \psi_y |y\rangle = |\Psi\rangle$$

which is literally what is expected of the identity operator (it's essentially an identity matrix).

2.4.8. Change of basis

If we have that

$$|\Psi\rangle = \psi_x |x\rangle + \psi_y |y\rangle$$

and we want to change the expression of $|\Psi\rangle$ in terms of the ± 45degrees basis, we simply make use of the identity operator as follows:

$$|\Psi\rangle = \psi_x \widehat{\mathbb{I}} |x\rangle + \psi_y \widehat{\mathbb{I}} |y\rangle$$

$$|\Psi\rangle = \psi_x (|+45^\circ\rangle \langle +45^\circ| + |-45^\circ\rangle \langle -45^\circ|) |x\rangle + \psi_y (|+45^\circ\rangle \langle +45^\circ| + |-45^\circ\rangle \langle -45^\circ|) |y\rangle$$

and now, using the fact that

$$\langle +45^\circ| |x\rangle = \frac{1}{\sqrt{2}}, \quad \langle -45^\circ| |x\rangle = \frac{-1}{\sqrt{2}}$$

$$\langle +45^\circ|\,|y\rangle = \frac{1}{\sqrt{2}}, \quad \langle -45^\circ|\,|y\rangle = \frac{1}{\sqrt{2}}$$

then

$$|\Psi\rangle = \psi_x(|+45^\circ\rangle\,\langle +45^\circ|\,|x\rangle + |-45^\circ\rangle\,\langle -45^\circ|\,|x\rangle) + \psi_y(|+45^\circ\rangle\,\langle +45^\circ|\,|y\rangle + |-45^\circ\rangle\,\langle -45^\circ|\,|y\rangle)$$

$$|\Psi\rangle = \psi_x(|+45^\circ\rangle\,\frac{1}{\sqrt{2}} + |-45^\circ\rangle\,\frac{-1}{\sqrt{2}}) + \psi_y(|+45^\circ\rangle\,\frac{1}{\sqrt{2}} + |-45^\circ\rangle\,\frac{1}{\sqrt{2}})$$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}[\psi_x(|+45^\circ\rangle - |-45^\circ\rangle) + \psi_y(|+45^\circ\rangle + |-45^\circ\rangle)]$$

$$|\Psi\rangle = \frac{\psi_x + \psi_y}{\sqrt{2}}\,|+45^\circ\rangle + \frac{\psi_y - \psi_x}{\sqrt{2}}\,|-45^\circ\rangle$$

2.4.9. Transmission probability through a single polaroid in Quantum Mechanics

We have that

$$|\Psi\rangle = \psi_x\,|x\rangle + \psi_y\,|y\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

Let's think our state vector represents a single photon. Because it is the particle with the quantum of energy associated to lightwaves in that frequency, when a single photon interacts with a polaroid, it either passes the polaroid or is absorbed by it. We cannot measure a "fraction of a photon", as it has the minimal quantum of energy. We can account for the probability of transmission through, say, an $x-$polaroid, which is given by

$$|\psi_x|^2 = |\,\langle x|\Psi\rangle\,|^2$$

This is why we have called the coefficients of the state vector probability amplitudes: the $x-$amplitude, is, for example,

$$\langle x|\Psi\rangle = \psi_x\,\langle x|x\rangle + \psi_y\,\langle x|y\rangle = \psi_x$$

2.4.10. Quantum Mechanics as "Amplitude Mechanics"

QM is a wave theory (which can explain "particle-like" discrete quantized properties via eigenvalue problems) that follows a linear PDE known as the Schroedinger equation. As a wave theory, we have the following

- Probability of outcome of an experiment is squared magnitude of the respective quantum amplitude
- Quantum amplitude for experiment composed of 2 alternate experiments is the SUM of the amplitudes for each individual experiment (superposition)
- Quantum amplitude for experiment made of 2 succesive experiments is the PRODUCT of amplitudes for each individual experiment (independence of probability / amplitudes for independent events).

2.4.11. Some Postulates of Quantum Mechanics

1. Hilbert Space postulate: The state of a physical system is represented by a vector $|\Psi\rangle$ in a Hilbert space (denoted as $\mathbb{V}$) of possible states. All vectors that represent physical (bounded, nonzero) states are normalized.
2. Measurement Hypothesis: An idealized measurement apparatus is associated with a respective orthonormal basis $\{|v_i\rangle\}$ of states of the system. After the measurement, the apparatus will randomly, with probability

   $$P_i = |\langle v_i|\Psi_{\text{before}}\rangle|^2$$

   point to one of the states $|v_i\rangle$, where we denote $|\Psi_{\text{before}}\rangle$ the state of the system before the measurement. The system, if it's not destroyed, will be projected ("collapsed") onto the state $|v_i\rangle$.

Remark 14. The probabilistic aspect of QM regarding measurement outcomes needs to make use of the abstraction of "ensembles" (a concept of Statistical Mechanics where it's imagined that one can study the state of a system by creating idealized physical "copies" of the system) over which to perform the measurement different times in order to talk about possible outcomes probabilistically.

2.4.12. Einstein's quote: "God does not play dice".

The measurement process abovementioned is called "projective measurement". There are other types of measurements, called weak measurements, that have been lately achieved experimentally.

The relation $P_i = |\langle v_i|\Psi_{\text{before}}\rangle|^2$ is called Born's rule (Max Born). In quantum mechanics, measurement seems inherently probabilistic. Einstein didn't like this aspect of QM, so he came up with a thought-experiment (gedanken-experiment) trying to show a paradox in QM (Einstein-Podolsky-Rosen, 1935), however up to experiments (such as Alain Aspect ones in 1984) QM has stood the test.

2.4.13. Randomness

Example 15. Light polarized at $+45°$, incident over polaroid transmitting vertically polarized light: half absorbed, half transmitted.

Example 16. Light arriving at polarizing beam splitter: divided into 2 polarizations (modulus squared of amplitudes).

Problem 17. Quantum case of a single photon: superposition of both polarizations before measurement. Measure of half-photon never happens, so one of two polarization outcomes occurs after measurement.

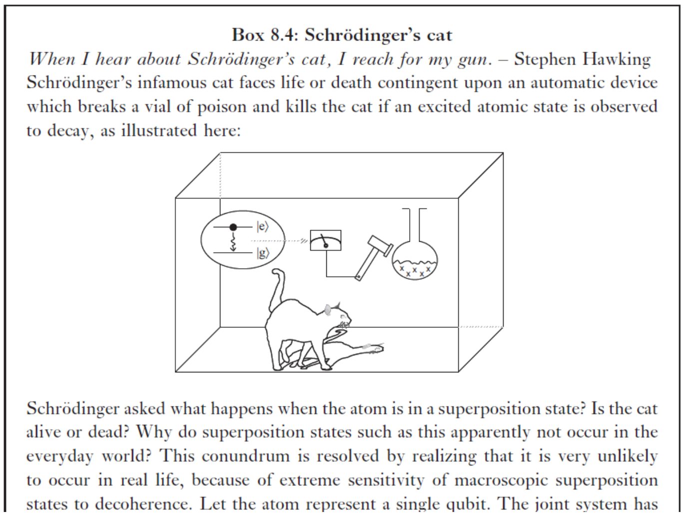
Solution 18. Nature's answer to the wave-particle duality problem: measurement outcomes are random. We only have probabilities of outcomes for the photon measurement above.

QM is deterministic w.r.t. the time evolution of the wave function before measurement (following Schroedinger's equation), but it is random (probabilistic) regarding the measurement outcomes.

Remark 19. No photon is divided in the problem above. At the same time, after a large number of experiments, the result approaches the "classical" (wave theory) one.

Hidden variable theories can be discarded (proved by Bell's inequalities).

2.4.14. Schroedinger's cat


Box 8.4: Schrödinger's cat

When I hear about Schrödinger's cat, I reach for my gun. – Stephen Hawking

Schrödinger's infamous cat faces life or death contingent upon an automatic device which breaks a vial of poison and kills the cat if an excited atomic state is observed to decay, as illustrated here:

Schrödinger asked what happens when the atom is in a superposition state? Is the cat alive or dead? Why do superposition states such as this apparently not occur in the everyday world? This conundrum is resolved by realizing that it is very unlikely to occur in real life, because of extreme sensitivity of macroscopic superposition states to decoherence. Let the atom represent a single qubit. The joint system has

Fig. From Quantum Computation and Quantum Information, by Michael A. Nielsen and Isaac L. Chuang (Cambridge University Press).

Decoherence is the reason we don't see "macroscopic superpositions" suggested by the "Schroedinger's cat" thought-experiment.

2.4.15. Interaction free measurement

This is a summary of the paper by A.C. Elitzur and L. Vaidman, Foundations of Physics 23, 987 (1993). Look at the following experimental diagram.

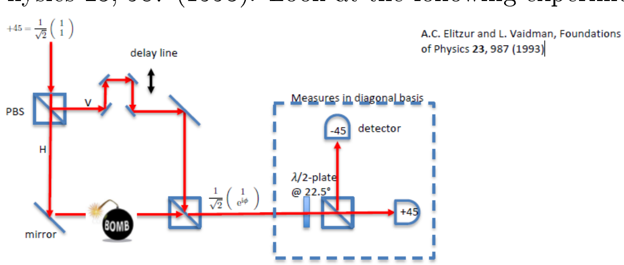


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

- Bomb explodes if even a single photon interacts with it
- Set delay line so $\phi = 0$; if bomb absent photon will exit interferometer at +45°, and the -45° detector will never click

- If bomb is present it may or may not explode: it implements a "which-way" (welcher-weg in German) measurement and photon goes randomly into upper or lower path. If it takes lower path the bomb explodes. If it goes into upper path the bomb remains intact and photon exits in V state: when measured in diagonal basis the photon has equal probability to be detected in ± 45°
- So if bomb is present there is a non-zero probability to get a click from -45° detector! In which case we have detected bomb without interacting with it.
- Not a perfect scheme... but can achieve close to 100% efficiency in a high finesse cavity where photon will likely pass through cavity if bomb absent but reflect if bomb present.

2.5. Quantum State Tomography

Problem 20. Can we measure the full quantum state of any qubit?

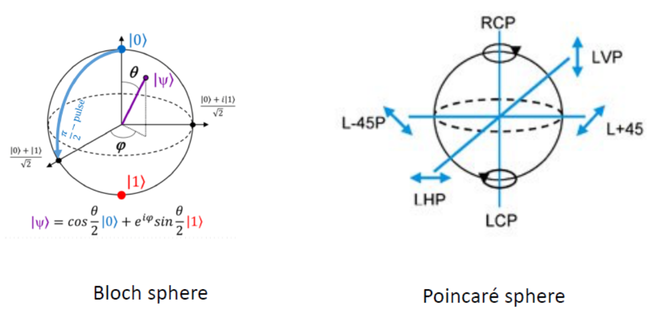


Fig. From "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

Solution 21. Yes, provided that we have many identical copies of the state, and that we can measure in 3 different bases (such as {H,V}, {+45,-45}, {LCP, RCP}).

Let's assume we have a given qubit, expressed in the HV basis as

$$|\Psi\rangle = \psi_H |H\rangle + \psi_V |V\rangle$$

and, since the probability amplitudes are complex coefficients, they can be expressed in the polar form ($r \geq 0$, $\phi \in \mathbb{R}$)

$$\psi_H = r_H \exp(i\phi_H)$$

$$\psi_V = r_V \exp(i\phi_V)$$

As we have sketched before, an overall ("global") phase is not measureable, so we can write

$$|\Psi\rangle = r_H\,|H\rangle + r_V \exp(i\phi)\,|V\rangle = \begin{pmatrix} r_H \\ r_V \exp(i\phi) \end{pmatrix}$$

with

$$\phi = \phi_V - \phi_H$$

Remark 22. A sketch of why an overall global phase is not measureable is that it won't appear (or make a difference) in the prediction of physical observables (measureable quantities). For example, if $\psi(x,t) = |\psi(x,t)| \exp^{i\alpha(x,t)}$ one wants to predict the particle density as a measureable quantity, we have

$$\rho(x,t) = \psi(x,t)\psi^*(x,t) = |\psi(x,t)|^2 \exp^{i\alpha(x,t)} \exp^{-i\alpha(x,t)} = |\psi(x,t)|^2 \geq 0$$

and ultimately the phase has no effect in the actual measureable quantity of interest, namely the density ρ.

Going back to our problem, our goal is to find the 3 variables determining the qubit, namely $r_H, \quad r_V, \quad \phi$. We follow the procedure below.

1. Peform a measurement in the $\{H, V\}$ basis: Then, the probability of detecting an H-polarized photon is

$$P(H) = |\,\langle H|\Psi\rangle\,|^2 = |(1,0)\begin{pmatrix} r_H \\ r_V \exp(i\phi) \end{pmatrix}|^2 = |r_H|^2$$

   Therefore,

$$r_H = \sqrt{P(H)}$$

   and normalization implies

$$1 = r_H^2 + r_V^2 \implies r_V = \sqrt{1 - r_H^2} = \sqrt{1 - P(H)}$$

2. It remains to determine ϕ. The probability of detecting a photon in the $+45$ degrees polarization is

$$P(+45) = |\,\langle +45|\Psi\rangle\,|^2 = |\frac{(1,1)}{\sqrt{2}}\begin{pmatrix} r_H \\ r_V \exp(i\phi) \end{pmatrix}|^2 = |\frac{(r_H + r_V e^{i\phi})}{\sqrt{2}}|^2 = \frac{|r_H + r_V e^{i\phi}|^2}{2}$$

$$P(+45) = \frac{|r_H + r_V\cos(\phi) + i r_V \sin(\phi)|^2}{2} = \frac{(r_H + r_V\cos(\phi))^2 + (r_V\sin(\phi))^2}{2} =$$

$$P(+45) = \frac{r_H^2 + 2r_H r_V \cos(\phi) + r_V^2\cos^2(\phi) + r_V^2\sin^2(\phi)}{2} = \frac{r_H^2 + 2r_H r_V\cos(\phi) + r_V^2}{2} =$$

$$P(+45) = \frac{1 + 2r_H r_V\cos(\phi)}{2} = 1/2 + \cos(\phi)\sqrt{P(H)}\sqrt{1 - P(H)}$$

   Therefore

$$\cos(\phi) = \frac{P(+45) - 1/2}{\sqrt{P(H)(1 - P(H))}}$$

Since the azimuthal angle $\phi \in [-\pi, \pi]$,there are at least two possible solutions:

$$\phi = \pm \arccos(\frac{P(+45) - 1/2}{\sqrt{P(H)(1 - P(H))}})$$

We need then to measure in the circular polarization measurement (which was not distinguished by the previous measurements as both left & right circularly polarized states have $P(H) = P(+45) = 1/2$).

3. The probability of detecting a right circularly polarized photon is

$$P(RCP) = |\langle RCP|\Psi\rangle|^2 = |\frac{(1, -i)}{\sqrt{2}}\begin{pmatrix} r_H \\ r_V \exp(i\phi) \end{pmatrix}|^2$$

$$= |\frac{(r_H - ir_V \exp(i\phi))}{\sqrt{2}}|^2 = \frac{(r_H - ir_V \exp(i\phi))(r_H + ir_V \exp(-i\phi))}{2}$$

$$= \frac{(r_H^2 + ir_V r_H[\exp(-i\phi) - \exp(i\phi)] + r_V^2)}{2} = \frac{(r_H^2 - i^2[2\sin(\phi)]r_V r_H + r_V^2)}{2} =$$

$$\frac{1 + 2r_V r_H \sin(\phi)}{2} = \frac{1}{2} + \sqrt{P(H)(1 - P(H))}\sin(\phi)$$

Therefore

$$\sin(\phi) = \frac{P(RCP) - 1/2}{\sqrt{P(H)(1 - P(H))}}$$

Now that we know that

$$(\cos\phi, \sin\phi) = (\frac{P(+45) - 1/2}{\sqrt{P(H)(1 - P(H))}}, \frac{P(RCP) - 1/2}{\sqrt{P(H)(1 - P(H))}})$$

we can find the unique angle ϕ satisfying the relations. This way, we have determined the state in terms of the probabilities $P(H)$, $P(+45)$, $P(RCP)$, the latter able to be found experimentally.

References

[1] "Communications theory", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

[2] Quantum Information, by Stephen M. Barnett (Oxford University Press, ISBN 978-0-19-852763-3).

[3] "Physical qubits: spinning particles and photons", by Prof. Duncan O'Dell, Introduction to Quantum Information course, McMaster University, 2019.

[4] Shannon, Claude Elwood (October 1948). "A Mathematical Theory of Communication". Bell System Technical Journal. 27 (4): 623–666. doi:10.1002/j.1538-7305.1948.tb00917.x.

[5] By Tatoute - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=34095239

[6] By Lookang - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16874302

[7] Young's double-slit experiment. Encyclopædia Britannica, Inc. https://www.britannica.com/science/light/The-electromagnetic-spectrum#/media/1/340440/91986

[8] http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/slits.html

[9] https://en.wikipedia.org/wiki/Polarization_(waves)

[10] Quantum Computation and Quantum Information, by Michael A. Nielsen and Isaac L. Chuang (Cambridge University Press).

[11] A.C. Elitzur and L. Vaidman, Foundations of Physics 23, 987 (1993)