

## Class notes - MT tutorial 5

This week we are discussing operators and measurements. Together with what we have learnt in previous weeks this completes, at least in its simplest form, the mathematical description of quantum mechanics.

### 1 Postulates of quantum mechanics

The mathematical framework of quantum mechanics can be summarised quite succinctly by 3 postulates. We have already encountered the first two which are

1. Associated to any isolated physical system is a complex vector space  $\mathcal{H}$  (often called a Hilbert space in reference to David Hilbert who worked out the details of the continuum limit). The state of the system is then described completely by a vector  $|\psi\rangle \in \mathcal{H}$  of unit length  $\langle\psi|\psi\rangle = 1$ . A slightly sharper definition might say that the state is a *ray* in the Hilbert space since  $|\psi\rangle$  and  $e^{i\phi}|\psi\rangle$  differing by a global phase are the same physical state.
2. The time-evolution of a closed isolated quantum system between any two times  $t = t_0$  and  $t = t_1$  involves connecting two states as

$$|\psi(t = t_0)\rangle \longrightarrow |\psi(t = t_1)\rangle,$$

and is described by a unitary transformation  $U(t_0, t_1)$  so  $|\psi(t_1)\rangle = U(t_0, t_1)|\psi(t_0)\rangle$ . For a system with a time-independent Hamiltonian  $H$  its evolution is governed by the time-dependent Schrodinger equation

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

which has a formal solution  $|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(t = 0)\rangle$  so the time-evolution unitary connecting the state of the system at times  $t = t_0$  and  $t = t_1$  can be readily identified as  $U(t_0, t_1) = e^{-iH(t_1 - t_0)/\hbar}$ .

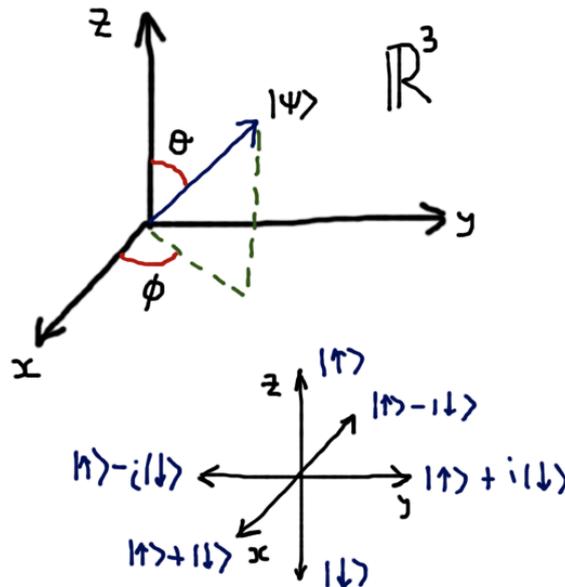
### 2 The simplest example - a qubit

The simplest possible quantum system possesses a two-dimensional Hilbert space  $\mathcal{H}_q = \mathbb{C}^2$ . This case is highly relevant since it may describe the intrinsic spin degree of freedom associated to a spin- $\frac{1}{2}$  particle like an electron (more on this next term), or a two-path optical interferometer like in our first tutorial, or more abstractly a quantum bit (qubit) in a quantum computer. Despite its limited dimensionality quite a bit of basic quantum phenomena is contained in a qubit.

Let us take two states, denoted as  $|\uparrow\rangle$  and  $|\downarrow\rangle$ , as a complete basis to  $\mathcal{H}_q$ . We know that any state of this system can be expressed in this basis as  $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ , where  $\alpha$  and  $\beta$  are complex numbers. At face value this suggests that our system is described by four real numbers, but in fact as postulate 1 describes  $|\alpha|^2 + |\beta|^2 = 1$  and the arbitrary global phase make two of these redundant. Instead we can parameterise any state as

$$|\psi\rangle = e^{i\gamma} \left( \cos\left(\frac{1}{2}\theta\right) |\uparrow\rangle + e^{i\phi} \sin\left(\frac{1}{2}\theta\right) |\downarrow\rangle \right) \quad (1)$$

where  $0 \leq \theta \leq \pi$ ,  $0 \leq \phi \leq 2\pi$  and  $0 \leq \gamma \leq 2\pi$ . For any values of these parameters the state is properly normalised and moreover we can discard  $\gamma$  by simply setting it to zero for convenience. Thus any state of a qubit is described by only two angles which are geometrically equivalent to the spherical polar-coordinates of a point on the surface of a unit-sphere in  $\mathbb{R}^3$ .

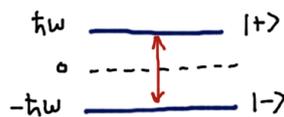


This is called a *Bloch sphere*. Notice that the states at the poles of this sphere along the  $x$ ,  $y$ , and  $z$  axes are the  $\pm 1$  eigenstates of the  $\sigma_x, \sigma_y, \sigma_z$  Pauli matrices, which you can recall are the  $2 \times 2$  complex matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where we have used the conventional choice of  $|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $|\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  as the eigenstates of  $\sigma_z$ . The Pauli matrices appear in many places in quantum mechanics so it is good to remember them and that they are traceless so  $\text{tr}(\sigma_j) = 0$ , hermitian so  $\sigma_j = \sigma_j^\dagger$ , and unitary which implies that  $\sigma_j^2 = \mathbb{1}$ , with  $j = \{x, y, z\}$ .

Having outlined the Hilbert space of a qubit let us now consider its dynamics. Suppose our qubit is a spin- $\frac{1}{2}$  particle in a magnetic field orientated along the  $x$ -axis. The spin degree of freedom of a particle is an intrinsic magnetic moment so the Hamiltonian describing this has the form  $H = \hbar\omega\sigma_x$ , where  $\hbar\omega$  quantifies the energy of this coupling. The eigenstates of  $H$  are simply those of  $\sigma_x$  equal to  $|\pm\rangle = (|\uparrow\rangle \pm |\downarrow\rangle)/\sqrt{2}$  with a corresponding energy  $\pm\hbar\omega$ .



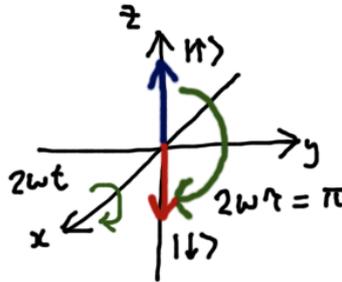
We may thus write the Hamiltonian in its diagonal form as

$$H = \hbar\omega|+\rangle\langle+| - \hbar\omega|-\rangle\langle-|. \quad (2)$$

So what does time-evolution of a qubit look like with this Hamiltonian? To be more concrete suppose at time  $t = 0$  our system is orientated upwards in the  $z$ -axis as  $|\psi(0)\rangle = |\uparrow\rangle$ , what happens to the state as a function of time? Firstly we construct the time-evolution unitary for this system which, because it is a qubit, is quite straightforward as

$$U(t) = e^{-iHt/\hbar} = e^{-i\omega\sigma_x t} = \cos(\omega t)\mathbb{1} - i\sin(\omega t)\sigma_x,$$

where I have used  $\sigma_x^2 = \mathbb{1}$  to simplify the operator exponential. If we evolved the system for a time  $\tau = \pi/2\omega$  then  $|\psi(\tau)\rangle = U(\tau)|\uparrow\rangle$  but since  $U(\tau) = -i\sigma_x$  this gives  $|\psi(\tau)\rangle = |\downarrow\rangle$  up to a global phase, since  $\sigma_x$  flips the spin in the  $z$ -axis. Roughly speaking we can think of unitaries as “rotations” in the complex vector space since they preserve norms and scalar-products. However for the time-evolution of a qubit this is literally a real-space geometrical rotation of points on the Bloch sphere. In this case we see that our qubit state has been rotated by  $\pi$  around the  $x$ -axis after evolving for a time  $\tau$ .



More generally time-evolution caused by this  $H$  causes a rotation of any state, not just  $|\uparrow\rangle$ , about the  $x$ -axis by an angle of  $2\omega t$  on the Bloch sphere. Evidently the states  $|\pm\rangle$  remain unaltered by the evolution (up to a trivial global phase factor) since they are stationary states.

### 3 Quantum measurements

We will now add a final postulate to our formulation which details how we describe measurements of quantum systems

- Given that an observable  $M$  is represented by an hermitian operator a (projective) quantum measurement of  $M$  is defined by its eigen-decomposition

$$M = \sum_m a_m \mathbb{P}_m,$$

where  $a_m$  is the  $m$ th (real) eigenvalue and  $\mathbb{P}_m$  is a projector on to the set of eigenstates sharing this eigenvalue  $a_m$ . The projector can be written as

$$\mathbb{P}_m = \sum_{j=1}^{d(m)} |m, j\rangle \langle m, j|,$$

where  $j$  indexes over the degeneracy  $d(m)$  of the  $m$ th eigenvalue  $a_m$ , so  $\mathbb{P}_m$  projects on to a  $d(m)$ -dimensional subspace (eigenspace). The possible outcomes of a measurement of  $M$  are its eigenvalues  $a_m$ . Given an initial state

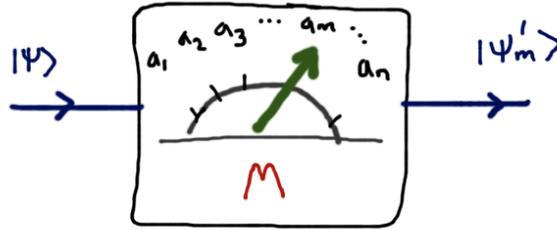
$|\psi\rangle$  the probability of measuring  $M$  and obtaining an outcome  $a_m$  is

$$P_m = \langle \psi | \mathbb{P}_m | \psi \rangle.$$

Having obtained the result  $a_m$  the state of the system immediately after the measurement is

$$|\psi'_m\rangle = \frac{\mathbb{P}_m |\psi\rangle}{\sqrt{P_m}}.$$

This formalism might be visualised by the following picture



in which we envisage the quantum system in an input state coming in one end, some macroscopic “classical” dial indicating the measurement outcome for some observable  $M$ , and then the output state of the quantum system emerging from the other end. Notice that from this definition the probability  $P_m$  is simply the expectation value of the projector  $\mathbb{P}_m$ , and similarly the state after the measurement<sup>1</sup> is obtained by apply this projector to the initial state and subsequently renormalising the resulting state (note that  $\mathbb{P}_m$  is hermitian and idempotent  $\mathbb{P}_m^2 = \mathbb{P}_m$  giving  $\langle \psi | \mathbb{P}_m \mathbb{P}_m | \psi \rangle = P_m$  so  $\sqrt{P_m}$  is precisely the norm of the projected state). For this reason this kind of measurement is called a *projective* measurement. It is the application of a projector which mathematically implements the “collapse” of a wavefunction since it kills off all the components of the state outside of the eigenspace it projects on to.

For projective measurements the final state after of a measurement of  $M$  is always an eigenstate of  $M$  corresponding to the measured outcome. As we have seen before the presence of degeneracies can make this point a bit more subtle than is first apparent. Suppose we had an initial state expanded in the basis  $\{|m, j\rangle\}$  of  $M$  as

$$|\psi\rangle = \frac{1}{\sqrt{20}} (|1, 1\rangle + 2|1, 2\rangle + |1, 3\rangle + |5, 2\rangle + 3|5, 4\rangle + 2|6, 1\rangle).$$

Where we to measure  $M$  we would obtain an outcome  $m = 1$  with a probability

$$P_1 = \langle \psi | \overbrace{(|1, 1\rangle\langle 1, 1| + |1, 2\rangle\langle 1, 2| + |1, 3\rangle\langle 1, 3|)}^{\mathbb{P}_1} | \psi \rangle = \frac{3}{10},$$

and in this case obtain a final state which is an  $m = 1$  eigenstate of  $M$  as

$$|\psi'_1\rangle = \frac{1}{\sqrt{6}} (|1, 1\rangle + 2|1, 2\rangle + |1, 3\rangle).$$

Notice that any superpositions between different degenerate eigenstates of  $M$ , labelled here by the index  $j$ , are retained since the projector itself is a sum over these states. In

<sup>1</sup>We only apply this procedure for working out the final state for outcomes that can occur, i.e.  $P_m > 0$ .

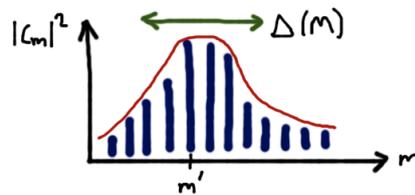
fact the extra label  $j$  is really just another observable, which we might denote as  $J$ , that commutes with  $M$  and whose eigenvalue index  $j$ , together with  $m$ , specifies completely and uniquely a basis of the Hilbert space as  $\{|m, j\rangle\}$ . The eigenvalues  $a_m$  and  $a_j$  for  $M$  and  $J$  form a what is called a complete set of quantum numbers. While this might seem like a burdensome over-complication at this juncture we shall frequently encounter this situation next term in the context of angular momentum and its application to atomic systems.

Since  $\mathbb{P}_m$  is not unitary it indicates that postulate 2 does not seem to be a complete description of quantum mechanical evolution. A projector is an irreversible operation since it annihilates lots of information in the initial state (many initial states  $|\psi\rangle$  map to the same final state  $|\psi'\rangle$ ). Moreover the actual projector which gets applied is probabilistic. In contrast unitary time-evolution is reversible and deterministic. The need for this third postulate is considered by many to be a major flaw in quantum mechanics. After all the measuring device, and surrounding environment including us, are supposed to all be described by quantum mechanics, so in principle if we include enough of the universe in our definition of the “system” then our evolution should be unitary. The jury is still out as to whether postulate 3 is derivable from postulate 2. Evidently the missing link is understanding how quantum mechanics works at the macroscopic level, like that typified by the Schrodinger cat thought experiment.

Let’s check that the definitions given in postulate 3 makes sense with what we already know. Since we obtain outcomes from our measuring device  $a_m$  with probability  $P_m$  the average value  $\langle M \rangle$  obtained by measuring  $M$  many times on systems identically prepared in the state  $|\psi\rangle$  is

$$\langle M \rangle = \sum_m a_m P_m = \sum_m a_m \langle \psi | \mathbb{P}_m | \psi \rangle = \langle \psi | \sum_m a_m \mathbb{P}_m | \psi \rangle = \langle \psi | M | \psi \rangle,$$

thus we recover the usual expression for the expectation value of  $M$ . To make things a little more familiar we can consider the special case where  $M$  is non-degenerate with  $M = \sum_m a_m |m\rangle \langle m|$  and so we can now identify a one-dimensional projector  $\mathbb{P}_m = |m\rangle \langle m|$ . Now the probability of the  $m$ th outcome reduces to the square-modulus of its probability amplitude as  $P_m = \langle \psi | m \rangle \langle m | \psi \rangle = |\langle m | \psi \rangle|^2$ . For any state  $|\psi\rangle = \sum_m c_m |m\rangle$  composed as a superposition of eigenstates  $|m\rangle$  of  $M$ , which for this non-degenerate case all have different eigenvalues  $a_m$ , the observable  $M$  will not have a definite value.



Indeed  $\langle M \rangle$  may or may not actually coincide with a possible measurement outcome  $a_m$ , but even if it happened that  $\langle M \rangle = a_{m'}$ , for some  $m'$ , the state of the system would not necessarily be in the eigenstate  $|m'\rangle$ . This is revealed by looking at the statistical dispersion  $\Delta(M) = \sqrt{\langle M^2 \rangle - \langle M \rangle^2}$  which quantifies the spread of outcomes obtained when measuring  $M$  on many identically prepared systems. Only for eigenstates  $|m\rangle$  is the measurement outcome certain to be  $a_m$  with  $\Delta(M) = 0$  and does  $M$  have a definite value.

## 4 Back to a qubit

We can apply this new postulate to our simple two-level quantum system. In this case we want to measure  $\sigma_x$  which is equivalent to a Stern-Gerlach type measurement of the  $x$ -axis spin-component of a spin- $\frac{1}{2}$  particle. Writing  $\sigma_x = |+\rangle\langle+| - |-\rangle\langle-| = \mathbb{P}_+ - \mathbb{P}_-$  we can see the link to the general formalism for this non-degenerate operator example. What happens if we measure an arbitrary state  $|\psi\rangle$  as parameterised in Eq. (1)? Applying the formalism in postulate 3 gives

$$\begin{aligned} P_+ &= \langle\psi|+\rangle\langle+|\psi\rangle, \\ &= \frac{1}{2}|\cos(\frac{1}{2}\theta) + e^{i\phi}\sin(\frac{1}{2}\theta)|^2, \\ &= \frac{1}{2}[1 + \sin(\theta)\cos(\phi)], \end{aligned}$$

along with

$$\begin{aligned} P_- &= \langle\psi|-\rangle\langle-|\psi\rangle, \\ &= \frac{1}{2}|\cos(\frac{1}{2}\theta) - e^{i\phi}\sin(\frac{1}{2}\theta)|^2, \\ &= \frac{1}{2}[1 - \sin(\theta)\cos(\phi)], \end{aligned}$$

which reassuringly gives  $P_+ + P_- = 1$ . If we obtained  $+$  as the measurement outcome then the final state is simply

$$|\psi'_+\rangle = \frac{\mathbb{P}_+|\psi\rangle}{\sqrt{P_+}} = \frac{1}{\sqrt{2}} \frac{[\cos(\frac{1}{2}\theta) + e^{i\phi}\sin(\frac{1}{2}\theta)]|+\rangle}{\sqrt{P_+}} = |+\rangle,$$

up to a global phase. That we get a definite state out from a measurement of a non-degenerate observable means that we can use projective measurements as a non-deterministic procedure for preparing eigenstates of  $M$  by filtering out the outcome we desire.

## 5 Heisenberg's uncertainty principle

Perhaps one of the most famous aspects of quantum mechanics here we shall derive it as a consequence of the mathematical structure imposed by our description, namely complex linear vector spaces stated in postulate 1. Firstly, let's consider two observables represented by hermitian operators  $A$  and  $B$ . Given a state  $|\psi\rangle$  then we can compute the expectation value  $\langle\psi|AB|\psi\rangle = z$  which will in general be complex. The product of operators  $AB$  can be written as  $AB = \frac{1}{2}[A, B] + \frac{1}{2}\{A, B\}$ , where  $\{A, B\} = AB + BA$  is the anti-commutator. This means that

$$\frac{1}{2}\langle\psi|[A, B]|\psi\rangle + \frac{1}{2}\langle\psi|\{A, B\}|\psi\rangle = \langle\psi|AB|\psi\rangle,$$

We then note that the expectation value of the commutator is purely imaginary

$$\langle\psi|[A, B]|\psi\rangle = \langle\psi|AB|\psi\rangle - \langle\psi|BA|\psi\rangle = \langle\psi|AB|\psi\rangle - \langle\psi|AB|\psi\rangle^* = 2i\text{Im}(z),$$

while the expectation value of the anti-commutator is purely real

$$\langle\psi|\{A, B\}|\psi\rangle = \langle\psi|AB|\psi\rangle + \langle\psi|BA|\psi\rangle = \langle\psi|AB|\psi\rangle + \langle\psi|AB|\psi\rangle^* = 2\text{Re}(z),$$

after using the basic property of our scalar-product  $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$ . This means that

$$|\langle\psi|[A, B]|\psi\rangle|^2 + \underbrace{|\langle\psi|\{A, B\}|\psi\rangle|^2}_{\text{drop}} = 4|\langle\psi|AB|\psi\rangle|^2.$$

We now drop the anti-commutator term indicated and since it is  $\geq 0$  this acts only to decrease the size of the lefthand side, leaving the inequality

$$|\langle \psi | [A, B] | \psi \rangle|^2 \leq 4 |\langle \psi | AB | \psi \rangle|^2. \quad (3)$$

Notice that the inequality can achieve equality since for any given operators  $A$  and  $B$  there exists some state  $|\psi\rangle$  where  $\langle \psi | AB | \psi \rangle$  is entirely imaginary. To proceed further we work on the righthand side of Eq. (3) by appealing to another basic property of complex vector spaces called the Cauchy-Schwarz inequality. This states that for any two vectors  $|u\rangle$  and  $|v\rangle$  we have that

$$|\langle u | v \rangle|^2 \leq \langle u | u \rangle \langle v | v \rangle.$$

In  $\mathbb{R}^n$  this follows quite trivially from the dot-product in the plane containing  $\vec{u}$  and  $\vec{v}$  with an angle  $\theta$  between them as

$$|\vec{u} \cdot \vec{v}| = \|\vec{u}\| \|\vec{v}\| |\cos(\theta)| \leq \|\vec{u}\| \|\vec{v}\|,$$

and so this theorem is merely established this geometric picture more generally to complex vector spaces. From this we see that

$$|\langle \psi | AB | \psi \rangle|^2 \leq \langle \psi | A^2 | \psi \rangle \langle \psi | B^2 | \psi \rangle,$$

and by feeding this into Eq. (3) we get

$$|\langle \psi | [A, B] | \psi \rangle|^2 \leq 4 \langle \psi | A^2 | \psi \rangle \langle \psi | B^2 | \psi \rangle.$$

We now define our observables  $A$  and  $B$  as having a zero mean with respect to  $|\psi\rangle$  which implies that they must have the shifted form  $A = C - \langle \psi | C | \psi \rangle \mathbb{1}$  and  $B = D - \langle \psi | D | \psi \rangle \mathbb{1}$  where  $C$  and  $D$  are two new observables. This is a subtle and somewhat strange definition<sup>2</sup> for an operator given that it depends on  $|\psi\rangle$ , but for any observables  $C$  and  $D$ , and state  $|\psi\rangle$ , we can nonetheless always form the pair  $A$  and  $B$ . Since we only shifted the operators we have that  $[A, B] = [C, D]$ . Using this we see that  $\langle \psi | A^2 | \psi \rangle = \Delta(C)^2$  and  $\langle \psi | B^2 | \psi \rangle = \Delta(D)^2$  so

$$|\langle \psi | [C, D] | \psi \rangle|^2 \leq 4 \Delta(C)^2 \Delta(D)^2,$$

thus the product of the statistical dispersions of two observables  $C$  and  $D$  for any quantum state  $|\psi\rangle$  obey the following inequality

$$\Delta(C)\Delta(D) \geq \frac{|\langle [C, D] \rangle|}{2}. \quad (4)$$

which is Heisenberg's uncertainty relation. We shall show shortly for some special cases that this inequality is tight. A simple consequence of this result is that if two observables have a non-zero commutator then it is not possible for  $\Delta(C)$  and  $\Delta(D)$  to both be close to zero. Only when operators commute do states exist with zero statistical dispersion simultaneously for those observables, i.e. they can be simultaneously well defined.

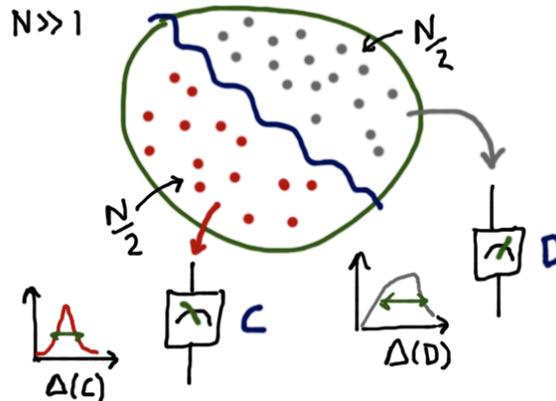
<sup>2</sup>If you don't like this definition you can always define  $A = C - \alpha \mathbb{1}$ , where  $\alpha$  is real, and carry it through the calculation until the end. The final result will be independent of  $\alpha$  and you are free to choose  $\alpha = \langle C \rangle$ .

## 6 What does the uncertainty principle mean?

In Q5.7 on the QMMT problem sheet some of the ideas behind the uncertainty principle are reviewed including some common misnomers. Let's discuss a few of these here. In particular the consider the following statement:

“By measuring  $C$  to some accuracy  $\Delta(C)$  causes the value of  $D$  to be disturbed by an amount at least as large as  $\Delta(D)$  in such a way that the Heisenberg uncertainty principle is satisfied.”

While it is true that measurements in quantum mechanics do disturb the measured system (we saw this above in the form of the projectors applied to get the final state), this is emphatically **not** the content of Heisenberg's uncertainty principle. To understand its correct interpretation we need to consider an ensemble of  $N \gg 1$  identical systems all in some arbitrary state  $|\psi\rangle$ . Now suppose we measure  $C$  on half this ensemble and  $D$  on the other half. On the first subset we will obtain a spread of measurement outcomes for  $C$  whose probabilities are given by  $|\psi\rangle$  and these will in general have a statistical dispersion  $\Delta(C)$ . On the second subset we will similarly obtain a spread of outcomes for  $D$  with a statistical dispersion  $\Delta(D)$ . Overall we might picture this setup as



The two standard deviations obtained,  $\Delta(C)$  and  $\Delta(D)$ , when multiplied together rigorously satisfy the Heisenberg uncertainty principle in Eq. (4). Notice that in this construction no single quantum system was ever measured twice, rather each system was measured only once by  $C$  or  $D$ . Thus the picture of “disturbance” is misleading since it seems to suggest that  $\Delta(C)$  and  $\Delta(D)$  are conceptually avoidable if we could only devise cleverer experimental setups. It also describes a distinctly classical picture where the observables  $C$  and  $D$  are presumed to have a well defined value and it is our ability to access them simultaneously which is limited.

A classical example of Heisenberg's uncertainty relation is for position and momentum where  $[\hat{x}, \hat{p}] = i\hbar\mathbb{1}$ . Inserting this into Eq. (4) gives the well known result

$$\Delta(\hat{x})\Delta(\hat{p}) \geq \frac{\hbar}{2}.$$

This relation is sometimes explained by physical arguments involving some specific measurement setups. For instance by determining the position of a particle by scattering light off it with a wavelength  $\lambda$ . Better resolution in position is then found only by

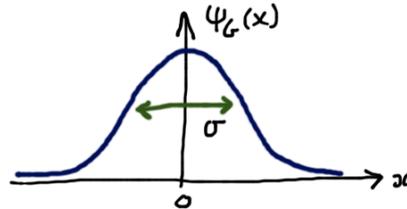
reducing  $\lambda$  and correspondingly increasing the photon energy and momentum causing a bigger momentum recoil. Although such heuristic arguments are intuitive they do not represent the central principle. The Heisenberg uncertainty principle does not depend on a particular physical realisation of a measurement (indeed there is no mention of this in Eq. (4)) nor on any subsequent disturbance. Our ensemble interpretation instead tells us that it is about the basic structure of quantum states and observables. The commutator between two observables  $C$  and  $D$  indicates the trade off in finding quantum states which have an intrinsic uncertainty  $\Delta(C)$  and  $\Delta(D)$ . These uncertainties are therefore nothing to do with us not knowing the answer, but rather they are a fundamental indeterminacy of the quantities  $C$  and  $D$  for the quantum state in question. Our ability to predict the results of these two measurements is limited by Eq. (4) *however we actually come to obtain the information.*

## 7 Gaussian wave-packets

We know that eigenstates of position are ultra-thin delta-function spikes in space while those for momentum are entirely delocalised plane-waves. It is therefore instructive to consider states in between which are not localised in position or momentum. An obvious choice are Gaussian wave-packets whose wave-functions are given by

$$\langle x|G\rangle = \Psi_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

This gives a wave-function centred on  $x = 0$  with a width characterised by  $\sigma$  as depicted here



Note that  $\int_{-\infty}^{\infty} dx |\Psi_G(x)|^2 = 1$  so it is properly normalised. Due to its symmetry we have that  $\langle \hat{x} \rangle = 0$  while  $\langle \hat{x}^2 \rangle = \frac{1}{2}\sigma^2$ , thus  $\Delta(\hat{x}) = \frac{\sigma}{\sqrt{2}}$  gives a quantitative measure of how localised around  $\langle \hat{x} \rangle = 0$ .

Now let's work out the wave-function of  $|G\rangle$  in momentum space by expanding it as  $|G\rangle = \int_{-\infty}^{\infty} dp |p\rangle \langle p|G\rangle$ . The momentum wave-function is then  $\tilde{\Psi}_G(p) = \langle p|G\rangle$  and can be worked out from the position space wave-function using the resolution of identity as

$$\tilde{\Psi}_G(p) = \int_{-\infty}^{\infty} dx \langle p|x\rangle \langle x|G\rangle = \int_{-\infty}^{\infty} dx \langle p|x\rangle \Psi_G(x).$$

Since position eigenstates  $|x\rangle$  have momentum wave-functions

$$\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp(-ikx),$$

where  $p = \hbar k$ , then

$$\tilde{\Psi}_G(p) = \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{\sqrt{\pi\sigma^2}} \underbrace{\int_{-\infty}^{\infty} dx e^{-ikx} e^{-\frac{x^2}{2\sigma^2}}}_{\text{FT}}$$

The momentum space wave-function is found from the position space one by performing a Fourier transform. This can be worked out as

$$\tilde{\Psi}_G(p) = \sqrt{\frac{\sigma}{\hbar\sqrt{\pi}}} \exp\left(-\frac{k^2\sigma^2}{2}\right),$$

which shows that the momentum space wave-function is also a Gaussian centred around  $p = 0$ . Again the symmetry of the wave-function means that

$$\langle \hat{p} \rangle = \int_{-\infty}^{\infty} dp p |\tilde{\Psi}_G(p)|^2 = 0,$$

while  $\langle \hat{p}^2 \rangle = \frac{\hbar^2}{2\sigma^2}$ . This means that  $\Delta(\hat{p}) = \frac{\hbar}{\sigma\sqrt{2}}$  and the spread of the wave-function around  $\langle \hat{p} \rangle = 0$  is inversely proportional to  $\sigma$ . This illustrates the expected trade-off between forming a quantum state that is localised in position verses one localised in momentum. However, we then observe that for Gaussian wave-packets

$$\Delta(\hat{x})\Delta(\hat{p}) = \frac{\hbar}{2},$$

showing that they actually saturate the Heisenberg uncertainty principle and are the minimum uncertainty wave-packets. Gaussian wave-packets will form an important part of next weeks class on the harmonic oscillator.

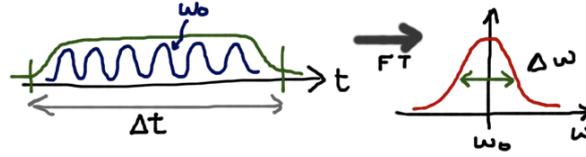
## 8 The energy-time uncertainty relation

Around the same period when Heisenberg's uncertainty relation was derived another uncertainty relation was proposed for energy and time of the form

$$\Delta E \Delta t \geq \frac{\hbar}{2}, \quad (5)$$

which has a deceptively similar form to the position-momentum relation. From a signal processing perspective the origin of these relations seems to be identical and based simply on Fourier transforms. Indeed to build a wave-packet that is localised in position  $x$  we require a superposition of many wave-vectors  $k$ , and similarly to build a wave-packet whose duration is short in time  $t$  many frequencies  $\omega$ . Fourier transforms are thus related by  $\Delta x \Delta k \geq \frac{1}{2}$  and  $\Delta \omega \Delta t \geq \frac{1}{2}$ . We have already seen that for momentum and position it was exactly this Fourier transform property which limited how well we can construct a state with both these quantities well defined. What about energy and time?

This classical signal-processing perspective suggests that to determine with precision  $\Delta \omega$  the frequency of a pulse requires that you spectrally analyse it over a time window (i.e. an envelope which truncates the wave-train) of size  $\Delta t \geq 2/\Delta \omega$ . In this context it makes perfect sense since you can only determine whether a pulse contains extremely low frequency components by waiting long enough to see them oscillate.



This intuition is not quantum mechanical however. The content of Eq. (5) has nothing to do with how long you must measure for in an experiment to determine the energy to some accuracy. Moreover Eq. (5) has a different and more indirect origin than all other uncertainty relations.

A first indication that things are not straightforward is that in non-relativistic quantum mechanics time is a parameter and not an observable. There is no hermitian operator  $T$  representing time (an argument as to why this is the case will be given below) and  $\Delta(T)$  is not therefore the statistical dispersion of the results of some measurement process of time. Nevertheless, we do make measurement of time. We do this by building *clocks*, which are systems whose dynamics causes their state to change in a predictable way, e.g. like a pendulum. A measurement of time is then a measurement of the clock from which we can infer the approximate present value of the parameter  $t$ . In our case the clock is a quantum system whose evolution is given by a Hamiltonian  $H$ , which is assumed to be time-independent, and the clock variable is some observable  $A$ . We know from Ehrenfest's theorem that the expected value of our clock will be governed by

$$\frac{d\langle A(t) \rangle}{dt} = \frac{i}{\hbar} \langle [H, A] \rangle,$$

but we also know that  $\Delta(E)\Delta(A) \geq \frac{1}{2} |\langle [H, A] \rangle|$  from Eq. (4) since the Hamiltonian  $H$  is the observable for energy, so

$$\Delta(E)\Delta(A) \geq \frac{\hbar}{2} \left| \frac{d\langle A(t) \rangle}{dt} \right|.$$

Thus we then define an uncertainty in time indirectly based on our uncertainty of our clock variable  $A$  and the rate at which it is changing as

$$\Delta t = \frac{\Delta(A)}{\left| \frac{d\langle A(t) \rangle}{dt} \right|}, \quad (6)$$

thereby giving Eq. (5) via an indirect route. Roughly speaking one can think of  $\Delta t$  as characterising the amount of time required for the expectation value of the clock variable  $\langle A \rangle$  to change by one-standard deviation  $\Delta(A)$ . One might therefore interpret it as the shortest timescale in which we will be able to notice changes in the state  $|\psi\rangle$ , caused by the time-evolution, by monitoring the observable  $A$ . Strictly speaking though Eq. (6) is actually an error-propagation relation between the uncertainty  $\Delta t$  in the inferred time given the uncertainty  $\Delta(A)$  in the clock variable  $\langle A \rangle$  used for the inference. This relation is completely classical, for example if our reading of an analog clock hand suffered from some angular uncertainty  $\Delta\theta$ , then an uncertainty in time  $\Delta t$  would follow from a similar equation. This shows that Eq. (5) is a different kind of uncertainty relation in that  $\Delta t$  is not directly quantum mechanical.

We might ask why is there no observable for time? Let's suppose there was one. For reasons which will become clear it would be an observable  $T$  which is canonically conjugate to the Hamiltonian  $H$ , by which we mean that it has a commutation relation

$[H, T] = i\hbar$ . Using the standard uncertainty relation Eq. (4) we then get that our desired relation Eq. (5) directly. Moreover from Ehrenfest's theorem we see that

$$\frac{d\langle T \rangle}{dt} = 1,$$

which shows that the observable  $T$  would function as a clock in the simplest and most perfect possible way since it would increase at a constant rate of one per unit of time, regardless of that actual state of the system  $|\psi\rangle$ . The problem is that for physically realistic Hamiltonians  $H$  you can prove that there is no operator  $T$  which is canonically conjugate and so there is no time observable. A glossed over reason for this is that any pair of operators with a commutation relation of the form  $[p, q] = i\hbar$  are essentially disguised versions of position and momentum (proven via the Stone-Von Neumann uniqueness theorem). However, these operators are, and are required to be, unbounded below, so their spectrum extends all the way down to  $-\infty$ . Any physical Hamiltonian  $H$  must be bounded below since it has to possess a ground state.

The above observation regarding  $T$  is perhaps not surprising. Hidden in our expressions above we have a state  $|\psi\rangle$  of the system. Requiring that the clock operates perfectly regardless of  $|\psi\rangle$  is a bit unnecessarily restrictive. We would be happy to get to find an operator which has an almost constant rate of change for some particular special clock state, leading us back to our previous argument. Were we to choose this state to be an eigenstate of  $H$  we would obtain a completely **useless** clock since none of its observable properties would change with time. A clock in a stationary state is therefore a *stopped* clock. A working clock must instead be in a superposition of energy eigenstates and possess a spread in its energy values. In Eq. (5) we have a quantitative trade-off between the indeterminacy of the system's energy and its possible resolution as a measure of time.

We can make a concrete calculation to this effect by showing that  $\Delta t$  is a measure of how long it takes for the system to evolve to a new state that is sufficiently distinguishable to indicate a new clock time. By distinguishable we mean that  $\langle \psi(\Delta t) | \psi(0) \rangle = 0$  so that the clock's state is orthogonal to its original state (non-orthogonal states are never perfectly distinguishable). The energy spread of the initial state  $|\psi(0)\rangle$  puts a speed-limit on how fast it can evolve into an orthogonal state. To illustrate this let's take a qubit system with the Hamiltonian given in Eq. (2) earlier. This had eigenstates  $|+\rangle$  and  $|-\rangle$  with corresponding energies  $\hbar\omega$  and  $-\hbar\omega$ . As with our earlier example we start the system in an initial state  $|\psi(0)\rangle = |\uparrow\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$  and compute the time-evolved state as

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}(|+\rangle + e^{-2i\omega t}|-\rangle),$$

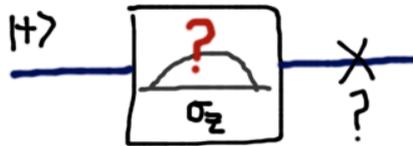
The energy difference  $2\hbar\omega$  determines how quickly the phase in the superposition oscillates. We can compute from  $|\psi(t)\rangle$  the quantities  $\langle E \rangle = 0$ , and  $\langle E^2 \rangle = (\hbar\omega)^2$ , so the intrinsic uncertainty in the energy of the system in this state is  $\Delta(E) = \hbar\omega$ , independent of time. So what is the minimum amount of time  $\Delta t$  for the system to satisfy  $\langle \psi(\Delta t) | \psi(0) \rangle = 0$ . For this we need  $\omega\Delta t = \pi$  so that  $|\psi(\Delta t)\rangle = |\downarrow\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$ , and since  $\omega = \Delta(E)/\hbar$  we have  $\Delta(E)\Delta t = \pi\hbar$ . The energy-time uncertainty relation is satisfied, but not saturated. It is left as an exercise for you to consider what the equivalent relation is for a system with  $d$ -levels, with a Hamiltonian  $H = \sum_{n=1}^d n\hbar\omega |n\rangle\langle n|$ , starting in an initial state  $|\psi(0)\rangle = \frac{1}{\sqrt{d}} \sum_{n=1}^d |n\rangle$ . What happens as  $d \rightarrow \infty$ ?

## 9 Tutorial question

In this weeks class the question for the tutorial was *How do measurements change the state of the system if the outcome is not observed?*. Let's suppose we have a qubit in the state

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

If  $\sigma_z$  is measured then we have probabilities  $P_{\uparrow} = P_{\downarrow} = \frac{1}{2}$  for each outcome. Now suppose that our Stern-Gerlach machine crashed right at the end so the outcome of the measurement was erased. I know that a measurement was made, and further that whatever the outcome was the system was projected into that eigenstate of  $\sigma_z$ . If a second measurement of  $\sigma_z$  is made with a working device the probabilities for the outcomes are again  $P_{\uparrow} = P_{\downarrow} = \frac{1}{2}$ . How does the state after the first measurement differ from that before the first measurement? Both give the same probabilities for measurements of  $\sigma_z$  but are they the same state?



The short answer to this question is **no**, the state after the first faulty measurement is very different from the initial state. The initial state  $|+\rangle$  is in a coherent quantum mechanical superposition of  $|\uparrow\rangle$  and  $|\downarrow\rangle$ . After the faulty measurement however, the system is for definitely in one of the  $\sigma_z$  eigenstates  $|\uparrow\rangle$  and  $|\downarrow\rangle$ , but which one is governed by a probability  $\frac{1}{2}$ . This situation has *classical ignorance*, which is the usual origin of probabilities, rather than quantum probability amplitudes. The difference is highlighted by considering a measurement of  $\sigma_x$ . For  $|+\rangle$  a  $+$  result is obtained with certainty, whereas for the state after the faulty measurement outcomes  $+$  and  $-$  occur with probability  $\frac{1}{2}$ . This difference is because  $|+\rangle$  has “coherence” and will exhibit quantum interference, whereas a classical mixture of  $|\uparrow\rangle$  and  $|\downarrow\rangle$  is incoherent and cannot produce interference. This interference property can be well illustrated by the Mach-Zehnder interferometer we considered back in the first tutorial. If we measured which path the photon took but forgot the result no interference would ever be observed. To add this kind of classical ignorance (of what state the system is actually in) to our mathematical framework it would need to be enhanced so that the state of the system is not described by a single vector  $|\psi\rangle$  but by a so-called density operator  $\rho$ . Unfortunately this takes us far beyond the syllabus of this course so we cannot discuss it further.