

PHYS301

Quantum Physics

Semester 1 2009

J D Cresser

jcresser@physics.mq.edu.au

Rm C5C357

Ext 8913

A new theory of physics

- In the early years of the 20th century there was developed a new theory of physics: **quantum mechanics**.
- The most successful theory of physics ever devised.
- It is not a 'physical theory'!!!!
 - It is a set of 'rules' by which all physical theories are constructed.
 - Believed to apply to **all** physical systems from subatomic particles to the universe itself.
- The successes of quantum mechanics include:
 - Explains structure of atoms and molecules (and all of chemistry??)
 - The operation of a laser
 - Explains properties of the solid state (responsible for ~ 40% of gross national product of USA)
 - Provides the 'standard model' of particle physics
 - Quantum gauge field theory \leftrightarrow the forces of nature
 - Quantum fluctuations in the early universe \rightarrow density fluctuations \rightarrow galaxies
 - Provides a link between general relativity and thermodynamics
 - Will play a central role in describing how the universe came into being.

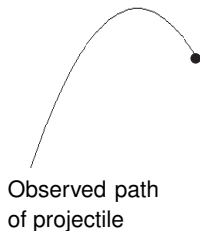
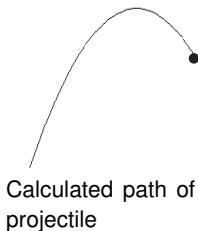
What do we 'lose' in this new theory?

- Quantum mechanics success comes at a 'cost':
 - Quantum mechanics forces on us a new way of looking at the world.
 - The world-view implicit in pre-quantum (classical) physics has to be abandoned.
 - Events appear to occur *without a cause*.
 - Physical systems appear to be able to do a number of mutually exclusive things *at the same time*.
 - We are even forced to ask:

Is there a 'real' world out there independent of us as observers?

- But first . . . what is 'classical physics'?

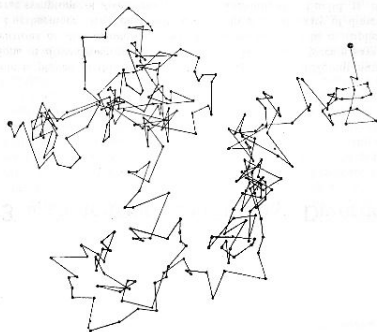
- 'Classical physics': a name for pre-quantum (or better, non-quantum) physics.
- Examples include: Newtonian mechanics & Maxwell's Theory of the Electromagnetic Field & General Relativity . . .
- According to classical physics, there is an objectively real world 'out there' with properties and behaviour that are completely independent of us as observers.
- The equations of physics describe what is 'really happening' with a physical system.
- E.g. every particle has a definite position and velocity.
- The solution to Newton's equations for a particle in motion is a perfect representation of what the particle is 'actually doing':



What about limits to accuracy of measurement and/or calculation?

- In principle, there is no limitation to how well we can measure the properties of a physical system.
- E.g. for a particle we can measure
 - either the position x or the momentum p or **both x and p at the same time**
 - **and** with a precision limited only by our experimental ingenuity. There is no law of classical physics that says we can't.
- There is an implicit belief that by refining our experiments — e.g. by measuring to >100 th decimal place — we are getting closer and closer to the values of the position and momentum that the particle 'really' has.
- We can, in principle, calculate, with unlimited accuracy, the future behaviour of any physical system by solving using Newton's equations, Maxwell's equations and so on.

- Realistically, we cannot measure everything . . .
- We cannot measure the position and velocities of each molecule in a gas of 10^{26} molecules!
- Or predict the motion of a pollen grain suspended in a liquid:



Brownian motion (random walk) of pollen grain due to collisions with molecules of liquid.

- Random behaviour only **appears** random because we do not have enough information to describe it exactly.
 - It is not really random because we believe that if we could repeat an experiment under **exactly** identical conditions we ought to get the same result every time.
 - So we use methods of probability and statistics to work out averages values — average energy or average velocity and so on . . .
 - We use probability methods to compensate for our **ignorance** of information — if we could get at this information then we could solve everything exactly and all would be known.
 - We cannot predict the winner of a horse race because we do not have all the relevant information – we can just lay odds as to which horse will win.
 - If a punter knew all the information than (s)he could predict the winner with absolute certainty !!

Does classical physics and its world-view always work?

- It works fine at the everyday (macroscopic) level
 - though there are things at the macroscopic level that cannot be understood using classical physics e.g.
 - The colour of a heated object . . .
 - Why there exists solid objects . . .
- Where does classical physics come unstuck?
 - Non-classical behaviour is most readily observed for microscopic systems – atoms and molecules, but is in fact present at all scales!
- What sort of non-classical phenomena do we observe?
 - Intrinsic Randomness
 - Interference phenomena (e.g. particles acting like waves)
 - Entanglement

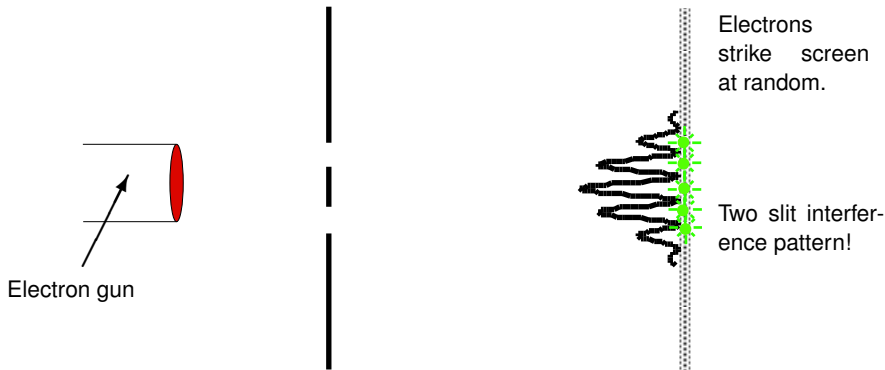
- Impossible to prepare *any* physical system in such a way that *all* its physical attributes are precisely specified – e.g. we cannot pin down both the position *and* the momentum of a particle at the same time.
 - For experiments repeated under exactly identical conditions there will always be some physical quantity which, when measured, will always yield randomly varying results.
 - E.g. prepare a particle with a precise momentum, measure its position, then repeat . . .
 - We get a randomly varying result for its measured position.
 - This is **NOT** because we do a lousy job of measuring the particle's position.
 - This randomness is **irreducible**: it **cannot** be totally removed by improvement in experimental technique.

What is the meaning of this?

- Nature places limits on how much information we can gather about any physical system.
 - We apparently cannot know with precision as much about a system as we thought we could according to classical physics.
- But is the missing information still there?
 - Does the particle have a precise position, but we simply cannot access it?
- Apparently not: the randomness cannot be removed by digging out missing information – there is no missing information.
- The randomness is not caused by our ignorance of some information!
- It is apparently 'uncaused' random behaviour.

Interference

- Physical systems behave as if they are doing mutually exclusive things at the same time.



e.g. a **particle** going through two separate slits simultaneously producing a **wave-like** interference pattern.

- Two widely separated systems can have properties more highly correlated than classical physics permits.
 - Experimentally testable prediction via Bell's inequalities.
 - Overwhelmingly confirmed!
- Seems to suggest two systems 'communicating' instantaneously, i.e. faster than the speed of light.
 - This is inconsistent with Einstein's theory of relativity.
- If we believe Einstein is right then something else has to change:
- Must assume that physical systems acquire some (maybe all?) properties only through the act of observation, e.g. a particle does not 'really' have a specific position until it is measured.

So what is quantum mechanics?

- Quantum mechanics is a **meta-theory** . . .
 - It provides a framework, or a set of principles, that **all** other theories of physics must fit, or satisfy.
 - Another example of a meta-theory is relativity — both special and general — which places strict conditions on the properties of space and time. In other words, **space and time** must be treated in all (fundamental) physical theories in a way that is consistent with the edicts of relativity.
- To what aspect of **all** physical theories do the principles of quantum mechanics (whatever they are!!) apply?
 - The principles must apply to theories as diverse as
 - Newton's Laws describing the mechanical properties of matter;
 - Maxwell's equations describing the electromagnetic field;
 - The laws of thermodynamics
- What is the common feature?
- The answer lies in noting how a theory in physics is formulated.

- Modern physical theories are not arrived at by pure thought (except, maybe, general relativity) . . .
 - Data is collected, either by casual observation or controlled experiment on, for instance
 - the motion of physical objects;
 - the behaviour of charges and currents and electric and magnetic field strengths;
 - temperature, pressure, volume of solids, liquids, or gases . . .
 - Scattering experiments in high energy particle accelerators
- Reproducible regularities are searched for within this data, and, typically, reorganised into mathematical statements.
 - Eugene Wigner (a Nobel prize winning theorist) would often comment on the ‘unreasonable effectiveness of mathematics in the physical sciences.’

The laws of classical physics

- Classical physics says that we can freely gather information on any such physical quantity, and to whatever level of precision we like, limited by our own ingenuity.
- Any uncertainty or randomness is due to our ignorance of information about a system.
 - The familiar rules of probability apply to cover those situations in which there is lack of complete information.
- Moreover, classical physics says that this information is a faithful representation of what is 'really' going on in the physical world.
 - These might be called the 'classical laws of information' implicit in classical physics.
- As a consequence, the data can be organized as mathematical equations, such as
 - $\mathbf{F} = m\mathbf{a}$ — Newton's second law;
 - $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ — One of Maxwell's equations (Faraday's law);
 - $PV = NkT$ — Ideal gas law (not really a fundamental law)
- The symbols represent the true, real value of the corresponding physical quantity.
- **These equations are relationships between information gained by observation.**

Quantum laws vs classical laws.

- Quantum mechanics is (also) a set of laws governing the **information** that we can have about the physical world.
 - But these are **NOT** the same laws as implicit in classical physics
- The laws of quantum mechanics place limits on the information that can be gained about a physical system.
 - If we repeat an experiment under identical conditions, the answers will vary in a random way.
 - Quantum entanglement seems to imply that a physical quantity (such as the position of a particle) does not have a definite value until it is measured!
 - Measuring the value of one physical quantity will randomly affect the value of some other
 - Measuring the position of a particle with makes its momentum completely uncertain.
- So we cannot write down equations like $\mathbf{F} = m\mathbf{a}$ that relates the 'actual' values of force and acceleration.

- But there is still law there!!
 - E.g. quantum mechanics tells us that

$$\Delta x \Delta p \geq \frac{1}{2} \hbar \quad \text{the Heisenberg Uncertainty Relation}$$

- Quantum mechanics provides a set of rules that the information about a system must obey.
 - There are a set of laws *common to all physical systems*.
 - But the details depend on the nature of the system.
- Quantum mechanics tells us how this information is processed e.g. as a system evolves in time (the Schrödinger equation)
- Or what results might be obtained (in a randomly varying way) when performing a measurement.
- Quantum mechanics is a theory of information:

QUANTUM INFORMATION THEORY

What are the consequences?

- **BAD:** We lose the apparent certainty and determinism of classical physics.
 - This is replaced by uncertainty and randomness.
 - This randomness is not due to our inadequacies as experimenters — it is **built into** the very fabric of the physical world.
- **GOOD:** Quantum laws mean that physical systems can do so much more within these restrictions.
 - A particle with position or momentum uncertain by amounts Δx and Δp means we don't quite know where it is, or how fast it is going, and we can never know this.
 - **But** the particle can be doing a lot more things 'behind the scenes' as compared to a classical particle of precisely defined position and momentum.
- The result is infinitely richer physics — quantum physics.

- Viewing quantum mechanics as a theory of information has led to major new applications of quantum physics in the last 20 years:
 - quantum computers — infinitely more powerful than a classical computer, if one is ever made!!!
 - quantum cryptography — perfect protection from eavesdroppers — now used in practice.
 - quantum teleportation (but not the Star Trek kind).
- But there have also been advances in the understanding of quantum mechanics itself:
 - The great mystery of ‘quantum measurement’ — God playing dice with the universe — is better understood, but still a long way to go.
 - How does a world governed by the laws of quantum mechanics end up behaving according to classical laws?
 - Where is the quantum-classical boundary?
 - The important role of decoherence in ‘washing out’ the quantum superpositions (a system doing many different things at once) is better understood (and even controlled), but still a long way to go.

- The mathematics should reflect the physics in a natural way.
- **Matrix Mechanics:** The very first version of quantum mechanics was due to Heisenberg (1925).
 - The observable properties of a physical system were written as matrices — **GOOD**
 - Little understood at that time — new mathematics, hard to work with and difficult to understand — **BAD**.
- **Wave Mechanics:** Then in 1926 along came Schrödinger

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x) \Psi(x, t) = i\hbar \frac{\partial \Psi(x, t)}{\partial t}.$$

- Directly reflects the observed wave properties of matter — **GOOD**.
- Solving the Schrödinger equation and working with the wave function $\Psi(x, t)$ was familiar mathematical territory — **GOOD**

- Useful when you are concerned about *where* a particle is to be found in space because:

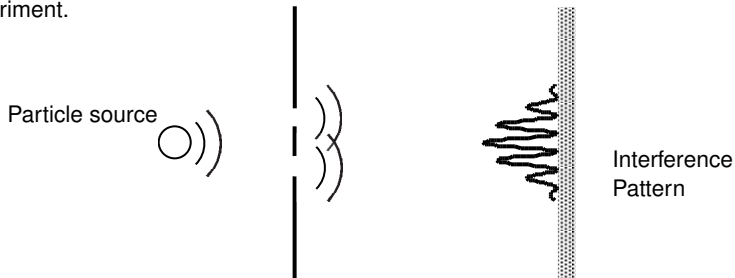
$$|\Psi(x, t)|^2 dx = \text{probability of finding the particle in region } (x, x + dx) \text{ at time } t.$$

- Wave mechanics was ‘the way’ to do quantum mechanics for almost 60 years.
- **Exceedingly successful** — almost all of chemistry, solid state physics, scattering theory . . . can be dealt with.
- But the wave function cannot do it all — e.g. neither particle spin nor photons can be described in terms of wave functions alone (for spin), or at all (photons) — **BAD**.
- What is required is a more general formulation of the quantum theory which can be applied to any physical system.

- This more general way is due to Heisenberg, Dirac, Jordan, Feynman and others.
 - Call it *quantum mechanics* to distinguish it from wave or matrix mechanics (to which it is closely related).
- The underlying mathematics of quantum mechanics is the mathematics of complex vector spaces i.e. **linear algebra** . . .
 - vectors, linear combinations, operators, matrices, eigenvectors, eigenvalues . . .
 - Fits naturally into the information interpretation of quantum mechanics.
- But is quantum mechanics *different* from wave mechanics?
 - Nope: wave mechanics emerges as one 'representation' of quantum mechanics: the **position representation**.
 - Involves working with $\Psi(x, t)$ which gives explicit information on the position of a particle, less directly all other information.
 - There are many other possible representations, (momentum, energy . . .) all superficially different, but all completely equivalent.

Constructing Quantum Mechanics

- To construct quantum mechanics, we start with, curiously enough, wave mechanics.
- We use the known properties of waves in producing interference patterns, and the fact that particles can exhibit wave-like behaviour to analyse the double slit interference experiment.



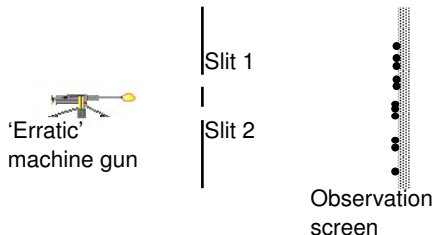
- From this we extract general principles which we assume to apply to all physical systems.
- This approach is due to Feynman who says . . .
 - 'This experiment embodies the essential mystery of quantum mechanics'.

THE TWO SLIT EXPERIMENT

- Shall be considered in three forms:
 - With macroscopic particles (bullets);
 - With 'classical' waves (light waves);
 - With electrons.
- The first two merely show us what we would expect in our classical world.
- The third gives counterintuitive results with both wave and particle characteristics that have no classical explanation.

An Experiment with Bullets: Both Slits Open.

- We notice three things about this experiment:



- The bullets arrive in lumps: for each bullet that gets through the slits, there is a single impact on the observation screen.
- The bullets arrive at random.

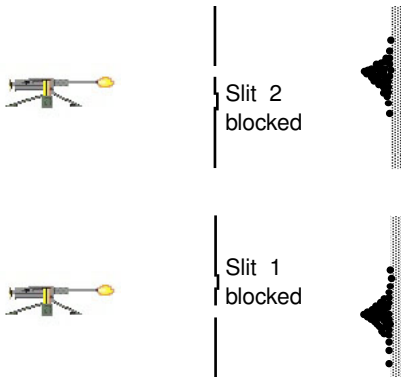
- If we wait long enough, we find that the bullet arrivals tend to form a pattern:



- This is more or less as we might expect:
 - The bullets passing through hole 1 pile up opposite hole 1 . . .
 - The bullets passing through hole 2 pile up opposite hole 2.

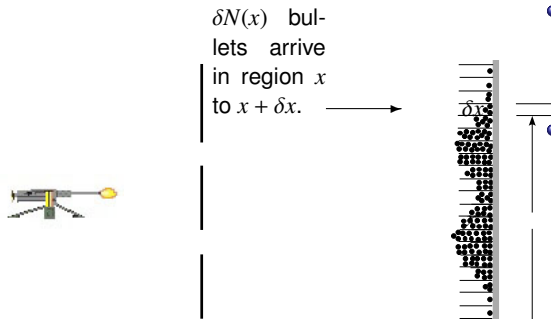
An Experiment with bullets: one slit open

- If we were to perform the experiment with one or the other of the slits closed, would expect something like:



- i.e. bullet arrivals accumulate opposite the open slits.
- From what we know about bullets — they are ‘classical’ objects — we expect that the result observed with both slits open is the ‘sum’ of the results obtained with one and then the other slit open.
- We can quantify this statement by constructing a *histogram* of how the bullets distribute themselves across the observation screen.

- Divide the observation screen into segments of length δx :



- Bullets that have passed through the first screen collected in boxes all of the same size δx .
- If N bullets are fired from the gun, then for N large

$$\frac{\delta N(x)}{N} \approx \text{probability of a bullet striking screen between } x \text{ and } x + \delta x. \\ = P_{12}(x)\delta x$$

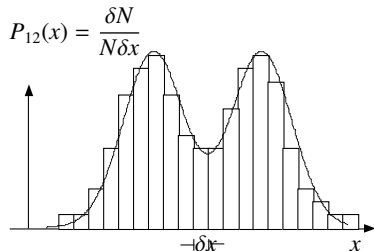
- Now form the ratio

$$P(x) = \frac{\delta N}{N\delta x}$$

for each segment along the screen and plot the result as a histogram.

Probability distributions of bullet strikes . . .

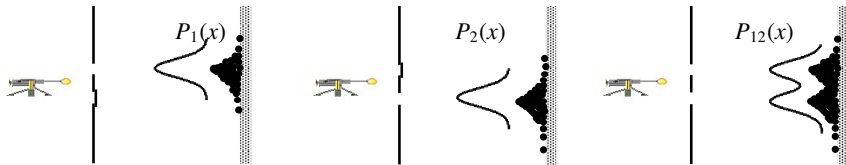
Histogram of bullet strikes



- Making δx smaller and N larger, can join the tops of the histogram to form a smooth curve, $P_{12}(x)$. Then

$P_{12}(x)\delta x =$ probability of a bullet striking screen between x and $x + \delta x$ when both slits are open.

- Can obtain the corresponding curves with one or the other slit open:



- Now make the claim that, for bullets: $P_{12}(x) = P_1(x) + P_2(x)$

- The fact that we add the probabilities $P_1(x)$ and $P_2(x)$ to get $P_{12}(x)$ is simply stating:
 - The probability of a bullet that goes through slit 1 landing in $(x, x + \delta x)$ is *completely independent of whether or not slit 2 is open*.
 - The probability of a bullet that goes through slit 2 landing in $(x, x + \delta x)$ is *completely independent of whether or not slit 1 is open*.
- The above conclusion is perfectly consistent with our classical intuition.
 - Bullets are particles: they arrive in lumps;
 - They independently pass through *either* slit 1 *or* slit 2 before striking the screen;
 - Their random arrivals are due to erratic behaviour of the source (and maybe random bouncing around as they pass through the slit), all of which, in principle can be measured and allowed for and/or controlled.

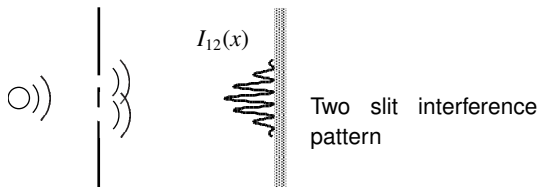
- Now repeat the above series of experiments with *waves* — shall assume light waves of wavelength λ .
- Shall measure the time averaged intensity of the light waves reaching the observation screen.
- Use complex notation for $E(x, t)$:
 - Electric field at x on the screen at time t is $E(x, t) = E(x)e^{-i\omega t}$
- Time averaged intensity of the light at x will be $I(x) = E(x, t)^* E(x, t) = E(x)^2$

An Experiment with Waves: One Slit Open



- $I_1(x)$ and $I_2(x)$ are the intensities of the waves passing through slits 1 and 2 respectively and reaching the screen at x . (They are just the central peak of a single slit diffraction pattern.)
- Waves arrive on screen 'all at once', i.e. they do not arrive in lumps, but . . .
- Results similar to single slits with bullets in that the intensity is peaked directly opposite each open slit.

An Experiment with Waves: Both Slits Open



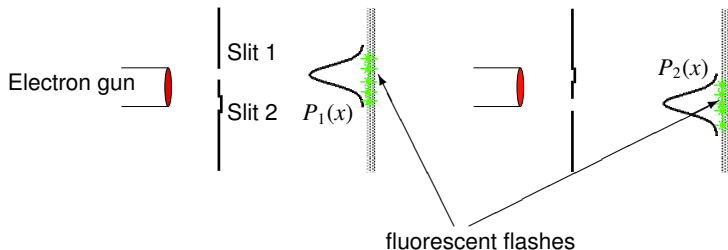
- Interference if both slits are open:

$$\begin{aligned} I_{12}(x) &= |E_1(x, t) + E_2(x, t)|^2 = I_1(x) + I_2(x) + 2E_1E_2 \cos\left(\frac{2\pi d \sin \theta}{\lambda}\right) \\ &= I_1(x) + I_2(x) + 2\sqrt{I_1(x)I_2(x)} \cos \delta \end{aligned}$$

- Last term is the *interference* term which explicitly depends on slits separation d — the waves ‘probe’ the presence of both slits.
- We notice three things about these experiments:
 - The waves arrive ‘everywhere at once’, i.e. they do not arrive in lumps.
 - The single slit result for waves similar to the single slit result for bullets.
 - We see interference effects if both slits open for waves, but not for bullets.

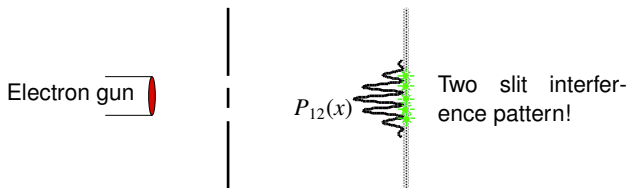
- Now repeat experiment once again, this time with electrons.
- Shall assume a beam of electrons, all with same energy E and momentum p incident on a screen with two slits.
- Shall also assume weak source: electrons pass through the apparatus one at a time.
- Electrons strike a fluorescent screen (causing a flash of light) whose position can be monitored.

An experiment with electrons: one slit open



- With one slit open observe electrons striking fluorescent screen in a random fashion, but mostly directly opposite the open slit — exactly as observed with bullets.
- Can construct probability distributions $P_1(x)$ and $P_2(x)$ for where electrons strike, as for bullets.
- Apparent randomness in arrival of electrons at the screen could be put down to variations in the electron gun (cf. erratic machine gun for bullets).

An experiment with electrons: both slits open



- The following things can be noted:
 - Electrons strike the screen causing individual flashes, i.e. they arrive as particles, just as bullets do;
 - They strike the screen at random — same as for bullets.
 - Can construct probability histograms $P_1(x)$, $P_2(x)$ and $P_{12}(x)$ exactly as for bullets.
- Find that the electron arrivals, and hence the probabilities, form an *interference* pattern — as observed with waves:

$$P_{12}(x) = P_1(x) + P_2(x) + 2\sqrt{P_1(x)P_2(x)} \cos\left(\frac{2\pi d \sin \theta}{\lambda}\right)$$

$\neq P_1(x) + P_2(x)$ — The expected result for particles.

An experiment with electrons: both slits open

Interference of de Broglie waves

- Problem: we have particles arriving in lumps, just like bullets, i.e. one at a time at localised points in space . . .
- but the pattern formed is that of waves . . .
- and waves must pass through both slits simultaneously, and arrive 'everywhere' on the observation screen at once to form an interference pattern.
- Moreover, from the pattern can determine that $\lambda = p/h$ — the de Broglie relation.
- Interference of de Broglie waves seems to have occurred here!

What is going on here?

- If electrons are particles (like bullets) then each one must go through either slit 1 or slit 2.
 - A particle has no extension in space, so if it passes through slit 1 say, it cannot possibly be affected by whether or not slit 2, a distance d away, is open.
 - That's why we claim we ought to find that

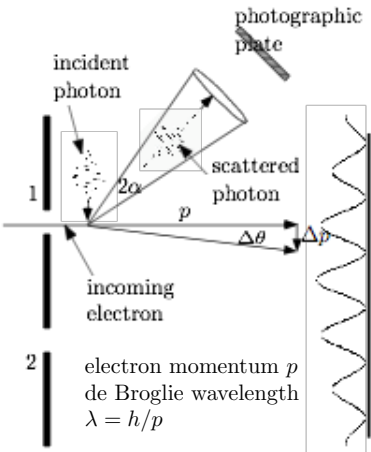
$$P_{12}(x) = P_1(x) + P_2(x)$$

— but we don't.

- In fact, the detailed structure of the interference pattern depends on d , the separation of the slits.
 - So, if an electron passes through slit 1, it must somehow become aware of the presence of slit 2 a distance d away, in order to 'know' where to land on the observation screen so as to produce a pattern that depends on d .
 - i.e. the electrons would have to do some strange things in order to 'know' about the presence of both slits such as travel from slit 1 to slit 2 then to the observation screen.
- Maybe we have to conclude that it is *not true* that the electrons go through either slit 1 or slit 2.

- We can check which slit the electrons go through by watching next to each slit and taking note of when an electron goes through each slit.
- If we do that, we get the alarming result that the interference pattern disappears — we regain the result for bullets.
- It is possible to provide an ‘explanation’ of this result in terms of the observation process unavoidably disturbing the state of the electron.
- Such explanations typically rely on invoking the Heisenberg Uncertainty Principle.

The Heisenberg Microscope



- Weak incident light field: minimum of one photon scatters off electron.
- In order to distinguish the images of the slits in the photographic plate of the microscope, require wavelength of photon at least $\lambda_l \approx d$. Momentum of photon is then $\geq h/d$
- Photon-electron collision gives electron a sideways 'momentum kick' of $\Delta p = h/d$.
- So electron deflected by angle of

$$\Delta\theta \approx \Delta p/p \approx (h/\lambda \cdot d/h) = d/\lambda$$

- $\Delta\theta$ is approximately the angular separation between an interference maximum and a neighbouring minimum!

An uncertainty relation explanation

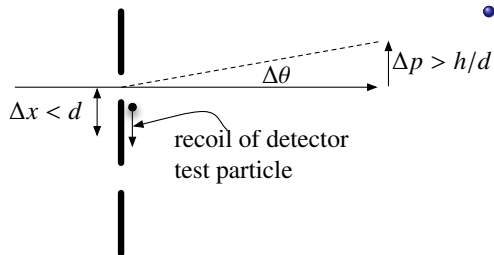
- Recall the Heisenberg uncertainty relation for a particle:

$$\Delta x \Delta p \geq \frac{1}{2} \hbar$$

where Δx is the uncertainty in position and Δp is the uncertainty in momentum.

- We could also argue that we are trying to measure the position of the electron to better than an uncertainty $\Delta x \approx d$, the slit separation
 - This implies an uncertainty in 'sideways momentum' of $\Delta p \geq \hbar/2d$
 - This is similar to (but not quite the same as) the deflection $\Delta p = h/d$ obtained in the microscope experiment — a more refined value for Δx is needed.
 - I.e., the requirement that the uncertainty principle (a fundamental result of wave mechanics) has to be obeyed implies that the interference fringes are washed out.
- What if we observed the position of the electron by some other means?

Other possible explanations



- E.g. make use of the fact that the electron has an electric charge
 - So, place a charge detector near each slit, for instance a test charge that moves when it feels the effect of the electric charge of the electron.
 - But the position of this charge must be specified with an accuracy of better than d .
- Otherwise, we will not know where the electron was (slit 1 or slit 2) that made it move.
 - So the detector test particle will have an uncertainty in momentum of $\Delta p \approx h/d$.
 - Since momentum is conserved in the interaction between the electron and the test particle, if the test particle has a momentum uncertainty Δp , the electron will acquire a momentum uncertainty of Δp .
- And we are back where we were with the microscope example.

- A similar kind of argument can be used for *any* means of measuring which slit the electron passes through.

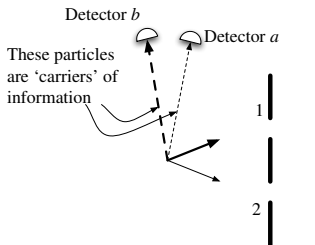
- The uncertainty relation 'gets in the way' somewhere.

- What we can claim is that:

Any attempt to measure which slit the electron passes through results in no interference pattern!

- The uncertainty relation appears to be a fundamental law that applies to *all* physical situations to guarantee that if we observe 'which path', then we find 'no interference'.
- But: it also appears that a physical intervention, an 'observation' is needed to physically scrambled the electron's momentum.
- In fact, this is not the case!!

The significance of 'which path' *information*



- The particle source sends out *pairs* of particles (typically photons in practice, but we'll stick with electrons)

- One heads towards one slit or the other
- The other heads off at right angles.

- There is *no interference pattern*

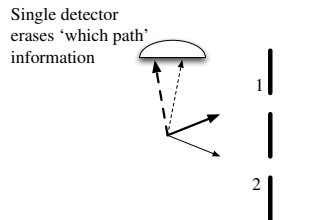
- In principle, can determine which slit an electron will pass through by *detecting* the correlated particle.

- I.e., if detect particle at *a*, then electron is heading for slit 1.

- No need for detection to get no interference.

- With a single detector, cannot tell which correlated particle is which – the 'which path' information is *erased*.

- There is now an interference pattern . . . but . . .



- **There has been no physical interaction with the electrons passing through the slits!!**

- This is an example of entanglement: the direction of the electron heading towards the slits is *entangled* with the direction of the ‘information carrier’ particles.
- **If there is information present as to which slit an electron passes through — even if we do not access this information — then there is no interference pattern!!**
- What happened to the uncertainty principle?
 - It's still there: the particle is either heading for slit 1 or slit 2 – we can determine which by checking which of detector *a* or detector *b* ‘goes off’.
 - The position of the electron is then known to an uncertainty $\Delta x \ll d$, and correspondingly will have a ‘sideways’ uncertainty of momentum $\Delta p \gg \hbar/2d$, more than enough for no interference to form.
 - But the role of the uncertainty principle fades into the background – it is the role of accessible information that matters.

What do we mean by 'observe'

- Have often used the phrase 'observing the slits' or 'observing the system' and so on.
 - Seems to imply the need for a conscious, sentient being to be observing, and aware of what is being observed.
- But, observation is really tied up with *irreversible creation of a macroscopic record*
 - It seems that provided monitoring a quantum system results in the permanent creation of a record of the outcome of the measurement
 - Whether in the brain of the experimenter
 - Stored in a computer memory
 - Carried irreversibly into outer space by a photon emitted by a decaying atom
 - Turned into a macroscopic current pulse
 - Lost as heat in a thermodynamic reservoir . . .
 - Then an 'observation' has been made.
- The need for irreversible dynamics, and the contact with the surrounding environment play essential roles in this process.

- Recall: the probability distribution of the electron arrivals on the observation screen is

$$P_{12}(x) = P_1(x) + P_2(x) + 2\sqrt{P_1(x)P_2(x)} \cos \delta$$

- This can be understood as arising from the interference of de Broglie waves, each emanating from slit 1 or slit 2, i.e.

$$\Psi_{12}(x, t) = \Psi_1(x, t) + \Psi_2(x, t) \quad \Psi(x, t) = \psi(x)e^{-iEt/\hbar} = |\psi(x)|e^{i\theta} e^{-i\omega t}.$$

- Then, according to the Born interpretation of the wave function

$$P(x, t) = |\Psi(x, t)|^2 dx = \text{probability of finding the particle in region } (x, x + dx) \text{ at time } t.$$

we get

$$\begin{aligned} P_{12}(x) &= |\Psi_{12}(x, t)|^2 = |\Psi_1(x, t) + \Psi_2(x, t)|^2 \\ &= |\Psi_1(x, t)|^2 + |\Psi_2(x, t)|^2 + 2\text{Re} [\Psi_1^*(x, t)\Psi_2(x, t)] \\ &= P_1(x) + P_2(x) + 2\sqrt{P_1(x)P_2(x)} \cos \delta, \quad \delta = \theta_1 - \theta_2 \end{aligned}$$

(this is steady state, so probabilities are time independent)

Adding probability amplitudes vs adding probabilities

- As the probabilities are calculated from the square of the wave function, the wave function is also known as a probability amplitude.
 - This is a terminology used even when we leave wave functions far behind.
- If we observe which slit the electrons pass through, then the interference terms are lost, and we get

$$P_{12}(x) = P_1(x) + P_2(x)$$

which is exactly what we expected for classical particles (e.g. bullets).

- A new rule is emerging here (one that will be generalized):

If the path followed by the electrons IS NOT OBSERVED, then the final probability of detecting the electrons is found by adding the probability amplitudes associated with each path, then squaring the result.

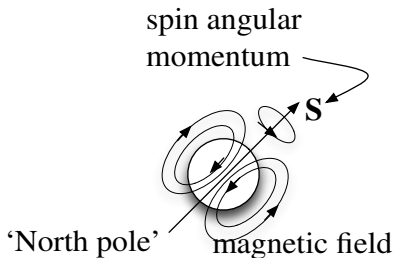
If the path followed IS OBSERVED, then the final probability is found by adding the probabilities associated with each path.

Is the wave function all that is needed?

- Wave mechanics, based on the use of the wave function, is ideally suited for analyzing and describing the properties of a particle (or particles) moving through space.
- But there are many other kinds of physical systems for which the wave function does not seem to work:
 - There is no wave function for the photon.
 - Particles can be created or destroyed – e.g. photons emitted or absorbed by atoms. It is the number of photons that is important here.
 - The spin of elementary particles cannot be explained by use of the wave function.
- If quantum mechanics is supposed to be a theory about the properties of information for all physical systems, then there must be a better, or more general mathematical language.
- But, first, a look at particle spin.

- Shall use particle spin to illustrate a kind of physical system whose quantum properties cannot be described in terms of a wave function . . .
- . . . and also to introduce a particularly simple system with which to illustrate and develop most of the quantum mechanics that follows.

Classical spin



- Classically, an electron can be looked on as a charged spinning sphere.
- As it is charged, the spin of the electron means that there is a circulating current and hence a dipolar magnetic field.
- The strength of the dipolar field is measured by the dipole moment $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \frac{-e}{2m_e} g \mathbf{S} \quad g = \text{gyromagnetic ratio} \approx 2$$

Comparison of classical and quantum spin of an electron

| Classical | Quantum |
|--|--|
| Electron ought to have a finite radius | Experiments indicate a point particle (radius $< 10^{-17}$ m.) |
| $g = 1$ for a solid sphere | $g \approx 2$ |
| \mathbf{S} can have any magnitude | $S = \sqrt{3}\hbar/2$ |
| S_x, S_y, S_z can have any value | Always find S_x etc = $\pm \frac{1}{2}\hbar$. In fact $\mathbf{S} \cdot \hat{\mathbf{n}} = \pm \frac{1}{2}\hbar$ for any unit vector $\hat{\mathbf{n}}$ |

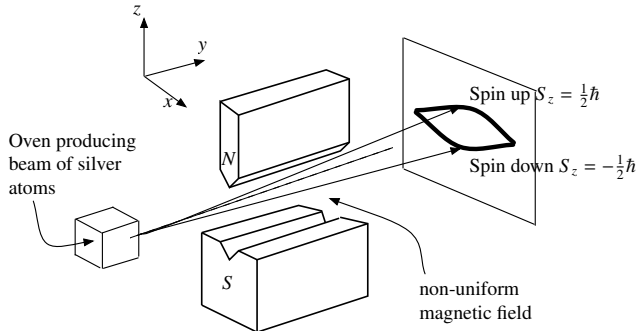
- For a classical sphere of radius 10^{-17} m, and spin angular momentum of $S = \sqrt{3}\hbar/2$, surface of electron moving faster than the speed of light!!!
- Treat spinning electron as a finite sized sphere, and calculate its spin using wave mechanics, find that $S_z = \frac{1}{2}\hbar$ is not allowed (only integer multiples of \hbar allowed!)
- Special relativity + quantum mechanics implies that electron is a point object. Cannot use a wave function description of electron spin – there is no motion of matter through space for a point spinning object.
- Must accept the quantum picture of electron spin as the intrinsic angular momentum of an object with no extension in space.

- Since the components of \mathbf{S} are *quantized*, so are the components of $\boldsymbol{\mu}$:

$$\mu_z = \frac{-e}{2m_e} g S_z = \pm \frac{e}{4m_e} g \hbar \approx \pm \frac{e\hbar}{2m_e}$$

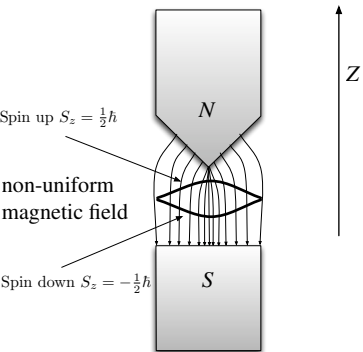
- Shall use this last property of electron spin in the context of the Stern-Gerlach experiment.
- This is an experiment originally designed to measure the orbital angular momentum of silver atoms
 - It confirmed an incorrect theory!!
 - Later recognized as confirmation of the existence of electron spin earlier predicted by the analysis of atomic spectra.

The Stern-Gerlach experiment



- Silver atoms have 47 electrons. Of these, the (spin + orbital) angular momentum of 46 of them add to give a total of zero, and no associated magnetic field.
- The remaining unpaired electron contributes a magnetic moment of $\boldsymbol{\mu} \approx \frac{-e}{m_e} \mathbf{S}$.
- Because of this magnetic moment, a silver atom will *precess* in a magnetic field, but it will also move in space in a non-uniform magnetic field.

Beam splitting by SG apparatus



- The potential energy of a magnetic dipole in a magnetic field is

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}(x, y, z) \approx -\mu_z B_z$$

for field (mostly) in z direction.

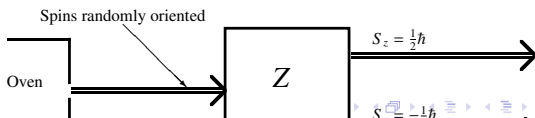
- So the force in the z direction is

$$F_z = -\frac{\partial U}{\partial z} = \mu_z \frac{\partial B_z}{\partial z}.$$

- Recall: $\mu_z \approx \frac{-e}{m_e} S_z = \pm \frac{e\hbar}{2m_e}$

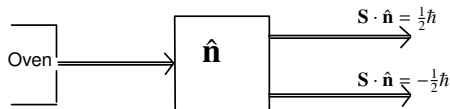
- So the beam will split into two, depending on the value of $S_z = \pm \frac{1}{2}\hbar$.

- Diagrammatic representation of SG apparatus

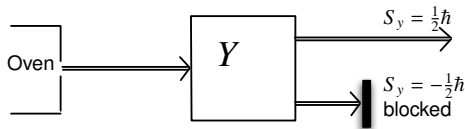


Typical SG diagrams

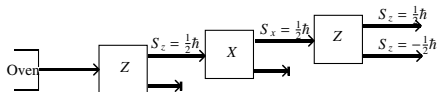
- Magnetic field in arbitrary direction represented by unit vector \hat{n} .



- Magnetic field in y direction ($\hat{n} = \hat{j}$) with one exit beam blocked.



- Can also arrange sequences of SG devices.



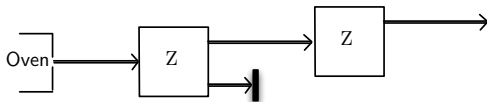
- Shall make use of the Stern-Gerlach apparatus to analyse the quantum properties of spin half in a way analogous to the two slit experiment. Shall consider
 - Repeated spin measurements
 - Quantum randomness for spin
 - Quantum interference for spin

The Stern Gerlach apparatus is a spin measuring device

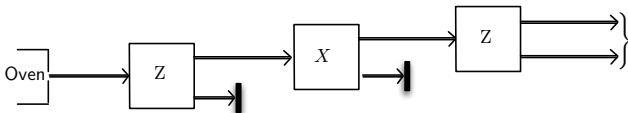
- If we pass one atom through the apparatus, it will emerge in either the spin up or the spin down beam.
 - Suppose it exits in the $S_z = \frac{1}{2}\hbar$ beam.
 - We now know the S_z component of the total spin of that atom
 - We have *measured* this spin component to have the value $S_z = \frac{1}{2}\hbar$.

Repeated Measurements & quantum randomness

- If an atom has been measured to have $S_z = \frac{1}{2}\hbar$, then an immediate remeasurement merely confirms this result.



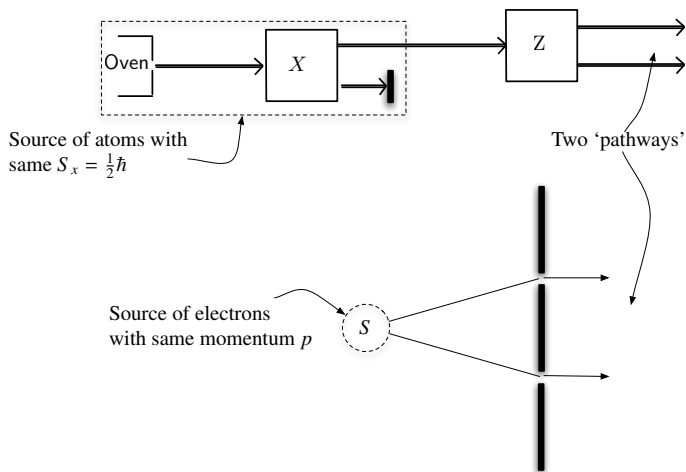
- But if there is an intervening measurement of some other component of the spin:



- The two possible outcomes, $S_z = \pm\frac{1}{2}\hbar$, occur at random, but with equal probability!
- Atom is measured to have definite values of $S_z = \frac{1}{2}\hbar$ and then $S_x = \frac{1}{2}\hbar$, but a remeasurement of S_z gives random results.
 - The intervening measurement of the X component of spin has randomly scrambled the previous measured result of $S_z = \frac{1}{2}\hbar$.
- This is reminiscent of what we have learned about measuring position and momentum of a particle.

Quantum interference for spin

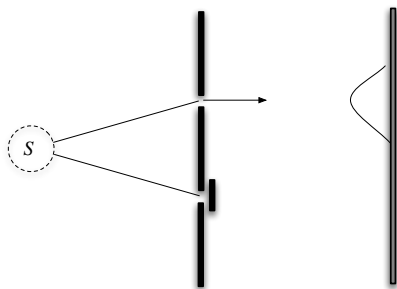
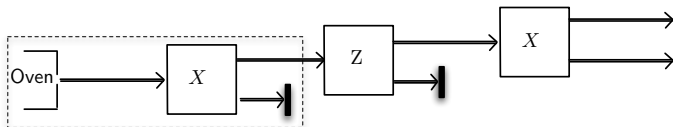
- Compare the SG experiment with the two slit experiment:



Quantum interference for spin

Blocking one slit/spin beam

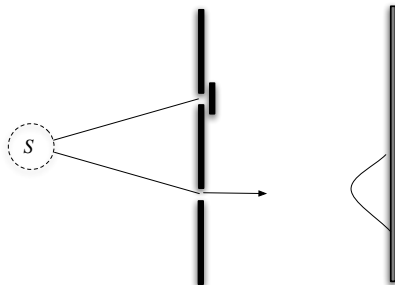
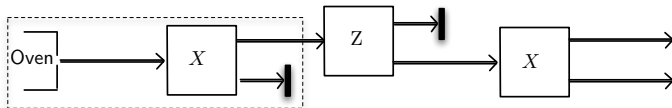
- In this case, there is no interference for the two slit experiment, and two output beams in the SG experiment.



Quantum interference for spin

Blocking one slit/spin beam

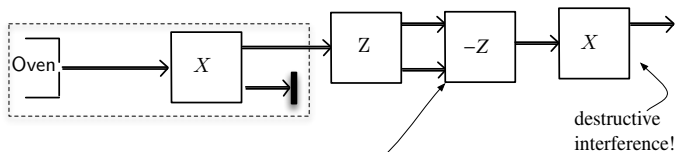
- Once again, blocking the other beam, there is no interference for the two slit experiment, and two output beams in the SG experiment.



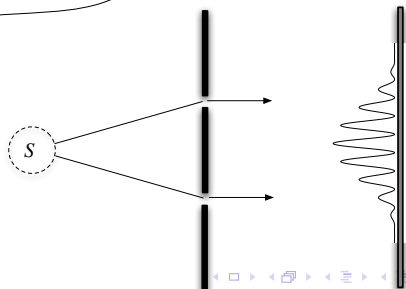
Quantum interference for spin

Now blocking none

- Now there is interference for the two slit experiment, and only one output beam in the SG experiment.
- The second beam has been cancelled out by 'destructive interference'.



Needed to bring the beams together



Quantum interference for spin

Summary of conclusions

- If we DO know which spin path the atoms follow, we find that atoms emerge with either of $S_x = \pm \frac{1}{2} \hbar$.
- If we DO NOT know the spin path of atoms, we get cancellation of the $S_x = -\frac{1}{2} \hbar$ beam.
- This can be taken as evidence of interference occurring, in the same way as for the two slit experiment.
- Suggests that we assign *probability amplitudes* for the different possible spin paths that the atoms follow through the Stern-Gerlach apparatuses.
- In fact, it is possible to show, from the observed properties of the detection probabilities as a function of the orientation of the final SG apparatus (which we have set to be in the x direction here), that a probability amplitude interpretation is 'built in' to the observed physical behaviour of spin half.

The Mathematical Language of Quantum Mechanics

- Have seen that two disparate physical systems:

- Two slit experiment
- Spin half system in SG experiment

have exhibited such behaviour as:

- intrinsic, irreducible randomness
- interference effects (which are dependent on information known about the system)

that have no explanation according to classical physics.

- The behaviour is suggestive of non-classical fundamental laws of physics in action.
- To state these laws, the language of classical physics is inadequate. For instance, for particle motion through space we need the notion of a wave function $\Psi(\mathbf{r}, t)$:
 - The wave function gives us information on position, momentum, energy, . . .
 - But Ψ cannot describe spin — spin discrete, position continuous

- A new mathematical language is needed that
 - Captures the content of these new physical laws (whatever they are)
 - Can take different mathematical forms capable of describing particle motion through space, or particle spin, or numbers of particles in a many-particle system . . .
- The mathematical formalism that does the job is that of complex linear vector spaces.
- First formulated heuristically by Paul Dirac, then formalised by John von Neumann and others
- The mathematical formalism is particularly suited to the ‘information content’ perspective on quantum theory
- Brings into sharp focus the need to be clear about the notion of the
state of a physical system

The state of a system

- Specifying the state of a system amounts to gathering together all the information that it is possible to know about the system at any instant in time.
- the information should be enough that, in principle:
 - it would be possible to reconstruct the system at some other place and time
 - and for the reconstructed system to behave in exactly the same way thereafter as the original.
- The information gathered usually takes the form of numerical data representing the measured values of e.g. for a single particle, its position or momentum at some instant in time.
 - I.e., the state of a system consists of a list of the values of various physical parameters that can be determined for the system!
 - Though there are more elegant ways of representing the state, e.g. *phase space*.

- **Classically** it is possible, in principle, to determine *exactly* all the quantities needed to specify the state of a physical system, e.g.
 - For a particle, the position **and** the momentum at some instant.
 - For a rigid body, the position **and** the momentum of its centre of mass, the orientation of the body in space, **and** its angular momentum.
 - For an electromagnetic field, the magnetic **and** electric field strengths at each point in space at some instant.
- **Quantumly** the uncertainty principle gets in the way.
 - For a particle, we can specify the position **or** the momentum
 - For a spinning particle, we can specify the magnitude of its spin angular momentum but only **one** component of its spin.
- So, we can still construct a list of physical parameters whose value is known, bearing in mind the above constraints.

- We shall adopt as our definition of the state of a quantum system the following:

| All the data concerning the system that can be known without mutual
interference or contradiction. |

- The symbol $| \ \rangle$ is known as a **ket**.
- Within the ket is a summary of all the data specifying the state of the system e.g.
 - $|x = 3 \text{ cm}\rangle$ would be the state of a particle known to be at the position $x = 3 \text{ cm}$.
 - $|x = 3 \text{ cm}, x = 5 \text{ cm}\rangle$ is NOT a possible state – it has a contradiction. A particle cannot be in two places at once, even in quantum mechanics.
 - $|S_z = \frac{1}{2}\hbar\rangle$ is the spin state of a spin half particle for which S_z has been determined to be $\frac{1}{2}\hbar$ e.g. in an SG experiment.
 - $|S_x = \frac{1}{2}\hbar, S_z = -\frac{1}{2}\hbar\rangle$ is NOT a possible state as we have seen that we cannot pin down the value of two spin components at the same time.
 - $|S_z = \frac{1}{2}\hbar, x = 3 \text{ cm}\rangle$ is the state of a particle for which $S_z = \frac{1}{2}\hbar$ and $x = 3 \text{ cm}$ (the two do not interfere with each other).

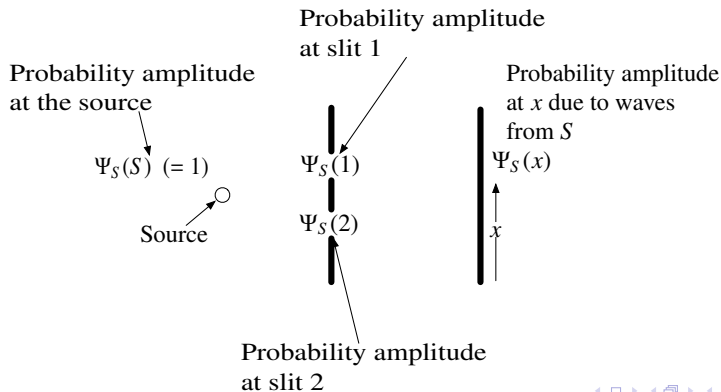
- Generic states would be:
 - $|x\rangle$ – a particle has a position x , left unspecified
 - $|p_x\rangle$ – a particle has an x component of momentum p_x , also left unspecified.
 - For an arbitrary state, we write $|\psi\rangle$ or $|\phi\rangle$ etc.
- A ket is also known as a ket vector, state vector, or sometimes just state.
- A companion to the ket is:

⌊ All the data concerning the system that can be known without mutual
interference or contradiction. ⌋.

- This is known as a bra or bra vector or also state vector or state.

The two slit experiment revisited

- Shall recast the 'sum of amplitudes' in the two slit experiment in a new way.
- The following argument is NOT rigorous, but it does suggest the way in which we will have to think about quantum mechanics.
- The final results can be set up in a more rigorous way for other physical systems, in particular the spin half system.



A consequence of linearity of waves

- Suppose the waves incident at slit 1 has an amplitude of unity
- AND suppose that in that case, the amplitude of the wave that arrives at x from slit 1 is $\Psi_1(x)$.
 - Note the notation convention continues!

- Then IF the wave incident on slit 1 has amplitude $\Psi_S(1)$, then the amplitude of the wave at x will be

$$\Psi_S(1)\Psi_1(x)$$

- Similarly wave from slit 2 will have amplitude $\Psi_S(2)\Psi_2(x)$
- Thus, the TOTAL wave amplitude at x due to waves from the source S will be

$$\Psi_S(x) = \Psi_S(1)\Psi_1(x) + \Psi_S(2)\Psi_2(x).$$

- This last result assumes that probability amplitudes are *linear* – we simply scale things up as appropriate, and add the results, no nonlinear steps at all.

- Have expressed the probability amplitude $\Psi_S(x)$ as a sum of two terms of the form:

$\Psi_S(n)\Psi_n(x) =$ Probability amplitude of observing the particle at slit n given that it originated from the source S .

× Probability amplitude of observing the particle at x given that it originated from slit n .

= Probability amplitude for the particle to pass from the source S to point x through slit n .

- Could also say that $\Psi_S(n)$ is
 - the probability amplitude of observing the particle at the position of slit n , given that it was originally at the position of the source S , or . . .
 - the probability amplitude of observing the particle to be in the state in which it is at the position of slit n , given that it was in a state in which it was at the position of the source S .

Sum of amplitudes in bra(c)ket notation.

- Now use the new notation:

$|S\rangle \equiv$ state of the particle when at the position of the source S

$|n\rangle \equiv$ state of the particle when at the position of slit n

$|x\rangle \equiv$ state of the particle when at the position x

- We can then write, for instance

$\Psi_S(n) =$ Probability amplitude of observing the particle in state $|n\rangle$ given that it was in state $|S\rangle$.

or

$$\Psi_S(n) = \langle n|S\rangle = \langle n|S\rangle$$

i.e. we have written the final state as a bra, and have replaced the double vertical bar by a single bar.

- Similarly $\Psi_S(x) \rightarrow \langle x|S\rangle$; $\Psi_n(x) \rightarrow \langle x|n\rangle$.

- In terms of this new notation we have

$$\langle x|S\rangle = \langle x|1\rangle\langle 1|S\rangle + \langle x|2\rangle\langle 2|S\rangle.$$

- At this point, we step into uncharted territory.

- The equation

$$\langle x|S\rangle = \langle x|1\rangle\langle 1|S\rangle + \langle x|2\rangle\langle 2|S\rangle.$$

is still saying nothing physically new, but . . .

- We note that x is a variable: it is the position on the screen and it is common to every term.
- So, let's 'cancel it out':

$$|S\rangle = |1\rangle\langle 1|S\rangle + |2\rangle\langle 2|S\rangle.$$

- This is a 'template' into which we can re-insert $\langle x|$ to give the complete expression.
- But this expression is also a brand new notation
 - seems to be saying that the state $|S\rangle$ is in some sense composed of the states $|1\rangle$ and $|2\rangle$
 - the weighting in this composition are the probability amplitudes $\langle n|S\rangle$

- The expression captures the essence of the idea that the particle leaving the source can, in some way, be present at both of the slits before it makes its way to the observation screen.

- The expression

$$|S\rangle = |1\rangle\langle 1|S\rangle + |2\rangle\langle 2|S\rangle.$$

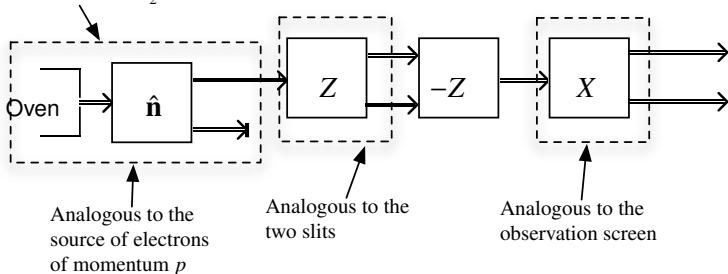
is known as a 'superposition of states'.

- This expression is also somewhat similar to an expression for the sum of two vectors to produce a third
 - This, in fact, is what we end up showing for the states of a quantum system.
 - However, this is not so easy for the two slit example
 - So, we will construct a similar result for the case of spin half.

Probability amplitudes for particle spin I

- Shall now argue by analogy with two slit experiment to set up a 'sum of probability amplitudes' result for spin half .

Source of atoms
for which $S = \mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar$



- Here $\hat{\mathbf{n}}$ is some arbitrary orientation of the magnetic field (in the XZ plane)
- The atom can emerge from the first Stern-Gerlach device in one or the other of two separate beams corresponding to the atomic spin component $S = \mathbf{S} \cdot \hat{\mathbf{n}} = \pm \frac{1}{2}\hbar$

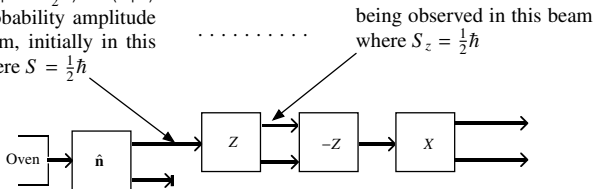
Probability amplitudes for particle spin II

- We will then assign a probability amplitude for an atom to pass along either of the $S_z = \pm \frac{1}{2}\hbar$ beams, written:

$\langle S_z = \pm \frac{1}{2}\hbar | S = \frac{1}{2}\hbar \rangle$ = Probability amplitude of observing the atom to have $S_z = \pm \frac{1}{2}\hbar$ given that originally it had an \hat{n} component of spin $S = \frac{1}{2}\hbar$.

= $\langle \pm | S \rangle$ [analogue of $\langle 1 | S \rangle$ or $\langle 2 | S \rangle$ for two slit experiment.]

$\langle S_z = \frac{1}{2}\hbar | S = \frac{1}{2}\hbar \rangle = \langle + | S \rangle$
is the probability amplitude
of an atom, initially in this
beam where $S = \frac{1}{2}\hbar$



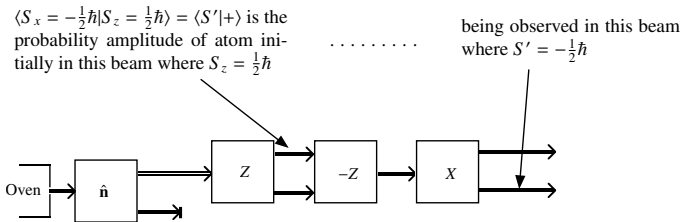
- Note simplification of notation: $\langle S_z = \pm \frac{1}{2}\hbar | S = \frac{1}{2}\hbar \rangle \rightarrow \langle \pm | S \rangle$.

Probability amplitudes for particle spin III

- The atomic beams are then recombined and finally, after passing through the last Stern-Gerlach device, emerge with $S_x = \pm \frac{1}{2}\hbar$. We then write

$\langle S_x = \pm \frac{1}{2}\hbar | S_z = \frac{1}{2}\hbar \rangle$ = Probability amplitude of observing the atom to have $S_x = \pm \frac{1}{2}\hbar$ given that it had a z component of spin $S_z = \frac{1}{2}\hbar$.

= $\langle S' | + \rangle$ [analogue of $\langle x | 1 \rangle$ for two slit experiment.]



- Here $S' = S_x$ can have the values of $\pm \frac{1}{2}\hbar$.

Probability amplitudes for particle spin IV

- A similar definition applies for $\langle S_x = \pm \frac{1}{2}\hbar | S_z = -\frac{1}{2}\hbar \rangle \rightarrow \langle S'|- \rangle$
- Now construct the probability amplitude of measuring the x component of spin to have the value S' given that initially the $\hat{\mathbf{n}}$ component of spin $\mathbf{S} \cdot \hat{\mathbf{n}} = S = \frac{1}{2}\hbar$

- Can be done by analogy with two slit case

$$\begin{array}{ccccc} \langle S'|S \rangle & \langle S'|+ \rangle & \langle +|S \rangle & \langle S'|- \rangle & \langle -|S \rangle \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \langle x|S \rangle & \langle x|1 \rangle & \langle 1|S \rangle & \langle x|2 \rangle & \langle 2|S \rangle \end{array}$$

- Can be done by a more formal mathematical route.
- Either way, this probability amplitude is given by

$$\langle x|S \rangle = \langle x|1 \rangle \langle 1|S \rangle + \langle x|2 \rangle \langle 2|S \rangle \quad \text{two slit experiment}$$

$$\langle S'|S \rangle = \langle S'|+ \rangle \langle +|S \rangle + \langle S'|- \rangle \langle -|S \rangle \quad \text{spin half}$$

- We now want to obtain from this result the idea of the spin state of the atoms emerging from the first Stern-Gerlach apparatus as being 'made up of' the states $|\pm\rangle$.

Superposition of States for Spin Half I

- As in the two slit result, we can argue that

$$\langle S'|S\rangle = \langle S'|+\rangle\langle +|S\rangle + \langle S'|- \rangle\langle -|S\rangle$$

holds for all final states $|S'\rangle$ so can try the 'cancellation' trick to give:

$$|S\rangle = |+\rangle\langle +|S\rangle + |-\rangle\langle -|S\rangle.$$

- Can reintroduce ' $\langle S'|$ ' back into this 'template' for any chosen final state.
- Note that $\langle \pm|S\rangle$ is the probability amplitude of the z component of spin having the values $\pm\frac{1}{2}\hbar$ given that initially the \hat{n} component of spin was $S = \frac{1}{2}\hbar$, i.e.

$$|\langle \pm|S\rangle|^2 = \text{probability of the atomic spin being in the state } |\pm\rangle \text{ given that it was in state } |S\rangle$$

- The coefficients $\langle +|S\rangle$ and $\langle -|S\rangle$ in some sense determine 'how much' of the state $|+\rangle$ is to be found in the initial state $|S\rangle$ and 'how much' of state $|-\rangle$ is to be found in $|S\rangle$.

Superposition of States for Spin Half II

- Again have the result that the state of a physical system can, in a sense:
 - be expressed in terms of other possible states of the system
 - with each other possible state having a 'weight' that determines the probability of observing the system in each of these other states.
- Thus, the state $|S\rangle$ is in some sense 'made up of' the two states $|+\rangle$ and $|-\rangle$ with a 'weighting' of $\langle+|S\rangle$ and $\langle-|S\rangle$ respectively.
 - The squares of these weights — $|\langle+|S\rangle|^2$ and $|\langle-|S\rangle|^2$ — are then the probabilities of observing the system in either of the two states $|+\rangle$ and $|-\rangle$ respectively.
- The expression

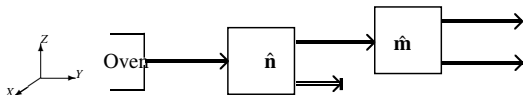
$$|S\rangle = |+\rangle\langle+|S\rangle + |-\rangle\langle-|S\rangle.$$

is known as a superposition of states.

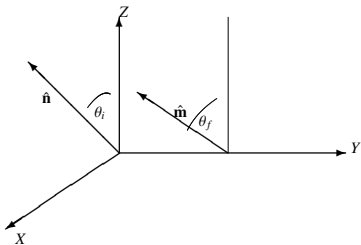
- But first, some expressions for the probability amplitudes $\langle\pm|S\rangle$.

Spin half probabilities I

- Consider the fairly general experimental arrangement



- Here, the first magnetic field makes an angle of θ_i with the z direction and the second an angle of θ_f with the z direction.



- The angle between the directions of the two magnetic fields is $\theta = \theta_f - \theta_i$.
- From experiment: probability of an atom emerging in the beam for which $S_f = \mathbf{S} \cdot \hat{\mathbf{m}} = \frac{1}{2}\hbar$, given that it entered the last Stern-Gerlach device with $S_i = \mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar$ is

$$P(S_f = \frac{1}{2}\hbar | S_i = \frac{1}{2}\hbar) = \cos^2(\frac{1}{2}\theta) = \cos^2[(\theta_f - \theta_i)/2].$$

- If $\theta = 0$ probability is unity
 - The magnetic fields in the two SG apparatuses are parallel.
 - An atom emerging from first apparatus with $\mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar$ is guaranteed to be measured to have $\mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar$ by second apparatus.
- $\theta = \pi$, $\hat{\mathbf{m}} = -\hat{\mathbf{n}}$ probability vanishes.
 - Atom emerges from first SG with $\mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar$ so it has zero probability of being measured to have $\mathbf{S} \cdot \hat{\mathbf{n}} = -\frac{1}{2}\hbar$ by second apparatus.

Spin half probability amplitudes

- Can *almost* work out probability amplitudes:

$$P(S_f = \frac{1}{2}\hbar | S_i = \frac{1}{2}\hbar) = |\langle S_f = \frac{1}{2}\hbar | S_i = \frac{1}{2}\hbar \rangle|^2 = \cos^2(\frac{1}{2}\theta)$$

\Downarrow

$$\langle S_f = \frac{1}{2}\hbar | S_i = \frac{1}{2}\hbar \rangle = e^{i\phi} \cos \frac{1}{2}\theta$$

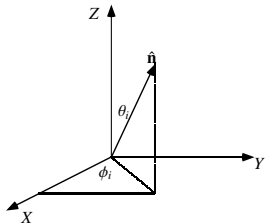
- Here $\exp(i\phi)$ is an **unknown** phase factor.
- For example, for $\langle S_z = \frac{1}{2}\hbar | S_x = \frac{1}{2}\hbar \rangle$, we have $\theta_f = 0$ and $\theta_i = \pi/2$

so that

$$P(S_z = \frac{1}{2}\hbar | S_x = \frac{1}{2}\hbar) = \cos^2(\pi/4) = \frac{1}{2}$$

and

$$\langle S_z = \frac{1}{2}\hbar | S_x = \frac{1}{2}\hbar \rangle = \frac{e^{i\phi}}{\sqrt{2}}$$



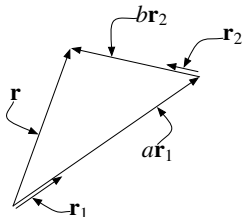
- For arbitrary $\hat{\mathbf{n}} = \sin \theta_i \cos \phi_i \hat{\mathbf{i}} + \sin \theta_i \sin \phi_i \hat{\mathbf{j}} + \cos \theta_i \hat{\mathbf{k}}$

$$|S\rangle = \cos(\frac{1}{2}\theta)|+\rangle + e^{i\phi} \sin(\frac{1}{2}\theta)|-\rangle.$$

- Shall ultimately show that the spin half superposition of states expression can be given a vectorial interpretation.
- But we need to look briefly at the concept of a vector to set the scene.
- The prototype of a vector is the position (or displacement) vectors in ordinary 3-D space.
 - Shall use 2-D vectors to illustrate the ideas.
 - But these need to be extended to the more abstract notion of a 'complex inner product vector space'.

- **Linear combinations – adding vectors**

- Consider two non-collinear vectors in 2-D space (i.e. a plane): \mathbf{r}_1 and \mathbf{r}_2 .



- Multiples of these pair of vectors can be added together by the *triangle addition rule* to form another vector:

$$\mathbf{r} = a\mathbf{r}_1 + b\mathbf{r}_2 \quad a, b \text{ real numbers.}$$

- Called a *linear combination*

- Conversely, any vector \mathbf{r} can be expressed in terms of \mathbf{r}_1 and \mathbf{r}_2 using appropriate values for the components a and b of \mathbf{r}

- Combining vectors produces other vectors, analogous to combining states produces other states.

- **Components**

- a and b are components of \mathbf{r} and a measure of how much of \mathbf{r}_1 and \mathbf{r}_2 respectively go towards making up the vector \mathbf{r} .

• Linear independence

- The vectors \mathbf{r}_1 and \mathbf{r}_2 are not collinear.
 - The only choice of a and b such that $a\mathbf{r}_1 + b\mathbf{r}_2 = 0$ is $a, b = 0$
 - The vectors \mathbf{r}_1 and \mathbf{r}_2 are said to be *linearly independent*.

• Basis vectors

- The two vectors \mathbf{r}_1 and \mathbf{r}_2 are such that *any* vector \mathbf{r} can be written as a linear combination of them.
 - They are basic items with which any vector can be constructed, hence the term *basis vectors*
 - An infinite number of choices for the basis vectors, including choosing three or more vectors.

• Dimension

- But if we insist on the basis vectors being *linearly independent* then the minimum number here is two.
- Identified with the dimension of the space in which the position vectors reside.

• Scalar product

- Position vectors in space have the properties of length and orientation.
- This leads to the idea of a **scalar product** of two vectors \mathbf{r}_1 and \mathbf{r}_2 :

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = r_1 r_2 \cos \theta$$

where r_1 and r_2 are the lengths of \mathbf{r}_1 and \mathbf{r}_2 respectively, and θ is the angle between them.

- Then $\mathbf{r} \cdot \mathbf{r} = r^2 = (\text{length})^2$ or $(\text{norm})^2$.
- If $\mathbf{r}_1 \cdot \mathbf{r}_2 = 0$ then \mathbf{r}_1 and \mathbf{r}_2 are **orthogonal**.

• Unit vectors

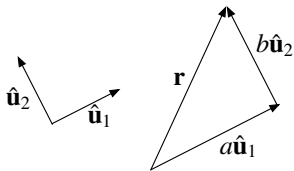
- Consider two vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ that satisfy

$$\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2 \cdot \hat{\mathbf{u}}_2 = 1$$

i.e. they have unit length or unit norm, hence they are known as unit vectors.

- And if $\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 = 0$ then they are **orthonormal**.

- **Orthonormal basis vectors**



- Any vector \mathbf{r} can be expressed using appropriate values for the components of \mathbf{r} , i.e.

$$\mathbf{r} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2.$$

- The components a and b represent 'how much' of the vector \mathbf{r} is made up of the vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$.

- Given by

$$a = \hat{\mathbf{u}}_1 \cdot \mathbf{r} \quad \text{and} \quad b = \hat{\mathbf{u}}_2 \cdot \mathbf{r}$$

- A well known example is to put $\hat{\mathbf{u}}_1 = \hat{\mathbf{i}}$ and $\hat{\mathbf{u}}_2 = \hat{\mathbf{j}}$:

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}.$$

Generalization of the primitive notion of a vector

- Can generalize the above ideas in a number of ways.
- Allow the dimension of the space in which the vectors reside to be three dimensional, or four, or five, . . . , or even infinite
 - Just add on more orthonormal unit vectors $\hat{\mathbf{u}}_3, \hat{\mathbf{u}}_4, \dots$ where

$$\begin{aligned}\hat{\mathbf{u}}_m \cdot \hat{\mathbf{u}}_n &= 0 & n &= m \\ &= 1 & n &\neq m\end{aligned}$$

- Keep the same rules (triangle addition rule) for forming linear combinations of vectors.
 - We are not dealing with ordinary space anymore, Toto.
- Allow the components in a linear combination to be complex numbers:

$$\mathbf{r} = z_1 \hat{\mathbf{u}}_1 + z_2 \hat{\mathbf{u}}_2$$

- **Definitely** not dealing with ordinary space anymore.
 - Absolute values $|z_1|$ and $|z_2|$ will then represent the extent to which \mathbf{r} is made up of the two vectors \mathbf{u}_1 and \mathbf{u}_2 .
 - Need to 'fix up' the idea of inner product.

Inner or scalar product for complex vectors.

- What is the 'length' of a complex vector?
 - Recall, for real vector \mathbf{r} : $\text{length} = \sqrt{\mathbf{r} \cdot \mathbf{r}}$.
- So, redefine scalar product of two complex vectors \mathbf{r}_1 and \mathbf{r}_2 by

$$\mathbf{r}_1^* \cdot \mathbf{r}_2 = (\mathbf{r}_2 \cdot \mathbf{r}_1)^* .$$

- Then, $\mathbf{r}_1^* \cdot \mathbf{r}_1$ will be real and positive (or zero), so can put

$$\text{length, or norm} = |\mathbf{r}| = \sqrt{\mathbf{r}^* \cdot \mathbf{r}} .$$

- An example: $\mathbf{r}_1 = 3\mathbf{i} + 4\mathbf{j}$
 $\mathbf{r}_2 = 4\mathbf{i} - 2\mathbf{j}$

Then $\mathbf{r}_1^* \cdot \mathbf{r}_2 = (3\mathbf{i} - 4\mathbf{j}) \cdot (4\mathbf{i} - 2\mathbf{j}) = 12 + 8i$

and the length, or norm of \mathbf{r}_1 is

$$|\mathbf{r}_1| = \sqrt{\mathbf{r}_1^* \cdot \mathbf{r}_1} = \sqrt{(3\mathbf{i} - 4\mathbf{j}) \cdot (3\mathbf{i} + 4\mathbf{j})} = 5 .$$

- The notion of complex vectors in a space of higher (even infinite) dimension is essential in quantum mechanics.
- The general idea of a vector is based on the primitive notion of a real position vector in 2 or 3 dimensional space. Mathematicians then
 - Extended the idea to more dimensions . . .
 - Allowed for complex valued components . . .
 - Extracted the underlying mathematical structure, and defined the concept of a **complex vector space** with an inner product . . .
 - And it turned out to be perfectly suited for describing the states of quantum mechanical systems.
- We do not need all of the associated mathematical machinery, just the easy parts of it.
- The idea is to show that the kets we have introduced earlier have all the properties of vectors in a complex vector space.

Complex linear vector spaces I

- A complex linear vector space \mathcal{V} is a collection of objects called vectors, written $\mathbf{u}, \mathbf{v}, \mathbf{w}, \dots$ which can be combined according to some rule involving the complex numbers, to produce other vectors.

- A possible combination of vectors is written

$$\mathbf{s} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w} + \dots \in \mathcal{V}$$

where a, b, c, \dots are complex numbers. This is a linear combination.

- The term 'linear' comes from the requirement that if

$$\mathbf{s}_1 = a_1\mathbf{u} + b_1\mathbf{v} + c_1\mathbf{w} + \dots \quad \text{and} \quad \mathbf{s}_2 = a_2\mathbf{u} + b_2\mathbf{v} + c_2\mathbf{w} + \dots$$

then

$$\alpha\mathbf{s}_1 + \beta\mathbf{s}_2 = (\alpha a_1 + \beta a_2)\mathbf{u} + (\alpha b_1 + \beta b_2)\mathbf{v} + (\alpha c_1 + \beta c_2)\mathbf{w} + \dots$$

- The vectors $\mathbf{v}_1, \mathbf{v}_2, \dots$ is said to be linearly independent if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + \dots = \mathbf{0}$$

can only occur if $a_1 = a_2 = a_3 = \dots = 0$.

- The smallest number of linearly independent vectors is called the dimension of the vector space. The dimension (N) can be finite or infinite.
- The inner or scalar product of any two vectors is a rule that associates a complex number with a pair of vectors.
 - The inner product of two vectors \mathbf{v}_1 and \mathbf{v}_2 is written either $\mathbf{v}_1^* \cdot \mathbf{v}_2$ or $(\mathbf{v}_1, \mathbf{v}_2)$.
 - The latter version is preferred by mathematicians, and turns into the bra(c)ket notation of quantum mechanics.
- The inner product has the properties

$$(\mathbf{v}_1, \mathbf{v}_2)^* = (\mathbf{v}_2, \mathbf{v}_1) \quad \text{i.e. order matters}$$

$$(a\mathbf{u}_1 + b\mathbf{u}_2, \mathbf{v}) = a^*(\mathbf{u}_1, \mathbf{v}) + b^*(\mathbf{u}_2, \mathbf{v}) \quad \text{easier to see with the } \mathbf{v}_1^* \cdot \mathbf{v}_2 \text{ notation}$$

$$(\mathbf{v}, \mathbf{v}) = \text{real number} \geq 0.$$

which we will see later in the bra(c)ket notation, but for now we will use the 'dot' notation

- More of the usual stuff follows:
 - Norm of a vector: $v = \sqrt{\mathbf{v}^* \cdot \mathbf{v}}$
 - A unit vector has norm = 1.
 - Orthogonal vectors: $\mathbf{v}^* \cdot \mathbf{u} = (\mathbf{v}, \mathbf{u}) = 0$.
 - Orthonormal vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ satisfy

$$\begin{aligned}\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 &= \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_2 = 1 \\ \hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_2 &= \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0.\end{aligned}$$

- An example in ordinary 2-D space:

$$\hat{\mathbf{u}}_1 = \frac{3\mathbf{i} + 4i\mathbf{j}}{5} \quad \hat{\mathbf{u}}_2 = \frac{4\mathbf{i} - 3i\mathbf{j}}{5}$$

- Find that, e.g.

$$\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 = \left(\frac{3\mathbf{i} - 4i\mathbf{j}}{5}\right) \cdot \left(\frac{3\mathbf{i} + 4i\mathbf{j}}{5}\right) = \frac{9 + 16}{25} = 1$$

- The unit vectors $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \dots, \hat{\mathbf{u}}_N$ then form a linearly independent set of orthonormal basis vectors that span the complex vector space \mathcal{V}
 - Any vector in the vector space can be written as a linear combination of these basis vectors:

$$\mathbf{v} = a_1 \hat{\mathbf{u}}_1 + a_2 \hat{\mathbf{u}}_2 + \dots + a_N \hat{\mathbf{u}}_N$$

where the components a_1, a_2, \dots, a_N are the components of \mathbf{v} with respect to this set of basis states.

- As there are N basis states, the vector space has dimension N (which could be infinite).
- The basis vectors are orthonormal:

$$\hat{\mathbf{u}}_m^* \cdot \hat{\mathbf{u}}_n = \delta_{mn}$$

where δ_{mn} is known as the *Kronecker delta* and has the property

$$\delta_{mn} = 1 \quad m = n \quad \delta_{mn} = 0 \quad m \neq n.$$

- Recall the 'sum over amplitudes' result obtained for spin half:

$$\langle S'|S\rangle = \langle S'|+\rangle\langle +|S\rangle + \langle S'|- \rangle\langle -|S\rangle.$$

- By 'cancelling' the common factor ' $\langle S'|$ ' from this we get

$$|S\rangle = |+\rangle\langle +|S\rangle + |-\rangle\langle -|S\rangle$$

- Aim is to show that there is a perfect analogy between this and corresponding relationship:

$$\mathbf{v} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2$$

for ordinary (complex) vectors.

- Further note we can also cancel the common factor $|S\rangle$:

$$\langle S'| = \langle S'|+\rangle\langle +| + \langle S'|- \rangle\langle -|.$$

- This turns out to be the 'complex conjugate' of the result for kets.

The Normalization Condition

- Start with the probability amplitudes $\langle S' | S \rangle$.
 - These are complex numbers in general.
 - For arbitrary spin directions $|\langle S' | S \rangle|^2$ is the probability of observing the spin to be in the state $|S'\rangle$ given that it was in the state $|S\rangle$.
- In particular, $\langle S | S \rangle$ is the probability amplitude of observing the spin to be in the state $|S\rangle$ given that it was in the state $|S\rangle$.
 - This will have to be unity, i.e. $|\langle S | S \rangle|^2 = 1$.
 - Must have $\langle S | S \rangle = e^{i\eta}$ where η is an arbitrary phase.
 - It turns out that this phase always cancels out in any calculation of observable quantities, so it is conventionally set to zero, and hence
$$\langle S | S \rangle = 1.$$
 - The state $|S\rangle$ is said to be **normalized to unity**.
 - In particular $\langle + | + \rangle = 1$.

- We can now consider the probability amplitude $\langle +|S\rangle$ obtained by replacing S' by $+$ in the expression for $\langle S'|S\rangle$:

$$\langle +|S\rangle = \langle +|+\rangle\langle +|S\rangle + \langle +|-\rangle\langle -|S\rangle.$$

We have seen that we can put $\langle +|+\rangle = 1$, so we have

$$\langle +|-\rangle\langle -|S\rangle = 0$$

Has to be true for all states $|S\rangle$, i.e. for all values of $\langle -|S\rangle$. Thus we conclude that

$$\langle +|-\rangle = 0.$$

- Thus $|\langle +|-\rangle|^2 = 0$: if the spin state is $|-\rangle$ then there is **zero probability** of it being measured to be $|+\rangle$.
 - The two states represent mutually exclusive possibilities.

A Neat Correspondence

- In the same way we can show that $\langle -|- \rangle = 1$ and $\langle -|+ \rangle = 0$.
- Thus we can set up a comparison:

$$\langle +|+ \rangle = 1 \quad \longleftrightarrow \quad \hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 = 1$$

$$\langle +|- \rangle = 0 \quad \longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0$$

$$\langle -|- \rangle = 1 \quad \longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_2 = 1$$

$$\langle -|+ \rangle = 0 \quad \longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0$$

- This comparison implies the following correspondences:

$$|+ \rangle \longleftrightarrow \hat{\mathbf{u}}_1 \quad |- \rangle \longleftrightarrow \hat{\mathbf{u}}_2$$

$$\langle +| \longleftrightarrow \hat{\mathbf{u}}_1^* \quad \langle -| \longleftrightarrow \hat{\mathbf{u}}_2^*.$$

- Let $\langle +|S \rangle = a$ and $\langle -|S \rangle = b$, a and b are just complex numbers. Then

$$|S \rangle = |+\rangle \langle +|S \rangle + |-\rangle \langle -|S \rangle \longrightarrow |S \rangle = a|+\rangle + b|-\rangle$$

and we establish a perfect correspondence with the expression $\mathbf{v} = a \hat{\mathbf{u}}_1 + b \hat{\mathbf{u}}_2$.

- The state $|S \rangle$ is a vector!?

- Further note the correspondences:
 - $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ are orthonormal basis vectors.
 - Similarly we have $\langle +|+\rangle = \langle -|-\rangle = 1$ and $\langle +|-\rangle = \langle -|+\rangle = 0$.
- Suggests the interpretation that inner product \leftrightarrow probability amplitude (shown later generally to be the case).
- In particular, suggests that the states $|\pm\rangle$ form a pair of orthonormal basis vectors, or basis states, and hence:
 - The ket or state $|S\rangle$ is to be understood as being a **vector**.
 - We can construct new spin states $|S\rangle$ by forming a linear combination of the two orthonormal basis vectors, or basis states, $|\pm\rangle$.

- Now use the fact that $\langle S|S \rangle = 1$, so that

$$\langle S|S \rangle = 1 = \langle S|- \rangle \langle -|S \rangle + \langle S|+ \rangle \langle +|S \rangle$$

- On the other hand, the total probability of observing the system in either of the states $|\pm \rangle$ *must* add up to unity, which means that

$$|\langle +|S \rangle|^2 + |\langle -|S \rangle|^2 = 1.$$

- By comparing the last two equations, and noting that

$$|\langle \pm|S \rangle|^2 = \langle \pm|S \rangle \langle \pm|S \rangle^*$$

we conclude that

$$\langle \pm|S \rangle = \langle S|\pm \rangle^*.$$

- More generally, find that $\langle S|S' \rangle = \langle S'|S \rangle^*$ (sometimes referred to as ‘time reversal’).

- If we now consider $\langle S' | = \langle S' | + \rangle \langle + | + \langle S' | - \rangle \langle - |$

- Use $\langle \pm | S' \rangle = \langle S' | \pm \rangle^*$, to write this as

$$\langle S' | = \langle + | S' \rangle^* \langle + | + \langle - | S' \rangle^* \langle - |$$

- expressed in terms of $a' = \langle + | S' \rangle$ and $b' = \langle - | S' \rangle$, we have

$$\langle S' | = a'^* \langle + | + b'^* \langle - |$$

- a perfect correspondence with an ordinary vector $\hat{\mathbf{v}}'$ in the form

$$\mathbf{v}'^* = a'^* \hat{\mathbf{u}}_1 + b'^* \hat{\mathbf{u}}_2.$$

- The bra $\langle S' |$ is a vector, a bra vector, the complex conjugate of the ket vector $|S'\rangle$ (?)

- Useful to treat a bra vector as the complex conjugate of a ket vector, but . . .

- A bra vector is a different mathematical object to a ket vector (it belongs to a different vector space, for one thing).
 - Bra vectors are 'linear functionals', a mathematical operator that acts on a ket vector to produce a complex number.

- Finally, to complete the correspondence, we note that the probability amplitude $\langle S'|S\rangle$ can be written

$$\langle S'|S\rangle = \langle S'|+\rangle\langle +|S\rangle + \langle S'|- \rangle\langle -|S\rangle = a'^*a + b'^*b$$

Can be compared with the inner product $\mathbf{v}'^* \cdot \mathbf{v}$, or written in the more formal notation:

$$(\mathbf{v}', \mathbf{v}) = a'^*a + b'^*b$$

- This tells us that the probability amplitude can be considered as being simply the inner product of the two vectors $|S'\rangle$ and $|S\rangle$, i.e.

$$\langle S'|S\rangle = (|S'\rangle, |S\rangle).$$

- We have a perfect analogy between
 - the two dimensional complex inner product space formed by linear combinations of the unit vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ and
 - a complex inner product space consisting of all the linear combinations of the states $|\pm\rangle$. The ket vectors $|\pm\rangle$ are referred to as *basis states*, analogous to $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ being referred to as basis vectors.

- Different spin states can be formed as linear combinations $|S\rangle = a|+\rangle + b|-\rangle$ of these basis states.
- But what about a state such as $|\widetilde{S}\rangle = |+\rangle + i|-\rangle$?

- The inner product of $|\widetilde{S}\rangle$ with itself is:

$$\langle\widetilde{S}|\widetilde{S}\rangle = (\langle+| - i\langle-|)(|+\rangle + i|-\rangle) = 2$$

- But we have seen earlier that the probability interpretation of $\langle\widetilde{S}|\widetilde{S}\rangle$ tells us that we ought to have $\langle\widetilde{S}|\widetilde{S}\rangle = 1$ – the normalization condition.
- We do not abandon such states. We can define a new state by

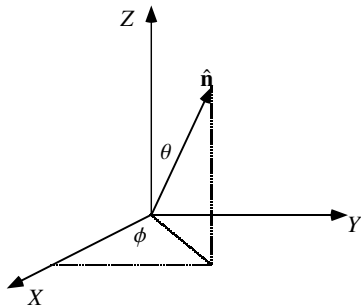
$$|S\rangle = \frac{|\widetilde{S}\rangle}{\sqrt{\langle\widetilde{S}|\widetilde{S}\rangle}} = \frac{1}{\sqrt{2}}|\widetilde{S}\rangle$$

so that $\langle S|S\rangle = 1$ — the state $|S\rangle$ has been ‘normalized’.

- We assume that any normalized state $|S\rangle$ or any of its multiples represent the **same physical state** of the system. But, we prefer to work with normalized states as the correct probabilities are already ‘built in’.

| | Classical vector | Quantum state vector for spin half system |
|--------------------|--|--|
| Basis Vectors | $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2$ | $ +\rangle, -\rangle$ |
| Inner Product | $(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^* \cdot \mathbf{v}_2$ | $(S'\rangle, S\rangle) = \langle S' S\rangle$ |
| Orthonormality | $\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_2 = 1$ $\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_2 = \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0$ | $\langle + +\rangle = \langle - -\rangle = 1$ $\langle + -\rangle = \langle - +\rangle = 0$ |
| Linear combination | $\hat{\mathbf{v}} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2$ a and b complex numbers | $ S\rangle = a +\rangle + b -\rangle$ a and b complex numbers |
| Normalisation | — | $\langle S S\rangle = 1$ |

- We now make the claim:
 - Any linear combination $a|+\rangle + b|-\rangle$ in which a and b are any complex numbers defines a possible spin state of the spin half system.
 - All physical states of the system can be written as a linear combination of $|\pm\rangle$.
- These are taken as postulates in the general formulation of quantum theory.
- All the possible linear combinations of $|\pm\rangle$, i.e. combinations of the form $a|+\rangle + b|-\rangle$ where a and b are complex numbers form a *complex vector space* known as the *state space* \mathcal{H} of the system.
- \mathcal{H} is a complex *inner product space*:
 - the inner product of $|S'\rangle$ and $|S\rangle$ is just the probability amplitude $\langle S'|S\rangle$.
 - In particular, the basis states are normalized to unity, i.e. they are unit vectors, and they are orthogonal to each other, i.e. they form a pair of *orthonormal basis states*.
 - The state space has dimension 2.



- By general arguments based on geometry and the properties of probability amplitudes, can derive an expression for an arbitrary spin state $|S\rangle$.
- Recall that $S = \mathbf{S} \cdot \hat{\mathbf{n}}$, and that $|S\rangle \equiv |\mathbf{S} \cdot \hat{\mathbf{n}} = \frac{1}{2}\hbar\rangle$.
- $\hat{\mathbf{n}}$ is a unit vector pointing in a direction defined by the polar angles θ, ϕ :

$$\hat{\mathbf{n}} = \sin \theta \cos \phi \hat{\mathbf{i}} + \sin \theta \sin \phi \hat{\mathbf{j}} + \cos \theta \hat{\mathbf{k}}$$

- The (normalized) state for which the atom the component of spin in the direction $\hat{\mathbf{n}}$ has the value $\frac{1}{2}\hbar$ is given by

$$|S\rangle = |\theta, \phi\rangle = \cos \frac{1}{2}\theta |+\rangle + e^{i\phi} \sin \frac{1}{2}\theta |-\rangle.$$

How do we generalize these ideas? I

- One way to approach this:
 - Develop a 'sum-of-amplitudes' approach by identifying the different possible 'pathways' that a system can follow from some given initial state to some final state.
 - Always working with probability amplitudes.
 - E.g. for an N slit interference experiment:

$$\langle x|S \rangle = \sum_{n=1}^N \langle x|n \rangle \langle n|S \rangle$$

- Leads to the Feynman path integral formulation of quantum mechanics: the most general way of formulating a quantum theory of any physical system.
- Or for atoms with spin s (not necessarily spin half):
 - Atom can emerge in any one of $2s + 1$ beams:

$$\langle S'|S \rangle = \sum_{n=1}^{2s+1} \langle S'|S_z = (n - s - 1)\hbar \rangle \langle S_z = (n - s - 1)\hbar |S \rangle$$

- Another way: work with state vectors right from the start.
 - That the state of a physical system acts like a vector is a direct consequence of the fundamental role of the probability amplitudes.
 - The 'sum-over-amplitudes' leads immediately to the vector formulation of quantum mechanics via the cancellation trick
 - In effect, we assume that the 'sum-over-amplitudes' is valid for a physical system of interest, and go straight to the vector description.
- The aim is to identify a set of basis states for the system.
- Can construct any state of the system as a linear combination of these states.

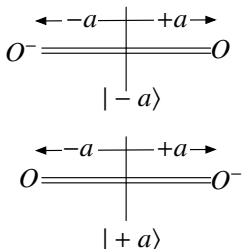
Properties of these basis states:

- Associated with a measurable property (an **observable**) of the system, e.g.
 - the **position** of an electron: at one slit or the other
 - a **component of spin** of an atom: a spin half atom emerging in one beam or the other of a Stern-Gerlach apparatus.
- They must each represent mutually exclusive possibilities, which means that different basis states are orthogonal e.g.

$$\langle +|- \rangle = \langle -|+ \rangle = 0.$$

- They cover all possibilities values for the observable.
 - The basis states are said to be *complete*.
- Each basis state is normalized to unity, e.g. $\langle +|+ \rangle = \langle -|- \rangle = 1$.
 - The basis states are said to form a **complete set of orthonormal basis states** for the system.

The O_2^- ion.



- Let $| - a \rangle$ be the state in which an electron is on the left hand oxygen atom
 - Let $| + a \rangle$ be the state in which the electron is on the right hand oxygen atom.
- Since $\langle -a | -a \rangle = 1$ then $\langle -a | -a \rangle = 1$.
Similarly $\langle +a | +a \rangle = 1$.
 - If the electron is definitely observed to be on the right hand oxygen atom, i.e. it is in the state $| + a \rangle$, then there is no chance of finding the electron on the left hand oxygen atom, i.e.

$$\langle -a | +a \rangle = 0.$$

And similarly $\langle +a | -a \rangle = 0$.

Examples II

- Thus the two states $|\pm a\rangle$ are orthonormal.
- Further these two states cover all the possibilities for the position of the electron – they are said to be *complete*.
- Shall use these states as a pair of basis states for the O_2^- ion.
- The state space is of dimension 2.

Example of state of the ion

- Consider the state $|\psi\rangle = 3i|+a\rangle + 4|-a\rangle$.
 - Is it normalized? Need to check $\langle\psi|\psi\rangle$.

$$\langle\psi| = -3i\langle+a| + 4\langle-a|$$

So

$$\langle\psi|\psi\rangle = (-3i\langle+a| + 4\langle-a|)(3i|+a\rangle + 4|-a\rangle) = 25$$

Therefore not normalized.

- In order to obtain correct probabilities we must normalize the state.

Examples III

- Introduce normalized state

$$|\tilde{\psi}\rangle = \frac{1}{5}|\psi\rangle = \frac{3i}{5}|+a\rangle + \frac{4}{5}|-a\rangle.$$

- Find that

$$\langle\tilde{\psi}|\tilde{\psi}\rangle = 1.$$

- Can now calculate the probability of finding the electron on the left hand oxygen atom:

$$\langle -a|\tilde{\psi}\rangle = \langle -a|(3i|+a\rangle + 4|-a\rangle)/5 = \frac{4}{5}.$$

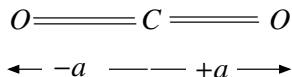
The required probability is

$$|\langle -a|\tilde{\psi}\rangle|^2 = \frac{16}{25} = 0.64$$

- And similarly for the right hand oxygen atom:

$$|\langle +a|\tilde{\psi}\rangle|^2 = \left|\frac{3i}{5}\right|^2 = \frac{9}{25} = 0.36.$$

Carbon Dioxide ion



- In this case the electron can reside on either oxygen atom or on the carbon atom in the middle.
- The corresponding states, taking the position of the electron as the observable will be

$$|+a\rangle \quad |0\rangle \quad |-a\rangle$$

corresponding to the electron being on the right hand oxygen atom, the middle carbon atom or the left hand oxygen atom.

- Here $\langle +a|0\rangle = \langle +a|-a\rangle = 0$ and so on, and $\langle 0|0\rangle = \langle -a|-a\rangle = \langle +a|+a\rangle = 1$.
- The state space is spanned by the three orthonormal basis states $|+a\rangle, |0\rangle, |-a\rangle$ and so has a dimension of 3.
- As there is no where else that the electron can be found (within the limits of the model), these states are complete.
- An arbitrary state of the CO_2^- will then be

$$|\psi\rangle = a|+a\rangle + b|0\rangle + c|-a\rangle$$

An infinite dimensional state space

- Infinite dimensional state spaces are the rule rather than the exception in quantum mechanics.
 - In principle, a particle free to move along, say, the x axis, has a continuously infinite range of possible positions.
 - This leads to mathematical difficulties that we will not concern ourselves with here.
 - But 'countable' infinities also arise.
- Consider a metal cavity designed to support an EM field of a single frequency and polarization.
 - This field is 'quantized' — it is made up of photons.
 - There could be 0, 1, 2, 3, ... photons in the cavity. with the corresponding states of the cavity field being
$$|0\rangle, |1\rangle, |2\rangle, \dots$$
 - These states are orthonormal $\langle m|n\rangle = \delta_{mn}$ and form a complete set of basis states

$$|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle \quad \text{e.g. a laser field} \quad |\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle.$$

- Recall for spin half:

$$\langle S'|S\rangle = \langle S'|+\rangle\langle +|S\rangle + \langle S'|- \rangle\langle -|S\rangle$$

- Expresses the fact that the total probability amplitude of finding the system in the state $|S'\rangle$ given that it was in the state $|S\rangle$ is the sum of the probability amplitudes of the system 'passing through' the 'intermediate' states $|\pm\rangle$.
- This idea can be generalised, based on our examples so far.
- Suppose we have a system and we have chosen an observable quantity (e.g. the position of an electron in an ion, the number of photons in a cavity, the possible energies of an atom, . . .) and identified a set of states $\{|1\rangle, |2\rangle, |3\rangle, \dots |N\rangle\}$ that are
 - Orthonormal: $\langle n|n\rangle = 1$ $\langle n|m\rangle = 0, n \neq m$
 - Complete: there are no other states possible for the chosen observable
- Then these states form a complete orthonormal set of basis states for the state space \mathcal{H} of the system. The state space has dimension N

The general case II

- Any state of the system can be written

$$|\psi\rangle = c_1|1\rangle + c_2|2\rangle + \dots + c_N|N\rangle.$$

- Now we use the orthonormality of the basis states:

$$\langle n|\psi\rangle = c_1\langle n|1\rangle + c_2\langle n|2\rangle + \dots + c_N\langle n|N\rangle.$$

- All of the quantities $\langle n|m\rangle$ will vanish except for $m = n$, so we get:

$$\langle n|\psi\rangle = c_n$$

and hence

$$|\psi\rangle = |1\rangle\langle 1|\psi\rangle + |2\rangle\langle 2|\psi\rangle + \dots + |N\rangle\langle N|\psi\rangle.$$

- Finally, take the inner product with some other state $|\phi\rangle$ to give

$$\langle\phi|\psi\rangle = \langle\phi|1\rangle\langle 1|\psi\rangle + \langle\phi|2\rangle\langle 2|\psi\rangle + \dots + \langle\phi|N\rangle\langle N|\psi\rangle.$$

- This result is of the same form as we saw for spin half
 - It suggests that the system 'passes through' the 'intermediate states' $\{|1\rangle, |2\rangle, \dots, |N\rangle\}$ as it makes its way from the initial state $|\psi\rangle$ to the final state $|\phi\rangle$.
 - It is known as the 'closure relation'.

- We have been solely concerned with states with a single label, e.g.

- $|S_z = \frac{1}{2}\hbar\rangle$ for a spin half particle for which $S_z = \frac{1}{2}\hbar$
- $|x\rangle$ for a particle positioned at the point x
- $|n\rangle$ for a single mode cavity with n photons.

- But recall the general definition of a ket:

| All the data concerning the system that can be known without mutual
interference or contradiction. |

so, for more complex systems, we can have knowledge of more than one (compatible) observable to specify the state, e.g.

- $|S_x = \frac{1}{2}\hbar, x\rangle$ for a spin half particle with x component of spin $S_x = \frac{1}{2}\hbar$ AND at the position x .
- Why 'compatible'? Because some observables interfere with others, e.g. momentum and position or different components of spin.

States with multiple labels II

- We could also have a system made up of more than one particle e.g. a system of two spin half particles.
 - With the state written as $|\text{data for first particle}, \text{data for second particle}\rangle$
 - Its possible states would then be:
$$|S_z = \frac{1}{2}\hbar, S_z = \frac{1}{2}\hbar\rangle, |S_z = \frac{1}{2}\hbar, S_z = -\frac{1}{2}\hbar\rangle, |S_z = -\frac{1}{2}\hbar, S_z = \frac{1}{2}\hbar\rangle, |S_z = -\frac{1}{2}\hbar, S_z = -\frac{1}{2}\hbar\rangle.$$
 - Can write this more simply as
 - $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ where $S_z = \frac{1}{2}\hbar \rightarrow 0$ $S_z = -\frac{1}{2}\hbar \rightarrow 1$.
 - The states are orthonormal: $\langle 00|00\rangle = 1$ $\langle 01|01\rangle = 1$ etc
 $\langle 00|01\rangle = 0$ $\langle 01|10\rangle = 0$ etc
 - An arbitrary state is $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$.
- These states form a set of orthonormal basis states for a state space of dimension 4.
- The idea can be extended to many particles, or other more complex systems e.g. whole atoms, or molecules or

- Note that a single spin half has the basis states $|0\rangle$ and $|1\rangle$ in our new notation.
 - An arbitrary state will be $|\psi\rangle = a|0\rangle + b|1\rangle$.
 - Compare with the two alternative states of a 'bit' in a computer memory: 0 OR 1.
 - The spin half particle is an example of a 'qubit': it can be in the states $|0\rangle$ and $|1\rangle$ simultaneously.
 - Any two state system qualifies as a qubit.
- For our two qubit system, the data 00, 01, 10, 11 appearing in the states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ are binary for the integers 0, 1, 2, 3.
 - The state $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$ can be considered a linear combination of the integers 0, 1, 2, 3
- This is exploited in quantum computers where computations carried out on one state vector $|\psi\rangle$ can be viewed as parallel processing four different values simultaneously.
- Can be made massively parallel by using N qubits which have 2^N basis states.
- The viability of this scheme relies on the linear combination not being destroyed by decoherence.

- Why can't we use these ideas for macroscopic systems/objects?

e.g. $|\text{car}\rangle = \frac{1}{\sqrt{2}} (|\text{mercedes}\rangle + |\text{bmw}\rangle)$

- The car is a superposition of two different possible cars.
- In some sense it is BOTH a mercedes AND a bmw at the same time.
- There is interference. Suppose the car could be red or black. The probability of it being red is

$$\begin{aligned} |\langle \text{red} | \text{car} \rangle|^2 &= \frac{1}{2} |\langle \text{red} | \text{mercedes} \rangle + \langle \text{red} | \text{bmw} \rangle|^2 \\ &= \frac{1}{2} (|\langle \text{red} | \text{mercedes} \rangle|^2 + |\langle \text{red} | \text{bmw} \rangle|^2 + 2\text{Re}[\langle \text{red} | \text{mercedes} \rangle^* \langle \text{red} | \text{bmw} \rangle]) \end{aligned}$$

- The last term is the quantum interference term.
- Contact of a system with its surrounding environment leads to *decoherence*

The macroscopic limit II

- Decoherence wipes out a linear superposition so that

$$|\text{car}\rangle \rightarrow \text{EITHER } |\text{mercedes}\rangle \text{ OR } |\text{bmw}\rangle$$

with a 50% chance of being either.

- We can see the effect of decoherence if we assume some values for the probability amplitudes:

$$\langle \text{red} | \text{mercedes} \rangle = \frac{1}{\sqrt{2}} \quad \langle \text{red} | \text{bmw} \rangle = \frac{1}{\sqrt{2}}$$

$$\langle \text{black} | \text{mercedes} \rangle = \frac{1}{\sqrt{2}} \quad \langle \text{black} | \text{bmw} \rangle = -\frac{1}{\sqrt{2}}$$

and the corresponding probabilities:

$$|\langle \text{red} | \text{mercedes} \rangle|^2 = \frac{1}{2} \quad |\langle \text{red} | \text{bmw} \rangle|^2 = \frac{1}{2}$$

$$|\langle \text{black} | \text{mercedes} \rangle|^2 = \frac{1}{2} \quad |\langle \text{black} | \text{bmw} \rangle|^2 = \frac{1}{2}$$

- If you determined only that the car was a mercedes then there is an equal chance of it being either red or black.
 - Similarly if you determined that the car was a bmw.

The macroscopic limit III

- Find that, for the superposition state $|\text{car}\rangle = \frac{1}{\sqrt{2}} (|\text{mercedes}\rangle + |\text{bmw}\rangle)$

$$|\langle \text{red} | \text{car} \rangle|^2 = \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2} + 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \right) = 1$$

$$|\langle \text{black} | \text{car} \rangle|^2 = \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2} + 2 \cdot \frac{1}{\sqrt{2}} \cdot \left(-\frac{1}{\sqrt{2}}\right) \right) = 0$$

i.e. because of the interference term:

- $|\langle \text{red} | \text{car} \rangle|^2 = 1$, but $|\langle \text{black} | \text{car} \rangle|^2 = 0$, so the car is guaranteed to be red.
- But, after decoherence, the car is either in the state $|\text{mercedes}\rangle$ or $|\text{bmw}\rangle$, and in either case
 - from the probabilities table above there is a 50:50 chance of the car being red or black, irrespective of what kind of car it turns out to be.
- Quantum interference (i.e. superposition of states) \implies car is *always* red.
- Decoherence (i.e. no superposition) \implies car has equal chance of being red or black.
- This example is the spin half interference experiment in disguise.

- We have seen how to construct the states of a quantum system.
- Now we need to see how to:
 - Describe how the state provides us information about the observable physical properties of the system it is describing.
 - How states change when something is 'done' to the system
 - Displace it in space.
 - Rotate it.
 - Allow it to evolve in time
- All these properties of a quantum system are described by making use of 'operators'
 - Operators are mathematical objects that act on state vectors to change them into other state vectors.

- Operators are always represented by a symbol with a 'hat' on top, e.g. \hat{A} , \hat{H} , \hat{S}_z , \hat{x} , \hat{p} , \hat{U} .
- An operator \hat{A} is a mathematical entity that acts on an arbitrary state $|\psi\rangle$ and maps it into some other state $|\phi\rangle$.
 - This is written $\hat{A}|\psi\rangle = |\phi\rangle$ **NEVER** $|\psi\rangle\hat{A}$.
- An operator is fully specified when its action on *every* state of a quantum system is known.
 - This is a major task in general as there are an infinite number of possible states.
- But operators in quantum mechanics have a very important property: they are *linear*.

- An operator is linear if $\hat{A}(a|\psi_1\rangle + b|\psi_2\rangle) = a(\hat{A}|\psi_1\rangle) + b(\hat{A}|\psi_2\rangle)$

where a and b are complex numbers and $|\psi_1\rangle$ and $|\psi_2\rangle$ are arbitrary states.

- There also occurs an 'antilinear operator' in quantum mechanics — the time reversal operator \hat{T} :

$$\hat{T}(a|\psi_1\rangle + b|\psi_2\rangle) = a^*(\hat{T}|\psi_1\rangle) + b^*(\hat{T}|\psi_2\rangle) \quad \text{— we won't be studying this.}$$

- Since any state can be written as a linear combination of basis states, we only need to know the action of a linear operator on a set of basis vectors.

- Suppose the basis states are $\{|1\rangle, |2\rangle, \dots\}$. Then any state can be written

$$|\psi\rangle = c_1|1\rangle + c_2|2\rangle + \dots$$

- The action of \hat{A} on $|\psi\rangle$ is then

$$\hat{A}|\psi\rangle = c_1(\hat{A}|1\rangle) + c_2(\hat{A}|2\rangle) + \dots$$

- So we only need to know $\hat{A}|1\rangle, \hat{A}|2\rangle, \dots$ to work out $\hat{A}|\psi\rangle$.

Example of a linear operator

- For example, consider a spin half system with basis states $|\pm\rangle$ (i.e. the old notation).
- Suppose an operator \hat{A} is defined by

$$\hat{A}|+\rangle = \frac{1}{2}i\hbar|-\rangle \quad \hat{A}|-\rangle = -\frac{1}{2}i\hbar|+\rangle$$

- Then for an arbitrary state $|\psi\rangle = a|+\rangle + b|-\rangle$ we have

$$\begin{aligned}\hat{A}|\psi\rangle &= \hat{A}(a|+\rangle + b|-\rangle) \\ &= a(\hat{A}|+\rangle) + b(\hat{A}|-\rangle) \\ &= a(\frac{1}{2}i\hbar|-\rangle) - b(\frac{1}{2}i\hbar|+\rangle) \\ &= \frac{1}{2}i\hbar(a|-\rangle - b|+\rangle)\end{aligned}$$

- So if $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$: $\hat{A}|\psi\rangle = \frac{1}{2}i\hbar\frac{1}{\sqrt{2}}(|+\rangle - |-\rangle) = |\phi\rangle$ so $|\phi\rangle \neq |\psi\rangle$.

So \hat{A} changes the state.

- But if $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle)$: $\hat{A}|\psi\rangle = \frac{1}{2}i\hbar\frac{1}{\sqrt{2}}(|-\rangle - i|+\rangle) = \frac{1}{2}\hbar\frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle) = \frac{1}{2}\hbar|\psi\rangle$.

So the new state $|\phi\rangle$ is simply a multiple of the original — the *physical* state of the system has not changed.

- Since $\hat{A}|\psi\rangle$ simply produces a new state $|\phi\rangle$, we ought to be able to take the inner product of $|\phi\rangle = \hat{A}|\psi\rangle$ with some other state $|\chi\rangle$:

$$\langle\chi|\phi\rangle = \langle\chi|\hat{A}|\psi\rangle.$$

- This notation tells us that we first let \hat{A} act on $|\psi\rangle$ then take the inner product with $|\chi\rangle$.
 - But could we first let \hat{A} act on the bra vector $\langle\chi|$ to produce a new bra vector and then take the inner product with $|\psi\rangle$, i.e. do it this way:

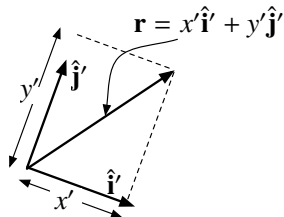
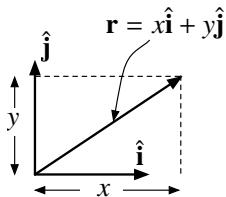
$$(\langle\chi|\hat{A})|\psi\rangle?$$

- It turns out that you can. More than that, you get the same answer, i.e.

$$(\langle\chi|\hat{A})|\psi\rangle = \langle\chi|\hat{A}|\psi\rangle$$

- So we drop the brackets and simply write $\langle\chi|\hat{A}|\psi\rangle$.
- To show this, we need to look at another way of writing vectors and operators — column vectors, row vectors and matrices instead of kets, bras and operators.

Representations of ordinary vectors I



- Consider an ordinary space displacement vector \mathbf{r} .

- In terms of the usual unit vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ we write

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$$

- Provided we remain aware of the choice of unit vectors $\hat{\mathbf{i}}, \hat{\mathbf{j}}$, we can focus on the components only and write

$$\mathbf{r} \doteq \begin{pmatrix} x \\ y \end{pmatrix}_R$$

i.e. we *represent* the vector \mathbf{r} by a column vector consisting of its components wrt to $\hat{\mathbf{i}}, \hat{\mathbf{j}}$.

- But the *same* vector \mathbf{r} can be written in terms of any other set of unit vectors $\hat{\mathbf{i}}', \hat{\mathbf{j}}'$ as $\mathbf{r} = x'\hat{\mathbf{i}}' + y'\hat{\mathbf{j}}'$

- Or, as a column vector:

$$\mathbf{r} \doteq \begin{pmatrix} x' \\ y' \end{pmatrix}_{R'}$$

- Same vector, different components.

Representations of ordinary vectors II

- Depending on the choice of axes, the same vector can be represented by different column vectors.
 - Just looking at a column vector won't tell you what the axes are.
 - But provided we have agreed on a choice of axes (or unit vectors), we can use the column vector unambiguously to represent the vector.
 - Note the use of the word 'represent', and the use of the symbol \doteq . The column vector merely represents the vector via its components, and is not the vector itself.

- Check out an ordinary scalar or inner product: $\mathbf{r}_1 \cdot \mathbf{r}_2 = x_1x_2 + y_1y_2$

- We have assumed the use of the $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ unit vectors.
- By the usual rules of matrix multiplication, we can write this inner product using as

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = x_1x_2 + y_1y_2 = \begin{pmatrix} x_1 & y_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

- i.e. have represented the first factor by a *row* vector.
- We now apply the same ideas to state vectors.

Representations of state vectors I

- Suppose the state space of a quantum system is spanned by a complete set of orthonormal basis states $\{|1\rangle, |2\rangle, \dots\}$.

- Any state vector can be written

$$\begin{aligned} |\psi\rangle &= |1\rangle\langle 1|\psi\rangle + |2\rangle\langle 2|\psi\rangle + \dots \\ &= \psi_1|1\rangle + \psi_2|2\rangle + \dots \end{aligned}$$

i.e. $\psi_1 = \langle 1|\psi\rangle, \psi_2 = \langle 2|\psi\rangle, \dots$

- With respect to this choice of basis states, we can represent this state vector by the column vector

$$|\psi\rangle \doteq \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}$$

- E.g. for $|\psi\rangle = \frac{1}{\sqrt{5}} (|1\rangle + 2i|2\rangle)$

$$|\psi\rangle \doteq \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2i}{\sqrt{5}} \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2i \end{pmatrix}$$

- In particular, for the basis states:

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{so that} \quad |\psi\rangle \doteq \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{2i}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Representations of state vectors II

- Now try the inner product $\langle\phi|\psi\rangle$

- First we have

$$\begin{aligned}|\phi\rangle &= |1\rangle\langle 1|\phi\rangle + |2\rangle\langle 2|\phi\rangle + \dots \\ &= \phi_1|1\rangle + \phi_2|2\rangle + \dots\end{aligned}$$

so that

$$\begin{aligned}\langle\phi| &= \langle\phi|1\rangle\langle 1| + \langle\phi|2\rangle\langle 2| + \dots \\ &= \phi_1^*\langle 1| + \phi_2^*\langle 2| + \dots\end{aligned}$$

- Next:

$$\begin{aligned}\langle\phi|\psi\rangle &= \langle\phi|1\rangle\langle 1|\psi\rangle + \langle\phi|2\rangle\langle 2|\psi\rangle + \dots \\ &= \phi_1^*\psi_1 + \phi_2^*\psi_2 + \dots\end{aligned}$$

or, in matrix notation

$$= \begin{pmatrix} \phi_1^* & \phi_2^* & \dots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}$$

- In other words, a bra vector is represented by a row vector:

$$\langle\phi| \doteq \begin{pmatrix} \phi_1^* & \phi_2^* & \dots \end{pmatrix}$$

Representations of state vectors III

- E.g. for $|\phi\rangle = \frac{1}{5} (3i|1\rangle + 4|2\rangle)$:

$$|\phi\rangle \doteq \frac{1}{5} \begin{pmatrix} 3i \\ 4 \end{pmatrix} \implies \langle\phi| \doteq \frac{1}{5} (-3i \quad 4)$$

- With $|\psi\rangle = \frac{1}{\sqrt{5}} (|1\rangle + 2i|2\rangle)$ we have

$$\langle\phi|\psi\rangle = \frac{1}{5\sqrt{5}} (-3i \quad 4) \begin{pmatrix} 1 \\ 2i \end{pmatrix} = \frac{i}{\sqrt{5}} \implies |\langle\phi|\psi\rangle|^2 = 0.2$$

- Note that a bra vector looks different from a ket vector
 - a ket is represented by a row vector, a bra by a column vector
- This emphasizes that a bra vector is a different mathematical object from a ket vector.
 - Ket vectors belong to the state space \mathcal{H} , bra vectors belong to the *dual* space \mathcal{H}^* .
 - A distinction that does not matter for finite dimensional state spaces, and won't matter for us.

Representation of operators as matrices I

- Now consider the operator equation $\hat{A}|\psi\rangle = |\phi\rangle$.

- Expanding the ket vectors (we'll assume only two basis states here):

$$\begin{aligned}\hat{A}(\psi_1|1\rangle + \psi_2|2\rangle) &= \phi_1|1\rangle + \phi_2|2\rangle \\ \Rightarrow \hat{A}|1\rangle\psi_1 + \hat{A}|2\rangle\psi_2 &= \phi_1|1\rangle + \phi_2|2\rangle.\end{aligned}$$

- Now take the inner product of both sides with $|1\rangle$:

$$\langle 1|\hat{A}|1\rangle\psi_1 + \langle 1|\hat{A}|2\rangle\psi_2 = \phi_1$$

- We should have written, e.g. $\langle 1|\hat{A}|2\rangle$, but we are assuming that we do not need the brackets.
- Next take the inner product with $|2\rangle$:

$$\langle 2|\hat{A}|1\rangle\psi_1 + \langle 2|\hat{A}|2\rangle\psi_2 = \phi_2$$

- Now put $A_{ij} = \langle i|\hat{A}|j\rangle$ to give

$$A_{11}\psi_1 + A_{12}\psi_2 = \phi_1$$

$$A_{21}\psi_1 + A_{22}\psi_2 = \phi_2$$

Representation of operators as matrices II

- This can be written as a matrix equation:

$$\begin{array}{ccc} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} & \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} & = & \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \\ \uparrow & \uparrow & \uparrow & \\ \hat{A} & |\psi\rangle & & |\phi\rangle \end{array}$$

- In other words, \hat{A} is represented by a matrix:

$$\hat{A} \doteq \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

- If there are more than two basis states i.e. $\{|1\rangle, |2\rangle, \dots, |N\rangle\}$:

$$\hat{A} \doteq \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{pmatrix} \quad \text{an } N \times N \text{ matrix.}$$

- So, an example:

- The operator \hat{A} for a spin half system with basis states $|\pm\rangle$, defined by

$$\hat{A}|+\rangle = \frac{1}{2}i\hbar|-\rangle \quad \hat{A}|-\rangle = -\frac{1}{2}i\hbar|+\rangle.$$

What is its matrix representation? Let's set $|+\rangle \equiv |1\rangle$ and $|-\rangle \equiv |2\rangle$

$$\hat{A} \doteq \begin{pmatrix} A_{++} & A_{+-} \\ A_{-+} & A_{--} \end{pmatrix}.$$

- We need, for example:

$$A_{++} = \langle +|\hat{A}|+\rangle = \langle +|(\hat{A}|1+\rangle) = \langle +|(\frac{1}{2}i\hbar|-\rangle) = \frac{1}{2}i\hbar\langle +|-\rangle = 0$$

$$A_{+-} = \langle +|\hat{A}|-\rangle = \langle +|(\hat{A}|-\rangle) = \langle +|(-\frac{1}{2}i\hbar|+\rangle) = -\frac{1}{2}i\hbar\langle +|+\rangle = -\frac{1}{2}i\hbar.$$

- Overall, we find that

$$\hat{A} \doteq \begin{pmatrix} 0 & -\frac{1}{2}i\hbar \\ \frac{1}{2}i\hbar & 0 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

- And hence $\hat{A}|\psi\rangle = |\phi\rangle$, with, e.g. $|\psi\rangle = \frac{1}{5}(3i|+\rangle + 4|-\rangle)$ becomes

$$\hat{A}|\psi\rangle \doteq \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 3i \\ 4 \end{pmatrix} = \frac{\hbar}{10} \begin{pmatrix} -4i \\ -3 \end{pmatrix} \implies |\phi\rangle = -\frac{\hbar}{10} (4i|+\rangle + 3|-\rangle).$$

Representation of operators as matrices IV

Action of an operator on a bra vector

- Using the matrix representation, can define an operator acting on a bra vector, i.e. $\langle\psi|\hat{B}$:

$$\langle\phi|\hat{B} \doteq (\phi_1^* \quad \phi_2^*) \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

- Multiplying this through clearly yields another row vector, i.e. a bra vector.
- For example, consider, for a spin half system, the operator \hat{B} defined by

$$\hat{B}|+\rangle = 2|+\rangle + 3i|-\rangle \quad \hat{B}|-\rangle = (1+i)|+\rangle - |-\rangle$$

and let it act on the ket vector $|\psi\rangle = \frac{1}{5}(3i|+\rangle + 4|-\rangle)$.

$$\hat{B}|\psi\rangle = \frac{1}{5} \begin{pmatrix} 2 & 1+i \\ 3i & -1 \end{pmatrix} \begin{pmatrix} 3i \\ 4 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 4+10i \\ -13 \end{pmatrix} \doteq \frac{1}{5} [(4+10i)|+\rangle - 13|-\rangle] = |\chi\rangle.$$

- Now try $\langle\psi|\hat{B}$:

$$\langle\psi|\hat{B} \doteq \frac{1}{5} \begin{pmatrix} -3i & 4 \end{pmatrix} \begin{pmatrix} 2 & 1+i \\ 3i & -1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 6i & -1-3i \end{pmatrix} \doteq \frac{1}{5} [6i\langle+| - (1+3i)\langle-|] \neq \langle\chi|$$

- So the result is, indeed, a bra vector, but note that
 - If $\hat{B}|\psi\rangle = |\chi\rangle$ then, in general, $\langle\psi|\hat{B} \neq \langle\chi|$.

Sums and products of operators

- The sum of two operators $\hat{A} + \hat{B}$ is found by simply adding their matrix representations.
- The product $\hat{A}\hat{B}$ of two operators is found by simply multiplying their matrix representations.

- It is important to note that, in general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$
- The difference $\hat{A}\hat{B} - \hat{B}\hat{A}$ is known as the commutator of \hat{A} and \hat{B} and is written

$$\hat{A}\hat{B} - \hat{B}\hat{A} = [\hat{A}, \hat{B}].$$

- If $[\hat{A}, \hat{B}] = 0$ the operators are said to commute.
 - The commutator of pairs of operators plays a central role in quantum mechanics.
- Some special operators:
 - The operator \hat{I} such that $\hat{I}\hat{A} = \hat{A}\hat{I}$ is called the unit operator. Its matrix has ones along the diagonal and zeroes everywhere else.
 - If $\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{I}$ then \hat{A} is said to be the inverse of \hat{B} , written $\hat{A} = \hat{B}^{-1}$.
 - Calculate the matrix \hat{A}^{-1} by the usual methods of matrix inversion (need $\det\hat{A} \neq 0$ so not all operators have an inverse.)

- We can use the matrix representation of an operator to define the important concept of a Hermitean adjoint.

- Start with an operator \hat{A} with the matrix representation

$$\hat{A} \doteq \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

- Construct a new matrix by taking the complex conjugate of all the elements, then transposing the matrix. The new matrix is:

$$\begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix}.$$

- This matrix represents a new operator, call it \hat{A}^\dagger , i.e.

$$\hat{A}^\dagger \doteq \begin{pmatrix} \langle 1|\hat{A}^\dagger|1\rangle & \langle 1|\hat{A}^\dagger|2\rangle \\ \langle 2|\hat{A}^\dagger|1\rangle & \langle 2|\hat{A}^\dagger|2\rangle \end{pmatrix} = \begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix}. \quad \text{i.e.} \quad \langle i|\hat{A}^\dagger|j\rangle = \langle j|\hat{A}|i\rangle^*$$

- The operator \hat{A}^\dagger is known as the Hermitean adjoint of \hat{A} .

- Example:

$$\hat{A} \doteq \begin{pmatrix} 1+i & i \\ 0 & -1 \end{pmatrix} \implies \hat{A}^\dagger \doteq \begin{pmatrix} 1-i & 0 \\ -i & -1 \end{pmatrix}$$

- This shows that, in general, $\hat{A} \neq \hat{A}^\dagger$.
- We can also prove (but we won't) that
 - Since $\langle i|\hat{A}^\dagger|j\rangle = \langle j|\hat{A}|i\rangle^*$ then for any states $|\phi\rangle$ and $|\psi\rangle$ we have $\langle\phi|\hat{A}|\psi\rangle^* = \langle\psi|\hat{A}^\dagger|\phi\rangle$.
 - So taking the complex conjugate is equivalent to reversing the order of all the factors, as well as taking the Hermitean conjugate of the operator.
 - It is sometimes helpful to think of \hat{A}^\dagger as the 'complex conjugate' of \hat{A} .
- This result also tells us (though we also won't prove it here) that if $\hat{A}|\psi\rangle = |\phi\rangle$ then $\langle\psi|\hat{A}^\dagger = \langle\phi|$.

Hermitean and unitary operators

- If $\hat{A} = \hat{A}^\dagger$ the operator \hat{A} is said to be Hermitean (or self-adjoint).

- E.g. if $\hat{A}|\pm\rangle = \pm\frac{1}{2}i\hbar|\mp\rangle$ then

$$\hat{A} \doteq \begin{pmatrix} 0 & -\frac{1}{2}i\hbar \\ \frac{1}{2}i\hbar & 0 \end{pmatrix} \doteq \hat{A}^\dagger$$

- Hermitean operators play the role of representing the physically observable properties of a quantum system, such as position, momentum, energy,

- If $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$ then $\hat{U}^{-1} = \hat{U}$ and the operator is said to be unitary.

- E.g., if $\hat{U}|\pm\rangle = i|\pm\rangle$ then

$$\hat{U} \doteq \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \hat{U}^\dagger \doteq \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \implies \hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \hat{I}.$$

- Unitary operators leave a normalised state normalised.

- If $\hat{U}|\psi\rangle = |\phi\rangle$ then $\langle\psi|\hat{U}^\dagger = \langle\phi| \implies \langle\phi|\phi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{I}|\psi\rangle = \langle\psi|\psi\rangle$

- So, if $\langle\psi|\psi\rangle = 1$, then $\langle\phi|\phi\rangle = 1$.

- Unitary operators represent 'doing something to a system'.

An example of a unitary operator I

The S matrix

- Consider a 'two level atom' that has two basis states corresponding to different energies, i.e.

$|g\rangle$ the ground state and $|e\rangle$ the excited state.

- Suppose that this atom suffers a collision with a passing electron. The results of the collision are as follows:
 - If the atom is in its ground state before the collision, it ends up in the state $(|e\rangle + |g\rangle)/\sqrt{2}$ after the collision.
 - If it is in its excited state before the collision, it ends up in the state $(|e\rangle - |g\rangle)/\sqrt{2}$ after the collision
- The effects of the collision can be represented by an operator \hat{S} with the properties

$$\hat{S}|g\rangle = \frac{1}{\sqrt{2}}(|e\rangle + |g\rangle) \quad \hat{S}|e\rangle = \frac{1}{\sqrt{2}}(|e\rangle - |g\rangle)$$

An example of a unitary operator II

The S matrix

- The matrix representation of \hat{S} is:

$$\hat{S} \doteq \begin{pmatrix} \langle e|\hat{S}|e\rangle & \langle e|\hat{S}|g\rangle \\ \langle g|\hat{S}|e\rangle & \langle g|\hat{S}|g\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

sometimes called the S matrix.

- The Hermitean conjugate \hat{S}^\dagger is

$$\hat{S}^\dagger \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

so that

$$\hat{S}^\dagger \hat{S} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Thus, \hat{S} is unitary: it describes the change in the state of the atom due to a collision

An example of a unitary operator III

The S matrix

- E.g. if the initial state is $|\psi\rangle = \frac{1}{5}(3|e\rangle - 4i|g\rangle)$ then the post-collision state is

$$|\phi\rangle = \hat{S}|\psi\rangle \doteq \frac{1}{5\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -4i \end{pmatrix} = \frac{1}{5\sqrt{2}} \begin{pmatrix} 3 - 4i \\ -3 - 4i \end{pmatrix}$$

i.e. $|\phi\rangle = \frac{1}{5\sqrt{2}} [(3 - 4i)|e\rangle - (3 + 4i)|g\rangle]$.

- Before the collision, the probability of finding the atom in its excited state was $|\langle e|\psi\rangle|^2 = 0.36$.
- After the collision, the probability is now $|\langle e|\phi\rangle|^2 = \frac{1}{50}|3 - 4i|^2 = 0.5$.
- This is, of course, a simplistic model, but it captures the essence of how collisional processes are described in practice.
 - The S matrix is a very much studied quantity, particularly as a function of the energy of the incoming particle.

- For some operators \hat{A} it may be the case that there is a state $|\phi\rangle$ say for which

$$\hat{A}|\phi\rangle = a_\phi|\phi\rangle$$

where a_ϕ is, in general, a complex number.

- Such a state is known as an eigenstate (or eigenket) of the operator \hat{A} , and a_ϕ is known as the associated eigenvalue.
- In general it is found that an operator can have
 - No eigenstates
 - One or more eigenstates
 - Real or complex eigenvalues.
- Of particular importance in quantum mechanics is the case of Hermitean operators.

- If \hat{A} is Hermitean then
 - the eigenvalues are all real (easy to prove)
 - Eigenvectors belonging to different eigenvalues are orthogonal i.e. if $\hat{A}|\psi\rangle = a_\psi|\psi\rangle$ and $\hat{A}|\phi\rangle = a_\phi|\phi\rangle$ then $\langle\phi|\psi\rangle = 0$ if $a_\phi \neq a_\psi$ (also easy to prove).
 - The eigenstates can be normalized to unity, i.e. so that $\langle\psi|\psi\rangle = 1$ (this is easy too).
 - The set of eigenstates of a Hermitean operator forms a complete orthonormal set of basis states for the state space of the system (not easy at all, and sometimes not true for infinite dimensional state spaces.)
- This last property means Hermitean operators are a convenient source of basis states for describing a quantum system.

Eigenstates and eigenvectors III

- Consider, as an example, the spin half operator \hat{A} defined such that $\hat{A}|\pm\rangle = \pm\frac{1}{2}i\hbar|\mp\rangle$. Its matrix wrt to the spin half basis states $|\pm\rangle$ is

$$\hat{A} \doteq \frac{1}{2}\hbar \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

- The usual methods shows that the eigenvalues of \hat{A} are $\pm\frac{1}{2}\hbar$ — i.e. both eigenvalues are real.
- The associated eigenvectors, or eigenstates, are then

$$|1\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \hat{A}|1\rangle = \frac{1}{2}\hbar|1\rangle$$
$$|2\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \hat{A}|2\rangle = -\frac{1}{2}\hbar|2\rangle.$$

- The eigenstates $|1\rangle, 2$ are orthonormal:

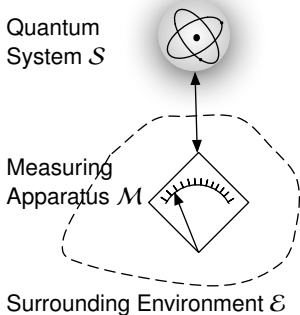
$$\langle 1|1\rangle = \frac{1}{2} \begin{pmatrix} 1 & i \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = 1 \quad \langle 1|2\rangle = \frac{1}{2} \begin{pmatrix} 1 & i \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{2}(1 - 1) = 0, \text{ etc.}$$

- There are two orthonormal eigenstates, enough to act as a basis for the two dimensional state space i.e. any spin state can be written

$$|S\rangle = |1\rangle\langle 1|S\rangle + |2\rangle\langle 2|S\rangle.$$

Measurement & observables I

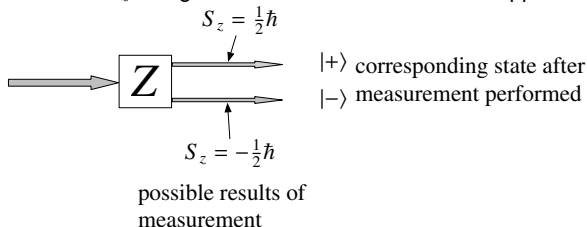
- One of the most important — and puzzling — parts of quantum mechanics is the measurement process.



- The system S being measured is coupled to the measurement apparatus.
 - Apparatus M is designed to measure, e.g. energy, spin, momentum, ...
- Apparatus is also interacting with its surrounding environment E .
 - The surrounding environment is a source of decoherence.
 - The apparatus ends up pointing at ONE of the possible results of the experiment (i.e. it is not in a linear combination of possible states.)
 - Still unexplained: how does it makes the 'choice'?

Measurement & observables II

- Apparatus is designed to measure some property of S , e.g. energy, momentum, position, spin, ...
- These properties are known as *observables*.
- The outcome of measuring an observable is *always* a real number.
- Below is the observable S_z being measured in a Stern-Gerlach apparatus:



- The state of the system is labelled by the result obtained in the measurement.
 - If the atom emerges with $S_z = \frac{1}{2}\hbar$, then the atom is assigned the state $|+\rangle$.

Measurement & observables III

- Likewise, if the atom emerges with $S_z = -\frac{1}{2}\hbar$, then it is assigned the state $|-\rangle$.
- Thus we have established a link between the information known about the state of the atom, i.e. a real number $S_z = \pm\frac{1}{2}\hbar$
- And the state assigned to the atom, $|\pm\rangle$.
- There is a clear correspondence here with Hermitean operators.

| Properties of a Hermitean Operator | Properties of Observable S_z |
|--|---|
| The eigenvalues of a Hermitean operator are all real. | Value of observable S_z measured to be real numbers $\pm\frac{1}{2}\hbar$. |
| Eigenvectors belonging to different eigenvalues are orthogonal. | States $ \pm\rangle$ associated with different values of the observable are mutually exclusive. |
| The eigenstates form a complete set of basis states for the state space of the system. | The states $ \pm\rangle$ associated with all the possible values of observable S_z form a complete set of basis states for the state space of the system. |

- This suggests associating a Hermitean operator with the observable S_z .
 - Call the operator \hat{S}_z .
 - Its eigenstates will be $|\pm\rangle$ and its associated eigenvalues $S_z = \pm \frac{1}{2}\hbar$:

$$\hat{S}_z|\pm\rangle = \pm \frac{1}{2}\hbar|\pm\rangle.$$

- Its matrix representation wrt the basis states $|\pm\rangle$ will be

$$\hat{S}_z \doteq \begin{pmatrix} \langle +|\hat{S}_z|+\rangle & \langle +|\hat{S}_z|-\rangle \\ \langle -|\hat{S}_z|+\rangle & \langle -|\hat{S}_z|-\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- This is a procedure that can be extended to any observable.
 - Though there are some mathematical subtleties if the observable has a continuous range of values, e.g. the position of a particle.
- We can now set up a procedure to construct a Hermitean operator to represent any observable.

Constructing an observable

- If a certain physical observable Q is measured to have the values q_1, q_2, q_3, \dots , then this observable is represented by a Hermitean operator \hat{Q} such that
 1. The eigenvalues of \hat{Q} are *all* the possible values q_1, q_2, q_3, \dots of Q that could be obtained in a measurement.
 2. If on performing a measurement of Q the result q_n is obtained, then the system will end up in the associated eigenstate $|q_n\rangle$ where

$$\hat{Q}|q_n\rangle = q_n|q_n\rangle$$

where we have labelled the eigenstate by its associated eigenvalue.

3. The eigenstates $|q_1\rangle, |q_2\rangle, \dots$ of \hat{Q} are assumed to form a *complete* orthonormal set of basis states.
 - This assumption is physically based, i.e. if we have exhaustively determined all the possible values for an observable Q (either by experiment or by theoretical argument), then we are claiming that any state $|\psi\rangle$ of the system can be written

$$|\psi\rangle = |q_1\rangle\langle q_1|\psi\rangle + |q_2\rangle\langle q_2|\psi\rangle + \dots$$

4. If the system is in the state $|\psi\rangle$ then the probability of getting the result q_n on measuring Q is

$$|\langle q_n|\psi\rangle|^2$$

- Note that the operator \hat{Q} itself is often referred to as an observable.

- The **position** \hat{x} of a particle in one dimension.
 - As the particle can have any position, its eigenvalues x have all values $-\infty < x < \infty$.
 - The eigenvalue equation is $\hat{x}|x\rangle = x|x\rangle$.
 - Completeness tells us that

$$|\psi\rangle = \int_{-\infty}^{+\infty} |x\rangle \langle x|\psi\rangle dx \quad \text{— an integral as } x \text{ is continuous.}$$

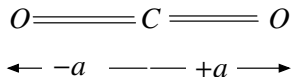
- Note that $\langle x|\psi\rangle$ is the probability amplitude of finding the particle at x , i.e.

$$\langle x|\psi\rangle \equiv \psi(x) \quad \text{the de Broglie wave function.}$$

- We have come full circle: the wave function idea led to setting up the mathematical structure of quantum mechanics, and now we see the wave function re-emerge from the formalism.
- We will not be dealing with observables with continuous eigenvalues — this is here for illustration purposes only.

Examples of observables II

- The **approximate position** of a particle.
 - This is a 'discrete' version of the position operator as in, e.g., the position of an electron on the atoms of a CO₂ ion:



- The position x has the three discrete values $x = -a, 0, +a$ so the associated operator \hat{x} has three eigenvalues $x = 0, \pm a$
- The eigenvalue equations and the matrix representing \hat{x} are

$$\begin{aligned} \hat{x}|0\rangle &= 0|0\rangle \\ \hat{x}|\pm a\rangle &= \pm a|\pm a\rangle \end{aligned} \quad \hat{x} \doteq \begin{pmatrix} \langle -a|\hat{x}|-a\rangle & \langle -a|\hat{x}|0\rangle & \langle -a|\hat{x}|+a\rangle \\ \langle 0|\hat{x}|-a\rangle & \langle 0|\hat{x}|0\rangle & \langle 0|\hat{x}|+a\rangle \\ \langle +a|\hat{x}|-a\rangle & \langle +a|\hat{x}|0\rangle & \langle +a|\hat{x}|+a\rangle \end{pmatrix} = \begin{pmatrix} -a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & +a \end{pmatrix}$$

- The state space has dimension three, so *all* other observables for the system will be represented by 3×3 Hermitean matrices.
- In particular, the energy and the momentum of the electron will be represented by 3×3 Hermitean matrices.

Examples of observables III

- The **momentum** p of a particle in one dimension can have any value $-\infty < p < \infty$.
- The operator is \hat{p} , its eigenstates are $|p\rangle$, and the eigenvalue equation is $\hat{p}|p\rangle = p|p\rangle$.
- Can also define an approximate momentum operators for approximate position models such as the CO_2 ion. The matrix representing \hat{p} turns out to look like:

$$\hat{p} \doteq \begin{pmatrix} \langle -a|\hat{p}|-a\rangle & \langle -a|\hat{p}|0\rangle & \langle -a|\hat{p}|+a\rangle \\ \langle 0|\hat{p}|-a\rangle & \langle 0|\hat{p}|0\rangle & \langle 0|\hat{p}|+a\rangle \\ \langle +a|\hat{p}|-a\rangle & \langle +a|\hat{p}|0\rangle & \langle +a|\hat{p}|+a\rangle \end{pmatrix} = \begin{pmatrix} 0 & ip & 0 \\ -ip & 0 & ip \\ 0 & -ip & 0 \end{pmatrix}$$

- The position basis states $|0\rangle, |\pm a\rangle$ have been used to construct the matrix.
- The matrix is said to be given in the position representation.
- This matrix is Hermitean — as it should be — and its eigenvalues are $0, \pm p$.
- Note that the position and momentum operators do not commute i.e.

$$\hat{x}\hat{p} - \hat{p}\hat{x} \neq 0$$

Examples of observables IV

- The **energy** of a particle is usually represented by the operator \hat{H} known as the Hamiltonian.
- A general expression for the Hamiltonian is

$$\begin{array}{ccccccc} \hat{H} & = & \frac{\hat{P}^2}{2m} & + & V(\hat{x}) \\ \downarrow & & \downarrow & & \downarrow \\ \text{Total energy} & = & \text{kinetic energy} & + & \text{potential energy} \end{array}$$

- Depending on the nature of the potential $V(\hat{x})$, find that \hat{H} has different possible eigenvalues and eigenstates, e.g.

$$V(\hat{x}) = \frac{1}{2}m\omega^2\hat{x}^2 \quad \text{SHO potential}$$

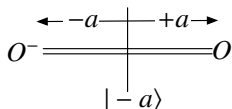
$$E_n = (n + \frac{1}{2})\hbar\omega \quad n = 0, 1, 2, 3, \dots$$

$$\hat{H}|n\rangle = E_n|n\rangle$$

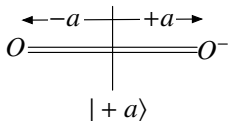
$$\langle x|n\rangle = \psi_n(x) \quad \text{wave function for simple harmonic oscillator}$$

The O_2 ion I

- The O_2 ion provides a useful model for illustrating many of the features of observables of a simple system.



- The positions of the electron are $x = \pm a$, so the position operator \hat{x} has eigenvalues $x = \pm a$, with associated eigenstates $|\pm a\rangle$, i.e. $\hat{x}|\pm a\rangle = \pm a|\pm a\rangle$



- As a matrix in the position representation:

$$\hat{x} \doteq \begin{pmatrix} -a & 0 \\ 0 & a \end{pmatrix}.$$

- As the state space has dimension 2, all observables will be represented by matrices of size 2×2 . In particular, the Hamiltonian matrix in the position representation:

$$\hat{H} \doteq \begin{pmatrix} \langle -a|\hat{H}|-a\rangle & \langle -a|\hat{H}|+a\rangle \\ \langle +a|\hat{H}|-a\rangle & \langle +a|\hat{H}|+a\rangle \end{pmatrix} = \begin{pmatrix} E_0 & V \\ V^* & E_0 \end{pmatrix}$$

- Note that *by symmetry* we *must* have $\langle -a|\hat{H}|-a\rangle = \langle +a|\hat{H}|+a\rangle$ as there is nothing to distinguish the oxygen atom at $-a$ from the one at $+a$.
- And as \hat{H} is Hermitean, the off-diagonal elements *must* be complex conjugates of each other.

- For simplicity, shall assume $V = -A$ where A is real.
 - This has no physical significance.
- Can easily find the eigenvalues and eigenvectors of \hat{H} :

Eigenvalue

Eigenvector

$$E_1 = E_0 + A \quad |E_1\rangle = \frac{1}{\sqrt{2}} (|-a\rangle - |+a\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$E_2 = E_0 - A \quad |E_2\rangle = \frac{1}{\sqrt{2}} (|-a\rangle + |+a\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

for example:

$$\hat{H}|E_1\rangle = E_1|E_1\rangle \implies \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (E_0 + A) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

- Thus the ion has two energy levels. Their energy difference, $2A$ is related to the coupling that allows the electron to 'tunnel' from one oxygen atom to the other.

- Note that we now have two possible sets of basis states: $\{|-a\rangle, |+a\rangle\}$ or $\{|E_1\rangle, |E_2\rangle\}$.
 - There are an infinity of other possible basis states.
- Each choice of basis states can be used to construct the matrix representing the operators and state vectors.
 - If we use the eigenstates of the position operator, i.e. $\{|-a\rangle, |+a\rangle\}$, we are using the *position representation*.
 - If we use the eigenstates of the Hamiltonian, i.e. $\{|E_1\rangle, |E_2\rangle\}$, we are using the *energy representation*.

$$\hat{H} \doteq \begin{pmatrix} \langle E_1 | \hat{H} | E_1 \rangle & \langle E_1 | \hat{H} | E_2 \rangle \\ \langle E_2 | \hat{H} | E_1 \rangle & \langle E_2 | \hat{H} | E_2 \rangle \end{pmatrix} = \begin{pmatrix} E_0 + A & 0 \\ 0 & E_0 - A \end{pmatrix} \text{ in the energy representation.}$$

- Another possible representation is the momentum representation, where we use the eigenstates of the momentum operator.
- Shall mostly use the position representation here. This and the energy representation are the most commonly used.

- Suppose we prepare the O_2 ion in the state

$$|\psi\rangle = \frac{1}{5} (3| - a\rangle + 4| + a\rangle) = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

and we measure the energy of the ion.

- We can get two possible results: E_1 and E_2 .

- We will get the result E_1 with probability $|\langle E_1|\psi\rangle|^2$ i.e.

$$\langle E_1|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{-1}{5\sqrt{2}} \Rightarrow |\langle E_1|\psi\rangle|^2 = \frac{1}{50} = 0.02.$$

- And the result E_2 with probability $|\langle E_2|\psi\rangle|^2$ i.e.

$$\langle E_2|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{7}{5\sqrt{2}} \Rightarrow |\langle E_2|\psi\rangle|^2 = \frac{49}{50} = 0.98.$$

- In general, for arbitrary systems in a state $|\psi\rangle$ and an observable Q is measured, the probability of getting the result q_n is $|\langle q_n|\psi\rangle|^2$.

- Any repeat of a measurement of some observable Q on identical copies of the same system prepared in the same state $|\psi\rangle$ will not necessarily yield the same result every time.
 - It therefore makes sense to use standard statistical tools to analyse the random data that would be collected.
 - These include the average (or expectation) value of the data, and the standard deviation (or uncertainty) of this data, this being a measure of the spread of the data about the average value.
- So suppose we have N identical copies of the same system all prepared in the same state $|\psi\rangle$.
- We then measure some observable Q , with the possible results q_1, q_2, q_3, \dots
- As the result q_n will occur with probability $|\langle q_n | \psi \rangle|^2$, the result q_n will occur, roughly, $N |\langle q_n | \psi \rangle|^2$ times.

- So the average value or *expectation value* of all these results will be

$$\begin{aligned}\langle Q \rangle &= \frac{N |\langle q_1 | \psi \rangle|^2 q_1 + N |\langle q_2 | \psi \rangle|^2 q_2 + N |\langle q_3 | \psi \rangle|^2 q_3 + \dots}{N} \\ &= |\langle q_1 | \psi \rangle|^2 q_1 + |\langle q_2 | \psi \rangle|^2 q_2 + |\langle q_3 | \psi \rangle|^2 q_3 + \dots\end{aligned}$$

- Now turn to the expression $\langle \psi | \hat{Q} | \psi \rangle$ which we can expand as follows:

$$\begin{aligned}\langle \psi | \hat{Q} | \psi \rangle &= \langle \psi | (\hat{Q} [|q_1\rangle \langle q_1 | \psi \rangle + |q_2\rangle \langle q_2 | \psi \rangle + \dots]) \\ &= \langle \psi | (\hat{Q} |q_1\rangle \langle q_1 | \psi \rangle + \hat{Q} |q_2\rangle \langle q_2 | \psi \rangle + \dots) && \text{using linearity of } \hat{Q} \\ &= \langle \psi | (q_1 |q_1\rangle \langle q_1 | \psi \rangle + q_2 |q_2\rangle \langle q_2 | \psi \rangle + \dots) && \text{using } \hat{Q} |q_n\rangle = q_n |q_n\rangle \\ &= q_1 \langle \psi | q_1 \rangle \langle q_1 | \psi \rangle + q_2 \langle \psi | q_2 \rangle \langle q_2 | \psi \rangle + \dots \\ &= q_1 |\langle q_1 | \psi \rangle|^2 + q_2 |\langle q_2 | \psi \rangle|^2 + \dots && \text{since } \langle q_n | \psi \rangle = \langle \psi | q_n \rangle^* \\ &= \langle Q \rangle.\end{aligned}$$

- Thus, the expectation value of the observable Q when the system is in the state $|\psi\rangle$ is

$$\langle Q \rangle = \langle \psi | \hat{Q} | \psi \rangle.$$

Expectation Values III

- For example, consider an O_2^- ion in the state $|\psi\rangle = (3i| - a\rangle + 4| + a\rangle) / 5$. The expectation value of the energy of the ion will then be

$$\langle\psi|\hat{H}|\psi\rangle = \frac{1}{25} \begin{pmatrix} -3i & 4 \end{pmatrix} \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} 3i \\ 4 \end{pmatrix} = E_0.$$

- This result can also be calculated via

$$\langle\psi|\hat{H}|\psi\rangle = E_1 |\langle E_1|\psi\rangle|^2 + E_2 |\langle E_2|\psi\rangle|^2$$

and working out the probabilities explicitly from the expressions for the vectors $|E_1\rangle, |E_2\rangle, |\psi\rangle$. E.g.

$$\langle E_1|\psi\rangle = \frac{1}{5\sqrt{2}} (\langle -a| - \langle +a|) (3i| - a\rangle + 4| + a\rangle) = \frac{3i - 4}{5\sqrt{2}}$$

$$\langle E_2|\psi\rangle = \frac{1}{5\sqrt{2}} (\langle -a| + \langle +a|) (3i| - a\rangle + 4| + a\rangle) = \frac{3i + 4}{5\sqrt{2}}$$

and hence
$$\langle\psi|\hat{H}|\psi\rangle = (E_0 + A) \left| \frac{3i - 4}{5\sqrt{2}} \right|^2 + (E_0 - A) \left| \frac{3i + 4}{5\sqrt{2}} \right|^2 = E_0$$

Collapse of the state vector I

- One of the central postulates of quantum mechanics is that of specifying the state of a system *after* a measurement is made.

- The postulate due to von Neumann is that:

If after the measurement of an observable Q the result q_n is obtained, then the system ends up in the eigenstate $|q_n\rangle$ of \hat{Q} immediately after the measurement.

- This is not the whole story, however, as there are measurements that *destroy* the system being measured.
 - Measuring the number of photons in a light field involves absorbing them, so if you count 10 photons, the state after the measurement is certainly not $|10\rangle$!
- However, for non-destructive measurements (e.g. the Stern-Gerlach apparatus), the von Neumann postulate is the way to go.
- It is possible, in fact, to relate *all* measurements back to a ‘von Neumann’ measurement.

Collapse of the state vector II

- So, returning to the O_2 ion example, and suppose we measure the energy of the ion.
 - The possible results are E_1 and E_2 .
 - If the result E_1 is obtained, then the ion ends up in the state $|E_1\rangle$.
- We could also measure the position of the electron.
 - The possible results are $+a$ or $-a$.
 - If the result $+a$ is obtained, then the ion ends up in the state $|+a\rangle$.
- This 'quantum process' is known as 'the collapse of the state vector'.
 - Is it a physical process?
 - The main school-of-thought is that it is *not* a physical process.
 - It represents an 'updating' of the information we have about the state of a physical system.
 - The measurement has given us new information about the state of the system, so we must assign a new state to the system in accordance with this new information.

- We can study consequences of alternate measurements of energy and position:
 1. Measure energy and get result E_1 say. New state of ion is $|E_1\rangle$.
 2. Now measure position of electron. Will get result $-a$ with probability $|\langle -a|E_1\rangle|^2 = \frac{1}{2}$, or result $+a$, also with probability $\frac{1}{2}$. Suppose we get $-a$. New state is $|-a\rangle$.
 3. Now remeasure the energy. We will get the result E_1 with probability $|\langle E_1|-a\rangle|^2 = \frac{1}{2}$ or the result E_2 also with probability $\frac{1}{2}$. Suppose we get E_2 . New state is $|E_2\rangle$.
 4. Measuring energy scrambles result for measurement of position, and vice versa.
- Note that the operators \hat{H} and \hat{x} do not commute:

$$\hat{H}\hat{x} - \hat{x}\hat{H} \doteq \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} -a & 0 \\ 0 & +a \end{pmatrix} - \begin{pmatrix} -a & 0 \\ 0 & +a \end{pmatrix} \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} = -2aA \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- As a general rule, if two observables do not commute, then the measurement of one scrambles the measurement of the other.
 - For instance, for spin $[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$
 - The most important commutation rule is $[\hat{x}, \hat{p}] = i\hbar$.

- Just as it is the case for classical systems, quantum systems evolve in time.
- In other words, the state vector is time dependent $|\psi(t)\rangle$.
- What is needed is an equation that describes this dynamics.
- To get an idea of what this dynamics might be, and hence what the equation will look like, we will consider the simplest possible case: an isolated system whose state does not change in time.
 - Such states are called *stationary states*.
- In spite of their name, stationary states do evolve in time, but in a particularly simple way that leads us to connect time evolution with energy, and hence the Hamiltonian.

Evolution of a stationary state I

- Recall that we can multiply a state vector by any number, and it still represents the same physical state of affairs. Thus, for a stationary state $|\psi(t)\rangle$ we can have

$$|\psi(t)\rangle = u(t)|\psi(0)\rangle \quad u(t) \text{ a complex number in general.}$$

- Further recall that it is states must be normalized to unity to have a correct probability interpretation. Thus we require

$$\langle\psi(t)|\psi(t)\rangle = |u(t)|^2\langle\psi(0)|\psi(0)\rangle$$

so if $|\psi(0)\rangle$ is normalized to unity, then so is $|\psi(t)\rangle$ provided $|u(t)|^2 = 1$.

- Thus we conclude $u(t) = e^{-i\phi(t)}$ where $\phi(t)$ is an unknown real valued time dependent function.
- Next, we make use of the fact that the system is *isolated*.
 - An isolated system is uninfluenced by anything external, nor does it influence an other systems.
 - An open system will exchange energy and matter with its surroundings.

Evolution of a stationary state II

- For an isolated system, it makes no difference when you start the clock — all that matters is how much time has elapsed.
 - I.e. we could allow the stationary state evolve starting at 10am for ten minutes, or start at 9:38pm and let it run for ten minutes. Either way, the system will end up being in the same state after ten minutes irrespective of when it started to evolve.
- What all this amounts to as that, for an isolated system, we can choose the initial time to be any arbitrary time t_0 , and the evolution of $|\psi(t_0)\rangle$ up to the time t would be given by

$$|\psi(t)\rangle = e^{-i\phi(t-t_0)}|\psi(t_0)\rangle$$

- If we now consider the evolution of the system in a stationary state over a time interval (t, t') , then we have

$$|\psi(t')\rangle = e^{-i\phi(t'-t)}|\psi(t)\rangle = e^{-i\phi(t'-t)}e^{-i\phi(t-t_0)}|\psi(t_0)\rangle$$

- But over the interval (t_0, t') we have

$$|\psi(t')\rangle = e^{-i\phi(t'-t_0)}|\psi(t_0)\rangle$$

- So by comparing the last two equations we find that

$$e^{-i\phi(t'-t_0)} = e^{-i\phi(t'-t)} e^{-i\phi(t-t_0)}$$

or

$$\phi(t - t_0) = \phi(t' - t) + \phi(t - t_0),$$

- This is an equation for $\phi(t)$ with the solution

$$\phi(t) = \omega t$$

where ω is a constant. Thus we can conclude that

$$|\psi(t)\rangle = e^{-i\omega t} |\psi(0)\rangle.$$

- The quantity ω appearing in $|\psi(t)\rangle = e^{-i\omega t}|\psi(0)\rangle$ has the units of angular momentum.
 - Recall the Einstein-Planck relation between energy and angular momentum: $E = \hbar\omega$.
- This suggests we write

$$|\psi(t)\rangle = e^{-iEt/\hbar}|\psi(0)\rangle \quad E = \hbar\omega.$$

- This further suggests that we identify the energy E as being an attribute of the stationary state $|\psi(0)\rangle$.
 - More than that, we are going to assume that for each possible stationary state of a system, there will be an associated energy E that determines its time evolution.
- Thus, we will assume that $|\psi(0)\rangle$ is an eigenstate of the energy observable, the Hamiltonian \hat{H} , for the system, with energy eigenvalue E , i.e. $|\psi(0)\rangle \equiv |E\rangle$ with

$$\hat{H}|E\rangle = E|E\rangle$$

and

$$|E(t)\rangle = e^{-iEt/\hbar}|E\rangle.$$

The Schrödinger equation I

- If we differentiate $|E(t)\rangle = e^{-iEt/\hbar}|E\rangle$ with respect to time we find

$$i\hbar \frac{d|E(t)\rangle}{dt} = E|E(t)\rangle = \hat{H}|E(t)\rangle$$

- Now consider an arbitrary state $|\psi(t)\rangle$.
- Since the eigenstates of \hat{H} , call them $\{|E_1\rangle, |E_2\rangle, \dots, |E_N\rangle\}$, form a complete set of basis states, we can write for the initial state

$$|\psi(0)\rangle = \sum_{n=1}^N |E_n\rangle \langle E_n | \psi(0)\rangle.$$

- The time evolved version is then

$$|\psi(t)\rangle = \sum_{n=1}^N e^{-iE_n t/\hbar} |E_n\rangle \langle E_n | \psi(0)\rangle$$

- We have, in effect, assumed that the time evolution can be obtained by replacing each of the basis states $|E_n\rangle$ by their time evolved form $e^{-iE_n t/\hbar}|E_n\rangle$.

The Schrödinger equation II

- On differentiating, this becomes

$$\begin{aligned}i\hbar \frac{d|\psi(t)\rangle}{dt} &= i\hbar \sum_{n=1}^N E_n e^{-iE_n t/\hbar} |E_n\rangle \langle E_n | \psi(t)\rangle \\ &= \sum_{n=1}^N \hat{H} e^{-iE_n t/\hbar} |E_n\rangle \langle E_n | \psi(t)\rangle \\ &= \hat{H} \sum_{n=1}^N e^{-iE_n t/\hbar} |E_n\rangle \langle E_n | \psi(t)\rangle \\ &= \hat{H} |\psi(t)\rangle.\end{aligned}$$

- Thus we end up with

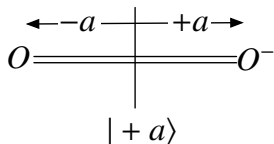
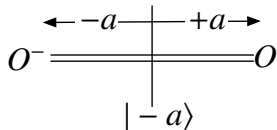
$$\hat{H} |\psi(t)\rangle = i\hbar \frac{d|\psi(t)\rangle}{dt}$$

which is the celebrated Schrödinger equation in vector form.

- Note that this equation does not explicitly tell us how the system is evolving, it is more correct to say that it tells us how the information we can gain about the system is evolving.

Time evolution of the O_2^- ion

- Shall illustrate how to set up and solve the Schrödinger equation for a simple system, the O_2^- ion.



- The Hamiltonian matrix for this ion, in the position representation is

$$\hat{H} \doteq \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}$$

- An arbitrary state of the system at time t can be written

$$\begin{aligned} |\psi(t)\rangle &= |-a\rangle\langle -a|\psi(t)\rangle + |+a\rangle\langle +a|\psi(t)\rangle \\ &= \psi_-(t)|-a\rangle + \psi_+(t)|+a\rangle \\ &\doteq \begin{pmatrix} \psi_-(t) \\ \psi_+(t) \end{pmatrix}. \end{aligned}$$

- Shall assume that initially, electron is on oxygen atom at $-a$, so that $|\psi(0)\rangle = |-a\rangle$.
- This problem can be solved either in the position representation or in the energy representation.

Evolution of the O_2^- ion

Solution in the energy representation I

- We do not actually have to solve the Schrödinger equation.
 - Instead, we use our knowledge of how energy eigenstates evolve, i.e. $|E(t)\rangle = e^{-iEt/\hbar}|E\rangle$
 - Doing this requires knowing what the energy eigenstates and eigenvalues of the Hamiltonian are.
 - Unfortunately, solving the eigenvalue problem for most Hamiltonians is a very difficult task, just as hard to do as solving the Schrödinger equation itself.
- Here we know the eigenstates and eigenvalues of the Hamiltonian:

$$|E_1\rangle = \frac{1}{\sqrt{2}} (|-a\rangle - |+a\rangle) \quad E_1 = E_0 + A$$

$$|E_2\rangle = \frac{1}{\sqrt{2}} (|-a\rangle + |+a\rangle) \quad E_2 = E_0 - A$$

- In terms of the basis states $\{|E_1\rangle, |E_2\rangle\}$, the state of the system at time t is

$$|\psi(t)\rangle = e^{-iE_1t/\hbar}|E_1\rangle\langle E_1|\psi(0)\rangle + e^{-iE_2t/\hbar}|E_2\rangle\langle E_2|\psi(0)\rangle.$$

- The initial condition tells us that

$$\langle E_1 | \psi(0) \rangle = \langle E_1 | -a \rangle = \frac{1}{\sqrt{2}} \quad \langle E_2 | \psi(0) \rangle = \langle E_2 | -a \rangle = \frac{1}{\sqrt{2}}$$

- Thus we have
- $$\begin{aligned} |\psi(t)\rangle &= \frac{1}{\sqrt{2}} \left(e^{-iE_1 t/\hbar} |E_1\rangle + e^{-iE_2 t/\hbar} |E_2\rangle \right) \\ &= \frac{1}{\sqrt{2}} \left(e^{-i(E_0+A)t/\hbar} |E_1\rangle + e^{-i(E_0-A)t/\hbar} |E_2\rangle \right) \\ &= \frac{1}{\sqrt{2}} e^{-iE_0 t/\hbar} \left(e^{-iAt/\hbar} |E_1\rangle + e^{iAt/\hbar} |E_2\rangle \right) \end{aligned}$$

- We are after the probability of finding the atom in either of the states $| -a \rangle$ and $| +a \rangle$, given by

$$P_-(t) = |\langle -a | \psi(t) \rangle|^2 \quad \text{and} \quad P_+(t) = |\langle +a | \psi(t) \rangle|^2 .$$

Evolution of the O_2^- ion

Solution in the energy representation III

- Thus, for instance
$$\begin{aligned}\langle -a|\psi(t)\rangle &= \frac{1}{\sqrt{2}}e^{-iE_0t/\hbar} \left(e^{-iAt/\hbar}\langle -a|E_1\rangle + e^{iAt/\hbar}\langle -a|E_2\rangle \right) \\ &= \frac{1}{2}e^{-iE_0t/\hbar} \left(e^{-iAt/\hbar} + e^{iAt/\hbar} \right) \\ &= e^{-iE_0t/\hbar} \cos(At/\hbar).\end{aligned}$$

- Consequently
$$P_-(t) = \left| e^{-iE_0t/\hbar} \cos(At/\hbar) \right|^2 = \cos^2(At/\hbar)$$

- Similarly, we find
$$P_+(t) = \sin^2(At/\hbar).$$

- Shall analyse these results later.

Evolution of the O_2^- ion

Solution in the position representation I

- It is more typically the case that the Hamiltonian is given in the position representation
- In this representation, the Schrödinger equation for this system is

$$\hat{H}|\psi\rangle = i\hbar \frac{d|\psi\rangle}{dt} \Rightarrow \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} \psi_-(t) \\ \psi_+(t) \end{pmatrix} = i\hbar \frac{d}{dt} \begin{pmatrix} \psi_-(t) \\ \psi_+(t) \end{pmatrix} = i\hbar \begin{pmatrix} \dot{\psi}_-(t) \\ \dot{\psi}_+(t) \end{pmatrix}$$

- Expanding out the matrix multiplication yields two coupled first order differential equations:

$$E_0\psi_- - A\psi_+ = i\hbar\dot{\psi}_- \quad (1)$$

$$-A\psi_- + E_0\psi_+ = i\hbar\dot{\psi}_+ \quad (2)$$

- There are a multitude of ways of solving these equations.
 - Deep down inside, all the methods are equivalent to finding the eigenvalues and eigenvectors of the Hamiltonian!!

Evolution of the O_2^- ion

Solution in the position representation II

- One method of solution is as follows.
- Add and subtract Eqs. (1) and (2):

$$(E_0 - A)(\psi_- + \psi_+) = i\hbar(\dot{\psi}_- + \dot{\psi}_+)$$

$$(E_0 + A)(\dot{\psi}_- - \dot{\psi}_+) = i\hbar(\psi_- - \psi_+)$$

- Defining new variables $X = \psi_- + \psi_+$ and $Y = \psi_- - \psi_+$ then gives

$$\dot{X} = (i\hbar)^{-1} (E_0 - A) X$$

$$\dot{Y} = (i\hbar)^{-1} (E_0 + A) Y$$

- These are readily solved to give

$$X(t) = e^{-i(E_0 - A)t/\hbar} X(0) \quad \text{and} \quad Y(t) = e^{-i(E_0 + A)t/\hbar} Y(0).$$

- We can recover $\psi_{\pm}(t)$ by using $\psi_- = \frac{1}{2}(X + Y)$ and $\psi_+ = \frac{1}{2}(X - Y)$

Evolution of the O_2^- ion

Solution in the position representation III

- Eliminating the quantities X and Y then gives:

$$\begin{aligned}\psi_-(t) &= \frac{1}{2} \left(e^{-i(E_0-A)t/\hbar} X(0) + e^{-i(E_0+A)t/\hbar} Y(0) \right) \\ &= \frac{1}{2} e^{-iE_0t/\hbar} \left(e^{iAt/\hbar} (\psi_-(0) + \psi_+(0)) + e^{-iAt/\hbar} (\psi_-(0) - \psi_+(0)) \right) \\ &= \frac{1}{2} e^{-iE_0t/\hbar} \left((e^{iAt/\hbar} + e^{-iAt/\hbar}) \psi_-(0) + (e^{iAt/\hbar} - e^{-iAt/\hbar}) \psi_+(0) \right) \\ &= e^{-iE_0t/\hbar} \left[\cos(At/\hbar) \psi_-(0) + i \sin(At/\hbar) \psi_+(0) \right]\end{aligned}$$

- Similarly we find

$$\psi_+(t) = e^{-iE_0t/\hbar} \left[i \sin(At/\hbar) \psi_-(0) + \cos(At/\hbar) \psi_+(0) \right]$$

- Now make use of the initial conditions $|\psi(0)\rangle = |-a\rangle$
- This tells us that $\psi_-(0) = 1$ and $\psi_+(0) = 0$ and hence

$$\psi_-(t) = e^{-iE_0t/\hbar} \cos(At/\hbar) \quad \text{and} \quad \psi_+(t) = e^{-iE_0t/\hbar} i \sin(At/\hbar).$$

Evolution of the O_2^- ion

Solution in the position representation IV

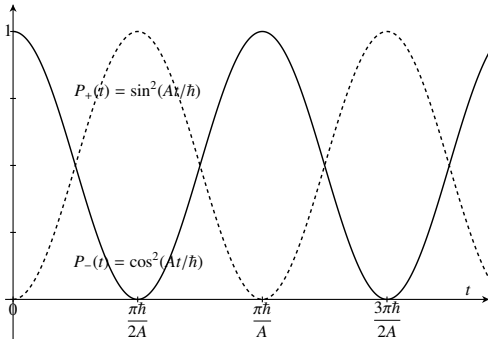
- The state vector for the ion as a function of time is then

$$|\psi(t)\rangle = e^{-iE_0t/\hbar} \left(\cos(At/\hbar) | -a \rangle + i \sin(At/\hbar) | +a \rangle \right)$$

- The probabilities $P_{\pm}(t)$ of finding the electron on the oxygen atom at $\pm a$ will then be

$$P_-(t) = |\langle -a | \psi(t) \rangle|^2 = \cos^2(At/\hbar) \quad P_+(t) = |\langle +a | \psi(t) \rangle|^2 = \sin^2(At/\hbar).$$

- Note that $P_-(t) + P_+(t) = 1$, i.e. probability is conserved.



- The probabilities oscillate with a frequency $f = A/(\pi\hbar)$, or with a period $T = \pi\hbar/A$.
- The electron is initially on the oxygen atom at $-a$, but eventually, after a time $\pi\hbar/2A$, it will be found with 100% certainty on the oxygen atom at $+a$.

- This oscillatory behaviour is found in all two state systems:
 - Magnetic dipole of a spin half atom in a magnetic field (i.e. precession of the magnetic dipole)
 - A two level atom in its ground state placed in a resonant single mode cavity containing one photon. The two states are then

$|g, 1\rangle$ atom in ground state, one photon present

$|e, 0\rangle$ atom in excited state, no photon present.

The oscillations are called vacuum Rabi oscillations.

- Two potential wells very close by (e.g. quantum dots). The electron can tunnel from one well to the other.
- Which brings us to: what is the meaning of A ?

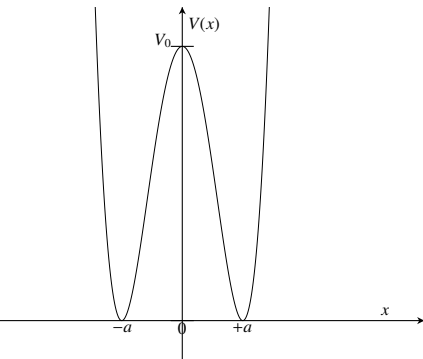
The meaning of the off-diagonal elements I

- To understand the meaning of the off-diagonal elements A in the O_2^- ion Hamiltonian, it is best to look at the particular case of $A = 0$ first.
- The Hamiltonian matrix looks like, in the position representation:

$$\hat{H} \doteq \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix}.$$

- The states $|\pm a\rangle$ are now eigenstates of the Hamiltonian, each with eigenvalue E_0 (i.e. E_0 is a degenerate eigenvalue).
- In other words, the states $|\pm a\rangle$ are now *stationary* states of the ion.
 - Put the electron on any one of the oxygen atoms, and it stays there.
- Physically, this will occur if the electron is trapped in an infinitely deep very narrow potential well from which it cannot escape.

The meaning of the off-diagonal elements II

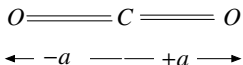


- If the barrier height $V_0 \rightarrow \infty$, then the electron at $+a$ will stay there, and similarly if the electron is at $-a$.
- If, however, V_0 is finite, then the electron in, for instance, the $+a$ well, can tunnel through to the $-a$ well.
 - The states $|\pm a\rangle$ are no longer stationary states, i.e. if $|\psi(0)\rangle = | + a \rangle$, then $|\psi(t)\rangle$ will not stay as $| + a \rangle$.
 - If the states $|\pm a\rangle$ are not stationary states, then the Hamiltonian must have non-zero off-diagonal elements.
- Thus $A \neq 0$ corresponds to V_0 finite.
 - Can show that approximately

$$A \sim e^{-\alpha V_0} \quad \alpha \text{ a constant.}$$

The meaning of the off-diagonal elements III

- Can generalise the idea to e.g. the CO_2^- ion:



- The basis states are $| -a \rangle$, $|0\rangle$ and $| +a \rangle$ corresponding to the three possible positions of the electron.

- The electron can tunnel from $x = -a$ to $x = 0$ (and back), or from $x = 0$ to $x = +a$ (and back). But we exclude tunnelling from $-a$ to $+a$.
- The Hamiltonian is then

$$\hat{H} \doteq \begin{pmatrix} E_0 & -A & 0 \\ -A & E_c & -A \\ 0 & -A & E_0 \end{pmatrix}$$

i.e. $\langle -a | \hat{H} | +a \rangle = 0$.

- In general, for any Hamiltonian \hat{H} , and for a set of basis states $\{|1\rangle, |2\rangle, \dots\}$, if $\langle i | \hat{H} | j \rangle \neq 0$ then neither of the states $|i\rangle$ or $|j\rangle$ will be stationary states.
- Such non-zero matrix elements indicate the prospect of the system 'making a transition' from the state $|i\rangle$ to $|j\rangle$ and back.

- A final example: a spin half particle will, in general, have a magnetic moment:

$$\hat{\boldsymbol{\mu}} = \mu \hat{\mathbf{S}}$$

where $\hat{\mathbf{S}}$ is a 'vector operator': $\hat{\mathbf{S}} = \hat{S}_x \hat{\mathbf{i}} + \hat{S}_y \hat{\mathbf{j}} + \hat{S}_z \hat{\mathbf{k}}$.

- The energy of the particle in a magnetic field \mathbf{B} is

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \hat{\mathbf{B}}.$$

- Using the spin states $|\pm\rangle$ as basis states, the Hamiltonian can be shown to be

$$\hat{H} \doteq \mu \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}$$

so if $B_z = B_y = 0$, i.e. the magnetic field in the x direction, we have the Hamiltonian of the same form as the O_2^- ion, so the initial spin state $|\psi(0)\rangle = |+\rangle$ evolves into

$$|\psi(t)\rangle = \cos(\mu B_x t / \hbar) |+\rangle + i \sin(\mu B_x t / \hbar) |-\rangle$$

- I.e. the spin 'precesses' with a frequency $f = \frac{\mu B_x}{\pi \hbar}$