

CfE Advanced Higher Physics – Unit 2 – Quanta

INTRODUCTION TO QUANTUM THEORY

1. Challenges to classical theory
2. Black body radiation
3. Photoelectric effect
4. Bohr model of the atom
5. Wave-particle duality
6. De Broglie waves
7. Uncertainty principle

PARTICLES FROM SPACE

8. Cosmic rays
9. Solar wind

Quantum Theory

An Introduction to Quantum Theory

Quantum mechanics was developed in the early twentieth century to explain experimental observations that could not be explained by classical physics. In many cases various quantisation ‘rules’ were proposed to explain these experimental observations but these did not have any classical justification. This results in physics that is difficult to reconcile with everyday experience or with normal intuition.

It led two of the most famous physicists of the twentieth century to make the following statements:

“If quantum mechanics hasn’t profoundly shocked you, you haven’t understood it yet.”
Niels Bohr

“I think it safe to say that no one understands quantum mechanics.”
Richard Feynman

Despite this, quantum mechanics has been one of the most successful theories in physics explaining many experimental observations such as blackbody radiation, the photoelectric effect, emission and absorption spectra, atomic structure and electron diffraction. It is these observations which provide the supporting evidence for Quantum Theory.

Blackbody Radiation

Towards the end of the nineteenth century there was interest in the frequencies (or wavelengths) emitted by a ‘black body’ when its temperature is increased. When an object is heated it can radiate large amounts of energy as infrared radiation. We can feel this if we place a hand near, but not touching, a hot object. As an object becomes hotter it starts to glow a dull red, followed by bright red, then orange, yellow and finally white (white hot). At extremely high temperatures it becomes a bright blue-white colour.

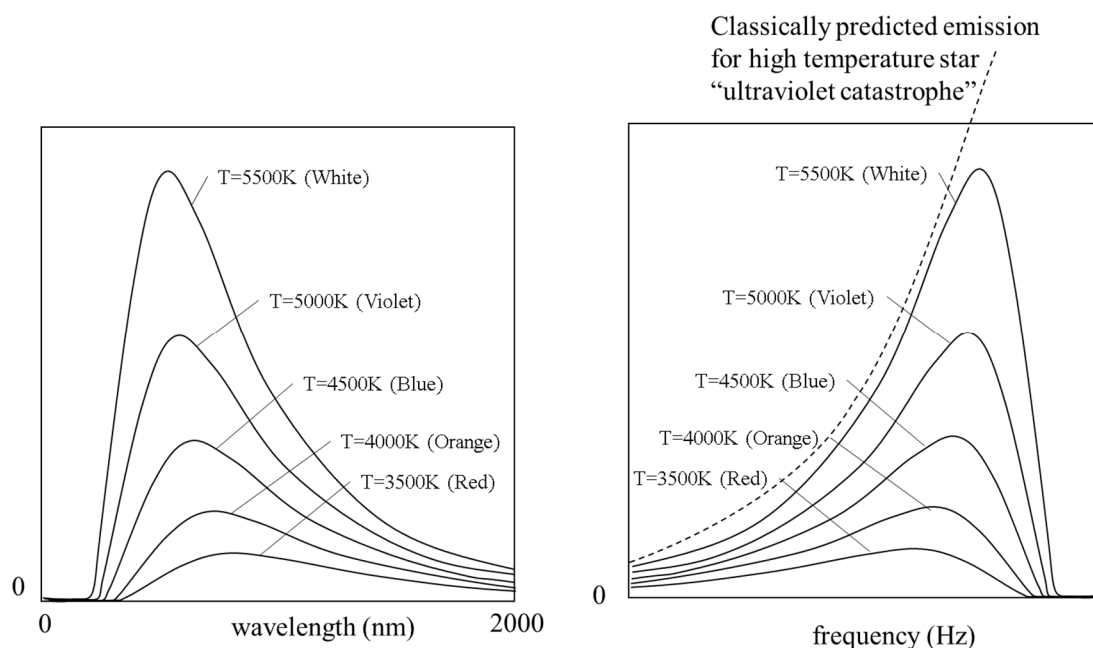
Measurements were made of the intensity of the light emitted at different frequencies (or wavelengths) by such objects. In addition measurements were made at different temperatures. In order to improve the experiment and avoid any reflections of the radiation, a cavity was used with a small hole, which emits the radiation: this is described as a black body.

A surface that absorbs all wavelengths of electromagnetic radiation is also the best emitter of electromagnetic radiation at any wavelength. Such an ideal surface is called a black body. The continuous spectrum of radiation it emits is called black-body radiation.

It was found that the amount of black-body radiation emitted at any frequency depends only on the temperature, not the actual material.

Specific intensity is a measure of the radiation emitted by a body. Irradiance is a measure of the radiation received by a surface.

Graphs of specific intensity against wavelength (or frequency) are shown. As the temperature increases, each maximum shifts towards the higher frequency (shorter wavelength) and hence becomes “bluer”.



Graphs of specific intensity against wavelength and against frequency.

Attempts to theoretically obtain the observed black-body graph failed when using classical mechanics. Wilhelm Wien obtained an equation that ‘fitted’ observations at high frequencies (low wavelengths). Later Lord Rayleigh obtained an equation that ‘fitted’ at low frequencies but tended off to infinity at high frequencies (see line on the above frequency graph). This divergence was called the ultraviolet catastrophe and puzzled many leading scientists of the day.

In 1900 Planck looked at the two equations and produced a ‘combined’ relationship, which gave excellent agreement with the experimental curve. However, initially this relationship could not be derived from first principles. It was a good mathematical ‘fudge’!

Planck studied his relationship and the theory involved and noticed that he could resolve the problem by making the assumption that the absorption and emission of radiation by the oscillators could only take place in ‘jumps’ given by:

$$E = nhf$$

where E is energy, f is frequency, h is a constant and $n = 0, 1, 2, 3 \dots$

Using this assumption he could derive his equation from first principles. The constant of proportionality h was termed Planck’s constant. The integer values n lead to the fact that the energy values are quantised, hence the introduction of the word quantum (plural for quanta, comes from the Latin ‘quantus’, meaning ‘how much’.)

It must be emphasised that Planck did this in a mathematical way with no justification as to why the energy should be quantised – but it worked! To Planck the oscillators were purely theoretical and radiation was not actually emitted in ‘bundles’, it was just a ‘calculation convenience’. It was some years before Planck accepted that radiation could be treated as energy packets.

The Photoelectric Effect

In 1887 Heinrich Hertz observed that a spark passed between two plates more often if the plates were illuminated with ultraviolet light. Later experiments by Wilhelm Hallwachs and Philipp Lenard gave the unexpected results we are familiar with, namely:

- (a) the non-emission of electrons with very bright but low frequency radiation on a metal surface, e.g. very bright red light
- (b) the increase in the speed of the emitted electron with frequency but not with intensity. Increasing the intensity only produced more emitted electrons.

The observations of these experiments are known as the Photoelectric Effect. These results were unexpected because energy should have been able to be absorbed continuously from a wave. An increase in the intensity of a wave also means an increase in amplitude and hence a larger energy.

In 1905 Einstein published a paper on the photoelectric effect entitled “On a Heuristic Viewpoint Concerning the Production and Transformation of Light”. He received the Nobel Prize for Physics for this work in 1921. The puzzle was why energy is not absorbed continuously from such a wave, e.g. any electromagnetic radiation, in a cumulative manner. It should just take more time for energy to be absorbed and for the electron to be emitted but this is not what is observed. Einstein proposed that electromagnetic radiation is emitted and absorbed in small packets. (The word ‘photon’ was unintentionally introduced by the physical chemist Gilbert Lewis in 1926). The energy of each photon is given by:

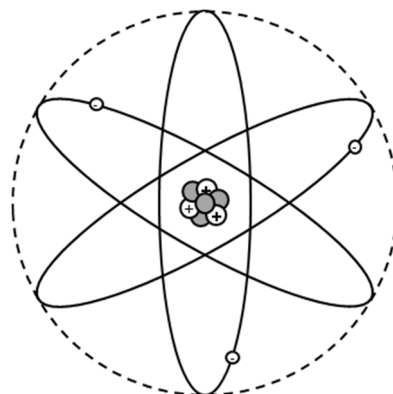
$$E = hf$$

where E is the energy of a photon of radiation of frequency f .

This proposal also explained why the number of electrons emitted depended on the irradiance (i.e. number of photons) of the electromagnetic radiation and why the velocity of the emitted electrons depended on the frequency. It did not explain the ‘packets’ or why they should have this physical ‘reality’.

Models of the Atom

Rutherford’s scattering experiment indicated that the majority of the mass of the atom was in a small nucleus, with the electrons ‘somewhere’ in the atomic space. He and his assistants could not ‘see’ the electrons. A picturesque model of the atom, similar to a small solar system, came into fashion. This model had many features to commend it. Using classical mechanics, an electron in an orbit could stay in that orbit, the central force being provided by electrostatic attraction. However, the electron has a negative charge and given that it has a (central) acceleration it should emit radiation, lose energy and spiral into the nucleus.



The current theory was insufficient. Why do the electrons ‘remain in orbit’? Do they in fact ‘orbit’?

In the late nineteenth century attempts were made to introduce some 'order' to the specific frequencies emitted by atoms. In 1885 Johann Balmer found, by trial and error, a simple formula for a group of visible lines in the hydrogen spectra.

$$\frac{1}{\lambda} = R \left(\frac{1}{2^2} - \frac{1}{n^2} \right)$$

where λ is the wavelength, R the Rydberg constant, n is an integer 3, 4, 5...

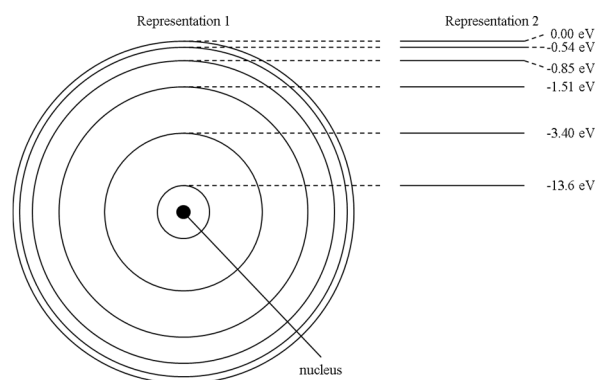
Other series were then discovered:

Lyman with the first fraction being $1/1^2$ (ultraviolet)

Paschen with the first fraction as $1/3^2$ (infrared)

However, this only worked for hydrogen and atoms with one electron, e.g. ionised helium, and moreover did not provide any theoretical reason why the formula should work.

In 1913 Niels Bohr introduced the idea of electron energy levels. Each atom has some internal energy due to its structure and internal motion but this energy cannot change by any variable amount, only by specific discrete amounts. Any particular atom, e.g. an atom of gold say, has a specific set of energy levels. Different elements each have their own set of levels. Experimental evidence of the day provided agreement with this idea and energy level values were obtained from experimental results. Electron transitions between energy levels give the characteristic line spectra for elements.



In order to solve the problem that an electron moving in a circular orbit should continuously emit radiation and spiral into the nucleus, Bohr postulated that an electron can circulate in certain permitted, stable orbits without emitting radiation. He made the assumption that the normal electromagnetic phenomena did not apply at the atomic scale! Furthermore he made an intuitive guess that angular momentum is quantised. It is said that he noticed that the units of Planck's constant (J s^{-1}) are the same as those of angular momentum ($\text{kg m}^2 \text{s}^{-1}$). The allowed orbit, of radius r , of an electron must have angular momentum of an integer multiple of $h/2\pi$:

$$mvr = \frac{nh}{2\pi}$$

where n is an integer 1, 2, 3...

The angular momentum of a particle, of mass m , moving with tangential speed v and circle of radius r is mvr .

Thus for any specific orbit, n , we can calculate the radius of that orbit given the tangential speed or vice versa.

Theoretical aside

For the hydrogen atom with a single electron, mass m_e revolving around a proton (more correctly around the centre of mass of the system), we can assume the proton is stationary since it is ~2000 times more massive. Hence, equating the electrostatic force and centripetal force $\frac{m_e v^2}{r}$:

$$\frac{1}{4\pi\epsilon_0} \frac{q_e^2}{r_n^2} = \frac{m_e v_n^2}{r_n}$$

where ϵ_0 is the permittivity of free space

q_e is the charge of the electron or proton (equal magnitude)

r is the radius of orbit

v is the tangential velocity

for the n th orbit.

Combining this with the previous equation, and solving simultaneously, we find:

$$r_n = \epsilon_0 \frac{n^2 h^2}{\pi m_e q_e^2} \quad \text{and} \quad v_n = \frac{1}{\epsilon_0} \frac{q_e^2}{2 \pi h}$$

for the n th orbit.

Calculating r_1 , the radius of the first Bohr orbit uses data given in assessments, namely h , m_e , q_e and ϵ_0 gives $r_1 = 5.3 \times 10^{-11} \text{ m}$.

These equations gave the values of the radii for the non-radiating orbits for hydrogen and the value of n was called the quantum number of that orbit.

For the hydrogen atom, calculating the velocity of an electron in the first Bohr orbit of radius $5.3 \times 10^{-11} \text{ m}$, using

$$mvr = \frac{nh}{2\pi}$$

with $n = 1$ we obtain $v = 2.2 \times 10^6 \text{ m s}^{-1}$.

Note: Bohr's theory only applies to an atom with one electron, e.g. the hydrogen atom or ionised helium atom.

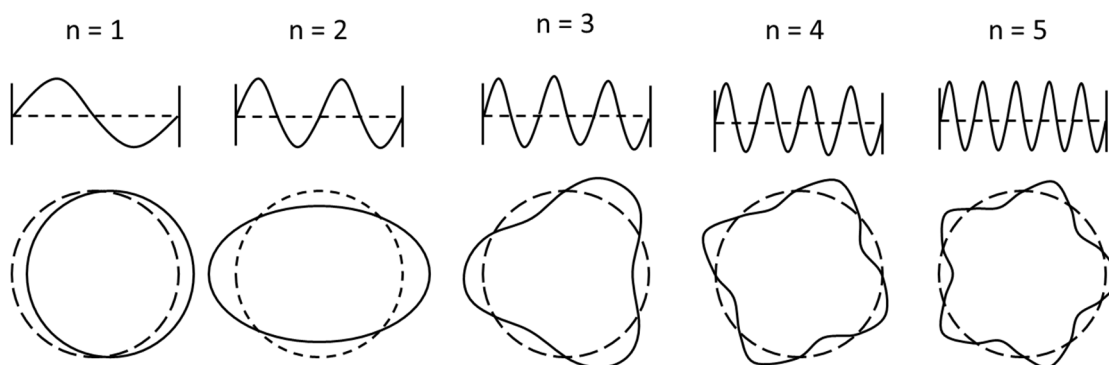
However, as mentioned above, the idea of energy levels can be extended to all atoms, not just hydrogen.

So this theory is not complete since it did not allow any prediction of energy levels for any specific element nor did it explain why angular momentum should be quantised or why electrons in these orbits did not radiate electromagnetic energy! Another question was, 'what happens during a transition?'

Wave-Particle Duality

Electrons as “Standing Waves”

An extension by Louis de Broglie suggested that electron orbits are standing waves. The electron, now behaving like a wave, forms a standing wave of an integer number of wavelengths that ‘fits’ into the circumference of an orbit. This is represented here:



The dashed lines represent the circumference, while the solid line represents the hypothesised wave-like path of the electron.

This is a picturesque idea but somewhat old-fashioned since an electron does not take a ‘wiggly’ path around the nucleus. This is an example of a model that should not be taken too seriously and could lead to poor understanding. Quantum mechanics goes on to show that we cannot describe the motion of an electron in an atom in this way.

De Broglie Wavelength

We use the word ‘particle’ to describe localised phenomena that transport mass and energy, and the word ‘wave’ to describe delocalised (spread out) phenomena that carry energy but no mass.

Experimental observations seem to suggest that both electromagnetic radiation and electrons can behave like particles and like waves. They exhibit both the properties of wave phenomena, such as interference and diffraction, and particle phenomena, for example photons causing electron emission in the photoelectric effect or electron ‘billiard ball type’ collisions.

So an electron can show wave-like phenomena. In the mid-1920s G P Thomson, in Aberdeen, bombarded a thin metal with an electron beam and obtained a diffraction ring. In 1927 Davisson and Germer directed a beam of electrons onto the surface of a nickel crystal and observed the reflected beam. They had expected to see diffuse reflection since even this smooth surface would look ‘rough’ to the tiny electrons. To their surprise, they observed a similar pattern to X-ray diffraction from a surface. Thomson and Davisson were awarded the Nobel Prize in 1937 for demonstrating the wave-like properties of electrons. (Thomson’s father, J J Thomson, won the Nobel Prize in 1906 for discovering the electron as a particle.)

In 1923 de Broglie suggested that since light had particle-like properties, perhaps the nature was dualistic and particles had wave-like properties.

From relativity theory, the energy of a particle with zero rest mass, e.g. a photon, is given by $E = pc$ and we know that $E = hf$, hence $p = h/\lambda$.

The apparent wave like behaviour of a particle is related to its momentum.

For a particle $p = mv$ and for a wave $p = h/\lambda$, giving a relationship $mv = h/\lambda$ or:

$$\lambda = \frac{h}{p}$$

p is the momentum

h is the Planck constant

We can calculate the de Broglie wavelength of a particle of velocity v .

Example problems

1. A neutron and an electron have the same speed.
Which has the longer de Broglie wavelength?

The electron, since the neutron has the larger mass. (The mass is in the denominator.)

2. An electron microscope uses electrons with a de Broglie wavelength of 0.04 nm.

What is the required speed of the electrons?

$$\lambda = \frac{h}{mv}$$

$$0.04 \times 10^{-9} = \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times v}$$

$$v = \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 0.04 \times 10^{-9}}$$

$$v = 1.8 \times 10^7 \text{ m s}^{-1}$$

On the very small scale it is only possible to observe details which are of the order of the light that is incident upon it. This means that to view the very small we need to use a wavelength of equivalent value. Notice that this wavelength of 0.04 nm is very much smaller than that of blue light (400 nm). Hence the use of electrons in microscopes can improve the resolution of the image compared to optical microscopes.

Quantum Mechanics

“The more you see how strangely Nature behaves, the harder it is to make a model that explains how even the simplest phenomena actually works.”

Richard Feynman

Matter was thought to be ‘atomistic’ with particles making basic interactions and the properties of these particles continually changing smoothly from place to place. Waves moved continuously from place to place.

Classical mechanics could not explain the various quantisation rules, those which attempted to give some limited agreement between observation and theory. The apparent dual wave-particle nature of matter could not be explained. With the introduction of quantum theory classical ideas would need to be revised.

There are various forms of quantum mechanics: Heisenberg’s matrix mechanics, Erwin Schrödinger’s wave mechanics, Dirac’s relativistic field theory and Feynman’s sum over histories or amplitude mechanics.

At this stage you don’t need to be able to work with any of these, however, you are expected to realise that in essence quantum mechanics provides us with the means to calculate probabilities for physical quantities. Exact physical quantities, e.g. position or velocity, do not have unique values at each and every instant.

“Balls in quantum mechanics do not behave like balls in classical mechanics... an electron between release and detection does not have a definite value for its position. This does not mean that the electron has a definite position and we don't know it. It means the electron just does not have a position just as love does not have a colour.”

Strange World of Quantum Mechanics, D Styer

Quantum theories incorporate the following concepts:

- (i) Transitions between stationary states are **discrete**. There is no meaning to any comment on a system in an intermediate state.

Depending on the experiment, particles may behave as waves or a wave as a particle. However, in a certain respect they act like both together. It is just not a sensible question in quantum mechanics to ask if matter is a wave or a particle.

- (ii) Every physical situation can be characterised by a wave function (or other mathematical formalism). This wave function is not directly related to any actual property of the system but is a description of the potentialities or possibilities within that situation. The wave function provides a statistical ensemble of similar observations carried out under the specified conditions. It does not give the detail of what will happen in any particular individual observation. The probability of a specific observation is obtained from the square of the wave function. This is an important and non-intuitive idea.

The quantum probability aspect is very different from classical physics, where we consider there is an actual state and any 'probability' comes from our inadequate measuring or statistical average.

In the quantum domain we can only calculate probabilities. For example, we cannot state when a particular nucleus will decay (although we can measure a half-life, based on group properties as opposed to individual nuclei) but we could calculate the probability of a particular nucleus decaying within a given time period. This is typical of the rules of quantum mechanics – the ability to calculate probabilities but not make specific, definite, predictions.

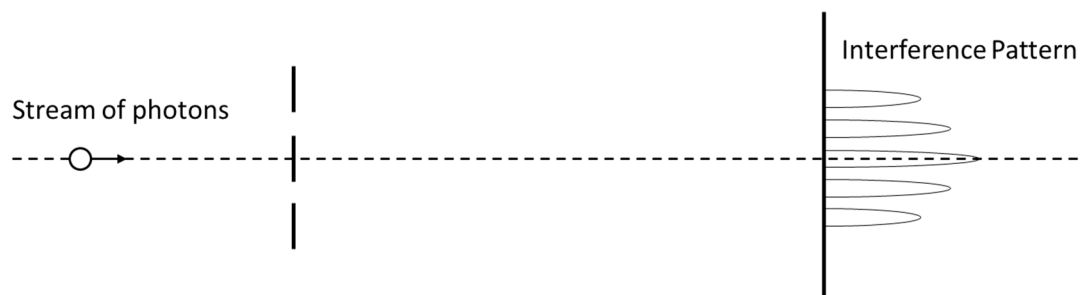
Quantum mechanics has enjoyed unprecedented practical success. Theoretical calculations agree with experimental observations to very high precision.

Quantum mechanics also reminds us that there is discreteness in nature and there are only probabilities.

Double-slit experiment

The double-slit experiment with light shows an interference pattern. This is a standard experiment to demonstrate that light demonstrates wave-like properties.

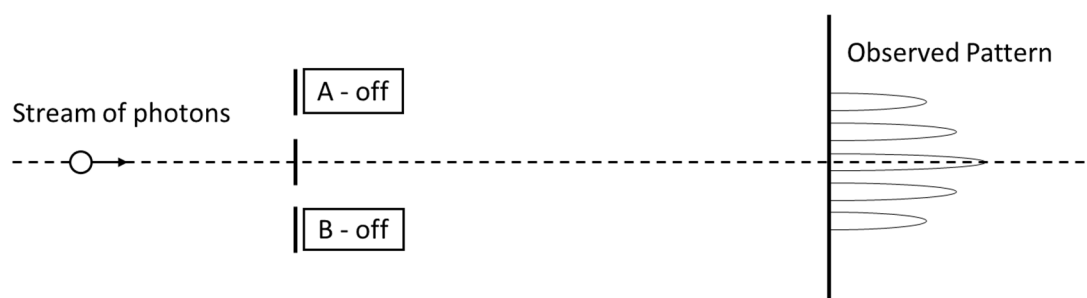
There is a central maximum on the central axis opposite the two slits as shown below.



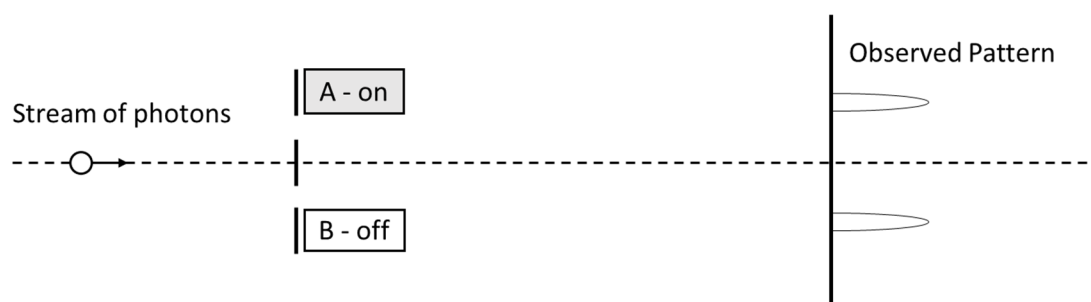
In more recent years this experiment has been performed with single photons (suggesting a particle-like behaviour) and a detector screen. Each photon reaches the screen and the usual interference pattern is gradually built up.

The question is: “how does a single photon ‘know about’ the slit it does not pass through?”

Let us place a detector near each slit as shown below. In this diagram the detectors are switched off and not making any measurements. The pattern seen remains the same.



Let us now switch on detector A.



We lose the interference effect and simply obtain a pattern for particles passing through two slits. We would get the same pattern if we switched on detector B instead of detector A or if we switched on both detectors.

It seems that if we ask the question ‘Where is the photon?’ or ‘Which slit does the photon pass through?’ and set up an experiment to make a measurement to answer that question, e.g. determine which slit the photon passed through, we do observe a ‘particle’ with a position but lose the interference effect, we lose the wave-like behaviour.

It appears that the single photon in some way does ‘know about’ both slits. This is one of the non-intuitive aspects of quantum mechanics, which now begins to suggest that a single particle must pass through both slits.

A very similar double-slit experiment can be performed with electrons. Again we can arrange for only one electron to ‘pass through’ the slits at any one time. The position of the electrons hitting the ‘screen’ agrees with our familiar interference pattern (demonstrating wavelike behaviour). However, as soon as we attempt to find out which slit the electron passes through we lose the interference effect. For electron interference the spacing of the ‘slits’ must be small. Planes of atoms in a crystal can be used to form the slits since electrons have a very small associated wavelength, the de Broglie wavelength.

These observations are in agreement with quantum mechanics. We cannot measure the wave and particle properties at the same time!

Conclusions

What we find is that we cannot predict the exact path that either an electron or photon will follow, and therefore predict where it lands on the screen, in this experiment.

We also find that we cannot make a direct measurement of the particles locations, as this will act in itself will interfere with the outcome of the experiment!

What we can do is say what probability the electron or photon has of landing at a particular point on the screen. The maxima are areas where there is a high probability of arriving, hence many electrons/photons arrive here and we see a bright patch. The minima represent areas of low probability, fewer electrons/photons arrive and it is dimmer/dark. The sum of all of these probabilities is 1. This ties in with the idea of Feynman’s sum over histories.

The Uncertainty Principle

A Theoretical Introduction

Using the wave theory of quantum mechanics we can produce a wave function describing the ‘state’ of the system, e.g. an electron. However, we find that it is not possible to determine with accuracy all the observables for the system. For example, we can compute the likelihood of finding an electron at a certain position, e.g. in a box. The position wave function may then be effectively zero everywhere else and the uncertainty in its position may be very small inside the box. If we then consider its momentum wave function we discover that this is very spread out, and there is nothing we can do about it. This implies that in principle, if we ‘know’ the position, the momentum has a very large uncertainty. If we reverse the situation we find that the opposite is also true, if we are more certain about the momentum of the particle we are less certain about its position!

Consider a wave with a single frequency. Its position can be thought of as anywhere along the wave but its frequency is uniquely specified.



Now consider a wave composed of a mixture of slightly different frequencies, which when added together produces a small ‘wave packet’. The position of this wave can be quite specific but its frequency is conversely non-unique.



Heisenberg’s Uncertainty Principle should more appropriately be called Heisenberg’s Indeterminacy Principle since we can measure either x or p_x with very low ‘uncertainty’ but we cannot measure both. If one is certain, the other is indeterminate.

Theoretical considerations also shows that the energy E and time t have this dual indeterminacy.

A thought experiment to illustrate Heisenberg’s Uncertainty Principle

In classical physics it is assumed that all the attributes, such as position, momentum, energy etc, could be measured with a precision limited only by the experiment.

In the atomic domain is this still true?

Let us consider an accurate method to determine the position of an electron in a particular direction, for example in the x direction. The simplest method is to use a ‘light gate’, namely to allow a beam of electromagnetic radiation to hit the electron and be interrupted in its path to a detector. To increase the accuracy we can use radiation of a small wavelength, e.g. gamma rays. However, we note that by hitting the electron with the gamma rays the velocity of the electron will alter (a photon-electron collision). Now the velocity or momentum of the electron in the x direction will have changed. Whatever experiment we use to subsequently measure the velocity or momentum cannot determine the velocity before the electron was ‘hit’. To reduce the effect of the ‘hit’ we can decrease the frequency of the radiation, and lose some of the precision in the electron’s position. We just cannot ‘win’!

Heisenberg's Uncertainty Principle is stated as

$$\Delta x \Delta p_x \geq \frac{h}{4\pi}$$

where Δx is the uncertainty in the position, Δp_x is the uncertainty in the component of the momentum in the x direction and h is Planck's constant.

Quantum mechanics also shows that there are other pairs of quantities that have this indeterminacy, for example energy and time:

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

where ΔE is the uncertainty in energy and Δt is the uncertainty in time.

We notice that the pairs of quantities in these relationships, termed “conjugate variables”, have units that are the same as those of h , namely J s. For energy and time this is obvious. For position and momentum we have:

$$\text{m kg m s}^{-1} \rightarrow (\text{multiply by s and s}^{-1}) \rightarrow (\text{kg m}^2\text{s}^{-2}) \text{ s} \rightarrow \text{J s}$$

The question ‘Does the electron have a position and momentum before we look for it?’ can be debated. Physicists do not have a definitive answer and it depends on the interpretation of quantum mechanics that one adopts. However, this is not a useful question since it is the wave functions that give us our information and there is a limit on what we can predict about the quantum state. We just have to accept this. It is worth reiterating that quantum mechanics does gives superb agreement with experimental observations.

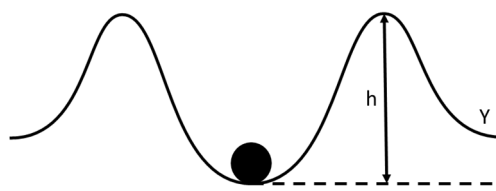
Using quantum mechanics the spectral lines for helium and other elements can be calculated and give excellent agreement with experimental observations.

More importantly, quantum mechanics provides a justification for the previously ad hoc quantisation ‘rules’ introduced earlier and gives us a very useful tool to explain observed phenomena through theory and make quantitative and accurate predictions about the outcomes of experiments.

Potential Wells and Quantum Tunnelling

Imagine a ball in a 'dip'. The exact shapes of the ups and downs is not particularly relevant.

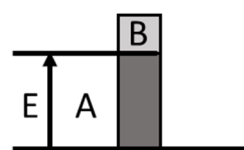
The ball cannot escape the well and get to position Y unless it receives energy $E = mgh$.



The ball can be said to be in a 'potential well' of 'height' mgh . This means that the ball needs energy E equal to or greater than mgh in order to 'escape' and get to position Y.

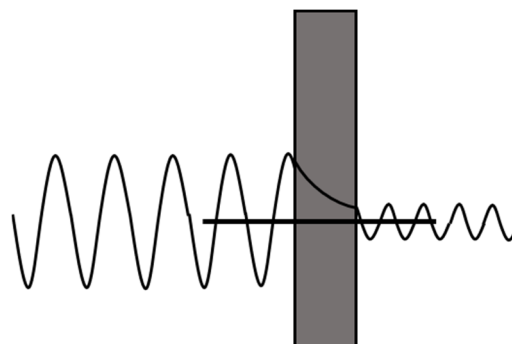
In the quantum world things are a little different, although the concept of a 'potential well' or 'potential barrier' is useful.

Now let us consider an electron with some energy E on the left-hand side of a potential energy barrier of magnitude greater than E . The electron is thus confined to side A. It does not have enough energy to 'get over' the barrier and 'escape'.



Quantum theory disagrees with this classical prediction.

The wave function for the electron is continuous across a barrier. The amplitude is greater in region A but it is still finite, although much smaller, outside region A (to the right of the barrier). The probability of finding the electron in region A is very high, but there is in fact a finite probability of finding the electron beyond the barrier.



The probability depends on the square of the amplitude. Hence it appears that the electron can 'tunnel out'. This is called quantum tunnelling and has some interesting consequences and possible applications.

Examples of Quantum Tunnelling

Alpha Decay

For some radioactive elements, e.g. polonium 212, the alpha particles are held in the nucleus by the residual strong force and do not have enough energy to escape. However, the quantum tunnelling effect means that they can escape and quantum mechanics can be used to calculate the half-life. In 1928 George Gamow used quantum mechanics (namely Schrödinger's wave equation) and the idea of quantum tunnelling to obtain a relationship between the half-life of the alpha decay and the energy of emission. Classically the alpha particle should not be able to escape.

Scanning Tunnelling Microscope

A particular type of electron microscope, the scanning tunnelling microscope, has a small stylus that scans the surface of the specimen. The distance of the stylus from the surface is only about the diameter of an atom. Electrons ‘tunnel’ across the sample. In this way the profile of the sample can be determined. Heinrich Rohrer and Gerd Binnig were awarded the Nobel Prize for their work in this field in 1986.

Virtual Particles

Another interesting effect of the Uncertainty Principle is the ‘sea’ of virtual particles in a vacuum. We might expect a vacuum to be ‘empty’. Not so with quantum theory.

A particle can ‘appear’ with an energy ΔE for a time less than Δt where

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

Do these virtual particles ‘exist’? This is not really a sensible question for quantum mechanics. We cannot observe them in the short time of their existence. However, they are important as ‘intermediate’ particles in nuclear decays and high energy particle collisions and if they are omitted theoretical agreement with practical observations may not be obtained. Virtual particles are important when using Feynman diagrams to solve problems.

Particles from Space

Cosmic Rays

In the early 1900s, radiation was detected using an electroscope. However, radiation was still detected in the absence of known sources. This was described as background radiation, which exists due to the sources that surround us in everyday life.

Austrian physicist Victor Hess made measurements of radiation at high altitudes from a balloon, to try and get away from possible sources on Earth. He was surprised to find the measurements actually increased with altitude. At an altitude of 5000 m the intensity of radiation was found to be five times that at ground level.

Hess named this phenomenon cosmic radiation (later to be known as cosmic rays).

It was thought this radiation was coming from the Sun, but Hess obtained the same results after repeating his experiments during a nearly complete solar eclipse (12 April 1912), thus ruling out the Sun as the (main) source of this cosmic radiation. In 1936 Hess was awarded the Noble Prize for Physics for the discovery of cosmic rays.

Tracks produced by cosmic rays can be observed using a cloud chamber. Charles T R Wilson is the only Scot ever to be awarded the Nobel Prize for Physics. He was awarded it in 1927 for the invention of the cloud chamber.

Robert Millikan coined the phrase ‘cosmic rays’, believing them to be electromagnetic in nature.

By measuring the intensity of cosmic rays at different latitudes (they were found to be more intense in Panama than in California), Compton showed that they were being deflected by the Earth’s magnetic field and so must consist of electrically charged particles, i.e. electrons or protons rather than photons in the form of gamma radiation.

Origin and composition

The term cosmic ray is not precisely defined, but a generally accepted description is ‘high energy particles arriving at the Earth which have originated elsewhere’.

Composition

Cosmic rays come in a whole variety of types, but the most common are protons, followed by helium nuclei. There is also a range of other nuclei as well as individual electrons and gamma radiation.

Nature	Approximate % of all cosmic rays
Protons	89
Alpha particles	9
Carbon, nitrogen and oxygen nuclei	1
Electrons	less than 1
Gamma radiation	less than 0.1

The energies of cosmic rays cover an enormous range, with the most energetic having energies much greater than those capable of being produced in current particle accelerators.

The highest energies produced in modern particle accelerators, including the Large Hadron Collider, are of the order of 1 tera electronvolt (10^{12} eV). Cosmic rays have been observed with energies ranging from 10^9 to 10^{20} eV. Those with energies above 10^{18} eV are referred to as Ultra High Energy Cosmic Rays (UHECRs).

The 'Oh-my-God' (OMG) particle with energy of 3×10^{20} eV was recorded in Utah in 1991.

Converting to joules (J), 3×10^{20} eV = $3 \times 10^{20} \times 1.6 \times 10^{-19}$ J = 48 J, i.e. ~ 50 J.

That is enough energy to throw a 2.5 kg mass (e.g. a bag of potatoes) 2 m vertically upwards. It is also approximately equal to the kinetic energy of a tennis ball served at about 90 mph (40 m s^{-1}).

Order of magnitude open-ended question opportunity here:

mass = 60 g = 0.06 kg

speed = 40 m s^{-1} ,

kinetic energy = $0.5 \times 0.06 \times 40 \times 40 = 48 \text{ J}$

The OMG particle was probably a proton and as such had about 40 million times the energy of the most energetic protons ever produced in an Earth-based particle accelerator.

Such UHECRs are thought to originate from fairly local (in cosmological terms) distances, i.e. within a few hundred million light years. Were they to originate from further away it would be hard to understand how they get all the way here at all, since the chances are they would have interacted with Cosmic Microwave Background Radiation (CMBR) photons along the way, producing pions.

Origin

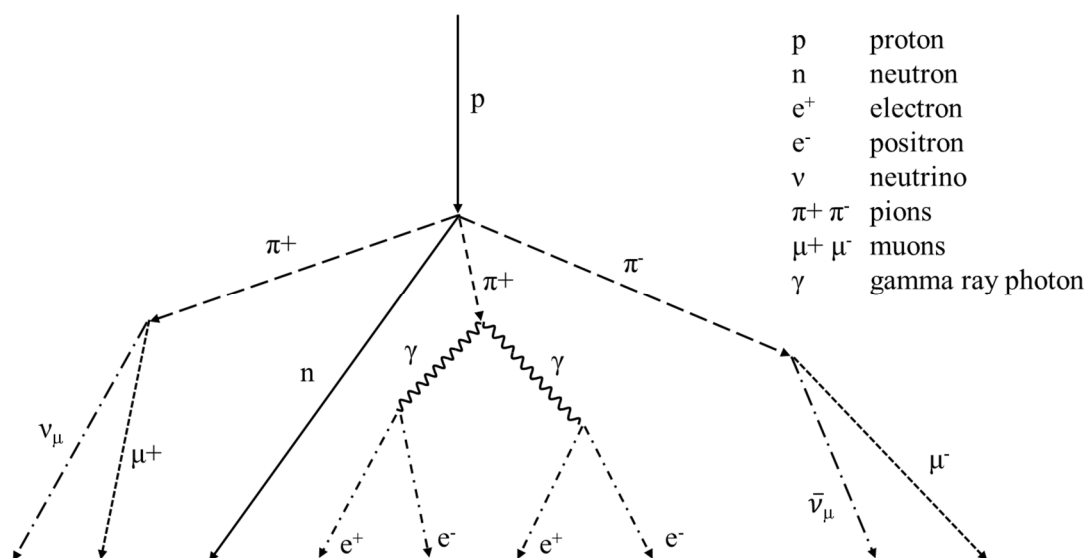
The lowest energy cosmic rays come from the Sun and the intermediate energy ones are presumed to be created elsewhere within the Milky Way, often in connection with supernovae. The main astrophysics (rather than particle physics) to come from the study of cosmic rays concerns supernovae since they are believed to be the main source of cosmic rays.

However, the origin of the highest energy cosmic rays is still uncertain. Active galactic nuclei (AGN) are thought to be the most likely origin for UHECRs. The compact galactic centres are associated with a much greater than expected luminosity over a large proportion of the electromagnetic spectrum. This is thought to be due to the accretion of matter by a supermassive black hole at the centre of that galaxy. A group of cosmologists (including Martin Hendry from the University of Glasgow) are working on the statistical analysis of apparent associations between the incoming direction of the highest energy cosmic rays and active galaxies.

Interaction with the Earth's Atmosphere

When cosmic rays reach the Earth, they interact with particles in the Earth's atmosphere, producing a chain of reactions resulting in the production of a large number of particles known as a cosmic air shower shown below. Air showers were first discovered by the French scientist Pierre Auger in 1938. Analysing these showers allows the initial composition and energies of the original (primary) cosmic rays to be deduced.

When cosmic rays from space (the primary cosmic rays) strike particles in the atmosphere they produce secondary particles, which go on to produce more collisions and particles, resulting in a shower of particles that is detected at ground level. The primary cosmic rays can usually only be detected directly in space, for example by detectors on satellites, although very high energy cosmic rays, which occur on rare occasions, can penetrate directly to ground level without making interactions.



(It may be worthwhile and even fun to investigate the interactions taking place here by referring to the Particle Physics section of the Higher Physics course. Can you justify to yourself that each interaction is possible, and even why the muon anti-neutrino makes an appearance in one instance but not in another? The primary interaction involves an initial proton leading to 3 pions and a neutron, in terms of quarks can you justify this?)

Detection

Consequently there are two forms of detector: those that detect the air showers at ground level and those located above the atmosphere that detect primary cosmic rays.

Cherenkov radiation

Air shower particles can travel at relativistic speeds. Although relativity requires that nothing can travel faster than the speed of light in a vacuum, particles may exceed the speed of light in a particular medium, for example water. Such particles then emit a beam of Cherenkov radiation – the radiation that causes the characteristic blue colour in nuclear reactors. (This is a bit like the optical equivalent of a sonic boom.)

Atmospheric fluorescence

When charged particles pass close to atoms in the atmosphere, they may temporarily excite electrons to higher energy levels. The photons emitted when the electrons return to their previous energy levels can then be detected.

The Pierre Auger observatory in Argentina was set up to study high-energy cosmic rays. It began operating in 2003 and at that time was the largest physics experiment in the world. It is spread over several thousand square miles and uses two basic types of detectors:

- 1600 water tanks to detect the Cherenkov effect
- four detectors of atmospheric fluorescence.

The Solar Wind and Magnetosphere

Structure of the Sun

The interior of the Sun consists of three main regions:

1. the core, within which nuclear fusion takes place
2. the radiative zone, through which energy is transported by photons
3. the convective zone, where energy is transported by convection.

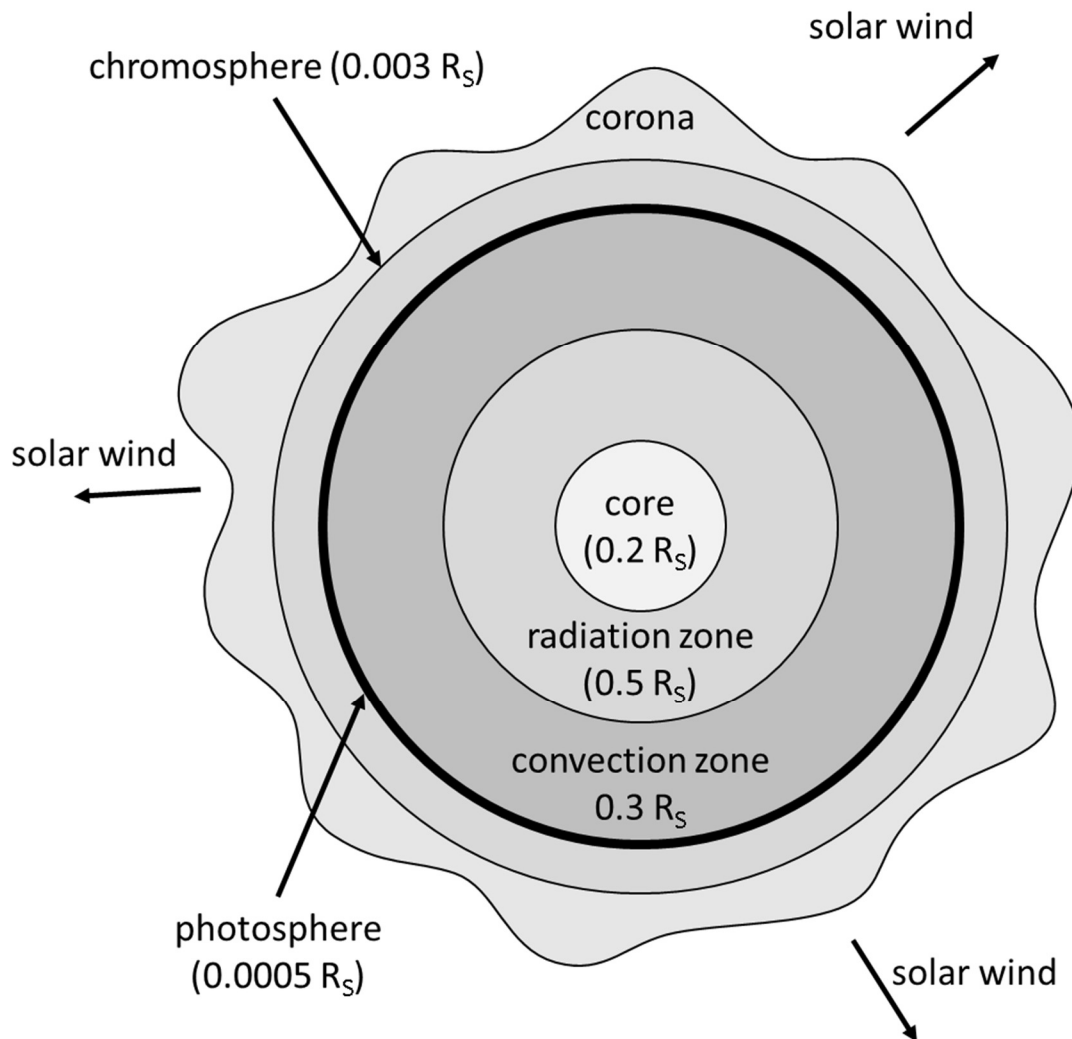
The extended and complex solar atmosphere begins at the top of the convective zone, with the photosphere.

The photosphere is the visible surface of the Sun and appears smooth and featureless, marked by occasional relatively dark spots, called sunspots.

Moving outwards, next is the chromosphere. Sharp spicules and prominences emerge from the top of the chromosphere.

The corona (from the Greek for crown) extends from the top of the chromosphere. The corona is not visible from Earth during the day because of the glare of scattered light from the brilliant photosphere, but its outermost parts are visible during a total solar eclipse.

The depth of each layer relative to the radius of the Sun (R_s) is shown. The photosphere is about 330 km deep ($0.0005R_s$) and the chromosphere is about 2000 km ($0.003R_s$) deep.



Coronagraphs are special telescopes that block out the light from the photosphere to allow the corona to be studied. These are generally used from mountain tops (where the air is thin) and from satellites. They have been able to detect the corona out beyond $20R_S$, which is more than 10% of the way to Earth.

The corona is permeated by magnetic fields. In particular there are visible loops along which glowing ionised gaseous material can be seen to travel. They have the shape of magnetic field lines and begin and end on the photosphere.

Information about magnetic fields in the corona has come from the study of emitted X-rays, obtained from satellites and space stations. The corona is the source of most of the Sun's X-rays because of its high temperature, which means it radiates strongly at X-ray wavelengths. The corona's X-ray emission is not even, however, with bright patches and dark patches. The dark areas hardly emit any X-rays at all and are called coronal holes.

The Solar and Heliospheric Observatory (SOHO) is a project of international collaboration between the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA) to study the Sun from its deep core to the outer corona and the solar wind.

In 2006 a rocket was launched from Cape Canaveral carrying two nearly identical spacecraft. Each satellite was one half of a mission called Solar TERrestrial RELations Observatory (STEREO) and they were destined to do something never done before – observe the whole of the Sun simultaneously. With this new pair of viewpoints, scientists are able to see the structure and evolution of solar storms as they blast from the Sun and move out through space.

The Solar Wind

There is a continual flow of charged particles emanating from the Sun because of the high temperature of the corona. This gives particles sufficient kinetic energy to escape from the Sun's gravitational influence. This flow is called the solar wind and is a plasma composed of approximately equal numbers of protons and electrons (i.e. ionised hydrogen). It can be thought of as an extension of the corona itself and as such reflects its composition. The solar wind also contains about 8% alpha particles (i.e. helium nuclei) and trace amounts of heavy ions and nuclei (C, N, O, Ne, Mg, Si, S and Fe).

The solar wind travels at speeds of between 300 and 800 km s^{-1} , with gusts recorded as high as $1,000 \text{ km s}^{-1}$ (2.2 million miles per hour).

Comet Tails

Although it was known that solar eruptions ejected material that could reach the Earth, no-one suspected that the Sun was continually losing material regardless of its apparent activity. It had been known for a long time that comet tails always pointed away from the Sun, although the reason was unknown. Ludwig Biermann (of the Max Planck Institute for Physics in Göttingen) made a close study of the comet Whipple-Fetke, which appeared in 1942. It had been noted that comet tails did not point directly away from the Sun. Biermann realised this could be explained if the comet was moving in a flow of gas streaming away from the Sun. The comet's tail was acting like a wind-sock. In the early 1950s Biermann concluded that even when the Sun was quiet, with no eruptions or sunspots, there was still a continuous flow of gas from it.

In 1959, the Russian space probe Luna 1 made the first direct observation and measurement of the solar wind. The probe carried different sets of scientific devices for studying interplanetary space, including a magnetometer, Geiger counter, scintillation counter and micrometeorite detector. It was the first man-made object to reach escape velocity from Earth, meaning it achieved a heliocentric orbit (orbit of the Sun rather than the Earth!).

Coronal Holes

Unlike most magnetic field lines in the corona, the magnetic field lines from coronal holes don't loop back onto the surface of the corona. Instead they project out into space like broken rubber bands, allowing charged particles to spiral along them and escape from the Sun. There is a marked increase in the solar wind when a coronal hole faces the Earth.

Solar Flares

Solar flares are explosive releases of energy that radiate over virtually the entire electromagnetic spectrum, from gamma rays to long wavelength radio waves. They also emit high-energy particles called solar cosmic rays. These are composed of protons, electrons and atomic nuclei that have been accelerated to high energies in the flares. Protons (hydrogen nuclei) are the most abundant particles followed by alpha particles (helium nuclei). The electrons lose much of their energy in exciting radio bursts in the corona. Solar flares generally occur near sunspots, which leads to the suggestion they are magnetic phenomena. It is thought that magnetic field lines become so distorted and twisted that they suddenly snap like rubber bands. This releases a huge amount of energy, which can heat nearby plasma to 100 million kelvin in a few minutes or hours. This generates X-rays and can accelerate some charged particles to very high, even relativistic, velocities.

The energies of solar cosmic ray particles range from milli electronvolts (10^{-3} eV) to tens of giga electronvolts (10^{10} eV).

The highest energy particles arrive at the Earth within half an hour of the flare maximum, followed by the peak number of particles, approximately 1 hour later.

Particles streaming from the Sun due to solar flares or other major solar events can disrupt communications and power delivery on Earth.

A major solar flare in 1989:

- caused the US Air Force to temporarily lose communication with over 2000 satellites
- induced currents in underground circuits of the Quebec hydroelectric system, causing it to be shut down for more than 8 hours.

The Solar Cycle

All solar activities show a cyclic variation with a period of about 11 years.

The three main features of the solar cycle are:

1. the number of sunspots
2. the mean latitude of sunspots
3. the magnetic polarity pattern of sunspot groups.

The number of sunspots increases and decreases with an 11-year cycle, the mean solar latitude at which the sunspots appear progresses towards the solar equator as the cycle advances and the magnetic polarity pattern of sunspot groups reverses around the end of each 11-year cycle (making the full cycle in effect 22 years).

Sunspots

When an image of the sun is focussed on a screen dark spots are often visible, these are called sunspots. Over a number of days they will be seen to move across the surface (due to the rotation of the Sun) and also change in size, growing or shrinking. The sunspots look dark because they are cooler than the surrounding photosphere. A large group of sunspots is called an active region and may contain up to 100 sunspots.

Individual sunspots may appear and disappear – short-lived ones only lasting a few hours whereas others may last for several months. The general pattern of the movement of sunspots shows that the Sun is rotating with an average period of about 27 days with its axis of rotation tilted slightly to the plane of Earth's orbit.

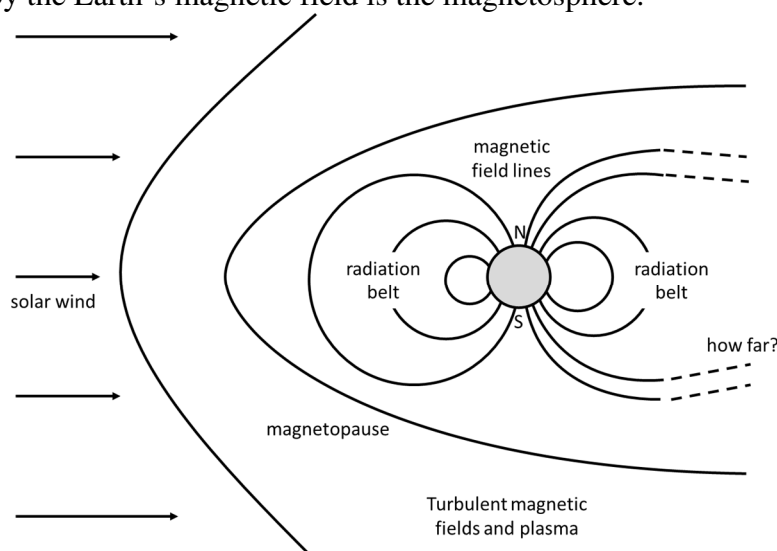
Unlike the Earth, the Sun does not have a single rotation period, since it is not a solid sphere. The period is 25 days at the Sun's equator and lengthens to 36 days near the poles. Sections at different latitudes rotate at different rates, this is called differential rotation.

The Magnetosphere

The magnetosphere is the part of the Earth's atmosphere dominated by the Earth's magnetic field. This region also contains a diffuse plasma of protons and electrons. The magnetic field resembles that of a bar magnet, tipped at about 11° to the Earth's rotational axis. However, the magnetic field is believed to be generated by the motion of conducting material, namely iron, in the Earth's outer core, like a giant dynamo. Geological evidence shows that the direction of the Earth's magnetic field has reversed on several occasions, the most recent being about 30,000 years ago. This lends evidence for the 'dynamo' model as the reversal can be explained in terms of changes in the flow of conducting fluids inside the Earth.

Interaction of the solar wind with the Earth's magnetic field

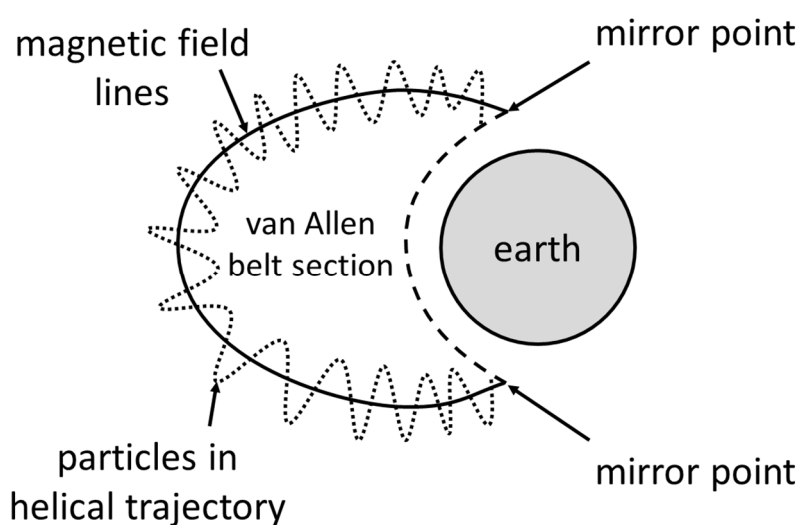
The solar wind interacts with the magnetosphere and distorts its pattern from the simple bar magnet model outlined above. The Earth's magnetic field also protects it from the solar wind, deflecting it a bit like a rock deflecting the flow of water in a river. The boundary where the solar wind is first deflected is called the bow shock. The cavity dominated by the Earth's magnetic field is the magnetosphere.



High-energy particles from the solar wind that leak into the magnetosphere and become trapped form the Van Allen belts of radiation. These are toroidal in shape and concentric with the Earth's magnetic axis. There are two such belts: the inner and the outer.

The inner Van Allen belt lies between one and two Earth radii from its axis, ($R_E < \text{inner belt} < 2R_E$) then there is a distinct gap followed by the outer belt lying between three and four Earth radii ($3R_E < \text{outer belt} < 4R_E$). The inner belt traps protons with energies of between 10 and 50 MeV and electrons with energies greater than 30 MeV. The outer belt contains fewer energetic protons and electrons.

The charged particles trapped in the belts spiral along magnetic field lines and oscillate back and forth between the northern and southern mirror points with periods between 0.1 and 3 seconds as shown.



Particles in the inner belt may interact with the thin upper atmosphere to produce the aurorae. These result from the excitation of different atoms in the atmosphere, each of which produces light with a characteristic colour due to the different energies associated with these transitions.

Charged particles in a magnetic field

The force acting on a charge q , moving with velocity v through a magnetic field B is given by:

$$F = qvB$$

where F , v and B are all mutually at right angles to each other.

Circular motion

As F is always at right angles to v , the particle will move with uniform motion in a circle, where F is the central force (assuming any other forces are negligible), so:

$$F = \frac{mv^2}{r}$$

Equating the magnetic force to the central force we get:

$$\frac{mv^2}{r} = qvB$$

so

$$r = \frac{mv}{qB}$$

Helical motion

If a charged particle crosses the magnetic field lines at an angle, then its velocity can be resolved into two orthogonal components: one perpendicular to the field and the other parallel to it.

The perpendicular component provides the central force, which produces uniform circular motion as described above. The component parallel to the magnetic field does not cause the charge to experience a magnetic force so it continues to move with constant velocity in that direction, resulting in a helical path.

Aurorae

The aurora (aurora borealis in the Northern hemisphere – the northern lights; aurora australis in the Southern hemisphere – southern lights) are caused by solar wind particles which penetrate the Earth's upper atmosphere, usually within 20° of the north or south poles. Between 80 and 300 km above the Earth's surface (aircraft fly at around 10 km altitude) these particles strike nitrogen molecules and oxygen atoms, causing their electrons to be excited into higher energy levels and subsequently emit light when the electrons drop back to their previous energy levels. The most common colours, red and green, come from atomic oxygen, and violets come from molecular nitrogen.