Space Weather

These notes are intended to be read in conjunction with the corresponding PowerPoint slides.

Introduction

My first lectures will be on the topic of *space weather*. To get an appreciation of what's going on we'll cover a pretty wide range of science. Space weather is less well understood than ordinary weather – Fewer people follow it, and it's also rooted in quite complicated phenomena concerning matter in a state that's not found on the ground. It's also a pretty young subject, though the Met. Office now has a space weather operations centre (MOSWOC) which, since 2014, gives space weather forecasts. See their web pages for useful background. I can't remember anyone using the phrase *space weather* before the 1990s. Much of the subject draws on understanding that's come about directly as a result of a vibrant space probe and satellite programme launched by ESA and NASA. You'd be hard pressed to find a subject where the cumulative impact of science shows better. The underlying behaviour of our space environment is ruled by basic laws of physics but the behaviour is so complex that we're only just now appreciating the important issues, after almost half a century of trying. What am I talking about? I'm talking about the environment beyond the Earth's inner atmosphere, beyond the mesosphere for those who remember their meteorology, where space stations, satellites and space probes operate.

The word 'space' almost suggests that these creations are working in a vacuum, that there's nothing there. It's true that if you look up you can't see anything. You'd think you could look all the way to the Sun through empty space. Well, you can look all the way to the Sun but you certainly aren't looking through empty space. You're looking through one of the marvels in the natural universe, an environment of energetic particles, complex electric and magnetic fields that was undreamt of before probes went out and measured what was there. It's an environment that may seem untouchably far away but which has a direct bearing on our human lives, through its influence on some of the highest tech international industries in the world, those of international communications and remote sensing. I think that every 21st century educated person should know about it. Space science is, by and large, the discovery of our times. We've every right to be enthusiastic about it.

Sunspot 9393

I'm going to begin with a lengthy, slightly edited, quotation from the book *Storms from the Sun* by Michael Carlowicz and Ramon Lopez (pp 2 – 4). "On March 22nd 2001 a large sunspot rotated around the Eastern edge of the Sun and into full view of the Earth. Scientists labelled it 'active region 9393' and for two weeks that group of sunspots roiled and pulsed. They swelled into a monstrosity that was visible to the naked eye at sunrise and sunset. The diameter of the active region was 140,000 km, more than 10 times the diameter of the Earth. The same spot lit up on April 2nd, producing a rare 'white-light' flare that was visible on Earth. The explosion was the most intense flare observed since scientists first began to keep records of the X-ray intensity in 1976. The flare was rated an X-22 event on a scale that's only supposed to go to X-20…. All told, the storm raged for nearly six weeks, blazing with dozens of solar flares and spitting several blobs of superheated gas, known as **coronal mass ejections**, or CMEs towards the Earth.

As a direct result of the storms on the Sun, a storm raged in space around the Earth and the northern lights danced as far south as El Paso, Texas and Southern California. Radio communications were distorted and occasionally blacked out over parts of the world for operators using high frequency radio signals, such as airlines, ship-to-shore radio and the BBC world service. At least 2 US military satellites and several commercial satellites suffered outages, hardware failures and computer errors. Some electric power companies worked to re-route power supplies so that their equipment would not be overwhelmed by surges of electric power from space, and several transformers tripped in New York and Nova Scotia. More than 25 flights between North America and Asia were diverted so as not to fly through polar regions…..” and so on.

Satellites

There have been bigger sunspots. There have been worse events. Since the account just quoted there was an even larger flare event in November 2003. At the beginning of this century there were over 600 working satellites in orbit costing over 100 billion pounds. By 2015 that number had increased to 1100 according to one source; about 500 in low earth orbit, about 100 in medium earth orbit and an amazing 500 in geo-synchronous orbit. In addition there are over twice this number of dead satellites still whizzing around the Earth. The slide shows the longitude and altitude of all Earth satellites at about 9 pm yesterday evening at the time of writing this sentence. Without them, life as we have come to expect it would cease: not only international business life but consumer life with its multi-channel TV, pretty cheap global phone-calls, whether from fixed phones or cell-phones, video news from every place no-matter how remote, the weather satellite images and the huge and diverse effort in environmental monitoring, and so on. There may be hardly any astronauts in space, but there’s plenty else. Space weather causes over one billion pounds of damage annually. The basic science behind it is left out of many degree programmes. For those who are taking this course voluntarily, well done – a smart move.

In this introduction, let me quote The UK Space Agency (yes, we have one). UK space industry has a turnover of about £14 billion per annum at the moment. Both upstream supply companies (those that build the technology) and downstream companies (such as satellite broadcasters and suppliers of remote sensing products) are increasing their turnover at a respectable rate. The UK space industry employs directly about 40,000 people, many thousands of whom are graduates. There are serious chances there for you to get employed and I’d like to think that they will be just that bit greater for you at the end of this course! A recent Minister of State for Science and Innovation commented “Space is increasingly important in all our lives. Applications from space underpin many of today’s major commercial sectors; they provide essential information to understand our planet’s environment, changing climate and weather.” If you pause and think about it, this is certainly true. The latest UK National Space Policy document of December 2015 can be found at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/484864/NSP_-_Final.pdf. At the risk of sounding political, I’ll quote “The UK is a global leader in space-related commerce, industry and research, and is host to some of the most profitable and technologically advanced enterprises in the world”. I’m tempted to use Michael Caine’s well-worn phrase: ‘not a lot of people know that’. Ask people what is the first word that comes into their minds when you mention ‘space’ and it may well be ‘NASA’. There’s a lot, lot more in space than has been launched by NASA. This course is about space science, not space exploitation as such, for space science underpins all of space-related commerce.

Part of the UK Space Agency is the National Space Academy at the Harwell Science and Innovation Campus in Oxfordshire. This centre is mainly educational, organising courses, meetings and competitions. The point of mentioning this here is to confirm that space research is not only seen as academically interesting but has real economic impact. The UK government anticipated UK space related turnover growing to £40 billion per annum by about 2030. So far, it seems to be on track. That's a lot of jobs, many of them for graduates!

As a historical footnote, let me say that Britain's involvement in space isn't a Johnny-come-lately entry. The first British-made satellite was launched almost 50 years ago – Ariel3 had 5 space environmental monitoring instruments on board and operated for two years from mid-1967. Ariel3 was launched in the States but Britain was the fourth nation after USA and Russia to launch its own satellite. The *Black Arrow* was an entirely British designed and built rocket that launched the Prospero satellite in 1971. Prospero is still circling the Earth, though now it is no more than an orbiting museum piece. The UK Conservative government of the time cancelled the rocket program without waiting to see if it was a success, claiming it was 'not value for money'. No-one in UK government, it seemed, appreciated that this was the dawn of the age of the communications satellite, the weather satellite, the remote sensing satellite and, indeed, modern space science. What should have been a lead of many years in the space business was thrown away. The UK, though, is back again now, a significant player on the international scene.

Out there

Ok, you're suited-up and tethered to your space-craft repairing the solar panel support, far away outside the Earth's atmosphere. Are you personally in any danger, for example of receiving a lethal dose of radiation? It's a tricky operation, requiring minute-by-minute instructions. Is there a significant chance that your telecoms with Houston or Darmstadt will suddenly go dead, leaving you feeling very lonely indeed? The answer to these questions is "yes" if your support team haven't done their homework well. There's more 'out there' than you probably think. The Earth's atmosphere protects life on Earth from a pretty harsh environment that starts, wherever you are on Earth, less than 100 km away, vertically upwards. That's about an hour at main road driving speed; only minutes away by rocket.

EM radiation from the Sun

Let's begin by looking at the electromagnetic radiation from the Sun. EM radiation is spread over a huge spectrum of wavelengths. You'll recognise the names for different parts of that spectrum. Visible light is only a tiny fraction of the whole spectrum, a very tiny fraction.

The disk of the Sun that we see is called the **photosphere**. Our view into the Sun extends only a few thousand kilometres into the body of the Sun, which isn't very far when you remember that the Sun has a radius of 7×10^5 km. The photosphere is therefore only a thin layer. I might be tempted to say that it's the surface of the Sun, which is what we all assume, but as you'll see in these next few lectures, the Sun's atmosphere doesn't stop there. That's a story I'm coming to. The radiation from the Sun does start there, at the photosphere. It streams out in all directions at a fantastic density. By the time this radiation gets to Earth, 150 million km away, it still has a power density of well over 1 kW per square metre.

Most of the Sun's radiation is spread over the electromagnetic spectrum of wavelengths much like that of any hot body. An ideal hot body emits as a 'blackbody'. A blackbody is one that absorbs all the radiant energy falling on it and doesn't reflect any. It emits its own radiation according to a theoretical distribution discovered by Max Planck, one of the pioneers of modern physics. It may seem odd to call the Sun a blackbody but it satisfies the criterion that any radiation falling on it isn't reflected off its surface but is absorbed into the Sun, to mix with the Sun's own radiation. The Sun is a blackbody, even though it's hot enough to glow pretty white hot. The graph shows the radiant energy from the Sun according to the blackbody law for a body at the temperature of the photosphere. A bit less than half of the emission from the Sun is in the visible. The real Sun emits more high energy radiation than this graph shows, as we'll see soon and as you may have guessed from the earlier quotation, which mentioned X-rays. There is no X-ray emission on our graph because bodies as hot even as 5780 K don't emit X-rays because of their heat. It needs a temperature of millions of degrees to generate X-rays from just the heat of a body. The Sun, though, emits a staggering amount of power, over 10^{26} W.

Flux hitting an astronaut on the Moon or around the Earth

Let's work out how much EM energy is hitting an astronaut on the moon, unprotected by any atmosphere. The slide shows that the astronaut has an area of about 2m^2. Hence he or she will intercept just over 2.5 KW of power. If the astronaut turns around by 90º, this may go down to 1 kW of power. It's a lot of energy and a big change in energy for a small re-orientation. Temperature control within the space suit is a serious difficulty. One strategy is to reflect away as much of this incident energy as possible so changes in it become less noticeable. See how the astronaut on the surface of the moon is wearing a white suit and metallic visor. The astronauts dressed for inside work within the space shuttle and space station don't need this elaboration. Notice that even the foot of the lunar lander is coated in reflective gold to cut down its heat absorption.

Digression on space suits

What does a space suit need to do? Here are 10 things:

- Keep in a pressurized atmosphere
- Give you oxygen
- Remove carbon dioxide
- Maintain a comfortable temperature despite strenuous work and movement into and out of sunlit areas
- Protect you from micrometeoroids
- Protect you from radiation to some degree
- Let you see clearly
- Allow you to move your body easily inside the spacesuit
- Let you talk with others (ground controllers, other astronauts)
- Let you move around the outside of the spacecraft

(list courtesy of 'howstuffworks.com').

You know that the Earth's atmosphere is just over 75% of nitrogen, a gas that's biologically inert as far as we're concerned. The oxygen is the real stuff we need to breathe. Astronauts can't afford to carry around gas cylinders where more than 75% of the gas is useless. Surely it makes sense to breathe largely oxygen. It does, but there is a problem. Oxygen at atmospheric pressure taken for any length of time will poison you. We haven't been built for that environment. Oxygen at lower pressure is fine but the nitrogen that's absorbed in our blood will boil out if we go into a low pressure environment. It's the same effect as CO_2 bubbling out of your coke in the glass. The CO_2 is dissolved at a higher pressure than 1 atmosphere in the can or bottle and when you pour it out, the gas bubbles out of the liquid, quickly at first but it keeps on doing so for quite a while. Divers experience the same effect if they surface too quickly, having increased the pressure of air in their lungs and the rest of their body to match the water pressure deep down. Upon surfacing, the gas at excess pressure dissolved in their blood bubbles out, giving rise to the painful and possibly fatal condition known as 'the bends'. To avoid getting 'the bends' in the reduced atmosphere within their space suits, astronauts need to acclimatize for several hours. This puts a damper on many scenes in SF films, where the hero dons his suit and immediately sallies forth.

The space suit is the astronaut's micro environment, their own totally enclosed life-support bubble. What happens to the CO_2 breathed out? It is absorbed by suitable chemicals to be recovered later. How is temperature controlled? By the substantial heat exchanger carried on the astronaut's back. I'll say more about temperature in the next few slides. As we'll see, there may be a lot of energy striking the astronaut on one side, but it's not far off zero degrees K in the other directions, which makes for a much harder temperature control problem than we have on Earth, where we're surrounded by air at the same temperature on all sides.

Reminder: Temperatures in K

I've twice mentioned temperatures in K, degrees Kelvin, named after the famous Scottish physicist Lord Kelvin. Kelvin degrees are the natural units to use in most physical circumstances. The Kelvin scale has the same intervals as the Celsius scale but absolute zero of temperature, the coldest of the cold, is 0 K. Notice that by international convention, we don't use the degrees sign with Kelvin, only with Celsius. The slide gives the relationship between degrees Kelvin and degrees Celsius. Last year's meteorology students will recognise this relationship.

These lectures aren't going to be all description. Science these days uses numbers in every discipline and we'll use numbers in this course too.

Radiant energy of a hot body

The meteorologists will have met the law that relates the total amount of energy radiated by a blackbody to its temperature. That's all it depends on. The law was discovered by Stefan and Boltzmann towards the end of the nineteenth century. The Stefan-Boltzmann law says that the energy radiated depends on the 4^{th} power of the temperature. That's a very high power to find in a law of nature. In symbols $E = \sigma T^4$, where σ is 5.67×10^{-8} W m^{-2} K^{-4}.

This law tells us how much energy the Sun is emitting. The answer is over 60 MW per m^2. That's an almost unbelievable rate to sustain energy emission. The amount emitted by a black body at 0 ºC is just over 300 W m^{-2}. This is the rate of loss of energy of a black body at 0 ºC placed in space. That's also pretty high. Interestingly enough we'll meet other 4^{th} power laws later in the course.

How hot is a body left in space?

We now have the physics on hand to answer the interesting and important question: "How cold will a body get when left in space?" The body could be a satellite, space-probe, an astronaut on a repair mission, an asteroid or even a planet. For example, what is the average temperature of Mars?

The physics going on is that the source of heat is the radiation from the Sun coming from one direction. In all directions the body re-radiates energy. Almost all bodies in space rotate or tumble around, thereby spreading the sunlight over the whole surface. We'll suppose that this happens. On Earth, it takes the atmosphere and oceans to spread a lot of the Sun's incident energy from equatorial regions poleward, but that's another story. This is a simple calculation of the kind that physicists are reputed do on the back of an envelope to get an idea of what's going on. We'll take the body as being spherical. I'm tempted to start by saying 'consider a spherical astronaut', which is not as crazy as it sounds because to solve this problem you need to put in all the crucial ingredients. What are these ingredients?

The body has radius r. We'll be slightly more subtle than assuming the body is a blackbody by assuming that it is a grey body, i.e. one that reflects a fraction a of the incident radiation. Most real bodies that aren't self luminous do reflect back some of the incident radiation. If they didn't, we wouldn't be able to see them in space. Venus, for example, which is covered in cloud, reflects back 60% of the visible light falling on it, which is one reason why it's so bright. We'll leave this fraction reflected back as the variable a. The Stefan-Boltzmann result will tell us how cold the body will be, along with the law of conservation of energy.

Energy conservation

When the body has reached its equilibrium temperature, it neither gains nor loses energy. If it gained energy over time, it would get hotter. If it lost more energy than was coming in, it would cool down. We suppose it's at its equilibrium temperature. This means that the total energy radiated through the Stefan-Boltzmann relation must equal the total energy received.

The Sun's energy is spread over a disk of area πr^2. The outgoing energy is spread over the surface of a sphere of radius r, which you may know is just $4\pi r^2$. The slide shows the result of equating these two energy flows. The temperature, T, of the body can be calculated for a given incident flux from the Sun. The graph shows the Excel spreadsheet result when the body is at 1AU (one astronomical unit, the average distance of the Earth from the Sun) for a range of reflectivities a.

The general trend of this curve is that the more radiation is reflected back, the colder the body will be, though it doesn't make that much difference to begin with. This is the effect of the T^4 term, which says that only a modest change in temperature is needed to compensate for a significant change in the absorbed energy reaching a body.

Notice that the temperatures are in degrees K, as they must be using Stefan-Boltzmann law. The particularly odd feature of the graph is that the values are pretty well all below freezing. Well, you may say, put a satellite above the Earth and it's not surprising that its average temperature is below freezing. We all know it gets colder as you climb a mountain. This is a specious argument. The reason it gets colder as we climb higher is entirely to do with the Earth's atmosphere. If there were no atmosphere, it would make no difference how high you climbed. You aren't significantly nearer the Sun at the top of a mountain than you are in the valley below. What's very odd about this result is that the Earth is a ball spinning in space and the average temperature of the Earth's surface is well above freezing. The Earth seems to have a natural global warming, caused by the presence of our atmosphere. Indeed, it actually does and the meteorology students will remember me spending some time on this last year. Satellites and space stations don't have this natural warming so if the satellite operators want the average temperature of their equipment to be above freezing, then they need to do something about it. One line of action is to fit large solar panels as wing-like structures and feed the power collected into the central equipment. Of course the panels themselves will have an average temperature well below freezing.

Notice in passing that we've built into our story that the reflectivity a that affects the mainly visible and near IR incident radiation doesn't affect the far IR emitted radiation. You will easily be able to see that if the emitted radiation were affected by the same factor then the reflectivity term would appear on both sides of the energy relationship and just cancel out. This would give the same result as a zero value for a.

Actual satellites around the Earth need to use some energy from their solar panels for heating, otherwise they would be even colder on average than the figures on the slide. The reason is that they spend some of their time on the night side of the Earth, receiving no solar energy at all. What fraction of sunlight they miss depends on the size of their orbit.

The further you are from Sun, the colder it is

We've used some powerful physics to reach some interesting conclusions. We can answer even more questions, such as "*what's the average temperature on a body as far away as Mars?*", if we know how quickly the Sun's radiation decreases with increasing distance. There's nothing special about the Sun's radiation as far as this is concerned. It obeys the general law that the radiation falls off as the inverse square of the distance from the source. Again, anyone who has been to my astronomy and meteorology lectures will have met this very slide before. It is essentially another consequence of the law of conservation of energy. Strictly it applies to 'point' sources of radiation, ones where the spread of the source can be ignored. The Sun is gigantic, hardly a point, but our distance from it is even more gigantic and so in fact even in this case the inverse square law works well.

Example of the inverse square law in action

If you like to see the inverse square law stated as a formula as well as in words, then the formula on the slide encapsulates what is going on. $R_d = R_1/d^2$, where R is the rate at which energy is received (in W m^{-2}) and d is the distance away from the source. The source, remember, is small compared with any distance away involved. R_1 is of course the rate at which energy is received at unit distance.

You can now calculate the energy received per second at any distance from the Sun if you know the energy received at the Earth. The example shows that the sunlight falling per second on a square metre at the distance of Venus from the Sun is about twice what we receive at the Earth. ESA's Venus Express probe that orbited Venus for many years had no shortage of solar energy to tap into.

How cold is the Cassini probe near Saturn?

Let's try out the inverse square law. Saturn is 9.54 AU from the Sun. Applying the inverse square law gives the flux of radiation from the Sun, R, as $1366/9.54^2 = 15$ W m^{-2}. Applying our previous method of determining the average temperature gives the results shown in the graph. It's very cold! 80 K is only just above -200ºC. Man-made electronics doesn't work at these temperatures, due to the properties of silicon and the materials we use for our circuitry. If we want the Cassini probe to work, then we have to provide on-board power. The Cassini probe is the last of the big NASA/ESA probes to be launched with a nuclear power plant inside it. It was launched in 1997 and arrived at Saturn in 2004. We are hearing a lot about this probe nowadays. More than a decade on, it's still working well.

One of Cassini's missions was to drop a remote sensing lander, the Huygens probe, though the atmosphere of Saturn's largest Moon, Titan. This activity was run by ESA with substantial British input. Titan is the only other body in the solar system that shares with the Earth the property that it has a substantial atmosphere that's mostly made of nitrogen. We knew almost nothing about Titan because its atmosphere shrouds it in haze. In this sense it's like Venus, with a surface that's completely hidden from outside. Titan is at the same distance as Saturn from the Sun and you'll see from the graph on the slide that its temperature can be only a little above the liquefaction temperature of nitrogen gas, namely 77 K. In fact a slight greenhouse effect from Titan's atmosphere gives the moon a surface temperature of 94 K. There is liquid methane on the surface (freezing point 90.6 K and boiling point about 112 K). If Saturn and its moons were a little further from the Sun, or the Sun a little less powerful, then Titan's nitrogen dominated atmosphere would be condensed into pools of liquid nitrogen on its surface and the methane vapour would deposit as methane frost. We would then be able to see the surface of this mysterious moon.

What would you expect should happen to the Huygens probe when it landed on the surface of Titan? Without any significant source of power to maintain its temperature, it would inevitably freeze to death. It landed on 14th January 2005 and did just that in a couple of hours or so. The escaping heat from the instruments evaporated some liquid methane in the soil.

The electromagnetic spectrum

We have looked so far at issues connected with the total power of the radiation coming from the Sun. Let's look at the spectrum of this power. The slide shows the electromagnetic spectrum with the conventional names for different parts of it. I've listed the approximate wavelengths at the boundaries of the names, which are there by international convention. The only boundaries which are conspicuous to us are the boundaries of the visible part of the spectrum. Our eyesight peters out at these particular boundaries. That of other animals doesn't necessarily peter out at the same places. For example, it's well known that bees can see in the near UV and rattlesnakes detect the near infrared.

The wavelength of radiation is particularly important, for wavelengths longer than those of visible light, in that it controls the size of equipment needed to generate or receive radiation. For example, medium-waveband radio waves are around 100 m in wavelength and efficient transmission and reception needs aerials at least tens of metres long. TV transmissions in this country used to be about 60 MHz, wavelength 5 m, but changed many years ago to frequencies nearer 600 MHz, wavelengths nearer 0.5 m, and the elements of TV aerials on chimneys shrank in size by a factor of about 10. Satellite TV works at still higher frequencies, around 10 – 12 GHz for European satellite TV, and the aerials are correspondingly different. Wavelength isn't so important at the visible or shorter wavelengths because the detection processes mostly involve the absorption of packets of energy at an atomic level.

It was Max Planck, again, who realised that energy in the electromagnetic spectrum couldn't be absorbed or emitted in smaller and smaller amounts but there was a natural limit to the size of the energy that could be extracted from electromagnetic radiation This limit, said Planck, was proportional to the frequency of the radiation. Einstein realised that Planck's limits on the absorption and emission of radiation applied to the very radiation itself. EM radiation only came in packets, said Einstein, of the size that Planck had found, namely packets of energy $E = hf$, where h is Planck's constant and f the frequency of the radiation. These packets are called photons. Photons are now part of the vocabulary of life. They weren't a century ago.

The convenient unit to measure photon energy is not Planck's constant, which is incredibly small in SI units, but in electron volts. Electron volts are also the convenient unit for changes in chemical energy per molecule or atom. I've listed the energy of the fundamental packets in the electromagnetic spectrum, in electron volts (eV). Photo-sensitive chemicals are affected by visible light photons of 1 – 3 eV. A great many molecules are affected by photons of more than 4 eV, which is sufficient energy to break the comparatively weak chemical bonds that make up the stuff of life. Life, if you think about it, can't be made of stuff that's as chemically stable as most minerals and rocks, for life is intimately dependent on chemical change for sustenance. If the bonds holding a structure together are too strong, then chemical change becomes too difficult. On the other hand, the chemical bonds of life can't be so weak that mere exposure to sunlight will break them down. Otherwise we couldn't exist in sunlight and ultimately it is sunlight that provides the energy for life. I'd say, therefore, that it's entirely predictable that life needs to be protected from most UV, which has photon energies large enough to break chemical bonds that are stable under visible light.

Visible emission of the Sun

With that background, look at the emission spectrum of the Sun. It is composed of a continuous background spread across the rainbow and beyond at both ends, but it's also covered in dark lines. These lines were discovered in the early 19th century and the stronger ones mapped out by Joseph Fraunhofer, the pioneer in this field. Modern instruments show many more lines than even Fraunhofer saw. The top of the slide simulates the general appearance of a solar spectrum as you might see it through quite a decent spectrometer. The bottom shows just a segment of the visible spectrum as it is seen through a high resolution recording spectrometer (i.e. one that shows particularly fine detail in the spectrum). As was discovered many decades after Fraunhofer's work, these lines are almost all due to absorption of light emitted lower in the Sun's photosphere as it passes out through the upper layers. The absorption lines are characteristic of the elements present in the Sun. I say more about this in our *Light Science* course. In the visible, at least, the absorption lines don't change the broad spread of intensity across the spectrum given by the blackbody curve, but they do change the amount of light present at particular wavelengths. What of other wavelengths outside the visible?

Broad spectrum of Sun

This slide shows the Sun's spectrum over a very wide range of wavelengths, with both wavelengths and intensities on a logarithmic scale covering many powers of ten. [A logarithmic scale is one where each successive scale division represents not the addition of 1, 2 or perhaps 5 units of a quantity but the multiplication by an extra amount. Here each division represents multiplication by a factor of 10. Another way of saying this is that each successive division represents the next power of 10]. You can see that resemblance to a blackbody curve is quite good into the UV but then in the extreme UV (EUV), where there should be next to nothing, it all goes pear shaped or, rather, no particularly smooth shape at all. The Sun is generating radiation that a hot body at 5780 K shouldn't do. Where's it coming from? We know from the physics of radiation that it has to be coming from the outside of the Sun, not the inside, because the Sun is opaque to radiation of this wavelength. Such wavelengths generated within the Sun wouldn't escape to reach us. Clearly our picture of the Sun as a very hot ball hanging in space is deficient. Something else is going on there that produces a lot of energetic radiation. Is this radiation enough to worry about in any practical context or is it just something for academics to spend their time on? In short, it is immensely relevant, as we'll see.

Sun in X-ray and Extreme UV

You've seen the graph showing total radiation in different spectral regions, but what does the Sun actually look like when imaged elsewhere in the EM spectrum? The most energetic radiations don't penetrate our atmosphere. We wouldn't be here if they did. To find out what the Sun looks like at these wavelengths needs satellite and space-probe technology. Two probes of the 1990s have made a huge contribution to our understanding: the Yokkoh and Soho probes. The Yokkoh probe produced the first X-ray pictures of the Sun. At high energy, short wavelength, the Sun appears much less passive than it does in the visible. Indeed, it *is* much less passive. It also appears much less well contained within a hard ball. It *is* much less contained. The boundaries of the X-ray and EUV Sun are not the spherical photosphere. The X-ray picture here was taken in wavelengths between 0.3 and 4.5 nm. All the radiation comes from the coronal region of the Sun, outside the photosphere. The source of the X-rays is very uneven, appearing in loops and arches and sometimes from apparently quiet regions of the corona.

The EUV picture taken at 30.4 nm, less than a tenth of the wavelength of the most violet light we can see with our own eyes, is dominated by light from the chromosphere of the Sun, a comparatively thin layer between the photosphere and the corona. The least light at this wavelength tends to come from near the solar poles, corresponding to the areas outside that are identified with 'coronal holes'. These areas are very active in producing a fast component of the solar wind, as we'll see.

At these wavelengths colour doesn't mean what it means in the visible, since colour is a phenomenon of the human visual perception system. Colour is simply used in these pictures to code for some other property. This technique of **false colour coding** is widely used when showing images obtained at non-visual wavelengths. If you look more closely at these pictures you'll see that they are really using colour to code the intensity of the signal, a narrow range of colour in the first image and the saturation of the yellow hue only in the second image. Some pictures of the Sun made in non-visible wavelengths use green or blue images, which are just as valid but somehow don't look quite so realistic. Maybe this is a good way of reminding us they are not real pictures of the Sun as we see it but are synthetic pictures derived from instrumental measurements.

Sun in visible, IR and microwave

Here are 3 ‘pictures’ of the Sun at longer wavelengths. The message comes across clearly that it is in visible light that the Sun looks most uniform, even bland. We perceive a steadiness in the Sun that’s deceptive.

To see structure at visible wavelengths, the Sun has to be looked at through suitable filters. E.g. the Calcium K filtered image in violet light at 393.4 nm highlights the regions of emitting calcium. The red hydrogen alpha image at 656.3 nm highlights the variable temperature of the most common element in the Sun, namely hydrogen.

Much of the Sun’s IR emission is absorbed by our atmosphere. The picture shown is in the near IR, at 1083 nm. Darker areas are where denser gas in the photosphere has absorbed some IR emitted from lower down. The microwave image is made at a wavelength of 17 mm. It particular highlights the transition region between chromosphere and corona, about 2000 km above the photosphere.

Radio flux from the Sun

That a significant amount of the variability of the Sun is associated with sunspots is illustrated by this slide of the solar 10.7 cm wavelength radio emission. This emission is generated by protons, namely hydrogen nuclei. The 11-year sunspot cyclic component is clearly visible. Those of you who haven’t been to the Astronomy course can read up about sunspots. The Sun certainly comes into our story of space weather as the driving cause of it all but I don’t want to spend too much time on solar physics. The slide shows monthly fluctuations on top of the general trend. The sunspot maxima underlying this plot occurred in 2002.

X-ray monitoring

This slide of the X-ray emission of the Sun shows fluctuations on a much shorter time scale, as received at the locations of two geostationary satellites that provide weather and environmental monitoring. GOES 10 is at a longitude of the Eastern Pacific in the northern hemisphere, ideal for monitoring weather approaching the western seaboard of North America, and GOES 12 is located at around the longitude of the eastern seaboard of the US. In a short space of time the X-ray flux increases in this trace by more than a factor of 10.

Look at the units of what’s being measured, namely W m^{-2}. It doesn’t look as if much power is involved. Indeed it’s not. The problem is not heating power but radiation damage by the X-rays. Let’s look at how many X-ray photons per second a flux of 10^{-4} W m^{-2} corresponds to.

Photon flux and power density

X-rays in the range covered by the red trace have energies of approximately 1000 to 10,000 eV. Let’s choose a sample figure of 3000 eV, corresponding to a wavelength of about 0.4 nm. 3000 eV is a huge amount of energy in one packet. Remember that a few eV is enough to break the chemical bonds we are made of. 3000 eV is enough to knock many atoms out of a crystal lattice of silicon and the other materials our electronic circuitry is made of. It's one advantage of the comparatively primitive technology we use to make equipment that it will still work fine with some atoms missing. The conducting tracks inside a decade old Pentium processor are so narrow that you can't see them in the most powerful optical microscope. Yet on an atomic scale they are still very wide indeed. Across the width of 150 nm you can stretch a line of at least 500 copper atoms. One X-ray photon isn't going to do much damage. However exactly how many are arriving if the flux is 10^{-4} W m^{-2}?

The slide shows the calculation. Because an eV is such a small energy, the number arriving is enormous. Every square metre of a satellite receives 200,000 million of these photons per second. More if the energy of the photons were lower, fewer if the energy of the photons is higher. This doesn't look like good news for satellite operators. However, that's not the end of the story. We think of X-rays as penetrating all before them, of showing up the bones in our skeleton as they pass right through us. Medical X-rays can penetrate our bodies, or at least some of them do. However, our bones stop quite a few, which is why our bones show up dark on the X-ray picture. Medical X-rays, though, are exceedingly energetic, with photon energies not uncommonly 100,000 eV. 3000 eV X-rays hardly have any punch. They are absorbed by thin layers of material. Given enough of them, in the end the shielding material will become radiation damaged but the shielding can stand a lot higher dosage than the circuitry behind. Photon flux isn't everything. What counts is the capacity of the photons to do damage and to do damage they must be energetic enough to penetrate into tissue or instrumentation and be absorbed. The damage caused by photons is very energy dependent.

That's all I want to say at the moment on the subject of electromagnetic radiation from the Sun. EM radiation, though, was only one of the components in our environment 'out there'. It is the component we know most about and whose behaviour is easiest to predict.

Solar wind

If all that came from the Sun was electromagnetic radiation, I wouldn't be giving lectures on space weather. The Sun blows the solar wind towards us and it's a very different wind from the one we're used to on the ground. Comet tails were the visible evidence that first suggested this wind was present but it was not until L Biermann examined some comet tails in detail in the 1950s that anyone made the connection. It isn't the conspicuous curved tail of a comet that's the evidence but the much finer, often bluish 'ion tail' that goes almost straight out from the direction of the Sun.

The solar wind has two main components: **plasma** and **magnetic field**. Plasma is a gas whose constituents have equal amounts of mobile positive and negative charge. The internal energy of a plasma is so high that the charges are not bound to each other but free to move by themselves. Plasma is actually the most common state of matter in the universe. Stars are mostly made of plasma. Even the matter between stars is mostly plasma. Not all plasma is out in space. Plasma is the state of matter in a fluorescent tube that's shining and, as we'll see, the Earth is surrounded by plasma. The magnetic field associated with the solar wind is known as the **interplanetary magnetic field**. It plays a crucial role in space weather. Hopefully in a couple of lectures time you'll know more about the interplanetary magnetic field than you do about that other IMF, the international monetary fund.

The positive charges in the solar wind are mainly protons (hydrogen nuclei), supplemented by helium ions and indeed ions of heavier elements. The negatively charged component is electrons. The atomic constituents represent the elements in the outer atmosphere of the Sun, but not necessarily in the same proportions as found in the Sun.

Positive ions in the solar wind

The slide shows the range of particles encountered in the solar wind as measured by the two ion spectrometers on the ACE space probe that I'm coming to shortly. Different masses are plotted up the y axis. Each mass is found in a range of charge states, spread over the horizontal axis. The ellipses indicate the positions of the commonly detected ions, including H^+, He^{2+}, C^{4+} to C^{6+}, O^{6+} to O^{8+}, Ne^{8+}, Mg^{10+}, Si^{8+} to Si^{9+}, Fe^{6+} to Fe^{16+}.

Fe+ ions observed by ACE

The graphs on this slide focus on Fe^+ ions. The basic data that instruments like the one flown give, which are also widely used in labs across the globe for chemical species analysis, is the ratio of charge/mass of the ions, or as it's plotted on the top graph, mass to charge. If all atoms of a given element had the same atomic mass, then identification would be straight-forward. However, there are a range of masses for one element, or to use the word that was coined by a man who used to be a lecturer at Aberdeen University, a range of *isotopes* for each element. [The name was coined by Frederick Soddy, who won the Nobel Prize for Chemistry in 1921 just after he left Aberdeen]. Even when you have identified an element, the peaks are spread out a bit because of the range of masses that might be found.

ACE

NASA's ACE probe is a huge success at monitoring the solar wind. It was launched in 1997, and has enough fuel to continue normal working to beyond 2020 though its notional scientific goal was only 5 years. It sits about 1.5×10^6 km from the Earth directly in line with the Sun, orbiting around the first Lagrangian point associated with the Earth-Sun system. Astronomers in the class have met these points before. Unlike a satellite, ACE is never shielded by the Earth. One motivation for ACE was to give warning on Earth of stormy space weather approaching. How long have we got? Taking an average wind speed of 500 km s^{-1}, to traverse the 1.5×10^6 km to Earth will take 3000 s or 50 minutes. That's enough time for many operators to take emergency action, such as turning the solar panels on a satellite away from the Sun, or preparing for big surges in transcontinental electrical power feeders. The scientific motivation was to find out a lot more about the solar wind and, to this end, there are 9 instruments aboard. See the ACE web-site for further details.

Variability of the solar wind

The solar wind and related particle flux from the Sun is the most variable component of space weather. The measurements recorded by ACE are put on the web for immediate use by the world in what they call browse mode. If you want to use figures for accurate calculations, then you need to get confirmed data that's been subject to calibration checks and other validation. In browse mode you can see the proton and electron fluxes, the IMF and other parameters running over the past 4 days.

Output from ACE probe

Weekly output from ACE is shown from their home page. Highlighted in coloured plots are:

1. B_z, the magnetic field component perpendicular to the ecliptic. Its value is measured in nano-Tesla (nT). The magnetic field at the surface of the Earth is a few hundred micro-Tesla in comparison. This component, it turns out, has a significant effect on whether the solar wind will penetrate beyond the Earth's own magnetic shield.
2. φ gives the angle of the magnetic field in the ecliptic plane. I'll show you the coordinate system shortly.
3. The density of the solar wind is obvious, measured in particles per cm^3 in many places but in particles per m^3 in other places. There are 10^6 cm^3 per m^3 so you have to multiple ACE's figures by 10^6 to get the corresponding concentration in particles m^{-3}.
4. Speed, in km s^{-1}, is very fast. I'll say more about solar wind speed later.
5. Temperature, in K of course, though it doesn't make much difference since temperatures are typically over 10^5 K. I'll say a bit more about temperature shortly.

Coordinate systems

There are quite a range of co-ordinate systems used to plot information in the Earth sciences. Everyone is familiar with the latitude and longitude system and you know it's different from the UK Ordnance Survey system for quoting map references, which is an X/Y system. One system used in space is the GSE system or 'geocentric solar ecliptic' system. The ACE data on the previous slide was given in this system.

Particle density and flux

The **density** of particles looks almost insignificant at some 4 per cm^3. In air under normal conditions N_o (Avogadro's number) of molecules occupy 22.4 litres, giving a density of 2.7×10^{19} particles per cm^3. However, what counts is more likely to be the **flux** through a square metre, i.e. the number of particles crossing an area of 1 m^2 per second. It is this, for example, that determines possible rate of radiation damage to a spacecraft by the solar wind. What is the relationship between particle density and flux?

The slide shows how to convert particle density into flux when you know the speed of the particles. The key to the idea is to see that all the particles in a cylinder of length v, the speed of the particles, will pass through an area in one second. The numbers on the slide show that if the speed of the particles is 500 km s^{-1}, then the flux of particles is pretty substantial. It is very common to use cm^2 as the unit of area and not m^2.

Another term that you'll see used commonly is **fluence**. Fluence is the total number of particles per unit area summed over (i.e. integrated over) time. Sometimes fluence is applied to energy too, then becoming the (total) energy received per unit area.

Temperature and thermal speed

The next two slides introduce the quite fundamental concept of temperature. Just what are you measuring when you put a thermometer into anything? Well you are measuring the average behaviour of the constituent particles of the system. If you had a thermometer in the air in this room, you're measuring the average kinetic energy of the air molecules. The air molecules are moving at random with varying velocities in all directions. The average kinetic energy is proportional to the temperature of the room and for one molecule that constant of proportionality is $3k/2$, where k is the Boltzmann constant. The source of this knowledge is the same Ludwig Boltzmann whom we met before. A man of great insight who felt tortured by the lack of understanding his contemporaries had of his ideas. So much so, that he committed suicide – a very sad story. The slide shows how temperature is therefore related to the average of v^2 for the particles. Applying the idea to protons in the Sun's corona, produces temperatures of a couple of million K.

Temperature of a stream of particles

ACE measures the solar wind streaming towards the Earth. Does the high speed motion of the particles flying past make them very hot? If I had a container of air and moved it through space as fast as the solar wind, the mere act of moving all the molecules at 250 km s^{-1} or even 500 km s^{-1} wouldn't change their temperature at all. Therefore temperature isn't quite as simple as the average kinetic energy of the particles in the gas. Looked at another way, if I was in a fast moving spaceship peering at the Earth's atmosphere in the distance, the molecules would all be moving very quickly with respect to my spacecraft but that doesn't make the Earth's atmosphere any hotter. In the graph at the bottom of the slide, the molecules have the same distribution of velocities as those at the top; the difference is that every molecule has an extra speed of 10 units. The temperature of the two sets of molecules is just the same. Temperature, therefore, is related to the spread of kinetic energies about the average velocity. For the air in this room, the average velocity is zero because the air isn't going anywhere.

Solar wind temperatures

Solar wind temperatures are therefore dependent not on the speed of the solar wind as such but on the range of speeds within the wind. If you look at the ACE results, you'll see that a temperature of about 10^5 K corresponds to a total KE spread of about 50 km s^{-1}. Since electron temperatures are in fact comparable to proton temperatures, the electrons must have a much larger average of v^2 since their masses are smaller by a factor of 1836.

Additional solar wind indicators

You will come across several sites on the web showing ACE solar wind dials that are frequently updated. On the ACE homepage clicking on the dials runs a history loop over the last few days. The significance of Bz, the component of solar wind magnetic field perpendicular to the ecliptic, is that if Bz points south then more solar wind particles enter the Earth's upper atmosphere. I'll come back to this. It's the basic reason that Bz gets a dial to itself.

The solar wind triggers auroral disturbances. Around the globe there is a permanent auroral oval. The web site from NOAA referred to on the slide shows the estimated energy in the auroral oval. In times of high geomagnetic activity, the auroral oval expands southward and we see the northern lights.

Geostationary satellite environment

One of the reasons for monitoring the solar wind is its direct and indirect effect on satellites. There are some 500 geostationary satellites around the world. We saw the X-ray flux at 2 of NOAA's meteorological satellites earlier. Here is the monitored environment at one of them. I'm mainly showing the slide to illustrate what is considered worth monitoring. You can follow up the detail of what is plotted by visiting the web site. I'd like to point out the K_p index shown at the bottom. K_p is the so-called planetary K index, introduced in the 1940s, many years before satellite monitoring, to act as a measure of geomagnetic variability. It is on a logarithmic scale from 0 to 9 (a bit like the Richter scale for earthquakes). Values of 5 and above are treated as serious and flagged as red here.

Public prediction of Kp index from ACE's solar wind data

The final slide of this section shows the K_p index prediction, closing the loop in a way, in the sense that one of the motivations for solar wind monitoring is to predict geomagnetic disturbance. This web page shows the prediction.

The Earth's electronic atmosphere

Space begins pretty close to the ground. The meteorology students will know how quickly the atmosphere peters out, at least in terms of decreasing air pressure. By the time you get to the top of La Mano in Argentina, 5.5 km up, the air pressure has gone down by a half. At 100 km, it's fallen by about $0.5^{(100/5.5)} = 0.5^{18}$ or by more than a factor of 300,000. Air pressure has dropped from about 1000 mb at the Earth's surface to 3 microbars. At 200 km the pressure has dropped by a further factor of 300,000. If there's next to no atmosphere above 100 km, why is it that the solar wind, that very energetic stream of plasma from the Sun, is rebuffed by the Earth if it tries to approach even as close as 50,000 km?

The answer is that the Earth is surrounded by its own plasma that behaves very differently from the lower atmosphere. There's enough of it to cause serious trouble to satellites. Its behaviour is very complicated and after more than half a century of trying to understand the Earth's upper atmosphere, there are still plenty of unanswered questions. I want to leave the solar wind for the moment and start looking at space from the Earth outward.

That the Earth is surrounded by a conducting medium was deduced from the behaviour of radio waves early last century. One of the properties of a plasma is that it conducts electricity because the charges within it are free to move. Before radio, no-one was really interested in what was tens of km up. No planes could reach that high, for the simple reason that there were no planes. No-one could survive that high in a balloon. With the advent of radio, particularly short-wave radio, it became clear that radio waves were able to travel substantially around the Earth. When I was a schoolboy I built my own radio transmitter and in the evenings communicated regularly with people right across Europe at a frequency of 7 MHz and, on one memorable occasion, with someone in New Zealand. During the day, 7 MHz radio waves got down to England and that was about all. What was going on?

My radio waves at night-time were reflected from the ionosphere, whereas during the day they weren't. What happened was that as the day wore on, a radio reflecting mirror was formed in the upper atmosphere. As night wore on, the mirror faded away. That such a layer should exist was predicted by Oliver Heaviside as early as 1902 but not confirmed until the 1920s. It was the famous Robert Watson-Watt, local man from near here who is credited with the invention of the first effective working radar, who, I believe, coined the term *ionosphere* that is now used for the several reflecting layers that exist. Heaviside's layer is the lowest, or D layer of the ionosphere. You can see from the sketch on the slide that some radio waves are reflected by the ionosphere and some transmitted but deviated. Why is this? The ionospheric behaviour depends on the frequency of the radio waves.

Probe the atmosphere from below

The ionosphere is probed by sending radio pulses upwards over a range of frequencies. Timing the delay before the pulses are returned tells the operators how far above the reflecting layer is. If there's no reflection, their pulses never come back. The type of results obtained is shown on the slide in what is called an **ionogram**. Look at the picture on the slide and you will see 3, if not 4, layers from which radio waves have been reflected. These are the layers of the ionosphere, labelled rather uninspiringly D, E, F_1 and F_2.

Interpretation of reflections seen

The layers represent layers of plasma around the Earth, layers in which positive ions and electrons exist in a free state and can respond to the stimulus of incoming radio waves, not unlike the free electrons inside a metal that allow the metal surface to reflect radio waves. The common satellite receiving dish uses this effect to reflect the incoming signals onto a feed placed at the focus of the dish. The metallic, shiny, surface of a telescope mirror does the same for the even higher frequency electromagnetic radiation that is light.

The density of the plasma in the ionosphere is much higher than the density of the plasma in the solar wind. This means that the chance of collisions between positive ions and electrons is very much higher. The temperatures in the ionosphere are also much lower than in the solar wind. Both these aspects combine together to make recombination of the ions and electrons into neutral atoms much more likely, particularly in the lower atmosphere. The only reason that the plasma keeps its presence is because sunlight is constantly supplying enough energy to create ions. During the day, then, the plasmas lower down gradually build up. Overnight they fade away. This happens in all the layers except the very highest, the F_2 layer, which I've shown in the sketch as circling the Earth. The F_2 level is so high that the collision rate between ions is too small to deplete the layer much at night. The slide shows what kind of heights we're speaking about.

The ionogram shows that there is a cut-off in the frequency of radio waves reflected. Hence the important but simple conclusion is that if you want to communicate with satellites, it must be done at high frequencies. This is true for applications like GPS, mobile phone relays, Sat TV and all the remote sensing applications.

Plasma oscillations

When a radio wave enters the ionosphere it forces the electrons to move backwards and forwards. This is simply because the electrons are charges and the radio wave involves an alternating electric field. If you force something to move, that something will respond. How it responds depends on its own natural vibrational properties. Now, the electrons have a natural frequency of oscillation called the plasma frequency, ω_p. When electrons are pushed together in increased concentration, they tend to repel each other. If a space appears in the electron gas with fewer than average electrons in it, there is a tendency for the electrons to move in and fill the space. All this happens pretty quickly because electrons have so little mass and are easily accelerated by forces. In fact only neutrinos have less mass than an electron.

These electron density oscillations are somewhat analogous to sound waves in air, only sound waves involve the motion of molecules. Molecules are more massive than electrons and a typical sound wave frequency is about a kHz. Plasma frequencies in the upper atmosphere are MHz. For sound, the natural resonance depends on the size of the container of air, so the analogy is not quite exact. Working out the natural frequency of vibration of the plasma is the kind of argument an Hons physics student could follow. The result is shown on the slide. The key ingredient in this quite simple formula is that the (natural frequency)^2 is proportional to the concentration of electrons, *n* in the formula.

Understanding the evidence

Now to the crucial bit. If you push a system regularly at a frequency that is lower than its natural frequency of vibration, then the system responds in sympathy. Think of a car going over a series of wide bumps. Inside the car you ride up and down as high as the bumps are because the stimulus on the wheels is at a lower frequency than the natural frequency of oscillation of the car. As the frequency of the stimulus approaches the natural frequency the response of the system becomes ever larger. In the car you don't notice this so much because of the presence of very effective shock absorbers that damp the car's response. Above the natural vibration frequency, the system doesn't respond much. Travelling over cobbles produces much less vertical motion. Actually, a better analogy is being in a small boat at sea. If the waves are slow and long, you ride up and down on them. If the waves are quick and short, they pass under you and splash on the side of the boat. If the waves have the natural frequency of roll of the boat, you are in for a very uncomfortable time.

Applying these general ideas to the ionosphere tells us that the electrons respond to low frequency incident radiation by vibrating in sympathy. Vibrating electrons radiate EM energy and as those who go to my optics course know, this re-radiation will produce a reflected beam. As the frequency of the radio waves is raised, the electrons vibrate furiously and this has the non-obvious effect of making the radio waves within the plasma travel more slowly. That delays the return pulse. Finally, above the plasma frequency, the electron response is feeble, radio waves penetrate and aren't reflected back.

What is perhaps surprising to see when you look at the ionogram is that it is the highest layers that reflect at the highest frequencies. This means that the electron density increases as you go up.

7 MHz signals will reach central Europe from Aberdeen from a single reflection at a height of say 300 km. To reach New Zealand, the signals must undergo multiple reflections without significant loss of too much strength, both from the ionosphere and from the Earth when they reach the Earth again. That very rarely happens.

Confirmatory electron densities

The final slide in this sequence shows electron densities deduced from measurements. They really do increase with height up to the F layer.

The magnetosphere

The ionosphere is not the end of our story. It turns out that there is much more out there beyond the ionosphere, almost all of which would be a surprise to anyone who had followed the story only until the middle of the twentieth century. The key new concept is that of the Earth's **magnetosphere**.

Close to the ground the Earth's magnetic field is much like that of a bar magnet. Far above the ground the Earth's field is strongly distorted by its interaction with the solar wind. This compresses the field on the side nearest the Sun and produces a huge tail to the field on the night side. The interaction is caused by the trapped magnetic field within the solar wind. Everyone knows that when you try to bring two similar magnetic poles together, their fields repel each other and you experience quite a strong force of repulsion. The two fields don't want to mix. This happens with the solar wind and the Earth's field. The particles in the solar wind follow the bent field lines and are deviated past the Earth. The vast majority don't rain down on us. Our magnetic field is therefore like a permanently raised umbrella that is always shielding us from the solar flux. Go to the Moon and the flux rains down directly on you. Neither Mars nor Venus has a magnetic shield and hence a protective umbrella.

Magnetic storm effects

I'm going to say a lot more about the magnetosphere. First, the next slide is just a quick graphic reminder of some of the effects of space weather and why it's necessary to bother about it.

What's out there?

Satellites, the international space station and all of near-Earth space business operate within the magnetosphere. Understanding what's out there is an essential challenge. What is out there is plasma, extended magnetic field and in various places currents of millions of amps. It's a complex structure that's harder to understand even in broad outline than the atmosphere below.

Very variable external magnetic field

This slide, courtesy of a NASA presentation, highlights that the magnetic field around the Earth has two major components. The first is the 'dipole' field, originating inside the Earth and created by the motion of electric currents in the molten iron core of the Earth. This is the main influence on a compass needle at the surface of the Earth. In addition, currents caused by the complex motion of charged particles originating both in the Sun and in galactic cosmic rays (more later on these) produce a magnetic field of external origin. This slides separates the two fields. Of course they both occur together. The internal field changes slowly over a period of years at any one place, this long period making the compass a useful tool. The external field can change significantly in a time measured in hours, or even quicker on occasions. It also, of course, rotates relative to the surface of the Earth whereas on a day-to-day basis the internal field doesn't.

Modelling the fields is very important to most satellite operators because both satellites and spacecraft move through these fields. The craft are full of wires and circuits moving at km s^{-1} through these magnetic fields get significant induced emfs in them, voltages that are significant in comparison with the few volts needed to operate modern integrated circuits. You can see that there is an important practical issue here but I'll not go into any detail.

More detail of the magnetosphere

The first notice that the approaching solar wind has of the Earth is a shock wave, a single stationary wave like that in front of a speeding bullet. In this case the Earth is at a fixed distance from the Sun and the wind is whipping past. Behind the shock, the solar wind particles start to move out around the Earth.

- The *magnetopause* is the outer limit of the field associated with the Earth. It is typically about 10 Earth radii (R_E) away from the Earth, further out than geostationary satellite orbits. These satellites are normally therefore protected from the bombardment of the solar wind but at times of high activity, the bow shock can be pushed inside the geostationary limit of 5.65 R_E.
- The *neutral sheet* is a huge sheet of plasma stretching back over 100 R_E, which is beyond the orbit of the Moon.
- The *cusps* are regions where particles from the solar wind can penetrate towards the Earth down the magnetic field lines.
- The *trapping region* is an extensive region where substantial concentrations of energetic particles are trapped. The trapping region includes the Van Allen radiation belts, two doughnut regions around the Earth where there is a high intensity of energetic particles, trapped electrons and protons.

Van Allen radiation belts

James Van Allen discovered the lower of these radiation belts when he organised the first of a series of rocket probes during the IGY, the International Geophysical Year, that started in 1957. The IGY was an international year of terrestrial physics exploration that really kick started Earth Sciences as a high profile science that had to do with more than just geology. The year was timed to coincide with the solar sunspot peak, which happened to be a very active one. It also coincided with the coming of age of rocketry for scientific purposes and it was no coincidence that the first satellite, Sputnik 1, was launched before the 1950s had finished. Van Allen was responsible for the instrumentation aboard Explorer 1, all of one single Geiger counter on board. To his and the team's astonishment, the Geiger counter sent back higher and higher counting rates as it climbed, eventually cutting out. As they listened to the increasing counts, they didn't know if the counter had failed or simply been swamped by particles. "My God, space is radioactive" was the comment, according to scientific folklore. We know a lot more about the Van Allen belts now and **space isn't radioactive**. The belts do though contain substantial amounts of high energy particles.

Electrons with energies below about 1 MeV and protons with energies below about 10 MeV are not going to penetrate space suits or the outer shells of satellites. However, electrons with energies over 1 MeV have a flux above a 10^{10} m^{-2} s^{-1} from 1 to 6 earth radii (about 6,300 - 38,000 km), and protons over 10 MeV have a flux above 10^{9} m^{-2} s^{-1} from about 1.5 to 2.5 Earth radii (9,500 km - 16,000 km). The radiation dosage from these particles is significant for any one, or any thing, that stays in the belts for some time.

In short, the magnetic field around the Earth that fends off the solar wind is not completely beneficial. It traps high energy particles, creating regions of high radiation levels above the Earth.

Detail of the Van Allen belts

The inner belt is around 1.5 R_E, mainly energetic protons produced from cosmic rays. The trapping of particles by the Earth's field is very effective and the protons spiral backwards and forwards for a very long time (years have been quoted). We'll look at this effect shortly. Although the supply from cosmic rays is not very big, the loss mechanism is very small so pretty substantial fluxes build up.

The outer belt at 3 – 9 R_E is trapped magnetospheric plasma. This is strongly dependent on the solar wind, as we'll see, and as a consequence the particle density here fluctuates a lot. The drift of the particles around the Earth contributes a so-called ring current that can rise to millions of amps. This space current induces ground currents that upset long electricity transmission lines, sometimes to the point of causing huge surges. The main contributor to the ring current is low energy protons, at least low energy compared with those that may cause radiation damage.

Harmful effects of the belts:

- degradation of satellite components, particularly semiconductor and optical devices
- generation of spurious background noise in detectors
- cause errors in digital circuits
- create electrostatic charge-up within insulators
- a health threat to astronauts

Following a particularly intense flare in October 1989 that strongly enhanced the solar wind and later the radiation belts, some 13 geostationary satellites suffered permanent damage to their solar cells. One of the GOES satellites lost 6 years off its design life. In 1991 another storm took 2 years off NOAA's 3 GOES weather satellites.

The physics of motion in magnetic and electric fields

You have to work hard to understand the magnetosphere. Charged particles in the magnetosphere exhibit three types of motion in the presence of magnetic and electric fields found around the Earth.

1. They spiral around magnetic field lines. This is known as gyration. Many old fashioned TV tubes used to have electro-magnets along the neck of the tube in order to deflect the beam of electrons that came from the gun and fell on the fluorescent screen at the front of the tube. By varying the current in two sets of coils at right angles to each other, the beam could be scanned both across the screen and downwards as the picture was written out in about one fiftieth of a second. Around the Earth the magnetic field is permanent and there is enough space for the electrons to be not only deflected but go around in rapid circles.

2. If the field lines come together then the particles can be reflected back along the field line they came on. Since the field lines come together at both poles, then particles end up spiralling from pole to pole
3. Less obviously, if an electric field is present at right angles to a magnetic field, then the particles drift at right angles to both fields. This leads to a drift of particles around the circumference of the Earth.

Let's look in a bit more detail at the 3 effects above, to take some of the hand-waiving out of the description.

Charged particles and a uniform magnetic field - 1

The magnetic force on a charged particle moving in a magnetic field is well known to be proportional to all 3 quantities involved, namely the particle's charge, the speed of the particle and the strength of the magnetic field. In addition, the force is at right angles to both the velocity of the particle and the magnetic field. This last condition is very important in the context because it means that the force can do no work on the particle to change its energy. All it can do is change the direction of a particle, not its speed.

The slide shows that the force on a charged particle acts just like the tension in a string when you're whirling an object around your head. The string keeps the object at a constant distance from you; only the direction of motion changes as the object goes around in a circle. Likewise for charge circling around in a uniform magnetic field. If you've taken our first-year physics courses you should be able to work out that the frequency of rotation is given by $f = qB/(2\pi m)$. Applying this to motion of electrons in a field of 10^4 nT, gives a frequency of over 10^5 Hz. The radius of the circle depends on the energy or, stated more precisely, on the velocity of the particle and is also inversely proportional to the frequency of revolution. This follows from the basic relationship that the distance travelled by a particle in one second is $2\pi rf = v$. This result tells us that high-speed particles entering a given field travel around in bigger circles than low speed particles. In a different scenario, given particles travel in tight circles in a strong field but wide circles in a weak field. We'll need this result in a minute.

Charged particles and a uniform magnetic field - 2

In real life, the charged particles don't just move around the field lines, they move up or down them. When a particle enters a magnetic field, it is likely to have some of its velocity along the field lines and some at right angles to the field lines. Since the force on the charges is at right angles to both the velocity and the field, there is no component of the force parallel to the field lines and hence no force to alter the motion up or down the field lines. The particle therefore keeps its original speed parallel to the field lines. Hence the forced gyration round the lines in combination with the free motion along the field lines produces a spiral motion of charged particles.

Charged particles spiralling down converging field lines

We have to treat the complexity of the world as we find it. We haven't reached the end of the story yet. The convergence of the field lines towards the poles mean that the field lines down which the particles are spiralling gradually converge. It's a serious problem to work out what happens to the spiralling charges in this circumstance. The result isn't that obvious. As the lines converge, the field increases and the particles gyrate in smaller, faster circles. That you'd expect, from two slides back. However, since the total energy of the particle isn't changed because the magnetic force is always at right angles to the velocity of a particle, then if the speed in a circle is increased, there must be less speed down the converging field lines. This is so and eventually the velocity down the field becomes zero when the gyration speed consumes all the available speed from the particle's initial energy. Remember that kinetic energy is just $\frac{1}{2}mv^2$. The slightest wiggle in the backward direction sends the particle gyrating at a slightly slower speed again and the motion back along the field line starts to pick up. So the particle is reflected back along its tracks. The phenomenon is known as the **magnetic mirror**. It really works.

The same idea has been tried in machines that attempt to sustain nuclear fusion in a plasma in the lab. The idea is to use the magnetic pinch to contain plasma that's too hot to be contained in ordinary walls. To make fusion work, a temperature of millions of degrees must be achieved. Everyone knows that fusion power sources haven't been made yet so something must go wrong. What spoils the system in the lab is that the pressure of the plasma is sufficiently high that collisions knock the charged particles off course and too many are lost from the magnetic trap before fusion can be sustained. However, people are still working on it.

Add an electric field ***E*** *perpendicular to the magnetic field B*

That's still not the end of the story for our particles in the magnetosphere. Electric fields also exist in the magnetosphere, which is perhaps not surprising because electric fields are produced both by charges and by the motion of magnetic fields. An electric field on its own exerts a force on a charge in the direction of the electric field. This will tend to accelerate the charge or decelerate it, depending on whether there is acceleration in the direction of the velocity of the charge or against it. The charges, remember, are running up and down the magnetic field lines.

There is a hand-waving argument that says that the electric field is mainly at right angles to the magnetic field. The argument goes like this. If there are components of electric field parallel to the magnetic field lines, then they will accelerate the trapped charge particles and give them more energy. This energy has to come from the source that's creating the electric field and though it might happen briefly, there isn't a source of energy around capable of continually increasing the energy of all the particles in the magnetosphere. Hence in the main the electric field present is perpendicular to the local magnetic field because that way it does not get energy taken from it by the charged particles. Even this last statement isn't exactly obvious but the slide shows what happens when you have electric and magnetic fields at right angles.

We're accustomed to electricity flowing in what you would call an 'obvious' way. You create an electric field by having a voltage between two points. You can create this voltage by a battery or by an electrical generator. If there is a conducting path between the two points, then a positive charge flows from the higher voltage point to the lower voltage point. In most domestic appliances, we provide the conducting path using wires. The current generates heat in the wire, which might be the whole purpose of passing the current, as in a cooker, or perhaps the current generates a magnetic field that operates a motor to drive some DIY equipment like an electric drill. Strange but true, the electric fields in the magnetosphere have a different effect on charge. When the electric field **E** and the magnetic field **B** are at right angles to each other, the result is a motion of the charge that is at right angles to them both, a drift of charge at a speed depending on the magnitude of E/B. This drift operates to make the charge circle around the poles in a **ring current**.

The transformer principle

This ring current is strongest in the outer Van Allen belt. It is millions of amps and increases very considerably during periods of high geomagnetic activity. In fact it acts like the current in the primary circuit of a transformer with the ground and electric transmission lines forming the 'coil' that is the secondary. The transformer principle itself is used by electric supply companies all the time, every day. You see transformers sitting on poles all around the countryside. They are there to convert the higher voltage, lower current, electricity transmitted between pylons and poles to an appropriate lower voltage for local use, essentially the mains voltage. The great transformer in the sky with its primary ring current around the Earth induces currents in the ground. Changes of voltage are not really important. What's significant is that the Earth currents can be big enough to produce surges in power transmission lines. If the surges take place too quickly or are too big, then power transmission equipment is damaged, or at the very least power circuits that could supply power to consumers have to be switched out. Notice the point mentioned in the slide that it is changes of current that cause induction, not steady currents. This is one reason why it is changes in solar activity that are particularly monitored.

Motion of electrons in the magnetosphere

The ring current isn't the only flow of electrons. This slide shows a current of solar wind particles, a current across the plasma sheet and a current involving charged particles, mainly electrons, flowing towards the poles. These later produce the aurora when they get close enough to the Earth to impinge on a significant density of upper atmospheric ions and atoms, as is sketched in the inset. This collision process between energetic electrons and the constituents of the upper atmosphere excites some of the atoms and molecules present, giving them more internal energy. Notice that it's internal energy the molecules have, not extra kinetic energy. This raises the electrons to higher energy levels within the atom. Excited atoms don't stay excited for long, because they can give up their energy by emitting a packet of electromagnetic radiation, namely a photon. If the photon is in the visible part of the spectrum, then we will see it as light. I'll say about more about this shortly.

Currents comparatively close to the Earth increase a lot at times of enhanced auroral activity. One of NOAA's Polar Orbiting Environmental Satellites (POES) was knocked out by a combination of faulty construction and atmospheric charging. The bolts holding the solar panels on were supposed to be covered with a Teflon coating but some bolts were too long and stuck out through the coating. No-one had done anything about it and the satellite was launched. It worked fine for a couple of weeks, orbiting at just less than 1000 km in altitude, until the next geomagnetic storm. The projecting bolts picked up so much charge that shocks were transmitted to circuitry never designed to withstand them and tens of millions of dollars of hardware fell silent. POES is now a joint NOAA and EUMETSAT 'constellation' of weather satellites.

One of ESA's recent missions, the Cluster mission, has an on-board ion discharger designed to electrically discharge the satellite by emitting a beam of ions.

Permanent auroral oval around both poles

The aurora is generated at heights between 300 km and 80 km. Although we only see it on occasions at Aberdeen, there is a permanent **auroral oval** around both poles of the Earth. The slide shows the intensity of the auroral oval superimposed on both the dark and illuminated side of the globe. Notice that it is more intense on the dark side, particularly where it is late evening but not yet midnight. This is the time you will most likely see the aurora at Aberdeen, from say 8 pm to midnight, though it's certainly not exclusively seen within these hours. The auroral oval increases in intensity at times of high geomagnetic activity, initiated by solar wind events. It particularly expands southward (in the Northern hemisphere) and it is only then that we are likely to see it if the night is cloud-free. If the Sun is the cause, why is the aurora so strong on the night side of the Earth? The slide shows that it's not just that it's dark here and we can see light that's always faint compared with sunlight much better. The auroral intensity really is much higher on the night side of the Earth. You'll have to wait a few slides for an idea of the answer to this difficult question.

Auroral colours

The aurora is not like sunlight, which is spread over the whole visible spectrum. Auroral light is mainly confined to spectral lines whose location in the spectrum is characteristic of the molecules and atoms emitting the light. What are they? Mainly molecular N_2 and atomic oxygen, O. At the lower heights involved molecular nitrogen, N_2, is excited and ionised by the incident electrons. It emits characteristically in the blue and red (the combination of which gives magenta). The meteorologists in the class might remember that at higher altitudes the UV below about 250 nm in wavelength is capable of breaking oxygen molecules O_2 into atomic oxygen, O. Excited atomic oxygen emits strongly in the green. When atomic oxygen is excited at even higher altitudes it emits strongly in the red? Why the difference? You're not likely to guess but it has to do with the even lower pressures at higher altitudes. Very low pressures mean even longer times between atomic collisions. Excited states of O that are the source of the red radiation need a long time to emit this colour. So long, in fact, that at lower altitudes the O atoms lose their excitation energy in collisions before they have time to give it out. Hence this red colour can only come from high altitudes.

The fundamental lines from aurora at different heights are suitable primary colours that when mixed together in a whole range of ratios can produce almost any colour sensation you than think of. This mixing of primaries is discussed in our *Light Science* course. The net result is that you can get a very wide range of auroral colours.

Example Auroras

Here are two examples of auroras. The one on the upper right shows the characteristic green from lower altitude atomic oxygen, with magenta showing through from low level N_2. The lower slide shows a red auroral glow from high altitude atomic oxygen that can pretty well cover the sky at times. You can find many excellent auroral pictures taken near Aberdeen in the book *The aurora: an introduction for observers and photographers* by Jim Henderson & John MacNicol (Crooktree Images, Kincardine O'Neil, 1997).

Solar wind and space weather is crucial to the origin of the aurora

To complete this section of our story of space weather, we must go back to the solar wind to see how it links together the Sun and the night side of the Earth.

The solar wind originates outside the visible region of the Sun. It is, essentially, a streaming away in all directions of the Sun's corona. The corona has always been visible in total eclipses of the Sun but not many people on Earth have found a total eclipse of the Sun visible from their home patch and even fewer people in the past realised that during the eclipse you could see the outer atmosphere of the Sun, a part that is normally invisible. Perhaps, you may say, it is not surprising that the outer region of the Sun streams off. After all, hot gases rise from the fire right through the chimney and escape. What makes the solar phenomenon so much stronger is that the corona is not just at a temperature of some 5800 K but at a temperature of a few million K. How does it get to be that hot?

Why is the Sun's corona a few million degrees?

The detailed answer to that question is still occupying solar physicists around the world. The underlying cause is unstable and rapidly varying magnetic fields at the surface of the Sun. Varying magnetic field is necessarily accompanied by electric fields that accelerate ions in the photosphere. No doubt many are accelerated to plunge back into the Sun again but others are accelerated very strongly to fill the space a few solar diameters away from the photosphere with very energetic matter, coronal matter. The slide shows a model of a small part of the Sun's surface. The magnetic field there is vastly more complex than the Earth's simple bar-magnet style field. Moreover, and this is crucial here, it is not steady like the Earth's field but constantly changing. A steady magnetic field can only deflect charged particles without giving them more energy. A changing magnetic field induces electric fields that can accelerate charges. The animation shown in the lecture illustrates the magnetic field streaming unstably from the surface of the Sun. The animation is a computer calculation of the appropriate magneto-hydrodynamic equations and not an artist playing with graphics software.

We live in the Sun

The solar wind isn't just a gust that sometimes comes to us in times of trouble. It's always there. It's now widely recognised to have two components, one being fast and one being 'slow' (if 350 km s^{-1} can be called slow). The fast component is comparatively steady. It seems to be sourced in coronal holes in the more polar parts of the Sun. The slower component is sourced nearer the Sun's equator and is much more variable. Both components take a good many solar radii to get up to full speed, so the acceleration process isn't confined to the surface of the Sun. Indeed you may now be wondering where is the surface of the Sun? In one sense the Sun hasn't got a surface. To put it even more dramatically, we live in the outer atmosphere of the Sun. From that perspective it's not surprising we're affected by the Sun. It's all the more reason that we should understand the Sun really well, particularly its outer reaches.

Magnetic fields and plasma

Magnetic field and plasma are intimately linked. Close to the Sun, the strong magnetic field controls the motion of the plasma as the field lines loop round, so the plasma in the inner corona follows the field lines. You can see this in such splendid images as the one on the slide, showing huge loops of matter following the field lines. The Earth would disappear under the arch of the loop, hardly to be seen. Such loops really are gigantic.

Far from the main body of the Sun, the magnetic field is comparatively weak and is controlled by the plasma. This is the case in the solar wind.

'Frozen' or 'trapped' magnetic field

Electrons flowing down our domestic wiring are constantly finding the atoms of the wire in the way. There are billions of collisions per second between atoms and electrons in even the tiniest current and as a consequence some of the motion of the electrons is transferred to the motion of the atoms of the wire, which heats up as a consequence. This is the underlying reason why wires have electrical resistance. In a plasma like the solar wind, there is virtually no electrical resistance to the motion of charge. There is little that stands in the way of the travelling ions. We're talking about particles travelling in ultra-high vacuum here.

James Clerk Maxwell is certainly the greatest scientist who has ever been Professor at this University. Among his many achievements was to deduce from the experimental evidence in his day and diverse laws that people had put together before him the complete set of equations that governed the behaviour of electric and magnetic fields. These are known as *Maxwell's equations*. There are four of them, plus some subsidiary equations that say how fields interact with matter. Maxwell published his equations 150 years ago (1865), before the revolution in physics at the beginning of the 20th century that brought quantum mechanics, relativity and other completely new ways of looking at nature. Maxwell's equations have remained unchanged ever since he wrote them down. No-one has come up with any reason for changing them. In brief, they work.

Maxwell's equations of electricity and magnetism predict that magnetic field in a plasma like the solar wind is trapped and carried along unchanged by the plasma in motion. Hence magnetic field generated on the Sun reaches the Earth. This might sound very odd. Yet we're familiar with varying electric and magnetic fields travelling in the form of electromagnetic radiation. No-one thinks it's odd that the electric and magnetic fields generated by the motion of atoms in a light-bulb should travel across the room and reach our eye, or even that similar motion in far distant stars should travel across space so that the variations in electric and magnetic fields generated in the distant cosmos should reach us on Earth. So perhaps the solar wind bringing the Sun's magnetic field to Earth and beyond is not that weird.

What does this magnetic field look like?

The Sun's spiral magnetic field

As the Sun turns and matter is ejected from one region of the corona, the magnetic field is stretched out into a spiral form. You can see that this has to be so by drawing a diagram like the one on the slide (which is animated to show how it's built up). It's a curious phenomenon. The matter goes straight out from the Sun but the trapped magnetic field spirals out. Try drawing the diagram yourself and you'll see it has to be so. It's a bit like the spiral pattern of water from a circulating sprinkler. By the time the magnetic field reaches Earth, it typically makes an angle of about 45° to the Earth-Sun direction. This is, of course, picked up by the ACE probe. Look at the appropriate output on ACE's web pages.

When fields collide

I mentioned earlier the formation of the shock wave when the solar wind hits the Earth's magnetic field. The Earth's field shields us from this wind, as effectively as an umbrella keeps out the rain. Perhaps that's a good analogy. We don't stay completely dry under an umbrella and we certainly get wet lower down as rain splashes up from the pavement. There's an equivalent mechanism sending solar wind plasma back towards the Earth from the dark side, well away from the front of the umbrella.

The fluctuating magnetosphere

Diagrams always show the magnetopause as static, of course. In reality it's not. The animated gif here shows the magnetopause fluctuating over a time of 1 hour. You see the current state of the magnetosphere by going to the web site mentioned.

Within the magnetopause

Solar wind particles get in at the cusps, contributing to the background auroral oval. However, the main mechanism for geomagnetic storms arises from instability in the tail of the magnetosphere. The tail doesn't quite flap around like a beached fish but the field there is never still. Particles leak in from the solar wind around the back. Sometimes the tail stretches out much beyond its average length and then snaps back. The electric fields generated by this snapping accelerate both electrons and protons to high energies, feeding the existing currents around the Earth and substantially enhancing them. Many more protons are fed into the ring current in the Van Allen belts and many more electrons are funnelled down the field lines to interact with the upper atmosphere atomic oxygen and molecular nitrogen. That auroral interaction is the end point of a process that started in the Sun, gave rise to instabilities further away from the back of the Earth than the Moon and finally had very energetic particles arriving just a few hundred km above us. Most of our satellites and probes are circulating within this highly active and highly variable region.

Coronal mass ejections (CMEs)

The corona sheds mass not only in the comparatively steady flow of the solar wind but in great bursts, known as coronal mass ejections. They are big, massive, fast and energetic. The mass ejections can quickly expand to become larger in size than the Sun. Although they do have a lot of mass in them (see the slide), they account on average for about 5% of the continuous loss of mass that takes place from the corona. They are caused by huge magnetic loops on the Sun being stretched beyond the limit that they can hold themselves together. Most mass ejections will miss the Earth, because the Earth is a pretty small target 150 million km away but when one does, it is often the cause of a major geomagnetic storm. It takes about 100 hours for the matter to come from the Sun, so observing probes such as SOHO that see the mass ejections travelling through the corona can in principle give us some 4 days warning. The problem is to recognise the characteristic form of a CME that might hit the Earth.

The next few slides introduce some space missions that are particularly relevant. There are, and have been, others such as Geotail, Wind and the amazing IMP-8, the last 'Interplanetary Monitoring Platform', launched in 1973 and a source of valuable data for over 30 years. You can explore these and the probes I'm coming to in much more detail on the web.

*SOHO - **SO**lar & **H**eliospheric **O**bservatory*

SOHO is a hugely successful probe that is a joint project between ESA and NASA. It was launched in 1996 with about a three year life and looks set to achieve at least 20 years of transmitting fundamental data about our Sun. It is designed to observe the Sun and solar wind continuously, orbiting around the first Lagrangian point. 12 instruments image the Sun over a range of wavelengths, particularly in the extreme ultraviolet, image the photosphere and the corona, and look at the bulk motions of the surface as a clue to the internal oscillations of the Sun – the so called helioseismometry. Many images per day are available of the Sun 'now'. The next two slides show two such pictures.

Large Angle & Spectrometric Coronagraph

LASCO is an instrument aboard SOHO that has given us a continuous look at the corona over a period never obtained before by any means. The slide shows one of several pictures available daily. Never minding the science behind it, the image itself is amazing. The visible Sun is within the ring in the centre. The white dots are stars. A huge coronal hole is visible between 7 and 8 O'clock in the picture. You can't doubt after seeing this that the corona is a huge volume around the Sun where particles are streaming out.

Magnetograms

The spectral lines from the Sun's emission contain detail that is a direct result of the emission taking place in a magnetic field. Lines that appear single in ordinary sources on Earth appear split into several components, whose separation depends on the local magnetic field. This phenomenon can be used to work back to the size of the magnetic field in the region of emission. From this a magnetic map, or **magnetogram**, can be constructed. The slide shows an image generated from SOHO data when the Sun had two large sunspot areas on it, one just right (West) of centre and the other on the left-hand (Eastern) limb. The field is grey-scale coded, with black being one magnetic polarity and white the other. The range, 250 gauss (2.5×10^{-2} W m^{-2}) on the scale, is about 150 times the horizontal component of the Earth's field that affects a compass here. Notice the myriad of places that the field comes out of the Sun, as on the hairy-ball model shown earlier.

Cluster II

Cluster II is an ESA mission, salvaged after the initial Cluster probes were destroyed by the explosion of the first Ariane rocket at launch. It's the first-of-its-kind mission to explore the magnetosphere in 3D by the use of 4 identical craft orbiting in formation. Information is sampled not simply at one point in space but at related points at the same time to reconstruct a 3D picture of magnetospheric behaviour. The Cluster orbits take the 4 craft across both the bow shock and the magnetopause at regular intervals. The orbits are highly elliptical, covering a range of distances from Earth of 19,000 km to 119,000 km, 3 to 18.7 Earth radii. The Cluster records magnetic field, electric field, proton and electron fluxes. One recent discovery of Cluster is that the magnetospheric plasma contains waves, travelling very quickly. Discoveries like this have to be fed into the model that attempts to predict magnetospheric behaviour. Cluster II is scheduled to collect data at least until December 2016.

Extra-terrestrial magnetospheres

Five planets beside the Earth have magnetospheres. Mercury's magnetic field was a surprise to everyone because it was assumed that Mercury was too small to have a molten core that could generate such a field. It's about 200 nT, under one thousandth of the Earth's field. Also the solar wind is stronger at the distance of Mercury from the Sun, which is just less than 0.4 AU on average. As a result, Mercury's magnetosphere is pretty compressed, as shown in the diagram. Because Mercury has virtually no atmosphere, it won't have auroras.

Jupiter's magnetosphere is pretty well everything Mercury's is not. Jupiter has a larger magnetic field than the Earth's, by a factor of about 10 on the 'surface'. The solar wind is weaker at Jupiter's orbit of 5.2 AU. As a result, the magnetosphere is huge, extending beyond the orbit of Saturn. Jupiter's magnetosphere has been described as the largest thing in the solar system, bigger even that the Sun, at least if you don't include the solar wind. All Jupiter's moons orbit within its magnetosphere. The corresponding radiation belts are huge also, emitting a substantial amount of radio emission on their own. Jupiter has auroral ovals, first observed by the Voyager probes of the late 1970s, were imaged by the Hubble Space Telescope in the mid 1990s.

The 3rd ingredient in near space

A long time ago now, it seems, we mentioned radiation as the first ingredient 'out there'. Then there's the solar wind and all that comes with it. Finally, there are cosmic rays, and quite a lot of them too. The very name 'cosmic' suggests an origin outside the solar system and this is correct. The graphic on this slide suggests that there is a sphere of influence of the Sun, the *heliosphere*, included in which are all the planets, solar wind, solar and planetary magnetic fields, etc. Into this from outside come the cosmic rays. How do we know? We've known for about a century now, thanks to the pioneering efforts of Victor Hess and his successors.

Cosmic Rays

Cosmic rays cause Geiger counters to 'click' even in the absence of any obvious radiation source. Cosmic rays are the background radiation we are all exposed to. You might suppose that this background radiation originated in the Earth below us. In early 20th century there were no Geiger counters but people realised that if you charge up an object, it spontaneously discharges. Why? The air around the device is made conducting by ions created in the air and this conduction leaks away the charge. People assumed that the source of the ions was in the earth around. Marie Curie had not long before separated radium from a natural product found in the ground and it clearly had the property of ionising air and would rapidly discharge an electroscope, the device that showed how charged a body is.

Victor Hess, a relatively young Austrian physicist, became interested in this problem. Victor Hess did in his own way what Van Allen was to do 46 years later. In 1911 – 1913 he took a radiation detector up high. He couldn't send a rocket; there were no upper atmosphere rockets in those days. Hess couldn't even take a Geiger counter, since Hans Geiger was only just developing his counter at the University of Manchester where he worked in Ernest Rutherford's laboratory. Hess went up in a balloon with the radiation detector of the day, an electroscope that gave a visible indication of when it was charged. Cosmic rays spontaneously discharged the device and from the rate of discharge he could estimate the amount of radiation that was present. Hess found that the radiation rate increased very substantially with height and concluded that whatever was causing the radiation at ground level might be coming from the sky, not the earth.

Further measurements made later by Hess and others found out two very supportive facts. You get the same amount of radiation by day and by night. When directional detectors were made, it was found that you get pretty well the same amount of radiation coming from all directions. Both these facts suggested that the obvious source in the sky, the Sun, didn't seem to have anything to do with it, or only a little. There's a third reason the Sun isn't particularly involved. It's not powerful enough. This may sound extraordinary since it's a power-house on a scale that's hard to imagine but it still can't produce most of the cosmic rays observed. Cosmic rays are found to be very energetic indeed. Many of them consist of nuclei of atoms stripped of all their electrons. That takes a tremendous amount of energy for the heavier elements. In addition the particles have tremendous kinetic energy. Victor Franz Hess shared the 1936 Nobel Prize in physics *for his discovery of cosmic radiation*.

Nowadays we believe that primary cosmic rays are the products of violent accelerations in the exploding atmosphere of a star going supernova. This happens to very heavy stars close to the end of their life. They go ballistic, with a vengeance. Among the projectiles coming out are cosmic rays. It turns out that there are just about enough supernova in our galaxy to account for the frequency of cosmic rays received on Earth.

The very high energy primary cosmic rays are stopped in our upper atmosphere and the act of doing this smashes numerous atoms into pieces, pieces like muons and other sub-atomic particles, than rain down on us from above. It is these secondary particles that trigger our Geiger counters at ground level and discharge electroscopes. Most of them actually pass through us leaving a trail of ionisation and some debris, and just keep going. Life has evolved in this environment and we are to some extent 'radiation hardened' as the technologists put it. There are natural mechanisms in life for discarding faulty molecules and replacing them with new ones. Not so, of course in electronic circuitry that we send up in satellites. It's not self-repairing.

More on cosmic rays

Cosmic rays contain almost all the elements in the periodic table, though with a slightly different distribution from that found in the solar system. This is another clue about their cosmic origin.

The IMF (interplanetary magnetic field) is easily sufficient to bend the path of a cosmic ray. A very rough calculation goes like this:

The gyration frequency, *f*, of a charged particle in a magnetic field is given by:

$$\boxed{f = \frac{qB}{2\pi m}}$$

This we discovered earlier. First we work out what frequency the cosmic ray particle has around the IMF and then what radius of circle it travels in. If this is smaller than the distance between planets, then we know that such a particle will follow the field lines in big circles and not come straight in. Take the case of a heavy, energetic particle, an Fe^{26+} nucleus travelling near the speed of light, *c*. For this particle: q = 4.165×10^{-18} C; B = 10 nT (say); m = 9.37×10^{-26} kg. These figures give $f = 7\times10^{-2}$ Hz.

If such a particle is travelling in a circle with speed close to *c*, the speed of light, then first year physics tells us that radius *r* of the circle is such that $2\pi rf = c$. Since $c = 3.0\times10^{8}$ m s^{-1}, $r = c/2\pi f = 6.8\times10^{8}$ m, or 6.8×10^{5} km, about the diameter of the Moon's orbit. [Lighter ions will travel in tighter circles.] Hence even an energetic heavy particle will spiral round the interplanetary field lines as it comes into the solar system. In this way we lose all track of where it came from.

The influence of the field lines is greater than just destroying a way of satisfying our curiosity about the origin of cosmic rays. The field lines to some extent trap the cosmic rays, shielding the Earth. When the Sun is very active, we get fewer cosmic rays reaching Earth. Cosmic rays have several influences on us. They create new isotopes in the upper atmosphere. Carbon 14 is one of them and the amount of carbon 14 in the atmosphere is a measure of the cosmic ray flux reaching the lower atmosphere. Carbon 14 concentrations can be inferred for many decades into the past, even many centuries, and there is a clear correlation between carbon 14 and solar sunspot activity. The more the sunspot activity, the less the carbon 14. Since there are ways of deducing the carbon 14 content of the atmosphere thousands of years ago, we can deduce sunspot activity long before mankind was recording sunspots by looking at them.

Why are we interested in sunspot activity? There seems to be a climate link between sunspot activity and average temperatures. When there are no sunspots, and this happens for decades at a time, then the climate gets colder, certainly across Europe, where there are century's old weather records, and likely elsewhere too. What is the link? One effect of cosmic rays and the showers of secondary cosmic rays is that they leave extensive ion trails in the atmosphere. These ion trails form nuclei for the condensation of water droplets, cloud and fog droplets in fact. Without the creation of such nuclei there will be fewer clouds formed. The logical development of this argument is that at times of no sunspots, the solar wind is at its weakest (this is true) and the shielding from cosmic rays is least. The increased cosmic ray flux reaching the lower atmosphere creates more ions and more clouds. More clouds shield the ground from the Sun and produce, on average, a colder climate. The logical connections of this argument are all in the right order. Of course the opposite argument should also be valid. The more sunspot activity, the sunnier and warmer the climate becomes. Lots of sunspots produce global warming. Have we had lots of sunspots these past few decades? "Yes" we have. There is a good case that at least some of the global warming in recent times is linked to completely natural causes. All that said, the hard part of the case to make is to obtain a reliable figure for the size of the effect and compare it with the observed climate variation.

Wilson cloud chamber

CTR Wilson was a Scot, born near Edinburgh, who became interested in condensation phenomena when he worked for a short spell at the Ben Nevis meteorological observatory in the summer of 1894. This observatory was located on the top of Ben Nevis but it's now over a century since it was abandoned, largely on account of the expense of manning such an inhospitable outpost. Wilson devised the first version of his condensation chamber in 1895 and he quite soon guessed that ionisation was responsible for the nucleation of some drops visible in it. This connection really determined the fate of the cloud chamber, for it was used to study particle physics for more than half a century. In short, it was the particles that produced the ions that became the focus of attention for the cloud chamber, not the drops that condensed around the ions. Several Nobel prizes came out of it, including one for CTR Wilson himself in 1927. It has been the only Nobel prize in Physics awarded to a Scot. Particle physics has moved on and now particle tracks are detected by very elaborate electronic means but cloud chambers have a lasting place in the history of physics. The cloud chamber certainly contributed to the understanding of cosmic rays. PMS Blackett won the 1948 prize "*for his development of the Wilson cloud chamber method, and his discoveries therewith in the fields of nuclear physics and cosmic radiation*", to quote the Nobel Prize citation.

Radiation from cosmic rays

Radiation from cosmic drays is measured in the same way as radiation from radioactive sources. How radiation is measured is something we should all know about in this day and age. As far as I know, it is touched on in both school Physics in both Scotland and England.

The basic physical quantity measured is the energy absorbed per unit mass of absorber. The unit of radiation dose is the Gray (Gy), named after Hal Gray (1905 – 1965) who worked with Rutherford at the Cavendish lab on the absorption of γ rays in matter. There is a Gray Cancer Institute in London, named after him.

$$\boxed{1\ \text{Gy} \equiv 1\ \text{J kg}^{-1}}$$

A Gy is a big unit. Radioactive nuclei produce α, β and γ radiation that has energies typically of a few MeV per ‘particle’. Now there are about 10^{19} eV in one Joule and hence 10^{13} typical radioactive particles, 10 million million absorbed per kg of material, needed to produce a dose of 1 Gy. 10 Gy of radiation received by you is probably enough to kill you, but read on.

Unfortunately, measuring radiation is not that simple. Deposited energy has different effects in biological tissue depending on the bearer of the energy. In biological terms, alpha particles are very damaging, neutrons slightly less so, protons less so and electrons and γ rays even less so. To take this into account there is a simple *Radiation Weighting Factor* for each kind of radiation. Multiply this into the absorbed dose in Gray and you get the **Equivalent Dose** in Sieverts (Sv).

$$\boxed{1\ \text{Sv} \equiv \text{Radiation Weighting Factor} \times \text{absorbed dose in Gy}}$$

Rolf Sievert was a Swede who worked in the Karolinska Institute in Gothenburg, developing what became known as the Sievert chamber for measuring radiation.

Radiation dose is not only complicated by what is carrying the radiation but by what the radiation hits. Thus, for example, it’s easier to damage your eye with radiation than your leg muscle. Radiation hitting your gonads can seriously damage the next generation. The bottom line is that the exact result of a certain dose in Sv is not predictable.

The Sv, the Gy and the other radiation unit of the Bequerel (Bq) were added to the internationally agreed system of units, the SI weights and measures, in 1979. They replaced the older units of rem, rad and Curie respectively. (The rad was 10 mGy and you may still come across data given in rads). The Bequerel is the unit of radioactivity of a source, measuring the number of radioactive disintegrations per second. A. Henri Bequerel (1852 –

1908) shared the Noble Prize for Physics in 1903 with the Curies, Pierre and Marie, for his discovery of radioactivity.

One memory aid that may appeal to you is to make an analogy between radiation and rainfall. In this analogy, Bequerels measure the amount of rain from the cloud, Grays represent the amount of rainfall that falls on you and Sieverts how wet you get. The analogy isn't quite right but anything that helps remember units like these that we don't meet everyday can be helpful.

Average dose to population

We live in a world that where there is natural background radiation. Without background radiation, natural evolution probably would not happen, and certainly not at the rate at which it does. We therefore owe our existence to radiation. Some of that background comes from cosmic rays. How much can be seen on the slide. The average dose across the UK is 2.65 mSv, according to the National Radiological Protection Board, whose job it is to monitor it.

DNA damage

It is rather frightening to imagine what happens when energetic particle crash into the very molecules that we're made of. Of course lots of molecular constituents in our cells are present in large numbers and the loss of a few molecules doesn't harm the function of the cell. However, at the very heart of the cell, the genetic code in the nucleus that provides the basic formulae for the manufacture of all the proteins in the cell is not present in duplicate. If the DNA is damaged at an important place, then it will fail to code the relevant protein correctly. This raises the whole issue of how radiation sensitive are we?

We have evolved over several billion years in an environment that has a significant background radiation generated by naturally occurring radioactive material and secondary cosmic ray products. What repair mechanisms have evolved? This is a question for biologists. The new factor introduced by the space environment, highlighted in the slide, is that high energy, high atomic number particles (HZE is how they are designated in the literature, standing for high Z, high E) have not been part of the environment we have evolved in because they are screened out by our atmosphere but they cause more damage than simple gamma rays, for example. Can our natural radiation response mechanisms cope with this type of damage? I don't know the answer. One natural defence mechanism is simply to recognise dysfunctional cells and kill them off. We all have this mechanism and I expect everyone has heard of auto-immune diseases, caused by a malfunction of this process when our own healthy cells are mistaken considered dysfunctional and killed off by mistake.

On-going research into space radiation

Fred the phantom torso is a mannequin, or part of one, filled in with real bones and tissue equivalent material and with over 400 cumulative radiation sensors and some 5 active monitors that allowed the instantaneous dose to be monitored. Previous records showed that the average exposure to Apollo crews measured by onboard dosimetry was 4.1 mGy absorbed dose and 12 mSv dose equivalent for the mission duration. This is a few times the annual average for us on the ground but not notably health threatening. At an altitude of 10 km, cosmic rays can produce a dose of 20 mSv per year. Pilots, though, are limited to about 1000 hours per year, which produces an average dose per annum of 1.2 mSv.

Radiation exposure limits for manned interplanetary missions have not yet been defined. For a manned mission to Mars, just over 100 mm of aluminium shielding would be needed to bring the dose equivalent to the internal organs from solar-minimum cosmic rays below the current exposure limit for astronauts in low-Earth-orbit. For those intending to live on Mars, the UV exposure on the surface of Mars is about 1000 times that on the surface of the Earth. Sun cream alone is not the answer and of course can't be applied to flora and fauna exported to Mars.

Radiation exposure on the Moon and Mars, for example, is a real and serious hazard to mankind's exploration of these places. Any imagined future scenario that fails to take this into account isn't realistic. The two main problems are continual bombardment from galactic cosmic rays (GCRs in the literature) and the huge flux of solar wind particles that accompanies coronal mass ejections from the sun and other solar particle events (SPEs in the jargon). Interposing mass in the form of shielding is the traditional method of sheltering from radiation hazard. The effectiveness of shielding depends intimately on the kind of radiation, its energy and the material of the shielding. It is a well-known effect that high-energy particles, the main hazard in space, generate secondary radiation when they hit shielding material and can create atomic fragments too. This effect is minimised by the use of hydrogen containing compounds, or even pure liquid hydrogen, for hydrogen cannot be broken into secondary atoms since its nucleus contains just one proton. Thus materials like polythene can make effective high-energy shields.

Dose measured in orbit

The Earth achieves shielding by a combination of a magnetic shield and the atmosphere. Proposals have been made for several decades now to use either magnetic or electrostatic shielding in space. Magnetic shielding works by deflecting incident charged particles into orbits that miss the shielded volume. Magnetic shielding is only feasible to protect the occupants of a long-term spaceship if high-temperature superconductors can be fabricated into adequate current carrying conductors in space and be allowed to cool by natural radiation into space. This scenario isn't yet technically possible. Electrostatic shielding is another option that relies on potentially simple physics. Very highly positively charged electrodes are placed outside the region to be protected, so that the resulting electric field deflects the energetic positively charged particles away. Of course these electrodes will attract electrons in the solar wind that will discharge the electrodes unless the high voltage is maintained by a power supply. Calculations show that power supplies of many kW will be needed, which renders this method impractical except perhaps for static colonies on the Moon, Mars, etc. So the radiation problem involved in transporting people around the solar system is real and by no means solved, or likely to be solved in the near future. In brief, simple massive shielding needs to be too great for our current rocket capabilities, magnetic shielding needs technology we haven't yet got and electrostatic shielding will consume too much power. Space may beckon but it will seriously test the ingenuity of mankind if we have aspirations to make long-term space missions.

Let me close this section with a quotation from Fred Hoyle in an article he wrote in a Sunday newspaper: *space isn't remote at all. It's only an hour's drive away if your car could go straight upwards.* That hour's drive would take you to a place that is so different from the place we live in that you could be excused for thinking it's a million miles away. It's not. It's here and now, as this section of the course has tried to convey.

The end of section *JSR*