# Gresham College Main logo

3 December 2014

**The Sun, our Nearest Star**

Professor Carolin Crawford

# Introduction

The Sun is our nearest star; it dominates our Sky from a distance of ‘only’ 150 million km. Even though it appears to be the same size as the full Moon, it’s over 400,000 times brighter, and dictates when we have night and day here on Earth. With a diameter of 1.4 million km it’s the largest body in the Solar System (excluding the longer, but far more transitory tails of comets), and it’s also the most massive – containing 99.9% of the total mass (*ie* over 700 times) of all the planets, moons, dwarf planets, asteroids and comets combined. This concentration of mass – and the accompanying gravitational force – is why the Sun sits at the very centre of the Solar System, pulling all the other bodies in orbit around it. We are entirely dependent on the Sun for the habitability of our planet, as it bestows us with the energy (in the form of heat and light) that we require to survive. But it also brings many potential hazards: from the continual flow of dangerous radiation that always lurks just beyond Earth’s atmosphere, to the sporadic and violent space weather that threatens much of our society’s infrastructure. Solar physics is thus unusual in astrophysical research, in having immediate relevance for our everyday lives.

As the nearest representative of stars in general, and the only one that whose behaviour we can really study in detail, the Sun is of crucial significance in understanding and interpreting observations of all other stars, including where they crowd together to form galaxies. It serves as a useful comparison precisely because it is a fairly average star, typical in many of fundamental properties such as its mass, age, size, temperature and chemical composition. The Sun possesses an intense magnetic field, and so it provides a local ‘laboratory’ for the study of the dynamic interaction between matter and magnetic fields – again highly relevant for our interpretation of many exotic cosmic objects where magnetic fields are an active ingredient in their behaviour.

But of course the Sun is of interest in its own right, and tracking its ever-changing atmosphere keeps us fascinated. Solar behaviour changes on timescales of minutes, hours, days, weeks, months, years, decades and centuries – and no doubt a lot longer!

# The structure of the Sun

Given that the Sun has a volume that is over a million times that of the Earth, yet contains ‘only’ 330,000 times the mass, we can immediately deduce that its *average* density is far lower than that of a terrestrial planet. Indeed, the average density is about the same as that of water, and less than a quarter of the density of the Earth; the Sun cannot be wholly composed of the same ‘stuff’ as our planet. It is instead made mainly of the lightest elements, hydrogen and helium, in a gaseous form. The distribution of the colour of sunlight (its rainbow, or *spectrum*), and particularly the fact that it is most strong in the visible waveband, informs us that its surface is at a temperature of 5780 K. But this is by no means the hottest part of the Sun.

The immense gravitational force of the Sun pulls all the matter it contains inwards. The lower layers of the Sun’s atmosphere are squashed by the weight of all the outer overlying gas, and so the density of the Sun increases with depth. The Sun is gaseous throughout; it just becomes denser and hotter as you sink down through the atmosphere. The gas right at the centre has been compressed to a density of around 150 g/cm3 (13 times the density of lead) and experiences a pressure of over 265 billion times the air pressure at sea level here on Earth! Welcome to the *core* of the Sun, the most important region, despite being completely obscured from direct view. The sharp density gradient means that nearly half the mass of the Sun is squeezed into less than 2% of its volume; and under the intense pressure here, the temperature of the gas in the core has risen to close to 15.7 million degrees.

At the extreme temperatures within the Sun, we can no longer regard the matter it contains as a simple gas. It is in the form of a hot *plasma.* There is so much energy contained in the ‘gas’ that electrons do not remain bound to their atomic nuclei, and so it is composed completely of charged particles (it is *ionised*): the detached electrons, and the positively charged atomic nuclei (known as *ions*) they leave behind. The electrical charge inherent in the plasma is what makes life a lot more interesting … and complicated! The path that charged particles take is dictated by the direction and strength of any magnetic fields present. But additionally, the movement of charged particles within a plasma generates magnetic fields… which in turn reinforce and strengthen the ambient magnetic field. The plasma and magnetic field thus obviously strongly influence on one other, with their behaviour and dynamics closely coupled together; the magnetic field directs the motion of the charged particles, and moving plasma will drag the magnetic field along with it. Magnetism is thus fundamental to a proper understanding of what drives the behaviour of the Sun, and the solar magnetic field is the strongest of any object in our Solar System. At the point where it emerges through the surface of the Sun, the global field resembles that from a bar magnet: it has a north and a south pole, with field lines curving out and in between them.

# Historical observations

The history of solar observations of course stretches back over millennia. The Sun doesn’t exhibit much to excite the unaided eye, so there is not a vast repository of historical (pre-telescopic) records, except for the rare occasions when something out of the ordinary is noticed. Large dark marks could occasionally be observed on the solar disc when the light from the Sun was diminished , when it is viewed through cloud or thick fog, or when it is close to the horizon and seen through the thicker layers of the atmosphere. [*Please do note that it is never recommended to view the Sun with the unaided eye, and never observe it at all without specialist equipment.]* Known as *sunspots,* the black specks might be visible only for a couple of days at a time, and years and decades might pass without any being seen.

The earliest known depiction of sunspots dates from December 1128, in the form of a drawing by the English monk John of Worcester in his *Chronicles.* The accompanying notes describe that *…on Saturday, 8 December, there appeared from the morning right up to the evening two black spheres against the Sun.* As his description shows, observers did not know how to interpret the dark spots, assuming either they were silhouettes of something lying between us and the Sun, or some kind of dark storm clouds on its face. It’s worth noting that just five days after John of Worcester’s sunspot observation, Korean astronomers also recorded observations of a red ‘vapour’ that *soared and filled the sky.* This we can interpret as a sighting of the aurora, very rarely seen at such low latitudes.

Only exceptionally large groupings of sunspots could be seen until the invention of the telescope in 1609, whereupon it was realised that sunspots were a common occurrence. Projection of solar images with a telescope enabled features on the Sun’s face to be routinely studied. The first reported telescopic view came from the English astronomer Thomas Harriot who sketched his findings in his notebook dated from 8th Dec 1610. Soon after, other observers – such as Galileo Galilei and David & Johannes Fabricius – tracked the progress of individual sunspots across the solar disc to infer the Sun rotated, with a period of around 27 days. Dissent continued about whether sunspots were intrinsic features to the Sun’s atmosphere or surface, or simply silhouettes of undiscovered planets around it, but consensus was reached by around 1630. However, the progress in understanding sunspots stalled, and was unable to benefit from the next generation of more powerful telescopes as the Sun’s disc unexpectedly went completely blank. Sunspots disappeared entirely from the face of the Sun for the latter part of the 17th century (between 1645 and 1717) during what later became termed the ‘Maunder minimum’, and this lack of sunspot activity was accompanied by a marked absence of aurora sightings.

Sunspots, and their scientific analysis, did eventually resume, and by the 19th century solar observation was a standard part of astronomical research. The English astronomer Richard Carrington was one of the foremost solar observers of the time. His detailed long-term observations of sunspots showed the Sun to have *differential rotation*, in that the equatorial regions of the Sun rotate about 10 days faster than polar regions; this demonstrated that the Sun could not be a solid body, but must be made of gas.

Carrington was also responsible for making a fundamental observation that revealed the close connection between the Sun’s behaviour and events on Earth. On the 1st September 1859, during his routine recording of features on the Sun’s disc, Carrington witnessed two brilliantly bright flashes of white light. They appeared in the vicinity of a large sunspot group and rapidly intensified. Indeed so blinding was the light released that his first thought was that there was a hole in his apparatus letting direct sunlight leak through, but the white patches stayed in the same place on the Sun even when his shifted his experimental setup slightly. Realising that he was observing something important, Carrington briefly left his observation in search of an independent witness to the event and when he returned only a few minutes later, he found the flare of light was already fading, and it disappeared altogether shortly afterwards. This observation on its own represented enormous progress, but it was following happenings observed on Earth that revealed its true significance.

Around 17 hours later, spectacular aurorae erupted in the sky, and were observed in many low-latitude locations, even as far south as Cuba and Hawaii. At the same time, ground-based magnetometers (which record disturbances in the Earth’s magnetic field) recorded one of the largest ever geomagnetic storms. Worldwide telegraph systems were strongly affected in different ways: some were completely disabled, with telegraph lines sparking and reports of subsequent fires in some telegraph offices, and of the telegraph operators receiving electric shocks; other offices continued operating, transmitting messages better than ever, even when the usual batteries were no longer connected! These events were particularly noteworthy because the aurorae and telegraph disruption repeated similar events from only a few days earlier. The juxtaposition of these events on Earth with the flare that Carrington had observed allowed him to make the intuitive realisation that the flare he witnessed on the surface of the Sun was somehow causing the subsequent events on Earth.

# How the Sun keeps shining

Even by the end of the 19th century, progress in observations still could not explain the nature of the energy source responsible for the Sun’s enormous output of power (of around 400,000,000,000,000,000,000,000,000 Watts…). Given that geologists had discovered fossils dating back several hundred million years – fossils of life that had presumably required the warmth and light of the Sun to thrive – how could the Sun have sustained this luminosity over such a long period? Ordinary chemical burning would not suffice as it would exhaust the entire mass of the Sun in less than 10,000 years. Although gravitational compression causes an increase in temperature as it squeezes the gas of the Sun, the Sun can’t shine just because it shrinks. For this to work the Sun would have to have been much larger in the relatively recent past, and there is insufficient mass in the Sun for the process to work for more than about 30 million years, even if one incorporated extra energy released through the recently discovered process of radioactive decay. The source of the Sun’s energy remained a mystery until Einstein’s 1905 special theory of relativity highlighted the much more efficient promise of nuclear fusion.

Chemical reactions rearrange electrons between atoms, whereas nuclear reactions change the contents of an atomic nucleus and thus the actual chemical element itself. But the nucleus is bound by a much stronger force than that keeping electrons attached to an atom, so to break apart this part of the atom requires much greater energy. The force that binds the nucleus operates only over very short distances, and so atoms need to be close together before it can operate; but this brings the added complication as that the Sun is made out of electrically charged plasma, so the (mainly hydrogen) positively-charged atomic nuclei will repel each other at close quarters. Thus for nuclear reactions to occur at all, matter needs to be densely packed (*ie* under conditions of extreme pressure) and moving very fast (*ie* under conditions of extreme heat) so that the electric repulsion can be overcome, and the nuclei get close enough to each to smash into each other. It was the English astronomer Sir Arthur Eddington who realised in the 1920’s that the physical conditions within the core of the Sun were extreme enough to permit the necessary nuclear reactions.

The mainstay of this process is a series of nuclear reactions known as the *proton-proton chain,* which convert four hydrogen nuclei (each just a single proton) into one Helium nucleus (consisting of two protons and two neutrons). Each step in this chain liberates a tiny amount of mass, in that the mass input to the reaction slightly exceeds that which emerges. The energy equivalent to this ‘lost’ mass (through *E=mc2*) escapes to power the Sun. [As the speed of light *c* in this equation is so large, a miniscule mass *m* can provide a tremendous amount of energy *E*.] Using nuclear reactions to power the Sun is around 10 million times more efficient than the release of energy through chemical burning. Even so, the Sun converts 600,000 million kilograms of hydrogen to helium every second to sustain its phenomenal energy output. The improved efficiency of nuclear fusion also means that the Sun has enough mass for its lifetime to be consistent with that of the Earth; the age of the Sun is inferred to be about 5 billion years, slightly older than the Earth and the rest of the Solar System. Its fuel supply is by no means exhausted, with sufficient remaining maintain the Sun at its current luminosity billions of years into the future.

# The chemical composition of the Sun

The gases that make up the Sun are identified through spectroscopy of sunlight, a process developed during the 19th Century. The Sun radiates at all wavelengths to produce a baseline of light known as *continuous* spectrum; imprinted upon this are distinctive patterns of features due to the absorption of colour by atoms within the outer solar atmosphere which enable identification of the chemical elements present in the gas. Not only that, but the strengths of these absorption features reveal the relative fractions different element present, known as the ‘chemical abundance’. The strongest features are due to hydrogen, showing that it’s the most common element. A new (emission-line) feature was detected from spectroscopy of the light emitted from the edge of the Sun during a solar eclipse in August 1868. Previously unseen, it was ascribed to a new chemical element, which was named Helium (after the Greek sun god Helios). Helium was later discovered to be present on Earth by 1895, having eluded detection up to then only because it is so rare. Despite its rarity on Earth, helium turns out to be the second-most abundant element in the Sun. Other spectral features are due to the presence of heavier elements such as sodium, iron and calcium. The chemical abundance (by number of atoms) within the material in the Sun is observed to be 92.1% hydrogen, 7.8% helium and with only 0.1% of all other, and heavier, elements. This ratio is very representative of the abundance of the chemical elements elsewhere in the Universe.

# The lifecycle of the Sun

Our understanding of nuclear reactions at the heart of the Sun informs us how other stars also shine. The fusion of hydrogen to helium is what powers all stars though the majority of their lifetime spent on the *main sequence* (see *The lives of stars* for more information). In return, observations of much younger stars and the nebulae that trace ongoing star formation in the spiral arms of the Milky Way around us reveal how the Sun may have formed five billion years ago. From these observations we infer that the Sun was born from the tenuous cold clouds of gas and dust that occupy the space between stars (the *interstellar medium*). Cooler and denser pockets within these clouds collapse under the pulling force of gravity. The matter within the cloud heats up as it compresses until gravity has squeezed the cloud so tight that temperatures and densities become extreme enough to initiate nuclear reactions, and a (proto-)star is born. [Of course this is a gross simplification. There will be added complications such as how well heat generated during the collapse can be radiated away as the cloud contracts, and whether this is inhibited by dust particles in the cloud. Other factors to consider are the uniformity of the cloud, the presence of any rotation, or magnetic field within the cloud material….] In the spiral arms of our galaxy, we see young stars crowded together into small groups or larger clusters, containing anywhere between tens to thousands of stars, each star having condensed from a separate fragment of a single parent gas cloud. It is possible that our Sun was also born as part of such a group of stars, but its current isolation is because any young group of stars rapidly disperses (within several million years), as they are pulled apart from random motions and the gravity of other nearby objects.

The fact that our Sun contains a tiny fraction of elements heavier than helium – elements which can only have been created in the core of very massive stars – show it cannot have collapsed from an interstellar gas cloud made just of primordial material. Instead it is the product of several generation of stars, and the parent cloud must have been previously enriched with the iron, calcium, carbon etc both ejected during, and formed in, the final supernovae explosions of early populations of massive stars.

Our understanding of the eventual fate of the Sun is also informed by observations of more evolved objects around us in the galaxy. The Sun has sufficient hydrogen at the right temperature and density to continue creating helium for a further 6 billion years or so. There are then other, shorter cycles of gravitational compression (due to the weight of the outer layers of the Sun), re-heating and re-ignition of nuclear reactions that it will go through. But eventually the supply of fuel, and all possibility of future nuclear reactions, is exhausted. By this point the Sun will appear very different from how it does today. It will have become a *red giant*; a much cooler, redder and far more bloated version of itself, with an atmosphere puffed so large as to swallow up Mercury and Venus and make conditions pretty uncomfortable here on Earth. Eventually the outer envelope of the red giant will be lost, expanding away to form a transitory shell known as a *planetary nebula.* The remaining hot core of the star is left exposed as a *white dwarf*, which will slowly cool and fade over billions of years until finally fading into a cold, dark and dense ball of compressed matter.

# The transport of energy through the sun

The nuclear reactions that power the Sun can only occur in highly compressed material at temperatures over 10 million K. Thus they are confined to the core; outside this region the density and temperature drop too low to sustain hydrogen fusion. How then does this energy escape away into space as the heat and light released? A small amount of the energy generated in the nuclear fusion is carried away immediately into space in the form of neutrinos, one of the by-products of the reactions. Neutrinos have no electric charge and interact only very weakly with ordinary matter, and thus can easily move directly through the Sun. However, most of the energy produced in the core is released in the form of highly energetic gamma-ray photons of light, and then this energy then has to be transported up through the Sun’s atmosphere in order to be radiated away. The process to achieve this depends on how the physical conditions vary through the solar atmosphere.

Around the core, the high density of the solar interior inhibits the immediate outward progress of the high-energy photons. The photons are continually being absorbed and then re-emitted by atoms and electrons in the gas, so that any photon only travels an average distance of about a centimetre between successive absorptions. The outwards migration of energy from the centre to the surface can take up to 200,000 years, a rate that corresponds to about 50 cm/hour! This is *radiative diffusion*, and it operates out as far as about 70% of the solar radius.

Further out, the temperature within the atmosphere has dropped so that it is now cool enough for hydrogen atoms to form, which much more efficient at absorbing the photons. This means they don’t pass the energy on, so heat is stored, and requires a different method of onwards transportation. The build-up of heat at the top of the radiative zone leads to a pattern of circulation, similar to the way that water boiling in a pan moves from the heat at the bottom of the pan up to the cool surface. In the outer layers of the Sun, hot plasma rises vertically to the surface, whereupon it cools and sinks back in a physical movement of gas known as *convection.*

# The outer solar atmosphere

The energy created at the core is finally released as light at the ‘surface’ of the Sun, a thick layer of gas known as the *photosphere,* which is at a temperature of 5780K, and has a density of only about a ten millionth of a g/cm3. Even though the matter in the Sun is gaseous, the disc appears to have a sharp edge because the gas in the photosphere is not fully transparent. We can only see through it to a depth of about 400km before it becomes opaque; this might sound a long way, but it’s only 0.06% of the way down into the Sun, and so only a tiny distance compared to the Sun’s full diameter.

The convection currents boiling beneath the photosphere mean that the gas in this region is constantly in motion. Images of the surface (taken in only a particular narrow band of colour) reveal that it is blotchy, and partitioned into cells (called *granules*) that mark where columns of plasma are rising and sinking at the top of the convection zone. Each granule appears brighter at the centre where the hotter gas is rising; and is edged with a dark boundary where the cooler (by about 300K) and thus dimmer gas sinks down. Granules typically span about 1000km, and several million granules patchwork the whole of the solar surface in a dynamic and changing pattern. Cells drift across the surface, sometimes merging with neighbouring granules, appearing and fading over a period of about 20 minutes. Several hundred granules group together into larger convection cells known as *supergranules,* which can be up to 40,000 km across, and exist for a day or two. Within a supergranule, cells migrate out from the centre to the edge, the plasma within this flow trapping and carrying with it some of the magnetic field, which becomes more concentrated at the boundaries of the supergranules.

Above the photosphere sits the thin (only some 2,000 km thick) *chromosphere* , where the gas is ten-thousand times less dense. It is far fainter, and invisible to the eye except when the light of the solar disc is eclipsed – either by the Moon during a total solar eclipse, or by an artificial disc in a process known as coronography. When revealed, the chromosphere appears as a pink ring around the edge of the Sun; the pink colour (and the colour is what inspires the ‘chromo’ part of its name) is radiated by excited hydrogen. Although made of the same material as the lower photosphere, the gas in the chromosphere temperatures are much higher, rising up over 20,000 K at its outermost edge. The light from the chromosphere does not form a smooth spherical shell, but it is structured into long fingers or columns called *spicules* that stretch out from the disc over several thousands of km. These are only short-lived features, caused where by gas is lifted up and out from the Sun before falling down again under gravity. They originate from the edges of the supergranules, where the magnetic field lines have been piled up to emerge vertically from the surface. The magnetic field pulls up some of the chromospheric gas into a temporary, until it collapses back to resume its place in the convective flow back down into the Sun.

The temperature keeps rising with distance beyond the chromospere. The outermost layer of the solar atmosphere is the *corona,* and it forms the most tenuous part of the Sun. At temperatures of 2-3 million degrees, most of the light the corona radiates is in the form of low-energy X-rays, and the incredibly faint visible light given off is again only apparent when the Sun’s disc is completely obscured. The plasma is at such low density that its distribution is heavily constrained by the Sun’s magnetic field, which shapes it into elegant wispy arcs and filaments that stretch over several million km well out into – and responsible for – the inter-planetary medium.

The material contained within a plasma as hot as that within the corona is moving at tremendous speeds. This means the particles it contains (particularly those in the most extended regions) can escape the gravitational influence of the Sun, to form a continual flow of plasma and magnetic field known as the *solar wind.* Travelling at speeds of around 400 – 800 km/s, the solar wind is composed of electrons and atomic nuclei (mostly hydrogen and helium, with only a tiny trace of anything more massive). The material is ferried away into space at a rate of about a billion kg/sec; while this may sound a phenomenal rate, even when summed over the billions of years of the Sun’s life, it still only represents a tiny fraction of the total mass available. The interplanetary medium it creates is incredibly sparse. At the Earth’s distance from the Sun it has dropped to only about 5 particles per cubic centimeter, with the density decreasing still further out into the Solar System. The Earth’s magnetic field is shaped by the solar wind. A bow shock compresses the magnetosphere on the day side of the Earth, where the solar wind is abruptly slowed. The solar plasma then continues to flow around the outside of earth’s magnetosphere, dragging out the night side into a tail more than a million kilometres long.

The Sun exerts an influence into space well beyond the orbit of Pluto, and out to a boundary termed the *heliopause.* This is the theoretical ‘edge’ to the Solar System, marked by the point when the solar wind is finally stalled by the pressure of the immediate interstellar medium (itself largely composed of the winds from neighbouring stars). In August 2012 measurements undertaken by the *Voyager 1* spacecraft of the direction of the ambient magnetic field, and the cosmic ray density, indicated that it had finally reached this boundary 18 billion km away (and 120 times further than the Earth’s distance from the Sun).

# The Sun’s magnetic field

The Sun’s magnetic field is produced in a dynamo-like process, where the motion of an electrically conducting fluid (*ie* the solar plasma) across a magnetic field generates an electric current in a process known as *induction.*  The Sun will always have a magnetic field, originating naturally from its parent cloud of interstellar gas. But this ‘seed’ field will have then been amplified and reinforced, as the induced flow of charge then creates more magnetic field; and the interplay between the plasma and the magnetic field continues in a self-sustaining feedback that maintains the global field of the Sun. There is currently some debate about exactly where in the Sun this process operates, most likely to be from a thin layer of turbulent motions sandwiched between the convective and radiative zones, where a sharp change in the speed at which the gases are moving is expected.

Although the global structure of the Sun’s magnetic field may well resemble that from a gigantic bar magnet, it is far more convoluted on a local level. It is also changing rapidly with time. Different latitudes of the Sun rotate at different rates, and the revolving plasma tugs along the magnetic field embedded within it. The magnetic field lines thus gradually become more tightly wrapped and concentrated around the equator. But as well as the global rotation of the Sun, the convective zone is moving plasma in and out along a radial (vertical) direction. The combination of all these motions produce complicated kinks, twists and tangles that distort the magnetic field at local level. Field lines erupt from and fall back into the photosphere in loops, some of which become so extended that the field lines rise directly out of the surface into space. It is these local entanglements that are responsible for many of dramatic surface features we observe on the Sun.

# Manifestations of solar magnetic activity

Much of the surface activity – the granulation, the spicules and the solar wind – is ongoing. Other, much more sporadic features such as sunspots, solar flares and large looping prominences which are all strongly tied to the magnetic behaviour of the Sun on smaller scales, and their occurrence increases in frequency when the Sun is ‘active’. Magnetic fields are able to store energy; think of the way that an elastic band that is stretched, or a spring that is coiled, can store energy which is released very rapidly when it relaxes back to its original, lower-energy state. Similarly as magnetic fields knot, tangle and kink they are storing energy. However, sharp changes in the direction of a magnetic field are unstable, and the field lines can suddenly reconnect to a simpler, more stable configuration. If oppositely directed magnetic fields come close to each other, they can cancel each other out – and when they do this then all the energy stored within the field is suddenly released in an explosive fashion.

We have already mentioned the transient dark *sunspots* sometimes visible on the surface of the Sun. Any individual sunspot might last anywhere between a few days to several weeks, during which time it can change in size, shape, even blend with others before it eventually disappears. Sunspots look dark because the gas contained within them is about 1,500K cooler that the rest of the photosphere. Whilst this is still a high temperature, the cooler gas is less luminous than the surrounding, hotter surface and so appears dark in comparison to the disc.A single sunspot can span a region several times the diameter of the Earth. It has an irregular shape, with a (cooler) darker *umbra* surrounded by a lighter (warmer) border known as the *penumbra.* Sunspots trace regions with the most intense magnetic fields on the surface of the Sun, and mark where the magnetic field lines emerge directly out through the surface. The concentration of the magnetic field lines locally inhibits the convective motion below the photosphere, thus preventing the heat from the interior reaching the surface in that location. This is why the centre of the sunspot remains cooler than its surroundings. The connection of sunspots to regions where the magnetic field is concentrated explains why sunspots nearly always occur in pairs (and in groups of pairs), as a loop of magnetic field will emerge and return in an upside-down U-shape. The pattern of magnetization within pairs of sunspots shows that they have opposite *polarity* (the polarity simply indicates which way its magnetic field is directed). One sunspot occurs where the field lines erupt out through the photosphere, and the other is where it threads back through on its return. Overall, there is an ordered pattern to this polarity, with the sunspots pairs in the same solar hemisphere sharing the same configuration. All the sunspots move coherently across the disc as part of the Sun’s rotation, with all the leading sunspots of pairs in one solar hemisphere showing the same polarity, which is the same as the polarity of the gobal field in that hemisphere… and thus opposite to the polarity of the leading sunspots in pairs located in the other solar hemisphere.

The magnetic activity can also explain the unexpectedly high temperature of the corona. It does not make immediate sense for the part of the solar atmosphere most remote from the source of the heat generation to be so much hotter than the surface. It seems most likely that magnetic field activity – the abrupt release of stored magnetic energy during reconnection of the field lines – could be responsible for the coronal heating. Sunspots in the photosphere trace where the magnetic field lines extend out through the chromosphere and corona, and indeed the brighter regions of the corona are located above sunspots.

# The Sunspot cycle

By the 19th Century, precise sunspot records accumulated over many years revealed the number of sunspots visible on the solar disc to follow a pattern with an eleven-year periodicity. The *sunspot cycle* regularly undergoes a *maximum* sunspot number followed later by a *minimum* when the Sun is completely bare. Richard Carrington also discovered that sunspots do not appear at random over the surface of the Sun, but are concentrated into two bands to either side of the equator, and that the latitude of these bands varies through this cycle. As the solar cycle progresses, sunspots migrate down from mid-latitude regions, reaching to within about 15 degrees away from the equator. Sunspots trace local variations of direction and intensity within the Sun’s magnetic field, and this pattern of cyclical behaviour shows how the global magnetic field resettles itself, unravelling the accumulated tangling and twisting that has built up over the previous eleven years. Over the cycle, the leading sunspots within groups/pairs move towards the Sun’s equator where they meet and cancel out the leading sunspots from the other hemisphere (which have the opposite polarity). The following members of sunspot pairs converge towards the global pole in that hemisphere (which has the opposite polarity), gradually first nullifying and then reversing the pole over the cycle. This causes the North pole to become the South pole and vice-versa, every 11 years – although the swap in both poles doesn’t necessarily happen at the same time. It’s possible for one hemisphere switch polarity before the other, leaving the Sun temporarily with two north, or two south poles! It thus takes a 22 years for the Sun to complete a full cycle of magnetic field reversal, *ie* through two sunspot cycles.

Over the four centuries that the variation in sunspot levels have been tracked, there are hints of other longer-term cycles and variations that affect the number of sunspots at maximum, such as when the Sun entered a 70-year period of absolute quiet during the ‘Maunder minimum’ of the 17th Century. We are currently on the 24th solar cycle since reliable recording of sunspot numbers started in 1755, and it’s been an exceptionally quiet one. The preceding minimum was unusually deep, lasting between 2007 to 2009; the maximum arrived late, bringing with it fewer sunspots at maximum than since 1906; and there have been two peaks to the maximum (Feb 2012, and one this year now; which is not unique, except that it is unusual for the second peak to be higher than the first). Even though a downwards trend has been indicated for a while, the depth of suppression of sunspot numbers this maximum has been a bit of a surprise. There are suggestions that the next solar cycle could be yet even weaker, with speculation that this could be the start of a new decades-long minimum.

It is noticeable that the decades-long period of the Maunder minimum of solar activity coincided with Europe entering a ‘Little Ice Age’, where much lower global temperatures than average were recorded; much cooler weather also accompanied a second, shorter ‘Dalton minimum’ from 1790 to 1830, suggesting that behaviour of the Sun could affect the global climate. No clear correlations have been found to exist between the the 11- (or 22-) year solar cycle and the Earth's weather. However, it’s possible that the Sun undergoes longer-term variation in behaviour. Obviously the presence of sunspots themselves doesn’t affect the weather, but they are a useful tracer of when the Sun is most magnetically active, which is also when it emits more of the highly energetic ultraviolet and X-ray light. This radiation impacts on, and is absorbed by, Earth’s upper atmosphere, with the potential to heat up the upper layers. Scientists have been using satellite experiments to measure the total amount of energy (in all wavebands, so before any absorption) that arrives at the top of our atmosphere. Observations to date of this *solar constant* (the measure of total solar irradiance) show that although it does track the sunspot cycle, the level in its variation is only 0.1% between maximum and minimum, and influences the global temperature to produce variations of only 0.5oC to 1.0oC. The changes in solar irradiation thus have only a very small impact on Earth’s climate, which is easily incorporated into climate models and predictions. The changes in solar energy cannot account for the global warming witnessed during the last few decades: not only has the Sun has become less active as the Earth’s average global temperature has risen, but global temperature increase is happening far more rapidly than any change in the Sun. Climate models can only account for the recent warming of the atmosphere with the addition of greenhouse gases.

# Prominences, flares…

Sunspot groupings reveal the most magnetically active regions on the Sun’s surface, and so host many secondary phenomena whose frequency of occurrence also vary with the Sun’s magnetic activity cycle. The vertical magnetic field lines that emerge from a sunspot pull out clouds of photospheric plasma into long filaments suspended above the disc, pinning it into fantastic looping arcs. Taking about a day to form, each can contain some 100 billion tonnes of solar material; as this gas is slightly cooler and denser, it appears as dark arcs called *filaments* when seen in contrast against the brighter disc of the Sun, or as bright *prominences* luminous against the black of space if seen in profile at the edge of the Sun. These structures extend much longer than the size of the Earth (some tens of thousands of km), yet they can be observed to change and evolve on timescales of tens of minutes to hours.

If the strong but tangled magnetic fields associated with sunspots reconnect, they release the stored energy very rapidly to produce a dramatic brightening within the immediate environment of the sunspot group. This was the cause of the bright white light flare observed by Carrington in 1859. The energy released immediately heats the plasma to temperatures of tens of millions of degrees, which then radiates light at all wavelengths, but mainly in the form of energetic UV radiation and X-ray emission. The flare can last between several minutes and several hours. The radiation crosses space at the speed of light to arrive at Earth only eight and a half minutes later. Thankfully we are shielded from the harmful effects of this radiation by our atmosphere. As they are intimately associated with sunspot activity, flares occur most often during the maximum of the solar cycle.

# …and coronal mass ejections

The stored magnetic energy released in an exceptionally powerful flare can also heat and accelerate a huge cloud of charged particles – hurling a bubble containing several billion tonnes of solar plasma (along with its associated magnetic field) out from the Sun at speeds of hundreds of km/s to form a *coronal mass ejection.* The cloud produced by such an eruption escapes away out into interplanetary space; but should cause concern if one is directed towards us. When a coronal mass ejection reaches the Earth two to three days later, it rattles the earth’s magnetic field to generate what is known as a ‘geomagnetic storm’. The occurrence of the flare gives us advance notice of this event, that it will arrive between 15 hours and a couple of days later, depending on how fast it’s moving, and how clear the passage between Sun and Earth is. The geomagnetic disruption of the 1859 Carrington event happened so soon after the flare was observed only because it followed hot on the heels of an event a few days earlier; the previous cloud would have cleared the way through the interplanetary medium allowing swift passage for the second.

One of the most spectacular side effects is the creation of aurorae across the sky. The everyday solar wind enables some charged particles to leak into our magnetosphere, but these are rarely energetic enough to produce aurora. The particles have to be accelerated to much higher energies, and such an injection of energy can come from the interaction between the magnetic fields of the plasma cloud and the Earth. When the fields meet at the magnetosphere, if they are oriented in the same direction (eg both point south) then there is not much interaction; but if the field lines are oriented in opposite directions they will reconnect to move to a lower energy state, and release a lot of energy to create a strong disturbance in the magnetosphere.

The rearrangement of the magnetic fields leaves two parts to the cloud of plasma from the coronal mass ejection. One contracts down to earth and accelerates the charged particles trapped within it, which spiral around and down the magnetic field lines on the day time side of the Earth. The other part of the cloud expands away to become part of the solar wind again, but the magnetic fields are stretched further back and finally couple together again with a further release of energy. The charged gas from the solar bubble again streams along the magnetic lines of the Earth on the night time side. Many particles plunge down into the upper atmosphere to collide with, and excite, the atoms and molecules of oxygen and nitrogen which glow to form the aurorae. The light is most often see where the magnetic field lines are most concentrated at the polar regions, and are only seen at lower latitudes during major disturbances.

There are, however, other and far more serious consequences to consider from one of these geomagnetic storms. The initial powerful burst of high-energy radiation reaches the Earth very quickly. Whilst it may not be able to reach the ground, astronauts outside of the earth’s magnetosphere are vulnerable. [Indeed, in August 1972, and inbetween the *Apollo* 16 and 17 missions, an intense solar flare produced radiation that would have seriously harmed astronauts were any outside of a spacecraft at the time.] Modern crewed spacecraft are now made with heavy shielding, which is okay as long as you’re inside when the flare erupts! The high-energy radiation can affect Earth's ionosphere, disrupting GPS and communication signals, and causing long-range radio blackouts. When the plasma bubble reaches the Earth’s magnetosphere, the electronics on satellites can be badly damaged, knocking them out of operation. The rapidly moving magnetic fields induce enormous electric currents that have the potential to damage modern power grids. A worst-case scenario puts such an event as a several threat to the infrastructure of modern society and its reliance on technology, with the economic fallout estimated at billions, if not trillions of dollars.

Since the 1859 Carrington event, we’ve had a small taste of what might happen. For example, a strong solar flare on 13 March 1989 was followed by a severe geomagnetic storm 3 ½ days later. Bright aurorae were observed, communications and control of some satellites were disrupted, and the power grid in Quebec was knocked out causing a 9 hour blackout. Since then preventative strategies have been put in place, and with prior warning of such an event, some precautions can be taken. Just over two years ago, on 22nd/23rd July 2012, we experienced a near-miss when the Sun spat out two coronal mass ejections one after another. These travelled out through the Solar System much faster than normal, their path having been cleared by another coronal mass ejection only four days earlier. The plasma in the bubble had a magnetic field oriented in the way that would cause maximal disturbance. Thankfully the clouds moved only *towards* rather than directly *at* the Earth, and passed through the location that Earth had been just 9 days earlier.

# The future of solar research

There is much still to understand and learn about the Sun – as is often the case with astronomical object, the more intensely it is studied, the more questions there are to answer! The probable consequences of a major geomagnetic storm hitting Earth has added impetus to continual monitoring of the Sun and its activity. Although we may now understand the basics of flares and coronal mass ejections (and the potential for major disruption to our lives that they might bring), we remain unable to reliably predict exactly when or where a flare/coronal mass ejection will be released, or how strong it will be. Monitoring of the Sun’s surface for flares brings the possibility of warning of geomagnetic storms. We also study the evolution of the coronal mass ejections as they disperse away into interplanetary space: for example, the *Stereo* mission consists of two identical spacecraft that travel ahead and behind the Earth in its orbit, affording a stereoscopic view of the plasma clouds as they leave the Sun, and particularly those that might impact the Earth. Attempting to predict the day-to-day behavior of the Sun can be compared to trying to predict the next day’s weather –understanding the dynamo process that drives the sunspot cycle is then akin to making a long-range weather forecast of what next year’s seasons might bring. And as yet, both are imprecise sciences.

© Professor Carolin Crawford, December 2014