GRESHAM COLLEGE

31 January 2018

The High Energy Universe

Professor Joseph Silk

X-rays

View the sky through an x-ray telescope and our conception of the universe changes dramatically. Exploding stars are the most dramatic objects one sees. These are visible at many wavelengths, but the most dramatic images are in the x-ray band, where stars are typically extremely faint. Explosions generate hot plasma that glows in x-rays. We see the aftermath of explosions that occurred hundreds to thousands of years ago, and even further back in time.

Ordinary light is reflected by mirrors. The wavelength of light is much longer than the separation of the atoms in the mirror surface, hence optical photons act like waves. X-rays are another story, as they penetrate most materials and get absorbed. One needs to change their direction in order to focus an x-ray beam. Their wavelength is very small, and x-rays act more like particles than waves when interacting with mirrors, since the separation of mirror atoms is similar the photon wavelength. The main technique for x-rays to change direction is to use what is called grazing incidence scattering or scattering off a mirror at very small angles. This is reminiscent of throwing a stone that skips off water. The small angle scattering is repeated until focusing is achieved.

X-ray astronomy began in 1962 with the launch of a rocket-borne x-ray detector. Intended to search for x-rays scattered off the Moon by impinging cosmic rays, the experiment discovered the brightest x-ray source in the sky, Scorpius X-1, as well as a diffuse background of cosmic x-rays. Subsequent developments led to satellites with focusing telescopes, culminating with the Chandrasekhar X-Ray Observatory, launched in 1999, a 1.2 m x-ray telescope with 0.5 arc-sec angular resolution over a half degree field of view. This is as good as the best resolution achievable by terrestrial optical telescopes.

Black holes

Black holes are intrinsically highly compact objects that emit no light directly. However accreting gas clouds heat up dramatically as they impact the black hole and emit x-rays. This is because impinging gas collides with the black hole at near light speed. This results in intense x-ray and even gamma ray emission. Optical light is almost an afterthought, even though it plays an indispensable role in discovering black holes. The most luminous objects in the universe are sources that emit bursts of gamma radiation. The gamma ray flux is a flash that only lasts a few seconds but leaves traces of an afterglow as faint optical emission. This suffices for the identification and follow-up as the gamma rays and other electromagnetic signals fade away. Bursts of high energy radiation are occasionally associated with unprecedented violence on a cosmic scale, the ripping apart of stars by massive black holes.

Black holes involve extreme physics. Understanding their interface with the visible universe requires pushing our modelling and observations to span a dynamic range of unprecedented extent. The outlook is bright for observational advances but less clear for the progress of simulations. Penetrating the interface of AGN activity and star formation is our current challenge. We see very massive black holes in the early universe. They are quasars, the most luminous objects in the universe, visible because they accrete interstellar gas. The gas heats up, glows in x-rays, and feeds the growth of the black hole. The mass doubling time for a typical massive black hole is about fifty million years. This has two implications. Firstly, we expect to find massive black holes a few hundred million years after the Big Bang. And indeed, black holes as massive as ten billion solar masses are found when the universe was a billion years old. Secondly, we need seeds, smaller black holes, to accrete gas or merge to make the massive ones. If we started from typical stellar black holes, around ten or twenty solar masses, there would not be time enough to grow the cosmic monsters we find, at least in the standard approach to black hole growth.

Black hole seeds

How the seeds are formed is somewhat of a mystery. We believe that they are formed from the first generation of million solar mass clouds formed after the Big Bang. These clouds were chemically pure, no heavy elements having yet formed in supernovae. This means that cooling occurs by hydrogen atom excitations. Electrons jump from one atomic orbit to a higher energetic level, due to a collision with another atom or electron, and then the energy level deexcites by emitting a photon. This is how cooling occurs. Atomic cooling provides a powerful channel for losing energy and guarantees that the clouds will undergo direct collapse to allow formation of black holes, typically of ten thousand solar masses.

The only obstacle is that too much cooling may occur. Hydrogen molecules are a catalyst for cooling as they are more easily excited than hydrogen atoms. Trace amounts of hydrogen molecules can form and encourage fragmentation of the cloud into stars. To inhibit fragmentation, one can resort to the following trick. Once a small black hole forms, it generates enough x-rays to destroy any molecules in its vicinity, and now it can grow to form the first generation of intermediate mass black holes. These are expected to be present in all dwarf galaxies. It is these intermediate mass black holes that undergo catastrophic accretion as dwarf mass halos merge to eventually form supermassive black holes in massive galaxies.

Every galaxy most likely has a lurking giant in its centre. Long ago, they played a crucial role in galaxy formation. Star formation in the forming galaxy was quenched. The radiation from the central black hole limited the amount of infalling gas that fragmented into stars. Long ago, galaxies formed stars prolifically, while today, most massive galaxies are not forming any stars. The black hole eventually reaches maturity. If the black hole were to get any more massive, the gas is mostly blasted away, thereby maintain a universal ratio of black hole to stellar mass. Remarkably, we observe such a relation from the smallest to the most massive galaxies, telling us that old stars and central black holes self-regulate and formed early in the history of galaxies.

Supermassive black holes at the centres of galaxies are generally dormant giants. They were hyperactive in their youth, long ago. They become reactivated only when fresh gaseous fuel is provided. This may happen after billions of years when a merger occurs with a nearby galaxy. Merger fuelling is often observed when we peer into the distant universe with our largest telescopes and observe galaxies in their youth.

Black hole activity can be a dramatic event. We occasionally see powerful jets of plasma that drive giant radio-emitting lobes. These are vigorous outflows that collide with and eject interstellar gas clouds into the surrounding circumgalactic medium, where the gas eventually cools and ends up as intergalactic clouds, to enrich new generations of galaxies. Powerful radio jets are produced very near the black hole by release of energy arising from the rotation and winding-up of magnetic fields anchored in the space-time near the black hole. This space rotates, and the supermassive black hole acts like a gigantic flywheel, with the spin of space providing the momentum. Gripping the flywheel is presumably done with powerful magnetic fields, thought to be omnipresent. The fields initially have a dipole pattern but soon tangle up because of the differential spin and turbulence. When magnetic lines of force of opposing polarity (north versus south for a magnet) intersect, the magnetic field annihilates. Magnetic reconnection releases huge amounts of energy that is initially channelled along the axis of rotation, emerging as a collimated jet and continuing for thousands of parsecs.

Black hole growth

Perhaps the most intriguing result of studies with the largest telescopes show that intense black hole activity and extreme rates of star formation occur in the same objects. They are also driving powerful outflows of gas. This is one of the biggest mysteries of structure formation. Why are these phenomena coinciding at early epochs, especially reaching extreme rates of star formation and mass loss that are rarely seen in the absence of supermassive black holes? Is black hole feeding, with its immense energy release, a consequence of star formation, with the stellar debris feeding the black hole? If so, this fails to address the extreme intensity of star formation seen in many of the most distant galaxies.

Both activity around massive black holes in the nuclei of galaxies and star formation bursts may be the collateral damage from a merger between two galaxies that helps shed angular momentum and allows gas to pour into the vicinity of the most massive central black hole. The galaxy merger provides the gas supply responsible both for star formation and supermassive black hole fuelling. In a merger, gas cloud orbits are perturbed, and some clouds are directed into the capture zone of the supermassive black hole, to refuel its activity. The increase in gas mass stimulates star formation at the same time. Or have the powerful outflows from the black holes compressed nearby clouds and triggered an intense burst of star formation in the surrounding clouds?

Of course, the outflows show that the star formation rate is being quenched. But this may have been preceded by a phase of triggering that induced the winds. The jury is out on any resolution to these questions. New observations with instruments such as the ALMA radio interferometer, now taking data at high angular resolution, and the James Webb Space Telescope, to be launched in 2018 and able to take exquisite images in the near infrared, will eventually elucidate many of these issues.

As for the origin of supermassive black holes, the ultimate window on building and observing them is gravity waves, detected from stellar mass black holes for the first time in 2016. Gravitational wave experiments are being planned to search for traces of the elusive signatures of supermassive black holes. These include the approved ESA space experiment LISA, a three-satellite interferometer with million km-long arms, spacings between the triangle of satellites connected by laser beams. An interferometer looks at time delays between light waves transmitted in orthogonal directions. Tiny shifts result in adding of intensity and bright spots if wave crests align, or in cancellations and blackness if peaks cancel out. In this way, one can measure the strength of a traversing gravity wave. LISA can measure gravity waves that produce the displacement amounting to the thickness of a human hair over a million kilometres. One can measure gravity waves produced by black hole mergers that occur billions of light years away.

Many of these supermassive black holes are in binary pairs and will orbit together and eventually merge, emitting gravity waves as they do. The LISA window for gravity wave detection includes mergers of black holes in the mass range 10^4 to 10^7 solar masses. This will provide the ultimate proof of the existence of supermassive black holes. We may have to wait for LISA, currently scheduled for launch in 2034, before our understanding of the formation of supermassive black holes is ultimately achieved.

Gamma rays

Photons that are a hundred times more energetic than x-rays are called gamma rays. These are notoriously penetrating through, for example, human tissue. Produced copiously in thermonuclear explosions, gamma rays are a prime cause of cancer. Chernobyl is a recent example of such an effect, Hiroshima was another. The sun is a source of gamma rays produced in occasional flares. So also, are the compact remnants of supernovae, neutron stars. More distant galaxies with very massive black holes activated by accreting clouds of gas are especially strong gamma ray sources. Fortunately, the earth's atmosphere shields us against cosmic gamma rays.

Gamma rays are very rare compared to cosmic x-rays and are notoriously difficult to focus because they are highly penetrating. The trick is to use a collimator that only allows gamma rays to pass into the detector from a narrow direction. Using a honeycomb of tube-like collimators, one can build up a crude image. Of course, one has to do this above the earth's atmosphere, which meant that the first cosmic gamma ray sources were detected by space satellites.

The most exciting development came in the late 1960s and 1970s with a series of Russian, then NASA, defence satellites that discovered bursts of gamma rays from space. These were eventually realised not to be clandestine nuclear tests, as originally feared, but true cosmic sources. However, the poor resolution of the early gamma ray telescopes made their distance, and hence their nature, difficult to ascertain. Gamma ray bursts were identified only decades later when gamma ray telescopes were launched with improved resolution, culminating in the 2008 launch of the Fermi Gamma Ray Telescope. They are now recognized to be a phenomenon involving remote stellar death and accretion onto highly compact objects, both neutron stars and black holes. They are the most luminous objects in the universe, during the few seconds that they flare. Hence although rare in our own galaxy, occurring at most once in a million years, many occur every day emanating from the remote depths of the universe, typically billions of light years away.

Gamma ray bursts are of very short duration, typically a fraction of a second, or of longer duration, typically tens of seconds. The long duration events are found in star-forming galaxies associated with the collapse of massive stars and often to be supernovae. The short duration bursts are only identified in older galaxies and thought to be neutron stars being disrupted and swallowed by black holes.

Cosmic rays

Particles of matter are accelerated to high energies in cosmic explosions. These are called cosmic rays. We measure them over a huge range of energies, from 50% of the speed of light to 99.9999999999% of light speed. They are mostly protons, although at the highest energies, their composition changes to an admixture of iron nuclei.

Cosmic rays are blocked by the earth's atmosphere and were first discovered in a pioneering balloon flight by German physicist Victor Hess in 1912. Most cosmic rays are produced in the remnants of massive exploding stars, or supernovae, in our own galaxy. The lowest energy cosmic rays are accelerated by solar flares. The most energetic cosmic rays are accelerated in more energetic environments, including the centres of distant galaxies where there are supermassive black holes generating strong outflows. These outflows are cosmic particle accelerators.

The energies of cosmic rays extend from the mildly relativistic range, around a GeV, to energies a trillion times larger. The energy distribution is that of a power-law, with 2 or 3 kinks. These kinks demarcate transitions for example steepening when cosmic rays escape from the galaxy, and flattening when extragalactic cosmic rays kick in. This way one can better understand the sources of injection of the cosmic rays.

Cosmic rays are detected by the ionization tracks that they leave behind. They were first discovered using stacked layers of photographic emulsions. At high energies they produce showers of ionizing particles in the earth's atmosphere. At the highest energies, the air showers extend over kilometres, and cosmic ray telescopes need to span huge areas to track the events that can be seen as nanosecond duration light flashes produced by muons.

In fact, muons were first discovered in cosmic rays. This is because a muon is a short-lived particle, lasting only a trillionth of a second before it decays. But thanks to Einstein's theory of relativity, when the muon is travelling very close to the speed of light it lives for much longer. Cosmic ray muons survive for the time needed to traverse the upper atmosphere where they were produced until they impact cosmic ray detectors on the ground.

The largest cosmic ray telescope in the world is the Pierre Auger Cosmic Ray Observatory in the Argentinian pampas, in the Mendoza region. There are 1600 water tanks spread over 3000 square kilometres designed to look for blue light flashes from air showers triggered by the most energetic cosmic rays. The events are rare, about one per square kilometre per century is expected, hence the huge area that is covered by the Auger Observatory. Cosmic ray is detected at energies up to the ultimate limit of about a billion trillion GeV, where they are blocked by absorption against the cosmic microwave background radiation.

Neutrino astronomy

Ghostly elementary particles that carry energy and spin, but essentially have no mass, are called neutrinos. These are produced in nuclear reactions. Our sun is a source of neutrinos, which provide the only way of "seeing" into its very centre. In this way, the theory of the origin of solar power by burning hydrogen into helium was verified. Solar neutrinos have energies of a few MeV. But high energy cosmic rays are a source of cosmic neutrinos when they strike intervening interstellar or atmospheric protons and produce other energetic particles, including neutrinos.

In fact, neutrinos can possess a tiny mass. This is the subject of intense searches as discovery would shed light on what lies beyond the standard model of elementary particles. The origin of neutrino mass requires new physics, hence searching for evidence of neutrino mass is a priority in experimental physics. There are two different approaches to determining neutrino mass. One involves cosmology, one direct measurement.

In cosmology, neutrinos move at light speed just like photons. The universe long ago was very hot and contained only relativistic particles, predominantly photons and neutrinos. If neutrinos possess even the tiniest mass, they can affect the rate at which the universe is expanding. This leaves an imprint on the temperature fluctuations in the cosmic microwave background, the fossil radiation from the Big Bang.

Direct measurement of neutrino masses is performed by examining radioactive decays of tritium. This naturally occurring isotope of hydrogen has a half-life of only 12 years, and this means that many decays every second occur in even a gram of tritium. The distribution of beta decays, or energetic electrons produced when a tritium atom spontaneously disintegrates, is affected by neutrino mass.

The world's largest neutrino detector is at the South Pole, under more than a kilometre of ice. The IceCube Neutrino Observatory has thousands of sensors that are lowered into holes two kilometres deep made by powerful hot water drills. The phototubes are spread over a cubic kilometre of ice and send digital data to the ice surface above the array. The high energy neutrinos generate muons in the ice. So, do cosmic rays. In complete darkness, a kilometre below the surface, the ice is transparent, and the muons are seen to decay, producing light flashes. Data from a hundred two-kilometre-long strings of sensors, spaced a hundred meters apart, provides triangulation that enables the directions of the incoming neutrino events to be inferred. Essentially, we have a neutrino telescope.

Most neutrinos are atmospheric in origin, produced by cosmic rays in the earth's atmosphere, and rain down from above. But a few are cosmic neutrinos. We are confident of this because cosmic rays are blocked by the earth. Down going muons are of cosmic ray origin in the earth's atmosphere. But upgoing muons give light flashes that can only be generated by the highly penetrating cosmic neutrinos that interact with the earth's mantle and the deep and dark Antarctic ice layer.

Another type of neutrino telescope is based a kilometre below ground in a disused zinc mine under Mount Ikeno near the city of Hida in Japan. A huge cavity is filled with 50 tons of highly purified water, and the sides of the cavity are lined with some ten thousand phototubes. The principle here is that neutrino interactions with the water, while very rare, occur occasionally and produce fast-moving charged particles that generate flashes of blue light. The main goal of this experiment is to detect a supernova explosion in our galaxy. We may have to wait a while: the average time between galactic supernovae is expected to be a century or thereabouts. However, we are probably overdue for one, as the last supernova visible from the earth occurred in 1604 and is named after astronomer Johannes Kepler.

In fact, there was a detection in 1987 of 25 neutrinos from a supernova in the Large Magellanic Cloud, our nearest neighbour galaxy. SN1987a was seen in neutrinos some 3 hours before the optical light reached the earth. These were high energy neutrinos from collapse of the predecessor star that preceded the explosion which gives rise to the visible light from the supernova. The parent star was a massive blue supergiant star, known to be on the verge of exhaustion of its nuclear fuel. Its explosion resulted in a million-fold increase in brightness that lasted for a few days and gradually faded over the next year.

Three neutrino observatories participated in the discovery, two purified water detectors in mines, in Japan at Kamiokande near Hida in a disused zinc mine and in the US in a salt mine at Lake Eyrie, and in Russia in an underground neutrino scintillation telescope at Baksan in the Caucasus mountains. The water-based detectors measured more than a dozen antineutrinos, and the scintillator-based detector detected a few neutrinos, all from SN1987a. The high energy neutrinos interact with the water by scattering off electrons. The electrons are scattered up to the speed of light, and it is the electron neutrinos that dominate. The electron speed is faster than the speed of light in water, and a sort of sonic light boom is produced as the electrons slow down. This produces what we call Cerenkov flashes of blue light in the water, with a duration of nanoseconds, to give the observed signal. The exact moment of the stellar death and collapse was not predicted but was inferred after the optical detection.

The sun's fuel supply involves thermonuclear burning of hydrogen into helium. A by-product is neutrinos of a few MeV energy. These are the only particles that can directly escape from the core of the sun. These lower energy neutrinos were first detected from the sun in 1970, with an experiment involving 100000 gallons of cleaning fluid in a tank 1500m underground in the Homestake Gold Mine in South Dakota, by Ray Davis (Nobel laureate in 2002). Trace atoms of chlorine converted to argon in this experiment, which however found a shortfall of a factor of 3 in the number of neutrinos predicted. This was later predicted to be due to in-flight oscillations between electron neutrinos, produced by thermonuclear reactions in the centre of the sun, and the other types of neutrinos, muon neutrinos and tau. In this way, only one-third of the neutrinos arriving at earth from the sun would be electron neutrinos.,

The missing neutrino problem was resolved with the detection of the different neutrino types. These were first measured in an experiment monitoring 1000 tons of heavy water at the Sudbury Neutrino Observatory, 2km underground in a Canadian nickel mine, led by Art Macdonald in 2001 (Nobel laureate in 2015). In heavy water, a hydrogen nucleus is replaced by a nucleus of deuterium, a heavy isotope, with one more neutron, of hydrogen. Deuterium provides the critical ingredient for detecting all of the neutrino species emitted by the sun, again by Cerenkov light flashes in the water now from the disintegration of neutrons into protons and very fast electrons. Neutrino studies are now mainstream physics.