Session 20

Astrophysics

Sarah Symons, Naomi Banks,

Cheryl Hurkett, Derek Raine

Physics Innovations Centre
for Excellence in Learning and Teaching

πCETL (Leicester)

Department of Physics and Astronomy

University of Leicester

Contents

Welcome

Welcome to session 20, our final session. We've organised this one a little differently from the others: in this session four of the team from the Physics Innovations Centre for Excellence in Teaching and Learning at Leicester will write about one of their areas of interest.

Session Author

Sarah Symons, Naomi Banks, Cheryl Hurkett, Derek Raine, University of Leicester.

Session Editor – Tim Puchtler

Positional Astronomy

The historical beginnings of astronomy deal with the appearance of the sun, the stars, and the planets as seen from Earth. Ancient and classical civilisations developed co-ordinate systems for describing the positions of celestial bodies in order to record and, later, to predict events in the sky.

Some ideas from ancient times persist into modern life. The Egyptians developed the concept of the twenty-four hour day. From the Babylonians, who used a base 60 number system, we have inherited 360 degrees in a circle.

In this section, we look at how positions of celestial objects are described.

Learning Objectives

- Alt-azimuth co-ordinates
- RA and dec
- Great Circles
- The ecliptic

The Problem

You have a friend who is an avid bird-watcher, who is now becoming interested in astronomy. She understands she needs a different sort of telescope to look at stars, but she wonders whether she can save money by using the same tripod for both telescopes. Her bird-watching tripod is heavy duty and has a smooth pan-and-tilt head. Would you advise her to invest in a new tripod or keep her old one?

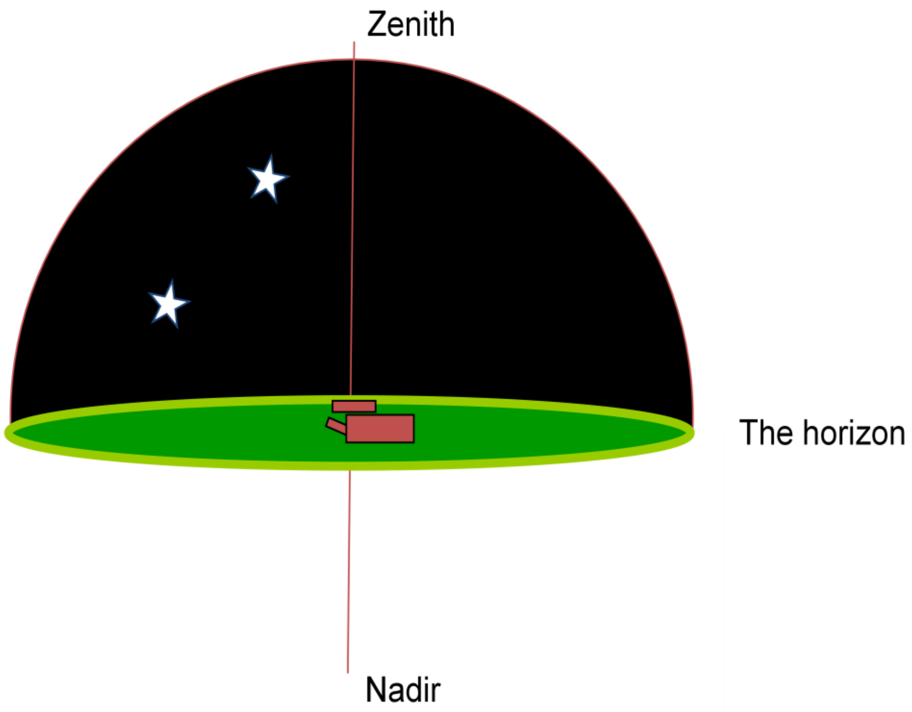
Image[1]

Pan and tilt

Imagine you are standing outdoors at night, on a flat plain. It's a dark night and you can see stars in every direction. Imagine a point on the celestial sphere directly above your head, a line from it, through you, and extending on through the Earth.. The point above your head is your zenith. The imaginary point below you on the other side of the celestial sphere is the nadir. Points on the horizon all around you are equidistant from the zenith and the nadir.

Now imagine setting up a telescope on your friend's pan-and-tilt tripod. If you lock the tilt with the telescope level, and just use the pan control, you can pan around the horizon a full 360 degrees.

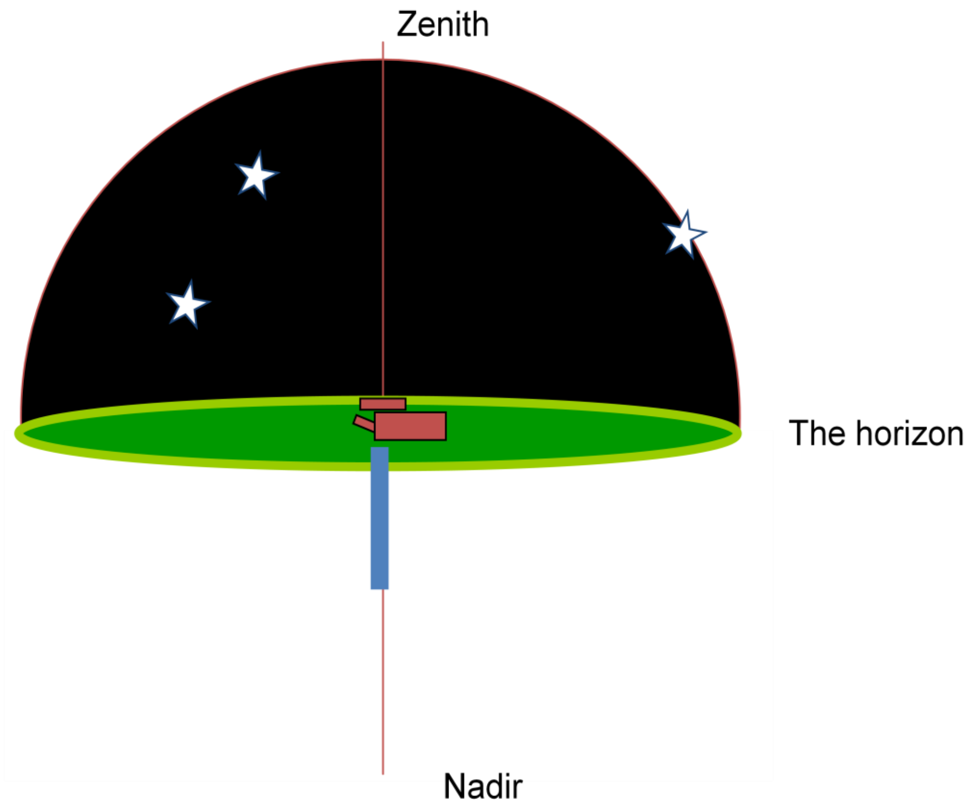
[1] WF Fancier 535 carbon fiber tripod, by F 5.6, as posted on www.flickr.com. Creative Commons Licensed



If you point the telescope due north, say, lock the pan, but release the tilt, you can tilt the telescope up from level (0 degrees) upwards until it becomes uncomfortable to look through the eye piece, or the tilt mechanism reaches its stops. You might even be able to point the telescope all the way up to the zenith (90 degrees).

Altitude and azimuth

These two degrees of freedom – pan and tilt – relate to a co-ordinate system which can be used for stars. "Azimuth" – corresponding to pan – ranges from 0 degrees (due north) through 90 degrees (east), south is 180 degrees, and west is 270 degrees. "Altitude" corresponds to the tilt of the telescope – from 0 degrees at the horizon to 90 degrees at the zenith.



If the tripod allows you to point the telescope towards any azimuth and any altitude, it means that you can view any star or planet that is above the horizon. But is there a better system? If you buy a (reasonable quality) astronomical telescope, you will not be supplied with a pan-and-tilt tripod. Let's see why.

The Pole Star

Let's return to the idea of pointing the telescope due north (0 deg azimuth) and tilting upwards from 0 deg altitude. As you tilt from due north on the horizon up to the zenith, you will notice the Pole Star "Polaris" in your field of view. Most people can find Polaris and know that it can be used to find due north. You may also know that if you point a camera at Polaris and take a long exposure frame, the surround stars will make circular trails on the sky, centred around Polaris, which seems to remain practically stationary. Why does Polaris have these special properties?

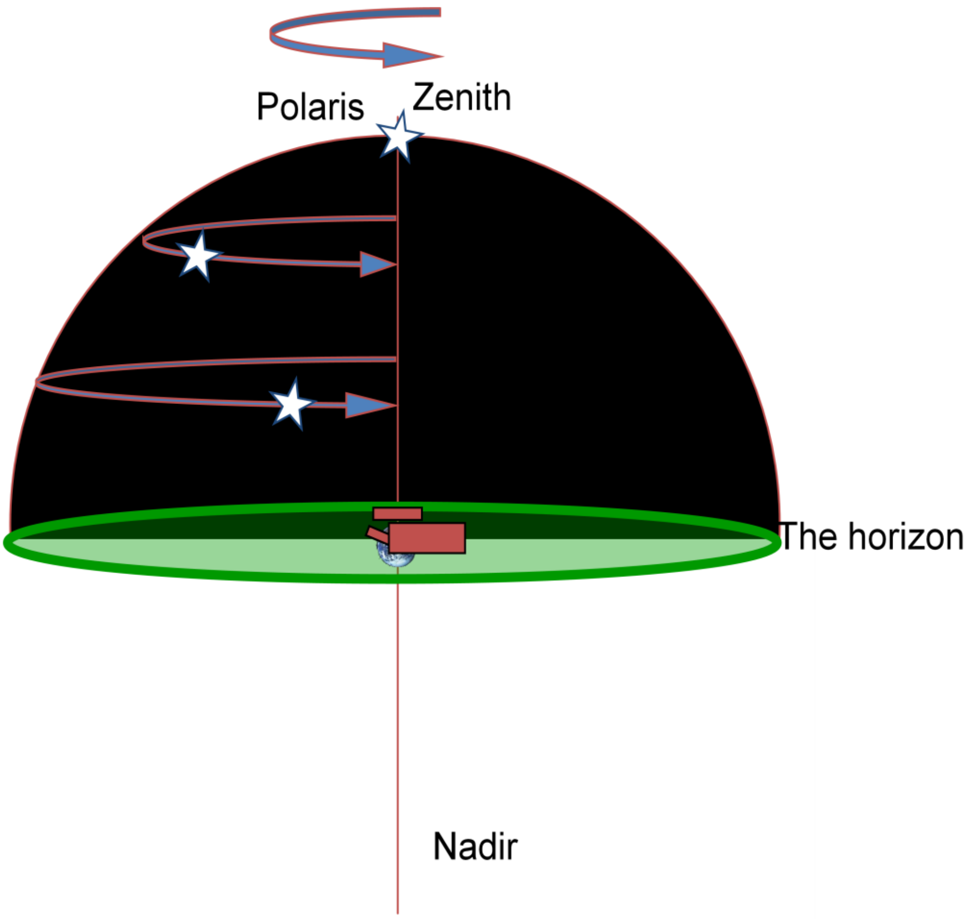
Image[2]

At the north pole

In fact, there is nothing unusual about Polaris itself. The behaviour is caused by the spin of the Earth. The Earth's axis of rotation in fact points at (or very near) Polaris.

Imagine that you are standing at the North Pole at night. The Earth's axis is running vertically through your body, and the Pole Star is directly overhead at your zenith. All the other stars appear (from your point of view) to rotate around the Pole Star, parallel with the horizon, never rising or setting. Each star has a fixed altitude, and travels 15 degrees of azimuth every hour.

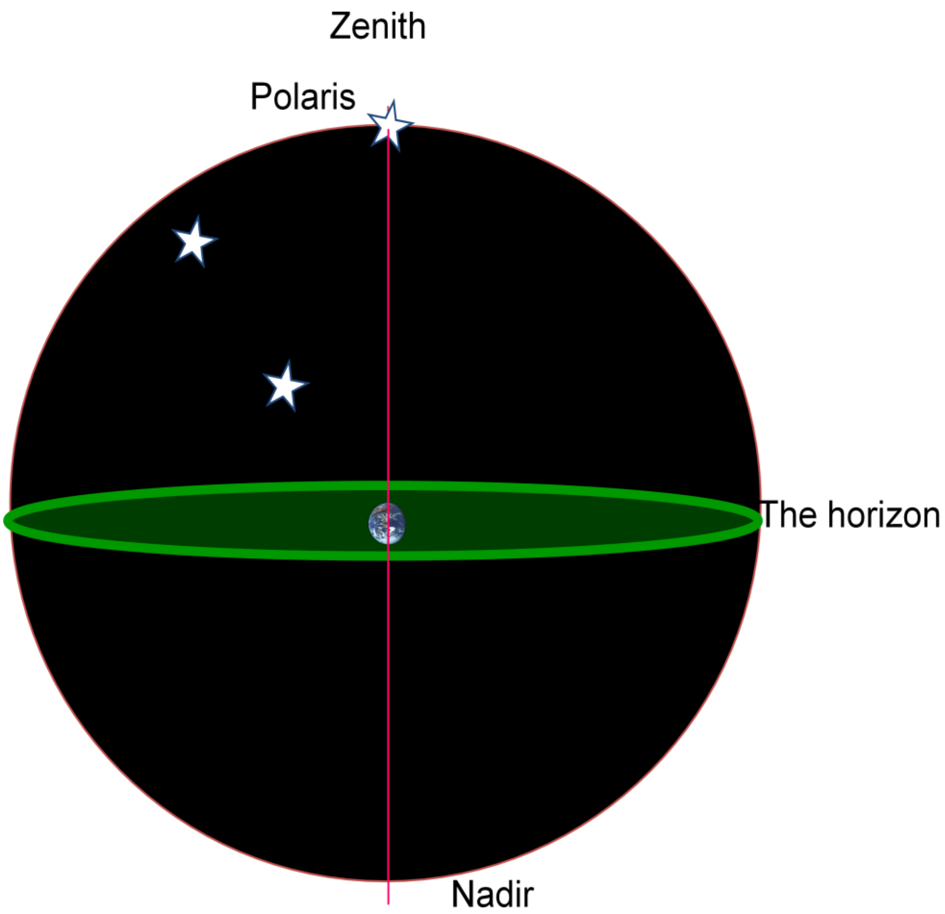
[2] Polaris and Firepit, by dkeros, as posted on www.flickr.com. Creative Commons Licensed.



If you set up your pan-and-tilt telescope at the North Pole and choose a particular star to observe, say Arcturus, you'll be able to lock the tilt to give you the correct altitude (about 19 deg altitude), and then just use the pan to track the star throughout the night. The tripod is ideal for this circumstance.

To warmer climates

Now let's leave the North Pole and walk south, all the way to the equator. While walking, let's think what is happening to the stars. As we move south, Polaris does not stay directly over head. It moves downwards (decreasing altitude) at the same rate that our latitude decreases. In fact, the altitude of Polaris always equals the latitude of the observer – one of those useful "survival" facts: if you ever find yourself on a desert island you now know how to calculate your latitude!



If you walked all the way to the equator, you would see that Polaris would disappear into the horizon, but let's not go that far, let's stop in Madrid. Polaris is at an altitude of around 40 degrees. Let's set up our telescope on its tripod and look at Arcturus again. We can find it by recognising the constellation, and then finding it in the viewfinder of the telescope. Let's observe it for a while.

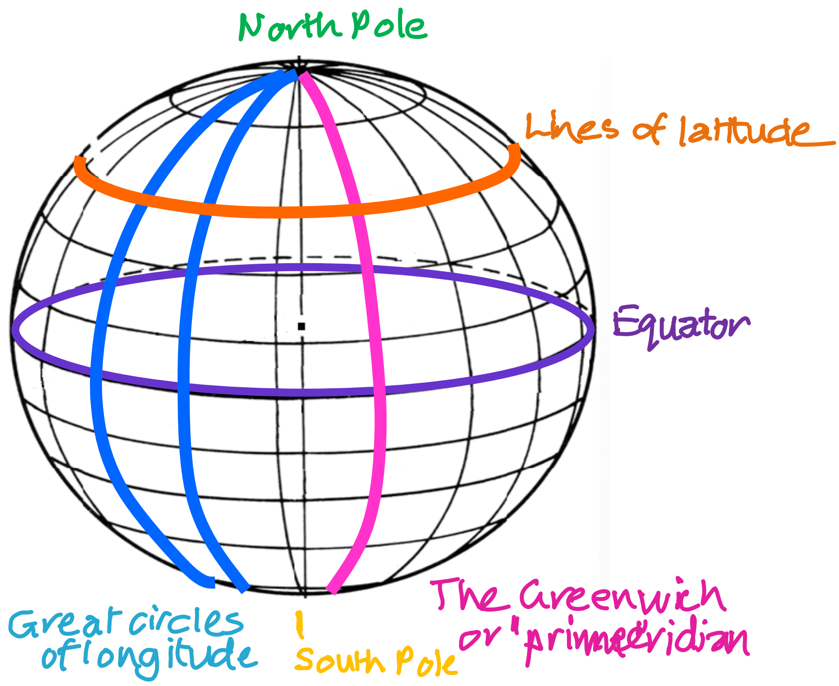
We find that when it moves, we have to use both the pan AND the tilt to keep Arcturus in the centre of the viewfinder. Why?

Arcturus is still in the same relative position to the Pole Star, rotating around Polaris in a circle 90 – 19 degrees away, but Polaris is not at the zenith – the whole sky is "tipped". The altitude and azimuth of each star except Polaris itself, is changing minute by minute. Our co-ordinate system, and our tripod mount, are no longer very useful to us. What other system can we use?

Terrestrial coordinates

Let's think of the system we use for describing location on Earth, latitude and longitude. This system uses the Earth's axis to give two points of reference: the north pole and the south pole. Next, we define the equator as the line of points on the surface of the Earth which are equidistant to the two poles. We designate points on the equator to have zero degrees latitude, and the poles themselves to be at 90 degrees of latitude north and south. All points in between the equator and the poles have a latitude between 0 and 90, either north or south. Lines of equal latitude form circles on the earth's surface which get smaller in diameter as they approach the poles. The equator itself is largest possible line of latitude, and (thinking of the Earth as perfectly spherical) is a circle of the same diameter of the Earth itself. This type of circle inscribed on a sphere is called a "great circle".

The other terrestrial co-ordinate, longitude, is also measured in degrees. Each line of equal longitude is a great circle which passes through both the north and south poles. We use 360 degrees to measure the full range of longitudes, usually express as 0-180 degrees east and 0 to 180 degrees west. Here, the zero point has to be chosen arbitrarily. For historical reasons, the line representing 0 degrees of longitude is the great circle passing through both poles and through Greenwich.



Celestial coordinates (RA, dec)

We can use a similar system for celestial co-ordinates by imagining the starry sky as a celestial sphere with a Pole near the Pole star, positioned precisely at the intersection of the celestial sphere with the Earth's axis, i.e. at the North Celestial Pole. The South Celestial Pole will be positioned opposite, and the celestial equator can then be defined as the line of points on the celestial sphere which are equidistant to the two poles.

We call celestial latitude "declination" written dec or delta, and celestial longitude is called "right ascension", R. A. or alpha.

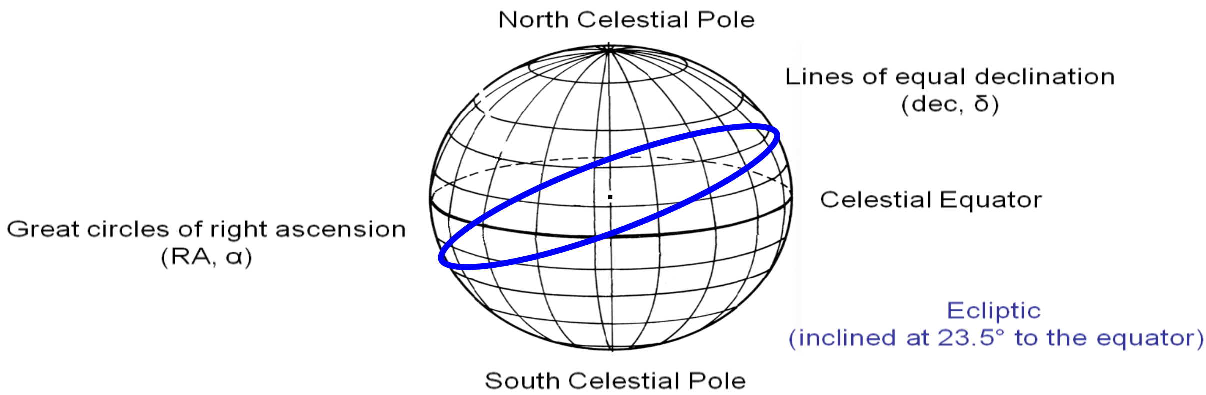
A star on the celestial equator has a declination of 0 degrees. Stars to the south have negative declinations, and those to the north have positive ones. Declination therefore runs from -90 through to +90 degrees. Degrees of declination can be subdivided into minutes and seconds to express greater accuracy.

Right ascension can be measured in 360 degrees, but is often expressed in "hours" – 0 through to 24. The conversion is easy: 1 hour equals 15 degrees. Degrees of right ascension can also be subdivided into minutes and seconds to express greater accuracy.

The zero point for right ascension (analogous to the Greenwich meridian) is not arbitrarily chosen. The mean path of the sun, called the ecliptic, is a great circle which is inclined when compared with the celestial equator and therefore intersects the equator twice.

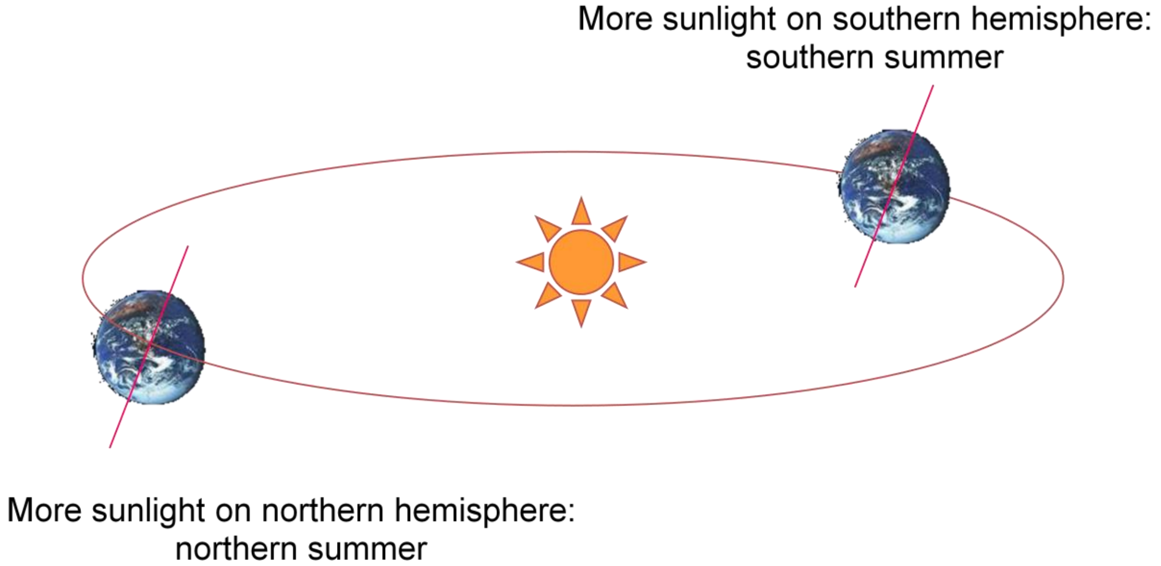
The sun lies on one of these points at the spring equinox and when it does, it has right ascension 0 and declination 0, marking the origin of the RA and dec coordinate system.

Let's digress for one moment to look at the ecliptic.



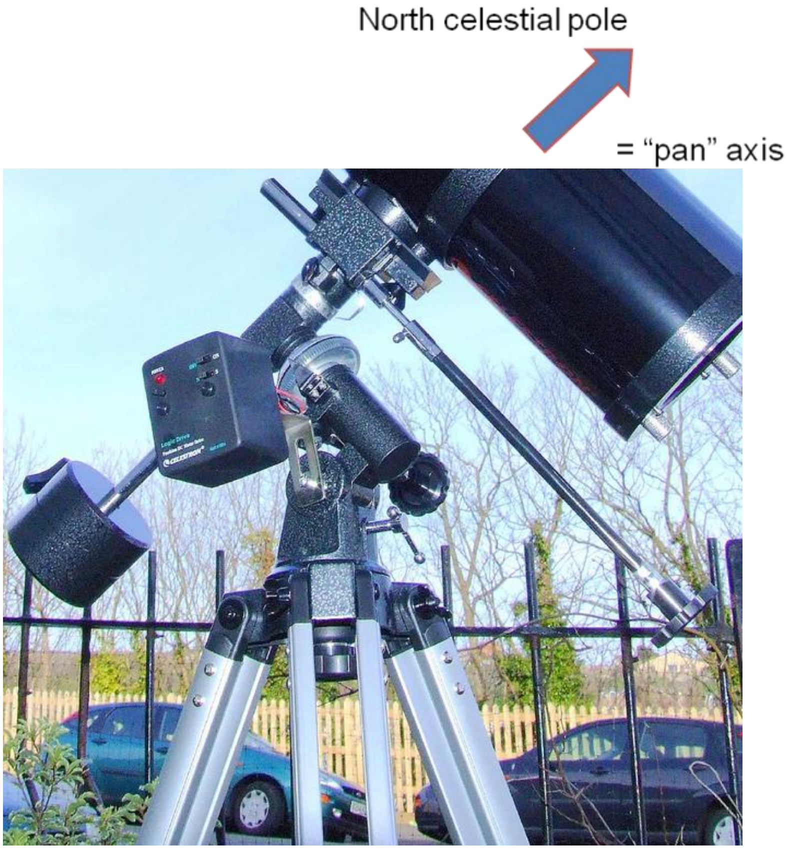
Seasons

The ecliptic is inclined because the axis of the earth is inclined with respect to the plain of its orbit around the Sun. The effect of this inclination is that the Earth has seasons. In summer, the nights are shorter and the day is longer. The extra sunlit time causes warmer weather.



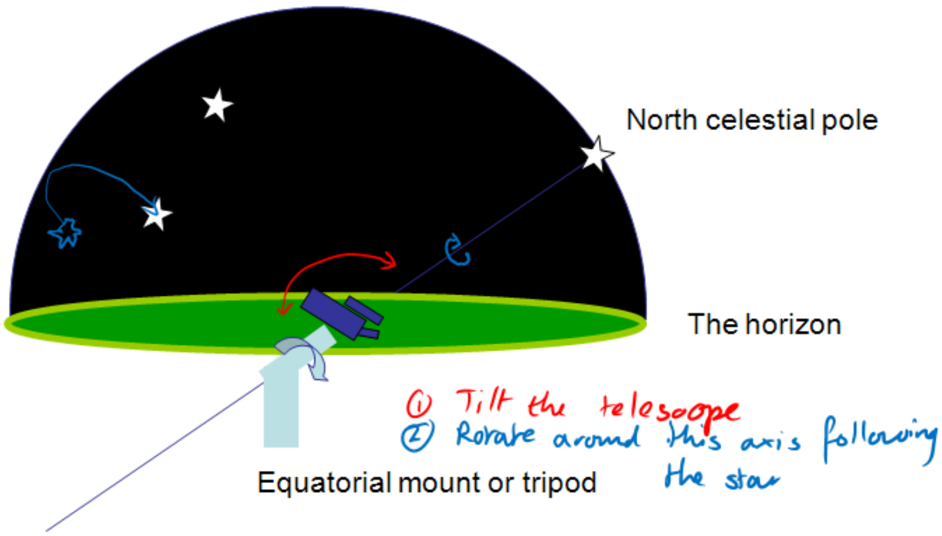
Note that the seasons do not occur because the Earth is "closer to the Sun" in summer. This is a common mis-conception and is wrong. If it were true, both the northern and southern hemispheres would experience summer at the same time.

Equatorial tripod



Image

Let’s return to RA and dec. A star’s RA and dec coordinates do change over time, but for amateur astronomy they can be considered constant. The equatorial mounting system for astronomical telescopes uses this to provide a tracking mechanism which (like the pan-and-tilt tripod at the North Pole) can track a star by rotating on just one axis.



The tripod head is set at an angle to the horizontal so that the axis of "pan" rotation points to the North Celestial pole. A motorised telescope mount contains a small but very smooth and accurate motor which rotates the telescope at precisely the same rate as the stars appear to move. Once the telescope is pointed at a target, it will track automatically, keeping the object in view.

So our advice to our bird watching friend is easy – unless you want to observe in the Artic, we strongly recommend an equatorial mount for your new telescope!

Summary

To identify star positions we use coordinates of RA and dec on the celestial sphere, which play the role of longitude and latitude on the Earth

SAQs

1. At a latitude of 50^0 N where is the North Celestial Pole?

   (a) At an altitude of 50^0

   (b) At an altitude of 40^0

   (c) Overhead

2. (a) True or (b) false?

   (i) The Sun traverses the ecliptic once a day

   (ii) The RA of a star changes by 24 hours over the course of a day

   (iii) The seasons are caused by the wobble of the Earth’s axis

The answers appear on the following page

Answers

1. (a) Correct: the pole star moves down from the zenith as the observer moves away from the pole to lower latitudes

   (b) Incorrect: Altitude is measured up from the equator just like latitude

   (c) Incorrect: the pole star appears to move down from the zenith (overhead) as the observer moves away from the pole to lower latitudes

2. All False

   (i) The ecliptic is the path traced out by the Sun over a year

   (ii) The RA of a star is fixed. This is the point of the celestial coordinate system. (Small changes in RA do occur over very long times as a result of precession of the Earth's axis.)

   (iii) The Earth's axis wobbles only on time scales of thousands of years. To all intents and purposes the axis is fixed in space. The seasons are caused by the changing inclination of the Earth's axis to the Sun throughout the year.

Stars

Learning Objectives

- Describe the Hertzsprung-Russell diagram
- Describe what is meant by the main sequence
- Explain qualitatively why most stars lie on the main sequence
- Describe the stages of evolution of stars
- State the end points of stellar evolution

The Problem

Star Formation

We start with the formation of stars. We show two well known regions of star formation: the Horsehead nebula and the Orion region.

Image[3]

Image[4]

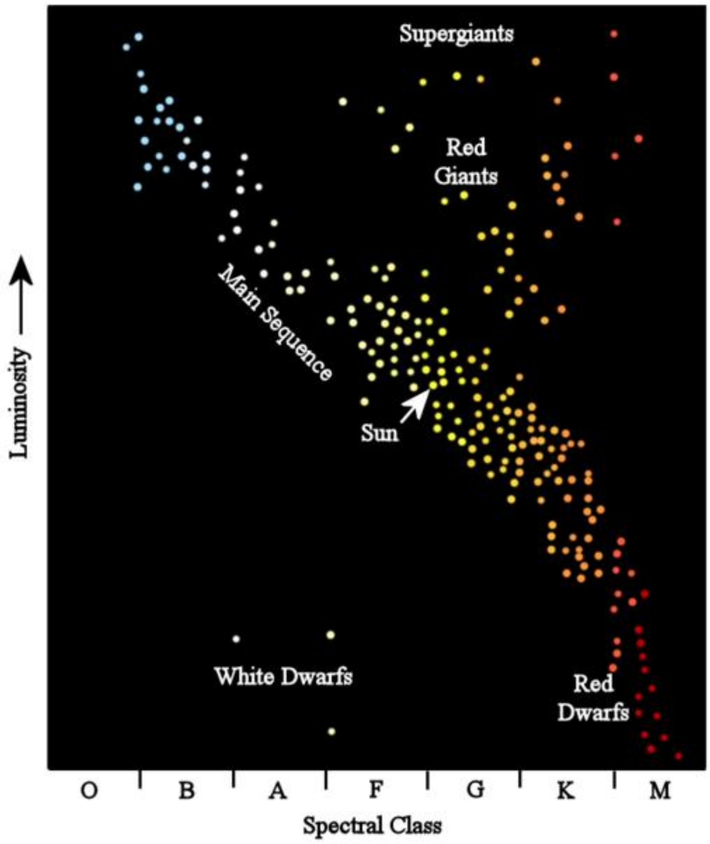
Here clouds of interstellar gas and dust are collapsing under their own gravity. The dark regions are the result of absorption of light by heavy elements (called dust by astronomers). Eventually the clouds will break up into objects that are hot and dense enough to undergo nuclear fusion producing their own light. These objects are stars. The gas and dust that are absorbing starlight will be blown away and the stars themselves will become visible.

The Hertzsprung-Russell or H-R Diagram

[3] A reproduction of a composite colour image of the Horsehead Nebula and its immediate surroundings, European Southern Observatory, as posted on commons.wikimedia.org. Creative Commons Licnesed.

[4] Hubble Panoramic View of Orion Nebula Reveals Thousands of Stars, NASA Hubble Space Telescope Collection, Nase, ESA, T. Megeath, M.Robberto, as posted on nasaimages.org.

Not all stars are the same. They have different masses and radii, and shine with different luminosities at different surface temperatures. The figure below shows a plot of luminosity against temperature for a representative sample of stars. This is one of the most important diagrams in astrophysics: it is called the Hertzsprung-Russell diagram, or H-R diagram for short, after the two astronomers who independently discovered it in the early 20th century. The important and obvious feature of the diagram is that the stars are not spread uniformly across the area. Apparently, stars tend towards certain combinations off luminosity and temperature. In fact most stars, some 90%, lie on the main sequence, and most of the others in the white dwarf region. Why is this? We'll look at the answer to this question in this section.

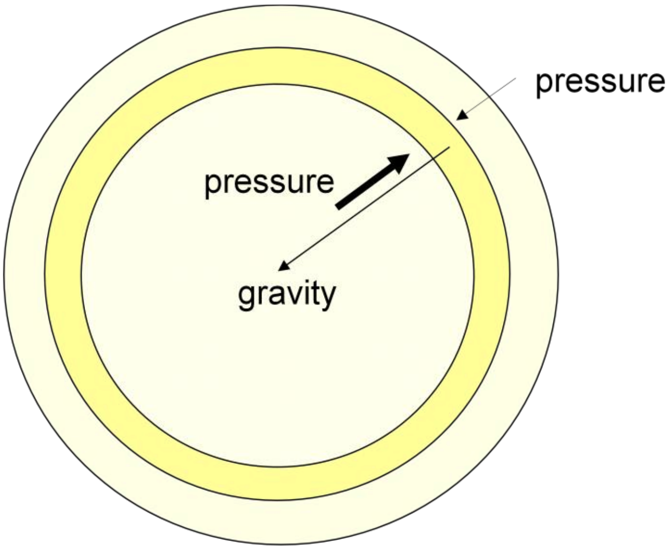


Image[5]

[5] Hertzsprung_Russell diagram, by RJ Hall, as posted on commons.wikimedia.org. Creative Commons Licensed.

Stellar Structure – Holding it up

We will explain the H-R diagram by looking at the physics of stars: what holds them up and what makes them shine. We'll do this qualitatively by looking at the relationships between the various physical quantities.



Hydrostatic Equilibrium: The pull of gravity on the shell is balanced by the difference in gas pressure P on the inner and outer surfaces.

Star: mass M, radius R

Let's start with the way a star is held up against into own gravity. At each radius the pressure must be sufficient to hold up the weight of the star above it. So the pressure gradient is related to the pull of gravity by the underlying material. This means that the pressure at any point must change if we alter the mass of the star or its radius for the star to remain stable. So Pressure is a function of mass and radius. We can get a quantitative estimate of the central pressure if we recall that we can think of pressure as an energy per unit volume (as well as a force per unit area). The energy in question here is the gravitational potential energy GM^2/R and the volume is $4/3\pi\ R^3$, which, leaving out numerical factors, gives the expression:

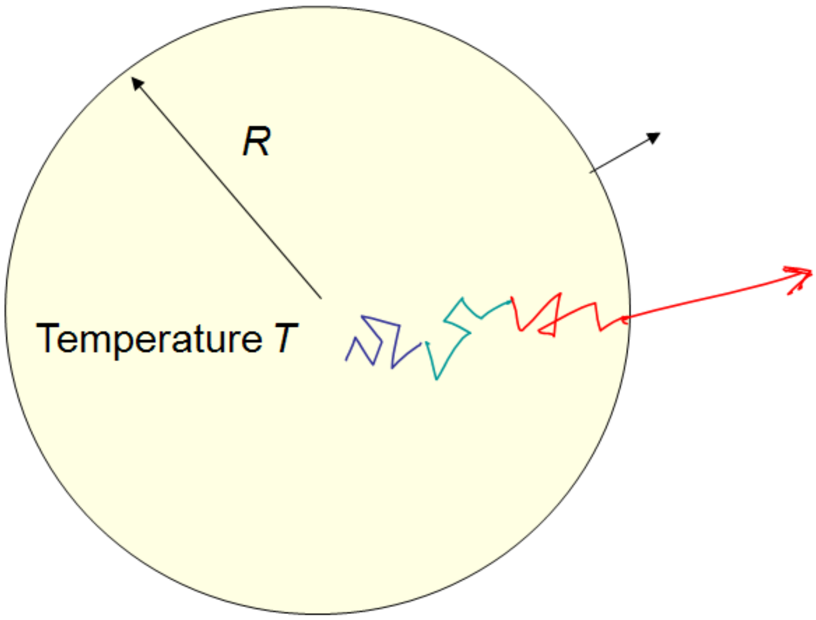
$$P \approx \frac{GM^2}{R} \times \frac{1}{R^3} \qquad \textit{(P is a function of M and R)}$$

The flow of energy

Now to how a star shines. There are two aspects to this. One is that radiation flows out of the star from the hot centre of the star to where it leaves the surface. This happens whether or not energy is being generated in the centre so long as the surface is cooler than the centre. The second aspect is the generation of radiant energy in the centre to replace that being lost. We'll take each in turn.

Luminosity L depends on the temperature gradient (Fick's Law)

Luminosity = energy per second



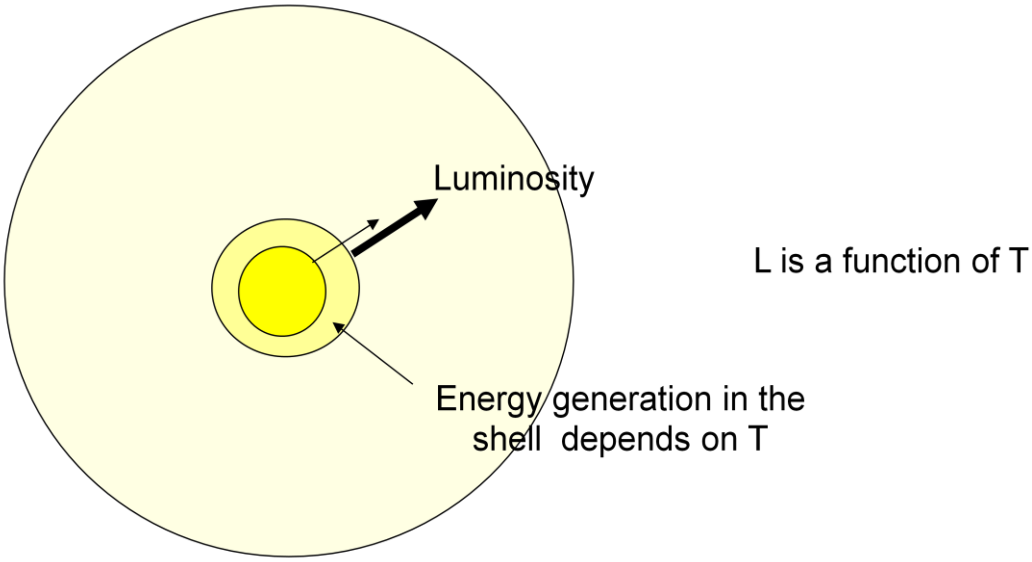
Because the radiation is absorbed and reemitted as it flows though the star, photons follow a random walk through the star, so this is a diffusion process. In order for there to be a net flow outwards, there must be a temperature gradient. The radiation then flows down the temperature gradient. Thus the energy leaving the star each second will depend on the central temperature and the radius. It will also depend on the area of the surface, but this again depends on the radius. This gives us the proportionality:

$$L \propto 4\pi R^2 \times \frac{T}{R} \qquad \text{(L is a function of T and R)}$$

The central temperature of a star is in the region of 10^7 – 10^8 K. Explain why the luminosity does not depend on the surface temperature.

Energy generation

Finally, for the star to continue shining, the rate at which energy is being lost must be balanced by the energy that is being generated by nuclear reactions in the centre of the star.



The rate of nuclear reactions depends on the density of material, but mainly on the temperature. So the luminosity also depends on the temperature through the nuclear reaction rate. The fact that the single quantity, the luminosity is determined separately both by the nuclear reactions and by the temperature gradient puts a constraint on the internal structure of a star. The temperature must be just right so that the two ways of determining the luminosity agree. This is basically why the temperature and luminosity can't take just any values in the H-R diagram. The next section sets out the argument in more detail.

Explaining the H-R diagram

Putting this together we can say that in order to hold up the star against gravity the gas pressure at the centre must be right for the mass and radius of the star. This is step 1.

Hydrostatic equilibrium $\Rightarrow$

P is a function of M and R (1)

Gas law $\Rightarrow$

P is a function of T (2)

T is a function of M and R (eliminate P) (3)

Flow of energy $\Rightarrow$

L is a function of T and R (4)

Eliminate T $\Rightarrow$

L is a function of M and R (5) L = f(M,R)

Energy generation $\Rightarrow$

L is a function of T (6)

Eliminate T $\Rightarrow$

L is a function of M and R (7) L = g(M,R)

Eliminate R $\Rightarrow$

$\Rightarrow$ L is a function of M

Eliminate M $\Rightarrow$

$\Rightarrow$ R is a function of L

Stefan's law $\Rightarrow$

L is a function of R and $T_{surface}$ $\Rightarrow$ L is a function of $T_{surface}$

Step 2, the gas laws tell us that the pressure will depend on the temperature, so, step 3, the *central temperature* must also be right for the mass and radius. Next, step 4, the luminosity flowing from the star depends on the central temperature and radius so, step 5, it must also be right for the mass and radius. Finally, step 6, the luminosity is also determined by the central temperature, through the nuclear reactions, so, step 7, L is determined in a different way by the mass and radius. That there are these two ways of determining the same L from M and R, means that the mass and radius must be just right to give the right luminosity: once the mass, say, is known, the radius is fixed. So the luminosity is a function of only one parameter, the mass of the star. Stars must therefore fall on a line in any plot of luminosity. This explains the reason behind main sequence: it shows stars of different mass, from low mass at the bottom right to high mass at the top left.

Energy generation

The pp chain

To explain stars that do not lie on the main sequence we turn to how the energy of stars is produced. This shows one of the main nuclear reactions in the centre of the Sun:

$$^{1}H + {}^{1}H \rightarrow {}^{2}H + e^{+} + \upsilon_{e} + 0.42\ \text{MeV}$$

$$e^{-} + e^{+} \rightarrow 2\,\gamma + 1.02\ \text{MeV}$$

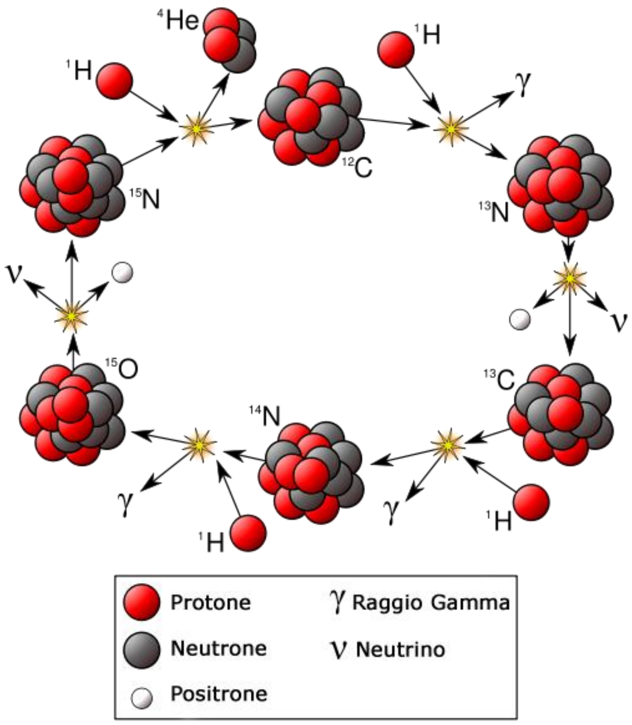
$$^{2}H + {}^{1}H \rightarrow {}^{3}He + \gamma + 5.49\ \text{MeV}$$

$$^{3}He + {}^{3}He \rightarrow {}^{4}He + 2\,{}^{1}H + 12.86\ \text{MeV}$$

Hydrogen is converted to helium by a number of routes collectively called the p-p chain. In stellar physics the consumption of hydrogen as a fuel is referred to as hydrogen burning, and likewise for other fuels.

The CNO cycle

In stars more massive than the Sun the central temperature is higher and the main energy generation is via the CNO cycle, so called because carbon, nitrogen and oxygen act as catalysts for the reaction.



Image[6]

Overall however, this is just an alternative way of converting hydrogen to helium. Below we have written the sequence of reactions. We also show a comparison of the rates of energy generation by the pp chain and CNO cycle at various temperatures. Note how the energy generation depends steeply on temperature: roughly T^4 for the pp chain and T^{17} for the CNO cycle.

$^{12}C + {}^{1}H \rightarrow {}^{13}N + \gamma +1.95$ MeV

$^{13}N \rightarrow {}^{13}C + e^{+} + \upsilon_e + 2.22$ MeV

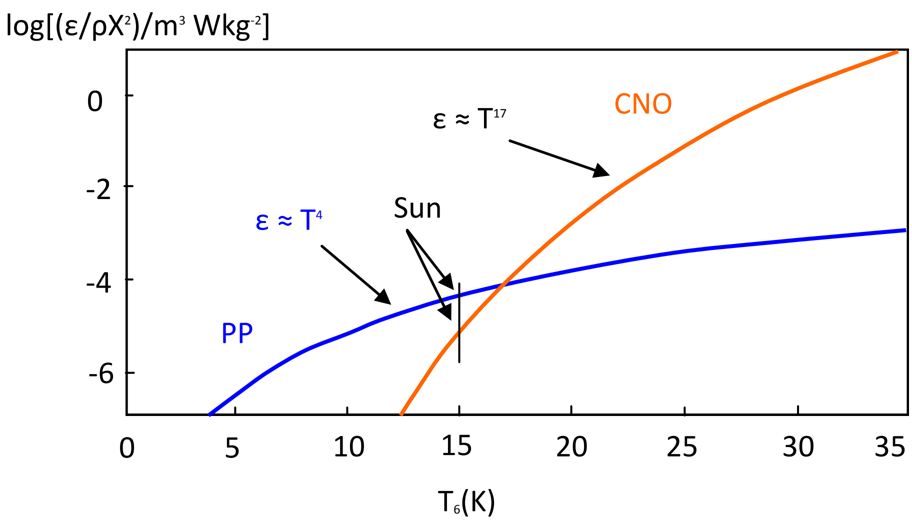
$^{13}C + {}^{1}H \rightarrow {}^{14}N + \gamma +7.54$ MeV

$^{14}N + {}^{1}H \rightarrow {}^{15}O + \gamma + 7.35$ MeV

$^{15}O \rightarrow {}^{15}N + e^{+} + \upsilon_e + 2.75$ MeV

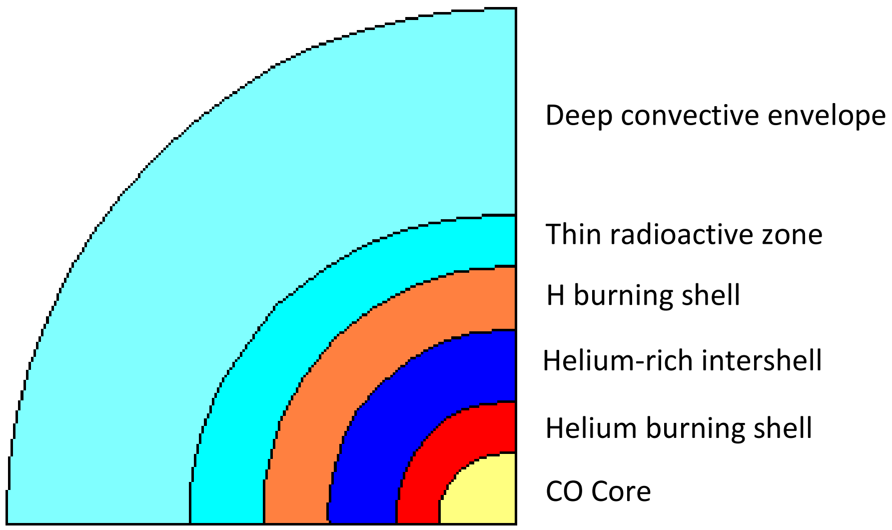
$^{15}N + {}^{1}H \quad {}^{12}C + {}^{4}He + 4.96$ MeV

[6] CNO_Cycle, by HeNRyKuS, as posted on commons.wikimedia.org. Public Domain.

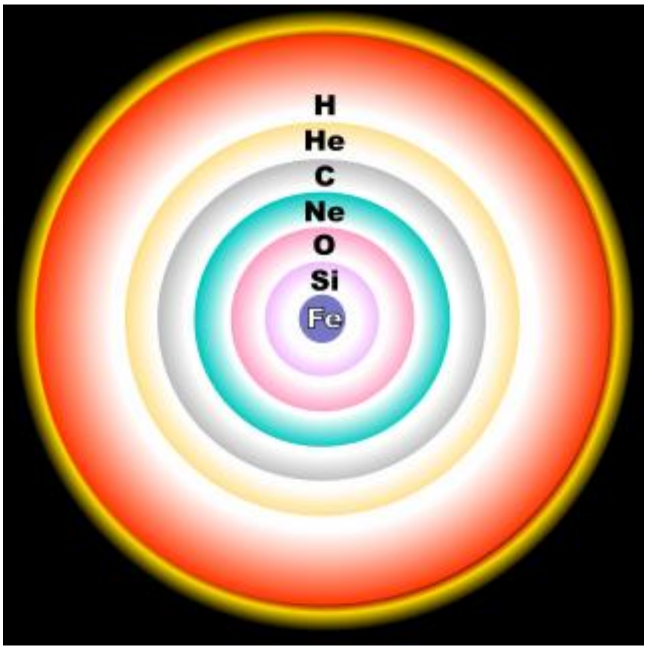


Star death: What happens when the energy runs out?

Once a star has run out of hydrogen at its centre, the central core will shrink. Part of the loss of gravitational potential energy goes into heating up the stellar material. This has several effects. First, the shell of material around the core becomes hot enough to ignite hydrogen burning, known as hydrogen shell burning. Second, the core itself becomes hot enough to ignite helium burning converting helium to carbon. Finally, the potential energy lost by shrinking the core goes partly into heating the core material, but also into gravitational potential energy of the outer layers, so the star expands into the red giant region of the H-R diagram.



In lower mass stars, the process stops at this point where the star runs out of helium in the core. The star then cools and shrinks to the white dwarf region of the H-R diagram. Here, at high densities, electrons behave quantum mechanically, just as they do in metal. The energy of the electrons provides a pressure to hold the star up against gravity that is independent of the temperature. So the star just continues to cool. These stars are known as white dwarfs.



Image[7]

When higher mass stars run out of helium at their centre, the core then shrinks and, in a repeat of the previous process, heavier element burning starts. This can repeat with progressively heavier elements creating "onion" shells of nuclear burning. This doesn't happen in lower mass stars as the gravitational potential energy is insufficient to heat the core enough for higher element burning to begin. Once the core has been converted to iron no more energy can be extracted. At this point a sudden collapse of the core results in the transfer of energy to the outer layers, which are blown off in a supernova.

Interactivity: In the hottest central region the reactions can proceed all the way to iron. Why not to heavier elements?

The effect of nuclear reactions, apart from supplying energy, is to change the chemical composition of the stellar material. It is this that drives the star to different regions of the H-R diagram.

[7] Evolved Star Fusion Shells.svg, RJ Hall, as posted on commons.wikimedia.org. Creative Commons Licensed.

Formation of the elements

How are elements beyond Fe produced?

A supernova returns stellar material to the interstellar medium where it can be reused to form new stars. This means that each generation of stars is increasingly enriched with heavier elements. In the supernova itself neutrons are released from the stellar core as most of the iron nuclei are broken down by the high temperatures, around 10^9 K, in the collapsing core. These neutrons react with the stellar debris to produce the elements beyond iron in the periodic table. All elements above lithium have been produced in stars or supernovae.

Image[8]

Crab nebula: the result of a supernova in 1054

[8] The Crab Nebula, European Southern Observatory, as posted on commons.wikimedia.org. Creative Commons Licensed.

The remnant of the stellar explosion is the result of the collapse of the core of the star. It therefore has a high density. In fact the density is of the order of that of an atomic nucleus. At this density neutrons are stable – in fact more stable than protons in an overall neutral material. So the remnant is an object made of neutrons, or a neutron star. This is held up against gravity by the pressure of the neutrons… unless the mass of the object is so large that this pressure is insufficient. In that case, nothing can support the star against gravity and the result is a black hole.

Final stages

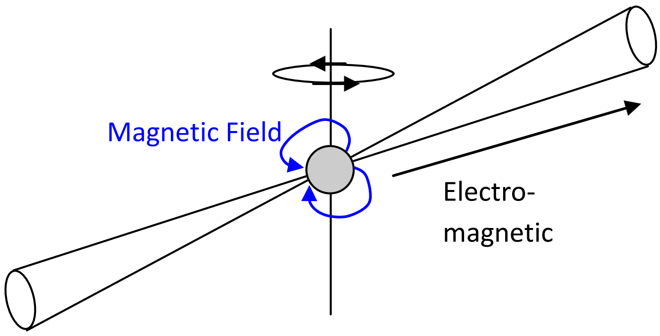
So to summarise the final stages: What is left behind as a result of stellar evolution depends on the mass of the star. If the star is less than 1.4 times the mass of the Sun, the final stages of stellar evolution yield a white dwarf. For more massive stars the result is a neutron star or a black hole depending on the mass of the core. The dividing line is not known exactly because of the difficulty of computing models of neutron stars that take account of the interactions between neutrons.

Mass < 1.4 solar masses -> white dwarf

Remnant < 2-3 solar masses -> rotating neutron star (pulsar)

Remnant > 3 solar masses -> black hole

A pulsar's 'Lighthouse effect' can be represented as shown below:



Summary

- Stars are formed from collapsing gas clouds in interstellar space
- The Hertzsprung-Russell diagram is a plot of stellar luminosity against colour or temperature
- The majority of stars lie along a curve in the H-R diagram called the main sequence
- The main sequence can be explained by combining the condition of hydrostatic equilibrium, the perfect gas, the diffusion of radiation and the rate of nuclear energy generation
- When a star runs out of nuclear fuel in its centre the core shrinks, heats up and may ignite new nuclear reactions
- Nuclear reactions generate the elements heavier than hydrogen and helium
- The end points of stellar evolution are white dwarfs, neutron stars and black holes depending on the mass of the star.

SAQs

1. A star has a mass M, radius R. Its central pressure is of order

   (a) GM/R^2

   (b) GM^2/R

   (c) GM^2/R^4.

2. A solar mass stellar core shrinks from 1 solar radius to the radius of a neutron star, about 10 km. How much gravitational potential energy is released? Give your answer as the power of 10 in Joules.

The answers appear on the following page

Answers

1. (a) Incorrect: this is the gravitational force at the surface.

   (b) Incorrect: This is the gravitational potential energy

   (c) Correct: Pressure is an energy per unit volume so it is of order the gravitational potential energy divided by R^3

2. The initial radius is so much larger than the final radius that we can treat this as a collapse from infinity. So the energy released is $GM^2/R \sim 3 \times 10^{46}$ J (accept 46 and 47)

The Violent Universe

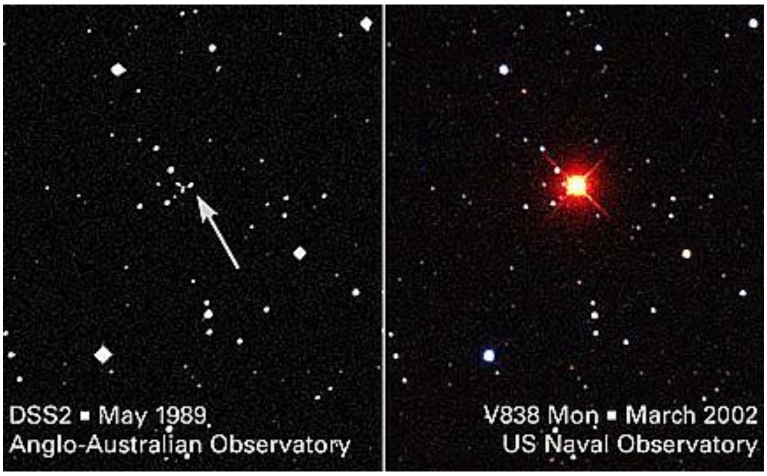
In this section we'll address the problem of how high energy radiation is produced in violent astrophysical events. First we'll set up the problem by looking at some violent events involving high energies.

Learning Objectives

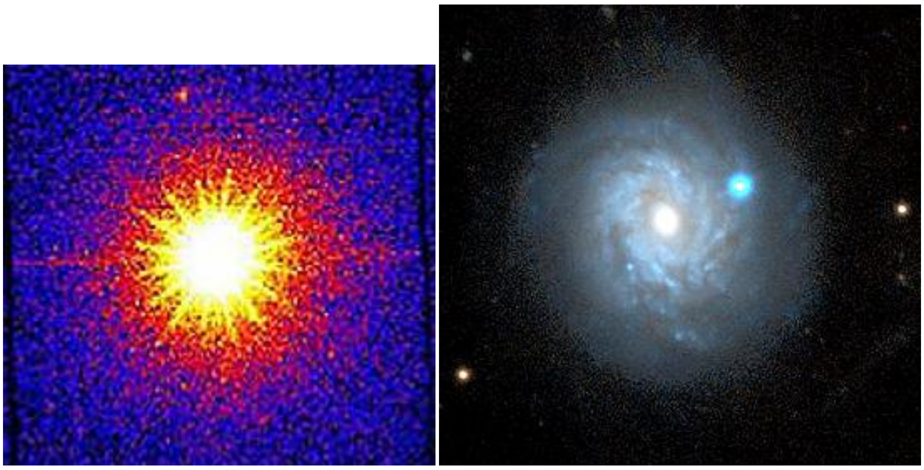
- Give examples of violent events in the Universe and the energies involved
- Describe the structure of a quasar
- Describe gamma ray bursts
- State what is meant by synchrotron radiation and inverse Compton radiation
- Distinguish between thermal and non-thermal emission

Novas

One would be forgiven for looking up at the night sky and thinking that the Universe is a placid, unchanging place. Far from it! Whilst the majority of stellar evolution occurs at a sedate pace there are many astrophysical phenomena that happen on extremely short timescales and release large quantities of energy over the whole of the electromagnetic spectrum. Below we have placed some images of stars called novas that flare up suddenly in optical light as well as a nova captured flaring up in X-rays.



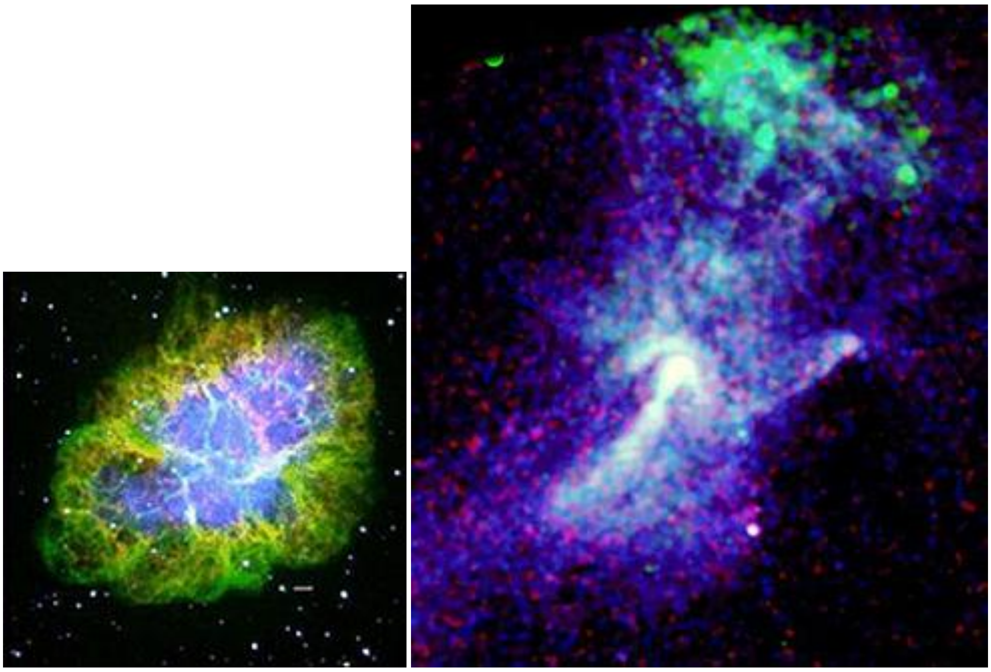
A Nova[9]



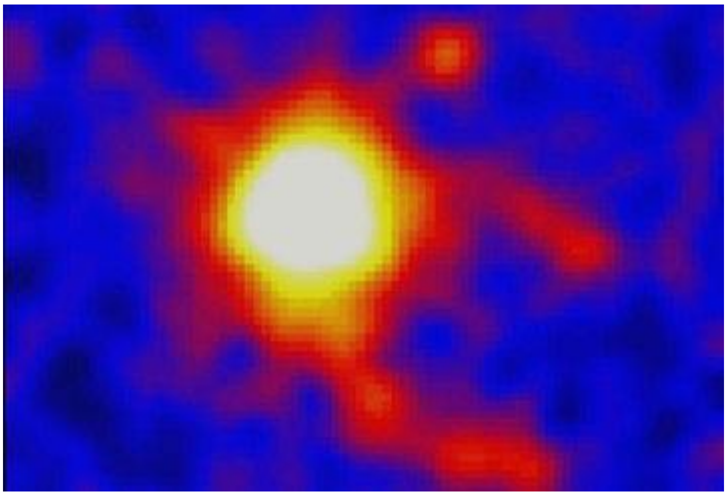
A false colour image of an X-ray nova (left), Supernova in the galaxy NGC3982 (right)[10]

The Crab Nebula is the remains of an event first observed in 1054 A.D. as the sudden appearance of a bright new star within the constellation of Taurus. Modern telescopes show filaments streaming at speeds up to half that of light from what has turned out to be an exploding star or supernova. This Supernova is an extremely energetic event emitting radiating from short wavelengths (gamma rays) through the visible and into the long wave radio region. A Pulsar, which is a rotating neutron star remnant from the explosion, lies in the central region rotating at 30 times a second.

[9] A Nova. Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/v838mon_1.jpg
[10] Images property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/nova1.jpg and http://rst.gsfc.nasa.gov/Sect20/NGC3982.jpg respectively.



The debris from a star in the Crab that exploded in 1054 leaving behind a central pulsar (left) and a pulsar showing gas illuminated by the jets of X-rays (right)[11]



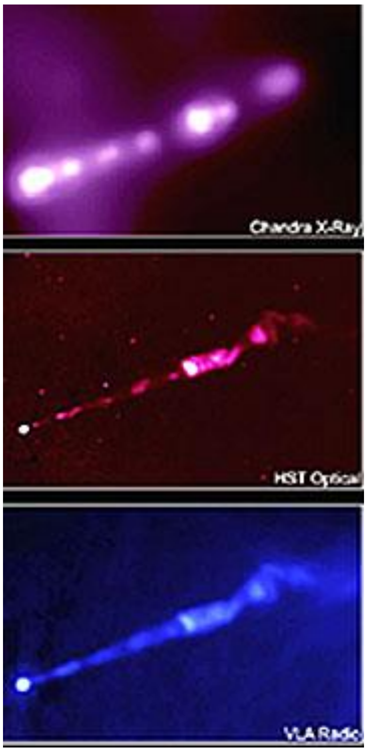
A Gamma ray image of the neutron star, Geminga, 500 light years beyond the Solar System[12]

[11] Images property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/9909282.jpg and http://rst.gsfc.nasa.gov/Sect20/chandravolt.jpg respectively

[12] Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/040715_hot_spot_A_02x.jpg

Supermassive black holes

Here we see energetic phenomena associated with galaxies. Three views of the jet from the galaxy called M87 shows it emitting in the X-ray region, in visible light, and in Radio waves.

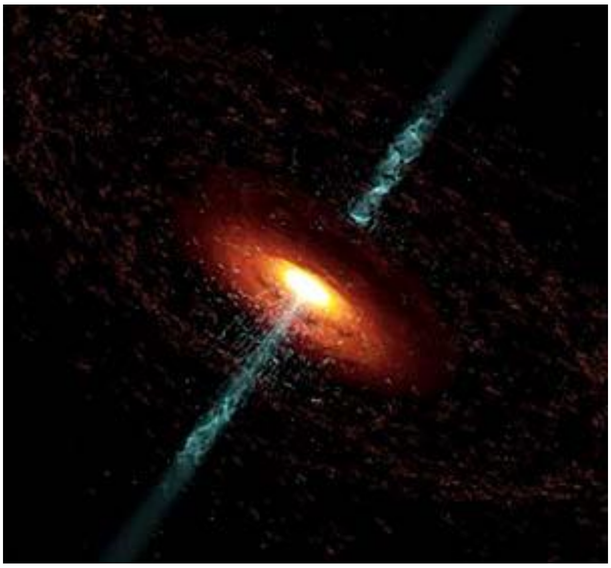


The M87 Galaxy observed in 3 electromagnetic regions.[13]

These jets are thought to have their origin in the strong, directional electromagnetic fields that surround a supermassive black hole in the centre of the parent galaxy.

We have also shown below an image of a paired jet associated with a supermassive Black Hole in the quasar called 3C120, captured by the Hubble Space Telescope.

[13] Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/m87-2.jpg



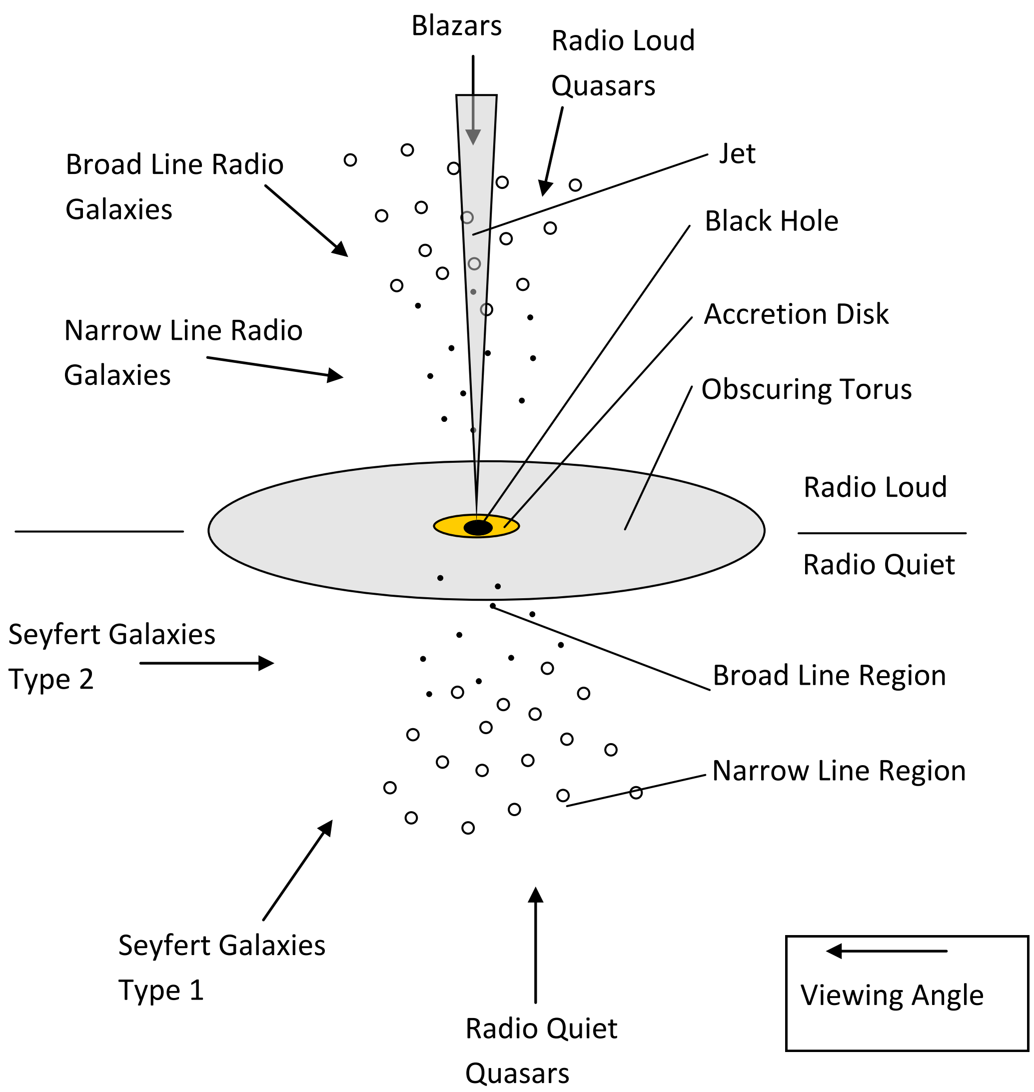
Quasar 3C120[14]

Quasars are galaxies with point-like sources of high energy at their centres. The jet is made of electrons, which emit from radio to X-rays as they follow lines of strong magnetic field.

Types of "active galactic nuclei"

Galaxies with supermassive black holes at their centres appear in various guises depending on the viewing angle:

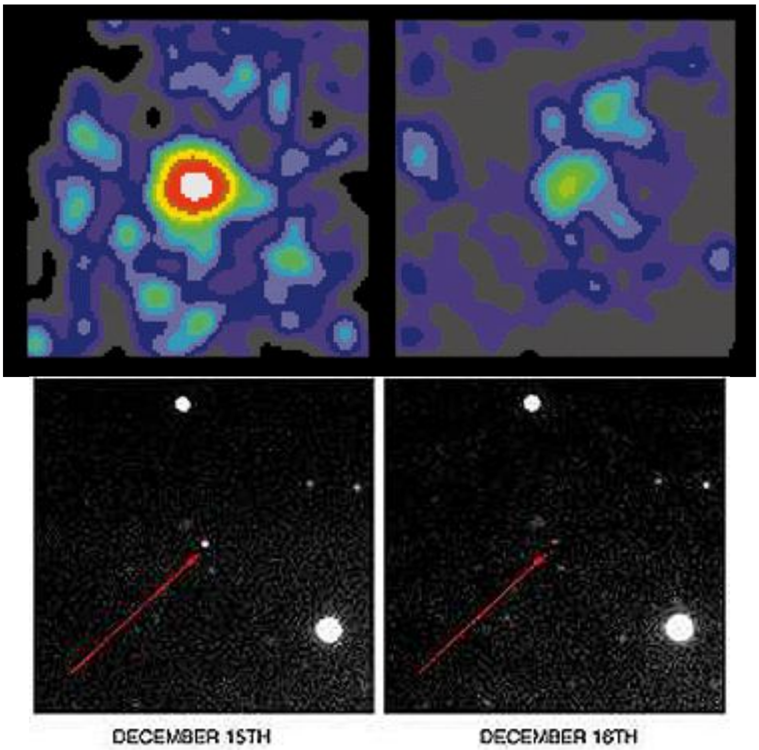
[14] Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/h_accretion_disk_02.jpg



The innermost region of the diagram shows a disc of material around a black hole. The accretion of this material by the black hole leads to the emission of radiation as gravitational potential energy is converted into heat. In some cases material is also blown out in a jet. Various structures surround the black hole giving rise to optical emission lines. Seen edge on the innermost region is obscured by a torus of gas.

Gamma Ray Bursts

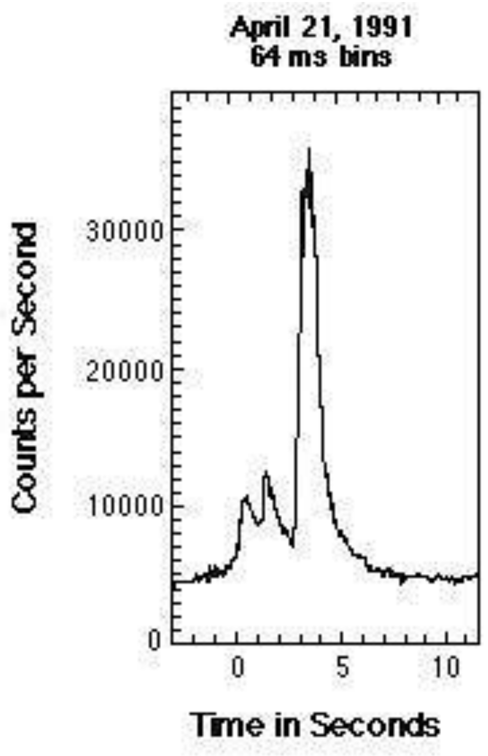
A recently discovered type of energetic event are the Gamma Ray Bursts. These are my area of research. They illustrate many of the features of high energy events and we'll look at them in more detail in the next few sections. They are extragalactic events that occur, on average, once per day from random positions on the sky; they are the signatures of the most powerful explosions seen in the Universe since the Big Bang.



A gamma ray burst seen in X-rays (top) and optical (bottom) as it fades[15]

As their name suggests these events are intense bursts of high energy photons called gamma-rays, which are associated with the births of stellar-sized black holes. After the initial, intense flash of gamma-rays has faded away they are still detectable for months or years after the event by their long lasting 'afterglows', which are visible over all electromagnetic wavelengths. They are also sources of ultra-high energy cosmic rays, high energy neutrinos and gravitational waves.

[15] Images property of NASA. As posted on http://rst.gsfc.nasa.gov/Sect20/grbsm.jpg



Data from a gamma ray burst. The total energy released is 10^{44} Joules[16]

Their typical duration is between fractions of a second up to several hundreds of seconds during which time they release approximately 10^{44} Joules of energy. This is comparable to the total amount of electromagnetic energy released in a supernova.

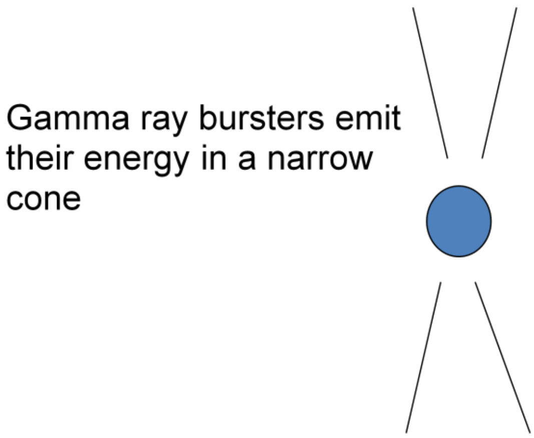
Energy scale

If Gamma Ray Bursts emit a comparable amount of energy to a supernova, what makes them so special?

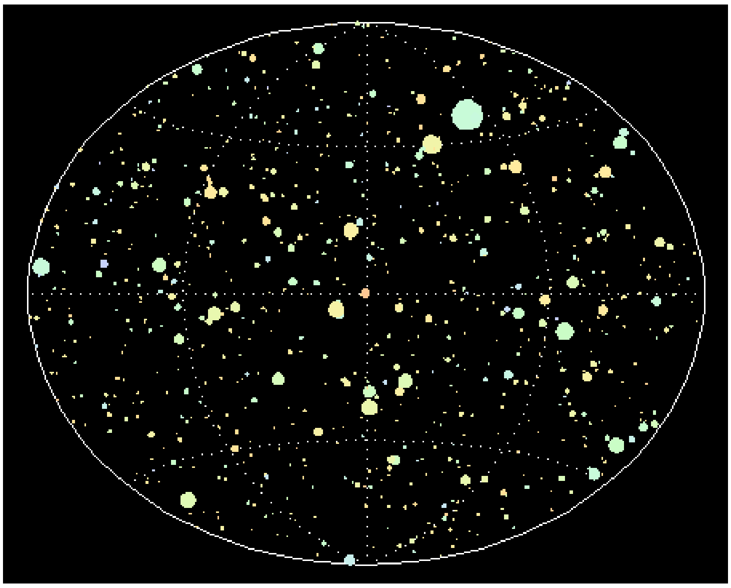
Supernovae release their energy and ejected matter isotropically (i.e. equally over all angles).

Gamma Ray Bursts, conversely, channel their energy and ejected matter along two polar beams or 'jets'. The effect of this beaming, plus their short duration, means that they have an *apparent* isotropic luminosity of the order 10^{37} to 10^{38} Joules per second making them temporarily the most luminous objects in the sky, outshining their own host galaxies and our Sun in gamma-rays!

[16] Gamma ray burst data property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/grb_1.jpg



To put it another way the energy output in a single Gamma Ray Burst event is the equivalent of burning up the entire mass-energy of the Sun … or to emitting the same amount of energy as our entire Milky Way does in a hundred years … all in the space of a few seconds!



The distribution of gamma ray bursts on the sky is isotropic implying that they cannot be local events.[17]

Estimate the energy output of the Sun over its lifetime

[17] Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/grb3map.gif

Why are gamma ray bursts of interest?

Most importantly the fact that they are so bright means that they can be seen across large distances. The most distant (confirmed) Gamma Ray Burst occurred 12.7 billion years ago, 0.89 billion years after the Big Bang. It has been suggested that Gamma Ray Bursts could be observable out to a mere 0.2 Gyr after the Big Bang, and indeed may have been already – we just haven't been able to confirm it as yet!

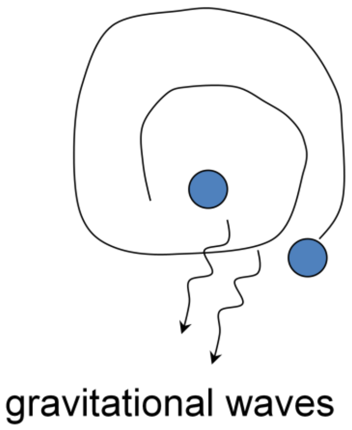
As the electromagnetic energy from the GRB travels towards us it picks up the imprint of the material it is passing through providing us with a unique opportunity to probe the environment of the early Universe.

Furthermore they are invaluable laboratories to study high energy astrophysical processes occurring in extreme environments; ones that we can't possibly replicate on Earth and are unlikely even to be matched in the core of our Sun.

Types of Gamma Ray Bursts

In fact there are thought to be two types of Gamma Ray Bursts, which are distinguished by the duration of their gamma ray activity. Short bursts typically last for a few tens of seconds up to a few seconds. Long bursts typically last between a few tens of seconds to several hundreds of seconds.

Short bursts are thought to be the result of the collision between a binary system of 'compact' objects, for example, a pair of Neutron stars or a Neutron star – White Dwarf system. The two compact objects orbit each other but over time the system's energy and angular momentum is carried away by gravitational waves. This causes the binary system's orbital separation to get smaller and smaller until the two objects merge. In the last minutes before coalescence the gravitational wave signal from the binary should sweep through the frequency range that is accessible to ground based gravitational wave detectors such as LIGO and VIRGO.



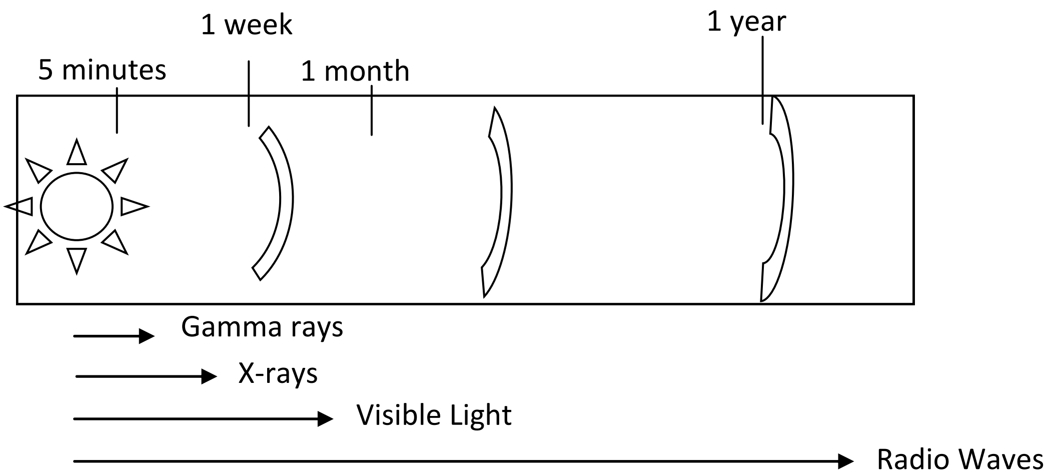
Long bursts, on the other hand, are thought to be the result of the death of a hypermassive star with a Main Sequence mass of the order 25 to 30 Solar masses; most likely a 'Wolf Rayet' star. As these stars collapse under their own gravity they undergo a 'hypernova', which is also known as a special type of supernova called a Type Ic-BL event. The signature of the hypernova can be seen in the Gamma Ray Burst's optical lightcurve after several days.

The outcome of both Long and Short Gamma Ray Bursts is the formation of a Black Hole (approximately 10 Solar masses) surrounded by a torus of material of a few Solar masses that accretes onto the Black Hole.

How much energy is required to remove 15 solar masses of material from the surface of a star?

The structure of a gamma ray burst

The mechanism for producing the jets seen in Gamma Ray Bursts, and other astrophysical phenomena such as Active Galactic Nuclei, is thought to be the result of a complex interaction between the newly formed black hole, the torus of material accreting onto it and the extremely strong magnetic field that surrounds the system.

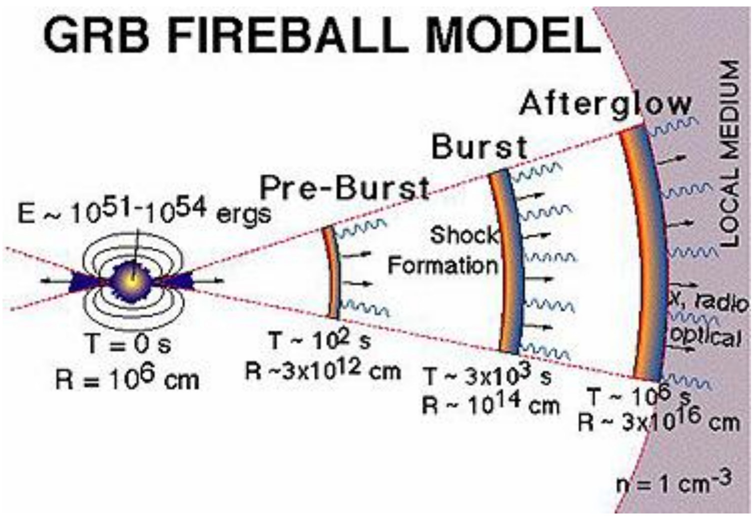


The Observability of Gamma Ray Burst Fireballs

In gamma ray bursts the majority of the energy released during the formation of the Black Hole-torus system and subsequent accretion processes is converted into neutrinos and gravitational waves, whilst a significantly smaller fraction goes into a high temperature 'fireball' containing electron-positron pairs, gamma-rays and baryons. This 'fireball', once injected into the jet pathway, expands at relativistic speeds.

Unlike other astrophysical phenomena, Gamma Ray Burst jets are not steady, that is energy is not injected at a constant rate. In reality it is perhaps better to describe the jets as a series of 'flying pancakes' (or 'shells') of material that trace out a cone shape as they progress further from the black hole.

In addition each fireball-shell may be ejected with a different speed. Over time the faster shells collide with the slower shells producing a series of 'shocks' that are thought to be responsible for the burst signal.



The GRB Fireball Model[18]

The gamma-rays observed from Gamma Ray Bursts do not necessarily come from the photons initially injected into the fireball: in fact, the vast majority of the internal energy of the initial fireball is converted into the kinetic energy of the shell as the shell material expands out adiabatically.

Instead the gamma-rays are produced from the internal shocks. The shocks accelerate the electrons within the colliding shells to close to the speed of light. These electrons emit radiation by the synchrotron and inverse-Compton processes which we'll look at next. This will answer our initial question as to how this high energy radiation is produced.

How are the gamma-rays produced?

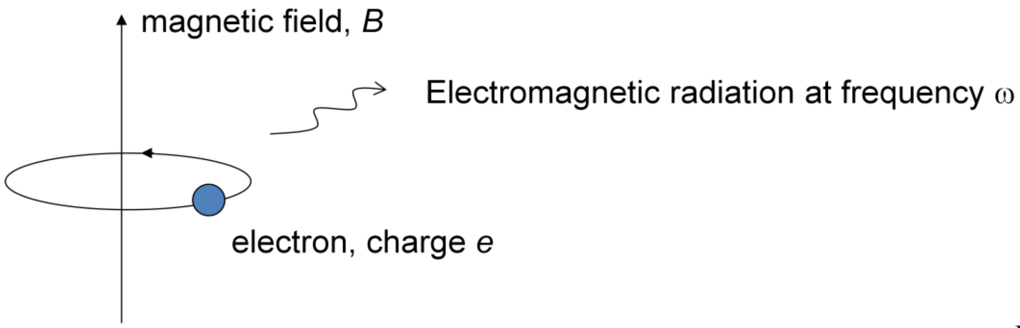
Synchrotron Radiation

What is the frequency of the radiation emitted by an electron moving in a magnetic field? We met this question previously. If the motion is non-relativistic, that is at speeds much less than that of light, the radiation peaks at the cyclotron frequency, eB/mc.

If the electrons are relativistic, that is they move at close to the speed of light, there are two effects to consider:

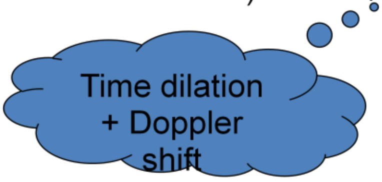
First, the time dilation effect of the electron's relativistic motion and second, that the radiation is Doppler shifted in the laboratory frame as the particle moves towards and away from us. These combine to give a factor γ^2 boost in the frequency, and hence in the energy, of the radiation. Radiation from electrons moving relativistically in a magnetic field is referred to as synchrotron radiation.

[18] Image property of NASA, as posted on http://rst.gsfc.nasa.gov/Sect20/fireball.jpg



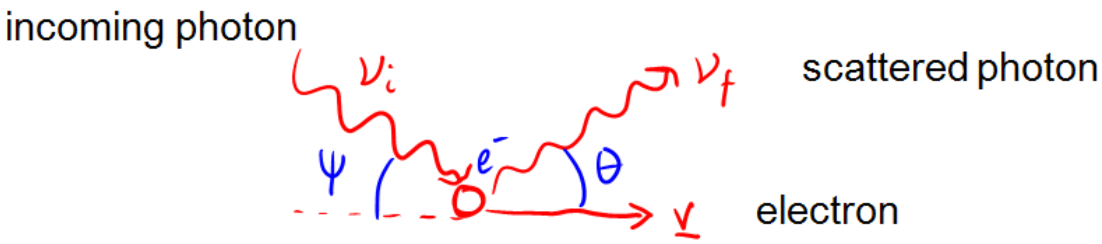
Cyclotron radiation: (Non-relativistic electron $v << c$) $\omega = \frac{eB}{mc}$

Synchrotron radiation: (Relativistic electron) $\omega = \gamma^2 \frac{eB}{mc}$



Inverse Compton Scattering.

Another way in which the photons produced within the internal shocks can be boosted into the gamma-ray range is by Inverse Compton Scattering.



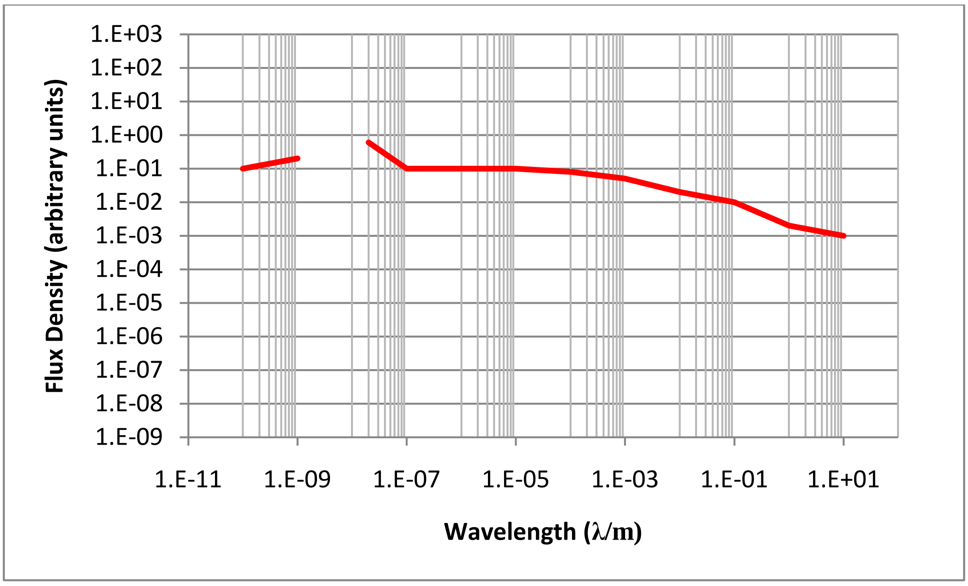
On average: $$\nu_f \approx \gamma^2 \nu_i$$

In the process known as Inverse Compton Scattering an incoming photon collides with a relativistic electron and gains energy from the electron, boosting its frequency. This occurs only when the electron has a significant amount of energy compared to the incoming photon: there is a net transfer of energy from the electron to the photon.

Inverse Compton scattering increases the photon energy by a factor of γ^2. So gamma-rays emitted by the synchrotron process are boosted to even higher energies.

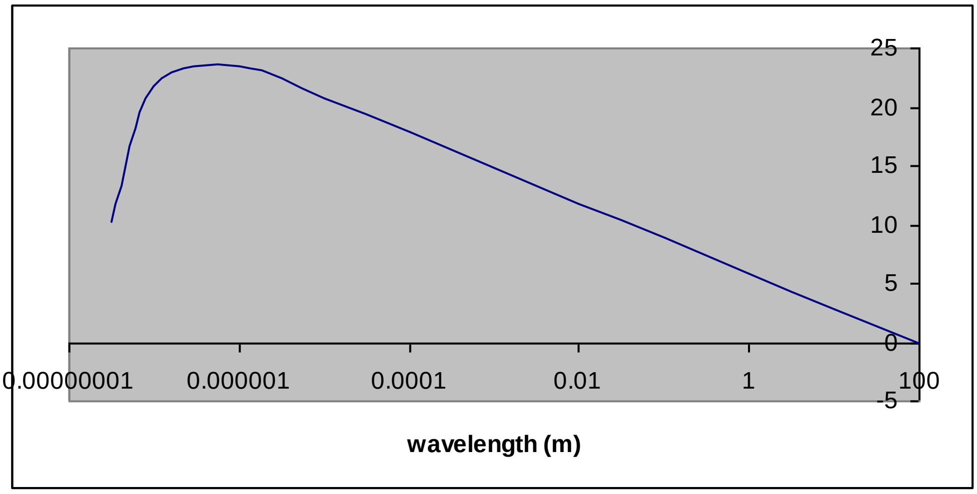
Non-Thermal Radiation

Let's summarise the problem we have looked at. This section shows first a typical quasar spectrum ranging from x-ray to radio wavelengths. It's plotted in a way that makes it fairly flat across the range which is a true reflection of the constancy of the energy emission per decade of wavelength. In other words a typical quasar is bright at all wavelengths.



Typical Quasar Spectrum

Below is the spectrum of a source of radiation from a body at 10 000K. It is plotted in the same way, with the absolute values of the vertical axis arbitrary, but the relative values showing the true shape of the curve.



Notice how it falls off at both large and small wavelengths. This is characteristic of a body in thermal equilibrium. So what the graphs show is that the emission from a quasar cannot be coming from a region in thermal equilibrium: the radiation must be non-thermal. We've seen that synchrotron and inverse Compton emission fit the bill here provided we can accelerate the electrons to relativistic speeds. We know that such energetic electrons can be produced under astrophysical circumstances because they rain down on the Earth continually in the form of cosmic rays. In the gamma ray bursts they are produced in the shocks that result from colliding clouds of rapidly moving gas.

Summary

- Many astrophysical object emit large amounts of energy from small volumes
- These objects include supernova, compact objects such as neutron stars (pulsars), and galactic black holes (quasars)
- The emission occurs across the electromagnetic spectrum and may be steady or in bursts
- These systems are powered by non-thermal processes such as synchrotron emission and Compton scattering

SAQs

Answer true or false?

1. From the radio emission of a quasar the temperature of the hot gas around the black hole can be deduced.
2. Highly energetic astrophysical sources are large and so cannot change on short time scales
3. Electrons never reach speeds approaching that of light in astrophysical sources

The answers appear on the following page

Answers

All false:

1. The temperature of a gas can be deduced only if it is in thermal equilibrium and the emitted radiation therefore has a black body spectrum. The radio emission of quasars comes from electrons accelerated to high speeds by non-thermal processes (such as in shocks).
2. We have seen that this is not true in the case of gamma ray bursts. The shortest time scale for change is the time it takes light to traverse the emission region which can be a matter of seconds or less.
3. Synchrotron emission and Inverse Compton emission arise from electrons moving close to the speed of light.

Cosmology

Cosmology is the study of the universe as a physical system. It seeks to answer such questions as what are the contents of the universe, how large is it, what is its past history and how will it evolve in the future.

Learning Objectives

- Describe the matter content of the Universe
- State the cosmological principle
- Give the evidence for the expansion of the Universe
- Describe the expansion in terms of a scale factor
- Describe the hot big bang model
- Describe the “horizon problem” and “inflation”

The Start of Cosmology

When we look at the night sky through a moderately powerful telescope we see myriad stars. It was established about 200 years ago by William Herschel that these stars make up a lens shaped system which today we call the Milky Way Galaxy. We now know that the Milky Way contains about 100 billion stars.

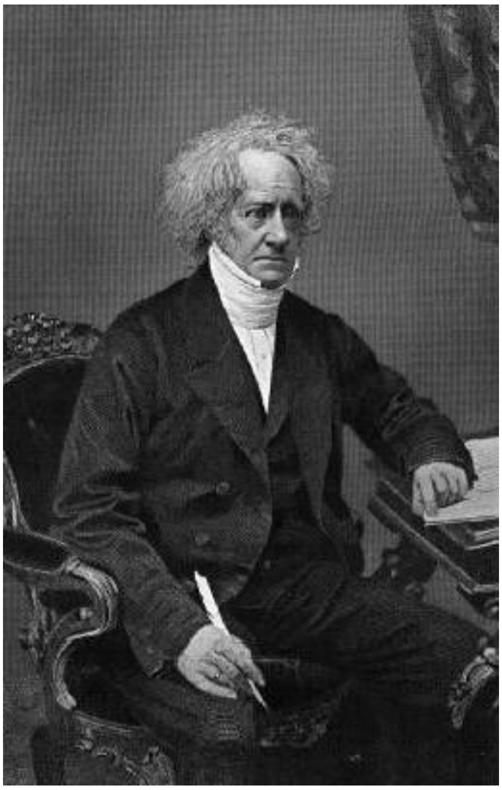
Our Galaxy, the Milky Way, in a clear night sky.[19]

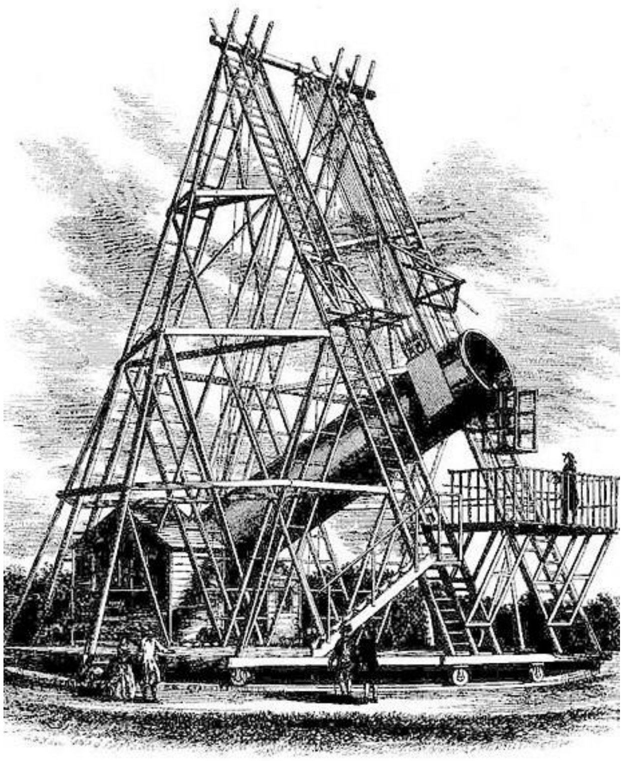
An image of what the Milky Way would look like if observed from outside.[20]

[19] Under the Milky Way, by jurvetson, as posted on www.flickr.com. Creative Commons Licensed.
[20] Galactic Twin. Image property of NASA, as posted on http://solarsystem.nasa.gov/multimedia/display.cfm?IM_ID=2625

Sir William Herschel and the Galaxy

Herschel was also interested in another class of visible object which were called nebulae at that time. He speculated that they were in fact other remote star systems like the Milky Way. Herschel was correct but it took until the 1920's to prove him right:



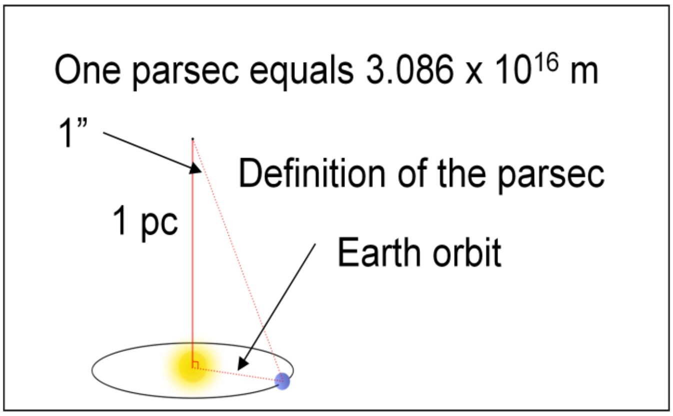


Frederick Willaim Hercschel (left)[21] and his largest, 40 foot telescope (right)[22]

Hubble determines the size of the Universe

Edwin Hubble using the newly commissioned 100 inch telescope on Mount Wilson showed that the M31 nebula (the Andromeda galaxy) was 700 kiloparsecs away from us. This places it well beyond the limits of the Milky Way. One parsec equals 3.086 x 10^{16} metres or about three light years.

[21] Friedrich_Wilhelm_Herschel, as posted by Duyckinick on commons.wikimedia.org. Public Domain.
[22] Wilhelm Herschel's 40-foot telescope as posted on commons.wikimedia.org. Public Domain.



The 100 inch Hooker Telescope used by Hubble to measure the expansion of the universe.[23]

[23] 100inchHooker, as posted by Solipsist, on http://en.wikipedia.org/wiki/Image:100inchHooker.jpg. Creative Commons Licensed.

The visible Universe

The distribution of galaxies stretches out from us in all directions to the limits of vision of the most powerful telescopes. We do not know if there is eventually an edge to the distribution, but we can say that there is an edge to the visible universe which we call the particle horizon. The visible universe is that part of the universe that we can in principle see. This limit to what we can see arises because the universe has a finite age, about 14 billion years, and light travels at a finite speed so the distance to our horizon is about 14 billion light years.

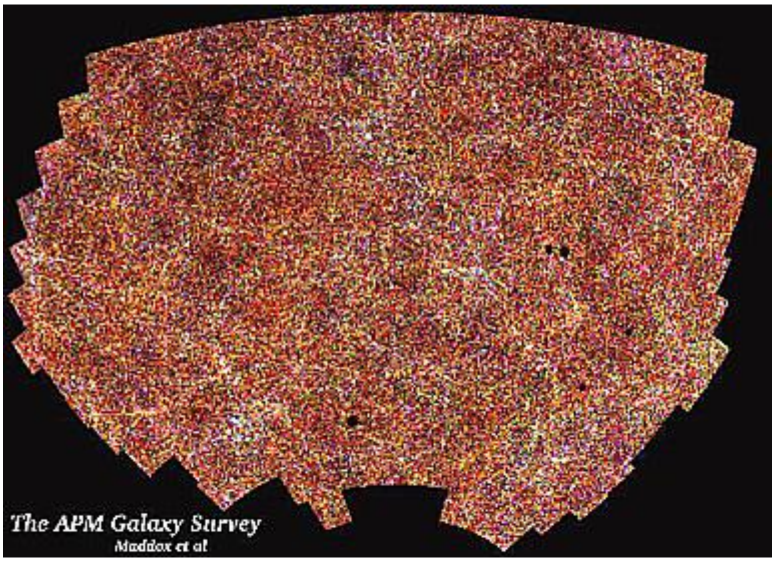
A cluster of galaxies – beyond this scale the distribution is uniform on average or "homogeneous"[24]

To get some feel for this distance, imagine that a galaxy is shrunk to the size of a penny then the typical distance between galaxies is about one metre and the horizon is 14 kilometres from us. There are about 100 billion galaxies in the visible universe.

[24] Hubble Views Distant Galaxies through a Cosmic Lens, part of the NASA Hubble Space Telescope Collection, by W.Couch. R. Ellis and NASA, as posted on nasaimages.org. Property of NASA.

The distribution of galaxies on large scales

Large scale surveys have been performed of galaxies we can observe. The distribution of the galaxies is isotropic about our position and the galaxy surveys reveal that on large scales it is homogeneous. We call this observation the *cosmological principle* it says in essence that the universe looks the same wherever you are in it. This is a very powerful statement that greatly simplifies theoretical modelling of the universe.



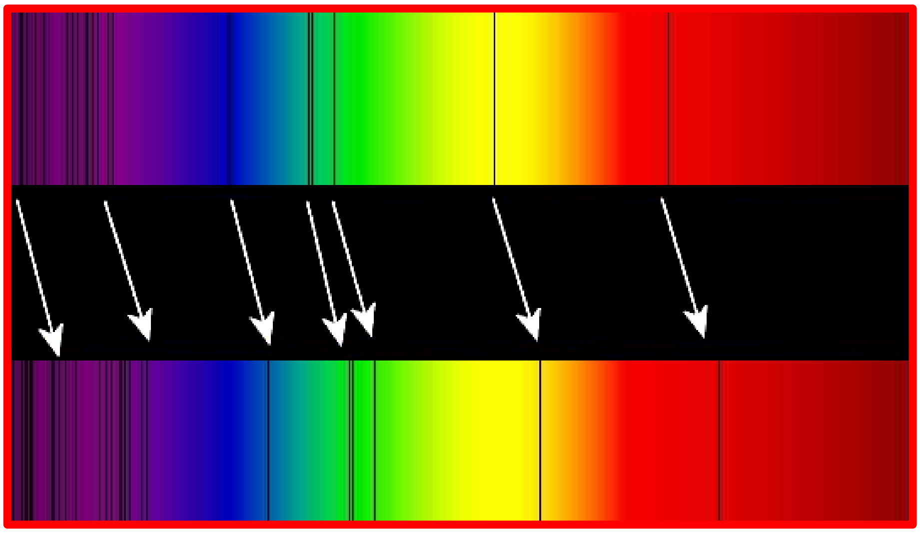
The APM Galaxy Survey, Maddox et al[25]

The cosmological principle says in essence that the universe looks the same wherever you are in it.

The expanding Universe

But before we turn to theory there is one further important observational feature of the universe that we must introduce, namely that the whole system of galaxies is expanding isotropically about our location. This realization came in 1930 after Hubble had shown that the light coming from galaxies is redshifted and the amount of redshift is proportional to the distance of the galaxy from us. Hubble's law is expressed by the equation $zc = H_0d$ where H_0 is the Hubble constant and d the distance.

[25] Image property of NASA. Image posted on http://rst.gsfc.nasa.gov/Sect20/galaxies2_apm.jpg

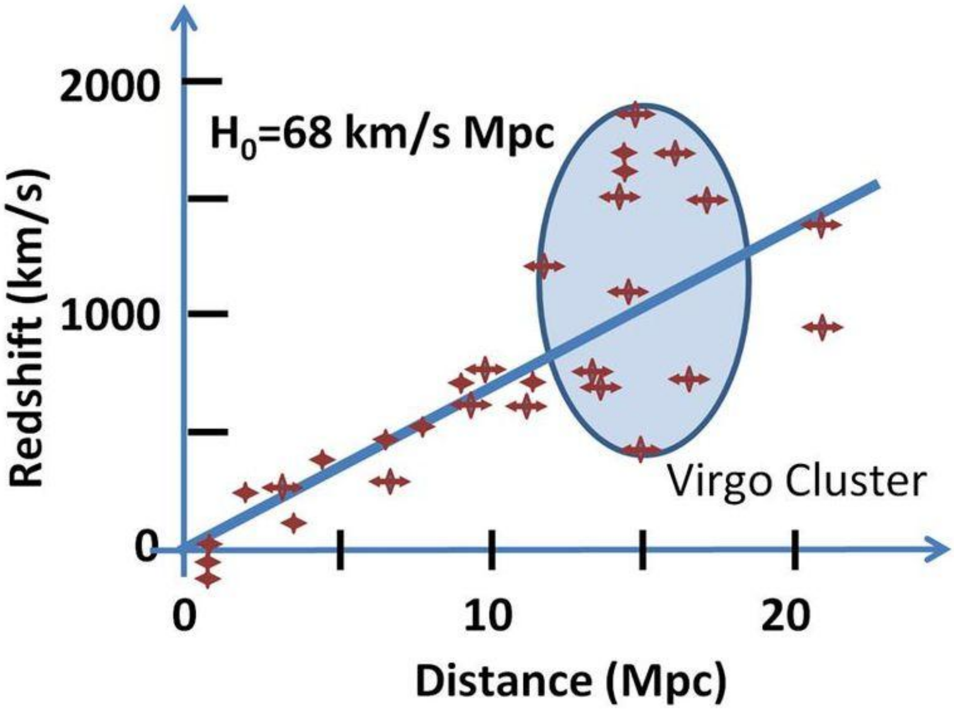


Redshift: spectral lines are shifted to the red[26]

For more information: http://en.wikipedia.org/wiki/Image:Redshift.png

We interpret the redshift as arising from the motion of the galaxy away from us. Combining this observation with the cosmological principle means we must conclude that the expansion has no centre: every galaxy sees every other galaxy receding away from it. We can think of the expansion as arising from the expansion of the space separating the galaxies. In reality not every galaxy is receding from every other galaxy as galaxies are usually found in clusters. Since clusters are bound systems the expansion only manifests itself on larger scales: to see the full Hubble flow, as it is termed, we have to look well beyond our local group of galaxies.

[26] Absorption lines in the optical spectrum of a supercluster of distant galaxies (BAS11) (right), as compared to those in the optical spectrum of the Sun (left). Based on a public domain image created by Harold T. Stokes, and amended by Ian Tresman. Original upload in English Wikipedia by Iantresman. Author Georg Wiora (Dr. Schorsch) created this image from the original JPG. The original image is PD this one is GFDL.



Hubble's plot of redshift (or velocity) against distance of nearby galaxies[27]

Hubble's Law: $v = cz = H_0 d$

These results show that the Universe is expanding

The expansion of the Universe

To describe the evolution of this expanding system of galaxies we need a mathematical model. The governing force is gravity and our theory of gravity is Einstein's general theory of relativity. It is here that the cosmological principle proves so important: it enables us to treat the universe as a fluid of uniform density subject to its own gravity. For such a universe containing matter, dark energy and radiation Einstein's equations reduce to equations (1), (2), (3) and (4). These can be solved to yield the evolution of the length scale R and the densities as functions of time. Equation (5) relates the scale length of the Universe R to the Hubble parameter H at any time t.

[27] Hubble Constant.jpg, by Brews ohare, as posted on commons.wikimedia.org. Creative Commons Licensed.

spatial curvature

Friedmann equation

$$\dot{R}^2 = \frac{8\pi G\rho}{3} R^2 - kc^2 \qquad (1)$$

Density ρ

$R(t)$ = size of Universe at time t

$$\rho = \rho_m + \rho_{de} + \rho_{rad} + \ldots..$$

Density of matter

Dark energy or "cosmological constant"

Density of radiation

Conservation of matter :

$$\rho_m R^3 = \text{constant} \qquad (2) \qquad \rho_{de} = \text{constant} \quad (3) \qquad \rho_{rad} R^4 = \text{constant} \qquad (4)$$

Relation of Hubble parameter to $R(t)$: $H = \frac{\dot{R}}{R}$ (5)

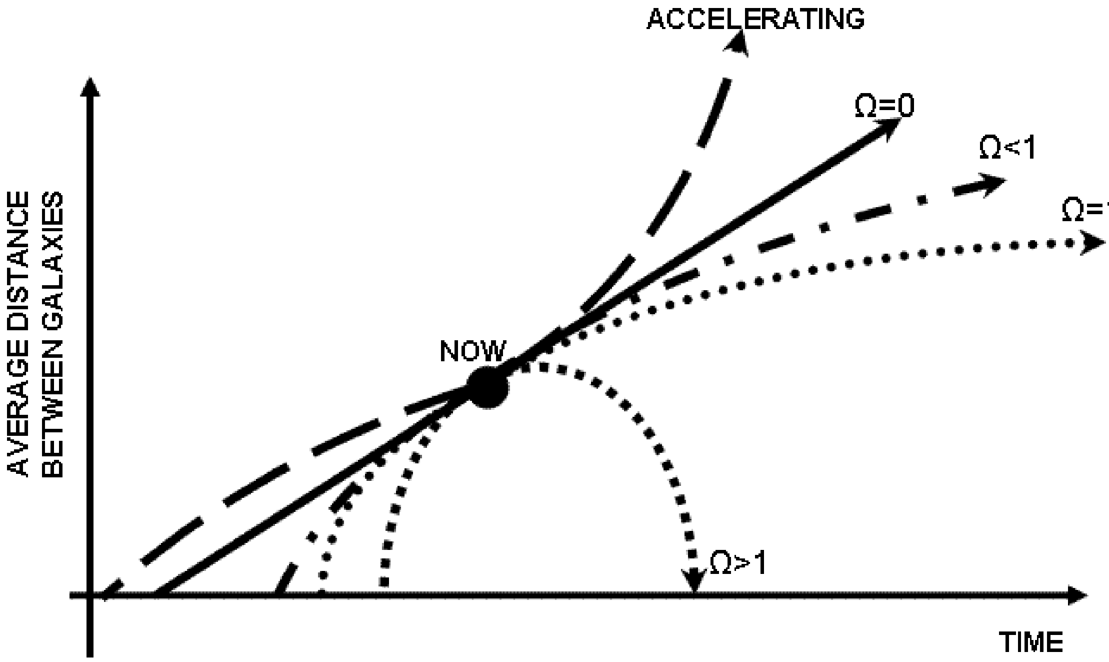
Critical density $\rho_c = \frac{3H^2}{8\pi G}$ (6) Density parameter $\Omega = \frac{\rho}{\rho_c} = \frac{8\pi G\rho}{3H^2}$ (7)

For a critical density universe (k=0) $\Omega_m + \Omega_{de} + \Omega_{rad} = 1$ (8)

The observational evidence points to our Universe having zero spatial curvature, that is k=0. This means that the spatial geometry is Euclidean: the angles of a cosmic triangle add up to 180^0. From equation (1) with k=0, and using the definition of the Hubble parameter equation (5), we get equation (6) for the definition of the critical density. Finally equation (7) defines the density parameter Ω which is often used instead of the density itself. A critical density Universe has Ω=1. For our Universe Ω is the sum of contributions from matter, dark energy and radiation, equation (8). At the present time we have $\Omega_m = 0.28$ and $\Omega_{de} = 0.72$ with a negligible contribution from radiation.

Models of the Universe

Here are lots of models obtained as solutions of the Friedmann equation under various assumptions. In order to find which (if any) correspond to our Universe we need to obtain the values for various parameters, namely the present values of the Hubble constant, the curvature constant and also the mean mass densities of the various contributing sources of mass energy.



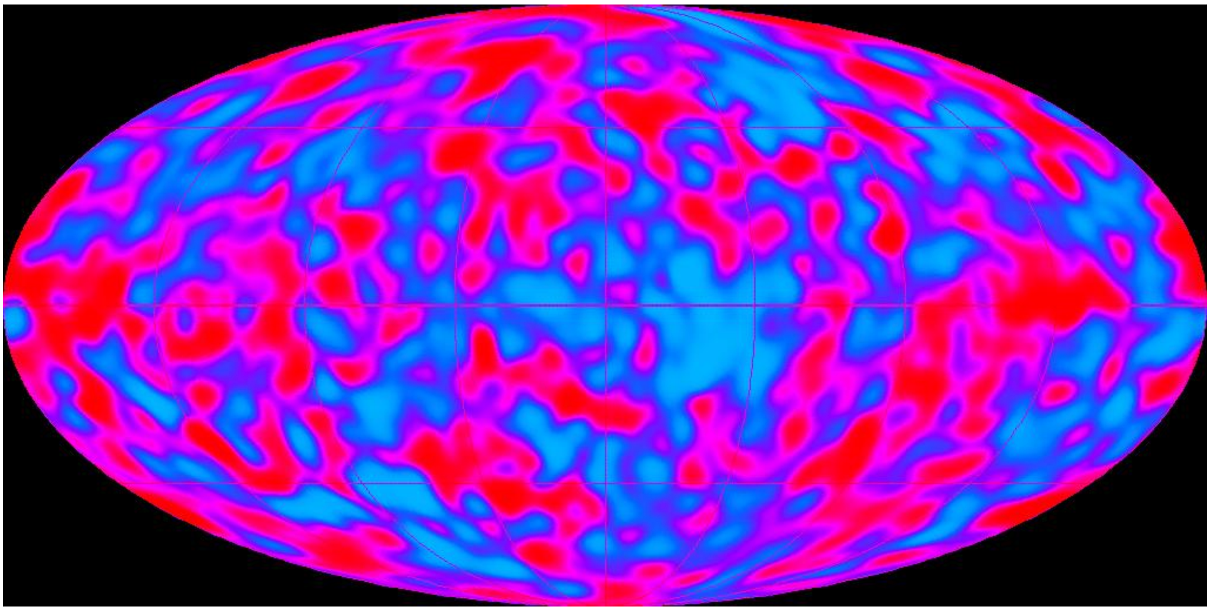
$R(t)$ plotted against t for various possible models of the Universe[28]

These quantities are now known to reasonable accuracy, so we can trace the evolution of the universe 14 billion years back into the past to a point where the density reaches infinite values. Thus general relativity predicts that the universe started from a singularity. In fact we believe that GTR will become invalid before the singularity is reached, because quantum effects will have been important at times earlier than 10^{-43} seconds. Only a theory of quantum gravity which does not yet exist can tell us about the earliest times.

The Cosmic Microwave Background radiation

Let us look now at the four sources of mass energy which govern the expansion, both past and future. They are radiation, baryonic matter, dark matter and dark energy. We will look at each of these in turn. The figure shows the small departures from a uniform temperature across the sky:

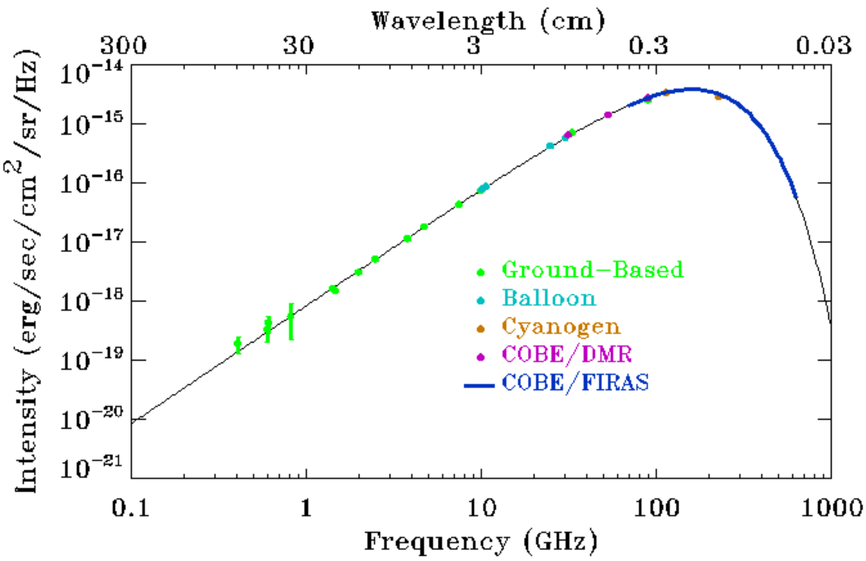
[28] Universos, as posted on commons.wikimedia.org. Public Domain.



Cosmic Background Radiation fluctuations[29]

The total mass density is close to the critical density so we will assume a critical density Universe. The four components of mass density always sum up to the critical density but their fractional contributions are functions of time.

The figure shows how closely the spectrum matches that of a blackbody at a temperature of 2.725 K:



CMB intensity[30]

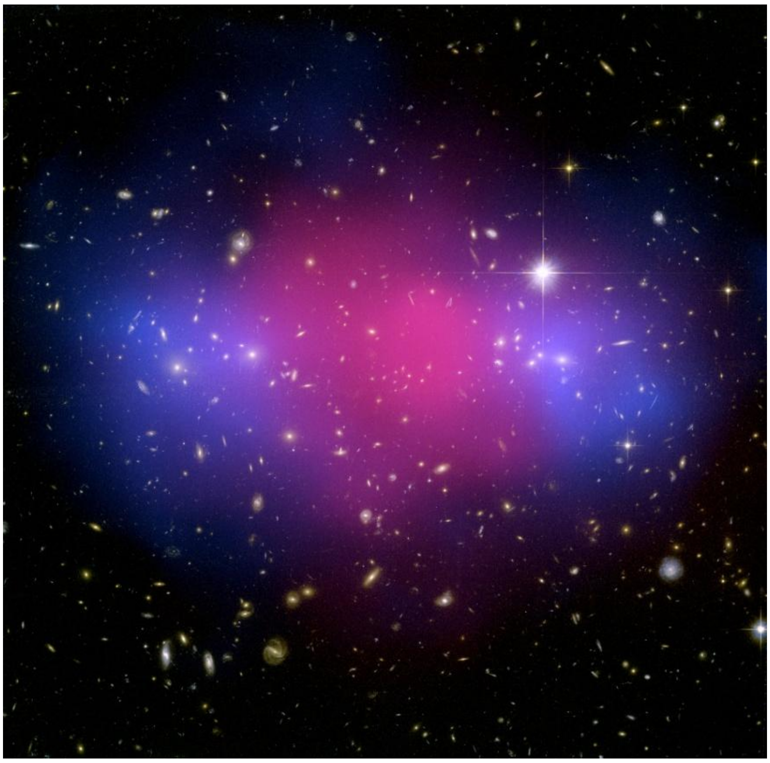
[29] COBE_cmb_fluctuations, NASA Goddard Space Flight Center, as posted on commons.wikimedia.org. Public Domain.

The temperature is proportional to $1/R(t)$ so the Universe was much hotter in the past.

A radiation background was discovered in 1965 by Penzias and Wilson; it has a black body spectrum at a temperature of 2.725 K and contributes only a fraction of a percent to the total mass density at the present time. However the temperature of blackbody radiation in an expanding universe is inversely proportional to the length scale, so at early times the universe would have been a very hot place and the mass density of the radiation would have dominated. This is why we talk about the "hot big bang".

Dark Matter and Dark Energy

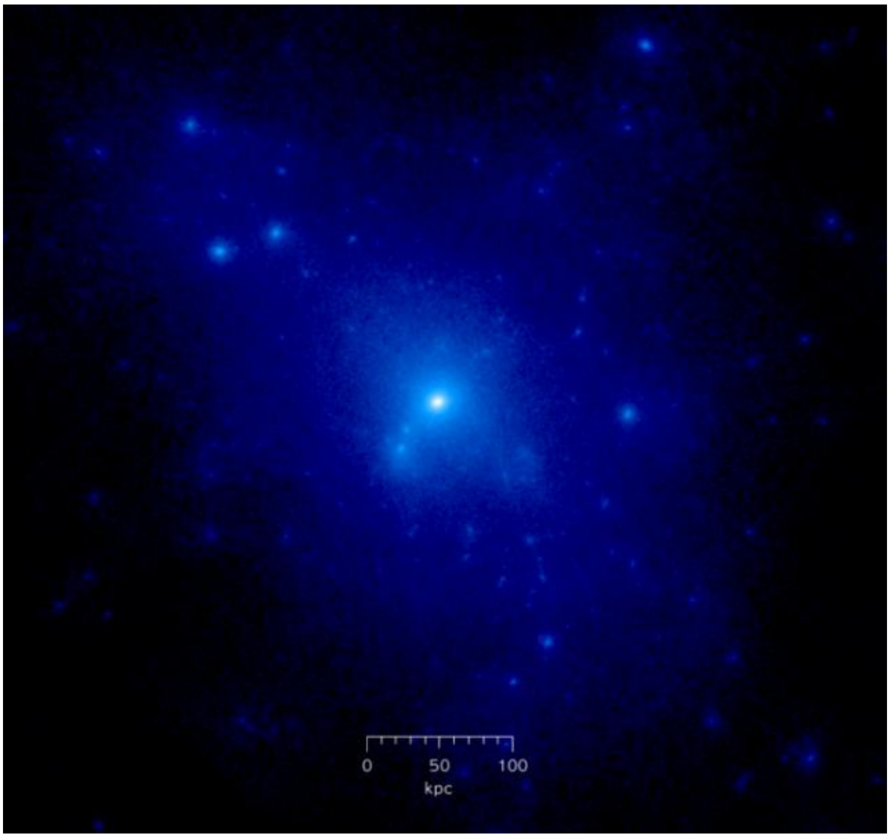
The baryonic matter is the familiar stuff composed of protons, neutrons and electrons that makes up ourselves, the Earth, the Sun and the stars, in fact it makes up everything that we can see in the universe. Surprisingly it only accounts for 4% of the total mass density at the present time.



[30] Property of NASA. As posted on http://arcade.gsfc.nasa.gov/images/cmb_intensity.gif

This picture of a galaxy is colour coded to show the dark matter (blue) deduced from the motion of the stars.[31]

The next component dark matter is believed to be made up of weakly interacting massive particles referred to as WIMPS. These are particles postulated by certain theories that aim to go beyond the standard model of particle physics. Dark matter plays several roles in cosmology: it holds galaxies together, without it they would fly apart and it mediates the formation of galaxies. So far it has evaded detection, which is worrying as it makes up 23% of the mass density. It is a major unsolved problem for cosmology.



This computer-generated image shows the simulated distribution of dark matter in a galaxy cluster formed in the universe with dark energy. The clumps are locations where galaxies form.[32]

Even more mysterious is the dark energy, it contributes the final 72% of the mass density. It is a uniform field having a time independent positive mass density and negative pressure. It was first introduced by Einstein who called it the cosmological constant. The negative

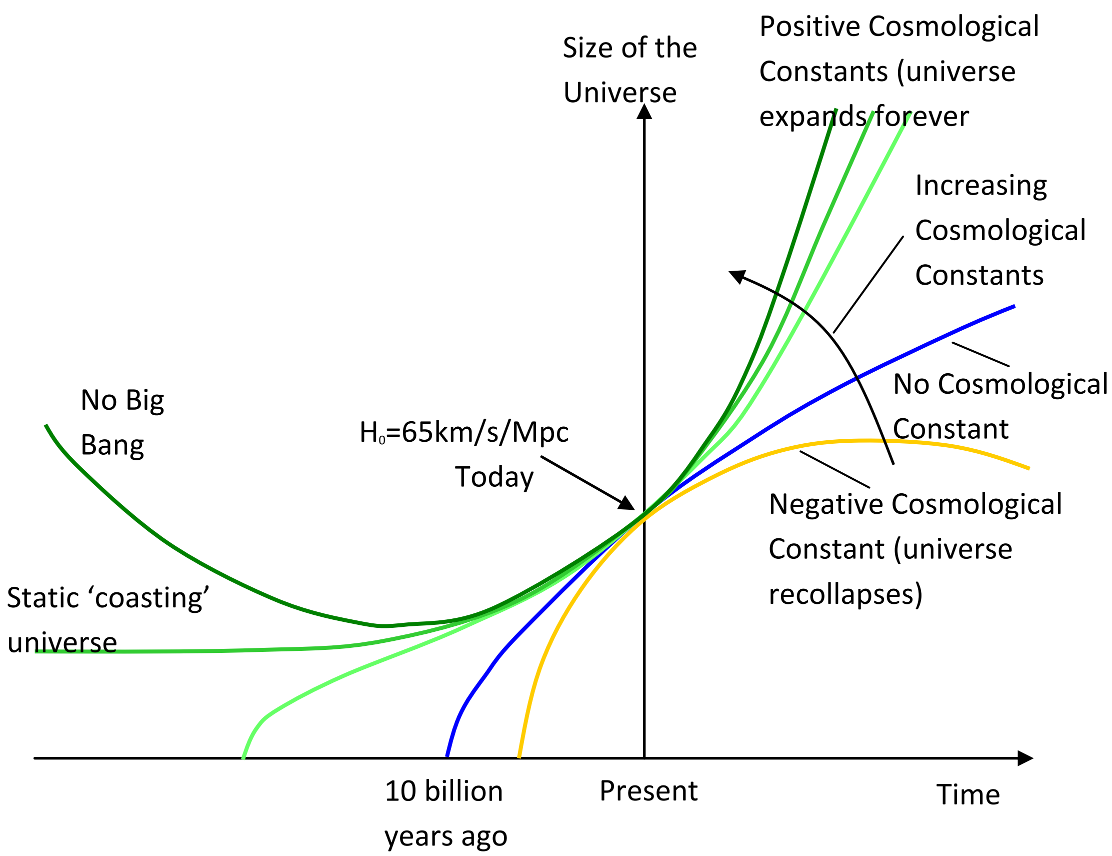
[31] Galaxy Cluster MACS, image from NASA, as posted on commons.wikimedia.com. Public Domain.
[32] Dark_matter_halo, as posted on commons.wikimedia.org. Public Domain.

pressure means that it gravitates repulsively, at the present time it is the dominant component of mass energy. Under its influence the Hubble expansion is accelerating rather that slowing down as would be expected to occur under the influence of matter only. What dark energy is and how it relates to the rest of fundamental physics is a mystery: it is the greatest challenge facing cosmologists.

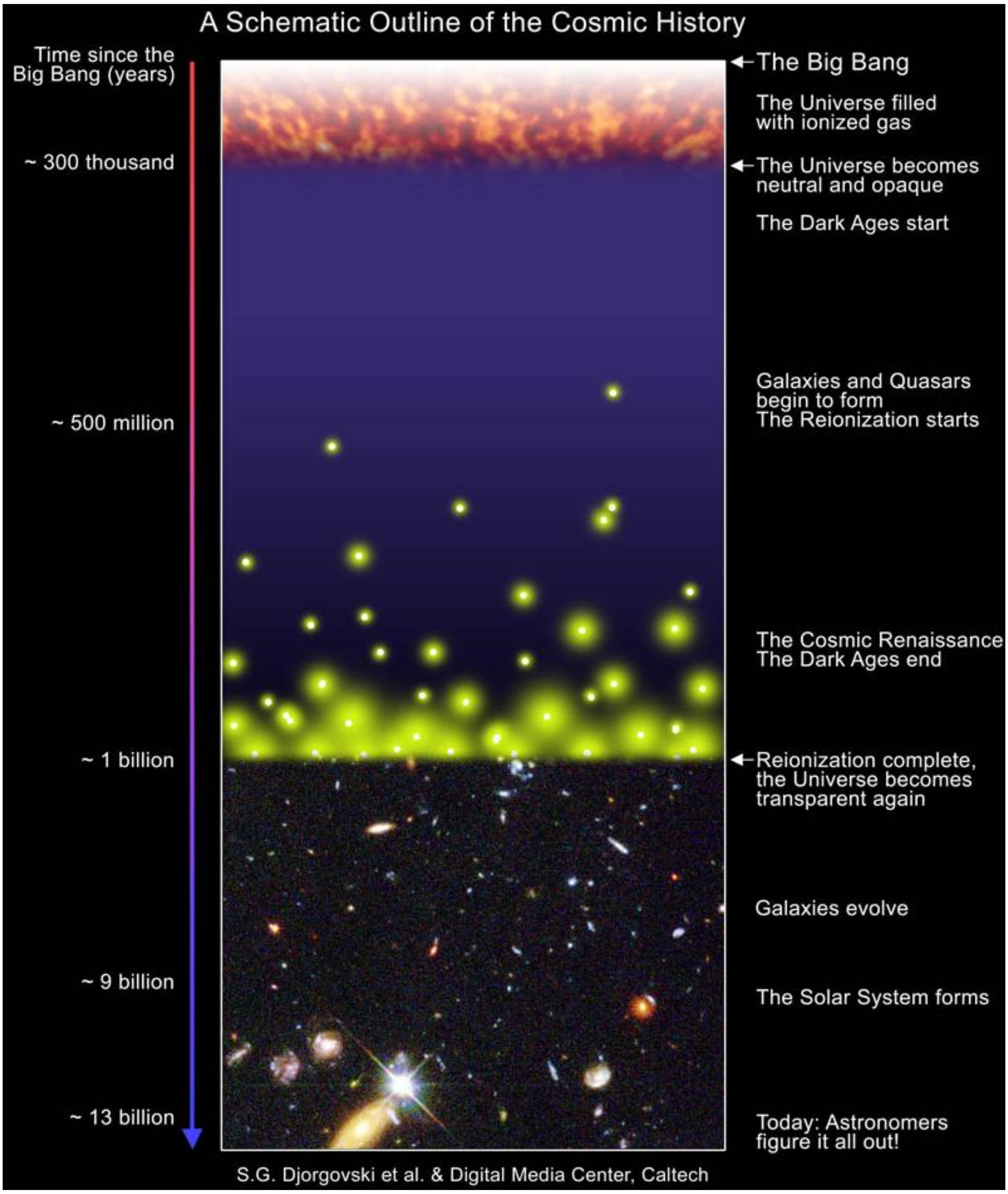
Baryonic matter	0.04
Dark Matter	0.23
Dark Energy	0.72
(Cosmological Constant)	
Overall Density:	$\rho = \rho_c$

Hubble parameter

$H_0 = 70.1 \pm 1.3$ km s $^{-1}$ Mpc^{-1}



What we have outlined is the standard big bang theory of cosmology. Summarising, it asserts that the universe came into existence about 14 billion years ago. The very earliest times were characterised by enormously high temperatures and densities. Since then the universe has been expanding and cooling. Once the temperature had dropped to about 3000 K neutral atoms could form for the first time. Also small primordial density fluctuations could start to collapse by gravitational condensation into stars and galaxies and subsequently galaxies clumped into larger scale structures bringing us to the present time.



An outline of the history of the universe.[33]

[33] Reion_diagram, NASA, as posted on commons.wikimedia.org. Public Domain.

Let's look more closely at the Universe at an age of a few seconds to a few hundred seconds. This is the time of nucleosynthesis, when the light elements were made from hydrogen.

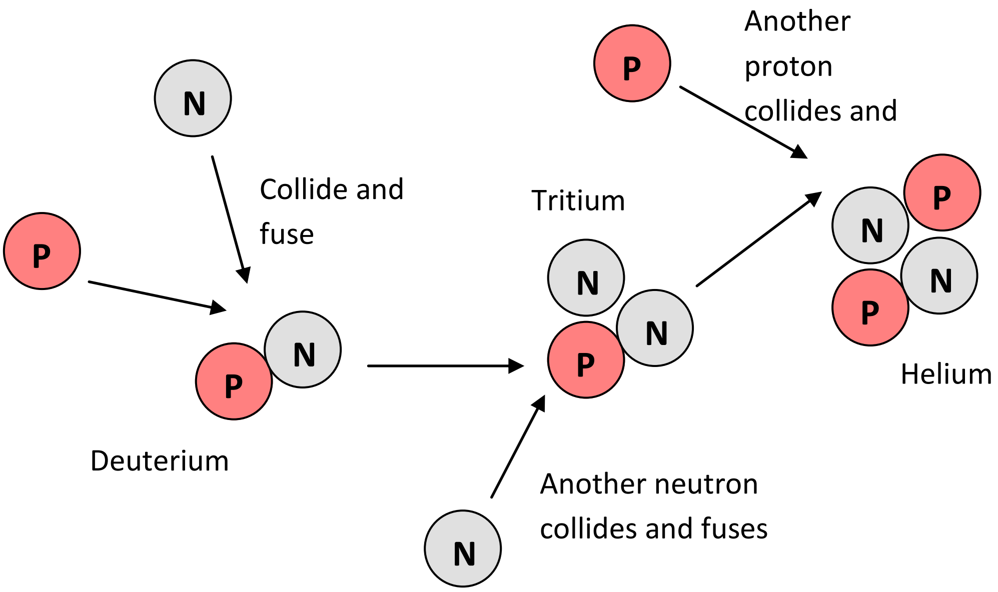
Nucleosynthesis

In the 1950's the theory of element building (nucleosynthesis) in stars was worked out by Hoyle, Burbidge, Burbidge and Fowler. It was assumed that the primordial element was hydrogen and that this was built up into all the other elements in the hot interiors of stars. This theory of element creation in stars is for the most part very successful. But in 1964 Hoyle and Taylor noticed a problem with it: the abundance of elements heavier than helium varies from one place to another but the abundance of helium-4 shows very little variation. Also there is too much helium, about 25% by mass, to be explained by stellar nucleosynthesis which can only have contributed a few percent. Then in 1965 the microwave background radiation was discovered by Pensias and Wilson and with this discovery came the realization that the early universe was a hot place, just the conditions needed for element building. High temperatures are needed to give a pair of nuclei enough kinetic energy to overcome the repulsive Coulomb force between them and bring them within the range of the short range nuclear binding force. Nucleosynthesis requires temperatures in the range 10^8 to 10^{10} K: if the temperature is too high the nuclei will be broken up by the photons if it is too low then they cannot get close enough to fuse.

As a rule of thumb the relation between temperature T and time t in seconds is $T = \frac{10^{10}}{t^{1/2}}$

Formation of Helium in the Early Universe

The initiating reaction is the fusion of a proton and a neutron to form deuterium. There is no coulomb barrier to be overcome in this reaction but the binding energy of the deuterium is low, so it is only stable at temperatures of 10^9 K and below. Then as soon as stable deuterium can form, the reactions can occur leading rapidly to the formation of helium 4; no elements beyond helium 4 are produced apart from traces of lithium 7.



There are several reasons why element fusion stopped at helium 4:

- there are no stable nuclei at mass numbers 5 and 8
- the density of the baryons was much lower than it is at the centre of a star
- there is only a short time interval, a few minutes, before the temperature has fallen below 10^8 K.

Where do the neutrons come from?

In the centre of a star like the Sun deuterium is formed by the fusion of two protons as there are no free neutrons present. Why then are there neutrons present in the early universe? We can understand this as follows.

At temperatures above 10^{10} K the photons of the background radiation have enough energy to create electron positron pairs as in equation 1.

Above 10^{10} K

$$\gamma + \gamma \leftrightarrow \boxed{e^-} + e^+ \qquad (1)$$

$$e^- + e^+ \leftrightarrow \nu + \boxed{\bar{\nu}} \qquad (2)$$

$$p + \boxed{e^-} \leftrightarrow n + \nu \qquad (3)$$

$$p + \boxed{\bar{\nu}} \leftrightarrow n + e^+ \qquad (4)$$

The electron positron pairs can also annihilate into neutrino anti neutrino pairs as in (2). So we have photons, electrons, positrons, neutrinos and antineutrinos all in thermal equilibrium. This means that protons and neutrons can be interconverted as in reactions (3) and (4). Thus neutrons were present in thermal equilibrium with the protons.

Below 10^{10} K at $t = 1$ second all the above reactions ceased. The first reaction ceased when the photon energies fell below the pair production threshold and subsequently the pairs annihilated into photons. The remaining reactions ceased because they are weak processes and their mean free time became longer than the age of the universe. The neutrinos could not annihilate so they are still present today. The neutron however is unstable with a mean lifetime of 887.5 s so after 100 s at temperatures of 10^9 K there were still neutrons present. In fact the neutron proton ratio was at this point 1/7. Effectively all these neutrons ended up in helium 4 nuclei.

Below 10^{10} K

$$e^- + e^+ \rightarrow \gamma + \gamma$$

$$n \rightarrow p + e^- + \bar{\nu}$$

Abundances of the light elements

One of the greatest successes of the big bang model is that it can account for the observed helium abundance. It also predicts small amounts of primordial Helium 3, deuterium and Lithium 7. The amounts of these 3 light elements produced depends on the baryon density at the time of nucleosynthesis. So if we know the primordial density of any of these light elements we can use it to calculate the baryon density at the present time. The amount of deuterium present in distant and therefore relatively pristine clouds of hydrogen gas has been measured with the Keck telescope. This gives a value for the baryon density at the present time of 4% of the critical density. Other independent methods for estimating the density in baryons give similar values.

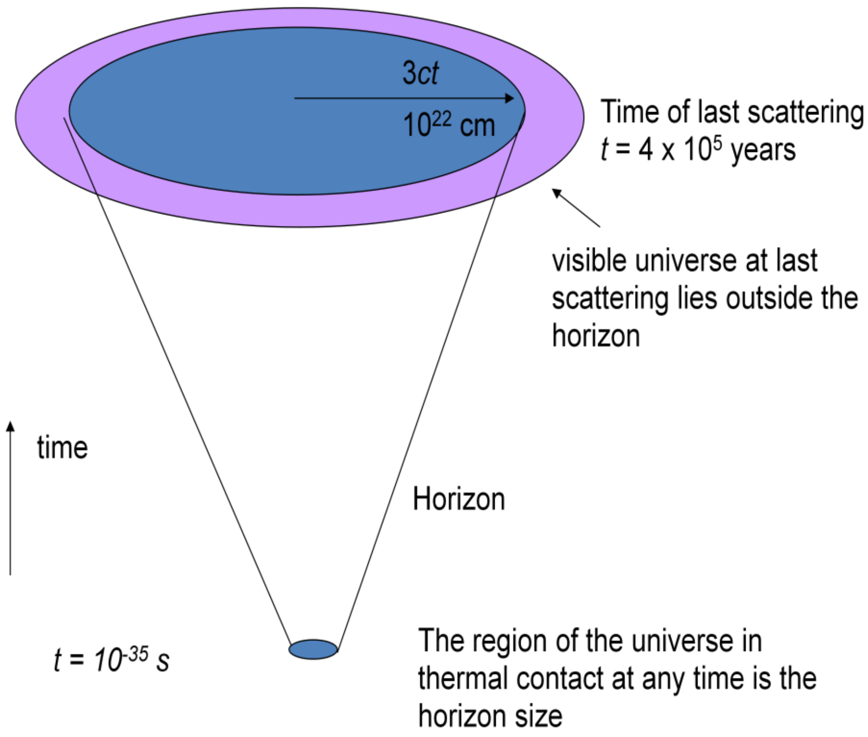
Propagation of radiation

Let's end by looking at one problem with this picture which will introduce the idea of inflation models of the early Universe.

At the present time the background radiation propagates freely with a mean free path greater than the length scale of the horizon. As we trace back in time the temperature of the radiation increases until it reaches 3000 K. Prior to this event neutral atoms did not exist and the matter formed a plasma; in such an environment radiation suffers frequent scattering: in other words the plasma is opaque. Thus when we look at the microwave background radiation we are looking back to the time when the radiation last interacted with matter i.e. to the time when neutral atoms formed and the radiation was able to propagate freely we refer to this time as the epoch of last scattering. The universe had an age of about 400,000 years when this event took place. A similar situation exists when we look at the Sun: we only see as far as its surface as this is the last time the photons interacted with the Sun's matter.

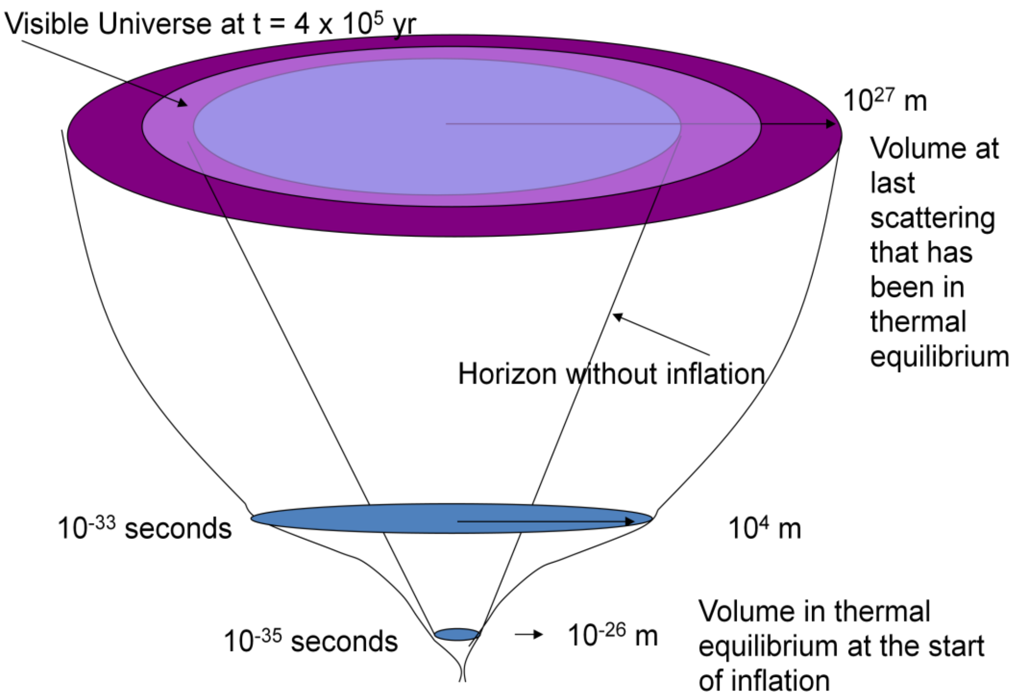
The Horizon Problem

For simplicity we will consider a critical density universe; for such a universe the length scale and the distance to the horizon at time t is $3ct$. Now the temperature of the background radiation is the same in all directions to within about one part in a hundred thousand. It seems reasonable to conclude that the whole volume within our particle horizon came into thermal equilibrium before the epoch of last scattering. However, according to the standard big bang picture this could not have happened! This is the horizon problem.



Inflation

The only explanation for the uniform temperature of the MWBG that the standard picture can offer is that this was an initial condition. However if we suppose that at a very early time s the universe underwent a brief period of exponential expansion or inflation, which increased the length scale by a factor of about 10^{30}, then very small regions that had come into thermal equilibrium before the burst of inflation would be very much larger than the volume occupied by the sources of the cosmic background radiation at the epoch of last scattering and the horizon problem would go away.



Inflation blows up a volume in thermal equilibrium at t = 10^{-35} s to reach a size larger than that occupied by the sources of the cosmic background radiation at the time of last scattering.

The future

Finally we will finish with a look into the future. In a universe containing no dark energy the visible universe will grow with time as new galaxies enter the horizon. So cosmology in such a universe would have a fascinating future. The opposite however will occur in our universe which is increasingly becoming dominated by dark energy. The red shifts of galaxies will increase, i.e the wavelength of light received from them will lengthen to the point where they will fade into invisibility. Trillions of years into the future the visible universe will consist of just our local group of galaxies which by then will have merged into a single super galaxy. By some happy accident we happen to live at a time when we can discover the hot big bang origin and subsequent evolution of our universe. The cosmologists of the far future will not be able to discover the past of their universe and it will seem to them to be a rather dull place.

Summary

- The Universe is homogeneous and isotropic on large scales
- The Universe is expanding on large scales, H = 72 km s^{-1} Mpc^{-1}
- The early Universe was hot
- Most of the ^{4}He and traces of ^{3}He, ^{2}H and ^{7}Li were produced at an age of a few minutes
- A very early episode of exponential expansion is needed to explain features like the extreme uniformity of the cosmic background radiation temperature

SAQs

1. (a) true or (b) false?

   (i) There is nothing beyond the visible Universe:

   (ii) (ii) The expansion of the Universe has a centre from which everything emerged

2. What is expanding in the expanding Universe?

   (a) everything

   (b) any object larger than a single galaxy (c) scales larger than clusters of galaxies

3. What is the distance to a galaxy of redshift 0.1?

   (a) 4.16×10^5 Mpc

   (b) 416 Mpc

   (c) 1.25×10^{25} Mpc

The answers appear on the following page

Answers

1. (i) false: the visible universe encompasses just those galaxies from which light has had time to reach us since the big bang.

   (ii) false: there is no centre to the universe. The expansion is uniform so each observer is equivalent to any other and appears as the centre to that observer.

2. (a) false: the expansion applies only to objects that are not held together by attractive forces

   (b) false: clusters of galaxies are held together at least partly by their own gravity and so do not expand with the Universe.

   (c) true: on this scale the clumping of matter dose not create a significant attraction to impede the universal expansion

3. (a) incorrect: you may have substituted a value for c in m s^{-1} and used H in km s^{-1}

   (b) true: $d = cz/H$ and $H = XX$ km s^{-1} Mpc^{-1}

   (c) Incorrect: you may have confused metres and Mpc.

Meta tags

Author: Sarah Symons, Naomi Banks, Cheryl Hurkett, Derek Raine.

Owner: University of Leicester

Title: Enhancing Physics Knowledge for Teaching – Astrophysics

Keywords: Positional astronomy; Stars; Galaxies; Cosmology; sfsoer; ukoer

Description: In this session in this session four of the team from the Physics Innovations Centre for Excellence in Teaching and Learning at Leicester will write about one of their areas of interest..

Language: English

Version: 1.0

Additional Information

This pack is the Version 1.0 release of the module. Additional information can be obtained by contacting the Centre for Interdisciplinary Science at the University of Leicester. http://www.le.ac.uk/iscience