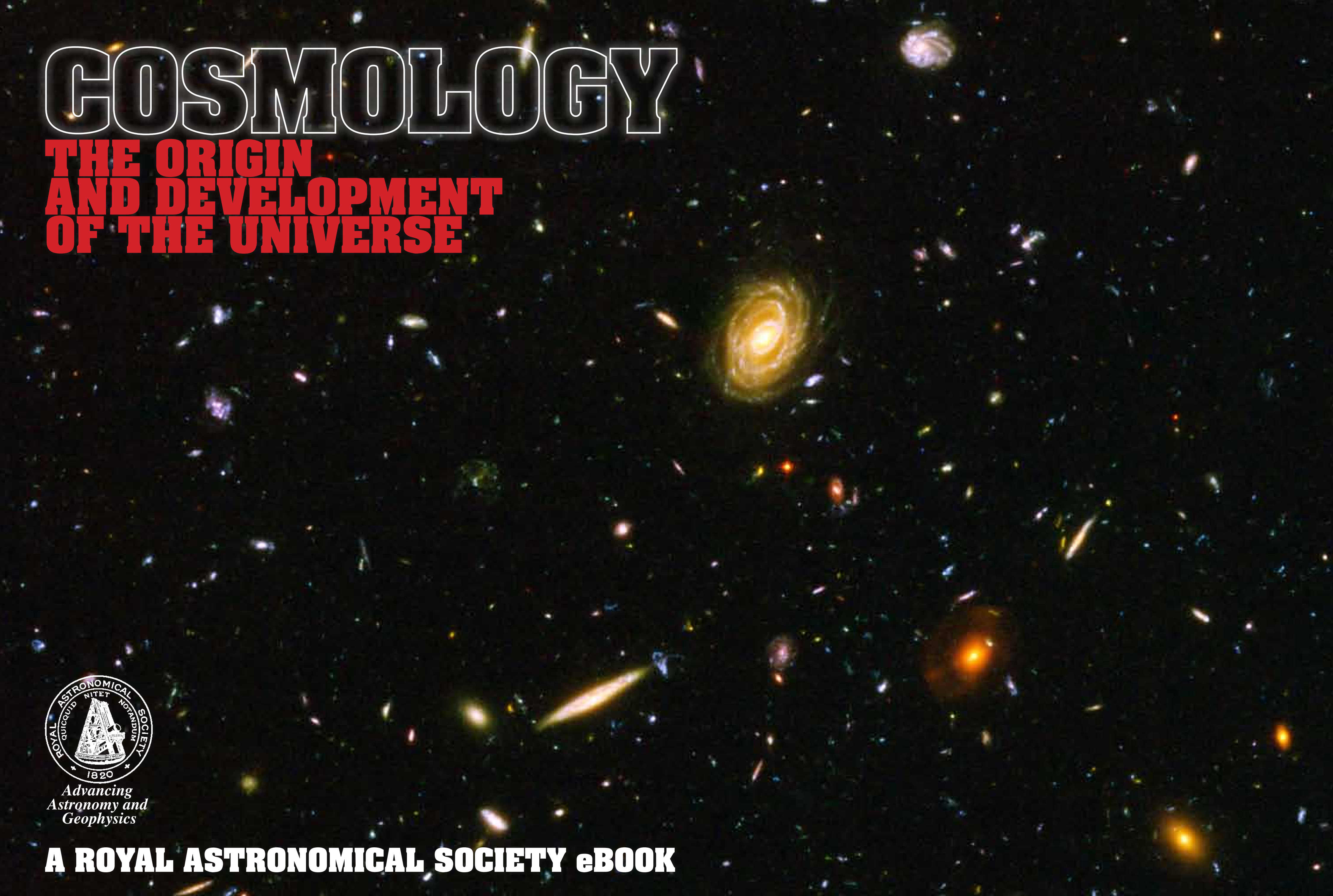
COSMOLOGY

THE ORIGIN AND DEVELOPMENT OF THE UNIVERSE



Advancing Astronomy and Geophysics

A ROYAL ASTRONOMICAL SOCIETY eBOOK

CONTENTS

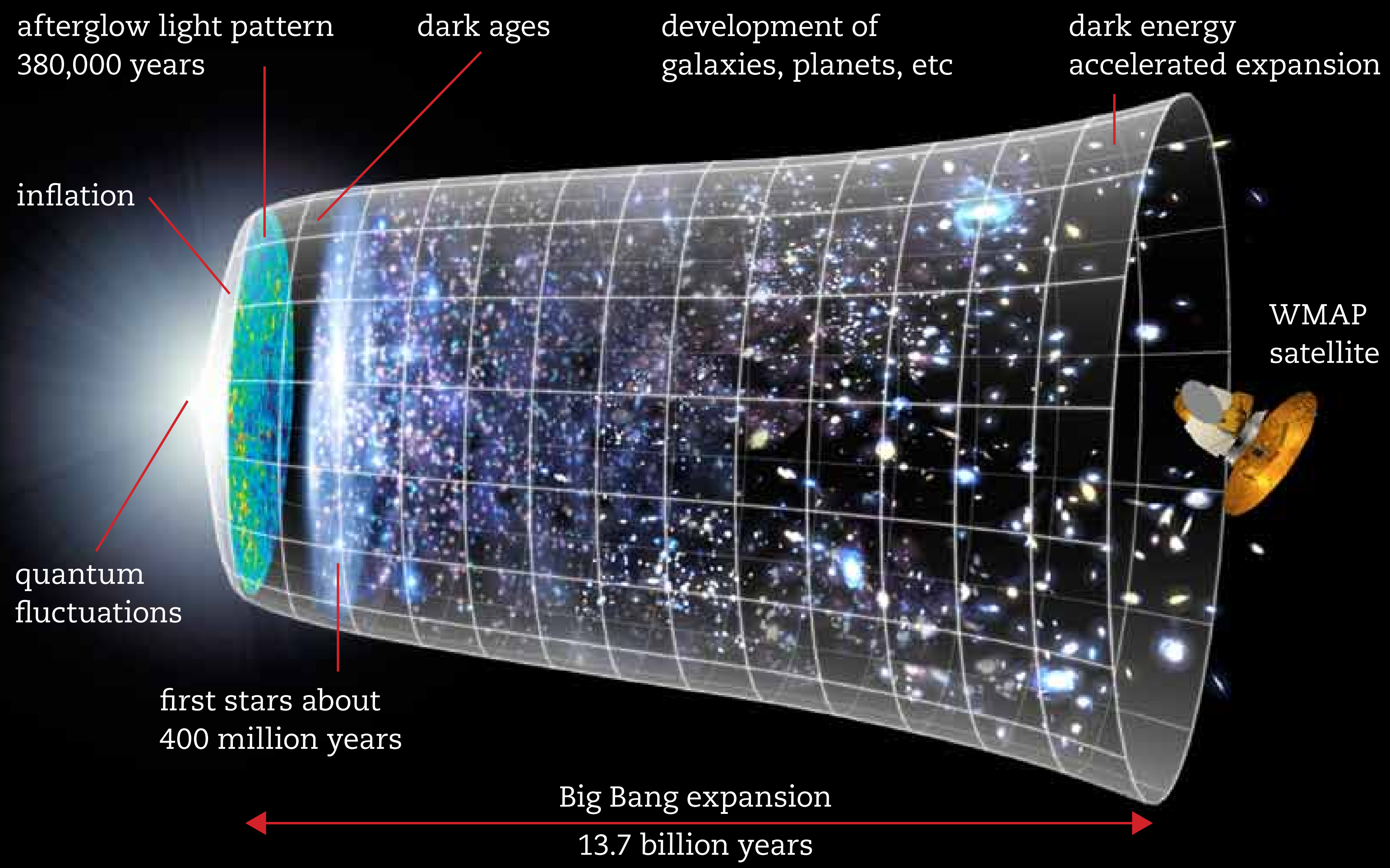
Cover: This image of very distant, faint galaxies shows part of the Hubble Ultra Deep Field. (NASA/ESA/S Beckwith [STScI] and the HUDF team)

INTRODUCTION

A remarkable fact about the universe we find ourselves in is that it is capable of sustaining a planet like the Earth and the complex chemistry of life.

On this planet we have achieved an understanding of the vast universe we inhabit, but this has been achieved only in the past century. It was only as recently as the 1920s that we began to get a glimpse of the vastness of the universe of galaxies. The discovery of the **microwave background radiation** and the realization that the universe began in a hot **Big Bang** dates back only to 1965.

And it is only since the beginning of the new millennium that cosmology has become a precision science, with a strong consensus emerging about what kind of universe we inhabit. ■



*A schematic illustration of the history of the universe, with time running from left to right. The earliest microscopic fluctuations in space-time are on the far left, followed by the Big Bang and the rapid expansion known as **inflation**. After 400,000 years the universe cools enough to become transparent but dark, until the first stars form about 400 million years later. Over the next 13 billion years the universe evolves, with bodies such as galaxies and planets forming, into the cosmos we see today. In the long term the expansion of the universe is accelerating, driven by recently discovered **dark energy**.* (NASA)

FIRST STEPS

First steps on the distance ladder

Aristotle (384–322 BC) was the first to estimate the size of the Earth, using the angle of the shadow of a pole at noon at a location 100 miles south of the equator. **Eratosthenes** and **Posidonius** later used a similar method. These latter estimates are within about 10% of the modern value.

In the second century BC, **Hipparchos** used a solar eclipse method to estimate the distance of the Moon and deduced a value of 59 Earth radii, a little under the modern value of 60.3. **Aristarchos** tried to estimate the distance of the Sun using a lunar eclipse, but was out by a factor of 20.

The Greeks also gave us Euclidean geometry (**Euclid** 300 BC), the idea of absolute, uniform time (Aristotle), and the idea of an infinite physical frame (the atomists, **Epicurus**). Interestingly, and contrary to the picture held by medieval thinkers, Aristotle believed that the stars were at a range of distances.

The Aristotelian Earth-centred (geocentric) universe reached its epitome in the detailed system of **Ptolemy** (first century AD), described in *The Almagest*. The Arabic-derived title reflects the crucial role of Arab scientists in preserving and extending the achievements of the Greeks. Certainly **Nicolaus Copernicus** (1473–1543) was aware of Arab work in his development of a Sun-centred (heliocentric) model of the solar system. A discovery of Copernicus that is less well-known is that he gave, for the first time, the correct relative distances of the Sun and planets. His values were within 5% of the modern values and the absolute scale of the solar system was not determined more accurately till the 19th century. The Copernican system also implied a huge increase in the minimum distance of the stars.

Nicolaus Copernicus's suggestion in 1543 that the planets circle the Sun is considered the origin of modern astronomy. Here he is imagined by the painter Jan Matejko in 'Conversation With God' (1872).

Galileo Galilei's discovery in 1609 of the moons of Jupiter lent weight to Copernicus's picture of the planets orbiting the Sun. His discovery of mountains on the Moon showed this was another world like the Earth. And his resolution of the Milky Way into stars was the first step into the universe of galaxies. Galileo's work on kinematics demonstrated the limitations of Aristotle's physics and paved the way for the system created by **Sir Isaac Newton** (Newtonian mechanics) and still used for most everyday situations today. In his interesting dialogue with the philosopher Richard Bentley, Newton also discussed the idea of an infinite universe of stars. ■

HOW BRIGHT, HOW FAR?

The inverse square law

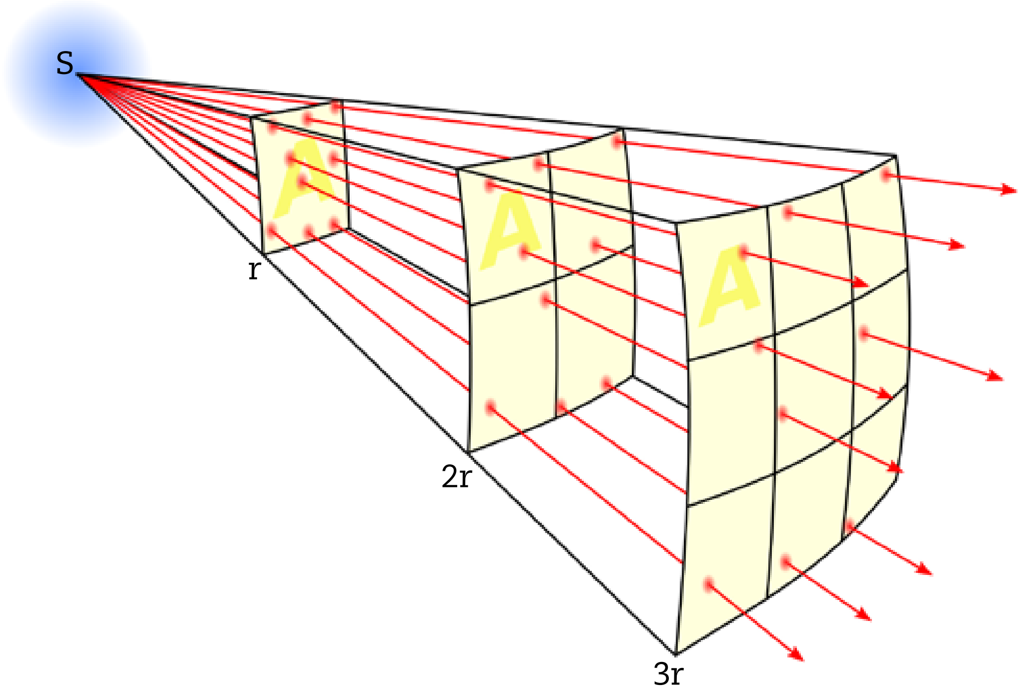
Newton tried to estimate the distances of the stars using the inverse-square law for light. Suppose that a star has luminosity L watts, so the total power of the light radiated from the surface of the star in all directions is L. At distance r (metres) from the star, this light is spread out over a sphere of area $4\pi r^2$ (m^2).

So at a distance r from the star the intensity of radiation I that we measure in a detector, for example the human eye, is $I = L/4\pi r^2$ (watts m^{-2}). Since I is proportional to the inverse square of the distance, this is called the **inverse-square law** for light (see figure).

Newton did not have a satisfactory way of estimating the intensities of light from stars, so he did not get sensible results, but this problem was solved during the 19th century. The ancient Greeks classified the stars visible to the naked eye on a magnitude scale from 1 to 6, where the brightest stars are magnitude 1 and the faintest magnitude 6. This system turned out to be an approximately logarithmic scale, with a change of about 2.5 for each change of 1 magnitude, so the **stellar magnitude scale** was defined as

$$m = \text{constant} - 2.5 \log_{10} I$$

where the constant depends on the observing waveband. (The apparent brightness of a star will appear different depending on the filter – red, yellow, blue, etc – used to observe it.) ▶



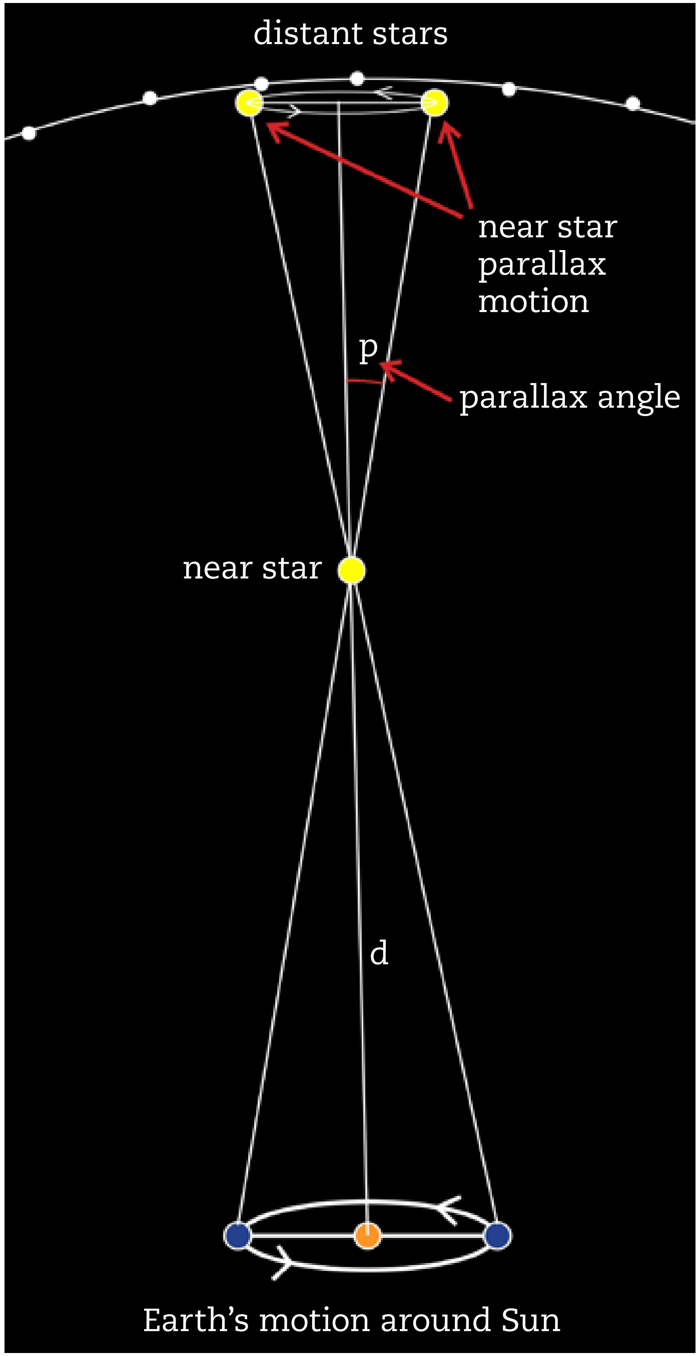
As light travels away from its source star S, it spreads out and weakens in intensity. Moving out from a distance of r to a distance of 2r, the light that previously covered a square of area r^2 spreads out to cover an area of $2r \times 2r = 4r^2$. When the same light reaches a distance of 3r, it covers an area of $3r \times 3r = 9r^2$. The intensity is inversely proportional to the area covered, which expands with the square of the distance, a rule known as the inverse square law.

Parallax

The first step on the distance ladder beyond the solar system was taken by the German astronomer **Friedrich Wilhelm Bessel** in 1838. He measured the parallax of the nearby star 61 Cygni – that is, the change in its apparent direction on the sky resulting from the movement of the Earth in its orbit around the Sun. This **parallax** phenomenon was the final proof of the Copernican system. Astronomers adopted a unit of distance related to parallax, the **parsec**. One parsec (1 pc) is the distance of a star when the half-angle subtended by the Earth's orbit at the star is one second of arc.

In 1781 the French comet-finder **Charles Messier** made a list of 110 objects which looked fuzzy or extended through a small telescope. Using much larger telescopes the British astronomer (and discoverer of Uranus) **William Herschel** showed that many of these are in fact clusters of stars.

In the 19th century astronomers used prisms and then gratings to disperse the light from astronomical objects into a **spectrum**. They found that many spectra had characteristic



Parallax is the small change in direction of nearby stars with respect to much more distant stars that results from the motion of the Earth around the Sun. Parallax allows us to estimate the distances of thousands of nearby stars. The parallax angle p is half the apparent change in position of the star at a distance d, with respect to more distant stars, seen in separate observations six months apart.

dark (absorption) or bright (emission) lines superimposed on them which indicated the presence of particular elements. This **spectroscopy** technique showed that some of the fuzzy objects were hot clouds of gas heated by young stars.

The nature of one remaining class of nebulae, the **spiral nebulae**, remained uncertain. Were they simply gas clouds in our Milky Way system, or could they be distant island universes (vast collections of stars now known as galaxies), as suggested by Christopher Wren, Thomas Wright and Immanuel Kant? ▶

Doppler shift

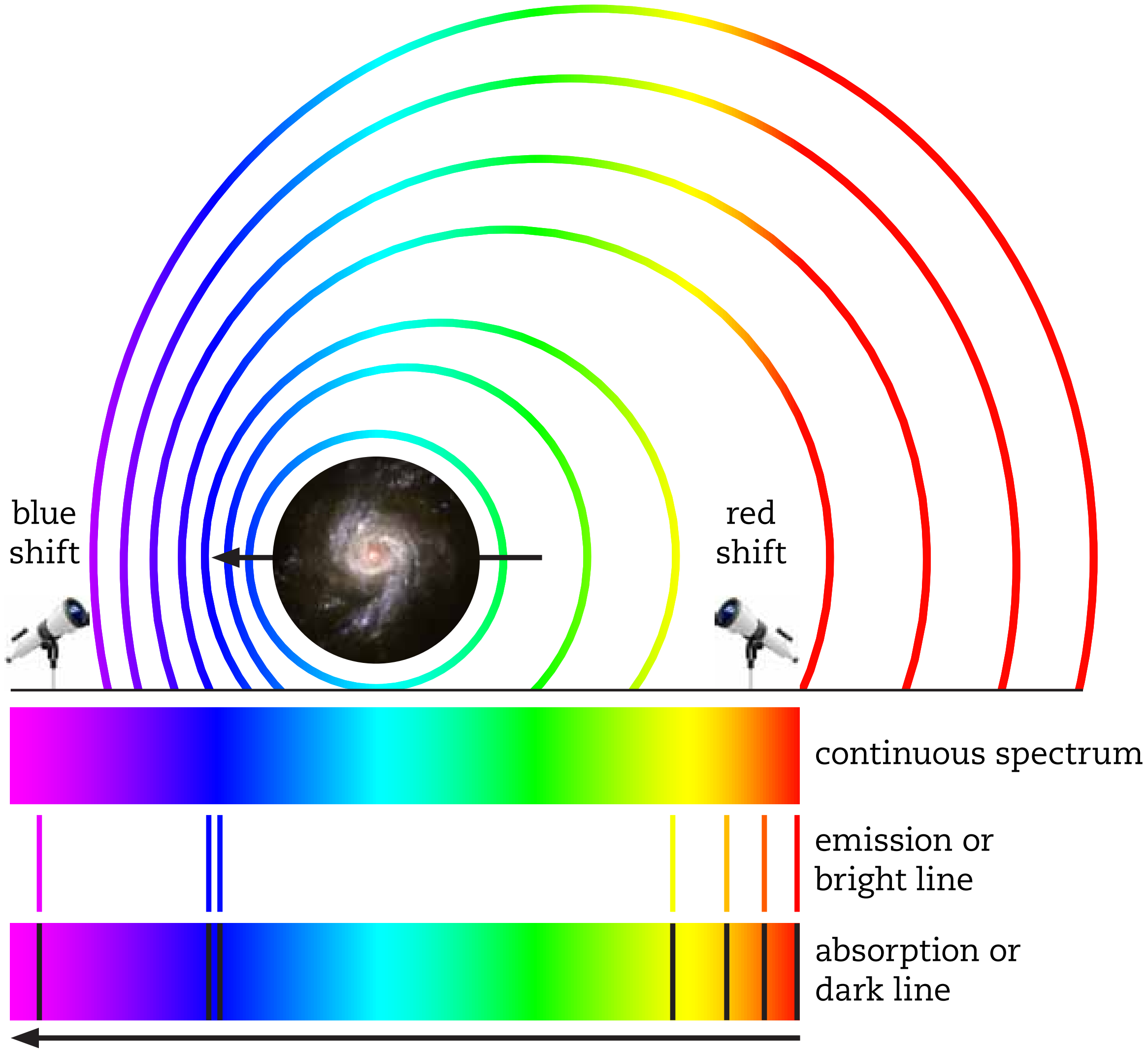
The advent of spectroscopy also allowed measurement of the velocity of stars in the line of sight, via the Doppler shifting of emission and absorption lines in their spectrum.

The Austrian physicist **Christian Doppler** made his discovery using sound waves in 1842. If a source emitting sound of wavelength λ is moving away from us, the wavelength λ will be increased by $\Delta\lambda = v/s$, where v is the velocity of the source away from us and s is the speed of sound. Similarly a source moving towards us has its wavelength shortened by the same amount. This happens because the waves get bunched together or stretched apart by the motion of the source relative to the observer.

For light the shift is $\Delta\lambda = v/c$, where c is the speed of light and the **redshift or blue shift** $z = \Delta\lambda/\lambda$.

The Doppler studies of stars in our galaxy showed that in addition to the broad circular motion of stars expected in a rotating disc, there were also random motions up and down, which broaden out the disc. So our galaxy is different from the solar system, where the planetary orbits are confined near a single plane.

Looking at the spectra of galaxies, astronomers can use this technique to measure the velocity at which they are receding from or sometimes approaching the Earth. ■



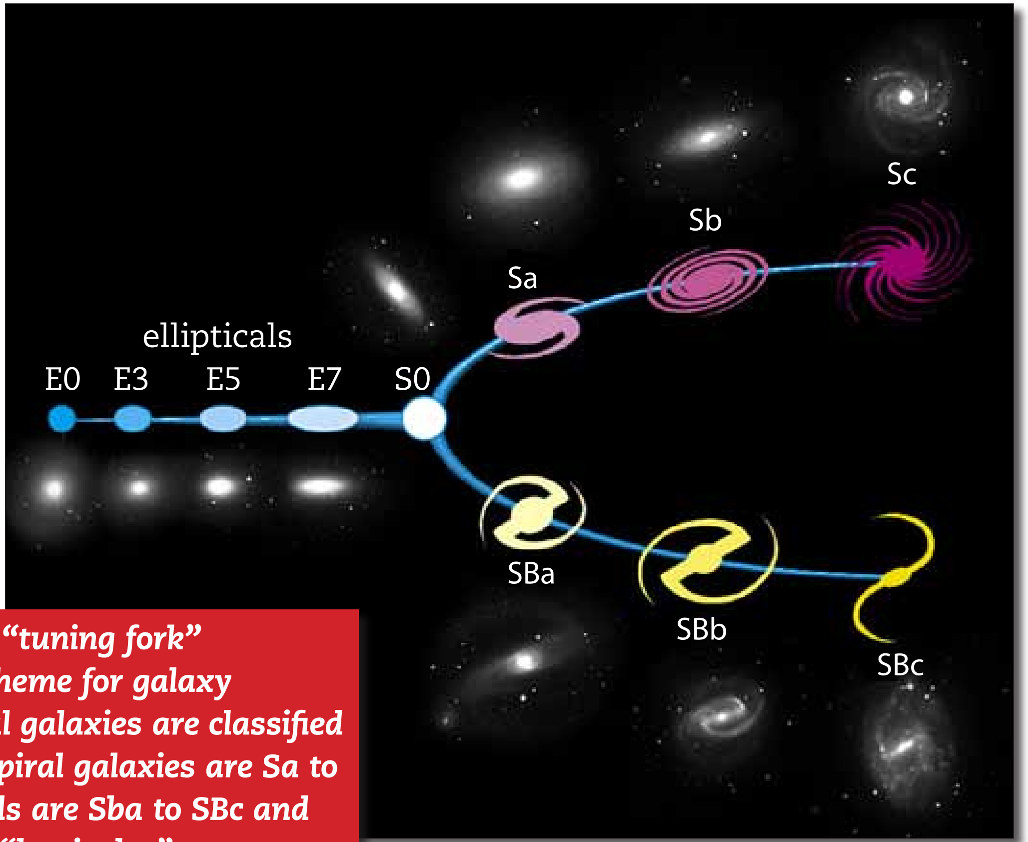
The Doppler shift, in which the wavelength of light (or sound) is shifted, depending on whether a source is approaching or receding. If a source is approaching the observer, the characteristic emission or absorption lines are at a shorter wavelength and shifted towards the blue end of the spectrum (blue shift). If the source is moving away, the lines are shifted to a longer wavelength towards the red end of the spectrum (redshift).

A UNIVERSE OF GALAXIES

Cepheid variable stars

The next crucial step on the distance ladder, still of prime importance today, was the work by **Henrietta Leavitt** on **Cepheid variable stars**. In 1912, while working at the **Harvard Observatory**, Leavitt discovered that these stars vary in luminosity in regular cycles, with the period of variation increasing with luminosity. This means that measuring the duration of the period of a Cepheid allows astronomers to deduce its luminosity. By comparing this with the observed brightness, the distance to the star can be calculated.

In 1924 **Edwin Hubble** used Leavitt's discovery to estimate the distance of Messier 31, the Andromeda Nebula (now known as the **Andromeda Galaxy**). It clearly lay far outside our own Milky Way Galaxy, thus resolving the long-standing controversy about spiral nebulae and opening up the universe of galaxies.



Edwin Hubble's "tuning fork" classification scheme for galaxy shapes. Elliptical galaxies are classified from E0 to E7, spiral galaxies are Sa to Sc, barred spirals are Sba to SBc and S0 galaxies are "lenticular". (NASA/ESA)

Dark matter

Hubble classified galaxies as **spiral, elliptical or irregular** (see figure). Modern spectroscopic studies showed that many galaxies, especially the spirals, are rotating systems, but the orbital speeds of the stars were found to be too large for the amount of visible matter they contain. Some other form of **dark matter** must be attracting the stars gravitationally.

Dark matter is also implied by the velocities of galaxies seen in clusters of galaxies. The final line of evidence for dark matter is provided by an effect known as **gravitational lensing**. When we see a distant background galaxy behind a foreground galaxy or cluster of galaxies, the gravitational bending of light predicted by Albert Einstein's **General Theory of Relativity** distorts the image of the galaxy, acting like a lens. This lensing effect allows us to map the dark matter in clusters of galaxies. ■

THE EXPANDING UNIVERSE

HUBBLE'S LAW AND THE EXPANSION OF THE UNIVERSE

In the first two decades of the 20th century, astronomers began to measure the velocities of nearby galaxies in our line of sight using the Doppler shift (**page 7**). The majority of these velocities were away from us, i.e. the galaxies are receding from us. A few galaxies, such as the Andromeda Galaxy, M31, are moving towards us. Edwin Hubble was also trying to estimate the distances of the galaxies, using Cepheid variable stars (**page 8**).

In 1929 Hubble announced, based on the distances of 18 galaxies, that the more distant a galaxy, the faster it is moving away from us:

velocity / distance = constant.

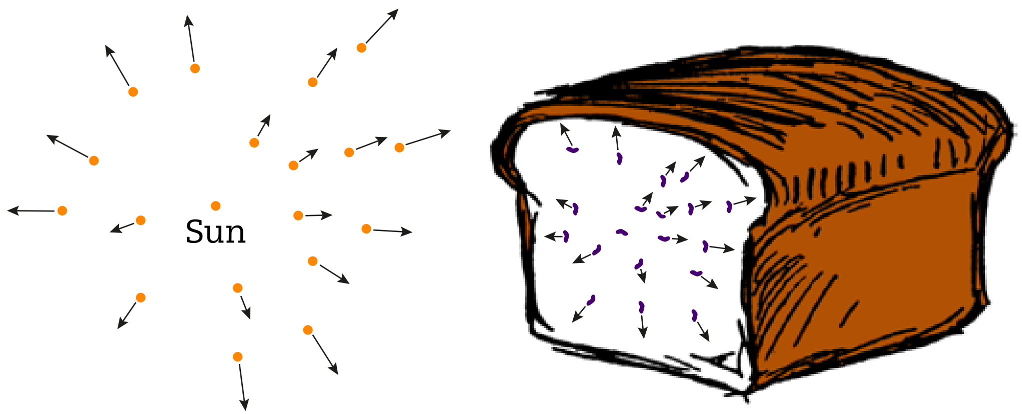
This constant is called the Hubble constant, or H_0 (**page 14**). **Hubble's Law** is often called the redshift–distance law, because the recession velocities of galaxies are measured by the shifting of their spectral lines towards the red end of the visible spectrum. This turned out to be just what would be expected in an **expanding universe**, the simplest possible model for the universe in Einstein's General Theory of Relativity (**page 10**).

For very high velocities (i.e. galaxies receding at a substantial fraction of the speed of light) the full and more complicated version of the redshift formula is used. Here the true distance to and velocity of the object being observed depends on the model of the universe in use.

The redshift z does give a relatively simple indication of how the scale of the universe has changed. If R_{now} is the scale factor (one measure of distance between two galaxies) in the present day and R_{then} the scale factor at the time when light left the galaxy being observed, then in this case the two are related by

$$1 + z = R_{now} / R_{then}$$

(1 + z) is then the factor by which the universe has expanded in the time taken for the light to travel from the galaxy to the telescope looking at it. ■



In the simplest model of the expanding universe, every galaxy moves away from us at a speed proportional to distance. And every other galaxy sees the same picture. An analogy is the raisins in a loaf of raisin bread which is expanding as it cooks. Over time the distance between each galaxy increases in the same way that the distance between each raisin increases.

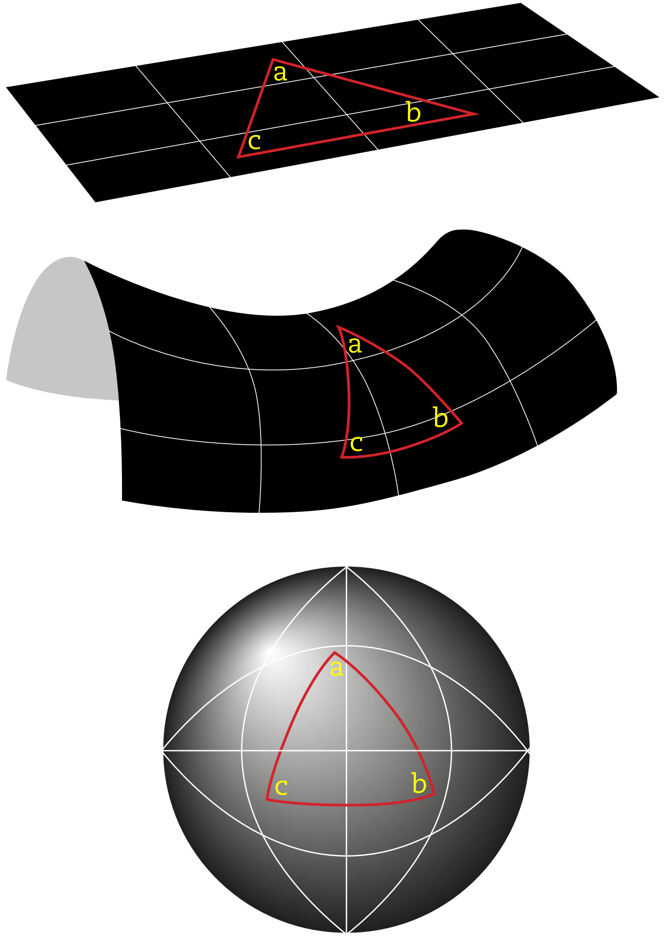
THE FLAT UNIVERSE?

Einstein's General Relativity

Einstein put forward his **General Theory of Relativity** in 1915. The main idea was that gravity is nothing more than the consequence of the bending of space-time around a massive body. So a mass distorts the space (and time) around it and a test particle then has to modify its orbit to follow the geometry of space. The geometry of space-time is curved by matter. In most situations, where the gravitational field is not too strong, Einstein's theory gives the same result as Newton's **inverse-square law of gravitation**.

In 1917 Einstein tried to apply his new theory to the whole universe. To do this he made a very dramatic assumption: the universe is (a) homogeneous (everyone sees the same picture) and (b) isotropic (the universe looks the same in every direction). Einstein used this unlikely assumption, the **cosmological principle**, to derive a static model of the universe in which the gravitational attraction between galaxies, which would tend to pull them all together, is balanced by a new force, the cosmological repulsion.

Einstein's inspired guess that the universe must be very simple (homogeneous and isotropic) is confirmed to very high accuracy today. However, his static model of the universe soon ran into trouble with Hubble's discovery that the universe is expanding. ▶



flat space:
a+b+c=180°
curvature=0

hyperbolic space:
a+b+c<180°
curvature=negative

spherical space:
a+b+c>180°
curvature=positive

If we take a slice of the universe at a fixed time, the 3D space may be (top) a flat, Euclidean space, in which the angles of a triangle add up to 180°, (middle) a space of negative curvature (hyperbolic space), in which the angles add to <180°, or (bottom) a space of positive curvature (spherical space), in which the angles of a triangle add to >180°.

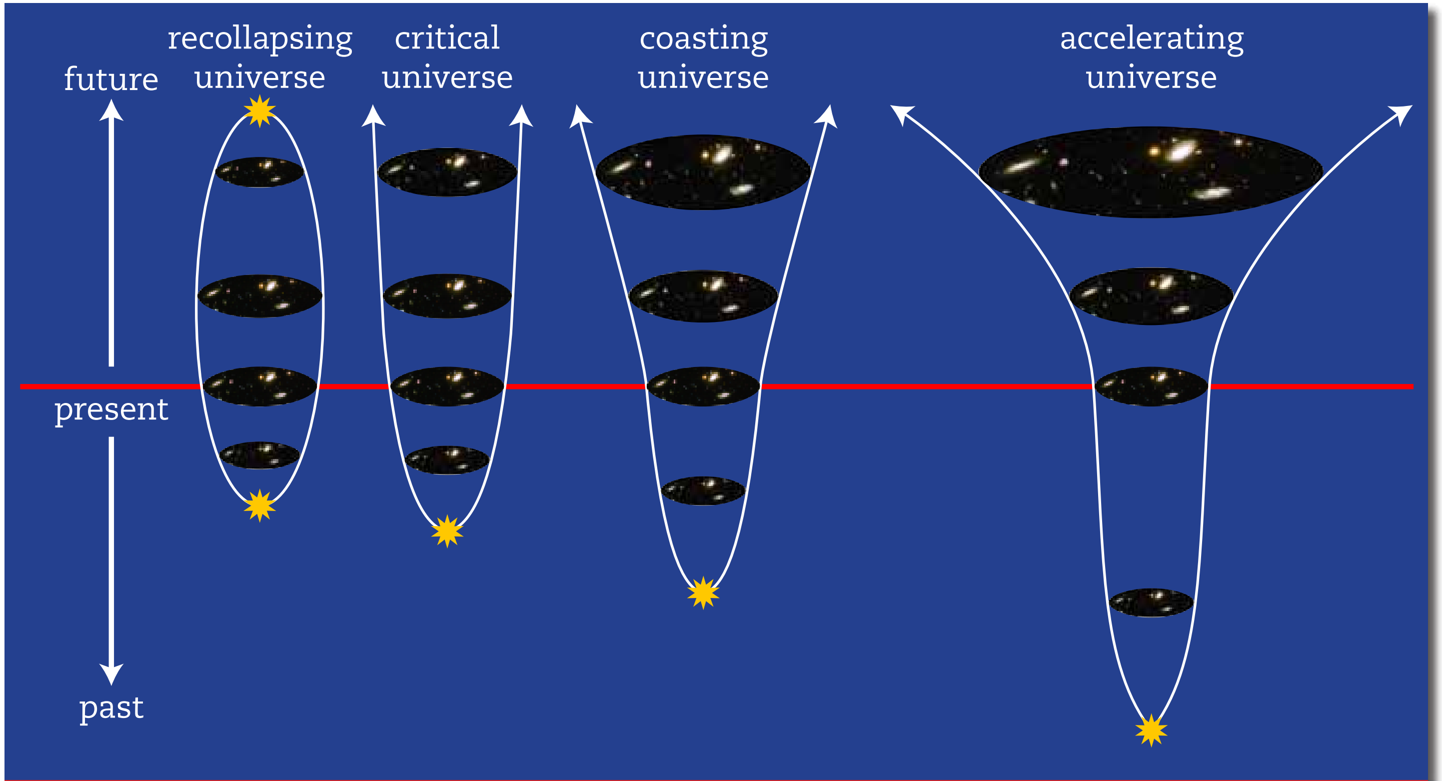
The Friedmann cosmological models

The Russian mathematician and meteorologist **Alexander Friedmann** had already shown in 1922 that expanding universe models are what would be expected according to Einstein's General Theory of Relativity, assuming the cosmological principle. Friedmann explored a whole class of models, including those with and without the cosmological repulsion.

The **fate of the universe** is determined by its composition. With just matter (including dark matter) the universe expands out from the Big Bang, gravity slowing down the expansion. Depending on the average density of matter in the universe, the universe may either reach a maximum extent and then fall back together in a "Big Crunch", or it may just keep on expanding for ever. There is a critical density that divides these two possibilities.

The fate of the universe is also connected, for a universe with matter only, with the spatial curvature of the universe. The critical density universe is spatially flat, while the higher density universe has positive spatial curvature and the lower density universe has negative spatial curvature.

In a universe with dark energy (the cosmological repulsion, **page 14**), the deceleration of the expansion due to gravity is eventually reversed and the universe begins an accelerated expansion. The observed universe seems to be just moving into this accelerated phase at the present epoch. Modern observations (**page 19**) suggest that the universe is close to being spatially flat. ■



The fate of the universe. With no cosmological repulsion, the expansion either keeps on going for ever, or reaches a maximum and recollapses to a Big Crunch. The cosmological repulsion generates an exponential acceleration of the expansion.

BUILDING PLANETS

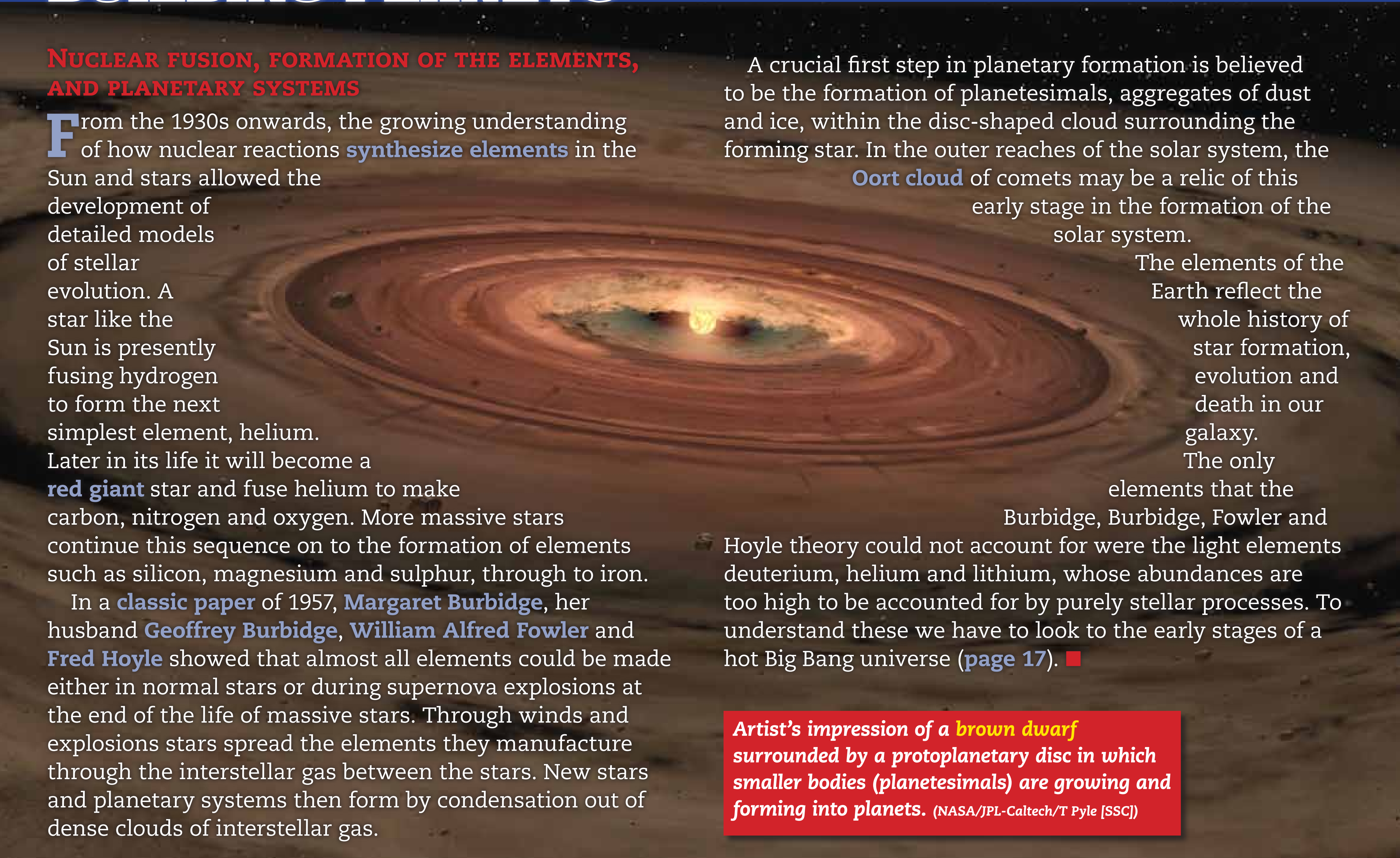
Nuclear fusion, formation of the elements, and planetary systems

From the 1930s onwards, the growing understanding of how nuclear reactions **synthesize elements** in the Sun and stars allowed the development of detailed models of stellar evolution. A star like the Sun is presently fusing hydrogen to form the next simplest element, helium. Later in its life it will become a **red giant** star and fuse helium to make carbon, nitrogen and oxygen. More massive stars continue this sequence on to the formation of elements such as silicon, magnesium and sulphur, through to iron.

In a **classic paper** of 1957, **Margaret Burbidge**, her husband **Geoffrey Burbidge**, **William Alfred Fowler** and **Fred Hoyle** showed that almost all elements could be made either in normal stars or during supernova explosions at the end of the life of massive stars. Through winds and explosions stars spread the elements they manufacture through the interstellar gas between the stars. New stars and planetary systems then form by condensation out of dense clouds of interstellar gas.

A crucial first step in planetary formation is believed to be the formation of planetesimals, aggregates of dust and ice, within the disc-shaped cloud surrounding the forming star. In the outer reaches of the solar system, the **Oort cloud** of comets may be a relic of this early stage in the formation of the solar system.

The elements of the Earth reflect the whole history of star formation, evolution and death in our galaxy. The only elements that the Burbidge, Burbidge, Fowler and Hoyle theory could not account for were the light elements deuterium, helium and lithium, whose abundances are too high to be accounted for by purely stellar processes. To understand these we have to look to the early stages of a hot Big Bang universe (**page 17**). ■



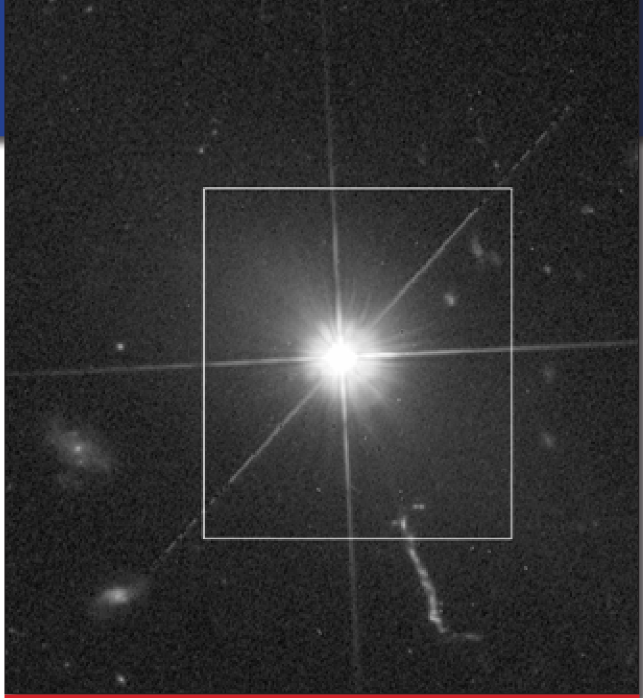
Artist's impression of a brown dwarf surrounded by a protoplanetary disc in which smaller bodies (planetesimals) are growing and forming into planets. (NASA/JPL-Caltech/T Pyle [SSC])

QUASARS

Controversy about the distance of quasars

The 1963 discovery of quasi-stellar radio sources – **quasars** – generated a new controversy about distance. First identified as bright radio sources apparently associated with stars, they soon turned out to have high redshifts. If the redshifts were interpreted according to Hubble's law, these were objects at huge distances, and hence with huge luminosities. Since some quasars were varying their light output on the timescale of months or even days, they had to be extremely compact systems, not much bigger than the solar system, yet generating much more power than our entire galaxy.

These strange properties led some astronomers to question whether the quasars really were at huge distances. Could the redshifts be caused by strong gravitational fields, or were quasars really fast-moving local objects thrown out in some spectacular galactic explosion?



The first quasar discovered, 3C273. Material leaves this quasar in a jet, part of which can be seen in the lower right. (NASA, J Bahcall [IAS], A Martel [JHU], H Ford [JHU], M Clampin [STScI], G Hartig [STScI], G Illingworth [UCO/Lick Observatory], ACS Science Team, ESA)

Black holes

Gradually evidence was accumulated that showed that at least some quasars are located in galaxies at the same redshift as the quasar. At the same time theorists began to model quasars as the product of massive **black holes**, 100 million times the mass of the Sun, in the nuclei of galaxies. Black holes originate as the remnants massive stars leave behind at the end of their lives and are so named because their gravitational fields are so strong that not even light travels quickly enough to escape their grip. Lower mass black holes are thought to be able to merge into larger ones, such as those found in the centre of galaxies.

Here the massive black holes might have formed initially from the collapse of a dense star cluster and have then grown by sucking in gas from their parent galaxies. Gas falling towards the black hole forms a rotating disc around it, which radiates at optical, ultraviolet and X-ray wavelengths. Today we have strong evidence for the black hole model and have even identified a smaller black hole, of one million solar masses, in the centre of our own galaxy.

Quasars have been detected out to redshift 7 (so far away that we see them as they were 12.9 billion years ago) and are among the most distant objects known in the universe. Today we think they are intimately connected with the formation and evolution of galaxies, with the black holes in their centres helping to regulate star formation. ■

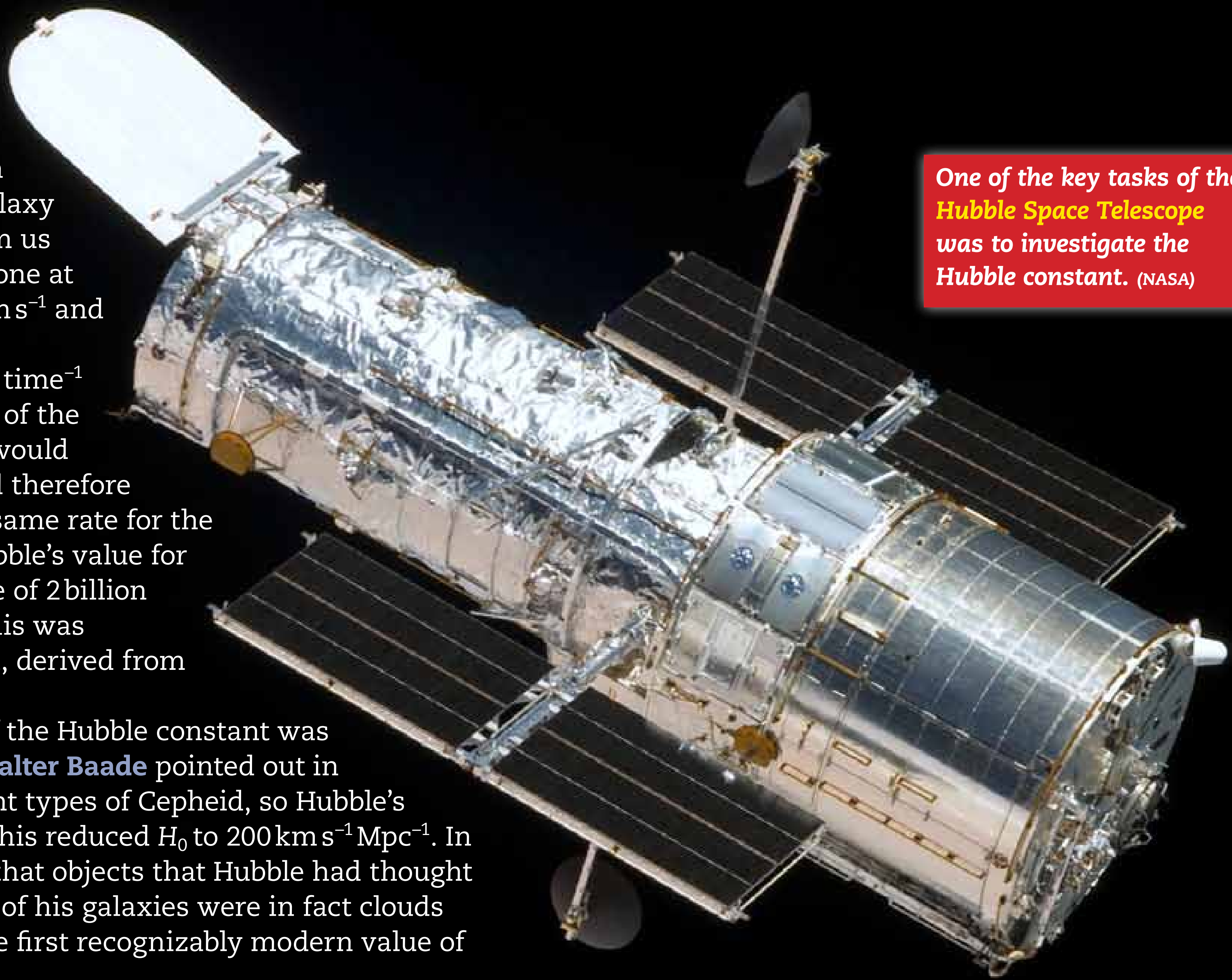
THE HUBBLE CONSTANT

The controversy over H_0

Hubble's 1927 estimate of H_0 (the "**Hubble constant**") was 500 km s^{-1} Mpc^{-1}, where 1 Megaparsec (Mpc) = 3.26 million light years. This means that a galaxy 3.26 million light years away from us would be receding at 500 km s^{-1}, one at 6.52 million light years at 1000 km s^{-1} and so on.

Now H_0 has the dimensions of time^{-1} and so $1/H_0$ is the expansion age of the universe – the age the universe would have if no forces were acting and therefore the expansion took place at the same rate for the whole history of the cosmos. Hubble's value for H_0 implied an age of the universe of 2 billion years and it was soon realized this was shorter than the age of the Earth, derived from **radioactive isotopes**.

From 1927 to 2001 the value of the Hubble constant was a matter of fierce controversy. **Walter Baade** pointed out in 1952 that there were two different types of Cepheid, so Hubble's calibration had been incorrect. This reduced H_0 to 200 km s^{-1} Mpc^{-1}. In 1958 **Allan Sandage** recognized that objects that Hubble had thought were the brightest stars in some of his galaxies were in fact clouds of hot gas; Sandage arrived at the first recognizably modern value of 75 km s^{-1} Mpc^{-1}. ▶



One of the key tasks of the Hubble Space Telescope was to investigate the Hubble constant. (NASA)

During the 1970s there was an acute disagreement between Sandage and **G Tammann**, on the one hand, favouring $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and **Gérard de Vaucouleurs**, on the other, favouring $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

HST Key Program

Following the launch of the **Hubble Space Telescope** (HST) in 1990, and the subsequent repair mission, substantial amounts of HST time were dedicated to measuring Cepheids in galaxies out to distances of 20 Mpc, to try to measure the Hubble constant accurately and to give the different distance methods a secure and consistent calibration. In 2001 the HST Key Program team, led by Wendy Freedman, announced their final result:

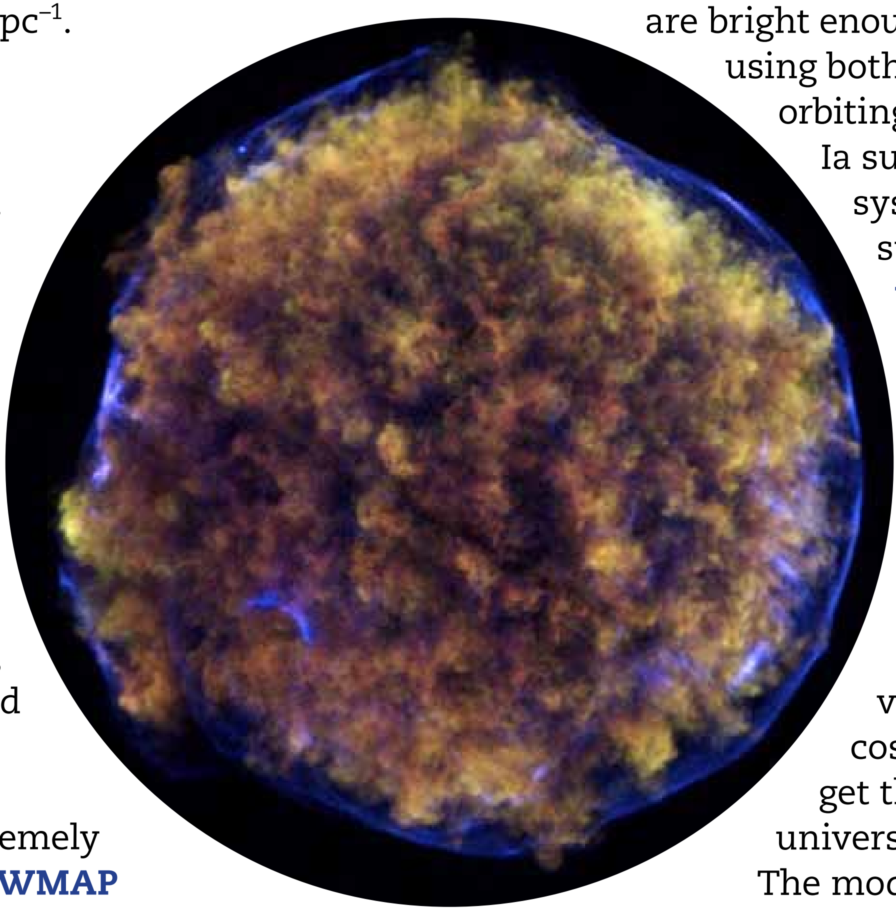
$$H_0 = 72 \pm 8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$$

This, as we shall see, agreed extremely well with the first results from the **WMAP** Cosmic Microwave Background mission ($72 \pm 5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$). It gave an age of the universe for an **Einstein–de Sitter model** of 9.1 billion years. This meant that a positive cosmological constant would be required for the age of the universe to be consistent with the age of the oldest stars.

Direct evidence for a positive cosmological constant came in 1998 from studies of **Type Ia supernovae** that are bright enough to be seen in distant galaxies, using both ground-based telescopes and the orbiting Hubble Space Telescope. Type Ia supernovae arise in binary star systems when the more massive star starts to dump gas on a companion **white dwarf**, the relic of a star like the Sun when it reaches the end of its life, which reaches a critical mass and explodes.

These supernovae explode with similar maximum luminosities, so measuring their brightness allows their distance to be calculated. The way this distance varies with redshift shows that a cosmological constant is needed to get the required flat geometry for the universe.

The modern interpretation of **Einstein's cosmological constant** is that it represents the energy-density of the vacuum, loosely called "**dark energy**", but a problem is that the observed value is 10^{120} times smaller than the value predicted by particle physics theories. ■



This is an X-ray image of Tycho's supernova remnant, the remains of a Type Ia supernova observed by Danish astronomer Tycho Brahe in 1572. Such explosions are used to measure distances because of their reliable brightness.

(NASA/CXC/Chinese Academy of Sciences/F Lu et al.)

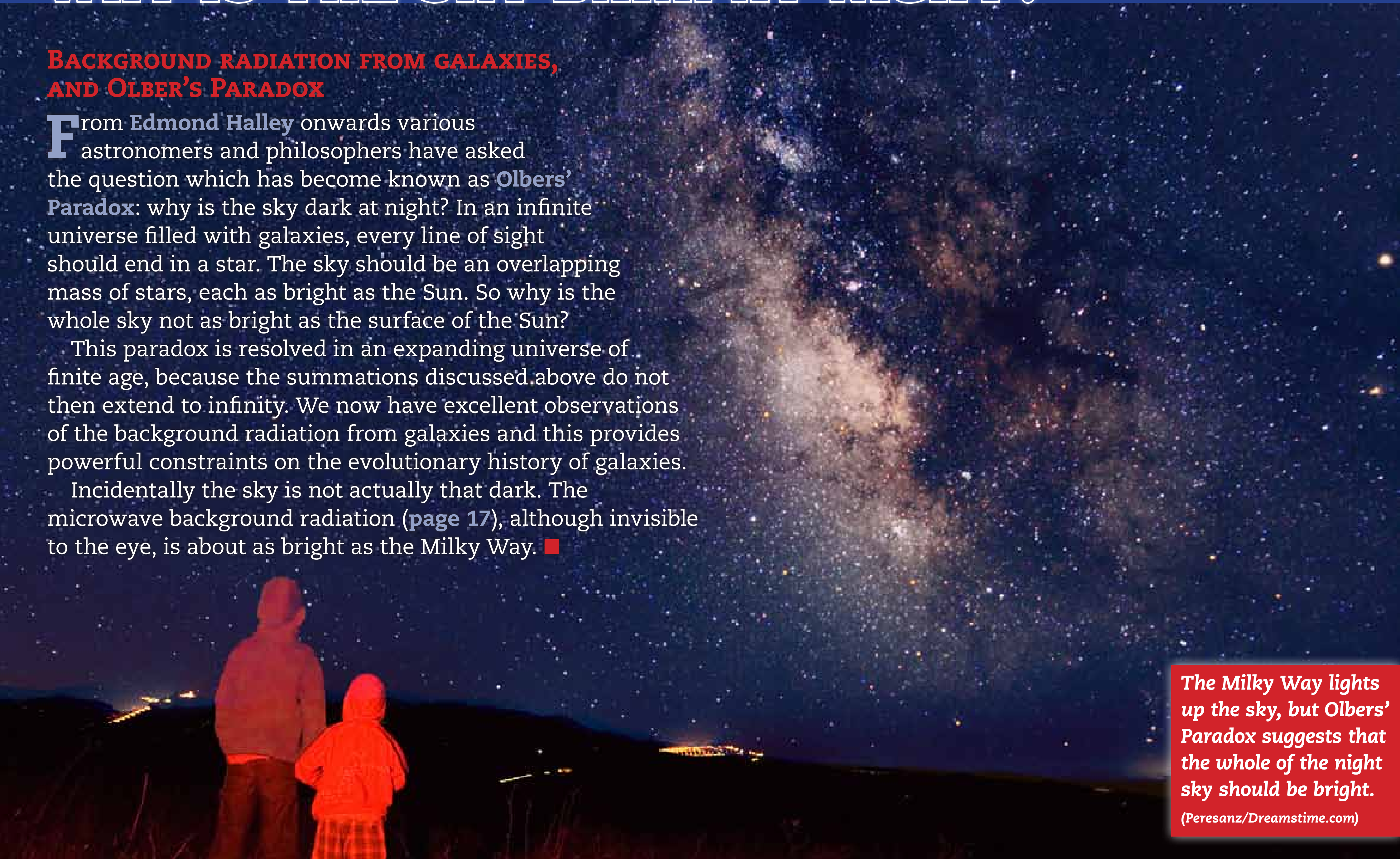
WHY IS THE SKY DARK AT NIGHT?

Background radiation from galaxies, and Olber's Paradox

From **Edmond Halley** onwards various astronomers and philosophers have asked the question which has become known as **Olbers' Paradox**: why is the sky dark at night? In an infinite universe filled with galaxies, every line of sight should end in a star. The sky should be an overlapping mass of stars, each as bright as the Sun. So why is the whole sky not as bright as the surface of the Sun?

This paradox is resolved in an expanding universe of finite age, because the summations discussed above do not then extend to infinity. We now have excellent observations of the background radiation from galaxies and this provides powerful constraints on the evolutionary history of galaxies.

Incidentally the sky is not actually that dark. The microwave background radiation (**page 17**), although invisible to the eye, is about as bright as the Milky Way. ■



The Milky Way lights up the sky, but Olbers' Paradox suggests that the whole of the night sky should be bright.

(Peresanz/Dreamstime.com)

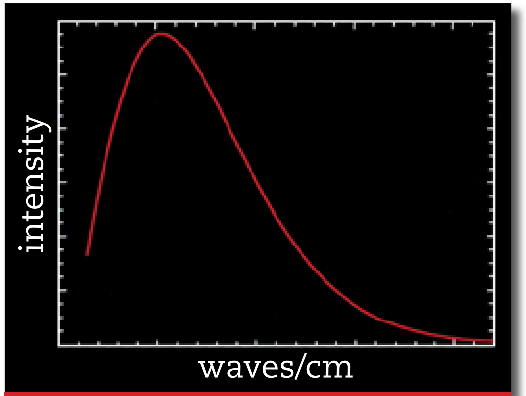
THE HOT BIG BANG

Blackbody radiation

In 1965 physicist **Arno Allan Penzias** and astronomer **Robert Woodrow Wilson** announced the discovery of the **cosmic microwave background** (CMB) radiation. This radiation is a relic of the “fireball” phase of the hot Big Bang universe, when the dominant form of energy was radiation. It has a perfect **Planck blackbody spectrum**, with a temperature of 2.7 Kelvin (2.7 K or –270.4 °C). The blackbody form of the spectrum tells us that at some time in the past the matter and radiation in the universe must have been locked together in thermal equilibrium.

If a piece of matter completely absorbs all the radiation falling upon it, or, conversely, behaves as a perfect radiator when heated, then the matter radiates as a black body, and the radiation has the characteristic blackbody spectrum, like the plot of intensity against wavelength in the figure. A practical example is the interior of an oven or furnace.

The spectrum peaks at a wavelength inversely proportional to the temperature of the black body. The Sun’s spectrum, approximately a black body of temperature 5800 K, peaks at a wavelength 0.6 µm, in yellow light. The Earth’s spectrum, at an average temperature of 288 K, peaks at 12 µm, at mid-infrared wavelengths. The cosmic microwave background peaks at 1 mm, corresponding to a temperature of 2.73 K.



The Planck blackbody spectrum of the CMB, measured by the COBE satellite team. (NASA)

Blackbody radiation is the signature of matter in thermal equilibrium with radiation, with the energy equally shared between the matter and the photons.

Particles and atoms

The universe appears to have formed in a single explosive event, the initial “singularity”, with essentially infinite density and temperature. As the universe expanded, the temperature and density dropped. Initially the universe was very simple, consisting of photons (particles of light), and particles of matter, divided into light particles (electrons and neutrinos) and quarks, which are the building blocks of heavier particles such as protons and neutrons. When the temperature dropped to 10^{12} K the quarks were confined to make protons and neutrons. When the temperature reached 10^{10} K, about 1 second after the Big Bang, **nuclear reactions** began and neutrons and protons fused together to make deuterium, helium and lithium. About 380,000 years after the Big Bang, when the temperature had dropped to 3000 K, electrons and protons combined together to make hydrogen atoms and the universe became transparent to radiation for the first time. This is the moment, called the **epoch of recombination**, we are looking at when we view the cosmic microwave background. Since this time the universe has expanded by a factor of about ▶

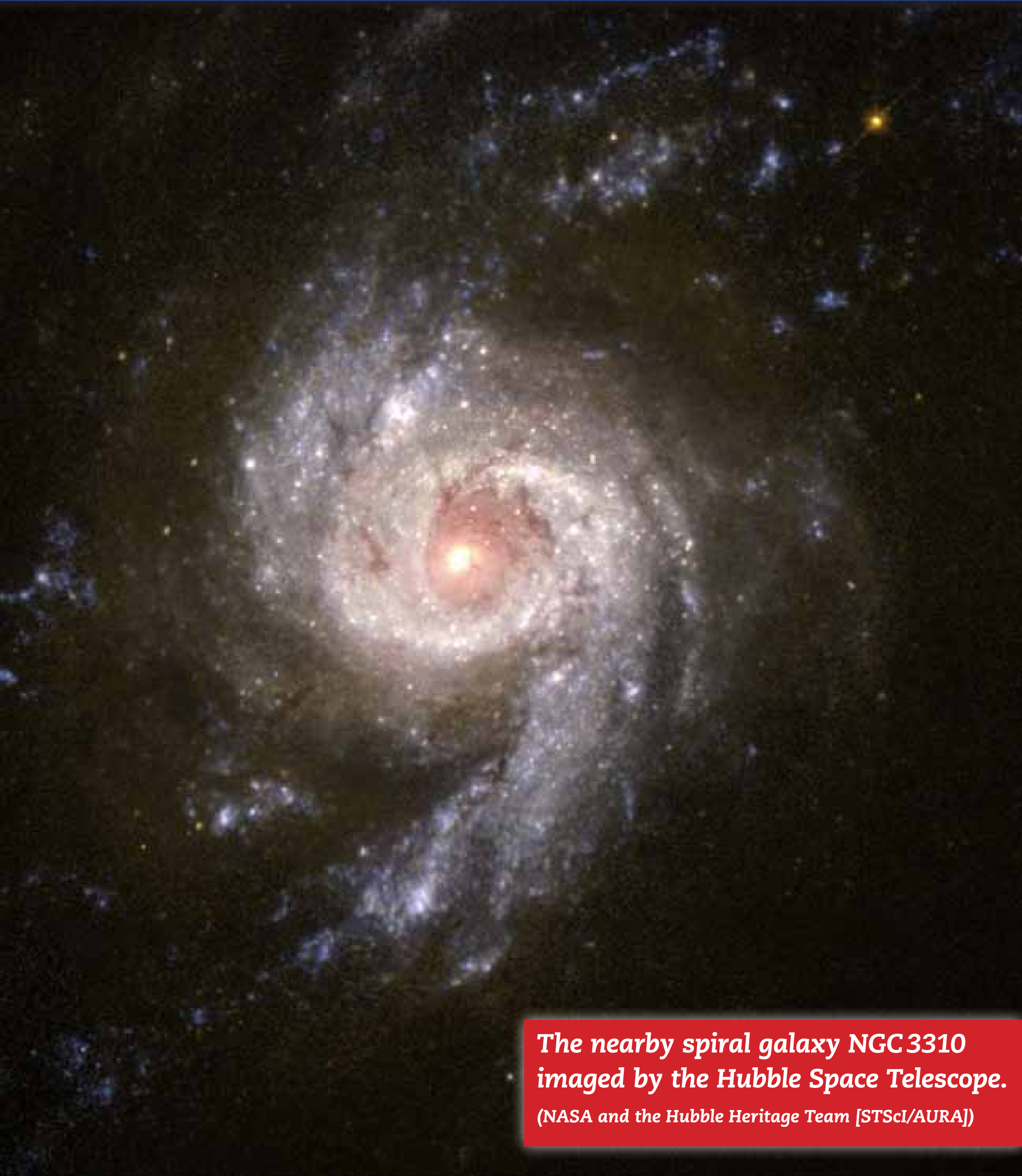
1100 in size and this accounts for the drop in the observed temperature of the microwave background radiation from 3000 K to 2.73 K.

Formation of galaxies

During the fireball phase, the ordinary matter that we and the Earth are made of – protons, neutrons and electrons – was extremely smoothly distributed, to better than one part in 100,000. For there to be planets, stars and galaxies today there had to be some small fluctuations in density present and these are believed to have been randomly distributed through the universe. These fluctuations were eventually shaped by gravity, which tended to concentrate matter to ultimately become galaxies of stars, planets, gas and dust.

To make galaxies, much stronger fluctuations must have already developed in the **dark matter**, which ceased to be controlled by radiation at a much earlier epoch. After recombination the ordinary matter responded to the gravitational attraction of the dark matter lumps and fell towards them, making proto-galaxies consisting of ordinary matter concentrations, in which stars started to form, embedded in halos of dark matter.

These proto-galaxies then merged together to make the galaxies we see today, with these mergers being accompanied by the vigorous formation of new stars. The gravitational assembly process also eventually collected galaxies together into groups and clusters. ■



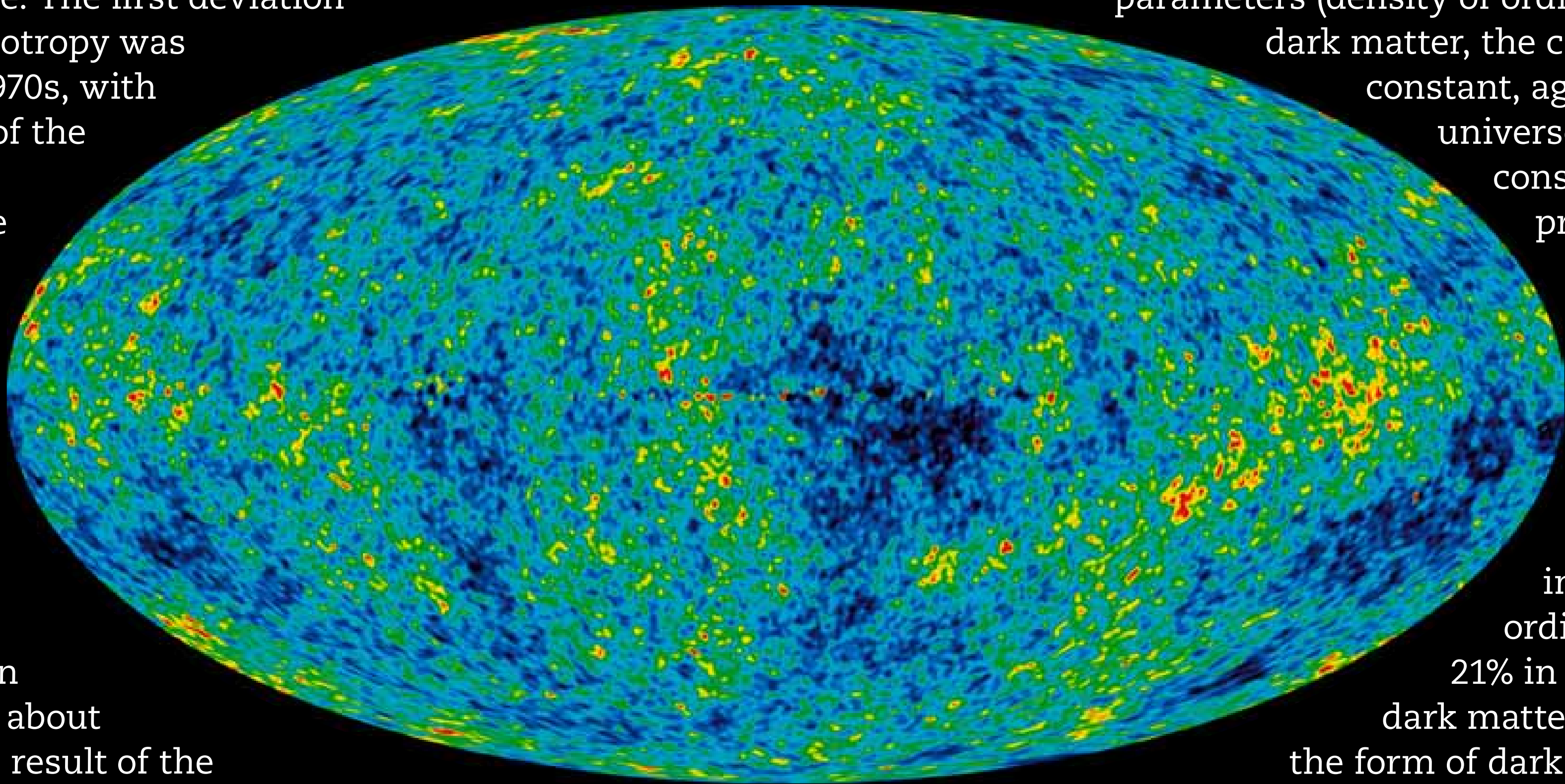
The nearby spiral galaxy NGC 3310 imaged by the Hubble Space Telescope.
(NASA and the Hubble Heritage Team [STScI/AURA])

THE COSMIC MICROWAVE BACKGROUND

Cosmic microwave background fluctuations

The remarkable **isotropy** of the **cosmic microwave background** (CMB) was the first real evidence that **Einstein's cosmological principle** is a good approximation for the universe. The first deviation from perfect isotropy was found in the 1970s, with the discovery of the CMB "dipole" anisotropy: the background is slightly hotter in one direction of the sky and cooler in the opposite direction. This is due to our galaxy's motion at a velocity of about 600 km s^{-1}, as a result of the combined gravitational pull of galaxies and clusters within about 300 million light years of us.

In 1992 the team working with COBE, the **Cosmic Background Explorer satellite**, found fluctuations in the temperature of the cosmic microwave background, at a level of about one part in 100,000. This was the crucial clue to how galaxies and clusters of galaxies formed in the universe. A later mission, **WMAP**, made precision measurements of the CMB fluctuations from which researchers were able to determine several cosmological parameters (density of ordinary and dark matter, the cosmological constant, age of the universe, Hubble constant) rather precisely. So today a tight consensus exists on the kind of universe we inhabit, with just 4% being in the form of ordinary matter, 21% in the form of dark matter, and 75% in the form of dark energy. ■



An all-sky map of the fluctuations in the cosmic microwave background radiation, made with the WMAP satellite. The radiation from our galaxy has been removed, as has the anisotropy produced by our galaxy's motion through the cosmic frame. The fluctuations have been highly magnified. In reality they are only one part in 100,000. (NASA/WMAP Science Team)

WHAT WAS BEFORE THE BIG BANG?

The puzzle of isotropy

It remains a puzzle why the universe is so isotropic and, at least in its early stages, so uniform in density. When we look at the cosmic microwave background in opposite directions on the sky, the regions we are looking at seem to have had no causal contact with each other – each lies far beyond the other's horizon – and yet they appear identical. A second problem is that the universe today seems to be quite close to being spatially flat. This requires the universe to have been incredibly close to flatness in its early stages. The horizon and flatness problems are partially solved by the idea that the universe went through a period of exponential expansion, "**inflation**", at or close to the Big Bang itself, driven by a phase with a very strong cosmological constant (**dark energy**). This period of inflation could have driven the universe closer to isotropy and uniformity.

The multiverse

Some cosmologists speculate that prior to the Big Bang the universe existed as a shadowy, chaotic quantum world. Some small region underwent a fluctuation and found itself with a very high vacuum energy, which triggered the exponential expansion of that region and the emergence of the universe we see today. In this picture there would be other expanding universes beyond our horizon, the **multiverse**. This allows scope for the **anthropic principle**, which argues that the properties of the universe have to be compatible with our presence, otherwise we would not be around to see it. In the multiverse model, most of the parallel universes would then be devoid of life.

Since the physics that we really know about extends back to the era of quark confinement, but no further, these ideas are essentially pure speculation. ■

Further reading

- **The Big Bang** J Silk (1980, W H Freeman)
- **Cosmology** M Rowan-Robinson (2005, 4th edition, OUP)
- **The First Three Minutes** S Weinberg (1993, Basic Books)
- **Nine Numbers of the Cosmos** M Rowan-Robinson (2001, OUP)
- **Life in the Universe** RAS leaflet (www.ras.org.uk/images/stories/ras_pdfs/Universe.pdf)

Useful websites

- **Cambridge Cosmology** (www.damtp.cam.ac.uk/research/gr/public/cos_home.html)
- **Ned Wright's Cosmology Tutorial** (www.astro.ucla.edu/~wright/cosmolog.htm)
- **WMAP's Universe** (map.gsfc.nasa.gov/universe)

Acknowledgments

- This eBook was compiled for the **Royal Astronomical Society** by the RAS Education Committee. © 2011. The RAS thanks its former President, Prof. Michael Rowan-Robinson, for supplying the original text. **www.ras.org.uk**
- The RAS is not responsible for the accuracy of the content of external material
- Designed by Paul Johnson at **Formula Media LLP** www.formulamedia.co.uk