Cosmology 4 – The Big Bang, the genesis of the Universe, the origin of the microwave background

The Big Bang

The Big Bang was totally silent. There was no-one about to hear it. There were not even any atoms and stuff made from atoms in the Universe at that time. Matter as we know it was formed later. In a way, the Big Bang wasn't big either. An almost infinitesimal fraction of a second after the Big Bang the entire universe was smaller than a full-stop, smaller than the smallest atom, smaller than the smallest nucleus, in fact incredibly small indeed. There was nothing big about it. The Big Bang wasn't even physics as we know it. It's what has come after the Big Bang that is physics. What was going on? You are about to find out.

a(t) = 0

Everyone has heard of the Big Bang theory of the origin on the Universe. It is the story of the Universe from the creation of what we know to the present day, as deduced by following the implications of the laws of physics back in time. The Big Bang is the origin, the statement that there was a time when the scale of the Universe was zero. In the notation of section 2 of these notes, $a(t) = 0$ at some time in the past, a time we now put at close to 13.8 billion years ago. The Big Bang theory used to be on stage with an alternative story called the steady-state theory, promoted by Fred Hoyle, one of the 20th century's great theoretical astronomers, and some other big names. Hoyle made major contributions to astronomy, not least pioneering the astrophysics of star evolution, but the steady-state theory has ad-hoc assumptions that don't fit the facts as we now know them. It was Fred Hoyle who coined the description 'Big Bang theory' in 1949 to describe the rival idea to his own. The name Big Bang has stuck. It is now the only scientific theory left standing and it has some remarkable successes. Hoyle eventually embraced Big Bang cosmology and was one of the pioneers in the mid 1960s who calculated the light element abundance predicted in a cooling universe.

[The photo on the slide shows Fred Hoyle as I remember him. There was a persistent rumour for decades in the Department that he had applied for a Professorship here in the 1950s and not got it. When he eventually came to Aberdeen as a visiting lecturer I asked him if the rumour was true. He said it was. The fact that his application was remembered here 40 years on quite tickled him, as he records in his autobiography. What a difference he would have made here! One of the observations that nailed shut the coffin of the steady-state universe of Hoyle and his colleagues Gold and Bondi was Martin Ryle's radio telescope result that galaxies in the earlier universe are closer together. As an encouragement to reading George Gamow's splendid book *Mr Tomkins in Paperback* I'll quote a few verses (written by his wife Barbara) from within:

'Your years of toil'
Said Ryle to Hoyle,
  'Are wasted years, believe me.
The steady state
Is out of date.
  Unless my eyes deceive me,

My telescope
Has dashed your hope;

  Your tenets are refuted.
Let me be terse;
Our universe
  Grows daily more diluted!'

Said Hoyle, 'You quote
Lemaître, I note,
  And Gamow. Well, forget them!
That errant gang
And their Big Bang—
  Why aid them and abet them?

You see, my friend
It has no end
  And there was no beginning,
As Bondi, Gold,
And I will hold
  Until our hair is thinning!'

'Not so', said Ryle
With rising bile
  And straining at the tether;
Far galaxies
Are, as one sees,
  More tightly packed together!'

There is more, but you'll have to look in the original book to see the next verse. (I'm not sure about the 1999 revision by Russell Stannard. Tinkering with a 'classic' text is bound to lose some of the appeal of the original).]

The Big Bang theory incorporates the expansion of the Universe, predicts the microwave background and its very accurately measured distribution with energy; it predicts the relative abundance of hydrogen and helium in the Universe that we find today and that almost all other elements were made within stars. It provides a story that has many testable features for how the Universe comes to be as we find it.

That said, the Big Bang theory is a tremendous extrapolation from our present day knowledge of the Universe to what happened, or could have happened, all those 13.8 billion years ago. Into that extrapolation goes all the particle physics we've learnt in the 20th century, the thermodynamics, the behaviour of radiation and matter, the dynamics of particles, General Relativity and the general experience of the success of theoretical physics in modelling systems by highlighting what is important and what can be ignored. However, as one famous cosmologist, William McRea, pointed out, '*From the observed properties of the present state of the Universe, we can infer less and less about earlier and earlier previous states*'. In other words, we're not going to get a view of the early Universe as if we were watching an archival film. Cosmology is going to produce only a very general picture that gets more blurred as we go further back in time. However, the general features of the Big Bang theory are so widely agreed that the theory is known as the 'concordance theory'. Cosmologists acknowledge that there are plenty of questions left to answer but the outline I'm going to give you will hold up to a lot of scrutiny.

Before I say some more about the beginning of the Universe, you have to remember that space and time don't exist independently of each other or independently of the matter in the universe. The Big Bang didn't happen at some particular point in a pre-existing space and time. The Big Bang was the creation of space, time, matter and motion - the whole Universe as we know it. Space and time are defined by measurements within the Universe; neither concept exists outside the Universe and hence the question of what happened before the Big Bang makes no sense. It's like a geographer asking "what place is north of the North Pole?" It's a pseudo-question that doesn't make sense in the context of the subject. Likewise, asking about what happened before the Big Bang.

How far back?

How far back in time can you push the laws of Physics? Very far, according to physicists. General Relativity doesn't include any quantum effects of space and time so you can push the laws of physics back until possible quantum effects on space and time might be important. That's not quite the instant of the Big Bang but to about 10^{-43} seconds afterwards. The implicit prediction by General Relativity that there was probably an origin to the Universe was made by Alexander Friedmann in the early 1920s, as mentioned previously, and again by a Belgian priest and cosmologist Georges Lemaître in the late 1920s. However, neither of these two tried to predict the conditions that would exist near the beginning of the universe. The Hot Big Bang theory, so called, was first described more than 50 years ago, in 1948, by a famous Russian émigré physicist, George Gamow, who was living in the US; the same Gamow who authored the Mr Tomkins stories. While in his 20s, Gamow had been a student of Friedmann in St Petersburg. Detail of the Big Bang theory has since been developed by various theoretical physicists, including Steven Weinberg, Roger Penrose, Alan Guth and others you may have heard of. The theory says that at the Big Bang the matter created was intensely hot and rapidly expanding. As the Universe expanded and cooled it ran through a range of epochs, most of which lasted a very short duration until the epoch we are now in. Each epoch was characterised by mass and energy in a different state. I'll shortly summarise the history of the Universe in the next paragraphs but before then there is a digression.

The Planck units of space and time

You may not have thought about it but our basic units of measurement, the metre, kilogram and the second, are not in any way fundamental to nature. They are pretty arbitrary units related to the size we are, the masses we tend to move around on a daily basis and how quickly we do things. They aren't in any way related to anything basic about the size, mass and timescale associated with the atoms of the universe, or the radiation in the universe, or the strength of fundamental forces in the universe, or anything like that. Do we not know enough about the universe to see that there are more fundamental units in nature? For example, the speed of light in vacuum seems to be a pretty fundamental quantity in the universe. Shouldn't we make its size unity?

Max Planck, the originator of the idea of the quantum of radiation, was not the first to ask such a question but his own answers are the ones that have been found useful. He asked himself if there were combinations of the fundamental constants of nature that provided basic and natural units for length, mass and time. He had come up with the relationship $E = hf$ for the quantum of energy, E, associated with radiation of frequency f. The constant h, not to be confused with the cosmological constant h associated with Hubble's constant, is known as

Planck's constant and has a value 6.64×10^{-34} J s in the SI system of units. Planck's natural units of length and time are formed from combinations of the key physical constants G, h and c, one from each of the central strands of physics, namely gravity, electromagnetism and quantum theory. They are l_{pl} and t_{pl} :

$$l_{pl} = \left(\frac{Gh}{c^3}\right)^{1/2} = 4.13\times10^{-35} \text{ m} \tag{4.1}$$

$$t_{pl} = \left(\frac{Gh}{c^5}\right)^{1/2} = 1.38\times10^{-43} \text{ s} \tag{4.2}$$

These units are ridiculously small to use for the activities of everyday life but they indicate that the universe naturally operates on a scale that we find hard to imagine. You may have noticed that light travels unit length l_{pl} in unit time t_{pl} and hence the speed of light in Planck units is indeed unity. If you measure G, h, and c in any consistent set of units you get the same physical sizes for l_{pl} and $t_{pl.}$ The numbers will be different in different sets of units but they will correspond to the same lengths and times. John Barrow has commented that the laws of physics properly formulated in universal units are the ultimate shared experience for everyone who inhabits our universe.

[It's not uncommon these days to hear people who have little experience of science branding science as just a social construct, like the plays of Shakespeare, the music of Beethoven or the poetry of T S Eliot. This is a viewpoint I strongly disagree with. If, or perhaps it's a case of when, humanity finally meets alien intelligence and can converse with it, then I'm sure the plays of Shakespeare, the music of Beethoven and the poems of T S Eliot will be news to the aliens but Newton's laws of motion, Maxwell's equations of electricity and magnetism, the laws of thermodynamics, Shannon's noisy channel communications limit and a slew of other knowledge we know as the laws of nature, as well as a great deal of mathematics, will be known to the aliens. In a trivial sense our articulation of the laws of nature are human constructs but I believe they embody truths about the Universe that any intelligence who shares this universe with us will also find.]

To return to Planck's units, the point of mentioning these units to you is that the Universe evolved through its early epochs in times that are ridiculously short when measured in seconds but are very significant when measured in Planck time units, as you'll see shortly. Secondly, Planck units are derived using the quantum of radiation. There is good reason to believe that General Relativity, which isn't a quantum theory, will work well provided we deal with dimensions much larger than the Planck length and events that take place in times much longer that the Planck time. This means that when we can expect to be able to extrapolate the laws of physics back to the first 10^{-40} s in the Universe, say, but not quite back to the origin. To get back further needs a theory of quantum gravity, which is a topic that has been getting a lot of attention these past 10 years from Stephen Hawking, Roger Penrose and other theoreticians at the cutting edge.

The hot Big Bang scenario - the beginning of time

The very, very beginning of the Big Bang is speculation, involving physics we have not yet established. There is plenty of speculation around, involving M-branes, string theory and other quantum gravity ideas. However, it doesn't take very long at all before we get into the realm where the physics we know starts to apply. Are you ready to hear the story?

From zero to our Universe in three-quarters of an hour is a bit of a roller-coaster ride. No-one could have invented this story. It comes from the science we understand, not from the imagination. I mentioned earlier that the view I'm going to tell you about is described nowadays as the 'concordance' view of cosmology, reflecting the widespread scientific agreement on the contents. It is essentially the starting point that science today tries to refine, to examine the weak points and to provide more detail. You will need to concentrate.

There's one more paragraph I want to insert before we go back in time as far as we can. The universe today is expanding and cooling. In the past it was hotter, much hotter. Matter as we know it doesn't like heat. Things fall apart in the heat. Paper, wood, plastics and so on burn and turn to ashes in chemical reactions that are too slow at room temperature but go like wildfire in the heat. Solids melt in the heat and turn to liquid; liquids turn to vapour with more heat. By the time the temperature reaches say 5000 degrees, no solids or liquids exist (at ordinary pressures). By 5000 degrees, atoms themselves are starting to fall part, losing electrons and ionising. Raise the temperature to say 1 million degrees, much lower than in the core of the Sun, and atoms have disintegrated into free nuclei and electrons. Raise the temperature to beyond 10^{14} K and the very protons and neutrons that make up the nuclei of atoms will disintegrate into elemental quarks and other particles. The story of the very beginning of the universe is therefore not a story of matter as we know it but of totally and utterly disintegrated matter. Now we can start.

Time: ~ 10^{-40} s or $1000t_{pl}$. The size of the Universe is $1000l_{pl}$. The universe was dominated by energy in the form of radiation. The Universe was small enough that radiation could travel across it, allowing it to be at one temperature, albeit an immensely high temperature of around 10^{30} K and at an incredible density. You can predict the temperature will be as high as this from the radiation physics that came into our *meteorology* or *light science* courses, namely Wien's radiation law. This law gives the wavelength, λ_{max}, where the radiation spectrum at a given temperature T has its maximum energy. The law says that T and λ_{max} are simply related by $T = 3\times10^{-3}/\lambda_{max}$, where λ_{max} is in metres. Putting $\lambda_{max} \approx 100\times l_{pl} = 4\times10^{-33}$ m, gives $T \approx 10^{30}$ K. It's a very rough figure but gives you a flavour that the early conditions of the Universe are found by taking the laws we know to the limit. The quick estimate works because the early Universe was dominated by radiation, not matter as we know it.

The energy was so high that particles and antiparticles were continually created out of the radiation only to disappear again by annihilation, turning back into radiation. This kind of situation is known in modern physics by the term **quantum fluctuations**. The number of particles of a given kind fluctuated as these transformations took place. This interchange between particles and anti-particles can represented by:

$$\text{particle + anti-particle} \leftrightarrow \text{radiation (a pair of photons).}$$

You may remember from our *astronomy* course that energy and matter are essentially the same thing, according to Einstein's famous relationship $E = mc^2$. We are talking here of energy that sometimes appears in the form of particles with kinetic energy and momentum and at other times just radiation photons. Radiation is the dominant form of energy. At this stage in the life of the Universe the 'particles' involved were not protons or neutrons, which are not truly elementary particles, but particles such as quarks, electrons and positrons.

The laws of physics as we find them applying to the elementary particles we know about conserve the balance of matter and antimatter in these reactions. This is a problem for cosmologists because in the observable universe matter is much more common than antimatter. How did it get that way? One of the curious features of our universe is that by far the most common objects in it are photons. Only a tiny fraction of the entities of the early universe will end up as matter. ‘Tiny’ here is of the order of 1 in a billion. So the imbalance in the creation of matter and antimatter that must have been present to produce the matter dominated universe we now know would have been very small indeed. But when did it occur? This question is one of the key questions in modern cosmology. It’s a question that hasn’t been clearly answered.

The commonly given story is that the imbalance between matter and antimatter occurred in the very first epoch of the Universe. The energy per particle was so high that 3 fundamental forces we identify in nature, **the strong force,** that holds protons and neutrons together, **the weak force**, that is responsible for phenomena like β decay, and **the electromagnetic force** are not differentiated. This fact is a consequence of today’s theories of very high-energy particles, known as ‘Grand Unification Theories’ or GUTs. Particle properties and interactions were certainly different in the conditions of the early universe and, the story goes, it is likely that the asymmetry arose then. Exactly how? The point just illustrates that modern cosmology is an evolving subject that doesn’t claim to have definitive answers to all the detail. Absence of all the detail doesn’t prevent the rest of the story being true.

Inflation

At around 10^{-35} s or $10^{8}t_{pl}$ a new process kicked in called **inflation**. Inflation involved the exponential increase in the space of the Universe, doubling its size every 10^{-35} s or every 10^{8} t_{pl}. Cosmological models didn’t used to have this concept in them but they could not explain the uniformity of the universe as we now find it. A theoretician called Alan Guth gets much of the credit for introducing the idea in 1981. Others had similar ideas at round about the same time and the detail has since been worked out at some length and is now generally accepted as part of the standard cosmological model. Inflation continued to double the size of the universe every 10^{-35}s for about a hundred times, finishing after 10^{-33} s. It looks better in Planck units to say that it finished in 10^{10} units, i.e. took 100 million units of time for every doubling and finished in 10 billion units.

The inflationary period involved an expansion of space faster than the speed of light. Let’s say that the universe started at 10^{8} units across. After 100 doublings it becomes $2^{100}\times10^{8}$ units, which is about 10^{38} units, or 10^{30} times bigger. At this size it would take radiation $10^{38}t_{pl}$ to travel across but the age of the Universe is ‘only’ 10^{-33} s or $10^{10}t_{pl}$ units. Remember that the speed of radiation is 1 in Planck units. So after inflation opposite sides of the Universe are not in radiation contact with each other but have all been derived from a pre-inflation universe that was in radiation contact and hence at the same temperature. This accounts for the fact that the background radiation in the Universe that we observe today is highly isotropic and homogenous, in fact to about 1 part in 100,000.

It’s hard to get a picture of how huge the inflation was from these figures. Let’s make an analogy. Suppose you are watching a ball expanding from almost nothing. After a short time, say 1 second, it has reached 1 cm across and you expect it to keep inflating so that when you look again about 100 seconds later it has expanded to 1 m across. If, however, inflation set in at the rate that seems to have happened in the early Universe, namely a factor of 2 for every time interval, then your ball would be 2^{100} times bigger, or 10^{28} m across when you looked back. That's about a million, million light years across. Now that's serious inflation!

Implications of inflation

Inflation has three other implications. It implies that the Universe as a whole is much bigger than the part we can explore with light or radio waves. There are other parts of it that will always be inaccessible to us. This doesn't at all imply that the Universe is infinite, just that there is matter not visible to us. Secondly, or really as a consequence of this, inflation accounts for the Universe being 'flat', neither open nor closed to a very high degree of accuracy. It does this simply by noting that the observable universe we inhabit is only a small part of the total universe created. Although General Relativity says that mass curves space-time, it also says that any small part of the totality of space will appear almost flat. According to inflation, our observable universe is only a small part of the totality and hence it is likely to appear flat. In fact this is the result we see experimentally. Finally, and this is more subtle, inflation expanded very small-scale inhomogeneities produced in the pre-inflation universe by quantum fluctuations into large-scale, but still very small, inhomogeneities. As the universe expanded after inflation, these now large-scale effects became in due course the seeds for the gravitationally induced structure we now see in the universe. Amongst other features to its credit, inflation explains how large-scale structure in the universe got started. This is a very significant question and one we'll come back to both later in this chapter and in the next chapter.

In short, inflation extends pre-inflationary cosmology so that it now makes sense of facts that couldn't previously be explained. However it does so by introducing a new concept as an add-on, rather finding it is a previously unrecognised consequence of fundamental principles. If an earlier theory is incomplete in some way then add-ons will be necessary. However, although inflation is seen as a cause of subsequent properties of the Universe, it is itself the effect of underlying principles that aren't fully understood. For example, is there another field present in the Universe that we haven't recognised and that necessarily drove inflation for the times it seems to have been necessary to account for our present day experience of the Universe? Inflation is an explanation that raises other questions not yet answered. This isn't necessarily a bad thing. Science evolves by processes like this.

First separation of forces

As the temperature and density of the Universe falls, the strong nuclear force and the so-called 'electroweak force' became distinct. The strong force is carried by particles called, appropriately enough, gluons. The electroweak force is carried by particles known as W and Z particles (bosons), both of which have been observed in high energy accelerator interactions at CERN. Indeed, Carlo Rubbia and Simon Van Der Meer won the Nobel Prize in Physics in 1984 for their contributions to this discovery. In 1979, Sheldon Glashow, Abdus Salam and Steven Weinberg had won the Nobel Prize in Physics for their theoretical work that showed how to unify the weak and electromagnetic interactions. So the physics of electroweak interactions is established.

The strong force (which the name alone suggests correctly becomes much bigger that the weak force as the Universe cools) governs the binding of quarks into the class of elementary particles called hadrons, of which the obvious ones are protons and neutrons. Both of these particles consist of triplets of 'up' and 'down' quarks. However, the temperature was still so high that the Universe was dominated by a soup of quarks and antiquarks, electrons and positrons (which are anti-electrons) and radiation. This effect is comparatively simple physics. If the temperature is high then the kinetic energy of collisions will break apart bound particles. It is just for this reason that there are no molecules in the Sun, for the temperature there is too high for them to exist. At the epoch we are talking about in the Universe, the temperature was too high for even protons and neutrons to exist.

Second separation of forces

At about 10^{-12} s, which is now over 10^{30} Planck times, a huge number and in one sense representing a very old universe, the temperature has decreased to about 10^{15} K and the electroweak interaction separates into the weak nuclear force and the even weaker electromagnetic force, carried by photons. The weak nuclear force is the one that controls radioactive β decay and also controls the fusion of hydrogen in the core of the Sun. It is therefore directly responsible for the energy we receive from the Sun. The fact that the force is weak and the fusion process goes comparatively slowly gives stars like the Sun a very long life, long enough for human life to evolve on the Earth. That is a digression but it helps to show that the physics concerned with the early universe is not some esoteric physics but is of great relevance to our existence. Back to our Big Bang scenario. All the time the Universe is expanding and cooling. At around 10^{-6} s and temperatures of 10^{13} K the quarks and antiquarks mainly annihilate each other leaving a residue of quarks (which slightly outnumber antiquarks) that can now bind to form protons and neutrons.

Electron volts (eV) ∝ temperature

That the creation of neutrons and protons should happen at such a temperature makes sense since the rest mass of a proton is about $E_p = 1$ GeV or 1.6×10^{-10} J. The KE of a particle at temperature T is about kT, where k is Boltzmann's constant of 1.38×10^{-23} J K^{-1} . Hence the temperature at which kT equals the energy of a proton is about $E_p/k \approx 10^{13}$ K. In some texts you'll find temperatures always quoted in eV or in GeV. No matter whether it's eV or K, the numbers are always very big for the early Universe. The relationship between them is given by T (in eV) ≡ $(k/e) \times$ T (in K), where e is the electronic charge of 1.6×10^{-19} Coulombs and k is given above.

[If you use GeV as a measure of temperature, then you can see why the electroweak interaction fades out at a temperature of about 10^{15} K, which is about 100 GeV. This is because the rest energy of the W and Z particles is about 100 Gev, as found experimentally at CERN. Hence below 10^{15} K there isn't enough energy to create these particles and hence produce the electroweak force carrier. Put another way, today's particle accelerators, perhaps I should say yesterday's particle accelerators, can create conditions that mimic conditions in the Universe at this time, where the individual energy per particle was some 100 GeV. The LHC at CERN runs at 7 TeV, 70 times as high.]

Formation of hadrons

Hadrons, as you know from earlier in the course, is the word used to describe particles made from quarks. Hadrons include baryons and mesons. Baryons, like protons and neutrons, are made from 3 quarks. Mesons, like π-mesons and K mesons (pions and kaons), are made of a quark and an anti-quark. At about 10^{-6} s, the temperature is about 1 GeV and that is too cool for free quarks to exist. However, protons and neutrons formed from quarks are in equilibrium with each other at such high temperatures because they interact via electron neutrinos, the other particles involved in the weak interaction. There is at this stage a soup of neutrons (n), protons (p), electrons (e^-), positrons (e^+), electron neutrinos (ν_e) and anti-neutrinos ($\overline{\nu}_e$), and of course lots of radiation. The particles interact like this:

$$e^+ + n \leftrightarrow p + \overline{\nu}_e ; \qquad \nu_e + n \leftrightarrow p + e^- .$$

~ 1 s: neutrino decoupling

However as the Universe cools the timescale of these interactions gets longer and longer until they can't take place, leaving a background of neutrinos in the Universe. The neutrinos are said to **decouple** from the neutrons and protons and if we could detect neutrinos more easily than we can the model predicts that we'd find a background of cosmic neutrinos (like the cosmic microwave background) that was created at a time of about 1 s after the Big Bang. The idea behind the concept of decoupling is that the two decoupled components from then on evolve separately, whereas before they interacted, they kept each other at the same temperature, for example. After decoupling they won't do so. Neutrinos are relativistic particles whose kinetic energies much exceed their rest energies and hence their temperature and density will evolve in a different way from matter like protons and neutrons which, for most of the time of the Universe, have had much smaller kinetic energies than their rest energy of m_0c^2. Neutrinos decouple when the average energy per particle is about 3 Mev. The neutrino background has expanded and cooled and the relevant physics tells us that the average temperature of this background is less than that of the microwave photon background by a factor of 1.4, giving it a temperature now of 1.949 K (< 1 meV) compared with the microwave background of 2.725 K.

One of the consequences of the decoupling of neutrinos is that the fraction of neutrons to protons starts to decrease. Neutrons are particles with more energy than protons, about 1.29 Mev in fact. Protons are essentially stable particles. Neutrons are not and spontaneously decay into protons with a half-life when free of around 13 minutes. Hence after the decoupling there will be increasingly fewer free neutrons than protons. We'll soon see that this has an implication for what comes next.

Primordial nucleosynthesis

At about 10 s, the first light-element nuclei were formed by the binding of protons and neutrons. The next lightest element to hydrogen in the periodic table is helium. ^4He contains 2 protons and 2 neutrons. Hence the formation of heavier elements and isotopes of hydrogen will mop up the neutrons. For the first minute or so the temperature is now about 10^{10} K and cooling fast. Conditions at this stage determine the ratio of different light elements formed at the beginning of the Universe. These temperatures are not much higher than those found in the core of the Sun and just as there, deuterium and helium are the most common elements formed. Why aren't all the heavy elements formed? The answer is simple. These need helium and more massive nuclei as starting ingredients, and higher temperatures and pressures, but by the time the helium is there the temperature and pressure has fallen too much for higher elements in the periodic table to be synthesised. A smattering of ^3He, tritium (^3H), *Be* and *Li*, the next lightest elements, were formed, but that is all. Elements with 5 nuclei and 8 nuclei are unstable, so we wouldn't expect these. Tritium is radioactive with a half-life for β decay of 12.3 years so it didn't last long either.

Although stars create helium and heavier elements too, in spite of all the nuclear synthesis that has taken place in stars since the early Universe, there hasn't been nearly enough time to generate the amount of helium that we observe in space. Observation shows that the Universe now has about 24% of helium and some light elements in comparison with the amount of hydrogen. These aren't expected to be produced as free elements by stars. The Big Bang theory explains these abundances well, accounting for them by **primordial nucleosynthesis** that took place in the first 3 minutes or so of the Universe.

You can understand why there is less helium than hydrogen in the Universe. The reason is that by the time that nucleosynthesis got going, there were fewer neutrons than protons, as mentioned a few minutes ago. The ratio of neutrons to protons was about 1:7. The 2 neutrons necessary to form a ^4He nucleus were accompanied by 14 protons, 2 of which were used in the ^4He nucleus. That leaves 12 protons. Hence of the original mass of neutrons and protons, 25% becomes helium and 75 % becomes hydrogen. When the neutrons were mostly used up, little more helium could be formed. In stars there is time to create neutrons from protons via the weak interaction but the Universe was cooling too rapidly for this process to take place outside stars. By the end of some three and a half minutes, the temperature had dropped too low for any further elements to be made.

The cosmic abundances of the light elements are a huge reality check on the Big Bang theory. If the Universe had not been as our laws of physics predict because, for example, physical constants were different in the first few minutes of the Universe or there was other matter in the Universe influencing its behaviour that we don't know about or, perhaps, the density of matter was not exactly as the model predicts then the nucleosynthesis time would not have lasted exactly as long as cosmologists say it did and the relative abundance of different isotopes would have been different. The light element abundances we find today strongly support the Big Bang scenario.

After about 3 minutes the temperature has fallen below 10^{10} K, about 1 MeV, the temperature at which electrons and positrons can be produced from the radiation. Hence the positrons have been annihilated by most of the electrons, leaving electrons only, because of the slight imbalance of matter over antimatter. The universe is dominated by radiation and a plasma of hydrogen and a few other light elements. Because electrons scatter radiation very well, the Universe was opaque at this stage. The temperature was more than enough to ionise any atoms that were formed, so no atoms as we know them yet exist.

~ 1000 s

The universe is still essentially a fireball, too hot for atoms to form, but it's cooling, though not as quickly as before. I haven't emphasised the point but until now most of the energy of the Universe has been radiation. Radiation obeys a well-known law, the Stefan-Boltzmann law, that says that the energy of radiation per square metre per second at temperature T is proportional to the 4^{th} power of the temperature. In symbols the law is written $E = \sigma T^4$ for a black body. Now as the Universe expands its temperature drops in proportion to the cosmic scale factor, namely the size of the universe, i.e. $T \propto 1/a(t)$. Hence the energy density in the black body radiation decreases as $1/a^4$. The density of matter, on the other hand just decreases because the volume it occupies increases and hence the density of matter decreases as $1/a^3$. In short the radiation energy density decreases faster than the matter energy density.

Matter dominates after 50,000 years

Observation of the relative amounts of radiation and matter in the universe tells us that today matter density dominates. There must have been a time in the past when they were both about equal and this was after about 50,000 years. From then on, the matter era began, with matter as the dominant component making up the energy of the universe. At 50,000 years, the Universe was still too hot for atoms to form permanently. Any electrons that got bound to nuclei were soon ionised off. The Universe was still opaque. It is just for the same reason that the Sun is opaque. All we see looking at the Sun is the thin layer of the photosphere. We can't see into the interior of the Sun.

The atomic universe

The beginning of the Universe we recognise occurred at about 380,000 years, when the temperature had cooled to about 3000 K. Now atomic hydrogen could form with a reasonable chance that it wouldn't be ionised. Effectively the electrons condensed out onto the protons. With few free electrons left, the Universe became transparent to light and other parts of the electromagnetic spectrum. The time is known as the **recombination era** (because protons and electrons now permanently combine) or sometimes the **decoupling era**. At this stage the radiation decoupled from the matter because they no longer kept each other at the same temperature. This happened when the Universe was about 1/1000 of its present size. The radiation has continued to cool by a factor of 1000 as the Universe has expanded. This is the radiation that we see as the cosmic microwave background. It began life mainly as light and IR but is now in the microwave region of the spectrum. The physics of expansion tells us that the Planck distribution of the radiation will be preserved. Put another way, observation tells us that to very high accuracy this radiation is distributed as the Planck black-body radiation at a temperature of 2.725 K. Hence it must have been black-body radiation at the time of the decoupling of radiation and matter close to 13.8 billion years ago, only at a temperature of about 1000 times greater or close to 3000 K.

The Dark Ages

The light that filled the very early universe became redder and redder as the universe expanded and cooled. After a few million years it was no longer visible light, but infra-red radiation. Had there been eyes to see, the universe would have become dark – hot and dark at first, like the inside of a cooling furnace, gradually becoming cooler and dark. This era is known as the 'dark ages'. No stars existed yet.

One focus of attention in modern astronomy is to work out when and how the first stars formed. Did they form as swarms of smaller stars in prototype galaxies, similar to the kind we see now in all directions around us, or perhaps they formed as isolated enormous stars? Condensation of the hydrogen in the universe is the key process. I'll pick up this story later. Suffice it to say now that the dark ages lasted some few hundred million years before the first starlight flashed out into the universe. The first view we have nowadays of the universe is not the light from the first stars, that is too feeble now for our present generation of detectors, but the radiation that filled the universe when radiation and matter first decoupled, the radiation that is now the cosmic microwave background.

Cosmic microwave background

The cosmic microwave background is the earliest view we shall ever get of the Universe, unless future science allows the neutrino background to be mapped. When we look in any direction, we see what used to be a yellowish light that filled an opaque Universe at about 3000 K, not far from the temperature of an incandescent light bulb. The expansion of the Universe has red-shifted this light by a factor of just over 1000 into the microwave region.

The microwave background is a view of the early universe as it was when it was less than 1000^{th} of its present age. A thousandth of the age of someone 20 years old is about a week. At least you were recognisably human then. The universe is almost unrecognisable – no pin-points of light from stars and galaxies because no stars or galaxies existed then, just an astonishingly uniform light. This isn't theory, its observation.

If the universe were utterly and completely uniform in its infancy, then it's hard to see how any stars and galaxies could have formed. There must, surely, have been some small irregularities. There were, and their origin was mentioned earlier in this chapter. They are at about the level of 1 part in 100,000, quite a technical challenge to measure accurately. The background is indeed astonishingly uniform but the tiny irregularities are the seeds from which the large-scale structure of the universe grew. The matter in the Universe we know is clustered into stars, galaxies, gas clouds, clusters of galaxies, super-clusters of galaxies, walls and filaments of super-clusters. All this clustering has taken place since the matter and radiation began their separate development from the time they were both in equilibrium. We know this because the cosmic microwave background is essentially a picture of the radiation at the time of recombination, some 380,000 years after the Big Bang and long before the first stars were born (which, as said above, was a few hundred million years after the Big Bang). The story of the Universe as we know it with clustered matter begins here.

When we look back 13.5 billion years and see structure in the background spread over $1/100^{th}$ of a radian (0.6°) in the sky, that structure covers 135 million light years of today's universe. Although the universe has evolved tremendously since it was $1/1000^{th}$ of its age, the presence of stars and galaxies being a conspicuous new feature, the scale of large structures in the universe matches that of the microwave background. To take one example, the universe is pretty homogeneous on a scale of 500 MLY and the microwave background has little structure on the corresponding scale of 2°. How the actual large-scale structure of the universe as it is observed today (see the next section of the course for illustrations) might have grown from the tiny seeds of inhomogeneity observed in the microwave background is the subject of intense study, largely carried out by computer simulation.

If the microwave background just told us something about large-scale structure of the universe, it wouldn't get all the attention I'm going to give it or astronomers are giving it. Measurements of the cosmic microwave background are the single biggest factor shaping modern cosmology. It is a red-hot topic in modern cosmology, if I can call something that's so close to absolute zero 'red-hot'. It's on the agenda of every cosmologist not because it's only recently been discovered (it hasn't) but because today's technology can measure the fine structure of the cosmic microwave background that tells us in numerical detail a very surprising amount about the early universe. The fine structure, the inhomogeneities described above, is the fossilised record of what the early Universe was like and we can interpret the detail to deduce a whole range of cosmological parameters. I say 'we' but it's not you or I but those who have taken the trouble to get their PhDs in the field. If you'd like a more religious metaphor, then the cosmic microwave background has been described as the *fingerprint of God*, the earliest imprint we will ever see of structure in the universe.

Let me go back to the beginning. The cosmic microwave background was discovered by Arno Penzias and Robert Wilson in 1964. It's quite a funny story, for not only were they not looking for it but they didn't recognise it for what it was when they heard it in their microwave receiver. Notwithstanding that, a Nobel prize for them also fell out of the sky in 1978. Some people are lucky. I shouldn't imply that Penzias and Wilson didn't deserve their prize, for they put a tremendous amount of effort into making good observations.

Penzias and Wilson at Bell Labs were planning a program of microwave radio astronomy using an aerial and receiver system had had been built only a few years earlier in 1960 for satellite telephone relay use, telephones being Bell Labs' main business. The wanted to use the system to make quantitative measurements of the strength of extra-galactic radio sources and to do this they needed to reduce the intrinsic background hiss that every radio receiver has to its theoretical minimum value. This was technically challenging. They kept measuring more background that they expected. On one occasion, the story goes, when they climbed up to look at their aerial to check that all was well they found that pigeons had left their droppings inside it and were nesting there. They are said to have reported words to the effect that *the aerial was contaminated with a white dielectric material*. The pigeons and their droppings were removed but even that made no difference. The noise was still there, no matter how they directed their aerial. In desperation they contacted Robert Dicke at Princeton, a well-known first-class physicist who had worked in his earlier years at the microwave Radiation Laboratory at MIT. Dicke had been studying Gamow's little known work on the Big-Bang and the subsequent prediction by two of his group that a cosmic microwave background should fill the universe. In fact Dicke and his own group had been planning to search for this radiation and very quickly realised from Penzias and Wilson's information that it had already been found. Dicke's group and Penzias and Wilson published two papers side by side, presenting the evidence and giving the explanation.

The first casualty of the observation was the 'steady-state' cosmology, a theory supported by leading astrophysicists such as Fred Hoyle, Hermann Bondi and Thomas Gold. This theory acknowledged the expansion of the universe but included the concept of the continuous creation of matter out of the vacuum to occupy the extra space that expansion produced. Only a modest rate of creation is required to keep the density of matter in the Universe at a steady value. One of their motivating factors for exploring a different theory was that the oldest globular clusters of stars seemed to be older than the age of the Universe as predicted from the Hubble constant. Continuous creation was a nice idea that got round this difficulty but the experimentally observed cosmic microwave background is nowhere to be found in steady-state cosmology. The mere existence of this background was enough to kill the theory. The age discrepancy has now been resolved with today's smaller and much more accurate value of the Hubble constant and a re-dating of the old globular clusters. Although increasingly accurate measurements of the microwave background spectrum were made over the next 25 years that confirmed its isotropy and temperature, the cosmic microwave background didn't make a really big impact on cosmology until the 1990s.

Looking back in time

This graphic from the WMAP team emphasises that the view of the cosmic microwave background is the view of the universe before stars were born, back to the time when yellow-hot radiation filled the universe and universe went from opaque to transparent. We will never see further back with electromagnetic radiation, for the universe was opaque before then. WMAP is a space probe that observed the microwave radiation from a vantage point beyond the Earth. More on WMAP soon.

Distribution of cosmic microwaves

If you measure the microwave flux coming from space you find a contribution from the Sun, from a spread of sources in our galaxy and from the cosmic background. To isolate the cosmic background, the foreground sources must be removed. In addition, you find that it's brighter in one half of the sky at a level of 0.1% than in the opposite direction, because of the motion of our galaxy within our local cluster and the motion of the whole cluster at several hundred km s^{-1} relative to the distant universe. This is interesting but it can all be allowed for to find the true background. The microwave background peaks at a wavelength of 1.8 mm and a frequency of 176 GHz in the electromagnetic spectrum. The uniformity of it is strong evidence for the cosmological principle, mentioned in my first lecture. What happened in the 1990s to make the cosmic microwave background into a hot topic was that technology developed to enable this background to be measured to very high accuracy from space, without interference from earth-based sources. As I've mentioned, at a level of about 1 part in 10^5, the cosmic microwave background contains a wealth of information about the early universe. This information is in the form of small fluctuations described in terms of their equivalent temperature fluctuations, ΔT, that produce the corresponding changes in microwave flux. These changes ΔT are in micro Kelvin (μK).

Space probe measurements

The first mission to report an all-sky survey of the cosmic microwave background fluctuations was the COBE (Cosmic Background Explorer) a satellite launched in 1989 and that reported results that hit the headlines in 1993. It showed that to astonishing accuracy the microwave background followed Planck's blackbody radiation curve for a temperature of 2.725 K. Also to astonishing accuracy, the microwave background is the same in whichever direction you look. By analysing the very faint structure of this radiation at a level of about 1 part in 10^5 over all directions in the sky, key cosmological constants could be measured, not merely guessed at. This was really the headline grabbing result.

2006 Nobel Prize in Physics

It not only seemed an outstanding piece of science at the time but with the passage of over a decade the achievement of the COBE team has been recognised by the award of the 2006 Nobel Prize for Physics to John Mather and George Smoot, key members of the team *for their discovery of the blackbody form and anisotropy of the cosmic microwave background radiation*. The COBE results suggested that the geometry of the Universe is flat, i.e. $\Omega_0 = 1$. The COBE picture of the detail in the distribution of the very small fluctuations in the microwave radiation was quite blurred and it was obvious that much more could be deduced about the universe as a whole if more detail were available.

→ *back*

The project that really delivered detail cosmologists of 10 years ago could only have dreamt of was the WMAP probe (**W**ilkinson **M**icrowave **A**nisotropy **P**robe). Wilkinson was one of the co-authors with Robert Dicke of the famous 1965 paper announcing the discovery of the cosmic microwave background. He was part of the WMAP team and died as the first results were coming in. The mission was renamed to mark his lifetime involvement with the cosmic microwave background. The first WMAP results were recorded over a period of a year from a vantage point 1.5 million km away from the Earth at the second Lagrangian point, shielded from the radiation of both Earth and Sun. WMAP data released in 2003 showed structure down to a scale of 2' (2 arc minutes). The results were rightly hailed in 2003 as *the science breakthrough of the year* and the papers giving these results have since been regularly among the most cited papers in the whole of science. Updated results based on three years of observation have been released in 2006 and a further update was published in 2008.

Ripples in the sky

What does it all mean? The analysis is mathematical. The trick is a bit like analysing a musical note into its various frequency components. If I played a few notes on an unseen musical instrument, many in the audience could tell me if the instrument was a piano, a violin, a clarinet, an oboe, a trumpet and so on. Telling the instruments apart is such a common 'trick' that you hardly think anything of it. It's done by analysing the ratios of the harmonic frequencies to the fundamental note played. Our brains do this instinctively, having learned the difference from experience. It wouldn't be that difficult to program a computer to determine these ratios if the sound waveforms to be analysed were fed into the computer. The computer would need to be given the characteristic signatures of a range of instruments from which it could deduce which instrument was being played. This physics is exactly that used by music synthesisers, which perform the trick in reverse by synthesising the notes from given harmonic signatures.

The same idea is used to analyse the 3D structure of the microwave anisotropy. The data is analysed into patterns of ever finer detail that are fitted to the observations over the celestial sphere. The patterns used are called 'spherical harmonics'. Each harmonic is specified by an integer ℓ. The bigger ℓ, the finer the angular detail in the picture. In fact the variations on an angular scale of θ are represented are given by $\theta = 60°/\ell$. The harmonic spectrum of the cosmic microwave background is a measure of how much you need of each harmonic to describe the fluctuations. The WMAP team analysed values of ℓ up to 2800, corresponding to details almost as small as θ = 0.02°. In plots, the horizontal axis essentially covers angular detail in the pattern on an ever finer scale as you move right. The fraction of higher harmonics needed just fades away with increasing ℓ and for technical reasons what is plotted is $\ell(\ell+1)\times$(amplitude of the ℓth harmonic). This is often called the 'power spectrum' of the ripples. This shows the very clear structure illustrated on the slide. The result is equivalent to the music of the cosmos. What instrument is being played? We can now tell. What does this structure tell us? Quite a lot!

The location of the first peak in the spectrum is directly related to the geometry of space, which has influenced the propagation of radiation from the time the radiation was emitted to the present. If space is 'closed', the ripples would be seen on a smaller angular scale and if space were 'open', they would be seen on a larger angular scale. Within a couple of percent, the results show that space is 'flat'. This is support for the inflationary version of the Big Bang. Details of the higher ripples depend on different combinations of cosmological parameters. If these ripples are well measured, then values for the different parameters can be extracted. These ripples are consistent with a simple picture of independent fluctuations that occurred in an early universe satisfying the cosmological principle. They tell us that clustering of matter that is so conspicuous in the universe now had not begun by the time the radiation and matter decoupled. This clustering might have happened if dark matter clustered before ordinary matter. In fact from an analysis of the ripples *in conjunction with other cosmologically sensitive measurements* comes the Hubble constant, the amount of dark matter in the universe, the density of neutrinos and more besides. I'll summarise the parameters of the Universe at the end of the next chapter on cosmology.

All action

Cosmologists are falling onto cosmic microwave background measurements like bears around a honey pot. Ground-based surveys, or rather balloon-based surveys, looking at part of the sky in detail have been undertaken in project BOOMERANG, based at the South Pole, and project MAXIMA. There is more to come, in that further information can be gleaned from the polarisation spectrum of the microwave background, which requires even more sensitivity to detect. Such information includes a measure of the background gravitational waves expected to fill the universe. New projects are CBI (**C**osmic **B**ackground **I**mager) based in the Atacama desert, CAPMAP (**C**osmic **A**nisotropy **P**olarisation **MAP**per), based in New Jersey, QUAD (being mounted at the South Pole, with UK input), VSA (the **V**ery **S**mall **A**rray in Tenerife), AMI (Arc-minute **M**icroKelvin **I**mager), CLOVER (another UK initiative in Antarctica), QUIET (**Q/U** Imaging **E**xperimen**T**, Q and U referring to the polarisation of the microwaves) and ESA's PLANCK mission launched in 2009 and as I update these notes has reported in 2013. PLANCK is ESA's follow-up mission to COBE and WMAP, aiming to produce and all-sky image with both greater sensitivity and finer angular detail. You can explore all these, and quite a few more initiatives, through the web. Starved of data for decades, cosmology has now become a truly observational science. This is perhaps the place to add that it's not just the cosmic microwave background that has made this possible. Cosmology is now informed by large-scale galactic positional surveys (involving hundreds of thousands of galaxies), red-shift surveys, rotational motion surveys, supernovae surveys and more besides. Four separate Nobel Prizes are mentioned in this chapter of the course. Cosmology is no longer the preserve of philosophers and theologians, a mere battleground for those with good intentions, imagination and finely honed words. Observational cosmology has now really been born. Watch it grow over the coming years.

Timeline of the Universe

The Universe has grown astronomically from its tiny beginning and has lasted an immense amount of time since the first changes took place in a fraction of a second. Why has it grown so big and lasted so long? Of course we have no other universes to compare with ours to see whether the behaviour of our universe is exceptional but there does seem to be a valid question lurking here that should be answered. Meanwhile, the timeline shown in the slide summarizes some significant events in the history of the universe. Timeline numbers in seconds look silly for the early stages of the Universe; timeline numbers in Planck units don't convey the huge stretches of time that passed during later stages of the Universe but I slightly prefer these, so the times are in Planck units, t_{pl}. The timeline is a summary of the stages outlined above. Notice that the first stars appeared around $10^{59}t_{pl}$. That's a long time from the beginning of the Universe. The Universe now is around $10^{60}t_{pl}$ old, on this scale not much older! [We don't think about it much but ridiculously short times are relevant to the chemistry of everyday life. Electrons can jump between atoms in chemical reactions and bonds re-arrange themselves on a timescale of a few tens of an attosecond (10^{-18} s). Nature really does work exceedingly fast. The are 10^{18} attoseconds in a second, a short time in our consciousness; 10^{18} seconds last 30 billion years, longer than the universe has existed. Pause for contemplation.]

It's perhaps worth reflecting that our presence as observers in the Universe already limits the type of universe that we can expect to observe. There are huge numbers of imaginable universes that could not sustain observers, perhaps because these universes are too small, too short lived, too hot and so on. To create the elements from which we are made, entire stars needed to live out their lives, taking billions of years. By the time this happened, the universe was of necessity old, vast, cold and had a very low density of matter. All life in the universe, including alien life, will find it so.

My final end-of-topic reflection is that the Universe is definitely on a one-way trip. The 'steady-state' idea is conceptually wrong, as well as physically wrong. All our experience on Earth tells us that our environment is evolving, that the Earth itself is evolving; it will never be the same in the future as it was in the past or indeed as it is now. The same is true of the Universe as a whole. It is an evolving system. Every star within it is evolving, every galaxy is evolving, the superclusters of galaxies are evolving, the structure of matter on the largest scale is evolving. Evidence of evolution is continuing to be revealed by observation. Look back at galaxies in the early Universe and they are different: smaller than many of the closer 'modern' galaxies, much fuller of young, blue, bigger stars, more likely to have active nuclei squirting out beams of radiation. Modern cosmology has that evolution built-in, because that is what observation of the Universe has told us is happening. In that way, at least, modern cosmology has got it right. Any cosmology that has built-in stasis or eternity has got it wrong.

I'll end by borrowing some words that were first said in relation to quantum mechanics. *If a thousand philosophers thought for a thousand years they would never have come up with* modern cosmology. In this case, many thousands of philosophers did indeed think about cosmology for thousands of years and they didn't come up with modern cosmology. What is different now is that nature, more than thought alone, has steered modern physicists in creating modern cosmology. Their story has been guided in every step by the fine detail of how we have found the cosmos and how we know its constituents behave. We know because of centuries of detailed observation and interpretation of experiment, though much of our knowledge has come in recent decades. Any viable cosmology of the future needs to fit all the facts.

The point of view of a sinner is that the church promises him hell in the future, but cosmology proves that a glowing hell was in the past. Yakov Zel'dovich, Russian astrophysicist, quoted by Joseph Silk.

In the beginning there was nothing, then something went horribly wrong.
Anon

JSR