THE BIG BANG - THE ORIGIN OF THE UNIVERSE

By
Colm O'Sullivan, Physics Department, University College, Cork.

1. Introduction

Galileo Galilei (1564 - 1642) was responsible for finally putting to rest the longest standing scientific controversy in history and in the process started the development of what we call science today. On and off for at least two millennia people had argued about whether the earth or the sun were at the centre of the universe. The establishment view in Galileo's time was that associated with Ptolemy, who lived in the second century AD but the theory had its roots in Greek science, which places the earth at the centre with the sun and the planets moving around it. The alternative picture, championed by Copernicus, but also traceable back to Greek philosophers (Aristarcus of Samos, for example), viewed the earth and the planets as moving in orbits around the sun. Galileo, and the other great scientist of his day, Johannes Kepler, with whom he corresponded, were persuaded by the Copernican model.

Galileo's difficulties with the Catholic Church on this issue are well known. The Church, over the years, had incorporated Ptolemaic astronomy and Aristotelian physics into its theology and world view and regarded the Copernican model as heretical. It must be said, however, that Martin Luther who was a contemporary of Copernicus also vehemently attacked Copernicus as, indeed, did John Calvin.

Galileo was a great believer in the use of observation and experiment to decide between competing theories. It seems that the story of him dropping things from the Leaning Tower of Pisa may be apocraphyl but there is no doubt he performed many careful and ingenious experiments. He was clearly up-to-date on the new technology of his time. The telescope seems to have first been used in the Netherlands (probably) in 1608. In 1609 we find Galileo using a telescope as a scientific tool - no doubt he had to design it himself and carefully oversee its fabrication in detail.

Once equipt with his state-of-the-art instrument, Galileo used it immediately to study the moon, sun, stars and, most importantly for our story, Jupiter and Venus. One can only imagine his excitement as he turned his telescope on Jupiter and clearly saw four of its moons. Here before his eyes was a Copernican system in miniature. Even more telling was his observation of the phases of Venus - unexplainable by the Ptolemy model. Galileo of course lost the battle, and his freedom of movement in the process, but won the war though he did not live to see himself vindicated. Most importantly, however, he established a new set of values for scientific methodology - particularly an insistence that only theories which are consistent with experimental observations can be considered to be scientific. One of the great advocates of this philosophy was Robert Boyle who was born and brought up in Lismore, County Waterford. Boyle received his early education in Lismore Castle before being sent to Eton he set off with his brother on 'the grand tour', as was customary for the sons of gentlemen at that time. In fact he was in Florence when Galileo died and his tutor there was a follower of Galileo. In his writings Boyle is very conscious of the importance of this experience for him and of his debt to Galileo.

I tell the story of Galileo to illustrate what I hope will appear as a theme throughout this lecture, namely the symbiotic-like relationship that exists between astronomy, on the one hand, and combination of the latest developments in 'pure science' and in technology, on the other.

2. Astronomy and New Technology in the 19th Century

Let us now jump forward almost two and a half centuries from the time of Galileo. Three examples of 'new technology' now come together to facilitate major breakthroughs in astronomy. These were (1) the technology to build large telescopes, (2) the development of spectroscopic techniques in science and (3) the invention of photography.

In fact the first really large telescope was built in the 18th century by William Herschel. With this telescope Herschel discovered over 2000 'nebulae', adding to the hundred or so of such objects previously discovered and catalogued by the French astronomer Messier. 'Nebulae' differed from 'stars' in that they were fuzzy extended objects and astronomers differed in their views on what they really were. Some argued that these were clouds of hot gas, others that they were clusters of large numbers of stars while still others considered them to be 'island universes' - a phrase coined by the philosopher Immanuel Kant - containing, perhaps, millions of stars. In the event, all were correct in that the nebulae did not turn out to be a single class of object.

In 1840, William Parsons, third Earl of Rosse, began the construction of a giant telescope at Birr in County Offaly comprising a metal mirror 2 m in diameter in a 16 m long tube. For three quarters of a century this remained the largest telescope in the world. With it Parsons was able to recognise the spiral structure of a number of nebulae of which he made excellent drawings. Unfortunately, the quality of the mirror was not quite good enough to identify any individual stars in any of the galaxies observed. Nevertheless, it was beginning to look increasingly likely that the spiral nebulae might indeed be 'island universes' but proof of this was a long time coming.

During the 19th century, also, spectroscopy became a routine tool in laboratory science. The basic principle of this technique is that every element when heated to a hot glowing gas gives off light which is characteristic of that element. By passing this light through a prism or grating these individual components can be seen as sharp spectral hues or 'lines'. Such a spectrum represents a fingerprint of the element and the amount of each element in a mixture of elements can be easily determined from the relative intensities of the spectral lines.

In the 1860s an Italian Jesuit priest Angelo Secchi attached a spectrometer to his telescope and pointed it towards the sun. What he saw were bright emission lines covered by sharper dark lines. The dark lines were understood as due to absorption by the cooler gas on the very outside of the star. All known elements were soon detected in the sun and other stars. Hydrogen was found to be far and away the most abundant element in almost all stars but the next brightest lines had never been seen in the laboratory - a new element had been discovered and was named helium. Soon afterwards helium was found on earth but both helium and hydrogen are much rarer on earth than in the universe at large.

But this is only half the story on the story of spectroscopy in astronomy. Also in the 1860s William Huggins used a spectrometer attached to his telescope on the roof of his London home. Huggins found that the spectral lines of Sirius, the brightest star in the sky, were not exactly where they should be - they were all slightly shifted towards the red end of the spectrum. Son afterwards he also found stars whose spectral lines were blue-shifted. Huggins had expected this effect as it had been predicted some twenty years earlier by a theory developed by Doppler in

Austria and Fizeau in France. Blue (or red) shifts in spectra are to be expected if a star is approaching (or receding) from the observer along the line of sight. Indeed by measuring the blue/red shift the approach/recessional speed can be determined. Huggins estimated that Sirius was receding at about 50 km/s.

The development of photography really revolutionised astronomy. Not only could stellar spectra be recorded on photographic plates and the intensities of the lines measured accurately but long exposure photographs of stars and, most importantly, nebulae meant that much fainter objects could be observed in detail.

The study and measurement of the large number of photographic plates which the new technology made possible meant that astronomy now became a labour intensive pursuit. Much of the most important work on spectra in the late 19th century was carried out at the Harvard College Observatory in the United States. The Observatory's director, Edward Pickering, decided that he needed additional manpower to attack the long and tedious measurements involved so he recruited a team of local women who were prepared to work for low pay and who, it was believed, were rather better at this type of work than men. It turned out, of course, that not only were these women better at making the careful measurements needed but they were also better and more imaginative astronomers. A number of Pickerings women became major figures in the history of astronomy in their own right. One of them, Henrietta Leavitt, plays a central role in the story that follows. She discovered that a particular kind of star, called a Cepheid variable, changes in brightness periodically in such a way that the period is directly related to its absolute brightness. The significance of this discovery was that if the apparent brightness and period of a Cepheid variable is known, then its distance away can be estimated, at least in principle.

3. Early Twentieth Century Astronomy

The two burning astronomical questions at the turn of the century were (1) how can the observed relative abundances of the elements be explained and (2) what are the distances to the spiral nebulae?

The action now switches to the western United States and, in particular, to the Mount Wilson Observatory where in the 1910s two great telescopes were built on the mountain overlooking Los Angeles. It was now possible to identify individual stars in the larger spiral nebulae which were already being called galaxies. The real breakthrough came one day in 1923 when Edwin Hubble, working at Mount Wilson, noticed what he first took to be a nova in the Andromeda nebula. Tracking back through photographic plates taken earlier, however, he realised that its brightness varied periodically - it was a Cepheid variable! It was very dim but its period indicated that it had a large intrinsic brightness from which Hubble was able to work out, using Henrietta Leavitt's discovery, that it (and hence the Andromeda nebula) was 900,000 light years away. This result astounded astronomers at the time, being much greater than anything observed before.

At last here was proof positive of the 'island universe' theory: The Andromeda nebula was a fully fledged galaxy. Most of the ordinary stars we see belong to a very similar galaxy (the Milky Way) - the two galaxies being the largest members of a small cluster of galaxies. Soon Hubble, and his assistant Milton Humasson, found Cepheid variables in other galaxies which turned out even more distant than Andromeda.

Meanwhile, since 1912, at the Lowell Observatory at Flagstaff, Arizona a skilled and patient spectroscopist called Vesto Slipher had managed to obtain the spectra of a number of these same galaxies. This was an extraordinary observational achievement; galactic spectra are necessarily very faint and complex since they originate simultaneously from millions of individual stars all very distant and each with its own blue/red shift due to its motion within the galaxy. Nevertheless when Slipher studied the spectrum of the Andromeda nebula as a whole he observed a blue shift which implied a speed towards us of around 300 km/s. He went on to study the spectra of other galaxies and found that most showed a red shift indicating recessional speeds of up to 1000 km/s.

Hubble was now able to combine his measurements of the distance of galaxies with Slipher results on their red shifts and came up with an extraordinary and, for most astronomers, an unexpected result. In the case of the majority of galaxies, the red shift, and hence the recessional speed, is in direct proportion to the distance away. After Hubble published this result in 1929 astronomy would never be the same again.

The implication of Hubble's result was revolutionary - the universe appears to be expanding! A consequence of this is that at some time in the distant past all the material in the universe was gathered together at a single point. How long ago that was depends on the value of the Hubble constant, i.e. *Ho* in Hubble's equation

recessional speed = *Ho* (distance away)

Even today the exact value of the Hubble constant (and hence the age of the universe) is a matter of dispute.

Hubble's results provided a field day for theoreticians. An obvious theory was that the universe started with a gigantic explosion (the 'Big Bang') and has been expanding ever since. Others preferred to believe that the universe always looked as it does now - it is indeed expanding but more matter is being continuously created at just the right rate (the so-called Steady State Theory).

In the 50 years that have passed each theory has had its ups and downs but today only the Big Bang model remains as a serious contender - not that it is without its difficulties as we shall see.

In the years that followed more and more galaxies were discovered and new methods were devised for measuring their distances away. Galaxies are found to come in all sorts of shapes and sizes and to group into clusters and into clusters of clusters. The development of radio astronomy, particularly after the Second World War, opened up the possibility of studying the emission from stars and galaxies of other parts of the electromagnetic spectrum; visible light is only a very small part of the total radiation emitted which includes not only radio but microwaves, ultra violet and infra-red radiation, x-rays and gamma rays. Some of the very strongest radio sources turned out to have only very faint optical counterparts but a detailed study of spectra of the latter showed that they have huge red shifts. During the early 1960s a whole range of objects which are now seen to be the nuclei of active galaxies were discovered, such as Seyfert and N-galaxies, BL Lacertae objects and, the brightest and generally most distant objects of all, quasars. Some quasars red shifts now indicate distances that could be as much as 12-18 billion light years away - we are looking at events that occurred in the first 10% of the lifetime of the universe!

4. Nuclear Physics and Big Bang Cosmology

At the same time as the great telescopes were being constructed on Mount Wilson, physics and chemistry were also developing in a revolutionary way. Partly as a result of an attempt to explain the characteristic spectra of elements, a new view of the nature of matter was emerging. Between 1910 and 1930 the idea of the nuclear atom as the basic building block of matter was being worked out. An atom comprised a very small nucleus with electrons in orbit around it; the nucleus in turn was seen to be composed of protons and neutrons. An Irishman, E. T. S. Walton, was awarded the Nobel prize for his part in the first experiment that 'split' the nucleus.

It had also been shown that when two light nuclei collided at great speed they could fuse together to form a heavier nucleus, the mass of the heavier one being a little less than the sum of the two lighter nuclei. Einstein's theory of relativity predicted that in these circumstances a significant amount of energy would be released, following the relationship

$$\text{energy released} = (\text{change in mass})c^2$$

where c is the speed of light in vacuum.

For such a fusion process to take place the particles involved need to collide at very high speeds, that is they much have very high energy. The temperature at the centre of the sun or a star, however, is high enough to give the hydrogen nuclei sufficient energy to collide and fuse - four hydrogen nuclei could fuse to form a helium nucleus with the release of a gigantic amount of energy. This in turn explains the source of the very high temperatures at the centres of stars.

During the Second World War, George Gamow at George Washington University in Washington DC set about investigating if the observed abundances of the elements could be explained by thermonuclear processes which went on at the very high temperatures in the early stages of the Big Bang. This was the first serious attempt at theoretical treatment of the Big Bang model. Gamow was disappointed to discover that only the abundances of the lighter elements (hydrogen, deuterium, helium, lithium) could be explained in this way. He also realised, however, that a great deal of radiation would be created in the process and while, in the early stages, the universe would have been opaque to this radiation, in due course as the universe expanded this radiation would be free to travel through the universe as a relic of the Big Bang. Because of the continued expansion this radiation would cool to a very low temperature and hence would have greatly increased in wavelength as time went by. Gamow estimated that the spectrum of this relic radiation today would be similar to the radiation from a body at a temperature of less than -250^{o}C.

Not too much attention was paid to Gamow's theory at the time. Other astronomers - Geoffrey and Margaret Burbidge, Fred Hoyle and William Fowler - found that the relative abundances of most of the elements could be explained as the result of thermonuclear fusion in ordinary stars. It turns out that Gamow was correct in the view that the lightest elements were formed in the primordial explosion while it is now believed the very heaviest elements were probably formed in supernovae. The alternative theory of the Steady State universe was pushed by Fred Hoyle of the University of Cambridge - a man with a real genius for publicity - throughout the 1950s and was a much more popular theory, at least with the general public if not with professional astronomers. All this changed, however, in 1965 as the result of a discovery by two scientists, Arno Penzias and Bob Wilson, working at Bell Labs in Holmdel New Jersey on the use of microwaves for communications. Microwave transmission is now a standard technique for telephone signals and for TV distribution but the technology was only at an early stage of development in the 1960s. Penzias and Wilson had a problem with their receiver: it seemed to have an excess of radio noise. They took it to pieces, replaced components, even cleaned off what they described in their later paper as "a dielectric substance caused by pigeons", but in the end came to the conclusion that the source of their problem could not be of terrestrial origin. Neither did it come from the direction of the sun or the galactic plane of the Milky Way. Indeed the radiation was exceptionally uniform in all directions.

Just up the road at Princeton University was a man who immediately understood the significance of this result. Robert Dicke did not know of Gamow's paper but he had independently done the same calculation and had concluded that the relic radiation from the Big Bang should be detectable as microwave radiation. He had predicted a temperature of around -270^{o}C for the current temperature (some 20^{o}C colder than Gamow had suggested). In fact Dicke and his collaborators were planning to build an antenna to try and detect the radiation when they heard of the Penzias and Wilson result. The two groups published the finding and the theoretical explanation simultaneously.

The discovery of the cosmic microwave background put the final nail in the coffin of the Steady State theory and from that day to this the Big Bang model has reigned supreme, even though in 1965 only the very bare outline of the theory had yet been worked out.

In 1989, NASA launched the Cosmic Background Explorer (COBE) designed to study the cosmic microwave background in great detail. The primary aims are to measure the temperature of the radiation and to determine the degree of isotropy. The former objective was easily achieved - the current value is 2.735 K not very different from Dicke's prediction - but determining the degree of anisotropy is an altogether more difficult task. Nevertheless, it is a most important result since the amount of clumping of matter in the early universe is very important in understanding how clusters of galaxies and super clusters of galaxies came into being and the amount of clumping of matter is in turn connected to anisotropy in the radiation. But the physics of what was going on in the early stages of the Big Bang is very different from that known to Gamow and Dicke.

5. Late Twentieth Century Physics

Our ideas on the constitution of atomic nuclei have undergone a paradigm shift during the last thirty years. One no longer thinks of the basic constituents as being protons and neutrons but rather of much smaller components called quarks held together by gluons. The term 'quark' has an Irish pedigree being coined by Murray Gell-Mann of Caltech from a quotation in Joyce's *Finnegan's Wake*:

> *Three quarks for Muster Mark*
> *Sure he hasn't got much of a bark*
> *And sure any he has it's all beside the mark*

Gell-Mann proposed that protons and neutrons were each composed of three quarks - it turned out there were more quarks than envisaged by Gell-Mann.

The current view is that matter is built up from two types of building block, quarks and leptons together with the 'gauge bosons' which are involved when quarks and leptons interact. There are in all six quarks and six leptons arranged in three 'families'. At low energies (i.e. low temperatures - such as that taking place in the centres of stars!) these particles interact through four distinct forces as follows, in order of decreasing strength.

1. Strong force (gluon interactions)
2. Electromagnetic force (photon interactions)
3. Weak force (weak boson interactions)
4. Gravitational force (gravitational interactions)

The current thinking however suggests that as one goes to higher and higher interaction energies these forces become progressively unified, first the Electromagnetic and weak forces become identical (the Electroweak force) and then the Strong and Electroweak force unify (Grand Unification). At the very highest energies involved immediately after the Big Bang all four forces are of equal strength. This latter suggestion is purely speculative at our present state of understanding. So, as we move back in time towards the moment of the Big Bang, our knowledge of the basic physics involved gets less and less and our ideas necessarily become more and more speculative. We can say nothing at all about what happens in the first 10^{-43}s and what little that can be said about the first microsecond remains somewhat speculative. As the universe evolved from there our knowledge of the basic physics is more well-founded but, as the system is very complicated, the calculations are not easy to carry out with any great confidence. Nevertheless the picture that emerges (outlined in the Appendix) is a long way from that imagined by Gamow, Dicke and their contemporaries.

So, the radiation detected by the COBE satellite has come to us directly from its release 300 000 years after the Big Bang, from the time that matter and radiation decoupled and from a time long before the formation of the first galaxies. Exactly how the matter we now observe evolved from the primordial hydrogen and helium is still very much an open question engaging the minds of astronomers at present. If the universe was completely isotropic at that time, however, it is hard to see how the universe as we now find it could have come about. This is why the COBE results are so important; it is too early to judge if the amount of anisotropy being reported by the COBE team is enough to explain the formation of the superclusters observed.

The amount and distribution of matter in deep space is also very important unanswered question. The problem here is that the only matter we can observe is that which emits radiation, i.e. hot glowing matter. It is assumed that there is at least as much (perhaps ten times more) invisible material - the so-called 'cold dark matter'. Many different candidates for this material have been proposed but verification is difficult. Clearly the amount. nature and distribution of such matter has to have an important bearing on any theory of the evolution of the universe.

The exact value of the Hubble constant , and hence the age of the universe, remains a matter of considerable dispute. Some calculations on the age of the Milky Way galaxy, for example, make it older than the suggested age of the universe! A very recent paper, however, gives a lower limit for the Hubble constant which is a factor of two lower than the previous lower limits; perhaps thing may be becoming clarified in this area also.

Clearly, the story of the Big Bang is not yet fully told.

The edifice described in this paper is far from complete but great progress has been made. It must surely rank among the great cultural and intellectual achievements of humankind - comparable, one would have thought, to the building of a great cathedral. It was Newton who said that he had achieved what he had because he had "stood on the shoulders of giants". Some intellectual giants have been mentioned in this story but all, including Newton, have stood on the massive shoulders of Galileo Galilei.

Appendix: a short history of the Big Bang

From 10^{-43} s to 10^{-35} s

Temperature = 10^{32} K

Strong and Electroweak forces unified

$\Downarrow$

hot soup of interacting quarks and leptons

From 10^{-35} s to 10^{-12} s

Temperature = 10^{27} K

Strong force decoupled from Electroweak force

$\Downarrow$

period of 'inflation'

$\Downarrow$

size of universe increases more than 10^{20}-fold

$\Downarrow$

quark-gluon soup + electroweak interactions

From 10^{-12} s to 1 μs

Temperature = 10^{15} K

Electromagnetic and Weak forces decouple

$\Downarrow$

quark-gluon soup
+ electromagnetic interaction of charged leptons
+ weak interaction of neutral leptons

From 1 μs to 100 μs

Quark confinement

$\Downarrow$

protons, antiprotons, neutrons, antineutrons, etc.

$\Downarrow$

$p+\bar{p} \leftrightarrow \gamma+\gamma$ and $n+\bar{n} \leftrightarrow \gamma+\gamma$
equilibrium

At t = 100 μs (Temperature < 10^{13} K):

$\gamma+\gamma \rightarrow p+\bar{p}$ and $\gamma+\gamma \rightarrow n+\bar{n}$
no longer possible

$\Downarrow$

$p,\bar{p}$ and $n,\bar{n}$ annihilation
(no $\bar{p}$ or $\bar{n}$ remaining)

From 100 μs to 1 s

$e+\bar{e} \leftrightarrow \gamma+\gamma$ equilibrium

At t = 1 s (temperature < 6 x 10^9 K):

$\gamma+\gamma \rightarrow e+\bar{e}$ no longer possible

$\Downarrow$

$e,\bar{e}$ annihilation
(no antimatter remaining)

$\Downarrow$

radiation dominated universe

From 1 s to 2 s

$p+e \leftrightarrow n+\nu$ weak interaction
$n+\bar{e} \leftrightarrow p+\bar{\nu}$ equilibrium

At t = 2 s:

universe becomes transparent to neutrinos

$\Downarrow$

cosmic neutrino background ($T \approx 2$ K)

From 1 s to 3 minutes

$p+n \rightarrow {}^2H$
$\gamma+{}^2H \rightarrow p+n$ (deuterium bottleneck)

At t = 3 minutes:

$\gamma+{}^2H \rightarrow p+n$ no longer possible

$\Downarrow$

thermonuclear production of ^2H, ^3He, ^4He, ^7Li
(Gamow theory)

From 3 minutes to 300 000 years

$e+p \rightarrow {}^1H$ atom
$\gamma+{}^1H$ atom $\rightarrow p+e$

At t = 300 000 years:

universe becomes transparent to photons

$\Downarrow$

cosmic microwave background ($T \approx 2.7$ K)

$\Downarrow$

matter and radiation decoupled

(This lecture was delivered during the 1993-1994 UCC Public Lecture Series and appeared in the book of the same lecture series ‘Town and Gown 2’, edited by Dr. Máirtín Ó Fathaigh and

Dr. William J. Reville, published by The Centre for Adult and Continuing Education Department, 1996.)