Section 1

3C36: COSMOLOGY AND EXTRAGALACTIC ASTRONOMY

Course notes, 2000[1]

Document source: Sec01.tex

This is an introductory course in the sense that it assumes only superficial background knowledge in the field, and elementary mathematical techniques. Specifically, the course does *not* require General Relativity (GR).

Recommended texts: The course is based on a variety of sources, both textbooks and the scientific press. It's noticeable that cosmology and active galaxies are dealt with quite extensively in textbooks, but 'ordinary' galaxies are surprisingly negelected.

• An Introduction to Modern Cosmology (Liddle, Wiley). The most appropriate book for the Cosmology section of the course; inexpensive (and discount copies are available from Don Davis).

• Active Galactic Nuclei (Peterson, CUP). If you intend to buy just one book, this is the one to go for, as it covers substantial parts of both the 'Cosmology' and the 'Extragalactic' portions of the course.. The AGN section of this course is largely based on Peterson's (reasonably-priced) book, which also includes a useful section on basic cosmology. Warning: (i) Peterson uses slightly different terminology in his cosmology sections than adopted for 3C36; (ii) do not confuse this book with others with very similar titles (e.g., by Robson – Robson's book is not at a level appropriate for this course).

• Principles of Physical Cosmology (Peebles, Princeton University Press). The fundamental text in the field. If you enjoy General Relativity, have a look at this weighty tome. The previous incarnation (Physical Cosmology) is much easier reading – no GR – but is out of print.

1.1 Introduction

At least when compared with traditional astronomical topics, cosmology is a relatively modern science. Some landmarks are:
• *1914:* Slipher discovered redshifts of nebulae
• *April 1920:* The 'Great Debate' between Shapley and Curtis. At that time, it was still reasonable to argue that our galaxy was the entire Universe! (We now know it is *at least* $\sim 10^6$ times bigger)

• *1923 Oct 6:* Hubble took a plate showing a Cepheid variable in M31, showing it to be an isolated, external galaxy
• *1924 December 30:* Hubble announced his result to the AAS
• *1920's:* Hubble & Humason found a relationship between velocity of recession, v, and distance, d
• *1929:* Hubble publishes what we now call 'Hubble's Law', $v = H_0 d$.
• *1952:* Baade showed our galaxy is 'typical', leading to the *Cosmological Principle:* we do not occupy a privileged or special position in the Universe.
• *1965:* Penzias & Wilson discover the Cosmic Background Radiation
• *1997:* Supernova cosmology results indicate a non-zero cosmological constant

Today, we believe that we have a reasonable understanding of the basic Big Bang model (the 'concordance cosmology'); however, it is worth recalling a quotation from Lev Landieu – "cosmologists are often in error but never in doubt"!

Section 2

The historical basis of cosmological theories

Document source: Sec02.tex

2.1 Olbers' Paradox

Many early astronomers (possibly Kepler, certainly Halley) realised that there were important consequences to the simplest observation of all: the night sky is dark. Olbers (1826) formulated the 'paradox' in the form that now bears his name.

We investigate this paradox by developing the simplest possible cosmological model,

Assumptions	(i)	Geometry – Euclidean (Olbers had no choice in this)
	(ii)	The distribution of matter in space is uniform on the large scale
	(iii)	The Universe is static
	(iv)	The Universe is time independent
Consequence	(A)	All observers see the same Universe, in all directions
	(i)	+ (A) $\Rightarrow$ (B) The Universe is boundless (infinite)
	(iv)	+ (A) $\Rightarrow$ (C) All observers see the same Universe, in all directions, at all times
		(C) $\Rightarrow$ (D) The Universe is boundless in time as well as space (infinitely old)
		(D) $\Rightarrow$ (E) Radiation is reaching us from the most distant sources

Assume a population of radiant sources (e.g., stars), each of luminosity L, number density n. Take a spherical shell of radius r, thickness dr; it contains $4\pi r^2 dr\ n$ sources, and their total observed flux per unit solid angle (i.e., their surface brightness) is

$$\frac{n 4\pi r^2 dr.L}{4\pi r^2} \equiv nLdr$$

Integrating over distance, we obtain $\int_0^\infty nLdr = nL\int_0^\infty dr = \infty$.
The sky brightness we observe is not, however, infinite! Why not? (Equivalently: why is the universe cold?) Possible explanations include:

(i) Stars have finite extension (they are not point sources) and block out stars behind them. This means that the observed intensity goes to surface brightness of stars, not ∞.

(ii) There is dust in the ISM. Could this block out radiation? The answer is no; for the dust to block out radiation it has to absorb energy, heating up dust – eventually to radiant temperature, and so dust will glow as brightly as star.

So simple galactic astrophysics doesn't help. There must be something wrong with our cosmological model; we must receive radiation only from sources at some distance $< R_{\max}$. Which assumptions were wrong?

(i) *Non-Euclidean geometry?* Curved spaces are possibly, but all sightlines still end up on the surface of a star

(ii) *Non-uniform universe?* On the large scale the distribution of matter *is* uniform, and we are not in a 'special' place

(iii) *Not static?* Observations show that, indeed, the universe is expanding; the radiation from increasingly distant sources is increasingly less, ν is less, $h\nu$ is less

(iv) *Time-dependent?* (Independent of (iii) – consider, e.g., the steady-state model) The weight of observational evidence is against (iv):
No stars older than $\sim 10^{10}$ years (or slightly more).
Evolution of sources with look-back (there are more quasars per unit volume at $z \sim 2$ than now.)

The Universe is dynamic and time dependent; a basic foundation of 'Big Bang' models. Cosmology today is concerned with the details of the Big Bang, and the Big Bang model will form a tacit underpinning of the bulk of our discussion.

2.2 Redshift and the Hubble flow

All galaxies outside the local group ($\sim$ 30+ galaxies within $\sim$ 1 Mpc) show redshifted spectral lines. The redshift, z, is *defined* as

$$1 + z \equiv \frac{\lambda}{\lambda_0} = \frac{\lambda_0 + \Delta\lambda}{\lambda_0} = 1 + \frac{\Delta\lambda}{\lambda_0}$$

where λ_0 is the emitted wavelength and λ is the observed wavelength.

If $v \ll c$ then

$$1 + z = \frac{\lambda}{\lambda_o} \simeq 1 + \frac{v}{c}$$

and the redshift is often given as a velocity, cz. If v is not $\ll c$ the relativistic form must be used:

$$1 + z = \frac{\lambda}{\lambda_o} = \frac{1 + v/c}{\sqrt{1 - v^2/c^2}}.$$

For redshifts less than a tenth or so, the redshift is often given as a velocity (in which case it is always just cz; *(Problem sheet 1 investigates this approximation.)* Larger values are given just as z (or $1 + z$).

Hubble (1929) found that $z \propto d$, the distance, for relatively nearby galaxies (i.e., $v \propto d$). We now write this dependence of velocity on distance as

$$v = H_0 d \tag{2.1}$$

where H_0 is Hubble's constant (note that this linearity breaks down at large z, unless special care is taken over the precise interpretation of 'distance' and 'velocity'). H_0 has dimensions of 1/time, but it is usually given in units of km s^{-1} Mpc^{-1} (i.e., v in km s^{-1}, d in Mpc)

Because the value of H_0 was, until fairly recently, quite poorly known (of order 100 km s^{-1}, but factor $\sim$2 uncertainty!), it has often been parameterized as

$$H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$$

Thus, for example, if a cosmological redshift of 500 km s^{-1} is measured for a galaxy, its distance may be expressed as

$$d = v/H_0 = v/100h = 5h^{-1}\text{Mpc}.$$

The last few years have seen a growing consensus that $H_0 \simeq 70 \pm 5$ km s^{-1} (i.e., $h \simeq 0.70 \pm 0.05$), through projects such as the HST Key Programme, Supernova Cosmology, WMAP etc. Because of increasing confidence that $h \neq 1$, it is becoming common to express results using other scaling factors (e.g., h_{70} for H_0 expressed in units of 70 km s^{-1}).

The expansion of the universe implies that at some time in the past everything was in the same vicinity. Suppose some galaxy, distance d, has been receding from us with its current velocity since the Big Bang (a very rough assumption). Then it has taken the age of the universe, t_0, to get to that distance, and

$$d \simeq vt_0;$$

but

$$v = dH_0$$

so

$$t_0 \simeq H_0^{-1} \equiv \tau_0^{-1}$$

where the 'Hubble time' , τ_0, is roughly the age of the Universe (more precisely, it is the time required for the Universe to double its size at the current expansion rate). We will obtain a more rigorous estimate of the age of the Universe, t_0, in Section 6 (H_0^{-1} proves to be an upper limit – i.e., $t_0 \leq \tau_0$ – if matter, of any nature, is the only important factor governing the expansion [i.e., if the cosmological constant is negligible; cf. section 6.4]).

It's of interest to note that Hubble estimated $H_0 = 530$ km s^{-1} Mpc^{-1}. This gives

$$\tau_0 = \frac{9.77 \times 10^{11}}{530}\text{yrs} \approx 2 \times 10^9\text{yr}$$

– implying that the age of the Universe is less than the age of old rocks!

2.3 The microwave background

In 1965 Penzias and Wilson accidentally discovered the microwave background (thereby winning the Nobel prize; poor old Dicke!). Subsequent observations, notably with the COBE satellite, have shown that after corrections for local motions the background is highly isotropic (to 1 part in 10^5!) on large scales; it is indistinguishable from black-body radiation (Fixsen et al. 1996 derive 2.728 ± 0.002 from COBE 4-year FIRAS results); and has it has low-level small-scale structure.

The natural interpretation of the evolution of the Universe is that it had some phase when it was extremely compact and hot; and that the '3K' background is the remnant of the explosion (Section 8).

2.4 Evolution of source counts

If a set of sources of the same luminosity L is distributed with number density n (per unit volume) then the number of sources per unit solid angle to distance r is

$$N(r) = \frac{4\pi r^3}{3} . \frac{n}{4\pi} = \frac{nr^3}{3} \tag{2.2}$$

and their brightness is greater than $\ell \propto L/r^2$ (i.e., $r^2 \propto L/\ell$).

The number brighter than ℓ is thus

$$N(\ell) \sim \frac{nL^{3/2}}{\ell^{3/2}}$$

i.e., $\log\ N(\ell) = A - 1.5\ \log\ (\ell)$ or, in magnitudes ($m \sim 2.5 \log \ell$)

$$\log N(m) = C + 0.6m \tag{2.3}$$

We have to make corrections to m as a function of z which arise simply because of the Doppler shift in spectral-energy distributions (so-called K corrections). After this, we find that observed source counts do *not* obey eqn. 2.3. Typically, for many classes of extragalactic sources, there were more in the past than now (e.g., Ryle). This constitutes an important disproof of steady-state theory – and a powerful tool for investigating source evolution.

2.5 Cosmic abundances

All stars in our galaxy appear to be formed from an initial composition which is ∼76% H, 24% He by mass. This 'primordial' composition appears to be universal. It is natural to suppose that this composition is a consequence of initial conditions in the Universe; and indeed, the H/D/He/(Li) abundances provide strong constraints on the evolution of the Universe (Section 9). (Note that heavy elements were formed subsequently, in stars – B^2FH).

2.6 Supernova Cosmology

Section 3

A brief history of the Universe

Document source: Sec07.tex

3.1 In the beginning...

We cannot locate any particle to a precision better than its Compton wavelength, h/mc, within which the particle takes on a wavelike character. If we go back far enough, the radius of the Universe becomes less than its Compton wavelength; at this epoch, we require (but don't have!) a quantum theory of gravity. We can equate this epoch the *Planck time,*

$$t_{\rm Pl} \sim \sqrt{Gh/c^5} \sim 10^{-43}\text{s} \tag{3.1}$$

We can't say *anything* sensible about earlier times, when $T > T_{\rm Pl} \sim 10^{32}$K ($z \sim 10^{32}$).

3.2 The era of elementary particles

Once temperatures are low enough that we can consider quantum gravity to be unimportant, we can – speculatively – extrapolate known physics.

The key point at this era is that conventional (matter) particles and photons become indistinguishable at high energies (as do the fundamental forces of nature). As the temperature drops, 'freeze out' occurs for a particle of mass m when

$$kT \sim mc^2. \tag{3.2}$$

At lower temperatures, on average individual photons no longer have enough energy to generate particle/antiparticle pairs spontaneously. Thermal equilibrium ensures $n_{\rm photons} \simeq n_{\rm particles}$ for particles of given mass; when $T < mc^2/k$, particle/antiparticle pairs can no longer be created, and they annihilate.

•$t \sim 10^{-35}$s: the electroweak and strong nuclear forces are unified in a 'Grand Unified' force. (A number of 'Grand Unified Theories' – GUTs – have been developed, although none is entirely satisfactory. Theories which combine electroweak, strong, and gravitational forces are 'Theories Of Everything' – TOEs.) It has been suggested that the phase transition associated with the separation of the strong and electroweak forces could leave the Universe with a colossal energy density, leading to a phase of enormously rapid expansion: 'inflation'. We will return to this in Section 11; for now, we proceed to an era when 'ordinary' physics starts to come into play.

•$t \sim 10^{-12}$s ($T \sim 10^{15}$K, $z \sim 10^{15}$): the electroweak forces decouple (the mediating W and Z bosons predicted by Weinberg and Salaam have now been confirmed at CERN). Somewhat prior to this, the observed particle/antiparticle imbalance must be generated (to give the matter-dominated Universe we observe), but at this time the Universe is a soup of quarks and leptons.

•$t \sim 10^{-6}$s ($T \sim 10^{13}$K): quarks and their antiparticles annihilate; the residue combines to form neutrons and protons.

3.3 The rest of time (1 ms → 10^{10} yr)

By $\sim 10^{-3}$s ($T \sim 10^{11}$K) the Universe is mostly photons, electrons/positrons, and neutrinos in thermal equilibrium ($h\nu \leftrightarrow e \leftrightarrow \overline{\nu}$). There are small numbers of neutrons and protons (∼1 particle in 10^9).

•$t \sim$ 1s: neutrinos decouple as the density drops; they have subsequently evolved pretty much isolated from the rest of the Universe. These primordial neutrinos permeate the present-day Universe with about the same density as microwave-background photons (roughly 10^8 m^{-3}; they are therefore about 10^9 times more numerous than atoms). However, their energies are only $\sim 10^{-3}$ eV (∼2K – 1.4× less than microwave-background photons), which are too small to excite any nuclear or atomic processes; this renders them effectively undetectable.

•$t \sim 10^2$s ($T \sim 10^9$K): electron/positron pair production ceases. The final recombination of free electrons and positrons contributes energy to the photon field *after* the neutrinos have decoupled; this is why the cosmic background photons are more energetic than the cosmic background neutrinos. A little later, nucleosynthesis occurs (Section 9).

•$t \sim 10^4$yr ($T \sim 10^5$K): the matter density finally exceeds the radiation density; the 'radiation era' ends and the 'matter era' begins.

•$t \sim 3{\times}10^5$yr ($T \sim$ 3000K): matter and radiation decouple; the microwave background we observe today forms (Section 8). The universe is neutral, and the 'dark ages' begin.

•$t \sim 10^9$yr ($T \sim$ 60K; $z \sim$ 20): galaxies start to form (the end of the 'dark ages').

•$t \sim t_0$($T \sim$ 3K; $z \sim$ 0): life evolves from the primordial slime, Jeremy Bentham dies, Ian Howarth is born (in that order)

Section 4

The Friedmann equation

Document source: Sec03.tex

Modern cosmological theory is rooted in GR (and beyond!), but many basic ideas can be presented using Newtonian mechanics. Remarkably, this classical approach is capable of producing essentially the same results!

In Newtonian gravity, the force between two objects is

$$F = Gm_1m_2/r^2$$

The acceleration is independent of mass; this means that the gravity can be expressed in terms of a potential, ϕ whose gradient gives the force: $F = -\nabla\phi$ (the minus sign arises because the force operates in the direction that reduces the potential fastest). For an isolated point mass m distance r from mass M, the gravitational potential energy is $V = -GMm/r$.

Consider an observer at point O located in a uniform medium of density ρ. A particle of mass m at distance r from the observer is attracted towards O as though there were a mass $(4\pi/3)r^3\rho$ located there. That is, the force is

$$F = \frac{GMm}{r^2} = \left(G\frac{4\pi}{3}r^3\rho m\right)\Big/ r^2$$

The particle's gravitational potential energy is

$$V = -\frac{GMm}{r} = -\frac{4\pi}{3}G\rho r^2 m$$

and the kinetic energy is just

$$T = \frac{1}{2}m\dot{r}^2.$$

Energy conservation requires that

$$U = T + V = \text{constant}$$

i.e.,

$$U = \frac{m\dot{r}^2}{2} - \frac{4\pi}{3}G\rho r^2 m$$

whence

$$\dot{r}^2 = \frac{8\pi}{3}G\rho r^2 + \frac{2U}{m}$$

We can switch coordinate systems by writing the *physical* co-ordinates, r, as

$$r(t) = x.R(t)$$

where $R(t)$ is the **scale factor** of the Universe and x represents a *comoving* co-ordinate system – i.e., one which expands (or contracts) along with the fabric of the Universe. On cosmological scales, the distance between two points is constant with time in comoving co-ordinates; the scale factor accommodates the changes in physical separation. Substituting, and showing time dependencies explicitly,

$$\left(\frac{\dot{R}(t)}{R(t)}\right)^2 = \frac{8\pi G\rho(t)}{3} - \frac{Kc^2}{R^2(t)} \tag{4.1}$$

where $K = -2U/mx^2c^2$ is a *constant* (since m, x, c, and U are all constant).

The basic relativistic form is functionally identical, but with the 'energy parameter', K, replaced by a 'curvature parameter', k.

$$\dot{R}^2(t) = \frac{8\pi G\rho R^2(t)}{3} - kc^2 \tag{4.2}$$

Equation 4.2 is our first version of the Friedmann equation.

Einstein noticed that the right-hand side of this equation is non-zero, implying non-zero $\dot{R}$ – in other words, a non-static Universe. He believed the universe *was* static, so he reviewed his equations and found that he was allowed to introduce an arbitrary additional term (essentially a constant of integration), so the final relativistic form of the equation becomes

$$\dot{R}^2(t) = \frac{8\pi G\rho R^2(t)}{3} - kc^2 + \frac{\Lambda}{3}R^2(t). \tag{4.3}$$

The 'cosmological constant' term, Λ, is becoming widely accepted; it is discussed further in Section 6.4.

The GR derivation of the Friedmann equation is normally constructed such that k takes allowed values of only 0 and ± 1, and characterizes the geometry of the universe. Note that $\rho(t)$ increases faster than $R^{-2}(t)$, so for small t, $k \approx 0$ (i.e., is negligible even if it's not zero).

Finally, we should note that cosmologists often work in units where $c \equiv 1$; the above equations will therefore sometimes be seen without the c term expressed explicitly.

The Friedmann equation, equations 4.2/4.3 describes the evolution of the Universe! Is it surprising that Newtonian mechanics gives fundamentally the same result for the Friedmann equation as GR? Not really – GR $\rightarrow$ Newtonian in the limit of low gravity, which applies globally through the Universe (though not locally!).

Section 5

Matter and radiation in the Universe

Document source: Sec04.tex

5.1 Density of the Universe – the fluid equation

The Friedmann equation describes the dynamical evolution of the universe. A complete description also requires a description of the physical characteristics of the matter/energy content. At the present epoch, the matter (energy) density exceeds the radiation (energy) density – by a factor $\sim 10^3$. However, this was not always so, and to understand the very early Universe we need to understand the role of radiation. We can discuss matter and radiation on an equal footing by using $E = mc^2$.

If we treat the Universe as a closed thermodynamic system ($dS = 0$) the first law of thermodynamics states

$$dE + pdV = 0 \tag{5.1}$$

(where E is the energy, p the pressure, V the volume). The combined (matter+radiant) energy is

$$E = (\rho_\mathrm{m} + \rho_\mathrm{r})Vc^2 = \rho_\mathrm{T}Vc^2 = \frac{4\pi}{3}R^3(t)\rho_\mathrm{T}c^2$$

Since $V \propto R^3(t)$ we differentiate eqn. 5.1 to obtain

$$\frac{d}{dt}\left(\rho_\mathrm{T}R^3\right) + \frac{p}{c^2}\frac{d}{dt}\left(R^3\right) = 0 \tag{5.2}$$

i.e.,

$$\dot{\rho}_\mathrm{T} + 3\frac{\dot{R}}{R}\left[\rho_\mathrm{T} + \frac{p}{c^2}\right] = 0 \tag{5.3}$$

This equation – the *fluid equation* – is the fundamental one relating density and pressure in a Universe containing a mixture of matter and radiation. (The full GR analysis gives the same result.)

5.2 Matter-dominated Universe

The energy density of matter exceeds that of radiation from $T \simeq 10^4$ yr onwards. It is a reasonable approximation to neglect (matter) pressure in today's matter-dominated Universe; i.e., to take $p = 0$. Then eqn. 5.3 becomes

$$\dot{\rho}_\mathrm{m} + 3\frac{\dot{R}}{R}\rho_\mathrm{m} = 0. \tag{5.4}$$

Multiplying both sides by R^3/R^3,

$$\left[\frac{R^3}{R^3}\frac{d\rho}{dt} + 3\frac{R^2}{R^3}\rho\frac{dR}{dt} \rightarrow\right] \qquad \frac{1}{R^3}\frac{d}{dt}\left(\rho_{\rm m}R^3\right) = 0$$

Multiplying both sides by R^3 and integrating gives $\rho_m R^3 =$ constant; i.e.,

$$\rho_{\rm m} \propto R^{-3} \tag{5.5}$$

This shows that the density in a fixed *comoving* volume varies with R^{-3} – not really very surprising! For the Einstein–de Sitter universe discussed in Section 7.1, we will show that $R \propto t^{2/3}$ (eqn. 7.2), so $\rho_{\rm m} \propto t^{-2}$ (and $\dot{R} \propto t^{-1/3}$).

5.3 Radiation-dominated Universe

The relationship between pressure and density of radiation is

$$P_{\rm r} = \rho_{\rm r}c^2/3$$

so in the radiation dominated limit eqn. 5.3 becomes

$$\dot{\rho}_{\rm r} + 3\frac{\dot{R}}{R}\left[\rho_{\rm r} + \frac{\rho_{\rm r}c^2}{3c^2}\right] = 0$$

$$\dot{\rho}_{\rm r} + 4\frac{\dot{R}}{R}\rho_{\rm r} = 0 \tag{5.6}$$

This is pretty much like eqn. 5.4 and can be rewritten as

$$\dot{\rho}_{\rm r}R^4 + 4\dot{R}R^3\rho_{\rm r} = 0$$

or

$$\frac{d}{dt}\left(\rho_{\rm r}R^4\right) = 0$$

Integrating gives

$$\rho_{\rm r} \propto R^{-4} \tag{5.7}$$

Unlike our previous result (eqn. 5.5), this *is* at first sight surprising – the radiation density falls off faster than the matter density!

Of the four powers of R, three are the usual volume dilution. The final power results from the 'stretching' of radiation. Since the wavelength 'stretches' $\propto R$, and $E = hc/\lambda$, the energy density drops by an extra power of R (thermodynamically, the Universe cools as it expands).

5.4 Mixtures

We have

$$\rho_{\rm m} = \rho_{\rm m,0}\left(\frac{R_0}{R}\right)^3 \qquad \rho_{\rm r} = \rho_{\rm r,0}\left(\frac{R_0}{R}\right)^4$$

(There is still only one Friedmann equation – eqn. 4.3 – but we have to write $\rho = \rho_{\rm m} + \rho_{\rm r}$; with different time dependencies for the matter and radiation densities, this gives messy solutions for $R(t)$!)

There must be some critical 'radius' (scale factor) at which $\rho_{\rm m} = \rho_{\rm r}$:

$$R_{\rm crit} = \frac{\rho_{\rm r,0}}{\rho_{\rm m,0}}R_0$$

Thus, even though the Universe is currently matter-dominated ($\rho_{\rm r}/\rho_{\rm m} \approx 10^{-3}$), at some point in the past $R(t) < R_{\rm crit}$. We conclude that the Universe began with a radiation-dominated era ($R(t) < R_{\rm crit}$) and is now (and hereafter) matter-dominated ($R(t) > R_{\rm crit}$).

Section 6

Cosmological parameters

Document source: Sec05.tex

6.1 The Hubble parameter

Since

$$v = \dot{r} = \frac{\dot{r}}{r}r = \frac{\dot{R}(t)r}{R(t)}$$

we see that Hubble's Law, $v = Hr$, means that

$$H = \frac{\dot{R}(t)}{R(t)} \tag{6.1}$$

So from Friedmann's equation (eqn. 4.3) we see that

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{k}{R^2(t)}c^2 + \frac{\Lambda}{3} \tag{6.2}$$

This equation describes the time evolution of the Hubble parameter, $H(t)$. Note that Hubble's 'constant' is the value of the Hubble parameter at $t = t_0$ (it is constant in space but *not* in time);

$$H_0 = H(t_0) = H(t = t_0)$$

where t_0 is the value of t now – i.e., the age of the Universe. (The subscript '0' will be used generally to indicate present-day values.)

6.2 The deceleration parameter, q_0

The rate of change of the rate of expansion is described by q_0. Consider a Taylor expansion of $R(t)$ about t_0 :

$$R(t) = R(t_0) + \dot{R}(t_0)(t - t_0) + \frac{1}{2}\ddot{R}(t_0)(t - t_0)^2 + \dots$$

i.e.,

$$\frac{R(t)}{R(t_0)} = 1 + \frac{\dot{R}(t_0)}{R(t_0)}(t - t_0) + \frac{\ddot{R}(t_0)}{2R(t_0)}(t - t_0)^2 \dots \quad \equiv \quad 1 + H_0(t - t_0) - \frac{q_0 H_0^2}{2}(t - t_0)^2 \dots \tag{6.3}$$

which defines the deceleration parameter algebraically as

$$q_0 = -\frac{\ddot{R}(t_0)}{R(t_0)}\frac{1}{H_0^2} = -\frac{\ddot{R}(t_0)R(t_0)}{\dot{R}^2(t_0)} \tag{6.4}$$

Generalizing,

$$q(t) = \frac{-R(t)\ddot{R}(t)}{\dot{R}^2(t)} \tag{6.5}$$

Noting that

$$\begin{aligned} \frac{dH(t)}{dt} = \frac{d}{dt}\left(\frac{\dot{R}(t)}{R(t)}\right) &= \frac{\ddot{R}(t)}{R(t)} - \frac{\dot{R}^2(t)}{R^2(t)} \\ &= -H^2(t)q(t) - H^2(t) \end{aligned} \tag{6.6}$$

then

$$1 + q(t) = \frac{-1}{H^2(t)}\frac{dH(t)}{dt} \tag{6.7}$$

That is, $q(t)$ physically describes the rate of change of $H(t)$

6.2.1 The acceleration equation

This is as good a place as any to introduce the *acceleration equation.* Differentiating the Friedmann eqtn., eqn. 4.3, and using eqn. 5.3 for $\dot{\rho}$, gives

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{3}\left[\rho_{\rm T}(t) + \frac{3p}{c^2}\right] + \frac{\Lambda}{3}. \tag{6.8}$$

We will make use of this result – the *acceleration equation* – later in the course.

Equation 6.8 shows that there are two terms contributing to the change in density:

(i) dilution in density because of increase in volume

(ii) loss of energy because pressure has done work expanding the element

6.3 The density parameter, Ω

Again, from eqn. 6.2 we have

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{R^2(t)} + \frac{\Lambda}{3}$$

For a given value of $H(t)$ there is a special, or *critical,* value of $\rho(t)$ where $k = 0$ (i.e., a geometrically flat universe), given by

$$\rho_{\rm c}(t) = \frac{3H^2(t)}{8\pi G}\left[-\frac{\Lambda}{8\pi G}\right] \tag{6.9}$$

If we consider a universe where the content is just matter, and substituting for H_0, π, and G,

$$\begin{aligned} \rho_{\rm c}(t_0) &= 1.88h^2 \times 10^{-26}\text{kg m}^{-3} \\ &= 2.78h^2 \times \frac{10^{11}M_\odot}{(\text{Mpc})^3} \end{aligned} \tag{6.10}$$

(Since $m_{\rm H} \simeq 3 \times 10^{-27}$ kg, $\rho_{\rm c}$ corresponds to a very low number density, $\sim$3 H atoms m^{-3}. On the other hand, a typical galaxy mass is $\sim 10^{11}M_\odot$ – a number we will justify later – and a typical separation is $\sim$1Mpc, so $\rho_0 \approx \rho_c$, within an order of magnitude or two.) Although we have expressed $\rho_c(t_0)$ in terms of matter density, any constituents (matter, radiation, neutrinos, 'dark energy') that add up to the same equivalent density will give a $k = 0$ universe.

The density parameter is defined as

$$\Omega(t) = \frac{\rho(t)}{\rho_c(t)} \tag{6.11}$$

with Ω_0 the present-day value and ρ to be understood to include *all* constituents (matter, 'dark energy', anything else). Substituting eqns. 6.2 and 6.9 into eqn. 6.11 gives, for matter *only*,

$$\Omega_M(t) - 1 = \frac{kc^2}{H^2(t)R^2(t)} \tag{6.12}$$

If $\Omega_0 = 1$ then $k = 0$, and hence $\Omega_M(t) = 1$ for all t (cf. Section 7.1).

From eqns. 6.9 and 6.11

$$\Omega_M(t) = \frac{8\pi G}{3H^2(t)}\rho(t) \tag{6.13}$$

but, from eqn. 6.8,

$$\ddot{R}(t) = \frac{-4\pi G}{3}\rho(t)R(t)$$

(neglecting the pressure term), giving

$$\Omega_M(t) = \frac{-2\ddot{R}(t)}{R(t)H^2(t)} \tag{6.14}$$

so, from eqn. 6.4,

$$q_0 = \Omega_0/2$$

or, more generally,

$$q(t) = \Omega_M(t)/2 \tag{6.15}$$

i.e., the deceleration and the matter density are related (big surprise!)

6.4 The cosmological constant

We have already seen that the cosmological constant appears as an extra term in the Friedmann equation

$$\left(\frac{\dot{R}(t)}{R(t)}\right)^2 = \frac{8\pi G\rho(t)}{3} - \frac{kc^2}{R^2(t)} + \frac{\Lambda}{3} = H^2(t). \tag{6.16}$$

Recall that Einstein's idea was that Λ could cancel the other terms, to give $H(t) = 0$ (i.e., a static Universe). His solution is actually unstable (so he *did* make a blunder) but there is currently renewed interest in the possibility of a non-zero cosmological constant.

The current interpretation of the cosmological constant is that it represents the energy density of the vacuum. It has two curious properties:

(i) It acts as a *repulsive* ('anti-gravity'!) term. (In Einstein's formulation, the cosmological constant exactly balanced the mutual gravitational attraction of matter in the universe.)

(ii) The energy density is *independent* of the size of the universe – doubling the volume of the universe does *not* half the energy density of the vacuum, so the product of volume times density *increases* with time (in an expanding universe).

The second point is especially remarkable. In the early days of the universe, the matter density was high but the total vacuum energy was (relatively) small. As the universe expands, the matter becomes attenuated, and the effects of gravity decrease; but the amount of vacuum energy increases. The repulsive effect of the cosmological constant steadily gains on the gravity's efforts to stop the unioverse expanding. If the cosmological constant becomes more important than gravity at any time, then it 'wins' for ever, and the universe will expand at an *accelerating* rate for the rest of time!

Just as it is convenient to parameterize ρ in terms of ρ_c, it is useful to normalize the cosmological constant by defining

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2}$$

Then the general condition for a flat universe (cf. Section 7) is

$$\Omega_{\rm M} + \Omega_\Lambda = 1.$$

(Strictly, we should include other constituents, such as radiation – $\Omega_{\rm R}$ – but these are believed ot be negligible compared to matter and 'dark energy').

6.4.1 Is there a non-zero cosmological constant?

The cosmological constant started life as a 'fudge' term, and until ca. 1997 largely remained as such. There are two recent observational approaches to measuring Ω_Λ:

- Recent observations of distant supernovae indicate that $\Omega_\Lambda \neq 0$. Although still in progress (and largely unpublished), this work has attracted considerable attention, and suggests that $\lambda \simeq 0.7$.
- Gravitational lensing (which we'll discuss later). If there *is* a non-zero cosmological constant, then the universe was bigger in the past than we'd otherwise expect. This increases the probability of gravitational lensing, so simple surveys of the numbers of gravitational lenses constrain the value of λ. (These 'simple' surveys need to carefully account for a wide variety of complex observational selection effects.) This work is also unpublished, but hints at a significantly smaller value of λ than do the SN studies.

There are alternative (and perhaps less convincing) arguments which argue for a zero-value cosmological constant:

- Quantum-field arguments allow an estimate of the approximate mass (energy) density expected for the vacuum, and hence for the size of the cosmological constant. Roughly, we expect one virtual particle per $[h/mc]^3$, where h/mc is the Compton wavelength of a particle of mass m; thus the expected mass density of the vacuum is $\rho_{\rm vac} \simeq m^4c^3/h^3$. For the largest-mass elementary particle usually considered, this density is 10^{94} kg m^{-3}(!). The observed density of the vacuum is $< 10^{-26}$ kg m^{-3} (roughly the density required to close the universe). Since the vacuum density is at least 120 orders of magnitude smaller than the naive quantum estimate, there must be an effective suppression mechanism in place; but why is it 'tuned' to 120 orders of magnitude (and not 119, or 121, [or 1, or 10000] which would produce very different.
- The SN data suggest that, at the present time, $\Omega_\Lambda \simeq \Omega_{\rm M}$ (to within a factor of $\sim$2). But at redshift 2 ($\sim$10Gyr ago), the vacuum density would have been only $\sim$10% of the total density (and undetectable), while in 10Gyr time it will be $\sim$95% of the total density (and dominant). What chance that right now we are at the epoch when the mass density and vacuum density just about balance?

The most important argument in favour of a *non*-zero cosmological constant comes from models of inflation (Section 11), which require $\Omega_0 + \Omega_\Lambda = 1$. Observationally, it seems very unlikely that the total amount of (baryonic + 'dark') matter in the universe exceeds $\Omega_0 \simeq 0.3$, requiring that $\Omega_\Lambda \simeq 0.7$. A non-zero cosmological constant also allows the universe to be older than do classical models (for the observed values of H_0 and q_0), helping address a long-standing problem in which it appeared that the oldest stars were older than the universe!

Section 7

Cosmological models and the age of the Universe

Document source: Sec06.tex

A cosmological 'model' is, essentially, a specific solution of the Friedmann equation; that is, a set of specific values for the various cosmological constants discussed in Section 6. We will explore some historically important (though probably unrealistic) models, in particular to explore the implications of different values of k.

7.1 $k = 0, \Lambda = 0$ (the Einstein-de Sitter model)

In the classical (Newtonian) model, $k = 0$ means that the total energy is zero; there is just enough to disperse matter to infinity, where it comes to rest (if $\Lambda = 0$). In terms of its geometry, the Universe is said to be *flat* (Euclidean).

7.1.1 Matter-dominated solution

From the Friedmann eqtn, eqn. 4.3, substituting for the matter-dominated solution ($\rho \propto R^3$, $\Lambda = 0$) gives

$$\dot{R}^2 = \frac{8\pi G}{3}\frac{\rho_0 R_0^3}{R^3}R^2 \quad = \text{(say)} \quad \frac{4}{9}\gamma^3 R^{-1} \tag{7.1}$$

where γ is a constant (rather odd-looking at this stage). Collecting 'R' terms, taking the square root, and integrating over t,

$$\int_0^t R^{1/2}\frac{dR}{dt}dt = \frac{2}{3}\gamma^{3/2}\int_0^t dt$$

giving

$$\frac{2}{3}R^{3/2} = \frac{2}{3}\gamma^{3/2}t$$

i.e.,

$$R(t) = \gamma t^{2/3} \tag{7.2}$$

Thus

$$\dot{R}(t) = \frac{2}{3}\gamma t^{-1/3} \tag{7.3}$$

$$\ddot{R}(t) = -\frac{2}{9}\gamma t^{-4/3} \tag{7.4}$$

Note that, since γ is positive, $R, \dot{R}$ are always positive, $\ddot{R}$ is always negative; and as $t \to \infty$, $R \to \infty$ but $\dot{R} \to 0$.

We also have

$$q(t) = -\frac{R\ddot{R}}{\dot{R}^2} = \frac{\gamma\frac{2}{9}\gamma}{\left(\frac{2}{3}\gamma\right)^2} t^{2/3\,-\,4/3\,+\,2/3} = \frac{1}{2}$$

i.e., $\Omega = 2q = 1$, for all t.

7.1.2 Radiation-dominated solution

From eqn. 4.3, with $k = 0$ and no cosmological-constant term (recall $k \approx 0$ for small t anyway),

$$\frac{\dot{R}^2}{R^2} = \frac{8\pi G\rho}{3} = \frac{8\pi G}{3}\frac{\rho_0 R_0^4}{R^4}$$

i.e.,

$$\dot{R}^2 = \frac{8\pi G\rho_0 R_0^4}{3}\frac{1}{R^2}$$

or

$$\dot{R} \propto R^{-1}$$

Integrating ($\int_0^t \frac{dR}{dt}dt \propto \int_0^t \frac{1}{R}dt$)

$$R(t) \propto t^{1/2}. \tag{7.5}$$

Combining eqns. 5.7 and 7.5 we see that

$$\rho_r \propto t^{-2}$$

(and $\dot{R} \propto t^{-1/2}$).

Thus the matter-dominated Universe expands at a different rate to the radiation-dominated Universe – but the density falls off at the *same* (t^{-2}) rate!

7.2 $k = -1, \rho = 0, \Lambda = 0$; the Milne model

This is a 'no-matter' model; $\Omega_0 = 0 = \Omega(t)$ for all t. From eqn. 4.3

$$\dot{R}^2(t) = +c^2$$

i.e.,

$$R(t) = ct \tag{7.6}$$

The Universe expands at a constant rate for ever (there is nothing to slow down the expansion!) – and 'open' Universe.

7.3 $k = -1, \rho > 0, \Lambda = 0$

At small t, $R(t)$ is small and $\rho(t)$ is large. The first term on the right-hand side of the Friedmann equation dominates, and so the initial properties are similar to the $k = 0$ case (Section 7.1), and $R(t) \propto t^{2/3}$.

At large t, $R(t)$ is large and $\rho(t)$ is small. The first term on the right-hand side of the Friedmann equation is negligible, and so the late-time properties are similar to the $\rho = 0$ case (Section 7.2); $R(t) \propto t$.

We see that $\dot{R} > 0$ for all R; R increases with t, but at a decreasing rate. The Universe is 'open' for all $k = -1$ models (i.e., $\Omega < 1$).

7.4 $k = +1, \Lambda = 0$

From eqn. 4.3,

$$\dot{R} = 0 \text{ for some } R = R_{\rm c} = \frac{8\pi G \rho_0 R_0^3}{3c^2} \tag{7.7}$$

and from eqn. 6.8 $\ddot{R}(t) < 0$ for all R. Thus the Universe reaches some maximum size $R_{\rm c}$, then contracts under its own gravity. The Universe is bound, or 'closed' (or may oscillate?!).

7.5 The age of the Universe

We know from eqn. 6.1 that

$$H(t) = \dot{R}(t)/R(t)$$

Then for the Milne (no-matter) model,

$$H(t) = c/ct \quad = \quad t^{-1}$$

and the present age of the Universe in that model is

$$t_0 = H_0^{-1} = \tau_0$$

where t_0 is the age of the Universe and τ_0 is the Hubble time, H_0^{-1}. This reproduces our earlier rough estimate (Section 2.2).

For the Einstein-de Sitter model, from eqns. 7.2 and 7.3

$$H(t) = \left(\frac{2}{3}\gamma t^{-1/3}\right) \Big/ \gamma t^{2/3} \quad = \quad \frac{2}{3} t^{-1}$$

and so

$$t_0 = \frac{2}{3} H_0^{-1} = \frac{2}{3}\tau_0$$

In general, in the absence of a Λ term, $t_0 \leq H_0^{-1}$, as in the Einstein-de Sitter model. This is simply because if there's matter in the Universe the acceleration must be decreasing, and the velocity of expansion must be decreasing, so the present rate of expansion is a lower limit to the average rate over the age of the Universe; hence H_0^{-1} must be an upper limit to the true age.

This represents something of a problem, since

$$\tau_0 = \frac{9.77 \times 10^{11}}{H_0} \text{ yr} = 9.77 \times 10^9 h^{-1} \text{ yr}$$

The best current estimates put $H_0 \simeq 70$ km s^{-1} Mpc^{-1}, whence, for an Einstein-de Sitter model,

$$t_0 \simeq \frac{2}{3}\frac{9.77 \times 10^9}{0.70} \text{ yr} \simeq 10^{10} \text{ yr}$$

which is younger than the youngest stars! (The most recent studies of globular-cluster ages, incorporating Hipparcos results, indicate that the oldest stars in the Galaxy have ages 11.5 Gyr; Chaboyer et al. link).

One option is to resort to a non-zero cosmological constant. To get $t_0 > H_0^{-1}$ requires $\Omega_{\rm M} < 0.26$, which in turn requires $\Omega_\Lambda > 0.7$

Section 8

The microwave background

Document source: Sec08.tex

8.1 Large-scale structure.

The Universe today is transparent – i.e., optically thin – and matter-dominated. A photon can travel several Hubble distances ($c/H_0 \sim 10^3$–10^4Mpc).

This could not always have been so. At previous epochs the energy density was higher (eqn. 5.7); that is, temperatures were higher. At some point the hydrogen would've ionized. We might expect this to occur roughly when

$$kT \sim 13.6\text{eV} \rightarrow T \sim 1.5 \times 10^5\text{K}.$$

However, recall that photons outnumber electrons by a factor $\sim 10^9$. (Note that this does *not* contradict the idea that the present-day Universe is matter dominated. The mass energy of one electron exceeds that of 10^9 3K photons!) Also, there is a distribution of photon energies at a given temperature, described by the Planck function. In consequence there were plenty of photons capable of ionizing hydrogen, even at temperatures well below 10^5K, and a better estimate of the temperature at which the radiation field just ionized hydrogen is

$$\sim (13.6\text{eV}/k)/\ln(10^9) \simeq 7600\text{K}.$$

More detailed calculations show that the hydrogen first recombined (combined?) at $T \sim 3000$K.

Because free electrons have much larger scattering cross-sections than electrons bound in hydrogen, the radiation and matter interacted strongly when the universe was hotter than this 'decoupling temperature'. The properties of radiation and matter were tightly coupled at this time, enforcing thermal equilibrium and giving the radiation a black-body spectrum.

When the electrons and protons combined, the Universe very rapidly became essentially transparent; radiation and matter decoupled. The energy density of radiation is

$$\rho_\mathrm{r} \propto T_\mathrm{r}^4$$

where T_r is the radiation temperature. Then from eqn. 5.7 ($\rho_\mathrm{r} \propto R^{-4}$) we have

$$T_\mathrm{r} \propto R^{-1}(t)$$

Clearly, as $R \rightarrow 0$, $T \rightarrow \infty$. Hence we did have a hot Big Bang; the early, optically thick phase prior to decoupling is often called the *fireball.*

that is, wavelengths get 'stretched' as the Universe expands. It is

8.1.1 Aside #1: on redshifts

Electromagnetic radiation leaves a point A at time t_0, and arrives at a receding point B, distance r, at time $t_0 + \delta t$. (We suppose the velocity of recession to be $v \ll c$, so $r \sim$ constant in interval δt.) Then

$$v = \dot{r} = \left(\frac{\dot{r}}{r}\right) r = \left(\frac{\dot{R}}{R}\right) r = \left(\frac{\dot{R}(t)}{R(t)}\right) c\delta t$$

If light is emitted with wavelength λ_0 at A, then at B it is seen with a redshift

$$\lambda(t_0 + \delta t) = \lambda(t_0)(1 + v/c)$$

i.e.

$$\lambda_0 + \delta\lambda = \lambda_0 \left(1 + (\frac{\dot{R}}{R})\delta t\right)$$

$$= \lambda_0 \left(1 + \frac{1}{R}\frac{dR}{dt}dt\right)$$

or

$$\frac{d\lambda}{\lambda} = \frac{dR}{R};$$

Integrating, $\ln[\lambda(t_1)/\lambda(t_2)] = \ln[R(t_1)/R(t_2)]$, or

$$\frac{\lambda(t_1)}{\lambda(t_2)} = \frac{R(t_1)}{R(t_2)};$$

that is, λ scales with R. Thus, since $z = \Delta\lambda/\lambda$,

$$1 + z = \frac{\lambda(t_0)}{\lambda(t)} = \frac{R(t_0)}{R(t)} \tag{8.1}$$

or $(1 + z) \propto R^{-1}$.

Since $\lambda \sim R$ and, by Wien's law, $\lambda_{\rm max} \sim 1/T$,

$$\frac{T(t)}{T(t_0)} = \frac{R(t_0)}{R(t)}$$

Since $T(t) \sim 3000$K and $T(t_0) \sim 3$K, $R(t) \sim 10^{-3}R(t_0)$ at the decoupling phase ($z \sim 1000, t \sim 3 \times 10^5$yr). We can never 'see' the Universe at earlier epochs, because it was optically thick previously.

8.1.2 Aside #2: on the cooling of the black-body curve.

We can write the Planck function at some radiation temperature T in terms of energy density per unit wavelength:

$$\frac{du_1}{d\lambda} = 4\pi B_\nu = \frac{8\pi hc}{\lambda^5}\left\{\exp\left(\frac{hc}{\lambda k T_1}\right) - 1\right\}^{-1}. \tag{8.2}$$

We now 'expand the Universe' by a factor R. The energy density at falls off by a factor R^4 (Section 5.3) to become

$$\frac{du_2}{d\lambda} = \frac{8\pi hc}{\lambda^5 R^4}\left\{\exp\left(\frac{hc}{\lambda k T_1}\right) - 1\right\}^{-1}.$$

We can write this in terms of the new, redshifted wavelengths (using eqn. 8.1, and $R_0 = 1$), $\lambda' = R\lambda$; i.e., $\mathrm{d}\lambda = \mathrm{d}\lambda'/R$. Substituting then gives

$$R\frac{du_2}{d\lambda'} = \frac{8\pi hc R^5}{\lambda'^5 R^4}\left\{\exp\left(\frac{hcR}{\lambda' k(T_1/R)}\right) - 1\right\}^{-1}.$$

or

$$\frac{du_2}{d\lambda'} = 4\pi B_\nu = \frac{8\pi hc}{\lambda'^5}\left\{\exp\left(\frac{hc}{\lambda' k(T_1/R)}\right) - 1\right\}^{-1}.$$

which is identical in form to eqn. 8.2. The only difference is that T_1 has been replaced by $T_2 = T_1/R$. Thus freely expanding black-body radiation is always described by the Planck function, but with $T \propto R^{-1}$.

8.2 Small-scale structure.

Although the CMB is smooth on large scales, it must carry an imprint of any small-scale irregularities that existed in the baryon-photon fluid at the time of decoupling. Anisotropies on angular scales larger than $\sim 2^\circ$ are dominated by the gravitational redshift the photons undergo as they leave the density fluctuations present at decoupling (the Sachs-Wolfe effect). Anisotropies on smaller scales are enhanced by *acoustic oscillations* of the photon-baryon fluid before decoupling. These oscillations are driven by the tendency of radiation pressure to resist gravitational compression, and their nature depends on the matter content of the Universe, and the numerical values of the cosmological parameters.

These small scale anisotropies in the CMB are characterized by the *angular power spectrum:* essentially, a plot of how much the temperature varies between points separated by angular distance θ, usually expressed in angular frequency, $\ell(= \pi/\theta$, with θ in radians; ℓ is the number of complete cycles round a great circle on the sky).

Numerical models show that the position of the peak of the power spectrum is sensitive to the total density, $\Omega_{\rm Tot}$. This is because the separation of clumps, as seen today, depends quite sensitively on the geometry of intervening space.

Results from BOOMERanG (Balloon Observations Of Millimetric Extragalactic Radiation and Geomagnetics – the geomagnetics didn't get done over the 259 hr 1998–1999 flight around Antarctica) show a peak in the power spectrum at $\ell = 197 \pm 6$, with amplitude $(68 \pm 8)\mu$K.

Section 9

The production of light elements

Document source: Sec09.tex

9.1 Baryogenesis

Before turning to the formation of the light elements, we briefly consider the origin of matter itself, and in particular the origin of baryonic matter (or, at a slightly deeper level, of quarks). We know that matter in the Universe is primarily matter: the Earth is matter, and the solar wind does not cause matter-anti-matter annihilation, so the sun is matter. The (matter) solar wind doesn't annihilate with planets (so the solar system is matter), nor the ISM (so the Galaxy is matter). High-energy cosmic rays of extragalactic origin are also matter.

Given the naive expectation that matter and antimatter should be formed (and destroyed) in equal quantities, how is it that we exist in a *matter* universe? This issue was considered by the Soviet physicist (and dissident) Andrei Sakharov. He showed that any theory for baryogenesis (and no complete theory yet exists) must obey the three 'Sakharov Conditions':

1. There must be reactions which violate baryon number conservation (the baryon number is the excess of baryons over antibaryons).

   [A characteristic of GUTs is that they are symmetric with respect to leptons and quarks; they contain leptoquark bosons X, Y, which transform quarks into leptons. Reactions involving leptoquarks *explicitly* violate baryon number conservation, since quarks have baryon number $B = +\frac{1}{3}$, but leptons have $B = 0$; e.g.,

   $$X \rightarrow \mathrm{e}^- + d \qquad \Delta B = +\frac{1}{3}$$

   $$Y \rightarrow \overline{u} + \overline{u} \qquad \Delta B = -\frac{2}{3}$$

   where u, d represent the up and down quarks. Electroweak theory also permits baryon-number nonconservation.]

2. There must be C/CP violation.

   [Why isn't violation of baryon number conservation enough? If the theory were C/CP symmetric, even baryon-violating reactions like 1 wouldn't produce an overall matter universe, as, on average, each reaction would be matched by a reaction with opposite ΔB. C/CP violation is necessary, as these operators turn particles into antiparticles. (CP violation is actually observed in kaon decays.)]

3. These processes (1) must occur out of equilibrium, because in thermal equlibrium the reverse processes occur with equal frequency.

As the universe cools, electroweak symmetry breaks in expanding 'bubbles'. Baryon-number-violating processes (rapid in the symmetric phase) are shut off, out of equilibrium, inside the bubbles. Finally, a plasma of quarks and antiquarks with CP violating interactions permeates the universe. Unfortunately, electroweak theory doesn't offer large enough effects to produce the observed baryon number – a GUTs origin may be necessary.

9.2 Primordial nucleosynthesis

As the temperature of the expanding Universe falls, a point is reached where neutrons and protons can combine into nuclei. The important points for our discussion of associated processes are:

- Protons are lighter than neutrons (938.3MeV vs. 939.6MeV; $\Delta m = 1.3$MeV)
- Free neutrons decay ($n \to p + e^-$), with a half-life of ~940s
- Bound neutrons (in nuclei) are stable

At high temperatures the numbers of neutrons and protons are practically identical (since the mass-energy difference between them is negligible compared to kT):

$$n \leftrightarrow p + e^- + \overline{\nu}_{\rm e}(+0.8\text{MeV})$$

(plus other processes). However, when $kT \sim 0.8$MeV neutrons can convert to protons, but not vice versa. Thus protons begin to outnumber neutrons, by a factor

$$\frac{n_{\rm p}}{n_{\rm n}} \sim \exp\left(\frac{\Delta m}{kT}\right) \sim \exp\left(\frac{1.3}{0.8}\right) \sim 5$$

The protons and neutrons combine to build up complex nuclei. Densities are too low at this stage for many-body collisions to be important, so nuclei build up through chains of two-body collisions, starting with $p{+}n \to^2 D{+}\gamma$ and then proceeding through a variety of routes, schematically as follows:

$$^2\text{D} + n \to^3 \text{T} + \gamma \quad ^2\text{D} + p \to^3 \text{He} + \gamma \quad ^2\text{D} +^2 \text{D} \to^3 \text{He} + n \quad ^2\text{D} +^2 \text{D} \to^4 \textbf{He} + \gamma$$
$$^3\text{He} + n \to^4 \textbf{He} + \gamma$$
$$^3\text{He} + n \to^3 \textbf{T} + p$$
$$^3\text{T} + p \to^4 \text{He} + \gamma$$

The key stage is the production (and destruction) of deuterium, which has a low binding energy (~2.2MeV), and so is destroyed at $T \geq 10^9$K. Nucleosynthesis actually occurs at around $kT \sim 0.1$MeV ($T \sim 3{\times}10^8$K), for a period around $t \sim 400$s. This timescale is long enough for the decay of free neutrons to be significant, but not complete; these decays reduce the neutron:proton number ratio by a further factor of $\sim \exp(400/940)$ to give a final neutron:proton ratio of

$$n_{\rm n}/n_{\rm p} \simeq 1/7.$$

The only elements produced in bulk are ^{1}H and ^{4}He (the only stable low-mass nuclei), with virtually all the neutrons in helium. Thus the number density of helium atoms is $n(^4\mathrm{He}) = n_\mathrm{n}/2$; each weighs 4 neutron masses (to a very good approximation) so the mass density is $2n_\mathrm{n}m_\mathrm{H}$. The total mass density is $m_\mathrm{H}(n_\mathrm{n}+n_\mathrm{p})$, so the mass fraction of helium is

$$Y_4 = \frac{2n_\mathrm{n}}{n_\mathrm{n}+n_\mathrm{p}} = \frac{2}{1+n_\mathrm{p}/n_\mathrm{n}} \simeq \frac{2}{1+7} = 0.25$$

A more detailed, numerical treatment keeps track of the whole temperature-dependent reaction network, and yields mass fractions of

^{2}H $\sim 10^{-4}$
^{3}He $\sim 2\times10^{-5}$
^{7}Li $\sim 10^{-10}$
with almost nothing else produced cosmologically.

9.3 Comparison with observations

The precise abundance of (especially) ^{4}He constrains the temperature (and its rate of change) at the epoch of nucleosynthesis, and thereby constrains the details of Big-Bang models. There are two important input parameters which affect the abundance:

- The number of neutrino species (which affects the expansion t–T relation and hence how the reactions go out of thermal equilibrium)
- The density of baryonic matter, from which the nuclei form. This is often expressed in terms of a quantity η, the ratio of baryon to photon numbers:

$$\eta = 2.76\times10^{-8}\Omega_\mathrm{B}h^2$$

  where Ω_B is the baryon density expressed as a fraction of the critical density.

The observed present-day abundance of helium, combined with standard Big-Bang models, led to a firm prediction that *only three neutrino species exist.* This prediction was subsequently confirmed experimentally by particle-physics experiments.

Overall agreement between light-element abundances and Big-Bang models is achieved for

$$0.010 \le \Omega_\mathrm{B}h^2 \le 0.022$$

Note the significance of this result: $\Omega_\mathrm{B} \ll 1$ **for any plausible value of** h**!** If $h \simeq 0.65$ then $\Omega_\mathrm{B} \lesssim 0.05$, which compares with $\Omega_0 \gtrsim 0.3$ suggested by other observations; the implication is that large quantities of non-baryonic matter must exist in the Universe. Simply in terms of mass content, neutrinos are a possible source of this matter; however, galaxy-formation models currently require 'cold' (slowly-moving) material to explain large-scale structure. At present, some sort of exotic non-baryonic mass is implied.

Section 10

Difficulties with the Big Bang

Document source: Sec10.tex

Over previous lectures we have examined various aspects of Standard Big-Bang cosmology. This model has obvious successes, as we have seen:

- It explains the Hubble flow
- It explains the microwave background
- It explains the primordial abundances of the light elements (consistently with the number of neutrino species)

It is not without its problems, however, and we will consider these problems in the following sections.

10.1 The flatness problem

Current observations suggest that $\Omega_{\rm M} \simeq 0.2$ – certainly this is true within an order of magnitude. From eqn. 6.12,

$$|\Omega(t) - 1| = \frac{kc^2}{H^2(t)R^2(t)}, \qquad (6.12)$$

and if $\Omega_0 = 1$ then $\Omega(t) = 1$ for all t. But what if $\Omega_0 < 1$, as presently appears probable (and as is certainly the case if all matter is baryonic)? We have

$$H = \frac{\dot{R}(t)}{R(t)} \qquad (6.1)$$

so

$$\begin{aligned} R^2(t)H^2(t) &= \dot{R}^2(t) \propto t^{-2/3} && \text{(matter – dominated; §5.2)} \\ R^2(t)H^2(t) &= \dot{R}^2(t) \propto t^{-1} && \text{(radiation – dominated; §5.3);} \end{aligned}$$

that is,

$$|\Omega(t) - 1| \propto t \quad \text{or} \quad t^{2/3}.$$

In other words, if $\Omega(t) \neq 1$ then it departs from 1 as an increasing function of time; if there is *any* departure from flatness, the Universe becomes more and more curved. If the Universe is *nearly* flat today, it must have been *very* nearly flat in the past:
Today ($t = t_0 \simeq 10^{17}$s) we have (very conservatively) $|\Omega - 1| < 1$; this implies that at
$t \sim 10^{12}$s (decoupling), $|\Omega - 1| < 10^{-3}$
$t \sim 10^{10}$s (matter/radiation equality), $|\Omega - 1| < 10^{-5}$

$t \sim 10^2$s (nucleosynthesis), $|\Omega - 1| < 10^{-13}$
$t \sim 10^{-12}$s (electroweak symmetry breaks), $|\Omega - 1| < 10^{-27}$
Thus at the time of nucleosynthesis (an epoch of well-understood physics) we had

$$0.9999999999999 < \Omega(100\mathrm{s}) < 1.0000000000001$$

Of all possible values of Ω $(0 - \infty!)$ this seems a very small and 'special' range. It would be nice to have an explanation!

10.2 The horizon problem

The universe has a finite age, so we can see only a finite volume (regardless of whether the universe as a whole is finite or not) – the Hubble volume. The microwave background was formed at $t \simeq 3 \times 10^5$ yr; when we look at it, we look back $\gg 99\%$ of the age of the universe. At first sight, the radiation we see from opposite sides of the sky cannot have interacted (the light-travel time is $\gg 2 \times 99\%$ the age of the universe!) – and yet the microwave background is isotropic on large scales, to $\sim$1 part in 10^5 *(COBE link).*

(In practice, the problem is even worse than this argument suggests, as different points in space would've had to interact before decoupling; when this is taken into account, the 'communication distance' is only $\sim 2°$.)

10.3 The monopole problem

Current Grand Unified Theories (GUTs) predict the existence of magnetic monopoles – very massive ($\sim 10^{16} m_{\mathrm{P}}$), stable particles. Their masses correspond to energies $\sim 10^{16}$GeV, and so they formed very early on in the evolution of the universe. Because they then diluted as R^{-3}, while everything else diluted as R^{-4} (section 5), and because they are predicted to be stable, they should dominate the universe; where are they??

Section 11

Inflation

Document source: Sec11.tex

Inflationary cosmologies were proposed in 1981, by Guth, Linde, Starobinsky, and others, as a possible solution to the problems discussed in section 10, and others. The ideas arose from investigations of Grand Unified Theories, which allow an equilibrium state at high temperatures ($\gtrsim 10^{27}$K), called the *false vacuum* (to distinguish it from the traditional 'true' vacuum; the false vacuum is an *unstable* state which eventually decays to the true vacuum). The vacuum has the curious property that its mass-energy density is constant (for the false vacuum, estimates are that $\rho_{\rm vac} \simeq 10^{70}$–$10^{80}$ gm cm^{-3}).

Consider a piston filled with false vacuum; pull out the piston to increase the volume by an amount V. The false-vacuum mass increases by $\rho_{\rm vac}V$, and its energy by $\rho_{\rm vac}Vc^2$, which must equal the work done on the piston, $-pV$. Thus $p = -\rho_{\rm vac}c^2$; the false vacuum has a *negative* pressure.

A *negative* pressure doesn't sound like a promising way to inflate the Universe, but it is not the pressure per se which is important. (Recall, in any case, that it is only pressure *differences* which have obvious effects.) Rather, under GR, not only mass, but also pressure has associated with it a gravitational field. Under normal circumstances the preessure term is negligible (air at room temperature exerts a gravitational field which is less than 10^{-11} of the field generated by the air's mass) but in the early universe pressures were so high that the associated gravitational fields were important.

A positive pressure creates an attractive gravitational field; the negative pressure of the false vacuum generates a repulsive field. From equation 6.8, when the vacuum dominates

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{3}\left[\rho_{\rm T}(t) + \frac{3p}{c^2}\right], \rightarrow -\frac{4\pi G}{3}\left[\rho_{\rm vac} - 3\rho_{\rm vac}\right]$$

whence

$$\ddot{R}(t) = 8\pi G\rho_{\rm vac}R(t)/3; \tag{11.1}$$

that is, inflationary models have a period when $\ddot{R}(t) > 0$ – the universal expansion *accelerates,* rather than decelerates. (A more general equation of state is $p = W\rho_{\rm vac}c^2$, with values other than -1 for W. This gives rise to so-called 'quintessence' models. Current evidence indicates that W is indistinguishable from the simple -1 case.)

The solution of equation 11.1 is

$$R(t) \propto \exp\left(\sqrt{\frac{8\pi G\rho_{\rm vac}}{3}}t\right), \qquad \text{i.e.,} \qquad \propto \exp(Ct), \tag{11.2}$$

so that during the inflationary phase the universe expands exponentially. For $\rho_{\rm vac} \simeq 10^{70}$–$10^{80}$ kg m^{-3} the time constant C in equation 11.2 is $\sim 10^{+34}$s so the e-folding time is $\sim 10^{-34}$s ($T \simeq 10^{28}$K, $r \lesssim 1$mm).

Finally, note that since $R(t) \propto \exp(Ct)$, then $\dot{R}(t) \propto \exp(Ct)$ also. This means that the Hubble parameter,

$$H(t) = \dot{R}(t)/R(t) \qquad (6.1)$$

is *constant* (with time as well as space) during inflation.

The energy of the false vacuum eventually decays to generate the mass and energy we observe in the universe today; in a sense, inflation therefore represents the start of the Big Bang. The decay is associated with a 'Grand Unified' phase transition that resolves the strongv and electroweak forces. We understand the Big Bang rather well, and inflation rather less well.

11.1 The flatness problem

Recall that

$$|\Omega(t) - 1| = \frac{kc^2}{H^2(t)R^2(t)} \qquad (6.12)$$

and that $H(t) = \dot{R}(t)/R(t)$, so the denominator in this equation is essentially $\dot{R}^2(t)$. The 'flatness problem' (section 10.1) was that $\dot{R}(t) \propto t^{-x}$, so $\Omega(t)$ diverges from unity. During inflation, however, we have

$$\ddot{R}(t) > 0 \equiv \frac{d}{dt}\left(\dot{R}(t)\right) > 0 \equiv \frac{d}{dt}\left(R(t)H(t)\right) > 0.$$

For the model described by equation 11.2 (exponential inflation) we have

$$|\Omega(t) - 1| \propto H^{-2}(t)R^{-2}(t), \quad \propto \dot{R}^{-2}(t), \quad \propto \exp(-2Ct)$$

which tends to zero as t tends to ∞. Inflation therefore forces Ω *towards* unity with increasing time – the reverse of 'ordinary' expansion!

Suppose inflation ends at $\sim 10^{-34}$s and is perfectly exponential; and (for the purposes of a rough calculation) that subsequently the universe has been radiation dominated, so that in the post-inflationary phase

$$|\Omega(t) - 1| \propto t.$$

As a very conservative assumption, we can safely state that Ω is greater than zero (and not more than 2) at the present day; i.e.,

$$|\Omega_0 - 1| \lesssim 1$$

at $t_0 \simeq 3 \times 10^{17}$s, implying

$$|\Omega(10^{-34}\mathrm{s}) - 1| \lesssim 10^{-51}.$$

During inflation, H is constant, and so, from equation 6.12,

$$|\Omega(t) - 1| \propto R^{-2}$$

which suggests inflation increases R by (at least) a factor $\sim 10^{25}$ – a huge amount!

Note that in solving the flatness problem, by driving $\Omega \to 1$, inflation introduced a new 'difficulty'. As we saw in section 9.3, the baryon density of the universe is only $\Omega_{\rm B} \simeq 0.02$; ~98% of mass in the universe appears to be 'missing' (or 'dark', or compensated by a cosmological constant, or 'dark energy'). Observations have resolved this issue in favour of inflation: the universe *is* dominated by dark matter and dark energy!

11.2 The horizon problem

Consider a pre-inflation region of radius $r_1 < cH^{-1}$ (where H^{-1} is the Hubble time), within which a smooth equilibrium is established. Inflation increases the size of any region of the universe while keeping the Hubble scale cH^{-1} *fixed* (because H is constant). Inflationary expansion can therefore increase the physical size of our region to radius $r_2 = r_1 R(t_2)/R(t_1)$. Because H (and c) are constant with time, the final state has $r_2 > cH^{-1}$. In other words, regions which can easily interact *before* inflation can subsequently be 'inflated' out of contact with each other.

11.3 The monopole problem

Inflation dilutes monopoles (and other fundamental particles which form before the inflationary era) like R^{-3}. Their (mass-)energy density therefore falls off rapidly compared to the *inflationary* energy density (which is constant!). They can therefore be diluted to insignificance (provided they exclusively form before inflation).

Section 12

Galaxy formation

Document source: Sec12.tex

A further consequence of inflation is that small density fluctuations get magnified to macroscopic scale. It is these fluctuations which eventually develop into galaxies.

Consider a 'wavelength' of 1Mpc today, which is a linear scale characteristic of the spatial 'cell' occupied by a typical galaxy. At the end of inflation (i.e., at the start of the 'normal' universe), time $t_{\rm f}(\sim 10^{-34}{\rm s})$, the corresponding scale would have been

$$\sim 1{\rm Mpc}\frac{t_0}{t_{\rm f}}^{-1} \sim 10^{-2}{\rm cm}.$$

This is small – but nonetheless much larger than the Hubble scale at that epoch, $cH^{-1} \simeq 10^{-24}$cm! This problem is similar to the horizon problem, and is addressed in a similar way, by inflation. Very small (quantum mechanical) fluctutations in the pre-inflationary universe could be inflated to provide the seeds of present-day structure.

12.1 Models of galaxy formation

Studies of galaxy formation form a very active research topic at present, and our understanding of the processes is changing rapidly, from both observational and theoretical viewpoints. The availability of results from the Hubble Deep Field has been a major factor in motivating these studies.

The formation of the first gravitationally bound structures is limited by radiation drag. This means that mass density fluctuations don't become free to grow until decoupling of matter and radiation (section 8). For galaxies to have formed by $z \simeq 10$ from a baryonic-matter universe, we require perturbations at the epoch of recombination of amplitude

$$\frac{\Delta\rho}{\rho} \simeq 10^{-4}$$

(the so-called 'Harrison-Z'eldovich spectrum), or $\Delta T/T \lesssim 10^{-4}$, on angular scales $\sim$0.5′–20′ to correspond to present-day galaxies and clusters. The subsequent detailed evolution of structure depends on the cosmological density parameters, and on the matter content of the universe.

Observations yield $\Delta T/T \lesssim 10^{-5}$ on arc-minute scales. (This result was initially obtained as a statistically significant excess noise in COBE results; direct 'imaging' is now available from several ground-based experiments, such as at Tenerife.) This means that either galaxies don't exist (observations suggest otherwise!), standard big-bang models are completely wrong (a conclusion we are reluctant to accept without further evidence), or that *galaxy formation is influenced by dark matter.*

We have seen that inflation strongly favours $\Omega_{\rm Tot} = 1$, while nucleosynthesis indicates $\Omega_{\rm B} < 0.1$. This implies considerable amounts of non-baryonic 'dark matter' – of unknown nature – which influences galaxy formation.

This 'cosmological' dark matter must have the following properties:

(i) it must interact only weakly with baryonic matter (to have avoided immediate detection);

(ii) it must not be subject to pressure from photons (otherwise we'd see it!)

(iii) it must have significant *gravitational* effects

If non-baryonic dark matter decoupled from radiation before $t_{\rm rec}$, the epoch of recombination, it could provied the gravitational wells necessary to 'seed' galaxy formation. Conventionally, dark matter is classified as 'hot' or 'cold' (depending on whether it moves relativistically or not), giving rise to rather different predictions for galaxy structure. The associated density fluctuations are categorized as:

i Adiabatic – dominant mass components (e.g., photons and baryons) have the same primordial spatial distribution;

ii Isothermal – matter is perturbed but radiation is not

iii Isentropic – matter and radiation vary with opposite phase (i.e., matter-dense regions have rarified radiation)

iv Turbulent – large-scale eddies in coupled matter and radiation.

12.1.1 CDM universes

Over the last several years, an adiabatic CDM universe has emerged as a sort of 'benchmark' model. The CDM condenses (before the baryonic matter, because it is independent of the radiation field), and sits there to form galaxy haloes. Baryonic matter settles in the potential wells to make the visible parts of galaxies. Galaxies then aggregate into clusters – i.e., small(ish) units form first, in a so-called 'bottom-up' scenario. (These models are sometimes called 'dissipationless'.)

12.1.2 HDM universe

'Hot' dark matter has $v \sim c$ at decoupling. A good case can be made for considering neutrinos as candidates for HDM, not least because they are one of the few forms of non-baryonic matter known to exist and to have, potentially, cosmologicaly significant mass!

(Most astronomers have considerable faith in the standard model of the Sun, which predicts many more neutrinos than observed. A good explanation of this 'neutrino problem' is that neutrinos change type, or 'oscillate', which requires that they have mass. Astronomers have known this for about 20 years. In 1998, particle physicists found the same result from experiments at SuperKamiokande, and made a lot of fuss about their 'discovery' that neutrinos have mass. There is still no measurement of what that mass is, and therefore no indication of whether or not neutrinos are likely to be an important contributor to 'dark matter'.)

In HDM models, fast-moving particles *smooth out* initial irregularities, on the Hubble scale. The mass distribution therefore has a large coherence length, and the first nonlinear structures to form are large 'Zel'dovich pancakes'. The characteristic scale corresponds to galaxy clusters, which would have then to fragment into galaxies – a 'top down' scenario.

12.1.3 Observational methods

How do we distinguish between these models? An important difference between them is in their predictions of the space distribution and clustering properties of galaxies. The space distribution is often characterised in terms of the *two-point correlation function.* If we write

$$\delta(x,t) = \frac{\rho(x,t) - \overline{\rho}(t)}{\overline{\rho}(t)}$$

then the (two-point) correlation function is

$$\xi(r) = \langle \delta(x)\delta(x') \rangle_r$$

where $\langle\rangle_r$ means the average over all points separated by a distance r. This quantity tells us how likely we are to find regions of under- (or over-) density next to each other.

Much effort in current cosmological research is directed towards empirically determining the two-point correlation function (and other descriptors which quantify and summarize the spatial distribution of galaxies). Numerically, observations suggest that the two-point correlation function for galaxies is well approximated by

$$\xi(r) = (r_0/r)^\gamma,$$

with $r_0 \simeq 5.4h^{-1}$ Mpc, $\gamma \simeq 1.8$ for 10kpc$\lesssim hr \lesssim$10Mpc.

Simulations of the observations, running large-scale N-body simulations for different cosmological models to determine which best describes the observations, support a ‘bottom-up’, or CDM-dominated model. (Very recent work suggests a significant fraction of HDM in a ‘mixed dark matter’ universe.) One simple argument along these lines is that the crossing time of the Local Group, ℓ/v, is of the order of the Hubble time, suggesting that it is still forming – but the galaxy is old (it has old stars). Although we see ‘pancakes’, they appear still to be forming. Finally, large structures should impose large CBR fluctuations – which are not obsered.