Chapter 1 - The Cosmological Critical Density

Last update: 19 June 2006

1. Introduction

Inevitably the story starts in the middle. It is not yet possible to start the story of the universe at the beginning, though it may be one day. Many of the world's most able physicists toil to extend physical theory to apply even at the moment of creation. To disguise their hubris, this endeavour is given the code name "quantum gravity". Our purpose is far more modest.

We shall pick up the story about one millisecond or so after the big bang. What exactly *was* the big bang? A description that will suffice for now is that the big bang was the origin of the universe in an extremely hot, dense state and in which the constituents of the universe were flying away from each other at very high speeds. The analogy with an explosion is helpful in some ways, but misleading in other respects (see the accompanying FAQ). Exactly what the constituents of the universe were at very early epochs will be discussed in later Chapters. Eventually – which is to say after a few hundred million years – the contents of the universe settled down sufficiently to form the first stars and galaxies. These initial stars and galaxies were probably quite different from the stars and galaxies which formed later in the life of the universe. However, even after many generations of stars, and the evolution of spiral galaxies, the universe on a sufficiently large scale continues the expansion initiated at the big bang. Will it expand forever?

In this initial Chapter we make the simplifying assumption that the cosmological constant is zero – which is to say, we shall ignore dark energy. In this case we shall see that the universe would expand forever if its mean density does not exceed a certain critical density. The existence of a critical density with this interpretation is sometimes presented as if it were a difficult outcome of general relativity theory. This is not so. We derive the critical density in this Chapter from very simple classical energy arguments. Experts will spot many shortcomings in this simplistic development. But I'm unrepentant[1].

2. The Universe's Total Energy

Envisage a region, R, large compared with the typical spacing between galactic clusters, i.e. sufficiently large that the mean density of matter enclosed (ρ) is representative of the universe as a whole. In addition to the rest-mass density (of both ordinary matter and any 'dark matter'), this ρ also includes the mass equivalent of the 'positive energies' (i.e. these energies divided by c^2).

The 'positive energies', E_+, are the radiation field energies (photons, neutrinos, etc) plus the kinetic energy of the massive particles. What is excluded from E_+ is the (negative) gravitational potential energy. E_+ also excludes the dark energy despite the fact that dark energy makes a positive contribution to the overall density of the universe (indeed, more than two-thirds of it at the present epoch, if current estimates are to be believed). For more on dark energy and dark matter see the accompanying FAQ.

[1] Be warned, though. It appears that our universe does contain dark energy, and lots of it. It would appear that this dark energy will cause the universe to expand forever despite the mean matter density being substantially less than the critical value. This is explained in detailed in Chapter 5B.

Hence, for the time being we shall take the galactic masses plus the positive energies as constituting our total mean density, ρ. The reason for including the positive energies in ρ is that they also contribute to causing gravitation. (This is one place where we *have* sneaked general relativity in). Thus, within a sphere of radius R, the gravitating mass is $M = (4\pi/3)R^3\rho$. The gravitational potential energy associated with a galaxy of mass m situated at the edge of the sphere of radius R is thus,

$$P.E. = -\frac{GMm}{R} = -G\frac{4\pi}{3}mR^2\rho \qquad (1.2.1)$$

Note that the exclusion of dark energy from our considerations at present is due to its peculiar gravitational characteristics. Whilst dark energy has a positive density, which contributes to attractive gravity in the usual way, it also has a negative pressure. In general relativity, pressure acts as a source of gravitation as well as mass. In particular, negative pressure causes repulsive gravitational forces. It is this property which leads to the hypothesis that dark energy exits. It is to provide an explanation for the observed accelerating expansion rate of the universe. Thus the essential difference between dark energy and the other constituents of the universe is not its 'darkness', nor its contribution to the overall density, but the fact that its pressure provides a significant – indeed apparently dominant – contribution to gravitation.

We shall now assume a Hubble relation between distance and velocity. This relation proposes that the relative recessional speed of two galactic clusters is proportional to their distance apart. At any epoch there is thus a 'constant' of proportionality, H, which relates the speed, v, to the distance, R,

$$v(t) = H(t)R(t) \qquad (1.2.2)$$

This relationship was suggested by observational evidence gathered by Hubble and his co-workers in the 1920's. Clearly, observational evidence is available only from the earth as an observational platform. (From the cosmic perspective, telescopes in earth orbit are not significantly different in terms of location). However, the Copernican principle, namely that the earth does not occupy a privileged position in the cosmos, suggests that (1.2.2) should be true with respect to any chosen origin. This leads to the interpretation of the 'Hubble law' as implying an expanding universe. The expansion in question is uniform and isotropic, the same sort of expansion that occurs when the temperature of a solid is increased. Note that 'curved space' is not necessary to understand this. For a little more about the history of the concept of an expanding universe see the accompanying FAQ.

Whilst H(t) is supposed to be the same throughout the universe at any given epoch, the notation emphasises that the Hubble 'constant' is not constant in time. For this reason we shall refer to H as the 'Hubble parameter'. Note that v is the velocity of a galaxy cluster at the edge of a sphere of radius R with respect to the centre of that sphere.

Using (1.2.1) and (1.2.2), the total energy of our chosen galaxy is thus,

$$E = K.E. + P.E. = \frac{mv^2}{2} + P.E. = mR^2\left[\frac{1}{2}H^2 - \frac{4\pi}{3}G\rho\right] \qquad (1.2.3)$$

Note that although we have called this the total energy, it is only the sum of the classical kinetic energy and the potential energy. There is no rest-mass energy included in (1.2.3). Hence it does not relate to the universe's density if divided by c^2. Nevertheless we shall see in Chapter 5B that a relativistic treatment leads to this same expression (in the absence of a cosmological constant). In the relativistic treatment, this expression is also a constant of the motion but, rather than being interpreted as energy, it is the curvature of space.

Also note that (1.2.3) is the energy of a galaxy with respect to some arbitrary chunk of the universe of radius R at some arbitrary reference time. If there are N galaxies in this chunk, then N times (1.2.3) will be the total energy of this 'R-chunk'. Since the universe may be of infinite size, we cannot assume that R is the size of the universe.

Equation (1.2.3) assumes that the Hubble velocity, v = HR, is the only significant motion of the typical galaxy. In addition to the tendency to move with the 'Hubble flow', stars and galaxies will have their own local motions superimposed. Thus, these additional local movements are assumed to be slow compared with the Hubble speed. Note that this is the opposite of what occurs in the kinetic behaviour of gases. For gases, the velocities associated with bulk gas motion are usually small compared with the random thermal velocities of the individual molecules.

[Aside: Dominance of the Hubble speed over the local speed is inevitable so long as a large enough distance, R, is considered. Consider these local speeds: the earth's speed in its orbit around the sun is 30 km/s; the sun's speed in its orbit around the milky way is about 250 km/s; the milky way's speed with respect to the centre of the local group of galaxies is about 300 km/s. Suppose these conspire to give a local speed of nearly 600 km/s. Since the Hubble parameter is 72 km/s/Mps = 22 km/s/Mlyr[2], it follows that the Hubble 'flow' speed will be comparable, i.e. 600 km/s, at a distance of 600 / 22 = 27 Mlyr. The typical diameter of a spiral galaxy is 100 – 150 klyr. The typical size of a cluster of galaxies is ~5 Mlyr, whilst clusters of clusters of galaxies are ~20 Mlyr in size. Finally, the thread-like filaments which comprise the largest non-uniform structure in the universe, the super-clusters, are several hundreds of Mlyr in length. Hence, the size scale on which the universe can be regarded as uniform is of the order of hundreds of Mlyr. Thus, if R is sufficiently large for the universe to be approximately homogeneous, it will be correct to assume the Hubble motion is dominant].

4. Dependence of the Universe's Expansion on Its Density

We now proceed on the basis of the energy given by Equ.(1.2.3) is constant as the universe expands, as a good energy should be. Firstly we note that the net energy, E, can be negative if the gravitational potential energy exceeds the kinetic energy in magnitude. If E is negative then the 'R-chunk' must have a maximum possible size. This is because, if we allowed R to tend to infinity then the density would decrease to zero, leaving only the first (kinetic energy) term in (1.2.3), which cannot be negative – a contradiction. In fact, from (1.2.2), it must be that H = 0 at the maximum size, since

[2] A parsec (ps) is 3.26 light-years (lyr). A mega-light-year (Mlyr) is 9.46 x 10^{18} km.

the velocity must then be zero (as the galaxies reverse direction and the universe begins to contract). Hence, for negative E, the maximum size of the 'R-chunk' follows immediately from (1.2.3), i.e.,

$$R_{MAX} = \frac{GM^2}{-E} \qquad (1.4.1)$$

where M is the total mass of the R-chunk.

Thus, as the kinetic and gravitational energies in (1.2.3) become closer to cancellation, i.e. as E becomes smaller, the R-chunk is destined to reach an ever greater maximum size. As the cancellation becomes exact, and the net energy of the R-chunk becomes zero, its radius can just expand to infinity (with velocities asymptotic to zero). From (1.2.3) the critical density at which this condition occurs is,

$$\rho_{\text{critical}}(t) = \frac{3}{8\pi G} H(t)^2 \qquad (1.4.2)$$

Note that the critical density varies with the age, t, of the universe, as does the Hubble parameter. Note also that if gravity were weaker (i.e. if G were smaller) then the critical density would be larger. This is not surprising, since, if gravity were weaker, it would obviously require a greater density to reverse the universe's expansion.

The value of H at the present time is 74 km/sec/Mparsec = 22 km/sec /10^6 light years = 2.3 x 10^{-18} /sec. This value for H derives from the WMAP/COBE satellite observations of the microwave background and is believed accurate to +/-4%. Since G = 6.67 x 10^{-11} $m^3 kg^{-1} sec^{-1}$ we get a critical density of

$$\rho_{\text{critical}} = 9.7 \times 10^{-27}\, kg/m^3 = 9.7 \times 10^{-30}\, g/cm^3 \quad \pm 14\% \qquad (1.4.3)$$

One hydrogen atom weighs 1.67 x 10^{-24} g. Hence the critical density corresponds to just one nuclear particle per volume 1.67 x 10^{-24}g / 9.7 x 10^{-30}g/cm^3 = 1.7 x 10^5 cm^3 = 0.17 m^3, or just 5.8 hydrogen atoms per cubic metre.

In sharp contrast, the best achievable vacuum in terrestrial physics laboratories is around 10^{14} molecules per cubic metre. Hence, an average cosmic density of material which would constitute a record breaking vacuum by some 13 orders of magnitude on earth, would be sufficient in its gravitational effects to eventually halt the expansion of the universe and cause it to contract.

5. Conclusion

We have discovered four remarkable things in this opening Chapter. Firstly, if the average cosmic density exceeds a certain critical density, the universe will stop expanding and contract instead. The situation is analogous to a projectile reaching the zenith of its trajectory. Conversely, a universe with an average density less than the critical density is analogous to a projectile whose velocity exceeds the escape velocity. In respect of the first conclusion, it is particularly important to recall that dark energy has been neglected in this Chapter.

Secondly, we have derived a very simple expression for the critical density, equation (1.4.2), in terms of the universal gravitational constant, G, and the Hubble parameter alone. We will see in Chapter 2 that the Hubble parameter is roughly the reciprocal of the age of the universe. Consequently, we have discovered in Equ.(1.4.2) the critical density's dependence upon time, apart from which it depends only on G.

Thirdly, in the case of a universe for which the gravitational potential energy is greater in magnitude than the kinetic energy, such a universe can only expand so far – and thereafter re-contracts[3].

Finally, using a measured value for the Hubble parameter (equivalent to a universe of age 13.7 billion years), the critical density is found to be remarkably small, namely around 6 hydrogen atoms per cubic metre.

All these conclusions follow from the simple energy equation (1.2.3). It almost seems to be cheating to conclude so much from so little. But we have not yet exhausted the possibilities inherent in equation (1.2.3), as we shall see in the next Chapter.

[3] Note that our simple treatment in this Chapter does not permit us to conclude that a universe which re-contracts is necessarily finite in spatial extent. This is true (for zero cosmological constant) but it requires a full relativistic derivation to prove it. This is provided in Chapter 5B.