15
Problems with the Big Bang

The hot big bang cosmology that we have discussed so far has been a very successful theory. It describes cosmic evolution starting from a fraction of a second after the big bang, accurately predicts the primordial nuclear abundances and the properties of the microwave background radiation, and explains how galaxies and clusters were formed over billions of years. And yet this theory fails to address some puzzling questions about our universe. Why is the geometry of the universe so close to being flat? Why is the universe so homogeneous on large scales? What is the origin of the small density fluctuations that seeded structure formation? And why is the universe expanding?

These questions do not have answers within the big bang cosmology. It simply postulates that the universe started out in a state of homogeneous expansion and was nearly flat from the start. But it is very hard to understand how such an initial state could arise, as we shall now discuss.

15.1 The Flatness Problem: Why is the Geometry of the Universe Flat?

The universe we observe today is close to having a flat, Euclidean geometry. This is equivalent to the statement that today the average energy density is nearly equal to the critical density, or that the present density parameter $\Omega_0 = \frac{\rho_0}{\rho_{c,0}}$ is close to unity. Observations indicate that Ω_0 deviates from one

D. Perlov and A. Vilenkin, *Cosmology for the Curious*,
DOI 10.1007/978-3-319-57040-2_15

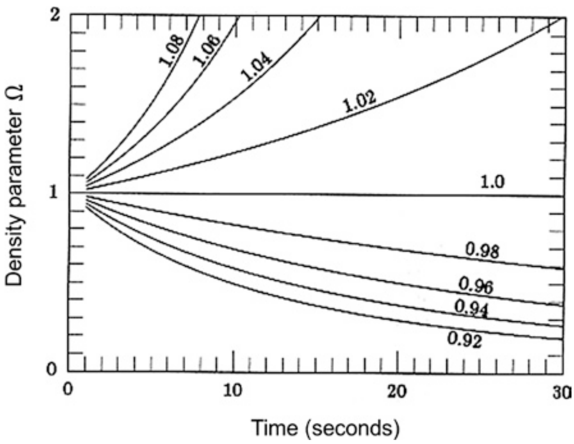

Fig. 15.1 Each line shows the evolution of the density parameter starting from the indicated initial value. In all cases, the evolution begins at one second after the big bang. If Ω was ever slightly greater than 1 in the past, then it will grow toward infinity. If Ω was ever slightly less than 1 in the past, then it declines towards zero. Thus, in order for Ω to be close to unity today it must have started out *extremely* close to unity in the past *Credit* Alan Guth

by no more than 1% (see Sect. 9.2). Trying to understand this major feature of our universe reveals a mystery, which is known as the flatness problem.

If the universe starts out with $\Omega = 1$, it will remain this way indefinitely. However, any slight initial deviation from unity will be amplified with time causing Ω to grow unchecked or plummet to zero[1] (see Fig. 15.1). In other words, $\Omega = 1$ is a point of unstable equilibrium. If, for example, we had $\Omega = 1.01$ at the onset of nucleosynthesis ($t = 1$ s ABB), then in less than a minute we would have $\Omega = 2$ and in a little over three minutes the universe would collapse to a big crunch. Similarly, if we started with $\Omega = 0.99$ at $t = 1$ s, then in about a year the density would be 300,000 times smaller than critical ($\Omega = 0.000003$). No galaxies or stars would ever be formed in such a low-density universe. In order for Ω to have its observed value at present, its value at $t = 1$ s must be set equal to one with an accuracy of one part in 10^{16} (see the Appendix).

Thus the flatness problem is the realization that the universe must be launched with an Ω that is finely tuned to unity, even though the big bang model cannot explain why this should be the case. It simply must be assumed as an initial condition.

[1]This is only true if the universe expands with deceleration, which is indeed the case for the radiation and matter dominated epochs of the standard big bang model.

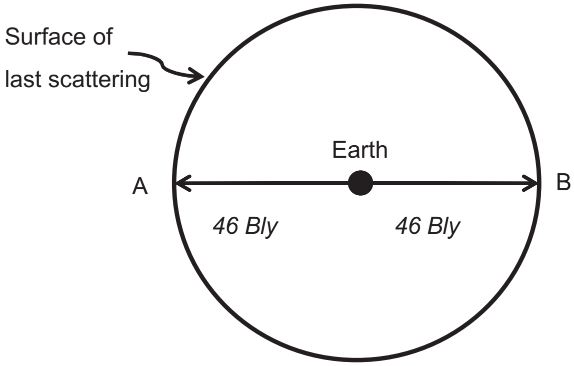

Fig. 15.2 CMB photons propagate to us from the surface of last scattering. In the standard big bang model, patches of the CMB at the points *A* and *B* have never been in causal contact. So why do they have almost identical temperatures?

15.2 The Horizon Problem: Why is the Universe so Homogeneous?

The near-uniformity of the CMB temperature over the sky tells us that the universe was extremely homogeneous at the time when the radiation was emitted. However, within the big bang model there is no reason why this ought to be the case. In fact, this ought *not* to be the case, unless the universe miraculously started out with very special initial conditions.

At first sight, a uniform temperature may not seem very surprising. A hot cup of tea left on a counter gradually cools down to room temperature. The CMB temperature could similarly equilibrate if there was some interaction between neighboring regions that emit radiation. However, when the CMB was emitted, the time that had elapsed since the big bang was too brief for such an interaction to have occurred. This is known as the *horizon problem.*

Consider the radiation coming to us from two small regions, A and B, which are diametrically opposite one another on the sky (see Fig. 15.2). The present distance to each of these regions is the distance to the surface of last scattering d_{ls}. Since the CMB radiation was emitted so early in the universe's history, there is not much difference between d_{ls} and the horizon distance d_{hor}. Thus the present distance to regions A and B is approximately equal to the horizon distance, $d_{hor} \approx 46\,\mathrm{Bly}$. The regions are therefore separated by twice this distance and cannot possibly interact. In particular, they cannot exchange heat to equalize their temperature—and yet they are observed to have equal temperatures, up to one part in a hundred thousand.

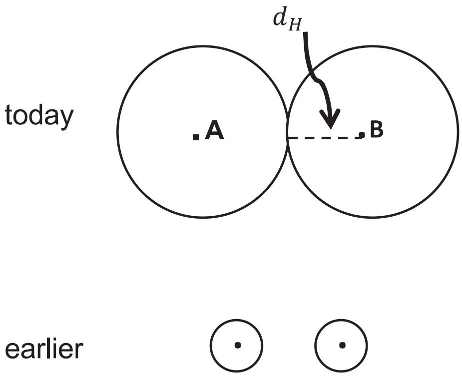

Fig. 15.3 Today regions A and B are separated by two horizon distances. *Circles* indicate the horizon distance from a region (*black dots*). As we go to earlier times, the regions get closer, but they are separated by an even larger number of horizon distances, because the horizon shrinks faster than the separation between the regions

Since the universe is expanding, regions that are distant today must have been in much closer proximity in the past. This, however, does not help to solve the problem. In fact, the horizon problem in the early universe is even more severe than it is today. To understand why, let us see how the separation between the regions d_{AB} and the horizon distance d_{hor} vary with time. Firstly, let's simplify the discussion by assuming that the universe has been in the matter dominated era[2] from the time the regions A and B emitted their radiation (at recombination) until the present. This means that their distance grows like the matter era scale factor $d_{AB}(t) \propto t^{\frac{2}{3}}$, and the horizon distance grows as $d_{hor}(t) \propto ct$. The horizon grows faster with time than the distance, which implies that as we go backwards from the present to the time of recombination, the horizon distance shrinks faster (see Fig. 15.3). It follows that if the separation distance d_{AB} exceeds the horizon, this excess could only be greater at earlier times. So, if two regions are now out of causal contact, they could not have been in causal contact before. For example, the regions A and B indicated in Fig. 15.2, which are now separated by $d_{AB} \approx 2d_{hor}$, were separated by approximately 80 horizon distances when the CMB radiation was emitted (see Question 5 at the end of this chapter). And at $t = 1$ s ABB they were separated by about 10^8 horizon distances.

[2]We disregard the recent period of accelerated expansion due to the dark energy. If we took this into account, our conclusions would remain the same.

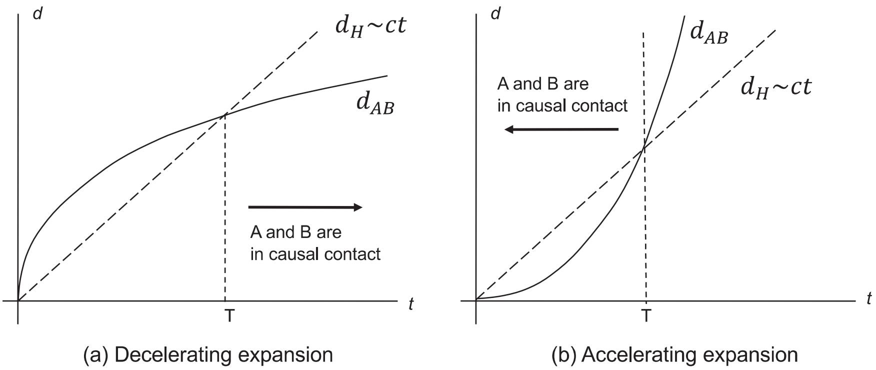

Fig. 15.4 **a** In a decelerating universe the slope of the curve $d_{AB}(t)$ decreases with time, while the horizon distance d_H has a constant slope. Regions A and B can only be in causal contact if $d_{AB} < d_H$. Initially, the universe expands so rapidly that $d_{AB} > d_H$. But, because of the decreasing slope of the curve $d_{AB}(t)$, there comes a time, denoted T, when it crosses the *straight line* representing the horizon. Once this happens, $d_{AB} < d_H$, and thus A and B are in causal contact at all later times. **b** In an accelerating universe, regions that are not in causal contact today (represented by times after T), were once in causal contact in the past (times earlier than T)

This means that hundreds of millions of causally separated regions would have had to have spontaneously started out in near perfect equilibrium in order for the CMB sky to have the uniform temperature distribution observed today.

Note that if one does assume such a special initial condition for the universe, then the hot big bang model is fully consistent. Neither the horizon nor the flatness problems are in contradiction with the big bang. They are problems in the sense that these remarkable features of the universe have no explanation within the theory.

The root of the horizon problem is that the expansion of the universe decelerates with time, so objects that are not in causal contact today could never have been in causal contact before. This makes one wonder what would happen if the universe underwent a stage of accelerated expansion. In such a universe, regions which are currently not in causal contact would in fact have been in causal contact at earlier times (see Fig. 15.4). So there would be no horizon problem. Moreover, the flatness problem would disappear as well! It can be shown that an accelerated expansion drives the value of Ω to one, even if initially it is significantly different from one. But what could cause accelerated expansion? You will have to wait until the next chapter to find out.

15.3 The Structure Problem: What is the Origin of Small Density Fluctuations?

We have marveled at the degree to which the universe is homogeneous. But even if we somehow explain the homogeneity, how do we account for the cosmic structures like galaxies, clusters of galaxies, and superclusters? In the big bang theory we have to postulate the existence of small density fluctuations, which gradually evolve into these structures. But what is the origin of the small initial density fluctuations?

15.4 The Monopole Problem: Where Are They?

As we discussed in Chap. 14, all GUT's predict that magnetic monopoles are produced during the big bang. Their initial density should be roughly one per horizon, which would result in a present density of about one monopole per cubic meter. This is comparable to the present number density of protons. But a monopole is much heavier than a proton (by a factor of 10^{16}). Thus, if monopoles were present at this density, their mass would far exceed the total mass in atomic and dark matter, in glaring conflict with observations.[3] This conundrum is known as the magnetic monopole problem.

Looming behind these and other problems is an even greater mystery: what actually happened at the big bang? What was the nature of the primordial force that launched the expansion of the universe and sent particles flying away from one another? All of these questions are addressed in the theory of cosmic inflation, to which we next turn.

Summary
The hot big bang theory is supported by a wealth of observational data, but it leaves unanswered some very intriguing questions about the initial state of the universe. For example, why is the geometry of the universe today so close to being flat? The flatness problem is exacerbated by the fact that geometry tends to veer away from flatness in the course of cosmic expansion. Hence, the universe had to be extremely close to flat at very early times.

[3]Our density estimate here is for GUT monopoles, but the problem persists even if the monopoles are formed at a lower energy scale.

Then there is the horizon problem: observations of the cosmic background radiation indicate that the early fireball was homogeneous on scales much greater than the horizon. Since no interactions can propagate faster than light, it seems that this homogeneity could not have been established by any causal process.

Other outstanding questions include: Why was the early universe expanding? What is the origin of small inhomogeneities that later evolved into galaxies? Where did all the magnetic monopoles go? What was the universe doing before the big bang?

Questions

1. In your own words, describe what the structure, horizon, and flatness problems are.
2. Do the flatness or horizon problems contradict the big bang theory? If so, explain. If not, explain the sense in which they are "problems".
3. Consider the graph in Fig. 15.1, showing how the density parameter Ω evolves with time. If $\Omega \approx 1$ today, what can we deduce about its value in the early universe?
4. In the hot big bang theory, does the universe undergo decelerated or accelerated expansion?
5 (a) Consider the regions A and B indicated in Fig. 15.2, which are now separated by 92 Bly. What was the distance d_{AB} between these regions at the time of recombination, $t_{rec} \approx 380{,}000$ yrs, when the radiation was emitted? (You may find the following facts useful: (i) the present CMB temperature is $T_0 \approx 3$ K; (ii) its temperature at recombination is $T_{\rm rec} = 3000$ K; (iii) the temperature changes with the scale factor according to $T \propto \frac{1}{a}$.)
   (b) Find the ratio d_{AB}/d_{hor} at the time of recombination, $t_{rec} \approx 380{,}000$ yrs. You may use the equation $d_{hor} \sim 3ct$ (introduced in Chap. 7) to approximate the horizon distance at recombination.