358

Which flatness problem does inflation solve?

Phillip Helbig[1,2] and James Overduin[3,4,5]

1. Thomas-Mann-Str. 9, 63477 Maintal, German
2. helbig@astro.multivax.de
3. Department of Physics, Astronomy and Geosciences, Towson University, Towson, Maryland, USA
4. Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland, USA
5. joverduin@towson.edu

Abstract

For several reasons, the flatness problem as originally proposed differs from the flatness problem today, both qualitatively and quantitatively; theory, observations, and the recent history of cosmology all play a role in this change.

Summary

The 'old flatness problem' was to explain why $\Omega + \lambda \approx 1$ (actually, at the time, just $\Omega \approx 1$, since $\lambda = 0$ was assumed) within a couple of orders of magnitude. Inflation was suggested as a solution to the flatness problem (*e.g.* Guth, 1981). Belief in the necessity of inflation, for this and other reasons, has made inflation a popular topic of study, even leading to predictions which have been confirmed. At the same time, several authors have claimed that the flatness problem is not really a problem, at least in the context of classical cosmology, though of course this is not evidence against inflation. It is still widely believed that the flatness problem is a real problem, perhaps because newer observations provide much better evidence for a flat universe: $\Omega+\lambda \approx 1$ not just to within a couple of orders of magnitude, but rather at the per-cent level. Whether some of the arguments against the old flatness problem also solve this 'new flatness problem', especially in the more realistic case of a perturbed Friedmann–Lemaître–Robertson–Walker model, is not completely clear. In any case, for a while after the formulation of the old flatness problem, cosmologists were perhaps right to believe in inflation but at least in part for the wrong reasons.

Notation

We assume that, at the level of detail necessary, the universe can be described by the Friedmann–Lemaître equation

$$\dot{R}^2 = \frac{8\pi G\rho R^2}{3} + \frac{\Lambda R^2}{3} - kc^2 \tag{1}$$

with the dimensionless constant k equal to $-1, 0, +1$ depending on spatial curvature (negative, vanishing, or positive, respectively); R is the scale factor (with dimension length) of the universe, G the gravitational constant, ρ the density of pressureless matter ('dust', including both baryonic and non-baryonic components), Λ the cosmological constant (dimension time^{-2}) and c the speed of light. (Our notation corresponds to that of Harrison (2000); in other schemes, *e.g.* that used by Heacox (2015), various terms can differ by factors of c^2.) It is useful to define the following quantities:

$$H := \frac{\dot{R}}{R}$$

$$\lambda := \frac{\Lambda}{3H^2}$$

$$\Omega := \frac{\rho}{\rho_{\mathrm{crit}}} \equiv \frac{8\pi G\rho}{3H^2}$$

$$K := \Omega + \lambda - 1$$

$$q := \frac{-\ddot{R}R}{\dot{R}^2} \equiv \frac{-\ddot{R}}{RH^2} \equiv \frac{\Omega}{2} - \lambda$$

which are all dimensionless except that H has the dimension time^{-1}. H is the Hubble constant, λ the normalized cosmological constant, Ω the density parameter, $k = \mathrm{sign}(K)$ and q is the deceleration parameter. For $\lambda = 0$ and $k = 0$, $\rho = \rho_{\mathrm{crit}} = \frac{3H^2}{8\pi G}$. This density is 'critical' in the sense that, for $\lambda = 0$, a greater (lesser) density implies a positive (negative) curvature and a universe—observed to be expanding now—which will collapse in the future (expand forever); similarly, for $k = 0$, a greater (lesser) density implies a negative (positive) cosmological constant and a universe—observed to be expanding now—which will collapse in the future (expand forever). However, in the general case ($\lambda \neq 0$ and $k \neq 0$), ρ_{crit} doesn't have any special meaning, though Ω remains a useful parameter. Equation (1) can be rearranged, using the definitions above, to give

$$R = \frac{c}{H}\frac{\mathrm{sign}(K)}{\sqrt{|K|}}, \tag{2}$$

thus R is positive for $k = +1$ and negative for $k = -1$; for $k = 0$, R can be defined as $\frac{c}{H}$. (Note that k and K have the same sign; the latter parameter, often written Ω_{k}, is sometimes defined with the opposite sign.)

It can be useful to express equation (1) with the values of the parameters as observed now, denoted by the suffix 0. This leads to

$$\dot{R}^2 = \dot{R}_0^2\left(\frac{\Omega_0 R_0}{R} + \frac{\lambda_0 R^2}{R_0^2} - K_0\right). \tag{3}$$

In general, H, λ and Ω all change with time. Note that since

$$\lambda = \lambda_0\left(\frac{H_0}{H}\right)^2, \tag{4}$$

the change in λ with time is due entirely to the change in H with time, since Λ is constant. Also, since the density ρ is inversely proportional to the cube of R,

$$\Omega = \Omega_0\left(\frac{H_0}{H}\right)^2\left(\frac{R_0}{R}\right)^3; \tag{5}$$

thus the variation in Ω is due both to variation in H and to the decrease in density as the universe expands.

In general, λ and Ω evolve with time. (They do not for the (λ,Ω) values of $(0,0)$ (Milne model), $(0,1)$ (Einstein–de Sitter model), $(1,0)$ (de Sitter model) and for the static Einstein model (in which λ and Ω are infinite (though Λ and ρ are not; $\Lambda = 4\pi G\rho$) since H is 0).) Thus, the history of model universes can be visualized as trajectories in the λ–Ω plane (see Fig. 1). For an excellent discussion of the evolution of λ and Ω (though expressed in the older notation using $\sigma = \frac{\Omega}{2}$ and $q = \sigma - \lambda$), see Stabell & Refsdal (1966).

The old flatness problem

In the original formulation of the flatness problem, Dicke & Peebles (1979) claimed that Ω must have been very close to 1 in the early universe, since otherwise it would not be of order 1 today. Helbig (2012) pointed out that there are two distinct issues: (1) Should we be puzzled that Ω was very close to 1 in the early universe? (2) Should we be puzzled that, *given* that Ω was close (not necessarily very close) to 1 in the early universe, it is still of order 1 today (Ω_0 being the present value)? It is important to remember that when Dicke & Peebles (1979) claimed that there is a flatness problem, Ω_0 was thought to be somewhere between 0.01 and 10 or so (*e.g.* Kolb & Turner, 1990) and λ was not mentioned at all (assumed to be 0, as was often the case then). Also, this 'old flatness problem' occurs in the context of *ideal* Friedmann–Lemaître–Robertson–Walker models (which does not mean that it does not occur in perturbed Friedmann–Lemaître–Robertson–Walker models). The original flatness problem thus sought an explanation for why Ω was equal to 1 *within a couple of orders of magnitude*, and was often discussed with the assumption that $\lambda = 0$ and usually in the context of ideal Friedmann–Lemaître–Robertson–Walker models.

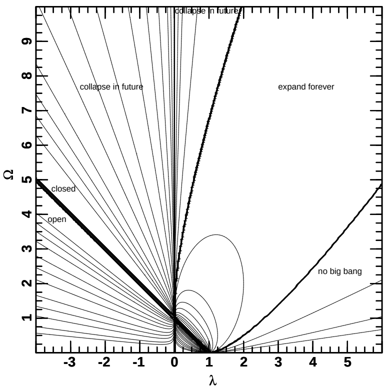

Figure 1: Evolutionary trajectories in the λ–Ω plane. The thick vertical line corresponds to $\lambda = 0$; the thick diagonal line corresponds to $k = 0$ with $k = -1$ below it and $k = +1$ above it. The thick curve near the vertical line separates models which will collapse (to the left) from those which will expand forever (to the right). Collapsing models retrace their trajectories in the opposite direction during the collapsing phase. Models on the curve start arbitrarily close to the Einstein–de Sitter model (like all non-empty big-bang models) and asymptotically approach the static Einstein model. The other thick curve separates big-bang models (to the left) from non–big-bang models (to the right); the latter contract from an infinite to a finite size then expand forever, retracing the same trajectory in the opposite direction (the initial and final points being arbitrarily close to the de Sitter model). Models on the curve (Eddington models) start at the static Einstein model and asymptotically approach the de Sitter model (the latter feature is shared with all models which expand forever and have $\lambda > 0$, except those, mentioned above, which asymptotically approach the static Einstein model). For non-empty models, trajectories between the two thick curves start arbitrarily close to the Einstein–de Sitter model and reach the de Sitter model (an attractor) after an infinite time. Empty models behave similarly, except that the Einstein–de Sitter model (a repulsor) is replaced by the Milne model (a saddle point). Due to time symmetry, big-bang models which expand forever have the same trajectories as big-crunch models which contract from infinity.

Many (*e.g.* Evrard & Coles, 1995; Coles & Ellis, 1997; Adler & Overduin, 2005; Helbig, 2012) have pointed out that the first problem is not really a problem at all, but simply a feature of the Friedmann–Lemaître–Robertson–Walker model: all non-empty big-bang models start off arbitrarily close to the Einstein–de Sitter model ($\lambda = 0$ and $\Omega = 1$). This is essentially an artifact of the mathematical definition of the critical density. One can ask *why* the universe is described by a Friedmann–Lemaître–Robertson–Walker model, but that is a different question. (Note that in this respect the flatness problem differs from the horizon problem: the latter can be solved, though unsatisfactorily, simply by demanding that the universe be described by a Friedmann–Lemaître–Robertson–Walker model; the former claims that there is a problem even *given* than the universe is described by a Friedmann–Lemaître–Robertson–Walker model, namely, an 'unlikely' value of the cosmological parameter Ω.) For *any* value of Ω at any time in a non-empty big-bang model, we can *always* find an earlier time when Ω differed from 1 by an arbitrarily small amount, so the fact that Ω was almost exactly 1 in the 'early universe' is not a problem at all, especially since there is no timescale in classical cosmology which could be used to define the 'early' universe (hence the value of Ω at the time of big-bang nucleosynthesis, say, is irrelevant).

Nevertheless, the second issue remains: Ω can become arbitrarily large (or small), so shouldn't we be puzzled that it is of order 1 today? No.

Consider first $\lambda \leq 0$ (of which the original formulation with $\lambda = 0$ is a special case, though not qualitatively different), which implies that the universe will collapse in the future (though for $\lambda = 0$, only if $\Omega > 1$; for $\Omega < 1$ see below). While Ω becomes arbitrarily large, this happens for only a relatively short time near the time of maximum expansion, so it is not puzzling that a typical observer does not observe a large value of Ω (Helbig, 2012).

If $\lambda > 0$, then, as pointed out by Lake (2005) (though he concentrated on the case that the universe expands forever), large values of Ω (and λ) are possible only in the case of fine-tuning between Ω and λ, namely $\alpha = \frac{27\Omega^2\lambda}{4K^3}$ must be ≈ 1 (this assumes $K > 0$ and hence $k = +1$, *i.e.* positive spatial curvature, but for other values of k, both Ω and λ are restricted to the interval $[0,1]$, so the question of large values does not arise). (There is a small portion of parameter space with $\lambda > 0$ and $\Omega > 1$ in which the universe has a maximum extent (for $\alpha = 1$ this is the static Einstein universe, which is reached after an infinite time; for $\alpha < 1$, the universe collapses in the future). For $\alpha \approx 1$, Lake's fine-tuning argument applies; for $\alpha \approx 0$, the same collapsing-universe argument applies as for $\lambda \leq 0$.) This is in marked contrast to the claim of the old flatness problem (which implicitly assumed that $\lambda = 0$) that fine-tuning is *required* to *avoid* large values of Ω. As mentioned above, this claim is misleading, since the fine-tuning is built into the model, but for $\lambda > 0$, large values never arise at all unless (another type of) fine-tuning is present. Related to this is the fact that in the original flatness problem with $\lambda = 0$, a nearly flat universe *now* nevertheless has arbitrarily large (or small) values of Ω in the future. In Lake's case, a 'solved flatness problem' (no fine-tuning, *i.e.* α is not of order 1) now implies that it is solved for all time in the sense that $|K|$ is never very large.

This demonstrates a *qualitative* difference, compared to the $\lambda = 0$ case, in the flatness problem if $\lambda > 0$ and $k = +1$. (Most discussions of the flatness problem do not take the theoretical progress in classical cosmology since the initial claim by Dicke & Peebles (1979) into account.) For $\lambda < 0$ (which implies that the universe will collapse after a period of expansion) and $\Omega < 1$ (now), the behaviour is also different: instead of asymptotically approaching 0, Ω evolves to a minimum (which always occurs at $q = \Omega/2 - \lambda = 1/2$ (Stabell & Refsdal, 1966)) then evolves to infinity (and then, during the contracting phase, back to the Einstein–de Sitter model along the same trajectory in the λ–Ω plane). However, except for the period when $\Omega < 1$, the behaviour is roughly the same as in the $\lambda = 0$ case with $\Omega > 1$ (where Ω increases monotonically during expansion and decreases monotonically during contraction). For $\lambda > 0$ and $\Omega < 1$ and for $\lambda < 0$ and $\Omega > 1$ (at all times), the behaviour of Ω with time is similar to the $\lambda = 0$ case.

Except for the case studied by Lake (2005) ($\lambda > 0$, $k = +1$), for $\lambda \neq 0$ the behaviour of Ω is thus by and large qualitatively similar to the $\lambda = 0$ case. The behaviour of λ is similar to that of Ω, except that for $\lambda < 0$ ('corresponding' to $\Omega > 1$) it evolves to $-\infty$ instead of $+\infty$ and for $\lambda > 0$ ('corresponding' to $\Omega < 1$) it evolves to 1 rather than 0 (which is the ultimate value of Ω if the universe expands forever, except for the Einstein–de Sitter model and those which asymptotically approach the static Einstein model). (Again, there is a small portion of parameter space with $\lambda > 0$ and $\Omega > 1$ where the final-state model is not the de Sitter model and the behaviour is similar to Lake's case or the $\lambda = 0$ case depending on the value of α, as mentioned above.) Also, the behaviour of K is similar to that of Ω, though here the special value is 0 instead of 1: if $K = 0$ exactly at any time, then $K = 0$ exactly at all times, otherwise $|K|$ can become arbitrarily large (K can be positive or negative), though if $\lambda > 0$ and $k \leq 0$, it is restricted to the interval $[-1, 0]$; for $k = +1$ and $\lambda > 0$, like λ and Ω, K can become large only in the case of fine-tuning. However, even if $K = 0$ exactly, Ω and λ individually have qualitatively similar behaviour to the $K \neq 0$ case.

The original flatness problem with $\lambda = 0$ was concerned not only with $\Omega > 1$ and hence arbitrarily large values of Ω, but also with $\Omega < 1$ and hence arbitrarily small values of Ω. However, the limit $\Omega = 0$ is reached only after an infinite time, so there is an obvious weak-anthropic argument as to why such extremely small values are not observed. (If the universe lasts forever, why are we near the beginning?) This is also the case for $\lambda > 0$. See Helbig (2012) for more details.

The new flatness problem

As explained above, it is not puzzling at all that Ω_0—or, more appropriate today, K—is observed to be between 0.1 and 10 or so. However, we now have much stricter bounds: $K = 0.003 \pm 0.004$ from *WMAP* (Bennett et al., 2013; Hinshaw et al., 2013) and $|K| < 0.005$ from *Planck* (*Planck* Collaboration, 2016) (CMB and other constraints in both cases). Should we be puzzled that $|K|$ is *this* small? If the universe were to collapse in the future, yes. In that case, the arguments above cannot explain the fact that the universe is flat to this degree of precision. If the universe will expand forever (as observational data indicate), the situation is less clear. Lake (2005) showed that the universe is 'nearly flat' as long as α is not fine-tuned to be of order 1. But how different from 1 must α be to avoid the impression of fine-tuning? 10? 100? 1000? Should we be surprised if α is much larger? The current constraints on λ_0, Ω_0, and K imply that $|\alpha|$ is larger than 3 million or so.

At first glance, $K = 0$ seems unlikely since the sum of λ and Ω must add up to exactly 1. If these are 'random' quantities, then of course $K = 0$ is a space of measure 0. However, λ and Ω in general change with time in a non-random way, which makes this probability measure less intuitive. Alternatively, $K = 0$ corresponds to an infinite radius of curvature, and of course a 'random' number between 0 and ∞ is likely to be very large, so seen this way $K \approx 0$ appears likely. Of course, both can't be right and, like λ and Ω, the (comoving) curvature radius in general also evolves with time. This aspect is avoided by using α as a parameter to distinguish models, since by construction it is a constant of motion. (Thus, a contour plot of α demonstrates trajectories in the λ–Ω plane; Fig. 1 shows contours for $\alpha = \{0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500\}$.) Like the radius of curvature, it can be anywhere in the interval between $-\infty$ and ∞. So does the fact that a 'random' value of α is probably very large automatically solve this 'new flatness problem' as well as the old? (Note that for $k = +1$ and thus a finite universe, α corresponds, up to a constant factor, to the ratio of Λ ($\Lambda = 3H^2\lambda$) to the square of the mass of the universe.) It does if one considers the observed value of α to be large enough so that it cannot be considered to be fine-tuned to be ≈ 1, at least in the limit of a completely homogeneous and isotropic universe; whether Lake's argument solves the flatness problem in a perturbed Friedmann–Lemaître–Robertson–Walker model is less clear.

Of course, this does not mean that inflation cannot occur, and does not mean that it did not occur. However, it means that the observed flatness of the universe cannot be used as an argument in favour of inflation (though of course it is compatible with it), at least in the case of pure Friedmann–Lemaître–Robertson–Walker models. Today, though, this argument is no longer needed, since there is other evidence for inflation, in particular, the robust prediction that the spectral index n describing fluctuations seen in the CMB power spectrum is slightly, but definitely, less than 1: $n_{\mathrm{S}} = 0.968 \pm 0.006$ (*e.g.* *Planck* Collaboration, 2016).

Historical aspects

When the old flatness problem was first formulated, constraints on Ω were weak; the problem was to explain why $\Omega + \lambda \approx 1$ (or, then, just $\Omega \approx 1$, since $\lambda = 0$ was assumed) within a couple of orders of magnitude, not the much flatter universe observed today. Due to the assumption $\lambda = 0$, at that time belief in a flat universe meant belief in the Einstein–de Sitter universe. Mainstream cosmologists in the period 1980–1995 or so tended to believe (a) that inflation makes the universe flat and (b) that the Einstein–de Sitter model (which is flat, but has $\lambda = 0$) is correct (*e.g.* Sandage, 1995). Current observations (*e.g.* *Planck* Collaboration, 2016) indicate that the universe is much flatter than was known during that period and also that the Einstein–de Sitter model is not correct. Although it took some time, the Einstein–de Sitter model is no longer considered a viable model for our universe, even though not that long ago it was considered almost crackpot to question it (*e.g.* Overbye, 1991; Kolb, 1998). At the same time, it is still widely believed that the old flatness problem is a real problem. Why the difference? Probably the main reason is that observations directly rule out the Einstein–de Sitter model while the flatness problem is more of a conceptual issue. Also, the observation that the universe is flat—even much flatter than was suspected 40 years ago—has tended to strengthen belief in the necessity of inflation, even though it is not needed to solve the old flatness problem, and, at least assuming that $\lambda > 0$ and that the universe will expand forever, might not be needed to solve the new flatness problem. There are, of course, other reasons to believe that inflation occurred, and this has tended to obscure the fact that one of the main reasons for interest in it in the past 40 years—namely that it can solve the old flatness problem—was actually not really there, though this belief played an important role in keeping the concept of inflation alive. So it turns out that in some respects cosmologists were right for the wrong reasons.

References

Adler R.J., Overduin J.M., 2005, *Gen. Relativ. Gravit.*, **37**, 1491

Bennett C.L., et al., 2013, *ApJS*, **208**, 20

Coles P., Ellis G.F.R., 1997, *Is the Universe Open or Closed?* **No. 7** in *Cambridge Lecture Notes in Physics*, Cambridge University Press, Cambridge (UK)

Dicke R.H., Peebles P.J.E., 1979, In Hawking S.W., Israel W., eds., *General Relativity: An Einstein Centenary Survey*, Cambridge University Press, Cambridge (UK), pp. 504–517

Evrard G., Coles P., 1995, *Class. Quant. Grav.*, **12**, L93

Guth A.H., 1981, *Phys. Rev. D*, **23**, 347

Harrison E.R., 2000, *Cosmology, the Science of the Universe*. Cambridge University Press, Cambridge (UK), 2nd edn.

Heacox W.D., 2015, *The Expanding Universe: A Primer on Relativistic Cosmology*. Cambridge University Press, Cambridge (UK)

Helbig P., 2012, *MNRAS*, **421**, 561

Hinshaw G., et al., 2013, *ApJS*, **208**, 19

Kolb E.W., 1998, In Vâ J.T.T., Giraud-Héraud Y., Bouchet F., Damour T., Mellier Y., eds., *Fundamental Parameters in Cosmology, Rencontres de Moriond*, Éditions Frontières, Paris

Kolb E.W., Turner M.S., 1990, *The Early Universe*. Addison Wesley, Boston

Lake K., 2005, *Phys. Rev. Lett.*, **94**, 201102

Overbye D., 1991, *Lonely Hearts of the Cosmos*. HarperCollins, New York

Planck Collaboration, 2016, *A&A*, **593**, A13

Sandage A.R., 1995, In Binggeli B., Buser R., eds., *The Deep Universe*, Springer, Berlin

Stabell R., Refsdal S., 1966, *MNRAS*, **132**, 379