8.5: Cosmological Solutions (Part 3)

The Vacuum-dominated Solution

For 70 years after Hubble's discovery of cosmological expansion, the standard picture was one in which the universe expanded, but the expansion must be decelerating. The deceleration is predicted by the special cases of the FRW cosmology that were believed to be applicable, and even if we didn't know anything about general relativity, it would be reasonable to expect a deceleration due to the mutual Newtonian gravitational attraction of all the mass in the universe.

But observations of distant supernovae starting around 1998 introduced a further twist in the plot. In a binary star system consisting of a white dwarf and a non-degenerate star, as the nondegenerate star evolves into a red giant, its size increases, and it can begin dumping mass onto the white dwarf. This can cause the white dwarf to exceed the *Chandrasekhar limit* (Section 4.4), resulting in an explosion known as a type Ia supernova. Because the Chandrasekhar limit provides a uniform set of initial conditions, the behavior of type Ia supernovae is fairly predictable, and in particular their luminosities are approximately equal. They therefore provide a kind of standard candle: since the intrinsic brightness is known, the distance can be inferred from the apparent brightness. Given the distance, we can infer the time that was spent in transit by the light on its way to us, i.e. the look-back time. From measurements of Doppler shifts of spectral lines, we can also find the velocity at which the supernova was receding from us. The result is that we can measure the universe's rate of expansion as a function of time. Observations show that this rate of expansion has been accelerating. The Friedmann equations show that this can only occur for $\Lambda \gtrsim 4\rho$. This picture has been independently verified by measurements of the cosmic microwave background (CMB) radiation. A more detailed discussion of the supernova and CMB data is given in Section 8.2.

With hindsight, we can see that in a quantum-mechanical context, it is natural to expect that fluctuations of the vacuum, required by the Heisenberg uncertainty principle, would contribute to the cosmological constant, and in fact models tend to overpredict Λ by a factor of about 10^{120}! From this point of view, the mystery is why these effects cancel out so precisely. A correct understanding of the cosmological constant presumably requires a full theory of quantum gravity, which is presently far out of our reach.

The latest data show that our universe, in the present epoch, is dominated by the cosmological constant, so as an approximation we can write the Friedmann equations as

$$\frac{\ddot{a}}{a} = \frac{1}{3}\Lambda$$
$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{1}{3}\Lambda.$$

This is referred to as a *vacuum-dominated universe* or the *de Sitter spacetime*. The solution is

$$a = e^{\sqrt{\frac{\Lambda}{3}}t}, \tag{8.2.5}$$

where observations show that $\Lambda \sim 10^{-26}$ kg/m^3, giving $\sqrt{\frac{3}{\Lambda}} \sim 10^{11}$ years.

The implications for the fate of the universe are depressing. All parts of the universe will accelerate away from one another faster and faster as time goes on. The relative separation between two objects, say galaxy A and galaxy B, will eventually be increasing faster than the speed of light. (The Lorentzian character of spacetime is local, so relative motion faster than c is only forbidden between objects that are passing right by one another.) At this point, an observer in either galaxy will say that the other one has passed behind an event horizon. If intelligent observers do actually exist in the far future, they may have no way to tell that the cosmos even exists. They will perceive themselves as living in island universes, such as we believed our own galaxy to be a hundred years ago.

When I introduced the standard cosmological coordinates earlier, I described them as coordinates in which events that are simultaneous according to this t are events at which the local properties of the universe are the same. In the case of a perfectly vacuum-dominated universe, however, this notion loses its meaning. The only observable local property of such a universe is the vacuum energy described by the cosmological constant, and its density is always the same, because it is built into the structure of the vacuum. Thus the vacuum-dominated cosmology is a special one that maximally symmetric, in the sense that it has not only the symmetries of homogeneity and isotropy that we've been assuming all along, but also a symmetry with respect to time: it is a cosmology without history, in which all times appear identical to a local observer. One way of checking this claim is by calculating curvature scalars, and we find, for example, that the Ricci scalar is a constant R = −12Λ (with the sign depending on the + − −− signature, example 25).

In the special case of this cosmology, the time variation of the scaling factor a(t) is unobservable, and may be thought of as the unfortunate result of choosing an inappropriate set of coordinates, which obscure the underlying symmetry. When I argued in section 8.2 for the observability of the universe's expansion, note that all my arguments assumed the presence of matter or radiation. These are completely absent in a perfectly vacuum-dominated cosmology.

For these reasons de Sitter originally proposed this solution as a static universe in 1927. But by 1920 it was realized that this was an oversimplification. The argument above only shows that the time variation of a(t) does not allow us to distinguish one epoch of the universe from another. That is, we can't look out the window and infer the date (e.g., from the temperature of the cosmic microwave background radiation). It does not, however, imply that the universe is static in the sense that had been assumed until Hubble's observations. The r-t part of the metric is

$$ds^2 = dt^2 - a^2 dr^2, \tag{8.2.6}$$

where a blows up exponentially with time, and the k-dependence has been neglected, as it was in the approximation to the Friedmann equations used to derive $a(t)$.[21] Let a test particle travel in the radial direction, starting at event A = (0, 0) and ending at B = (t', r'). In flat space, a world-line of the linear form r = vt would be a geodesic connecting A and B; it would maximize the particle's proper time. But in the this metric, it cannot be a geodesic. The curvature of geodesics relative to a line on an r-t plot is most easily understood in the limit where t 0 p is fairly long compared to the time-scale T = $\sqrt{\frac{3}{\Lambda}}$ of the exponential, so that a(t') is huge. The particle's best strategy for maximizing its proper time is to make sure that its dr is extremely small when a is extremely large. The geodesic must therefore have nearly constant r at the end. This makes it sound as though the particle was decelerating, but in fact the opposite is true. If r is constant, then the particle's spacelike distance from the origin is just ra(t), which blows up exponentially. The near-constancy of the coordinate r at large t actually means that the particle's motion at large t isn't really due to the particle's inertial memory of its original motion, as in Newton's first law. What happens instead is that the particle's initial motion allows it to move some distance away from the origin during a time on the order of T, but after that, the expansion of the universe has become so rapid that the particle's motion simply streams outward because of the expansion of space itself. Its initial motion only mattered because it determined how far out the particle got before being swept away by the exponential expansion.

Note

A computation of the Einstein tensor with $ds^2 = d\bar{t}^2 - a^2 (1 - kr^2)^{-1} dr^2$ shows that k enters only via a factor the form $(\ldots)e^{(\ldots)t} + (\ldots)k$. For large t, the k term becomes negligible, and the Einstein tensor becomes $G^a{}_b = g^a{}_b\Lambda$, This is consistent with the approximation we used in deriving the solution, which was to ignore both the source terms and the k term in the Friedmann equations. The exact solutions with Λ > 0 and k = −1, 0, and 1 turn out in fact to be equivalent except for a change of coordinates.

Example 19: Geodesics in a vacuum-dominated universe

In this example we confirm the above interpretation in the special case where the particle, rather than being released in motion at the origin, is released at some nonzero radius r, with $\frac{dr}{dt}$ = 0 initially. First we recall the geodesic equation

$$\frac{d^2x^i}{d\lambda^2} = \Gamma^i_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda}. \tag{8.2.7}$$

from section 5.7. The nonvanishing Christoffel symbols for the 1+1- dimensional metric $ds^2 = dt^2 - a^2dr^2$ are $\Gamma^r_{tr} = \frac{\dot{a}}{a}$ and $\Gamma^t_{rr} = \dot{a}a$. Setting T = 1 for convenience, we have Γ^r_{tr} = 1 and $\Gamma^t_{rr} = e^{-2t}$.

We conjecture that the particle remains at the same value of r. Given this conjecture, the particle's proper time $\int$ ds is simply the same as its time coordinate t, and we can therefore use t as an affine coordinate. Letting λ = t, we have

$$\frac{d^2t}{dt^2} - \Gamma^t_{rr}\left(\frac{dr}{dt}\right)^2 = 0 \tag{8.2.8}$$

$$0 - \Gamma^t_{rr}\dot{r}^2 = 0 \tag{8.2.9}$$

$$\dot{r} = 0 \tag{8.2.10}$$

$$r = constant \tag{8.2.11}$$

This confirms the self-consistency of the conjecture that r = constant is a geodesic.

Note that we never actually had to use the actual expressions for the Christoffel symbols; we only needed to know which of them vanished and which didn't. The conclusion depended only on the fact that the metric had the form $ds^2 = dt^2 - a^2dr^2$ for some function a(t). This provides a rigorous justification for the interpretation of the cosmological scale factor a as giving a universal time-variation on all distance scales.

The calculation also confirms that there is nothing special about r = 0. A particle released with r = 0 and $\dot{r} = 0$ initially stays at r = 0, but a particle released at any other value of r also stays at that r. This cosmology is homogeneous, so any point could have been chosen as r = 0. If we sprinkle test particles, all at rest, across the surface of a sphere centered on this arbitrarily chosen point, then they will all accelerate outward relative to one another, and the volume of the sphere will increase. This is exactly what we expect. The Ricci curvature is interpreted as the second derivative of the volume of a region of space defined by test particles in this way. The fact that the second derivative is positive rather than negative tells us that we are observing the kind of repulsion provided by the cosmological constant, not the attraction that results from the existence of material sources.

Example 20: Schwarzschild-de Sitter space

The metric

$$ds^2 = \left(1 - \frac{2m}{r} - \frac{1}{3}\Lambda r^2\right) dt^2 - \frac{dr^2}{1 - \frac{2m}{r} - \frac{1}{3}\Lambda r^2} - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2 \tag{8.2.12}$$

is an exact solution to the Einstein field equations with cosmological constant Λ, and can be interpreted as a universe in which the only mass is a black hole of mass m located at r = 0. Near the black hole, the Λ terms become negligible, and this is simply the Schwarzschild metric. As argued in section 8.2, this is a simple example of how cosmological expansion does not cause all structures in the universe to grow at the same rate.

Example 21: Conservation of energy-momentum

Suppose that we assume the de Sitter geometry, and ask what type of matter fields are necessary to create it. We know that a cosmological constant will do the job, but could we have some other matter field that would also work? Suppose that the matter field is constrained to be a perfect fluid. The total stress-energy is then of the form T_ν^μ = diag(ρ, −P, −P, −P) in Cartesian coordinates. (See example 4 for the signs, some of which depend on our use of the + − −− signature.) The divergence $\nabla_\mu T_t^\mu$ measures the rate at which an observer says energy is being created, and we need this to be zero. This expression is one of those tricky examples where the covariant derivative can be nonzero even when the thing being differentiated vanishes identically. The divergence is $\nabla_t T_t^t + \nabla_x T_t^x$, and the term that doesn't vanish is the second one, even though $T^x{}_t = 0$. Using the nonvanishing Christoffel symbols this becomes $\Gamma^x_{xt} T^t_t - \Gamma^x_{tx} T^x_x = \frac{\dot{a}}{a}(\rho + P)$, so that ρ + P = 0. This condition is satisfied by a cosmological constant. Our result is that the only way to get a de Sitter geometry is with matter fields that exactly mimic a cosmological constant. This is of some historical interest in the context of the steady-state cosmologies, section 8.4. It may seem mysterious that we have obtained this result by requiring conservation of energy-momentum, but we could also have done it using the Einstein field equations. In fact these are not two separate requirements, since the field equations require conservation of energy-momentum in order to be consistent.

The Big Bang Singularity in a Universe with a Cosmological Constant

Earlier we discussed the possibility that the Big Bang singularity was an artifact of the unrealistically perfect symmetry assumed by our cosmological models, and we found that this was not the case: the Penrose-Hawking singularity theorems demonstrate that the singularity is real, provided that the cosmological constant is zero. The cosmological constant is *not* zero, however. Models with a very large positive cosmological constant can also display a Big Bounce rather than a Big Bang. If we imagine using the Friedmann equations to evolve the universe backward in time from its present state, the scaling arguments of example 14 suggest that at early enough times, radiation and matter should dominate over the cosmological constant. For a large enough value of the cosmological constant, however, it can happen that this switch-over never happens. In such a model, the universe is and always has been dominated by the cosmological constant, and we get a Big Bounce in the past because of the cosmological constant's repulsion. In this book I will only develop simple cosmological models in which the universe is dominated by a single component; for a discussion of bouncing models with both matter and a cosmological constant, see Carroll, "The Cosmological Constant," http://www.livingreviews.org/lrr-2001-1. By 2008, a variety of observational data had pinned down the cosmological constant well enough to rule out the possibility of a bounce caused by a very strong cosmological constant.

The Matter-dominated Solution

Our universe is not perfectly vacuum-dominated, and in the past it was even less so. Let us consider the matter-dominated epoch, in which the cosmological constant was negligible compared to the material sources. The equation of state for nonrelativistic matter (example 4) is

$$P=0. \tag{8.2.13}$$

The dilution of the dust with cosmological expansion gives

$$\rho \propto a^{-3} \tag{8.2.14}$$

(see example 23). The Friedmann equations become

$$\begin{aligned} \frac{\ddot{a}}{a} &= -\frac{4\pi}{3}\rho \\ \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi}{3}\rho - ka^{-2}, \end{aligned}$$

where for compactness ρ's dependence on a, with some constant of proportionality, is not shown explicitly. A static solution, with constant a, is impossible, and $\ddot{a}$ is negative, which we can interpret in Newtonian terms as the deceleration of the matter in the universe due to gravitational attraction. There are three cases to consider, according to the value of k.

The Closed Universe

We've seen that k = +1 describes a universe in which the spatial curvature is positive, i.e., the circumference of a circle is less than its Euclidean value. By analogy with a sphere, which is the twodimensional surface of constant positive curvature, we expect that the total volume of this universe is finite.

The second Friedmann equation also shows us that at some value of a, we will have $\dot{a}$ = 0. The universe will expand, stop, and then recollapse, eventually coming back together in a "Big Crunch" which is the time-reversed version of the Big Bang.

Suppose we were to describe an initial-value problem in this cosmology, in which the initial conditions are given for all points in the universe on some spacelike surface, say t = constant. Since the universe is assumed to be homogeneous at all times, there are really only three numbers to specify, a, $\dot{a}$, and ρ: how big is the universe, how fast is it expanding, and how much matter is in it? But these three pieces of data may or may not be consistent with the second Friedmann equation. That is, the problem is overdetermined. In particular, we can see that for small enough values of ρ, we do not have a valid solution, since the square of $\frac{\dot{a}}{a}$ would have to be negative. Thus a closed universe requires a certain amount of matter in it. The present observational evidence (from supernovae and the cosmic microwave background, as described above) is sufficient to show that our universe does not contain this much matter.

The Flat Universe

The case of k = 0 describes a universe that is spatially flat. It represents a knife-edge case lying between the closed and open universes. In a Newtonian analogy, it represents the case in which the universe is moving exactly at escape velocity; as t approaches infinity, we have a → ∞, ρ → 0, and $\dot{a}$ → 0. This case, unlike the others, allows an easy closed-form solution to the motion. Let the constant of proportionality in the equation of state $\rho \propto a^{-3}$ be fixed by setting $\frac{-4\pi\rho}{3}$ = −ca^{-3}. The Friedmann equations are

$$\begin{aligned} \ddot{a} &= -ca^{-2} \\ \dot{a} &= \sqrt{2c}a^{-1/2}. \end{aligned}$$

Looking for a solution of the form a ∝ t^{p}, we find that by choosing p = 2/3 we can simultaneously satisfy both equations. The constant c is also fixed, and we can investigate this most transparently by recognizing that $\frac{\dot{a}}{a}$ is interpreted as the Hubble constant, H, which is the constant of proportionality relating a far-off galaxy's velocity to its distance. Note that H is a "constant" in the sense that it is the same for all galaxies, in this particular model with a vanishing cosmological constant; it does not stay constant with the passage of cosmological time. Plugging back into the original form of the Friedmann equations, we find that the flat universe can only exist if the density of matter satisfies $\rho = \rho_{crit} = \frac{3H^2}{8\pi} = \frac{3H^2}{8\pi G}$. The observed value of the Hubble constant is about $1/(14 \times 10^9$ years), which is roughly interpreted as the age of the universe, i.e., the proper time experienced by a test particle since the Big Bang. This gives $\rho_{crit} \sim 10^{-26}$ kg/m^3.

As discussed in section 8.2, our universe turns out to be almost exactly spatially flat. Although it is presently vacuum-dominated, the flat and matter-dominated FRW cosmology is a useful description of its matter-dominated era.

The Open Universe

The $k = -1$ case represents a universe that has negative spatial curvature, is spatially infinite, and is also infinite in time, i.e., even if the cosmological constant had been zero, the expansion of the universe would have had too little matter in it to cause it to recontract and end in a Big Crunch.

The time-reversal symmetry of general relativity was discussed in section 6.2 in connection with the Schwarzschild metric.[22] Because of this symmetry, we expect that solutions to the field equations will be symmetric under time reversal (unless asymmetric boundary conditions were imposed). The closed universe has exactly this type of time-reversal symmetry. But the open universe clearly breaks this symmetry, and this is why we speak of the Big Bang as lying in the past, not in the future. This is an example of spontaneous symmetry breaking. Spontaneous symmetry breaking happens when we try to balance a pencil on its tip, and it is also an important phenomenon in particle physics. The time-reversed version of the open universe is an equally valid solution of the field equations. Another example of spontaneous symmetry breaking in cosmological solutions is that the solutions have a preferred frame of reference, which is the one at rest relative to the cosmic microwave background and the average motion of the galaxies. This is referred to as the Hubble flow.

Note

Problem 5 shows that this symmetry is also exhibited by the Friedmann equations.

Example 22: Size and age of the observable universe

The observable universe is defined by the region from which light has had time to reach us since the Big Bang. Many people are inclined to assume that its radius in units of light-years must therefore be equal to the age of the universe expressed in years. This is not true. Cosmological distances like these are not even uniquely defined, because general relativity only has local frames of reference, not global ones.

Suppose we adopt the proper distance L defined in section 8.2 as our measure of radius. By this measure, realistic cosmological models say that our 14-billion-year-old universe has a radius of 46 billion light years.

For a flat universe, $f = 1$, so by inspecting the FRW metric we find that a photon moving radially with $ds = 0$ has $|\frac{dr}{dt}| = a^{-1}$, giving $r = \pm\int_{t_1}^{t_2} \frac{dt}{a}$. Suppressing signs, the proper distance the photon traverses starting soon after the Big Bang is $L = a(t_2) \int d\ell = a(t_2) \int dr = a(t_2) r = a(t_2) \int_{t_1}^{t_2} \frac{dt}{a}$. In the matter-dominated case, $a \propto t^{2/3}$, so this results in $L = 3t_2$ in the limit where t1 is small. Our universe has spent most of its history being matter-dominated, so it's encouraging that the matter-dominated calculation seems to do a pretty good job of reproducing the actual ratio of $\frac{46}{14} = 3.3$ between L and t_2.

While we're at it, we can see what happens in the purely vacuumdominated case, which has $a \propto e^{t/T}$, where $T = \sqrt{\frac{3}{\Lambda}}$. This cosmology doesn't have a Big Bang, but we can think of it as an approximation to the more recent history of the universe, glued on to an earlier matter-dominated solution. Here we find $L = [e^{(t_2-t_1)/T} - 1]T$, where t_1 is the time when the switch to vacuum-domination happened. This function grows more quickly with t_2 than the one obtained in the matter-dominated case, so it makes sense that the real-world ratio of $\frac{L}{t_2}$ is somewhat greater than the matter-dominated value of 3.

The radiation-dominated version is handled in problem 12.

Example 23: Local conservation of mass-energy

Any solution to the Friedmann equations is a solution of the field equations, and therefore locally conserves mass-energy. We saved work above by applying this condition in advance in the form $\rho \propto a^{-3}$ to make the dust dilute itself properly with cosmological expansion. In this example we prove the same proportionality by explicit calculation.

Local conservation of mass-energy is expressed by the zero divergence of the stress-energy tensor, $\nabla_j T^{jb} = 0$. The definition of the covariant derivative gives

$$\nabla_j T^{bc} = \partial_j T^{bc} + \Gamma^b_{jd} T^{dc} + \Gamma^c_{jd} T^{bd}. \tag{8.2.15}$$

For convenience, we carry out the calculation at r = 0; if conservation holds here, then it holds everywhere by homogeneity.

In a local Cartesian frame (t', x', y', z') at rest relative to the dust, the stress-energy tensor is diagonal with $T^{t't'} = \rho$. At r = 0, the transformation from FLRW coordinates into these coordinates doesn't mix t or t' with the other coordinates, so by the tensor transformation law we still have $T^{tt} = \rho$.

There are a number of Christoffel symbols involved, but the only three of relevance that don't vanish at r = 0 turn out to be $\Gamma^r_{rt} = \Gamma^\theta_{\theta t} = \Gamma^\phi_{\phi t} = \frac{\dot{a}}{a}$. The result is

$$\nabla_\mu T^{t\mu} = \partial_t T^{tt} + 3\frac{\dot{a}}{a} T^{tt}, \tag{8.2.16}$$

or $\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a}$, which can be rewritten as

$$\frac{d}{dt}\ln\rho = -3\frac{d}{dt}\ln a, \tag{8.2.17}$$

producing the proportionality originally claimed.