5
An Expanding Universe

5.1 Einstein's Static Universe

Shortly after completing the general theory of relativity, Einstein went on to apply his new theory to the universe as a whole. The structure of the universe beyond our Milky Way galaxy was then completely unknown, so Einstein had to make some assumptions. Following Newton, he assumed that on average matter is uniformly distributed in the cosmos. There are of course local variations, with the density of stars higher in some places and lower in others. However on very large scales the universe is well approximated as being perfectly homogeneous.

Einstein also assumed that the universe is isotropic on average. This means that it looks more or less the same in all directions. A homogeneous and isotropic (on average) distribution is illustrated in Fig. 5.1(a), with galaxies represented by dots. An example of a homogeneous distribution that is not isotropic is shown in Fig. 5.1(b), where the galaxies (dots) form a regular lattice. This distribution looks the same from every galaxy, but it looks different in horizontal, vertical, and diagonal directions. The actual distribution of galaxies, as revealed by modern astronomical observations, is more complicated than that in Fig. 5.1(a). Individual galaxies form clusters, which are in turn grouped into huge superclusters, typically 150 million light years across. But that is where the hierarchy of cosmic structure appears to end. If the distribution of galaxies is smoothed over distances of 300 million light years or so, it does appear to be homogeneous and isotropic.

D. Perlov and A. Vilenkin, *Cosmology for the Curious*,
DOI 10.1007/978-3-319-57040-2_5

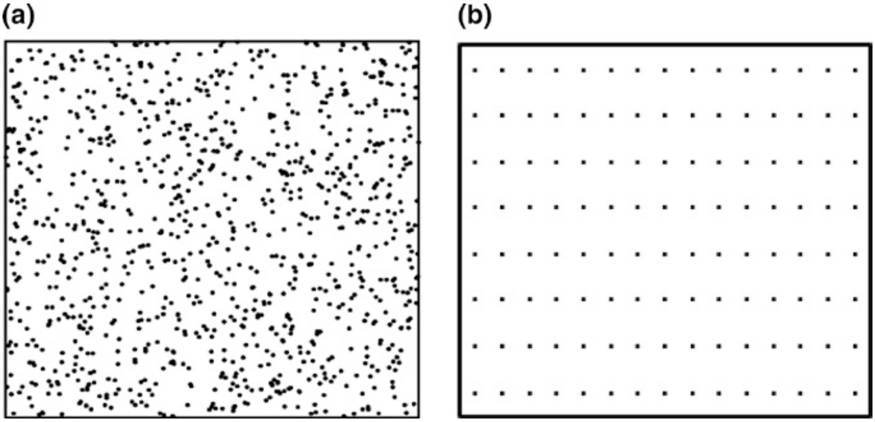

Fig. 5.1 **a** A homogeneous and isotropic (on average) distribution of galaxies. **b** This distribution is homogeneous, but not isotropic

The universe cannot be homogeneous and isotropic unless space itself has these properties. The curvature of space should be (on average) the same at all places and in all directions. As we discussed in Chap. 4, there are only three types of such spaces: a flat Euclidean space, a closed spherical space, and an open hyperbolic space. A homogeneous and isotropic universe should therefore have one of these three geometries.

Finally, Einstein assumed that the average characteristics of the universe, such as the average density of stars, do not change with time. His overall picture was thus that the universe looks more or less the same at all places, in all directions, and at all times. Einstein did not have much observational data to back up his assumptions, but philosophically he found this picture of a homogeneous, isotropic, static universe very attractive.

It turned out, however, that the equations of GR have no solutions with these properties. The problem is that masses distributed in the universe are pulled together by gravity and refuse to stay at rest. The theory seemed to suggest that the universe could not be static. But the preconception of an eternal, immutable universe was too deeply rooted. Reluctantly, Einstein concluded that the equations of GR had to be modified, by adding an extra term, to allow for the existence of a static world.

The effect of the new term was to endow the vacuum—that is, empty space—with energy and pressure. This may sound crazy, but we know that Einstein was not afraid of making counter-intuitive assumptions and following them to their logical conclusion. According to the modified equations, the energy density of the vacuum ρ_v is constant everywhere; Einstein called it the *cosmological constant*. The vacuum pressure P_v is related to the vacuum energy density ρ_v simply as

$$P_v = -\rho_v \tag{5.1}$$

Therefore, if ρ_v is positive, the pressure is negative. (We give a derivation of Eq. (5.1) from the work-energy relation at the end of this section.)

What does it mean for the pressure to be negative? The usual, positive pressure is an outward-pushing force, like the pressure of air in a balloon. Negative pressure is what we ordinarily call *tension*. It pulls inward, like the tension in a stretched piece of rubber. So, if the vacuum has tension, why does it not suck itself in and shrink? The reason is that in order to produce a force you need a *difference* in pressure: a balloon will expand if you increase the pressure inside it, but there will be no effect if the exterior pressure is increased by the same amount. The vacuum pressure is the same everywhere, and thus we should not expect any shrinkage (or expansion). The energy of the vacuum is equally elusive. There is no way to extract this energy; unfortunately we cannot solve the world's energy crisis by harnessing energy from empty space. The energy and pressure of the vacuum are thus completely unobservable—except for their gravitational effects.

The force of gravity in GR depends both on the energy (or mass) density, ρ, and the pressure, P. It is proportional to

$$\rho + 3P. \tag{5.2}$$

For ordinary matter, pressure is negligible, so we are used to thinking about gravity as being dependent only on mass. However, for the vacuum, the pressure has the same magnitude as the energy density (see Eq. (5.1)), and we find that the gravitational force of the vacuum is proportional to

$$\rho_v + 3P_v = -2\rho_v. \tag{5.3}$$

The negative sign here (in contrast to the positive sign for regular matter) indicates that the gravity of the vacuum is *repulsive*.

Einstein realized that by adding a cosmological constant to his equations, he could balance the gravitational attraction of matter with the gravitational repulsion of the vacuum. All he needed was a matter density with $\rho_m = 2\rho_v$, to perfectly balance the gravitational effect of the vacuum given in Eq. (5.3). He thereby obtained a solution that describes a static universe. This solution has a closed, spherical geometry, with the radius determined by the matter density. For the density given by recent measurements, the corresponding circumference is about 100 billion light years.

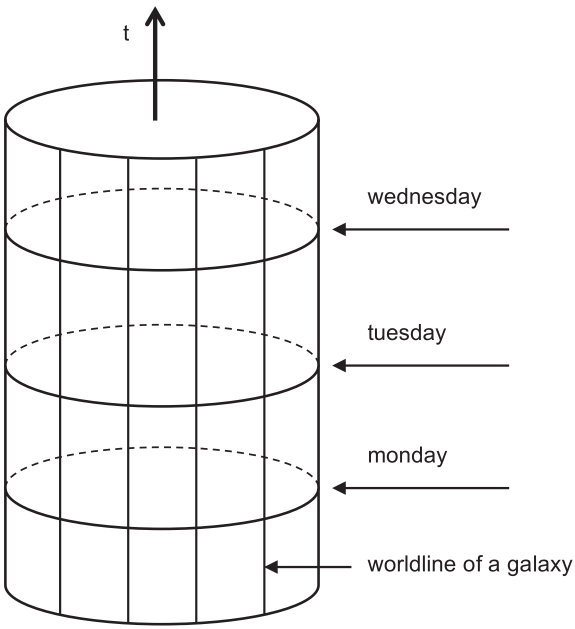

Fig. 5.2 Spacetime diagram of Einstein's static universe. Horizontal *circles* represent momentary snapshots of the universe. Two out of the three spatial dimensions are not shown

The spacetime of Einstein's static universe is illustrated in Fig. 5.2, with two out of the three spatial dimensions suppressed. It looks like the surface of a cylinder embedded in a 3-dimensional space, but only points on the surface belong to the spacetime. Time runs in the vertical direction, and horizontal slices give "snapshots" of the universe at different moments of time. In the figure these slices are circles, but in the four-dimensional spacetime the slices would be three-dimensional spherical spaces. The vertical straight lines are the worldlines of galaxies. In this universe, nothing changes with time, so all snapshots are identical and the positions of the galaxies do not change.

5.2 Problems with a Static Universe

Despite its philosophical appeal, it turns out that Einstein's static cosmological model is not acceptable. To see why, think about what would happen to the matter density, ρ_m, and the vacuum energy density, ρ_v, if the radius of the universe were slightly decreased. It doesn't take an Einstein to realize that

ρ_m would increase and ρ_v, by its very definition, would remain constant. This causes the balance scales to tip in the gravitational tug-of-war between the attraction of matter and the repulsion of the vacuum. Attraction prevails and the universe begins to contract. As it contracts, the matter density is further increased, so the contraction accelerates.

Similarly, we could ask: What would happen if the radius of the universe were increased slightly? In this case, ρ_m decreases, and the gravitational repulsion of the vacuum will win, causing the universe to expand ad infinitum. Small fluctuations in the radius of the universe cannot be avoided, and thus Einstein's universe cannot remain static for an infinite time.

Another problem with the idea of an eternal universe is that it is in conflict with one of the most universal laws of Nature—the second law of thermodynamics. This law states that an isolated physical system evolves from more ordered to more disordered states.[1] A gust of wind will lift papers from your desk and scatter them randomly over the floor, but you never see the wind picking up papers from the floor and organizing them neatly on the desk. A spontaneous ordering of this kind is not impossible in principle, but it is so unlikely that it is never seen to occur. A book sliding along the floor comes to a halt due to friction, and the energy of its directed, ordered motion turns into heat, that is, into the energy of the disordered motion of molecules. The inverse process would be for the book to cool down and start moving along the floor. This is forbidden by the second law of thermodynamics.

A mathematical measure of disorder is called *entropy*: the more entropy an object has, the more disordered it is. The second law says that the entropy of an isolated system can only increase. The evolution from ordered to more disordered states leads eventually to the state of maximum entropy, known as *thermal equilibrium*. In this state, all ordered motion ceases, all energy is converted into heat, and a uniform temperature is established throughout the system.

The universe can be regarded as an isolated system (since there is nothing outside of it). Therefore, if it existed forever, thermal equilibrium would have already been reached. The stars would have completely burnt out, cooling to the same temperature as interstellar space, and no life would be possible.[2] But this is not what is observed, so the universe could not have existed forever.

[1]In case you are wondering, the first law of thermodynamics is just a statement of energy conservation, generalized to include thermal processes. It states that the total energy of an isolated system, including its heat energy, is conserved.

[2]This bleak prediction was publicized by the German physicist Hermann von Helmholtz. He called it the "heat death" of the universe.

There is a caveat though. The Austrian physicist Ludwig Boltzmann realized that even in thermal equilibrium, spontaneous reductions of disorder occasionally happen by chance. They are called *thermal fluctuations*. So in order to reconcile the second law of thermodynamics with a universe that has existed forever, and an observable universe that is not in thermal equilibrium, we would have to conclude that we are living in a huge thermal fluctuation.

You may be concerned that such a huge fluctuation is extremely improbable. True. But if life can only exist in ordered parts of the universe, one can argue that this explains why we are observing such an incredibly rare event. Yet if we take this approach, we are still at a loss to explain why we don't find ourselves in a much smaller, and much more probable, fluctuation. It would suffice to turn chaos into order on the scale of the Solar System as opposed to the vastly larger scale of the observable universe.

Derivation of Eq. (5.1)

Consider a chamber of volume V filled with a vacuum of energy density ρ_V and pressure P_V. The volume of the chamber can be varied by moving a piston, as shown in Fig. 5.3. The total energy of the vacuum is

$$E = \rho_V V, \tag{5.4}$$

and the force it exerts on the piston is

$$F = AP_V, \tag{5.5}$$

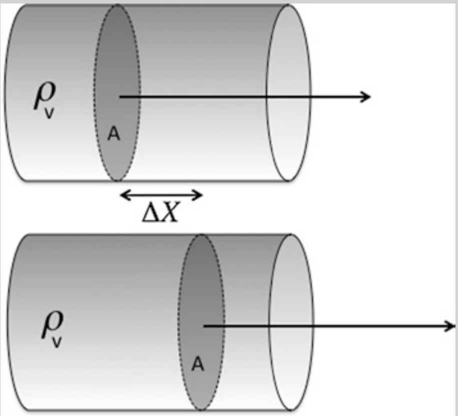

Fig. 5.3 Changing the volume of a chamber filled with constant energy density

where A is the surface area of the piston. (Recall that pressure is the force per unit area, $P = F/A$.) Suppose the piston is moved outwards by amount Δx, so that the volume is increased by

$$\Delta V = A\Delta x. \tag{5.6}$$

You may remember from elementary physics that the resulting change in the energy is

$$\Delta E = -\Delta W, \tag{5.7}$$

where ΔW is the work, which is defined as

$$\Delta W = F\Delta x. \tag{5.8}$$

(The work is positive if the force is in the direction of motion of the piston and negative otherwise.)

Thus, using Eqs. (5.7), (5.8), (5.5), and (5.6), we can show that the change in the energy of the vacuum is

$$\Delta E = -F\Delta x = -P_V A\Delta x = -P_V \Delta V. \tag{5.9}$$

Using $\Delta E = -P_V\Delta V$ (from Eq. (5.9)), and using $\Delta E = \rho_V \Delta V$ (from Eq. (5.4)), we find that the pressure of the vacuum is related to its energy density as $P_V = -\rho_V$.

5.3 Friedmann's Expanding Universe

The next breakthrough development in cosmology occurred in a rather unlikely place—the Soviet Petrograd, devastated by war and the Russian revolution. It took several years for Einstein's papers on GR to reach Russia. Once they got there, the young mathematician Alexander Friedmann voraciously studied the theory, focusing on what he thought was its central problem—the structure of the universe as a whole. He adopted Einstein's assumptions that the universe was homogeneous and isotropic and that it had a closed spherical geometry. Then he took a radical step: he did not require that the universe is static (Fig. 5.4).

With the requirement of a static universe lifted, Friedmann found that Einstein's equations did have a solution. The solution describes a spherical universe with a time-varying radius and mass density. It starts with zero radius, expands, comes to a halt, and then contracts back to size zero. If you were an observer living in some galaxy in such a universe, then during the expansion phase, you would see all the other galaxies moving away from your galaxy, whilst during contraction all galaxies would approach you. It might seem that you are located at some special cosmic center, but observers in all the other galaxies would see the same thing.

Fig. 5.4 The Russian mathematician Alexander Friedmann (1888–1925) was the first to find time-dependent solutions of Einstein's equations, describing an evolving cosmos. During World War I, while Einstein was completing his general theory of relativity, Friedmann served as a bomber pilot in the Russian air force. He was awarded a George Cross for bravery. Apart from his work in cosmology, Friedmann did groundbreaking research in hydrodynamics and meteorology. He died of typhoid fever at the age of 37

To understand how this is possible consider the surface of a balloon, which is a good 2D analogy to a 3D closed spherical geometry. Imagine that small dots, representing galaxies, are painted on the balloon (see Fig. 5.5). As the balloon is inflated, all the dots move away from each other as their relative distances increase. Conversely, as the balloon is deflated, all the dots get closer to each other. It doesn't matter which dot we focus on, the view is the same.

One limitation of this 2D analogy is that when a balloon expands, it expands into the volume of air surrounding it. So, what does Friedmann's universe expand into? Nothing. In the analogy, the surface of the balloon is all the 2D space there is. The amount of space (the area) grows as the balloon expands—but there is nothing outside or inside the surface. Similarly, the total volume of a 3D Friedmann universe grows during the expansion and decreases during the contraction.

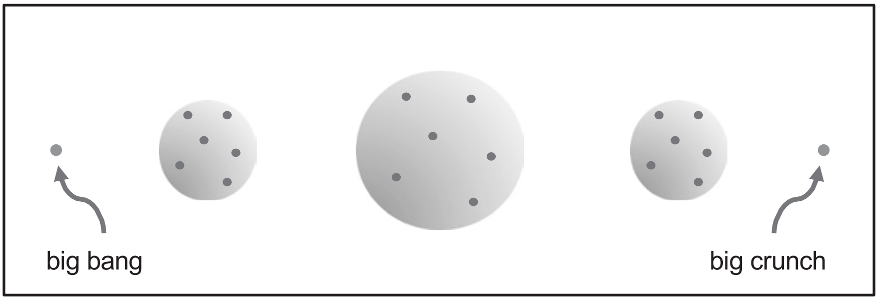

Fig. 5.5 The universe begins as a point and expands until gravity finally halts the expansion, and the universe collapses back to a point

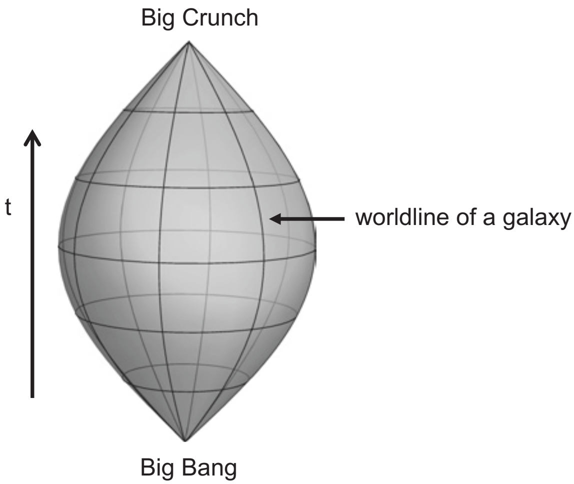

Fig. 5.6 Spacetime diagram of a closed universe. Horizontal *circles* are momentary snapshots of the universe, and the "meridian" lines are worldlines of galaxies

The history of an evolving closed universe is encapsulated in the spacetime diagram in Fig. 5.6. Here, time runs from the bottom up and horizontal circles represent instantaneous snapshots of the universe (with two spatial dimensions suppressed). The initial and final moments were later named, somewhat disrespectfully, the *big bang* and the *big crunch*. At these moments, all matter is compressed into an infinitesimal volume (a single point), so the density is infinite. This makes Einstein's equations mathematically ill defined, so the spacetime cannot be extended beyond these points. Such points are called spacetime *singularities*.

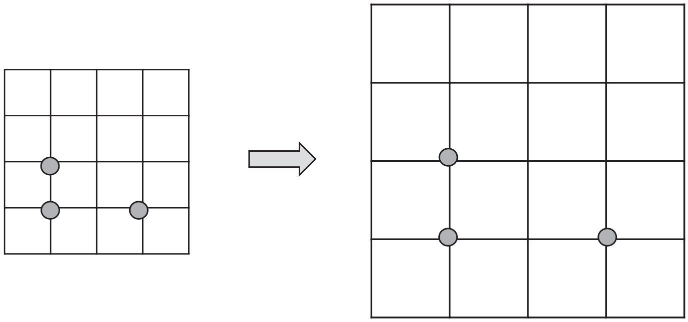

Fig. 5.7 2D stretched rubber sheet with galaxies represented by *circles*

According to Friedmann's solution, shortly after the big bang, the expansion of the universe is very fast. Then it is slowed down by the gravitational attraction between galaxies, and eventually comes to a halt, followed by contraction. This is similar to the motion of a projectile launched vertically upwards. The projectile is slowed down by gravity until it reaches some maximum height and then falls back to the ground. The greater the initial velocity, the higher it will go. Similarly, a Friedmann universe will expand to a larger radius if the initial expansion rate is increased.

Friedmann presented his solution in a paper that was published in 1922 in a German physics journal. Two years later, he published a follow-up paper describing an infinite (open) homogeneous and isotropic universe with hyperbolic geometry. Once again, he found that such a universe expands from a singularity of infinite matter density. The expansion slows down initially but it never stops completely, with galaxies approaching constant recession speeds at late times. This is analogous to a projectile launched at a speed exceeding the escape velocity (see Chap. 2). The gravitational pull of the Earth is not strong enough to turn it around, and the projectile permanently leaves the Earth.

The borderline case between open and closed solutions is a "flat" universe, having Euclidean geometry. Such a universe expands forever, but at ever decreasing speed, like a projectile launched at exactly the escape velocity.[3] A 2D analogue for an expanding flat universe is a flat rubber sheet that is being uniformly stretched in both directions. The distances between all "galaxies" are then stretched by the same factor (see Fig. 5.7). The sheet can be arbitrarily large, and we can imagine it to be infinite. When we say that a flat universe has expanded by a certain factor, what we mean is that distances between all galaxies have increased by that factor.

[3]Friedmann did not consider this borderline case. It was later studied by Einstein and Willem de Sitter.

Friedmann did not give preference to closed or open universe models. He wrote: "The available data is completely insufficient for any numerical estimates to find out what kind of universe is ours." Sadly, Friedmann died in 1925, before his papers had attracted much attention. The Belgian priest Georges Lemaitre rediscovered the expanding universe models in 1927, but his work also passed unnoticed. All this changed in 1929, when Edwin Hubble made what was arguably the most unexpected discovery in the history of science: he observed that the universe is indeed expanding! Friedmann and Lemaitre were immortalized (Fig. 5.8).

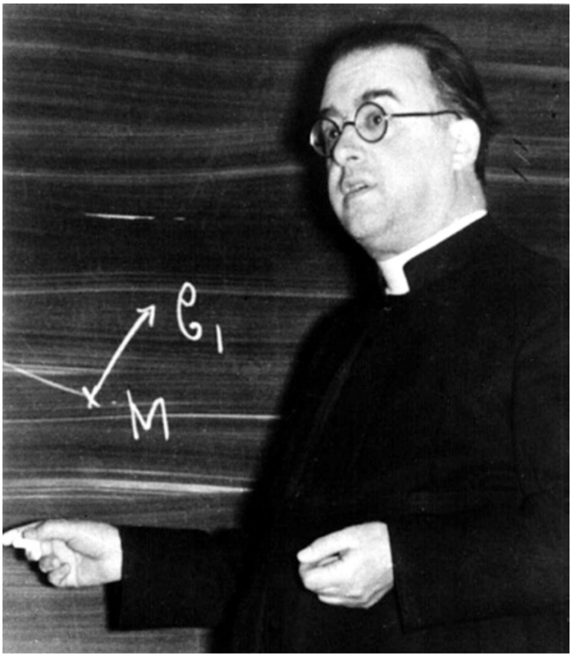

Fig. 5.8 Georges Lemaître (1894–1966) interrupted his undergraduate studies to serve in the Belgian army during World War I. After the war he went back to university and earned a Ph.D. in mathematics in 1920. He then went on to study for the priesthood, becoming ordained in 1923. By this time Lemaitre developed an interest in astronomy, which he pursued at Cambridge, Harvard, and then at MIT, where he earned his second Ph.D. In his Ph.D. thesis Lemaitre rediscovered Friedmann's solutions of Einstein's equations describing an expanding universe. He also showed that recession speeds of galaxies in such a universe should obey what is now known as the Hubble law—two years before Hubble's discovery. Lemaitre explained his ideas to Einstein at a conference in Brussels in 1927—to which Einstein replied: "Your calculations are correct, but your grasp of physics is abominable." A few years later Einstein changed his mind.
Being both a Catholic priest and a renowned scientist, Lemaitre saw no conflict between science and religion. He believed that religion should keep to the spiritual world, leaving the material world for science.

As for Einstein, he reportedly quipped that adding the cosmological constant to his equations "was the greatest blunder of my life". But even though the cosmological constant fell out of favor after the expansion of the universe was observed, it has since returned to the forefront of physics research, and we shall have much more to say about it in the coming chapters.

Summary

As soon as Einstein had completed the general theory of relativity, he applied it to the universe as a whole. Like Newton, Einstein believed that the universe was static and eternal, but he soon discovered that his theory did not admit such solutions. He then added an extra term to his equations, the so-called cosmological term, which endowed the vacuum with a nonzero (positive) energy density. According to general relativity, the vacuum then produces a repulsive gravitational force, which can balance the attractive gravity of matter. The modified equations had a static solution, describing a closed, spherical universe, but this model was seriously flawed. It was unstable to small perturbations and contradicted one of the most fundamental laws of Nature—the second law of thermodynamics.

In the meantime, the Russian mathematician Alexander Friedmann found dynamical solutions of Einstein's equations describing evolving universes that expand from a singular state of infinite density. His closed geometry solution describes a finite universe that starts out expanding rapidly, slows down, and eventually turns around and starts to collapse. The open geometry solution describes an infinite universe that starts out expanding rapidly, and although the expansion slows down, it never stops completely. Flat expanding universes are the marginal case, between open and closed. They are infinite and galaxies approach a recession speed of zero.

Questions

1. When Einstein first applied GR to the universe as a whole he assumed that the universe is homogeneous and isotropic. This is known as the "cosmological principle". Is this principle consistent with a universe which has a center or an edge?
2. Is the distribution of galaxies in Fig. 5.9 homogeneous? Is it isotropic about the central galaxy? Is it isotropic about any other galaxy?
3. True or false: If the universe is isotropic about every galaxy, it must also be homogeneous.
4. Why did Einstein add a cosmological constant to his equations of GR?

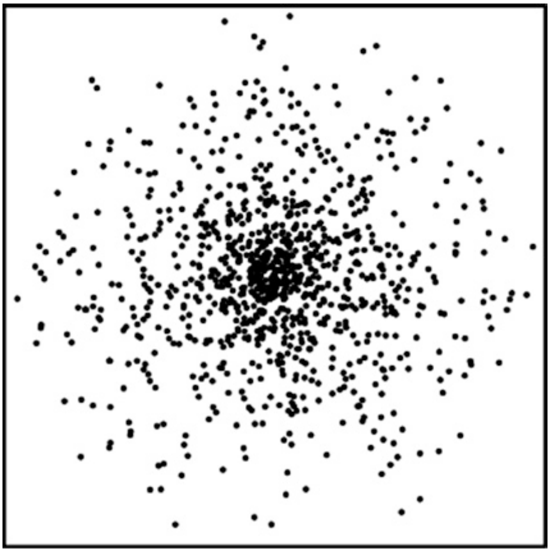

Fig. 5.9 A distribution of galaxies

5. Einstein included a positive cosmological constant ($\rho_v > 0$) in his equations. What would have happened if he had added a negative cosmological constant?
6. Einstein's cosmological constant endows the vacuum with negative pressure. Negative pressure acts like tension in a piece of rubber. So why doesn't the universe suck itself in? What effect does negative pressure have on the expansion rate of the universe?
7. What do we mean when we say Einstein's static model of the universe is unstable?
8. Is the following a correct statement of the second law of thermodynamics: "Any physical system evolves from more ordered to more disordered states"? If not, why not?
9. Why is Einstein's model of the universe in conflict with the second law of thermodynamics?
10. In a spacetime diagram for a static Einstein universe, like the one in Fig. 5.2, sketch the worldline of a flash of light emitted from some galaxy, which runs around the universe and returns to the same galaxy.
11. For a Friedmann closed universe, what is the density and radius of the universe at $t = 0$, the time of the big bang? Are the equations of general relativity valid at $t = 0$?
12. What is a spacetime singularity?

13. We used two-dimensional balloon and rubber sheet analogies to visualize closed and flat expanding universe models. Can a similar visualization be set up for an open, hyperbolic universe?
14. Is it possible to distinguish inertial motion from rest in Einstein's universe? In other words, is there any special class of inertial observers in such a universe, which can be characterized as being at rest?
15. Consider two twins who live in Einstein's static closed universe. One of the twins sets out in a rocket and heads away from his sibling at near light speed. Eventually the travelling twin returns to where he started due to the curvature of space. As he passes his twin and looks at him, which of them is older? (Note they both have maintained only inertial motion).