This curricular companion (“C5” for short) is developed in eight modules, with the idea that each module could support one of eight linked lectures on cosmology in an “Astro 101” course for non-majors. The modules are designed to supplement textbook material and to provide useful tips to instructors. They are not stand-alone lecture outlines. Most offer connections to more detailed information, for those interested in delving deeper. The modules devoted to new discoveries in cosmology, less well incorporated into standard texts, are longer. As currently structured, these modules would fit best towards the end of a one-semester or one-year survey of astronomy. The starting assumption of C5 is that students will already have been introduced to the observable properties of galaxies, including their sizes, and to spectral lines. If not, the latter topic could be taught as part of lecture 1. The only other knowledge required of students is an understanding of gravity as an attractive force.

A secondary aim of Modules 1, 2, 4 and 8 is to introduce students to graphs, and their power to represent and transmit complex ideas. Likewise, Modules 6 and 7 can serve as a refresher course on geometry. While some knowledge of geometrical results (such as the area of sphere \(A = 4\pi r^2\)) is assumed, the math level is algebra-only. Some suggestions for in-class thought exercises, demonstrations or calculations are embedded in the text [like this, in square brackets].

Lecture/Module 1  
Evidence for the Expansion of the Universe

We begin with three crucial properties of the Universe on a large* scale:

1.) It is isotropic (the same in all directions). Best evidence: the cosmic microwave background, or heat left over from the Big Bang, is the same temperature to ~0.01% in all directions (see Module 7 below)

2.) It is homogeneous (the same density and composition everywhere). One piece of evidence comes from counts of galaxies.

There is also a philosophical argument: if we are not special, and the Universe is isotropic for all its possible inhabitants, then it must also be homogeneous [an interesting claim to get your students to mull over].

3.) It is uniformly expanding, as shown by Edwin Hubble in the late 1920’s (see his The Realm of the Nebulae, 1936).

More details on the evidence for expansion: Vesto Slipher and Hubble measured the spectra of distant galaxies in an effort to understand their nature. Weak spectral lines were observed, the same lines seen in the spectra of most stars. But Hubble detected a systematic shift in the wavelengths of spectral lines of distant galaxies: they are in general slightly shifted to longer wavelengths. Since a shift to longer wavelengths is a shift towards the red end of the

* Certainly larger than the size of galaxy (about \(10^5\) light years); evidence suggests that scales of order \(10^8\) light years are big enough.
spectrum, the observed wavelength shift is called “redshift.” The fractional change in the wavelength is such that the change in wavelength $\Delta \lambda$ divided by the emitted wavelength is a constant number for all of the lines in the spectrum of a given galaxy. We write that constant as $z = \Delta \lambda / \lambda$, the proportional increase in the wavelength. Just such a shift is expected from the Doppler phenomenon, the change in the wavelength of light or sound emitted by a moving object. [For students who have not been exposed to the Doppler shift before, explain it in terms of the diagram below.]

![Figure 4.21](image)

**Fig. 1.** Explaining the Doppler shift. [from 21st Century Astronomy; Bennett et al.] Note that light from a receding object is redshifted.

The key discovery pointing to the expansion of the Universe was Hubble’s finding that the fractional shift in wavelength, $z$, seemed to increase linearly with the distance to the galaxy being studied. His distance measurements were very rough; for instance, to determine the relative distance to galaxies he could use the fact that more distant galaxies appear smaller to us. For galaxies that were not too distant, he had other and more precise measures of distance available based on the properties of individual variable stars in the galaxies. [Here’s an opportunity to dig down and discuss the properties of Cepheid variables if not already covered.] The data, while rough, all agree: $z \propto d$.

Shown in the figure below is what Hubble found, presented in graphical form displaying $z$ along the vertical axis, and distance along the horizontal axis. While there is lots of scatter in the data, the linear or proportional increase of $z$ with distance is clear: $z \propto d$. Now we make use of another linear dependence: Providing the speed, $v$, of a moving object is not too great, the Doppler shift is also a linear effect, $\Delta \lambda / \lambda = v/c$, where $c$ is the speed of light. Putting these two together, Hubble had shown that $v \propto d$.

The constant of proportionality is written as $H_0$, and called the Hubble constant: thus, we have $v = H_0 d$. Present observations favor $H_0 = 70$ km/sec for every Mpc of distance (or ~20 km/sec per million light years).
Since \( v \) is rate of change of distance with time, it is useful to rewrite this as \( \Delta d/\Delta t \propto d \) or \( \Delta d/\Delta t = H_0d \) (Hubble’s law).

\[ v = cz \]

**Fig. 2.** Hubble’s (rather shaky) evidence for linear expansion with \( v \propto d \), taken from *The Realm of the Nebulae* (1936).

A side remark: linear expansion, with \( v \propto d \), and *only* linear expansion, preserves homogeneity [another interesting conclusion for your students to ponder.] Hence we can call this *uniform expansion*.

[Demo 1: Consider a set of objects, like galaxies, homogeneously distributed (evenly spaced) in a one dimensional Universe.

Pick any two points—e.g., 1 and 2. Now allow this “toy” Universe to expand long enough until the distance has changed by 10% (so \( \Delta d = 0.1 \) d).

Repeat the same calculation for all other distances, 1-3, 2-3, …, and replot the new positions of the objects. They will again be evenly spaced.

Repeat the process with a *non-linear* expansion law, like \( \Delta d/\Delta t \propto d^2 \)—the resulting distribution of points after a cycle of expansion will be non-uniform even if you started with an evenly spaced set of objects.]

[Demo 2: Model uniform expansion using a Slinky, a set of paperclips attached to a long rubber band or, for a three-dimensional analogy, the image of raisins in bread puffing up as it is being baked. Or use these free interactive online demonstrations: [http://demonstrations.wolfram.com/TheExpandingUniverse/](http://demonstrations.wolfram.com/TheExpandingUniverse/) and [http://carma.astro.umd.edu/AWE/deploy/Hubble.html](http://carma.astro.umd.edu/AWE/deploy/Hubble.html) ]
Lecture/Module 2  The Scale Factor and the Big Bang

As the Slinky model illustrates, in a given interval of time, the increase in distance between two links of a Slinky—or between two galaxies in the Universe—is proportional to the initial distance between them \(\Delta d \propto d\). Thus instead of talking about how far galaxy A is from galaxy B, and how that distance changes, we can talk instead about a scale factor, \(a\). [This is a very hard concept to get across, but worth the effort.] Any distance increases by the same factor as the Universe (uniformly) expands. All distances increase by the same factor in a uniformly expanding Universe with \(\Delta d/\Delta t = H_0 d\), since \(\Delta d \propto d\).

That is why it makes sense to define a scale factor \(a\) to describe how (all) distances change with time. If \(a\) doubles, all distances double (like doubling a recipe written for 4, if you’re cooking for 8).

So write the distance between any two galaxies in the Universe as \(d_0\) where \(d_0\) is the present distance (so \(a\) is taken to be \(= 1\) now). In the future, given expansion, the distance will be bigger, so \(a(t > t_0) > 1\) later. In the past, the distances were smaller, so \(a < 1\).

The entire expansion history of the Universe is encoded in the function \(a(t)\): how the scale factor changes with time: \(a(t)\) is in effect the history of the expanding Universe.

The same curve \(a(t)\) describes at any time. the distance between us and a distant galaxy, or the distance between two remote galaxies picked at random. All the rest of modern cosmology is simply determining what path the Universe actually follows in this graph of \(a\) versus \(t\).

Now let’s start by graphing what we already know. Express and label the present moment as \(t_0\).

\[\text{Fig. 3. What we know so far: the scale factor, } a, \text{ is increasing with time at present } (t_0).\]

Recall that we took the current value of the scale factor \(a\) to be unity: \(a(t_0) = 1\). Since the Universe is expanding, we know the scale factor will be larger in the future, and we may infer that it was smaller in the past – hence the small slanting line segment at \(t_0\). [Here is a good time to start to drive home the point that this single graph, no matter how \(a(t)\) varies, summarizes the entire history of the Universe: graphs may be abstract, but they are very powerful.]
Introducing the scale factor also allows us to jettison the Doppler model. Instead of talking about galaxies rushing away from us, so that their light is Doppler shifted when we study them, it is more correct and more useful to think of galaxies as sitting still in an expanding space. As the Universe expands, so do the wavelengths of light traveling through it. Light emitted from a distant galaxy was necessarily emitted at some earlier time, given the finite speed of light. At earlier times, we have seen, the scale factor was smaller, so the wavelengths of light when the light was emitted were smaller than they now appear. The wavelengths have increased (that is, redshifted) by the later time at which the light reaches our telescopes, just as Hubble found. Thus

\[
\frac{\lambda_{\text{obs}}}{a(t)} = \lambda_{\text{obs}} = \lambda_{\text{emit}}/a(t)
\]

But we can also relate the wavelength of light we see now \( \lambda_0 \) to the wavelength at which it was emitted as: \( \lambda_0 = \lambda_e + \Delta \lambda = (1 + z) \lambda_e \), so that we find the connection between redshift and scale factor: \( a = (z + 1)^{-1} \). We no longer need the Doppler description or the restriction that \( v << c \). If we see a galaxy whose spectral lines are redshifted by as much as \( z = 1 \), we are looking back to a time when the scale factor of the Universe was only one half its present value. [Get students to work out what the density of that time would be compared to the present density.] The largest redshifts observed for quasars or galaxies now exceeds 7. When the light from these systems was emitted, the Universe was over 300 times its present density!

[ Next, as a class exercise, try to get groups of students to work out how the Universe would expand if no forces were acting on it. We know it is getting bigger with time, so that the line representing relation \( a(t) \) is sloping upwards to the right. Guide them into the view that the expansion would be a straight line if no forces act, and as a consequence that straight line must at some point in the past cut the horizontal axis, so that \( a = 0 \) and all distances in the Universe go to zero.]

\[\text{Fig. 4. “Running the film of expansion backwards.”} \]
\[\text{Fig. 5. Start time at the moment } a = 0 \text{ (the Big Bang)}\]

Spend some time on what it means to have \( a = 0 \): all distances go to zero. Having already introduced the way density depends on \( a \), point out that the moment \( a = 0 \) implies an infinitely dense Universe. At all times later than the particular moment when \( a = 0 \), the Universe appears to
be expanding away from an initial state of infinite density—hence the phrase “Big Bang” is used to describe the moment when $a = 0$.

Now we can make another reasonable assumption. Just as we fix the scale factor to be unity now, it makes sense to start counting time at the moment of formally infinite density; that is, the moment of the Big Bang. Hence we re-label the axes of the graph of Fig. 4, and select $t = 0$ at the point where $a = 0$. Thus $t_0$ takes on a new meaning; not only does it represent the present, it also is the age of the Universe, the time elapsed since the Big Bang.

Lecture/Module 3 Okay, Is There Any Evidence for a Big Bang?

It is helpful to start by spending some time exploring what happens when matter is expanded or compressed. Expanding matter—such as gas rushing out of an aerosol can—cools down. On the other hand, compressing matter—think of pumping air into a bike tire—heats it up. The same is true in the Universe. As it expands, it cools. Perhaps more importantly, if we run the film of expansion backwards and look back from $t_0$ towards the Big Bang as $t \rightarrow 0$, we expect the Universe to have been hotter in the past. [The same conclusion can be reached by thinking again about the wavelengths of photons; when the scale factor was smaller, the photon wavelengths were shorter, the corresponding frequencies and energies were higher, and this corresponds to a higher temperature.] The fundamental fact that the Universe may have started hot as well as dense was first pointed out in a series of papers by Gamow, Alpher and Hermann in the 1940’s. They even made rough calculations of what the present temperature of the Universe would be, arriving at a figure of about 5 K, or 9° F, above absolute zero. We can actually measure the current temperature of the Universe, and we obtain a number not far from the rough estimate by Gamow, Alpher and Hermann: the current best measurement is $2.725 \pm 0.001$ K, the value provided by the FIRAS experiment on NASA’s COBE satellite.* The detection of heat left over from the Big Bang is one strong piece of evidence that the Universe did in fact go through a stage of high temperature and high density—a Big Bang.

The argument given above may suffice, but you may also want to go into somewhat more detail and talk about the particular spectrum of the heat left over from the Big Bang, since a blackbody spectrum can be produced only by matter that is dense as well as hot. I will assume here that a blackbody spectrum has been mentioned earlier in the course, probably in connection with stellar emission. If so, remind students that the spectrum has a very characteristic shape, determined only by the temperature of the emitting material. In addition, this kind of spectrum can be emitted only from dense materials, whether solid, liquid or gas. We have already argued that at early times the Universe was both dense and hot. Thus it is reasonable to expect a thermal, blackbody spectrum to emerge. As the Universe cools down, the temperature drops, but the shape of the spectrum remains the same. As $T$ drops, the wavelength at peak emission shifts to longer wavelengths, reaching $\sim 2$ mm for the current temperature of $2.725$ K.

As noted above, we have known since 1964 that the Universe is full of heat left over from the Big Bang. Discovering this heat, in the form of millimeter wave or microwave photons, won the Nobel Prize for Arno Penzias and Robert Wilson several decades ago. Furthermore, observations by the COBE satellite have shown that the leftover heat radiation has a spectrum that is extremely close to blackbody. In Fig. 6, the COBE results are displayed. The little boxes represent the measurements, with $400\sigma$ error bars, and the curve that fits them so well is a
blackbody curve. The heat left over from the Big Bang does have the expected properties if the early Universe was dense as well as hot.

It is a complicated calculation (see my book *3 K: The Cosmic Microwave Background Radiation*), but it may be shown that the Universe needed to be something like $10^{19}$ times denser than it is now to ensure the high density needed to produce a thermal, blackbody spectrum. At that density, its temperature was $\sim 5 \times 10^6$ K, and the Universe was only a month or so old!

We will return to studies of the heat left over from the Big Bang in Module 7. Either then or at this point introduce the current nomenclature for this heat, the *cosmic microwave background* (CMB). It is certainly cosmic, and the wavelength at which it is easily observed falls in the microwave part of the electromagnetic spectrum. In Module 7, we will see that it truly is a background, arising at very large distances from the observer, so that the third word in the usual term for this left-over heat is appropriate.

![Graph of blackbody curve](image)

**Fig. 6.** Does the cosmic microwave background (measurements shown as points with 400σ error bars) have a blackbody spectrum, as expected? Yes! These are results from the FIRAS instrument on NASA’s COBE* mission.

Having established that the Universe did indeed have a Big Bang (or at least was initially very dense and hot), it is time to return to the plots of scale factor vs. time and to begin to think about how graphs like Figs. 4 & 5 might evolve in the future as well as the past.

* See [lambda.gsfc.nasa.gov/product/cobe/firas_overview.cfm](http://lambda.gsfc.nasa.gov/product/cobe/firas_overview.cfm) for the gory details. The Wikipedia article on COBE is quite weak. A better introduction to the CMB is provided by Wayne Hu’s website [background.uchicago.edu/~whu/beginners/introduction.html](http://background.uchicago.edu/~whu/beginners/introduction.html)
Lecture/Module 4 The History of the Universe in a Single Graph

Remind students of Fig. 5—how the Universe expands if no forces act. That graph describes the entire history of the Universe. When no forces act, there is constant expansion. More complex histories are possible; but all can be represented in a graph of scale factor (“size”) versus time, as suggested by various curves (imaginary histories) in cartoon form in Fig. 7.

![Graph](image)

Fig. 7. Fanciful, possible histories of the Universe; \( t_0 \) indicates the present moment.

Remember that Fig. 5 represents a special case: no forces act. [Get students to discuss how realistic that assumption is. Guide them to consider the effect of gravity.]

The role of gravity. Gravity is always attractive, pulling masses (like galaxies) together. Therefore gravity always slows expansion. [Get students to make sense of the assertion that a gradually slowing expansion introduces curvature in the plot of \( a(t) \), in the sense sketched.]

Point out that including gravity can’t eliminate a Big Bang, that is a moment in the past when \( a \rightarrow 0 \). But adding gravity does lower the age of the Universe, \( t_0 \). This is because the \( a(t) \) curve bends down to cut the \( t \) axis nearer \( t_0 \) (see Fig. 8B).

If gravity is strong enough, it will stop the expansion at some future time and cause the Universe to recollapse (as in Fig. 8A). [Get students to see why the expansion can only stop in the future, at a time \( t > t_0 \) reminding them of Hubble’s discovery]. If gravity is weak, it slows expansion, but not enough to cause recollapse (see Fig. 8B). Students may object that, since gravity is always attractive and the gravitational force always is present, it will slow expansion (and keep the \( a(t) \) line concave downwards). That is true. So why is unending expansion possible? Remind them that the gravitational force depends on \( 1/R^2 \), and distances are always
increasing in an expanding Universe. [For groups that have examined escape velocity, note the parallels.]

**Fig. 8.** Two possible histories of the expanding Universe.

*The role of density.* “Strong gravity” means recollapse. “Weak gravity” allows expansion forever. What determines if gravitational forces are strong or weak? The average *density* of the Universe: more matter ⇒ more gravity. Thus Fig. 8A corresponds to high density; 8B to low.

[FAQ: *Why doesn’t total mass matter, rather than density? Because we don’t know the absolute size of Universe. It is density that matters.]*

Next, work with students to understand that there must be a special value of the density—the *critical density*, Ωc—that separates the two cases. If the density happens to equal Ωc, the Universe expands more and more slowly and drifts to a halt. And if Ω > Ωc, we are in the high density, strong gravity case shown in Fig. 8A. Likewise, if Ω < Ωc, the Universe expands forever.

Point out that these futures are inescapable. If the Universe recollapses, our descendents get crushed—and/or baked, since the Universe will heat up as the matter in it is compressed. In a recollapsing Universe, the moment when the scale factor again approaches zero is called, for obvious reasons, the “Big Crunch.”

[One can show that we’ll be baked (T > 500 F) well before we get crushed (Ω > 1 gm/cm³). This can be shown (perhaps an interesting homework problem for advanced students) by noting that the temperature of the Universe was roughly 5 \times 10^6 K when its density was 10^{19} times the present density. While 10^{19} is a large factor, the current density in the Universe is so small (see below) that even increasing it by that factor would leave the density less than that of air. Thus even close to the time of the Big Crunch, when the temperature returns to 5 \times 10^6 K (surely be high enough to roast us), the density will not be high enough to crush us.]

Students may want to know whether the recollapse, in the case of high density, is a time-symmetric version of the expansion. It is, and we have just used that fact implicitly to show we will bake before being crushed. The time symmetry follows from the mathematical form of the equations derived from general relativity which specify a(t) [A clear but technical presentation, based on Newtonian Physics, rather than General Relativity, can be found in the paper by Callan, Dicke, and Peebles (*American Journal of Physics*, 33 p105, 1965).]
To repeat, low density means expansion forever, and high density means recollapse. Clearly, as noted above, there has to be one special value for the density that separates the two cases. In the box below we present a derivation of that quantity, and evaluate it in terms of \( \text{gm/cm}^3 \).

Hence DENSITY is DESTINY.

Wouldn’t it be nice to know if \( \rho > \rho_c \) or not?

---

**The critical density, \( \rho_c \).**

The following is a demonstration that the critical density just described is equal to \( 3H_0^2/8\pi G \). It is based on the idea of escape velocity, and hence can be presented if students have looked at that concept.

Consider a bunch of galaxies in a uniformly expanding Universe. Arbitrarily select any two galaxies, say 1 and 2. Now let’s consider how the distance between those two galaxies would change if the Universe expanded forever, or if it were to recollapse. In the case of expansion forever, the distance obviously becomes infinite. In the case of recollapse, the distance between galaxy 1 and galaxy 2 would increase at first and then decrease as the Universe recollapses.

If we now make the mental jump of treating galaxy 2 as a particle “escaping” from galaxy 1, the parallels with escape velocity become clear. In particular, let’s draw a sphere centered on galaxy 1 with a radius equal to the distance between galaxy 1 and galaxy 2—see the second panel of Fig. B1.

For galaxy 2 to “escape” from the sphere drawn around galaxy 1, it has to have sufficient outward velocity to allow it to overcome the gravitational attraction of the mass contained within that sphere. In particular, the escape velocity is given by

\[
v_{es} = \sqrt{GM/2R}.
\]

where \( M \) is the mass contained within the sphere of radius \( R \). In turn, given that the Universe is homogeneous, we can write \( M = 4/3\pi R^3 \rho \). But what outward velocity does galaxy 2 have? That
comes from the linear expansion encoded in Hubble’s Law: \( v = H_0 R \). Escape occurs if the outward velocity is greater than \( v_{es} \), that is, if:

\[
H_0 R > \sqrt{GM/2R}.
\]

Algebra then shows that escape occurs (so the distance between galaxies increases forever) if:

\[
\rho < \frac{3H_0^2}{8\pi G}.
\]

Thus the special, limiting case is

\[
\rho = \rho_c = \frac{3H_0^2}{8\pi G}.
\]

Since we know \( H_0 \) (see Lecture/Module 1) we can evaluate \( \rho_c \). The first step is to convert \( H_0 \) from astronomers’ units to physical units: \( 70 \text{ km/sec per Mpc} = 70 \text{ km/s divided by (3.086 \times 10^{19} \text{ km/Mpc})} \). Thus \( H_0 = 2.3 \times 10^{-18} \text{ sec}^{-1} \), the rate of increase of distances in the Universe. Yes, the Universe is expanding, but slowly! Consequently, \( \rho_c \) is going to be (perhaps surprisingly) small. Given this value of \( H_0 \), we find

\[
\rho_c = 9.7 \times 10^{-30} \text{ gm/cm}^3.
\]

Is the average density of the universe higher than this tiny value?

---

**Lecture/Module 5 Measuring Density, Directly and Indirectly**

Since DENSITY is DESTINY, wouldn’t it be nice to know if \( \rho > \rho_c \) or not? To find the density of the Universe, why can’t we simply add up the mass of all the galaxies in a big chunk of the Universe, and divide by the volume of that chunk? We can. But there are several reasons why this simple approach doesn’t work. First, we observe light, not mass. We can count galaxies, but the connection between their light and mass is varied enough to make estimates of the latter very shaky. A much stronger negative argument is the presence of Dark Matter—mass we cannot see because it neither emits nor absorbs light. Nearly 80 years ago, Fritz Zwicky showed that some form of Dark Matter was needed to hold clusters of galaxies together gravitationally. A couple of decades later, Vera Rubin showed that Dark Matter was needed inside galaxies as well as between them [see her elegant article in the 1998 *Scientific American*]. So you can’t calculate the average density of the Universe without knowing how much Dark Matter there is—and it is not possible to find this directly. Nor do we know what Dark Matter is. All we know is that the Dark Matter is more important than the luminous matter making up the visible stars and galaxies. [ Martin White provides a nice review and some useful references in: http://astro.berkeley.edu/~mwhite/darkmatter/dm.html. ]

*Note: we will soon see how much Dark Matter there is—but again not what it is. The estimate of the density of Dark Matter depends on the measurement we are about to describe. 

Fortunately, there is another approach to determining our destiny, one that can provide a crisp answer, and has. The approach is based on a fundamental property established by Einstein’s General Relativity: mass curves space. It does so radically in the vicinity of a Black Hole (a
Black Hole is a “hole” because space is so curved that there is no “outward” direction; all directions point inwards. The Sun’s mass curves space much less dramatically—and so does all the matter in the Universe.

If, on a large scale, the Universe is homogeneous (the same everywhere), then the curvature must be the same everywhere: a homogeneous Universe must have uniform curvature.

Gauss and others showed that there are only 3 classes of curvature that are uniform (“the same everywhere”). Zero curvature, or “flat,” or Euclidean space is one such case. Euclidean space has no curvature anywhere, so it is evidently uniform. But space can be uniformly curved and not flat. Consider the two-dimensional surface of a balloon. The tightly stretched rubber is a curved space, but the curvature is the same everywhere (if the balloon is spherical).

This class of uniformly curved space is called positively curved. A way to assess the sign of the curvature is to draw two perpendicular line segments on the balloon, crossing at a point; both line segments are concave, or curved inwards.

Harder to construct (or even to picture) is two-dimensional negatively curved space; here the curvature is concave in one direction, convex in the perpendicular direction. The surface of a saddle is a chunk of negatively curved 2D space—curved downwards to fit over the horse and upwards in the perpendicular plane to hold the rider’s bottom. Figure 9 shows a device made of a rubber membrane to represent a two-dimensional chunk of negatively curved space.

Thus, there are three classes of uniformly curved space. Here’s the beauty. It happens that if the density of the Universe is $> \rho_c$, cosmic space is positively curved, and if $\rho < \rho_c$, it is negatively curved. The critical dividing line of zero curvature corresponds to the critical density itself. So we do not need to measure density; we can measure curvature instead. We describe the results of this measurement in Module 7.

There is thus a direct connection between the future of the universe, density and curvature. As we will soon see, curvature is much easier to measure than the density (you don’t have to worry about that pesky and invisible Dark Matter; it acts just like visible matter to curve space).

In the following modules, we look at the geometrical basis of two tests for the curvature of the Universe, and then look at the observational evidence that tells us that the Universe is, to quite high precision, flat. And a zero-curvature Universe must have a total density equal to the critical density, whose numerical value we have already calculated.

**Lecture/Model 6  Geometrical Properties of Curved Space**

We will be looking at several geometrical properties of curved space. These include the sum of angles in a triangle; the connection between the circumference of a circle and its radius; and the connection between the area of a sphere and its radius.

All these familiar relationships change in curved space, and the amount of change depends both on the curvature of space and the size of the triangle or circle.

Let’s begin by taking a two-dimensional flat space and “curving” it, as described in the next demo. There are two things to note.
[Demo 1: Start with a large sheet of paper. Demonstrate ordinary geometrical properties like the area of a circle or the sum of angles of a triangle in this flat space. Then, with a student volunteer, bend all four corners down, producing a wrinkled box. This is a crude, but clearly not uniform, approximation of positively curved space. To make the space more uniformly curved we would need to stretch it. Next, try the same trick pulling up on two diagonal corners while the volunteer pulls the other two down (this is a distorted piece of negatively curved space—a saddle-shaped space—see Fig. 9 below). We’ve constructed a rubber membrane on a bendable frame to illustrate these properties of negatively curved space. Here we show a photo of the two authors as we demonstrate that curving space (negatively, in this case) also causes distortions and “wrinkles.”]

**Fig. 9.** The two of us demonstrating negative curvature: Nick pulls up two sides of the surface, while Bruce pulls down the other two. If you look carefully, you will see that “wrinkles” or distortions appear in the surface as it is curved.

Having played these games with an initially flat sheet of paper, guide the students to two conclusions:

1) Curving the two-dimensional space requires the use of the third dimension, that is, a higher dimension than the space you started with. Since we don’t have a handy fourth dimension, it
is hard to bend a three-dimensional space into that dimension, so we will stick with a two-dimensional model, which can be “bent” or curved into the third dimension.

2) The second crucial thing to notice is that to curve the space without wrinkles and other non-uniformities, the space must stretch. It is that stretching of space that changes the geometrical relations in uniformly curved space from their flat space values. Here is where you introduce the surface of a (hopefully spherical) balloon as an example of uniformly curved, positively curved, two-dimensional space.

Next, there is the issue of the magnitude of space curvature. The commonsense notion works: space has a bigger curvature when it is more sharply bent. More quantitatively, imagine fitting a circle to the curved space at a particular point. Figure 10 demonstrates this for the easy case of a one-dimensional curved space. Select a point on the curve. To measure curvature there, draw the “best-fit” circle, as shown by the dashed lines in Fig. 10. The smaller the radius of that circle is, the larger the curvature, $k$.

In this one-dimensional case, we can define the curvature as $k = 1/R$, with $R$ the radius of the best fit circle.

![Fig. 10. A one-dimensional curved space (a curved line), and the “best-fit” circle that allows us to define the curvature at or near a point on the curved space.](image)

In two dimensions, we need two best-fit circles, and it is reasonable to draw them so they lie in two perpendicular planes. I hasten to add that this is cutting a big mathematical corner—but it produces the right qualitative results, and makes intuitive sense. In the two-dimensional case, we elect to define $k$ as $\frac{1}{r_1 r_2}$ where $r_1$ and $r_2$ are the radii of the two perpendicular, best-fit circles. Demo 2 below provides an illustration.

[Demo 2: Back to the balloon. Let your students select an arbitrary point on the balloon, then through that point draw a small arc of a great circle on the surface. What is the radius of curvature of that small arc? Obviously, it equals the radius of the balloon. Then draw a second small arc perpendicular to the first, again through the selected point. The radius of the best-fit circle to it is also just the radius of the balloon, so $k = 1/R^2$ in this case.]

In 2D space, there is also the issue of the sign of curvature. If the two best-fit circles both lie on the same side of the surface, we have positive curvature; if on opposite sides, as in a saddle, we have negative space curvature. Clearly, the 2D surface of the balloon is an example of positively curved space.
Now we introduce an important restriction—let’s consider only spaces of uniform curvature, namely, spaces that have the same curvature everywhere. We introduce this for simplicity. We also introduce it because we are going to apply curved space notions to the Universe, and as we have seen, there is good observational evidence that the Universe, on a large scale, is homogeneous, so its space curvature must be uniform as well.

As we have already mentioned, there are only three classes of uniformly curved space. The easiest to consider is an infinite expanse of flat space. Space can also have uniform negative curvature. Remember a saddle. If the curvature of the side that fits the horse is the same as the curvature of the side that fits the rider, we have uniform, negatively curved space. To avoid the problem of edges, negatively curved space, like flat space, must be infinite if it’s uniform. In the case of positively curved space, on the other hand, we can construct a two-dimensional space of uniform curvature that is not infinite. It is the 2D surface of a 3D sphere—aka the rubber of a balloon.

At this point, it is fun to hand out balloons to the students, or small groups of students, so they can conduct their own measurements and experiments.

At some point it is worth reminding students that the rubber membrane of the balloon is the curved two-dimensional surface, and we’re interested only in the properties of that two-dimensional, curved surface. What is “inside” the balloon is not part of the space.

If the balloon is a true sphere, we have seen that the space is positive, with $k_0$ simply equal to $1/R^2$.

Now let the students investigate some geometrical properties of this positively curved 2D space. First, as is true in flat space, a line is the shortest distance between two points.

[Demo 3: Gently stretch a thread taut along the surface between two points on the surface to demonstrate what a line looks like in this 2D curved space. If there is time, get the students to guess and draw what they think might be a straight line in the curved surface. Then test their guess by using a taut thread.]
Next, we’re going to construct an equilateral triangle in this positively curved space, consisting of three such straight lines, then use it to investigate the sum of angles in a triangle.

[Demo 4 (for the students to do): Pick an arbitrary point, draw a straight line on the surface in any direction with a length equal to ¼ of the way around the surface of the balloon. Turn 90° and mark off an equal distance. Turn 90° in the same sense and again mark off an equal distance. You should arrive back where you started. And you have just constructed an equilateral triangle, the sum of whose angles is 270°: see Fig. 12 below.]

Conclusion: in positively curved space, the sum of angles of a triangle > 180°.

[Demo 5: By having them draw a really small triangle on the surface of the balloon, get students to see that a small triangle (whose sides are much less than the size of the balloon, R) would have a sum of angles much closer to 180° than a big triangle. This demonstrates that the departure from the flat space expectation depends on the size of the triangle as well as the curvature of the surface.]

[Challenge students to figure out the nature of a triangle with the largest possible sum of angles in positively curved space. It turns out to be a triangle with the sum of angles = 900°.]

Now let’s check the connection between the circumference of a circle and its radius. A circle is defined as the locus of points equidistant from a chosen point, and we will elect to draw a circle whose radius is ¼ of the way around the balloon. That radius is equivalent to the distance from pole to equator on the globe; from that it is easy to get the students to see that in this positively curved space, and for a circle of this radius, \(C = 4r\), not \(2\pi r\). The circumference of a circle in positively curved space is thus less than that in flat space. It is easy to show that the same is true for the area of a circle, and this might be a fine exercise to work out, again using the handy balloon.

In summary, in positively curved space, the sum of angles in a triangle is bigger than in flat space, and the circumference and areas of circles are smaller.

I assert that the same results apply to positively curved, three-dimensional space. In such a space, for instance, the area of a circle is still less than \(\pi r^2\), and the area of a sphere is less than \(4\pi r^2\).

It is a little harder to demonstrate these properties in uniformly but negatively curved space. A decent analog for a piece of negatively curved space of uniform curvature is a Pringles potato chip.

[Demo 6: With a magic marker, a potato chip and a bit of hand waving, try to demonstrate that the sum of angles in a triangle drawn in such a space is less than 180°. Negatively curved space has exactly the opposite behavior from positively curved space in this regard and also for the circumference and area of circles and spheres. In particular, the area of a sphere in negatively curved, uniform space is >4\(\pi r^2\) (and we will later make use of that result).]

* Hint—drawing a small triangle \(\Delta\) on the balloon actually generates two triangles. The smaller one has angles that add up to just over 180°; the rest of the surface is the second triangle.
Fig. 12. Nick Vechik demonstrating a triangle in curved space, one with three 90° angles.

**Lecture/Module 7 The Geometry of the Universe**

The cleanest test of the curvature of the Universe—and thus its future—is based on one of the geometrical properties we just developed. Suppose we draw the skinny (and isosceles) triangle shown in Fig. 13. It has two long sides of length D, and a short base of d. If d << D, and we’re willing to express the angle θ in radians, we know from the small angle formula that θ = d/D — *at least in flat space*. In positively curved space, however, the sum of angles in a triangle is larger, so we expect θ > d/D (and vice versa for negatively curved space).

![Diagram of triangle with sides D, d, and angle θ]  

**Fig. 13.** Knowing d and D allows a calculation of θ of in flat space.

*Thus, to determine the geometry of the Universe, and hence its future, all we need is a cosmic triangle resembling Fig. 13 with known values of both d and D.*
Triangles like this one, with $D$ and $d$ of known size, do exist in cosmic space, as shown below. Furthermore, the sides $D$ are very large (and recall that the differences between expected flat-space angles and those measured in curved space are helpfully larger for larger triangles).

**Detour: The surface of last scattering and fluctuations in the CMB**

In Module 3, the discovery of the cosmic microwave background (CMB) was cited as the best evidence we have for a Hot Big Bang.

Now let us ask: What do we “see” when we use radio telescopes to observe the CMB? The answer is, a sharply defined surface at a very large distance from us. This surface is called the surface of last scattering (SLS), because it marks a dividing line between regions where light scatters frequently and those that are transparent. There is an exact and useful analogy with what we see when we look at a cumulus cloud. Here, the answer is the surface of the cloud. Why?

![Fig. 14.](image)

In a cloud, light scatters off water droplets, until it reaches the surface of the cloud; this surface is what we see. In the Big Bang, light scatters from free electrons until they suddenly disappear by incorporation into atoms. Our telescopes “see” this surface of last scattering.

Between the viewer and the cloud, the atmosphere is transparent. Hence we see nothing in front of the cloud. We cannot see through the cloud because light rays are scattered by the water droplets in the cloud (there may be some absorption, but relatively little in a nice, white, cumulus cloud). What we **do** see is the surface where light breaks free, or last scatters, before traveling in a straight line to the eye. Exactly the same physics operates early in the Universe, where free electrons play the role of water droplets as scatterers. Light scatters from free electrons until they disappear. Then, without further scattering, the light is free to travel to our waiting telescopes, and this is what we see as the CMB. The time when free electrons disappear occurs abruptly;* thus there is a sharply defined surface of last scattering.

It remains to see where the free electrons come from, and why they disappear suddenly. Early in the history of the Universe, when temperatures exceeded 3000 K, all atoms (90% of them hydrogen) in the Universe are ionized. As the Universe expands and the temperature drops below 3000 K, protons combine with the free electrons to form neutral hydrogen. Once captured in atoms, the electrons no longer scatter (think of the transparency of oxygen gas, consisting of atoms with 8 electrons). That the free electrons can exist only at high temperatures ensures that

* Technical remark: the number of free electrons is governed by the Saha equation, and thus depends exponentially on the temperature, and thus on redshift $z$. 
they disappear into neutral atoms very early in the history of the Universe; that is, fairly close in time to the Hot Big Bang. Hence the SLS is very far away. The current estimate for the time when neutral atoms formed is 370,000 years after the Big Bang, and the Universe is now thought to be 13.8 billion years old. Thus the distance to the SLS is just under 14 billion light years. Let us call that that large distance D; it is evidently very large indeed. More to the point, the distance itself is known to an accuracy of something like 1%.

Next, theorists have predicted for years that this surface cannot be exactly smooth—like the surface of a cumulus cloud, it has small bumps and wiggles. There are regions of slightly higher density*, where the temperature is a bit higher, and regions that are less dense and cooler. Thus the surface is mottled or lumpy.

And it turns out that these lumps have a characteristic, physical scale. That size is set by the speed of light and the age of the Universe at the time of last scattering, 370,000 yrs. Regions bigger than 370,000 light years in size have not had time enough to reach the same value of density. Consequently, there can be lumps of a characteristic size roughly equal to this value. [For details see Hu and White in *Scientific American*, 2004.] We need also to take account of expansion from the moment of last scattering until now. When we do, we can calculate d exactly. And recall that we also know D to high precision. Thus we can calculate θ and apply our test of curvature to the Universe. If cosmic space happens to be exactly flat, the lumpiness would appear with a characteristic scale of ~1/60 radians, or just less than 1°. But in positively curved space θ would be larger (and in negatively curved space, smaller) than this value. Thus the observed size of lumpiness on the surface of last scattering, or fluctuations in the temperature of the CMB, provides a method of determining the curvature of space. The difference is displayed schematically in the lower panels of Fig. 15. This displays the lumpiness of the SLS using false color. The regions of slightly higher temperature (shown in red), like regions of lower temperature (blue), are bigger in the left-hand panel, which corresponds to positively curved space.

These panels show predictions. What does the observed lumpiness of the surface of last scattering look like? This is shown in the upper panel. These are actual measurements of ~400 deg.² of the sky carried out by a balloon-borne radio telescope called BOOMERanG (see [http://cmb.phys.cwru.edu/boomerang/press_images/cmbfacts/cmbfacts.html](http://cmb.phys.cwru.edu/boomerang/press_images/cmbfacts/cmbfacts.html)).

[ *Class activity.* Get your students to vote on the question of which panel most resembles the observations. There will be some “scatter,” but most will favor flat space. ]

You can repeat the process with the results from an entirely different experiment, MAXIMA (Fig. 16); this figure uses a different false color palette, but comes with a useful angular scale along the vertical axis. Again, we see that the characteristic scale of both hot and dark patches is ~1° (see [http://cfpa.berkeley.edu/group/cmb/](http://cfpa.berkeley.edu/group/cmb/)).

So here is the answer: the geometry of the Universe is indistinguishable from flat. More refined measurements by the *Planck* satellite released in 2013 show that any possible departures from flatness are at below the 1% level. If the Universe has flat geometry, its total density must be very close to ρc. Our future would seem to be gradually slowing expansion, forever.

* The denser regions eventually grow, by the action of gravity alone, into the large-scale structure we see in the Universe today.
Fig. 15. The three lower panels show how the lumpiness of the SLS would appear in, from right to left, $k < 0$ negatively curved space, flat space and $k > 0$ positively curved space. As expected, the angular scale of the lumps is larger in positively curved space. The top panel shows actual measurements of the temperature anisotropies in the CMB, or lumpiness on the SLS, as measured by the BOOMERanG experiment. We see it most closely matches the flat space case.

Fig. 16. A map of the CMB “lumps” made by a second experiment, MAXIMA. Again, the angular scale of the hot (and cold) spots is just less than $1^\circ$. 
Very well, it seems that the Universe has no measurable curvature. But why? In principle, the Universe could have had a curvature of, say, $10^2$ m$^{-2}$, in which case the Universe would be the size of a basketball. Why the Universe is in fact as big it is was one of the questions that troubled some theorists, among them Alan Guth and Andrei Linde. in the early 1980’s. Another was where the lumpiness observed on the SLS came from. Both questions were answered by a simple suggestion: the Universe went through a brief period of explosive, exponential expansion early in its history. Since this idea arose when some economies were going through an inflationary binge, this theorized moment of ultra-rapid expansion was called “inflation.” The phrase “early in its history” is perhaps understatement – we currently think inflation happened at about $10^{30}$ sec into the history of the Universe! In a fraction of $10^{30}$ sec, the Universe blew up in size by something like a factor of $10^{50}$ or more. That extraordinary change in the scale factor had two consequences. First, suppose the pre-inflation curvature did happen to be $10^2$ m$^{-2}$. Changing the scale factor by $10^{30}$ would inflate the radius of curvature by the same factor, thus reducing the curvature to $10^{-58}$ m$^{-2}$. Not zero, but way too small to measure. Inflation flattens the Universe. The same huge increase in scale factor can also scale up tiny quantum fluctuations to astronomical scale. Inflation does not affect the amplitude of fluctuations, but it can make them $10^{30}$ larger in linear dimension, inflating them from microscopic to astronomical size.

Inflation also resolves some other puzzles lurking in the picture of the Universe we have so far described: for details, see the suggested readings listed below. We stress that all available astronomical evidence is consistent with the suggestion that the Universe underwent an early period of inflation. We are even beginning to be able to set constraints on how it occurred, how long it lasted, and so on. This is an area where cosmology intersects with high energy physics – and we can look forward to rapid progress.

[ For an early description of inflation, see the article by Guth and Steinhardt in the 1984 Scientific American. The website http://www.ctc.cam.ac.uk/outreach/origins/inflation_zero.php also treats inflation (and some other topics we have presented, albeit in a quite different way). The Wikipedia article on cosmic inflation is very rich – but also very technical. Finally, in Sky & Telescope, 2005, Nadis describes some of the many variants of inflation theory. ]

Module 8 Observations of Supernovae Upset the Applecart

In 1998, two independent groups working on a different test of the overall geometry of the Universe reported an astonishing and quite unexpected result. The expansion of the Universe is not slowing down, but instead accelerating. This work was appropriately declared the scientific discovery of the year in 1998, and resulted in a Nobel Prize to the leaders of the two teams a dozen years later. In the box below, we lay out a sketch of the test they were attempting, and an outline of their results. The test was based on using one class of exploding stars (supernovae) as bright but “standard” candles – a class of objects all having the same intrinsic energy output. [See Reiss and Turner, Scientific American 2004, or pages 60-62 of the Cosmic Times Teachers’ Guide (http://cosmictimes.gsfc.nasa.gov/) for more details on Type Ia supernovae, the kind used for these studies. ]
In Module 6, we showed that, in positively curved space, the area of a sphere is less than the expected $4\pi r^2$. Likewise, in negatively curved space, the area is larger.

Keep that in mind as we briefly review the connection between the luminosity and brightness of an astronomical object. Brightness is defined as the amount of energy per second per square meter crossing a surface at distance $r$ from the object. If we assume an object with luminosity $L$ is emitting the same amount of energy in all directions, then the connection between the two is $B = L/4\pi r^2$—the inverse square law of light.

Now imagine an observer remaining a fixed distance $r$ away from an astronomical source of luminosity $L$. If we leave everything else constant, but merely change the overall geometry of the Universe to a positively curved geometry, the area of a sphere of radius $r$ is smaller than $4\pi r^2$, so the amount of energy traveling through each square meter of the spherical surface will be somewhat larger; that is, the brightness is larger in positively curved geometry than it is in flat geometry, and vice versa for negatively curved space. This result allows for a separate independent test of the geometry of the Universe. If we have sources of known $L$, we can measure $B$ at a distance $r$, and thus find the geometry.

Unfortunately, the measurement of distances in astronomy is fraught with problems—it is hard to measure $r$ exactly. As a consequence, we use a different trick. Imagine that we have a set of objects at different distances, but all having the same intrinsic luminosity. These are traditionally called “standard candles.” If we look at nearby objects, so that $r$ is small, the difference between flat, negative and positively curved geometry is small (remember the earlier demonstration, number 5 in Module 6). As $r$ increases, however, the non-Euclidean departures grow larger. That allows us to draw, qualitatively, a figure like A1 for the three cases. If we look

![Fig. A1. How brightness, $B$, depends on distance for a set of standard candles in spaces of three different curvatures. Only for flat space does $B$ vary as the inverse square of $r$.](image)

at a bunch of objects, again assuming they all have the same intrinsic luminosity, do their measured brightnesses as a function of distance fall along the upper, middle or bottom curve? That allows us to determine the geometrical curvature.
As already noted, measurements of $r$ are slippery in astronomy. Much more precise is the measurement of redshift, which can be used as a proxy for distance. An object with a bigger redshift is receding faster from us, and therefore is further away. Hence we can recast Fig. A1 in the form of a plot of brightness vs. redshift, as shown in Fig. A2. This is the so-called “brightness-redshift” test.

![Fig. A2.](image)

Fig. A2. As in Fig. A2, except that we use redshift as a proxy for distance. Once again, how $B$ changes with distance is different for the three values of space curvature. It is true that in positively curved space, $B$ begins to increase beyond a certain redshift, as shown.

The one crucial requirement is a set of objects of the same luminosity; “standard candles.” The two groups applying this test were both using a particular class of exploding star, supernovae Type Ia, as their objects of standard luminosity. Supernovae of this type are luminous enough to be individually detected in galaxies out to distances large enough so that the curvature effects become important. Members of the two teams naturally expected their observations to fit one of the three lines shown in Fig. A2. Instead, the results were quite different, and can be explained only if the expansion of the Universe is accelerating so that density alone is no longer destiny.

[There are excellent discussions of this Nobel-winning work in Adam Reiss’s 2004 article in Scientific American and Saul Perlmutter’s slightly more technical article in the 2003 Physics Today. It is also fun to read the Nobel organization’s description of the work that led to the 2011 Prize: http://www.nobelprize.org/nobel_prizes/physics/laureates/2011/advanced-physicsprize2011.pdf.]

As noted in the box above, the supernova teams expected their measured brightnesses to fit one of the three curves. They did not (for details, see the references listed in the box). The trend of brightness with redshift, $B(z)$, can be explained only if the Universe is expanding faster.
and faster – that is, if the expansion is *accelerating*. We earlier argued (correctly) that gravity acts to slow the expansion of the Universe. What can cause acceleration instead?

[Demo 7: Back to the graph of $a(t)$. If the expansion of the Universe is accelerating, rather than slowing down, the curve of $a(t)$ must be curving upwards. Lead students to this conclusion, and open the question of what could cause more and more rapid expansion, rather than deceleration.]

Interestingly, Einstein had supplied, then abandoned, an answer to that question many years earlier. When he first applied his new theory of General Relativity to the cosmos, Einstein, like all his predecessors and contemporaries, naturally assumed that the Universe was static. Hubble’s discovery of uniform expansion lay some years in the future. So Einstein was faced with a problem. How can the Universe be static if gravity acts always to pull matter together? Einstein’s answer was to introduce into the equations of General Relativity something called the cosmological constant. This is akin to a new “force,” to speak loosely, which serves to balance gravity. In his formalism, this new “force” depends only on distance, not mass. It allowed a precarious balance with the inward pull of gravity, and hence static equilibrium. (We will explore the physical properties of the cosmological constant below, and drop the potentially misleading description of it as a kind of “force.”)

Once the expansion of the Universe was discovered, however, the cosmological constant was no longer needed, since the Universe was no longer thought to be static, and Einstein abandoned the cosmological constant. The statement that the cosmological constant was “the biggest blunder of my life” is attributed – probably erroneously – to Einstein. In any case, it may be that declaring the cosmological constant a blunder was in fact a bigger blunder. Since it acts to counter gravity, the cosmological constant can produce acceleration. It is just what we need to explain the supernova observations described in the box above.

Introducing the cosmological constant, however, undermines several of the conclusions drawn in earlier modules. Since this “force” is independent of mass, its magnitude is unaffected by density. Thus density alone is no longer destiny as claimed in Module 4. Density and the cosmological constant take turns in determining the expansion rate of the Universe. Early in the history of the Universe, when the density was high, deceleration due to gravity dominated. Only later did the density fall enough so that the cosmological constant became the dominant effect. The result is to produce a far more complicated graph of the scale factor, $a(t)$, as shown schematically in Fig. 17. At early times, the expansion of the Universe is indeed slowed down by the action of gravity (so initially $a(t)$ behaves as shown in the earlier Fig. 8). The Universe was nevertheless still expanding, so the density was dropping. Thus the relative strength of gravity decreased compared to the (constant) cosmological constant. At a certain moment in time, the two “forces” became equal, but the expansion continued, reducing still further the effect of gravity. Once the Universe began to accelerate, of course, the density dropped so rapidly that the matter no longer matters. The expansion runs away. In the limit as density $\to 0$, the scale factor increases exponential with time. Also, once the expansion enters an accelerating phase, our future becomes clear – we live in a Universe that expands forever, no matter what the density is.

The crossover to a phase of accelerating expansion is now thought to be at redshift $z \sim 1$, i.e., when the Universe was several billion years old. Since the exploding stars Schmidt and
Perlmutter and their colleagues were observing were all relatively nearby, they necessarily probed an epoch when the expansion was accelerating.

---

**Fig 17.** Our current understanding of the expansion history of the Universe. Early on, gravity dominates, and the curve $a(t)$ is concave downwards (the expansion is slowed, as was the case in Fig. 8). At later times, but well before the present moment $t_0$, the cosmological constant takes over, expansion accelerates and the curve arcs upwards.

Since the effect of the cosmological constant does not depend on density, one casualty of this discovery is the simple connection we made in Module 4. As we have noted, density is no longer destiny. To know our future, we need to know both the density and the magnitude of the cosmological constant. Another consequence of adding a cosmological constant is to make the Universe older, that is to increase the span of time between the moment of the Big Bang and the present epoch.

[This is a useful if somewhat detailed exploration of the new graph of $a(t)$ in Fig. 17. If we draw a tangent line to the $a(t)$ curve in Fig. 17 at the present moment, and project it back to the moment when the scale factor goes to zero, we see that the backward projection cuts the horizontal axis much closer to $t_0$ than the actual $a(t)$ curve does—the actual Universe is older than we might expect by projecting the present rate of expansion into the past.]

Let us return to the fundamental question: what produces the observed acceleration? All astronomical observations currently available, including those of supernovae, are consistent with the view that the detected acceleration is caused by a truly constant cosmological constant, of just the sort envisioned by Einstein. The cosmological constant was introduced as an extra term in the equations of General Relativity, without any real physical explanation. But there is a more fundamental way to look at the cosmological constant, and that is in terms of a new constituent of the Universe, *Dark Energy*. This is emphatically different from Dark Matter, which "gravitates" or attracts like ordinary matter. Dark Energy instead has very different properties. First, since the cosmological constant is indeed constant, the amount of Dark Energy must
remain constant even as the Universe expands. This is in flat contradiction to the behavior of matter (whether Dark or ordinary): for these, the density drops as the Universe expands. This apparently paradoxical behavior of Dark Energy is allowed if we treat Dark Energy as a kind of vacuum. If you start with a small container of vacuum, and expand the container, you just end up with more vacuum. Hence the term “false vacuum” often applied to Dark Energy. As the Universe expands, and volume increases, so does the amount of Dark Energy, in just such a way as to keep the amount of Dark Energy per cubic cm constant. Yes, this is perplexing, but it is in agreement with the observations.

There are still more oddities of Dark Energy to explore. We have just argued that the amount of Dark Energy or false vacuum in each cubic cm of the Universe remains constant. Now we make use of Einstein’s best-known equation, \( E = mc^2 \). If there is a certain amount of energy in each \( \text{cm}^3 \), there must be an equivalent amount of mass, \( m = E/c^2 \). This mass, like any other, must gravitate. As a consequence, the Dark Energy does contribute to the deceleration of the Universe. Indeed, since we have argued that Dark Energy appears to be dominating the expansion of the Universe today, the amount of Dark Energy per \( \text{cm}^3 \) must exceed the density of both Dark Matter and ordinary matter. It does, as we’ll show below.

First, however, we’d better sort out an apparent contradiction. Dark Energy was introduced to explain the observed acceleration of the expansion of the Universe. But we have just argued that the Dark Energy contributes an equivalent mass that instead acts to slow the expansion. There must be something more….

There is. The expansion of the Universe is affected by pressure as well as density. So far, we have ignored that complication because the pressure of ordinary matter and Dark Matter is so small we can safely neglect its effect. Not so for Dark Energy. The effect of pressure is proportional to \( 3P/c^2 \), so it is this quantity that must be compared to density. The deceleration produced by a combination of density and pressure is proportional to \( G(\rho + 3P/c^2) \). For ordinary matter \( 3P/c^2 \ll \rho \), and we think \( P = 0 \) for Dark Matter. Interestingly, the major contributor to pressure in the Universe today, other than Dark Energy, is the radiation that makes up the CMB, but even that amount is a tiny fraction of even the density of ordinary matter. Only the pressure of Dark Energy matters.

In the case of Dark Energy (the “false vacuum”), however, the pressure is not only large but negative. This is the last paradox of Dark Energy we need to wrestle with. Let’s start this somewhat technical pair of paragraphs by writing down a general equation of state, which describes how \( P \) and \( \rho \) of a substance are related. We write \( P = w \rho c^2 \) where \( w \) is a dimensionless constant. In the case of an ordinary gas (say, air) which has a molecular weight \( \mu \) and is at temperature \( T \), we know \( P = kT\rho/\mu \), where \( k \) is Boltzmann’s constant. It is easy to show the quantity \( w = kT/\mu \) is very small, as claimed. Note, however, that \( w \) is \( > 0 \). The false volume of the cosmological constant, on the other hand, has \( w = -1 \), so \( P = -\rho c^2 \). If we plug that into the formula given above for deceleration, we find \( G(\rho + 3P/c^2) \rightarrow -2G\rho \). The deceleration becomes negative, i.e. an acceleration.

A value of \( w = -1 \) has another interesting consequence. Consider a random chunk of the Universe. I claim that energy \( Q \) neither enters nor leaves that chunk as it expands (where would this energy come from or go to?). Thus the first law of thermodynamics takes a simple form:
\[ \frac{dQ}{dt} = 0 = \frac{dE}{dt} + P \frac{dV}{dt}, \text{ or } \frac{dE}{dt} = -P \frac{dV}{dt}. \] For ordinary matter at constant pressure, as \( V \) increases, the internal energy \( E \) decreases. For \( w = -1 \), however, we have \( \frac{dE}{dt} = + \rho \frac{dV}{dt} \); as \( V \) increases, so does the internal energy in just such a way as to keep the amount of false vacuum per cm constant.

Whether or not we include the results of the last two, technical, paragraphs, it remains true that the properties of Dark Energy are both counter-intuitive and hard to explain. So it is worth repeating that these odd behaviors of Dark Energy follow from and are consistent with the observations that the cosmological constant is indeed a constant. It is not just the supernova observations that show this: the properties of fluctuations in the CMB are also consistent with a true cosmological constant. To quantify the agreement, we introduce a quantity \( w \), which is identically -1.00 for a pure cosmological constant (the negative sign arises because the pressure is negative). The newest results from ESA’s Planck mission (and related CMB experiments) produce an approximate value for \( w \): -1.1 \pm 0.1. Thus, as claimed, astronomical observations support the existence of Dark Energy. These same observations tell us how much Dark Energy there is. The newest Planck results favor the following mix:

- Ordinary matter – including all the stars, planets and us – a measly 4-5%
- Dark Matter – 27%
- Dark Energy -- 68%

These observations tell us how the Dark Energy acts (it speeds up the expansion of the Universe), and how much of it there is. What they don’t do is to provide a real handle on what the Dark Energy is. We suspect, by now, you or your students will be asking that question. What is this cosmological constant = Dark Energy = false vacuum? There is a crisp and definitive answer to this question: we have no idea. There are some speculative ideas, but no convincing and agreed upon explanation.

One possible hint is that the Universe went through a similar accelerated expansion, called inflation very early in its history (as we outlined at the end of Module 7). “Very early” indeed – this brief flash of accelerated expansion was over by something like \( 10^{-30} \) sec after the origin of the Universe. Is Dark Energy somehow related to whatever produced the much, much earlier inflation? If so, it must be of order \( 10^{100} \) times weaker. The explanation of Dark Energy remains an open question, the “what” of Dark Matter unexplained.

In many ways, this is an excellent place to end this “Curricular Companion.” Scientific exploration is an ongoing process. We don’t know all the answers. There is – fortunately for us scientists – much still left to figure out.

Acknowledgements. An early draft of “C5” was substantially improved thanks to the careful attention of another Planck Team member, Jatila van der Veen. Ben Walter, Haverford ’13 was a careful editor of the text and provided some of the figures. Eric Caliendo, Haverford ’16, also read the draft with care. Our work and theirs was supported in part by a NASA-Planck subcontract to Haverford College from JPL.