An Introduction to Our Universe

Lyman Page
Princeton, NJ
page@princeton.edu

Draft, August 31, 2018

The full-sky heat map of the temperature differences in the remnant light from the birth of the universe. From the bluest to the reddest corresponds to a temperature difference of 400 millionths of a degree Celsius. The goal of this essay is to explain this image and what it tells us about the universe.
# Contents

1 Preface ................................................................. 2

2 Introduction to Cosmology ........................................ 3

3 How Big is the Universe? .......................................... 4

4 The Universe is Expanding ......................................... 10

5 The Age of the Universe is Finite ................................. 15

6 The Observable Universe .......................................... 17

7 The Universe is Infinite ?! ......................................... 17

8 Telescopes are Like Time Machines .............................. 18

9 The CMB .................................................................. 20

10 Dark Matter .......................................................... 23

11 The Accelerating Universe ........................................ 25

12 Structure Formation and the Cosmic Timeline ............... 26

13 The CMB Anisotropy ................................................ 29

14 How Do We Measure the CMB? ................................. 36

15 The Geometry of the Universe ..................................... 40

16 Quantum Mechanics and the Seeds of Cosmic Structure Formation. 43

17 Pulling it all Together with the Standard Model of Cosmology 45

18 Frontiers .................................................................. 48

19 Endnote ................................................................... 49

A Appendix A: The Electromagnetic Spectrum ............... 51

B Appendix B: Expanding Space .................................... 52

C Appendix C: Significant Events in the Cosmic Timeline 53

D Appendix D: Size and Age of the Observable Universe 54
1 Preface

These pages are a brief introduction to modern cosmology. They were written for family and friends who at various times have asked what I work on. The goal is to convey a geometrical picture of how to think about the universe on the grandest scales.

One of the challenges in presenting any recent development in science is to connect with common knowledge. We have all talked about the universe since we were young. The vocabulary we use—“Big Bang,” “universe,” “infinity”—means different things to different people. We will give these terms specific meanings as the essay develops. At the more detailed level we will, for example, assume that the reader is familiar with the notion of light as a wave of electromagnetic energy of a certain wavelength and will just remind him or her of that fact instead of explaining it. Appendix A, entitled “The Electromagnetic Spectrum,” has a short guide to various sources of electromagnetic radiation and their wavelengths.

We expect that most readers know that the speed of light, indeed of all electromagnetic radiation, is finite and a fundamental constant of Nature. What is less widely appreciated is that no matter where you are in the universe or how fast you are moving you will measure that the speed of light is 186,000 miles per second. This is one of the foundations of Einstein’s special theory of relativity. We will not dwell on that but simply take it as a fact of Nature that allows us to define universal measures of distance.

To connect with daily experience we will at first work in lengths and distances measured in feet and miles but will for the most part transfer to the more natural units of light-years on the large end and microns (or millionths of a meter) on the small end. A human hair is typically between 20 and 100 microns in diameter. We explain light-years in Section 3. Later on we use centimeters. Recall that they are a little less than half an inch. By necessity we will be quantitative but the math needed will be at the level of distance= speed × time. When we feel that the precision of a number is important, we will give three of more non-zero digits, for example “186,000.” For the most part we will give round numbers as they are easier to grasp.

The subject of the universe is ripe for wild theories and speculation. In this essay the goals are modest. We will stick closely to what we can know through observation and measurement. We’ll try to explain not only the observations but also give a picture in which they can be tied together. The picture we give is not the only possible one, but it explains the data with a minimal set of assumptions. Continuing observations will tell if it is correct. In keeping with our simple goals, we will not delve into the details of what might have happened before the Big Bang nor speculate too much about what might happen billons and billions of years in the future.

My own path to learning cosmology came through making measurements of the Cosmic Microwave Background (CMB). In keeping with that, this essay focusses on understanding the CMB.

As this is a popular piece, there are no scientific references and the attribution of specific ideas and findings is minimal. I hope my colleagues forgive me!
2 Introduction to Cosmology

Cosmology is the study of the universe at the most extreme scales of space, energy, and time. The goals are to understand how the universe came to be, its composition, its geometry, how it evolved, and the laws of physics that describe it.

We now have a widely accepted Standard Model of Cosmology that agrees remarkably well with all measurements of the cosmos. The model is predictive, testable, and could easily be falsified or augmented if that were called for. Among other things, the model says that the universe is comprised of a little under 5% atomic material, the stuff of which we are made, about 25% “dark matter,” and 70% “dark energy.” You have probably heard these different terms. One of our goals is to explain what they mean. The model is based on Einstein’s theory of gravity or the “General Theory of Relativity” and specifies how the various components interact and evolve from the very earliest times to the present. We take from general relativity a way of thinking about space and describe how the cosmic components—the atoms, dark matter, and dark energy—fit together to make the universe we observe.

Perhaps the most amazing thing about the universe is that we can understand it at its grandest scales to percent-level accuracy through measurement. This has become possible only in the last couple of decades. As we shall see below, the universe at the largest scales and earliest times is remarkably simple. It is much easier to understand than, say, the Earth. With its atmosphere, oceans, moving continents, magnetic field, just to name a few attributes, the Earth is fascinatingly complex. That being said, although we have an excellent model of the universe, we do not yet have a fundamental understanding of most of its constituents.

It is an observational fact that thermal radiation suffuses the universe. It is everywhere. We absorb minute amounts of this radiation when we walk outside. It is called the Cosmic Microwave Background (CMB). It is electromagnetic radiation in the form of heat. In general characteristics it resembles radiant heat from the Sun or from an electric stove burner, but it corresponds to a much much colder temperature. In fact the temperature of the radiation is a mere 2.725°C above absolute zero or 2.725 K.\footnote{The CMB is often called the “3K background” because 2.725K is almost 3K. The number of °C above absolute zero corresponds to the Kelvin temperature scale. That is, 1°C above absolute zero is 1 K; there is no ° sign for Kelvin. A change of, say, 0.01°C is the same as a change of 0.01 K. In this system, which we’ll use from here on, absolute zero is −273.14°C, water freezes at 0°C or 273.14 K and boils at 100°C or 373.14 K. The Sun is about 5500°C or 5780 K.}

The CMB is interpreted as the thermal afterglow of the birth of the universe. The evidence in support of this interpretation is overwhelming.

There is much more to the CMB than its temperature. In fact most of what we know comes from tiny variations in its temperature and polarization from position to position across the sky. The CMB is ever so slightly different in temperature in, to pick two arbitrary directions, the north and south celestial poles. Because the CMB can be measured in such exquisite detail, our understanding of it is the foundation for our cosmological model. In following a path to explain observations of the CMB, we will develop our picture of how to think about the universe as a whole.
There are a number of tricky aspects to envisioning the universe. One of the fundamental ones is that the universe evolves in both space and time. To start off, we will treat the spatial and temporal aspects separately. As we build up the picture, you will hopefully get a feel for when it is more convenient to think in terms of space and when it is more convenient to think in terms of time.

The distances and time scales needed to quantify the universe are so large they can be difficult to imagine. To make them easier to grasp, we will count things in “billions.” One billion is 1,000,000,000, or a thousand millions. There are somewhat over 7 billion people on Earth. In the tip of your little finger there are about 1 billion cells. One billion M&M’s would slightly overfill a cubic box about 20 feet on a side.

3 How Big is the Universe?

It is really really big! More seriously, this is a deep question. Before we get to what the question even means, let us first consider some typical distances. In cosmology, the distances are truly vast. To set the scale we start locally and then work our way out.

The Moon is about 250,000 miles away. This is near by. It is close to the typical mileage on a car before it breaks down. With a really good car you could drive to the Moon and possibly even make it back.

A convenient way to measure distances is by saying how long it takes light to travel the distance. We noted above that in one second light goes 186,000 miles, and so in 1.3 seconds it goes 250,000 miles. Thus we say the Moon is 1.3 light-seconds away. Note that we are using time (light-seconds) when we are really talking about distance. The speed of light is a constant of Nature and so it is a convenient standard. To be specific, 1 light-second is the distance it takes light to travel in one second, or 186,000 miles.

The Sun is on average about 93,000,000 miles from us, or about 8 light-minutes away. Because the fastest information can travel is the speed of light, when something happens on the surface of the Sun we cannot observe it until 8 minutes have elapsed. We will come back to this concept but applied to cosmic scale. The Sun is a fairly typical star but most importantly, it is our star. The next nearest star is Proxima Centauri. It is 4.3 light-years away. It is actually part of a small star system. If there were a civilization there with which we wanted to communicate, it would take 4.3 years for them to get our message and then another 4.3 years for us to hear the response. Communication would be rather slow!

When you are next away from city lights and look up at the night sky, you will see a swath that is brighter than everything else. This glow comes from billions of stars that are part of the disk of the Milky Way, our galaxy. In a typical galaxy there are roughly one hundred billion stars. One way to remember the number is that Jeff Bezos and Bill Gates are worth about 100 billion dollars. That is, they have a dollar for every star in our galaxy. In a similar vein, our brains have about 100 billion neurons so there is a neuron for every star. The stars in the Milky Way are collected in sort of a dinner-plate shape called the galactic plane that is about 100,000 light-years in diameter and about 3,000 light-years thick. The solar system is about half the way out from the center of the plate. When we look toward

---

2That is 93,000,000 miles divided by 186,000 miles per second is 500 seconds or a little over 8 minutes.
the center of the plate we see many more stars than when we look well off to the side. It is a bit like living on the outskirts of a city. You are a part of the city, but you can still see all the tall buildings off in one direction.

Figure 1 shows a picture of the Milky Way. This image was taken with a CCD camera using visible light. That is, if our eyes were more sensitive and larger we would see the galaxy like this. Visible light is one form of an electromagnetic wave. Its wavelength is about 0.6 microns, or about one hundredth the diameter of a human hair. The dark swaths in this image are from dust in our galaxy that obscurs the starlight.

![Image of Milky Way with GC and GP labels](image)

Figure 1: The Milky Way, running diagonally up at roughly 45° from the lower left. This picture was taken by Alessandro Schillaci from just outside of San Pedro de Atacama, Chile. Cerro Lincancabur is in the lower left. The bright dots away from the galactic plane that suffuse the night sky are stars in our galaxy. The galactic plane lies between the arrows on the periphery labeled “GP.” The intersection of the galactic plane and the line indicated by “GC” marks the galactic center. The dark regions in the galactic plane are “dust lanes.” The dust blocks visible light but emits thermal radiation as shown in the next figure.

Figure 2 shows a picture of the Milky Way taken by “DIRBE,” an infrared telescope, and one of the three instruments on the COsmic Background Explorer (COBE) satellite. Unlike the other two instruments discovered the anisotropy in the CMB (DMR, leader George Smoot) and
the image in Figure 1, this one was made at “far-infrared” wavelengths, in particular at 100 microns. The “infrared” spans a large wavelength range of the electromagnetic spectrum, roughly from 1 micron to 1000 microns (see Appendix A). In this image we see primarily the thermal glow, in other words the emission of heat, from the dust that fills our galaxy. It turns out that a typical galaxy like the Milky Way has an average temperature of about 30 K so it isn’t very hot, but it still emits thermal energy.

There is a loose analogy with an incandescent light bulb. The bulb is most obvious to us because of the visible light it emits, as in Figure 1. However, much more energy comes out as heat that we can feel but not “see.” You may have seen pictures of houses taken in infrared light. These pictures tell you where the heat is leaking out (often at the windows). When you feel the heat from a hot body, it is mostly infrared radiation that you feel. Infrared radiation tells us how things emit heat.

Our galaxy is a member of the “local group” of roughly fifty galaxies as shown in Figure 3. The local group is some six million light-years across. The Milky Way is second in size to the Andromeda Galaxy. Whereas Andromeda has about a 1000 billion stars, the smaller “dwarf” galaxies have tens of millions of stars. The Large Magellanic Cloud (LMC, Figures 2 & 3) is a nearby small galaxy that orbits the Milky Way. Although there are no sharp boundaries for when something is said to be “cosmological,” we typically think in terms of the average properties of spheres or cubes about 25 million light-years across. The local group is just a fraction of this size.

Figure 4 shows an amazing image. It was taken with the Hubble Space Telescope by observing one region for a long time so that it could build up sensitivity to faint objects. It is called the Hubble Ultra Deep Field (HUDF). The most distant objects in it are billions of light-years away. The area covered by the image is about 1/60 the area of the full moon. The full moon is about 1/2 degree across. You can work out that it takes 200,000 full moons to cover the full sky. Here is the mind blowing thing about the image. Only a hand full of the objects in the image are stars, the large majority of objects are galaxies. And each of those galaxies typically has about 100 billion stars.

To get the number of galaxies in the image, you simply need to count them. With a full resolution picture you could do this by hand but it is easier to use computers. The HUDF team finds there are about 10,000 galaxies. This means that across the full sky there are about 100 billion galaxies. We say that in the observable universe there are roughly 100 billion galaxies, each typically with about 100 billion stars. It is a coincidence that the numbers are so close.

made the definitive measurement of CMB temperature (FIRAS, leader John Mather). Mike Hauser led DIRBE. The instrument is perhaps best known for discovering the combined thermal emission of all the galaxies in the universe.

Modern LED or CFBs have a higher ratio of visible light to heat which is why they are more efficient for lighting.

Although named after Ferdinand Magellan’s report of them in 1519, they were first recorded over five hundred years earlier by Abd al-Rahman al-Sufi Shirazi, a Persian astronomer.

This is \((10,000 \text{ galaxies per HUDF}) \times (60 \text{ Deep Fields per full Moon}) \times (200,000 \text{ full Moons in the full sky})=120,000,000,000\) which we round down to 100 billion. If we looked at much lower mass galaxies than are readily seen with the Hubble, we might get a factor of 10 more but each would have far fewer stars. The main point, though, is that there are a finite number of typically sized galaxies.
Figure 2: The glowing dust in the Milky Way as observed by the DIRBE instrument aboard the COBE satellite. This image is at a wavelength of 100 microns or about 200 times longer than that for Figure 1. This is called “far-infrared” radiation. While in Figure 1 the dust obscures starlight, at this wavelength you can see the dust glowing. The center of the Milky Way is in the center of the image. The blob above the center is the Ophiuchus Complex, a large dust cloud. The feature on the far left is the Cygnus region, and the bright spot to the lower right is the Large Magellanic Cloud (LMC). In Figure 1, the LMC was off to the right and below the horizon when that picture was taken. This image is 180° wide. This is the amount of sky you could see horizon to horizon if, say, you were out in the desert. The height of the image is half that. Credit: NASA/DIRBE.
Figure 3: The local group of galaxies. Some of the local galaxies are well known. Andromeda is about 2.5 million light-years away but can be seen with the naked eye in good conditions. In length it appears a few times as large as the full moon. The Magellanic clouds are readily visible in the southern hemisphere. The larger one, close to the Milky Way in this image and shown in Figure 2 emitting thermal radiation, is about twenty full moons across. The top and bottom wire grid “wheels” are six million light years in diameter. Credit: Andrew Z. Colvin, https://en.wikipedia.org/wiki/Local_Group.
Figure 4: The Hubble Ultra Deep Field. The vast majority of the objects in this image are galaxies. This picture was taken while observing in the direction of the constellation Fornax. Credit: NASA/HST.
We have made two huge jumps here. How do we know that there are not a whole bunch of faint objects that we cannot see? What do we mean by “observable universe?” If we look in the HUDF direction, can we look out to greater and greater distances with more and more powerful telescopes and see ever greater numbers of galaxies? We will come back to these questions but the quick answer is that the Ultra Deep Field shows us, to a reasonable approximation, all the Milky Way type galaxies that can be seen in that direction.

If we could tour the universe, what would we see? For a moment, think about space as static and endless. Let’s put aside the Big Bang and imagine that someone, say Alice, could go anywhere in the universe instantaneously and communicate with someone else instantaneously. We can think of galaxies as cosmic signposts. We can in principle give them names and know where they are in the universe. This has already happened in the local group as you can see in Figure 3, but we want to go to much larger distances. Let’s say Alice is in a distant galaxy that is ten billion light years away. We ask her to describe the local cosmic environment in broad terms such as the number and general appearance of the other galaxies near her. We then compare our description from our home in the Milky Way to Alice’s. We find the descriptions are similar. Although there would be a large variety of galaxies, no matter where we went, no matter how far away, no matter what direction, on average the galactic environment would look very much like it does right around us and the same laws of physics would describe Nature.

This is an important conceptual point and is worth repeating because we will build on it. At this instant in time, every place in the universe looks, in broad brush strokes, similar. We could call up someone near any distant galaxy and say “describe the galaxies within 25 million light-year diameter sphere centered on you” and we would find their general description described our galactic neighborhood.

That the universe is on average the same everywhere you go at a specific time is called the "Cosmological Principle." Another way to say it is that the universe is homogeneous when averaged over a large enough volume and looks the same in all directions, or is “isotropic.” The cosmological principle is fundamental and at the root of modern cosmology. The way in which we use it was put forward by Einstein.

This simple picture of an infinite, endless, and static universe that we can instantaneously explore is a starting point. We know, though, that the universe evolves and that the speed of light is finite, precluding instantaneous communication.

4 The Universe is Expanding

In the preceding section we painted a picture of the immensity the observable universe. In our picture we could think of it as static. We considered objects regardless of how far away they were. When we counted objects in the HUDF we wondered if there were even fainter objects like the ones around us but much further away that we just couldn’t see. To answer this question, we need to think of the universe in both space and time. We take the first steps here.

The universe is expanding. This is not a theory, or model, or anything like that. It is an observational fact. Once you get well past the local galactic group (Figure 3) and into
cosmological distances one observes that *the further away a galaxy is, the faster it is moving away from us*. This is called Hubble’s Law, after Edwin Hubble, who published it in 1929\(^7\). In day-to-day terms Hubble’s Law says that for every million light years away you observe an object, its recessional velocity increases by about 15 miles per second. This value is called “Hubble’s constant.”

Hubble’s observation immediately brings to mind a question: “Are we at the center of the universe?” The answer is “no.” We discuss this below. Hubble’s observation also implies that the universe was more compact than in the past. To be specific, we will use the term “compact” to mean smaller in length or distance as opposed to volume. If the diameter of a sphere were halved, we would say it is twice as compact even though its volume would decrease by a factor of eight. When the universe was twice as compact, everything was half as far away. We return to this in the next section.

Just because we see all galaxies rushing away from us, it does not mean that we are at the center of the universe! We are special but not *that* special. All observers on all galaxies anywhere in the observable universe see the same thing. This is because the expansion has a very particular form, namely that the recessional velocity is proportional to the distance. That is, if a galaxy is twice as far away, it is moving away from us twice as fast. Let’s be more concrete and imagine a sample string of galaxies each representing their local region of 25 million light years across. We’ll start with the Milky Way in the center. If the galaxy named “Nan” was 25 million light years away it would be moving away at 375 miles per second according to the value of Hubble’s constant ([15 miles/sec/million light-years] × [25 million light-years] = 375 miles per second). If another galaxy named “Orr” was 50 million light years away, it would be moving away at 750 miles per second, and if “Pam” was 75 million light years distant it would be receding at 1125 miles per second. These are depicted in a row in the top panel of Figure 5. Even with these enormous distances, the speeds are less than 1% the speed of light.

Now imagine in your mind that you could be instantly transported from the Milky Way, in the center of Figure 5, to Nan. That is, you would be at rest on Nan. Of course, if you looked back at the Milky Way while sitting on Nan it would be moving away from you at 375 miles per second. Here is a way to think about how the whole picture changes. If you were on the Milky Way and wanted to be at rest with respect to Nan, as long as it was moving 375 miles per second to the right but no matter where in space it was, you would need move at 375 miles per second to the right. This is shown in the middle frame. Moving next to something with the same velocity it has is the same as being at rest with respect to it. This is just like looking at a car next to you on the highway going the same speed. Relative to you, that car is stationary. In the bottom picture, we have just subtracted the velocities\(^8\)

---

\(^7\)Hubble’s original value, which was based on his observations of distances and velocities measured by Vesto Slipher, was about seven times the currently accepted value because of a flawed distance estimator. When you get very very far away, you need to account for the General Theory of Relativity but we will leave this aside. The value used in the scientific literature is 70 km/s/Mpc which is about 15 miles/sec per million light-years. The history of the discovery, like so many, is complex and involved many others including his assistant, Milton Humason, and a pioneer of theoretical cosmology, Georges Lemaître.

\(^8\)A velocity is a speed with a direction associated with it. To subtract velocities, take the arrows in the middle, reverse their directions, and add them to the top row.
Figure 5: The expansion of the universe in one dimension. The top row shows the expansion from our point of view. The “MW” stands for the Milky Way and the red indicates that it is reference point. The arrows indicate velocity. The galaxy Nan, at 25 million light years distance is moving away from the Milky Way at 375 miles per second. Since Orr is twice as far away as Nan, it is moving away twice as fast as indicated by the arrow being twice as long. The middle row shows the speed of Nan but at all points in space, not just at Nan’s location. Say you were moving at this velocity and near Nan. To you it would seem as though Nan was standing still. The bottom row shows the pattern of velocities from the point of view of someone on Nan. The red dot indicates it is now the reference. As you can see an observer on Nan seems to be the center of the universe and all the other galaxies are rushing away and obeying the same Hubble law. Even Jon, which is 100 million light years away from Nan, is receding at 1500 miles per second.
Figure 6: The expansion of the universe but now in two dimensions. Each point represents a galaxy. The arrows show the speed and direction the galaxy is moving as we see it. Of course, in reality the galaxies are much more irregularly spaced. The left side shows what we’d see looking out to greater distances than in the previous figure. Again it looks like the galaxies are all moving away from us in proportion to their distance. Imagine that instead we transport ourselves to the galaxy marked by the red arrow four galaxies in and five up from the lower left and so that we are at rest with respect to it. Again the overall picture is the same: we seem to be the center with all other galaxies rushing away from us with speeds proportional to their distance.
to see what the universe looks like from the perspective of someone on Nan. But now note that the bottom picture looks just the same as the top but from the perspective of someone on Nan. Here again the same Hubble law applies. Someone on Nan suspects they might be the center of the universe.

For the time being, just imagine that this line of galaxies can go on forever and that the velocities increase without limit. We will come back to this later. The main point here is that as long as the velocity of recession is proportional to distance, all observers in the universe see the same pattern of recession and to all it appears that they are in the center of the expansion. Although we have shown the expansion with a line of galaxies in one dimension, it works in two and three dimensions as well. Figure 6 shows the same process in two dimensions from the perspective of two widely separated galaxies.

There is a simple and yet radically different way of thinking about the expansion. In the picture we just gave, we had in our minds a fixed space in which the galaxies were moving. That is, the space was fixed and the galaxies were traveling through it at different velocities. We now want to make a huge conceptual jump. Let us again imagine that the galaxies are on a line as in the top row of Figure 5 but let us ignore the velocity arrows. Think of the galaxies as representing coordinates in space, just like mile markers on a highway (in one dimension) or as latitude and longitude positions on a two-dimensional map. Instead we want to add space between the mile markers. In the top row, this is equivalent to taking a pair of scissors, cutting the figure vertically between all galaxies, and taping and extra strip of paper of width, say, 0.2 cm wide. Let’s say it takes us 30 minutes to do this process. After we are done, Nan is 0.2 cm further from the Milky Way than it initially was, Orr is 0.4 cm further away than initially, and Pam is 0.6 cm away than initially. In this 30 minutes, Pam has moved three times as far as Nan, and Orr has moved twice as far as Nan. Pam who started out three times further than Nan has moved three times as far in the same 30 minutes and thus apparently has three times the velocity. We have reproduced the cosmic expansion but from a completely new perspective. Instead of space being fixed with the galaxies moving, the space between the galaxies is expanding.

From now on we want to think of space as fungible, not as a set stage on which the cosmic evolution unfolds but rather as the entity that is evolving. In the above Jon, Nan, and Pam don’t have to communicate with each other. They just sit still while space is created locally everywhere at the same time and at the same rate. In this picture, Hubble’s Law is just a statement about a specific rate of the expansion of space. In the two-dimensional case shown in Figure 6 we cannot cut strips of paper but can instead imagine the galaxies as painted dots on a rubber sheet. Here, expanding space would be like stretching a very very large sheet in both dimensions. In three dimensions we might imagine a very very large grid of Tinker Toys with the wooden hubs as the fixed coordinates and the struts between the hubs as the space that grows with time.

To reiterate, we can think about the expansion of the Universe as the expansion of space. The rate of expansion depends on the rate that we “make space.” We should not think about

---

9 For the experts, what we term “expanding space” is increasing the scale factor $a(t)$ in the metric. There are heated discussions about whether “expanding space” is a useful concept. In Appendix B we discuss some of the pitfalls.
the expansion of the Universe as galaxies flying away from each other in a pre-defined space. The “Big Bang” was not like a bomb exploding billions of years ago. The Big Bang marked the beginning of an explosion of space, everywhere at a fixed time in our distant past.

To recap, we started off this section by explaining Hubble’s Law and showed that no matter where you are in the universe, you appear to be in the center with the velocity of recession of other galaxies proportional to their distance. We then introduced space as the changing quantity, and realized that if the galaxies represent fixed coordinates, Hubble’s Law is a special case of expanding space at a specific rate. In general we can expand space at any rate and still not be in a special place. We continue to think of the universe as infinite in extent.

Our new way of thinking about space begs the question “What is space?” This is a deep question akin to “What is a vacuum?” Most physicists would say that we do not know. There are epochs in our cosmic history when an expanding space is the best description of Nature, and there are epochs when it can be misleading. Regardless, it is a conceptually unifying concept that aids in envisioning the expansion of the universe and dovetails nicely with the warping of spacetime described by general relativity theory. We consider other elements of expanding space in Appendix B.

Before moving on we note that on human scales the expansion is ignorable. We see it only because we can look out to such vast distances. In Woody Allen’s movie “Annie Hall,” Alvy is in a room with his mother and Dr. Flicker. Alvy is concerned that the universe is expanding and that “someday it will break apart and that would be the end of everything!” His mother, agitatedly, says “What is that your business? He stopped doing his homework.” Alvy replies “What’s the point?” To which his mom says loudly “What has the universe got to do with it? You’re here in Brooklyn! Brooklyn is not expanding!” She’s right! The reason is that the forces that hold the Earth together and that bind the Earth to the Sun completely dominate the effects of an expanding space. Even our galaxy is not expanding. We can be more quantitative. In 100 years, the width of this page would expand by about 0.001 microns, or roughly ten times the diameter of an atom if it partook in the cosmic expansion! This would be easily measurable. However, the forces that bind the molecules in the paper, which by measurement appear to be constant in time, would keep it at its current size.

5 The Age of the Universe is Finite

If the galaxies are all apparently moving away from us now, in the past they were closer together. Or, in our new way of talking there was “less space” in the past. At some point in the distant past, the galaxies were much much closer together. If we go back even earlier, the galaxies had not yet formed and instead of thinking of the space between galaxies we think of the space between the constituents of galaxies. As we go further back there was less and less space and, since there is the same amount of matter, the matter density becomes enormous. At some point in our extrapolation the laws of physics break down but we need not extrapolate back that far. The important point is that we can extrapolate back to a time when the universe was extremely dense and that we reach that epoch in a fixed amount
of time. In other words, the universe has a finite age.

The most accurate measure of the age of the universe comes from the Wilkinson Microwave Anisotropy Probe (WMAP) and the Planck satellites. The best estimate, taking into account everything we know about the expansion history, gives 13.8 billion years to roughly 1% accuracy.

We can get an approximate value of the age of the universe from the observations we explored in the previous section. As we saw, two galaxies 75 million light years apart are moving apart at 1125 miles per second. At this speed, assuming it is constant, 12 billion years ago the galaxies were on top of each other. Two galaxies 125 million light-years apart are now moving apart at 1875 miles per second, and so in the same amount of time, they too would be on top of each other. By extension all observers, no matter how far away they are from each other, would say the universe is 12 billion years old as all galaxies would be on top of each other when all the space is taken away. It is fortuitous that this simple estimate is so close to the more accurately determined value of 13.8 billion years as we explain next.

When we extrapolate back we do not know that the rate of expansion has been constant. In fact, we know it has not been constant. One way to think of this is that the gravitational attraction of all the galaxies to each other would tend to slow the expansion. While this makes sense intuitively, in saying it we are linking the presence of matter to the rate of expansion, a concept at the heart of the General Theory of Relativity. Since the expansion rate has not been constant, Hubble’s Law has not been constant throughout our cosmic history. It is a function of time. The value we gave above, 15 miles/sec per million light years, is only applicable now.

As recently as the 1990s there was no consensus on how to extrapolate back because we did not know the expansion rate throughout the history of the universe. However, as the web of observations has become ever tighter we have learned about the constituents of the universe we are now confident that we know the age.

In some models of the universe, the current expansion is just one of many, perhaps one of an infinite number, of expansion cycles. There are no observations as yet to rule in or out such models. Most cosmologists do not subscribe to the cyclic model because it has not been as widely scrutinized as the baseline “inflation model” which we discuss later. However, we should keep in mind that cyclic models are a possibility. If in fact the universe is cyclic, then the 13.8 billion years would refer to the age of this cycle and the discussion in this book would refer only to the latest cycle.

We now have the framework to be more precise about the term “Big Bang.” We will take it to mean the time at which the universe began to expand. It is when we start our clocks. The term has nothing to do with space. We do not yet know enough physics to extrapolate all the way back to the Big Bang even though many cosmologists are working on the problem.

Even though we have talked above about the distances of objects, this can become confusing because of the difference between the distance to the object when the light we observe  

\[ \text{A speed of 1125 miles per second is the same as a distance of 6 million light-years per a time interval of one billion years. The age is then } [75 \text{ million light-years}] / [6 \text{ million light-years/ 1 billion years} ] = 12.5 \text{ billion years. Similarly, a speed of 1875 miles per second is the same as 10 million light-years per billion years. The age is then } [125 \text{ million light-years}] / [10 \text{ million light-years/ 1 billion years}] = 12.5 \text{ billion years.} \]
was emitted and the distance at this instant accounting for the expansion of the universe. Instead of talking about how far away something is, we will mostly discuss phenomena in terms of their age since the Big Bang or equivalently how compact the universe was when the object emitted the light we see. By referring to age and compactness, we sidestep the fact that the universe expands while light from a distant object travels to us. Appendix C has a timeline of significant events and the associated cosmic compactness. For example, about 5.9 billion year after the Big Bang, about 8 billion years ago, the universe was twice as compact. We then say the “scale factor” is 0.5 because distances between objects when the universe was twice as compact are half the current size. The Earth and Moon formed about 9.3 billion years after the Big Bang or roughly 4.5 billion years ago at a scale factor of 0.71. Dinosaurs roamed the Earth 0.1 billion years ago and Homo Sapiens appeared a mere 100,000 years (0.0001 billion) ago when the universe was only negligibly more compact than it is today.

6 The Observable Universe

We just saw that the age of the universe is finite and that all observers agree on its value. The next key ingredient for developing our model is taking into account the speed of light. So far we have primarily used the finite speed of light to establish a distance, namely the light-year. We continue in that vein.

Because the speed of light is finite and we know the age of the universe, there is an upper limit on the size of the universe we can possibly observe. We can get an estimate of its size. To a first approximation, in one direction we cannot see further away than the age of the universe times the speed of light. It is as though each observer is in the middle of a spherical volume with a diameter of $2 \times 13.8 = 27.6$ billion light years. The actual diameter of the spherical volume is a little over three times larger because our approximation did not include the fact that the universe expands while light is traversing it. Nevertheless, the important conceptual point is that because information cannot travel faster than the speed of light there is a limit to how far we can see. This is the “observable universe.” Many times when cosmologists talk about the “universe” they really mean the observable universe. In Appendix D we discuss further the relation between the age, size, and compactness of the observable universe.

If we could instantaneously travel anywhere in the universe right now the galactic environment would look similar to the environment around us. Even at the “edge” of our observable universe, even beyond the edge. There would be a variety of galaxies, there would be local groups, no matter where we went we would compute the age of the universe to be 13.8 billion years.

7 The Universe is Infinite ?!

This is meant to be a provocative section heading. We do not know that the universe is truly infinite in the sense that it goes on forever in space. Far far beyond our observable universe,
space could be much different and even the laws of physics might be different. However, the observations tell us that an infinite universe with properties similar to those around the Milky Way is the best and most parsimonious description of the data. That is, we cannot tell the difference between what we observe and a universe that is infinite in spatial extent.

To put this in context, until a couple of decades ago there was no scientific reason to believe a priori that the universe was infinite. There were many possible models that fit the data. Observations could have told us that the universe was finite, that it contained a fixed amount of stuff. We could still have had an expanding universe with a finite age, but it could have been finite in extent. Instead the observations have told us that for all intents and purposes we should treat the universe as infinite.

Let’s picture for a moment the universe as an absolutely enormous container of chocolate chip ice cream. The chocolate chips are the galaxies and the ice cream is the space. Our observable universe would be like a very large scoop taken from somewhere inside the container, far far away from the container walls. Our scoop would have all sorts of different sized chips. But all scoops would be similar and recognized as the same chocolate chip ice cream no matter where we took them as long as they were well away from the walls. The container walls, if they exist, represent some new physics to which we have no access.

The extent of the universe is an active area of investigation. Every now and then someone comes up with a model for a finite universe. However, when the predictions of the model are compared to the data one finds that an infinite universe provides a better description of the observations. With this picture in mind, the question of “What is the universe expanding into?” is not answerable or even relevant.

8 Telescopes are Like Time Machines

We now add the next conceptual component to our picture. This one again is based on the speed of light but we are not here using light as a distance measure. As we noted in Section 3, since the Sun is eight light-minutes away we see it as it was eight minutes ago. Similarly, if an object is 20 million light-years away, when we observe it we see it as it was 20 million years ago. As we peer deeper and deeper into space we see objects as they were at earlier and earlier stages in their lives. Our whole cosmic history can be read by looking ever deeper into space, because as we do so we look ever more back in time. In other words, telescopes are like time machines.

First let’s think about what this means taking just a small step out. Stars can explode, releasing enormous amounts of light and particles in a “supernovae.” We can see these explosions. In 1987, a star exploded in the neighboring Large Magellanic Cloud (LMC, shown in Figures 2 and 3). The LMC is only about 160,000 light-years away. That is, the star exploded before Homo Sapiens first came on the scene but we only saw the light in 1987. This particular supernovae, called 1987A, is unique because in addition to the light, we detected elementary particles called neutrinos. from it. Neutrinos are elementary particles associated with nuclear interactions such as fission and fusion. They barely interact with matter and only recently were discovered to have mass. From this supernovae alone, roughly 100 billion neutrinos per square centimeter hit the Earth. Most passed right through.
Supernovae are so bright they can be seen to vast distances. With powerful telescopes astronomers can catch supernovae from relatively short-lived stars that exploded when the universe was twice as compact as it is now, 5.9 billion years after the Big Bang. In other words, the original star has not existed for 8 billion years! Mostly what is left is a roughly spherical shell of light and particles traversing the universe. We see this shell as it passes by Earth.

Figure 7: Telescopes are like time machines. As we look out in space we look back in time. With the Hubble Ultra Deep Field image we look back to when the galaxies began to form. Light from the first stars was emitted when the universe was roughly 200 million years old and has been traveling to us since then. We can think of it as coming from a shell out near the edge of the observable universe. The CMB comes to us from a shell essentially at the edge of the observable universe. On this picture it is the outer yellow ring. The label “Big Bang” marks the beginning of the time line. Credit: NASA. Get proper credit and 13.7− > 13.8 Gyr.

Although individual young stars at great distances are too small to see unless they explode, we can see nascent galaxies from when the universe was less than a billion years old. Let’s go back to the Hubble Ultra Deep field. Figure 7 shows us what we see as we peer deeper and deeper into space. The Hubble and other telescopes can just about look back to a time when galaxies were just forming. In section 3 we said that we could count all the galaxies in the universe. We can now see what this means. We can look back to a time before galaxies existed. This corresponds to an epoch when the universe was about twenty times as compact and about 200 million years old (Appendix C). Thus in the part of the
universe to which we have access, our observable universe, we can count all the galaxies. Again, there are about 100 billion similar to the Milky Way.

If we peer deeper still we could see the birth of the first stars. This has not yet been done but instruments are being built to detect them. Going back even further in time, we can see the remnant radiation of the Big Bang, the Cosmic Microwave Background. It is the light from the edge of the observable universe.

9 The CMB

We are now in a position to understand the CMB. There are three main aspects of it. The first is its temperature, 2.725 K. The second is the small temperature differences from place to place in the universe, or, as viewed by us, from position to position on the night sky. The third is its polarization. In this section we consider just the first.

That we can characterize the CMB as a temperature is a profound statement on its own. The CMB is thermal radiation, or radiant energy, of a very particular form. It is called blackbody radiation. A black piece of paper left in the Sun gets much hotter than a white one and hotter still than a perfect mirror. The black piece of paper absorbs the radiation that lands on it, the white piece of paper absorbs some of the radiation but scatters most of it away, and the perfect mirror reflects all the radiation that lands on it and doesn’t absorb any\textsuperscript{11}. From the laws of thermodynamics, it can be deduced that a good absorber of radiation is also a good emitter. Thus, if you put your hand over, not on, the black piece of paper that has been exposed to the Sun you will feel that it is radiating more energy than the white one or a mirror. Other good sources of blackbody radiation are the Sun or a pottery kiln.

Objects emit their thermal energy over a range, or spectrum, of wavelengths. However, even for blackbodies, most of that energy comes out over a limited portion of the spectrum; in other words blackbodies emit predominantly over a relatively small span of wavelengths. For example, almost half of the Sun’s energy comes out at wavelengths between 0.4 and 0.8 microns. It is no coincidence that this is the visible spectrum that our eyes detect. Our eyes likely evolved to take advantage of the Sun’s spectrum. We know the Sun also emits UV radiation. For example, “UVB,” the primary source of sunburn, is at 0.3 microns. But we cannot see that radiation. The Sun also emits in the “near-infrared” region but we cannot see that light either. And as we saw earlier in Figure 2, DIRBE showed our galaxy glowing in the “far-infrared.”

The lower the temperature of an object, the longer the predominant wavelength of emission. This is known as the “Wien displacement law.” It tells us that the dominant wavelength of emission of a blackbody in microns is roughly 3000 divided by the temperature in Kelvin. For example, for the Sun (with temperature 5780 K, see footnote 1), the predominant wavelength is 3000/5780 or approximately 0.5 microns.

\textsuperscript{11}It is difficult to make a perfect mirror. Aluminum is an obvious candidate but is a good absorber of ultraviolet radiation and so gets hot in the Sun. If our eyes could see ultraviolet radiation, which they cannot, an aluminum mirror would look dark.
It is a deep aspect of blackbody radiation that all you have to do is specify its temperature and you know how much energy is radiated at all wavelengths. That is, the temperature describes the entire spectrum. By definition, objects that emit blackbody radiation are in thermal equilibrium with that radiation. In other words, the temperature of the radiation corresponds to the temperature of the object. For example, you might embed a thermometer in the walls of a kiln well away from the radiation. The temperature that you would ascribe to the radiation in the kiln by measuring the amount of energy at each wavelength would be the same as the kiln’s wall temperature as read by your thermometer.

Max Planck derived the celebrated formula that describes blackbody radiation in 1900. The CMB has by now been measured at many wavelengths and to the limits of measurement we know that it follows the Planck formula. Thus, we know that it came from an era when the matter in the universe was in thermal equilibrium with the radiation. Figure 8 shows the measurement of the CMB blackbody spectrum from the COBE satellite. Many have tried to explain the spectrum of the CMB with different sources of radiation. For example, the CMB was once proposed to be the emission by distant clouds of cool dust. Such attempts have not succeeded because the predicted spectrum from alternative sources of radiation

\[ \text{Figure 8: The spectrum of the CMB. The } x\text{-axis shows the wavelength and the } y\text{-axis shows the emitted power as a function of wavelength. The thin black line shows Planck’s celebrated formula for a 2.725K blackbody. The continuous red line shows the measurement from the FIRAS instrument on the COBE satellite. The error bars are smaller than the thickness of the line. Some selected measurements at wavelengths longer than those from FIRAS are shown in red. The agreement between the observations and the blackbody formula is clear.} \]
does not match the observations. Searching for departures from a Planck spectrum is one of the frontiers of cosmology as we discuss in Section 18.

To derive his formula, Planck had to hypothesize that electromagnetic radiation was quantized. In other words, that radiation could be described as discrete packets of energy. The quantum of radiation is called a “photon” or a “particle of light.” This realization marked the birth of quantum mechanics. Part of the foundation of quantum physics is that the interaction of radiation and matter may be considered either as involving waves or as involving particles. At times it is easier to discuss radiation as photons instead of as electromagnetic radiation. For the CMB, there are 400 photons in every cubic centimeter of the universe.

We do not know \textit{a priori} that the universe started off in an incredibly hot state. In fact, the whole picture we have been developing of the expanding universe, the observable universe, looking back in time, etc. could in principle work with a relatively cool early universe. However, because the CMB exists, we know the early universe was hot and in thermal equilibrium. Here is how it works.

When the universe expands, the wavelengths of light are stretched in proportion to the expansion. Imagine you had a slinky. Think of each full turn as corresponding to a wavelength of light. Let’s say that the slinky is initially 10 cm long. Now stretch it to 20 cm. The total number of turns is the same but the space occupied by each turn has increased. This is analogous to the stretch of the wavelength of light as the universe expands a factor of two.

Why do the CMB wavelengths expand with the universe and yet Brooklyn doesn’t expand? The CMB is traversing the universe. With Brooklyn there are also a lot of forces at work. Earth, rocks, buildings and us are held together by strong chemical forces. The only force on the CMB is gravity and that force is very weak by comparison.

If we go back in time, the wavelengths that comprise the CMB decrease because space is more compact. Since the CMB has a temperature of 2.725 K now, the most prominent wavelength of emission is $\frac{3000}{2.725}$ or approximately 1100 microns (0.1 cm or 1 mm) as determined by the Wien displacement law. It turns out that with the expansion of the universe, the Planck formula keeps its same form as long as the temperature changes to account for the expansion. That is, when the universe was a factor of two more compact, as it was 8 billion years ago, the temperature of the CMB was double, or about 5.2 K, and its predominant wavelength was half or 500 microns.

We can keep going back to when the universe was much more compact. At 400,000 years after the Big Bang when the universe was 1000 times more compact, the CMB was 2725 K, about half as hot as the Sun but energetic enough to ionize hydrogen. That is, the radiation was intense enough that the electrons were ripped away from the protons.\footnote{Hydrogen, the simplest atom, has one proton in the nucleus and one electron orbiting it. The proton has a positive charge and the electron is negative. Protons are 2000 times as massive as electrons. Helium has two protons and two neutrons in the nucleus and two electrons in orbit. Neutrons are almost as heavy as protons and have no charge; they are neutral. A free neutron decays on average in 10 minutes to a proton, an electron, and an anti-neutrino.} Going back further to about three minutes after the Big Bang, when the universe was about a third of a billion times as compact as today and thus at billion Kelvin, the radiation was so intense that the nuclei of helium could just hold together. At yet another factor of three thousand times
more compact and hotter, about 25 millionths of a second after the Big Bang, neutrons and protons did not independently exist and the universe was a “quark-gluon plasma.” These are the elementary particle that make up protons and neutrons. This state of matter has been reproduced on Earth in the Relativistic Heavy Ion Collider (RHIC) on Long Island, NY. If one goes back to a state where the universe was about 50 million billion times as compact as today, about a hundred thousandth of a billionth of a second after the Big Bang, the energy in the photons was roughly the energy of a collision between protons in the Large Hadron Collider (LHC) at CERN. These are the highest energy elementary particles yet produced by humankind. Yet, with the universe we can explore even greater energies as we discuss later.

Not only does the light from the CMB get stretched on its way to us, but all light from all distant objects gets stretched. This phenomenon is called the “cosmological redshift” because longer wavelength correspond to redder light. Not only do we see distant objects as they were when they were younger but we also see them through stretched wavelengths.

Let us now tie this in with the picture we have been developing. If we could travel anywhere in the universe *instantaneously* right now, the temperature would be 2.725 K everywhere. We can call this the current temperature of the universe. Another way to think of this is that we are immersed in a gas of photons. If you could have instead traveled *instantaneously* around the universe when it was twice as compact you would have measured the temperature everywhere in space to be 5.2 K. We’d say the universe was twice as hot and that each photon was twice as energetic. A galaxy at this time would be 8 billion years younger. If we observed this same galaxy today, we would see it as much younger *and* the wavelengths would all be a factor of two longer because of the expansion of the universe.

The CMB defines an absolute cosmic reference frame. There is no special place in the universe, but at any place or any time, we can measure our motion relative to the CMB. This does not violate any laws of physics.

When we measure the CMB, where does the light come from? Let’s go back to Figure 7. The CMB light that lands on our detectors has been traveling to us since just after the Big Bang. It started on its path towards us before there were stars or galaxies and of course before the Earth existed. It is light from the edge of the observable universe. When it started on its path to us, it was young, was more energetic, and was described by the Planck function but with a much much higher temperature. On its way to us, the universe expanded, the wavelengths stretched, and the radiation cooled. We now see the remnant glow of the Big Bang 13.8 billion years ago in a conceptually similar way as we see the supernovae light from stars that no longer exist.

10 Dark Matter

If you looked up in the night sky and saw that, over a period of time, a distant star was following a circular path of, say, two full moons (a degree) in diameter, you would immediately conclude that it was in orbit around another object. For an object to move in a circle there *must* be a force on it. For the cosmos, that force is gravity. You might then train your telescope to look for the companion. It might be a black hole or a dim star that you had
not at first noticed but you know that something has to be applying a gravitational force on the star. There has to be some “missing matter.”

For many decades astronomers have been making observations like the one above, in spirit, although in much more clever ways with different systems and with less obvious geometries. In the cosmological context, the existence of missing matter was first proposed by Fritz Zwicky in 1933 based on observations of the Coma cluster of galaxies. Others extended his findings. Of particular note were observations of the orbital velocities of stars and star-forming regions as well as of diffuse hydrogen gas in the Andromeda galaxy, an excellent laboratory because, as shown in Figure 3, it is nearby and looms large. In 1970, Vera Rubin and Kent Ford showed particularly clearly that their stellar velocities agreed with earlier measurements of the diffuse hydrogen gas velocity. Models of the orbits of the observed stars and gas in Andromeda showed that there had to additional matter that was not in either luminous stars or diffuse gas in order to explain the velocity profiles. More generally, regardless of the size of the system, from nearby in our galaxy to distant galaxies and groups of galaxies astronomers found that there is not enough observable matter to account for the motions of stars and galaxies.

The amount of missing matter is not small nor is its effect subtle. By observation, there is more than about five times as much missing matter as observable matter. The best accounting of it comes from measurements of the CMB anisotropy which we discuss in Section 13. Here, though, we focus on its characteristics that are independent of the CMB.

The path from not being able to find the missing matter to concluding that there must be a new form of matter, or “dark matter,” has involved thousands of researchers and multiple lines of evidence. So far we have been able to tie our developing model directly to observations and basic physics. In the interest of brevity, here we will just have to state the results from the interpretation of many astronomical observations. In part the missing matter has been characterized by the process of elimination. We know what it is not. We know it cannot be an assemblage of planets, say “Jupiters,” that are just to dim to see. We know it cannot be atomic, such as hydrogen gas, or be the same as the stuff of which we are made. We know that it cannot be black holes. We know that it cannot be one of the three types of neutrinos even though there are almost as many neutrinos in the universe as CMB photons.

One assumption is that the missing matter is a new type of elementary particle, but it may just as well be a new family of particles, multiple families, or a combination of different types of particles. Generically we call these possibilities “dark matter.” If dark matter is a particle, we do not how it interacts with other particles or even, if two dark matter particles collide, with itself. We know that it does not interact significantly with photons which is why it is called “dark.” The dark matter apparently only interacts gravitationally. It’s character is a grand mystery. Yet it is unambiguous that dark matter exists in vast quantities and that it is not a form of matter we have encountered in our laboratories.

The search for dark matter is a very active area of physics. There are experiments around the world trying to detect it directly. There have been hints of possible detections and reported detections that have not withstood further scrutiny. As of 2018, there are no iron-clad direct detections. There are hopes that dark matter particles will be detected in the Large Hadron Collider.
The discovery of new elementary particles has mostly taken place in particle accelerators that were precursors to the Large Hadron Collider. There is an enormously successful “standard model of particle physics.” This model has seventeen different fundamental elementary particles, including the quarks that make up the protons and neutrons, the electrons and neutrinos, and most recently the Higgs. Although comprehensive, predictive, and well tested, we know the standard model of particle physics is not complete because there are measurements of elementary particles that it cannot explain. We can hope that the detection and characterization of the dark matter in the lab will show us how to advance our model of particle physics.

Might it be that there are no dark matter particles and that our laws of physics are incomplete? A lot of research has gone into investigating how general relativity might be wrong on large scales and so that, in fact, there is no missing matter and instead a new force accounts for the observations. These new theories generally go under the name MOdified Newtonian Dynamics or MOND. Fortunately, MOND makes predictions that can be tested and some of the predictions do not agree with observations. In contrast, there has yet to be an observation in disagreement with the General Theory of Relativity. Therefore MOND is disfavored by the large majority of cosmologists. Of course it could well be that there are other forces or laws of physics we simply have not yet discovered.

11 The Accelerating Universe

In Section 4 we discussed how the universe was expanding and gave the current rate as 15 miles per second per lightyear. In other words, a galaxy 10 million light years distant appears to be moving away from us at a speed of 150 million miles per second. In the late 1990s it was discovered that the expansion rate is increasing. In other words, the expansion is accelerating. In 1 billion years that same galaxy will move away at 156 million miles per second; one billion years ago it was moving away as 144 million miles per second.

This remarkable observation was made by two independent groups, the Supernovae Cosmology Project and the High-Z Supernovae Search Team and has been confirmed by others. As their names suggest, they used supernovae to look back to when the universe was just a few billion years old and about twice as compact. The trick was to find objects for which they could determine accurate distances and velocities and compare the expansion rate then to the current expansion rate.

The way to think about this observation in the context of Section 4 is that space is being made at an accelerated rate. Not only is the concept of “making space” everywhere a convenient way of thinking about the expansion, but we are now forced to think about it this way. In a static space, we can imagine that two galaxies could be moving apart at a constant velocity but there is no way for us to come up with a way for them to accelerate away from each other. Acceleration requires a force\(^\text{13}\) and in a static space the only force available is gravity which, if anything, would tend to decelerate the expansion.

So the question is, why is space being made at an accelerated rate? We do not know. What it means is that space, the vacuum, appears to have an energy density associated with

\(^{13}\)This is Newton’s second law.
it. This energy density acts like a pressure that expands the universe. The energy density is quantified as a cosmological constant denoted by the Greek letter “Lambda,” Λ. This is a new constant of Nature that may, in fact, not even be constant.

Einstein introduced Λ in 1917, before Hubble’s observations. He thought the universe was static, that is, not expanding. To see the motivation imagine two isolated galaxies in the universe. They are attracted to each other by gravity and would fall toward each other. The cosmological constant provides a new kind of pressure that holds them in place. By extension, this would apply to a universe full of galaxies. After Hubble’s realization, Einstein abandoned Λ. We now know this pressure exists but at a smaller value than Einstein thought required.

There are other explanations for the accelerated expansion that the cosmological constant. In general they posit some form of “dark energy” that is not constant. These alternatives make predictions for the acceleration versus age of the universe. Measurements are in progress to test these predictions. We do not know if the dark energy is, say, a substance, or if it is constant throughout space. It may be that there is some fundamental element of one of our theories that we are missing. At the moment, the most straightforward explanation that agrees with all the data is that space is described by a cosmological constant that is constant in space and over time. We adopt this point of view.

The mere existence of the cosmological constant is deep. It is not part of any fundamental theory in physics. It has no bearing on life or physics on earth. There doesn’t appear to be an need for it. There are no laboratory experiments that have been thought of that can measure it. It is a constant that allows us to quantify the cosmic acceleration and in so doing tells us, to reiterate, that there is an energy density or pressure associated with space.

What does this mean for the future? We will set aside the cautionary words in the introduction and extrapolate. If you are in your car on the highway and accelerate at a constant rate you of course go faster and faster. It is no different with the universe. In other words, the space between galaxies is expanding at an ever increasing rate. Galaxies that are widely separated now will soon apparently be moving apart faster than the speed of light. There is no contradiction with the special theory of relativity which requires only that information and massive particles cannot be transmitted from one place to another faster than the speed of light. For the galaxies, the space in their cosmic neighborhoods is just expanding at an ever increasing rate. No information is being transmitted over cosmic distances. From the point of view of someone on the Milky Way, distant galaxies that we can currently observe will no longer be visible in the future.

12 Structure Formation and the Cosmic Timeline

The night sky is full of magnificent “structure” ranging in size from planets to stars, to galaxies, to clusters of galaxies. By structure we mean objects that are held together by gravity. The space between the objects is vast and cold. In contrast, the early universe is a near uniform primordial soup of hot thermal radiation (CMB photons), electrons, protons, neutrons, neutrinos, and dark matter. How did the universe get from one state to the other? In other words, how did structure form and grow? Keeping in mind that this is an active area
of research, we just touch on the well-established key elements of the process as it pertains to the CMB and the standard model of cosmology.

We will pick up the story five minutes after the Big Bang. At this time the temperature of the universe was a little under a billion Kelvin, the diameter of the observable universe was about 5000 light-years, the expansion rate was 100,000 billion times what it is today, and the cosmological constant was irrelevant because its energy density was so much less than that of everything else. The properties of matter at this temperature, roughly five time hotter than the center of the Sun, are well understood. The atomic nuclear composition of the universe was roughly 75% hydrogen and 25% helium by mass (the same as it is today). This ratio was already set by nuclear reaction rates between the protons, neutrons, and neutrinos during the first three minutes. One of the triumphs of the cosmological model is that this composition can be computed based on standard physics applied to the early universe. We return to this again in section 17. These primordial nuclei were in a gas with the electrons and photons but the temperature was too hot for neutral atoms to form. Such a gas is called a “plasma,” sometimes termed the fourth state of matter after solid, liquid, and gas. In the cosmological plasma, for every electron there were almost two billion CMB photons. And, for every proton there was a little over five times the mass in dark matter particles, assuming that there is just one kind of dark matter particle. Other than their participation in nuclear reactions, the neutrinos did not directly participate in the formation of structure at this epoch.

With the ingredients and the state of the universe in hand we now turn to the physical process. Let’s put aside for a moment our concept of an expanding universe. Imagine an infinitely long one-dimensional string of equally-spaced and stationary masses. These masses are attracted to each other gravitationally. Let’s assume that gravity is the only force on them. This configuration is not stable because gravity is only attractive. Pick any mass and displace it ever so slightly to the right. Now, it is closer to its right-hand neighbor than its left-hand neighbor. Because the gravitational force is inversely proportional to the separation squared, the attraction to the right is even stronger than the initial attraction to the left: the mass and its right-hand neighbor fall towards each other. Once the spacing is changed anywhere, the whole string becomes unstable and the masses start clump together. The physical process behind the formation of cosmic structure is gravitational instability. We need something to start the process going, a “seed,” but once it gets going, a once-uniform gas of dark matter and plasma can form structure. Of course the one-dimensional string of masses is overly simplified. In the full picture we have to consider all the constituents in a rapidly expanding universe. Now let’s go through the process. We will put off the source of the seeds until section 16.

The seeds of structure formation initiate the clumping of the dark matter, but for the first fifty thousand years the universe is expanding too rapidly for structure to form. In our one-dimensional analogy the masses start to fall together but the universe is expanding too quickly for them to clump. During this time, the electrons and the CMB photons strongly interact. They constantly scatter off each other. The negatively charged electrons in turn interact with the positively charged protons (the hydrogen nuclei and in the helium nuclei) because opposite charges attract. These interactions are much stronger than the force of gravity and so even if the universe were not expanding so rapidly the plasma would not
clump. The electrons and thus protons are kept from clumping by the intense interaction with the radiation.

As the universe evolves, the expansion rate decreases and the radiation cools. Soon after 50,000 years, the expansion rate is slow enough for the dark matter to start clumping but it is still too hot for the plasma to clump.

After 400,000 years the universe cools to the point where hydrogen atoms can form. In a relatively short time, the electrons bind to protons. Once this happens, the electrons are no longer free and can no longer scatter the CMB. Without the photon scattering, the hydrogen can begin to clump. (The helium undergoes a similar process but slightly earlier.) But there are already large agglomerations of mass because clumps of dark matter have been forming. The atoms fall into the dark matter structure.

The time at which hydrogen atoms form is called “decoupling” because the CMB photons decouple from the atoms. The photons are then free to roam the universe. To a reasonable approximation, the photons that land on our detectors were last scattered in this process and have traversed the radius of the observable universe to get to us. Thus they bring to us a picture of the universe from 13.8 billion (minus 400,000!) years ago, very much like the light from a distant galaxy brings us an image of the galaxy in its youth. The main difference is that the CMB comes to us from a time before stars and galaxies, from a time when matter was just beginning to clump to form structure. This is why the CMB is sometimes called the “baby picture of the universe.”

After decoupling the universe is neutral and enters a period somewhat playfully called the “dark ages,” as in Figure 7, because there were no stars to shine and the CMB had been cooled enough from the expansion that it did not emit visible radiation. During this time, the atoms continued their fall into concentrations of dark matter. The clumping triggered by the gravitational instability took place on all scales, from stellar sizes to huge filaments containing countless protogalaxies. The first objects to form, though, were stars. They lit up the universe, ending the dark ages.

Star formation took place about 200 million years after the Big Bang. The first generation of stars was made of hydrogen and helium but in their cores they produced heavier elements such as carbon, nitrogen, and oxygen through the process of nuclear fusion. These stars aged and exploded in supernovae, spewing the heavy elements throughout the universe. We are made of these heavier elements. There are on-going searches to identify remnants and signatures of these stars; some could even have become black holes. Nevertheless, we know they have to exist because we see their ashes. More recently formed stars, such as the Sun, contain elements in their surfaces heavier than helium and these elements could not have been formed prior to the first generation of stars in the quantities we now observe in stars. As Joni Mitchell sang in 1969 “we are stardust, billion year old carbon...” Our knowledge of the universe has increased enormously since then but “we are stardust, 13.6 billion year old carbon” doesn’t have the same ring.

We also know that the first stars produced enough energy to tear the electrons from the hydrogen nuclei (the protons) by bombarding them with energetic photons. The process

---

14 The Sun is 4.6 billion years old and is expected to stay in its current form for roughly another 5 billion years. It is not massive enough to go through a supernovae.
of ripping an electron from its nucleus is called “ionization.” Thus the universe began as an ionized plasma with no structure, became a neutral gas of hydrogen and helium after decoupling, and then was reionized by the first stars predominantly at an age between 500 million and one billion years. By this point though, the universe had expanded enough, and the CMB was cool enough, that structure could continue to form. However, the newly freed electrons scattered roughly 3%-8% of the CMB photons, an effect which is observed in the CMB and described in Section 17. As with the formation of the first stars, the process of reionization is complicated and not yet well understood. Groups around the world are trying to measure the remnant light from the process. In any case, we know the process took place because we observe that intergalactic space is ionized today.

As the universe ages, new stars come on the scene, galaxies begin to form, and clusters of galaxies begin to grow. The largest structures are still forming today. Although we have laid the process out sequentially, all three processes are taking place to varying degrees at the same time. Telescopes of all kinds are mapping out the process by looking at different structures from different epochs.

In preceding sections, we described specific observations and how they are part of an overall cosmological model. In this section we have discussed how the expanding universe and the cosmic constituents work together to produce structure. The same model that describes the universe also specifies how all the components act together. The model is predictive and there are multiple on-going efforts to check those predictions. If gravity acts differently than we think, if the cosmological constant isn’t constant, if we don’t have the correct ratio between protons and dark matter, if neutrinos play a large role in structure formation we can see the effects through detailed measurements of how cosmic structure grows over time.

13 The CMB Anisotropy

We are now in a position to understand the second aspect of CMB mentioned in Section 9, namely the small temperature differences from position to position on the night sky. This is called the temperature anisotropy. The word “isotropic” means “having a physical property that has the same value when measured in different directions.” Anisotropic means not isotropic. The CMB is not isotropic, but the difference in temperature for different directions in the sky is tiny, typically one ten thousandths of a Kelvin.

The CMB anisotropy has been measured with exquisite precision over the entire sky by the WMAP and Planck satellites. The maps are usually shown in a Mollweide projection. This just specifies the manner in which you represent something that is intrinsically a spherical shell, like the Earth’s surface, on a flat piece of paper. Figure 9 shows the Earth in a Mollweide projection. The equator runs horizontally along the middle of the map, the North Pole is on top and the South Pole is on the bottom.

Figure 10 shows maps of the CMB anisotropy from both WMAP and Planck. Whereas Figure 9 is made looking down on Earth from space, the images in Figure 10 are made looking up into the sky. These maps are oriented so that their equators are aligned with the
Figure 9: A Mollweide projection map of the Earth’s surface. Image from http://upload.wikimedia.org/wikipedia/commons/9/9e/Mollweide.projection_SW.jpg

The center of the map corresponds to the center of the galaxy; the top is the “north galactic pole” and the bottom is the “south galactic pole.” The map from WMAP was made at a wavelength of 0.5 cm (5000 microns). The one from Planck was made at a wavelength of 0.2 cm. It is higher precision than WMAP’s but the similarity once one moves away from the galactic equator is striking.

Neither WMAP nor Planck measures the absolute temperature of the CMB. They can only measure deviations from the average temperature. The color bar shows the magnitude of the deviations. Some parts are hotter than the average, the reddest places are about 2.7253 K or more above absolute zero, and some are colder, the bluest, places are about 2.7247 K or less above absolute zero. We say “or more/less” because those colors are at the ends of the color bar. The broad hot stripe down the equator of each map is the emission from the Milky Way at these wavelengths. The map on the frontispiece has had a model of this “galactic emission” subtracted. We show the full picture here so that the next time you look at the Milky Way you can think about it at longer wavelengths and picture its relation to the CMB anisotropy.

To analyze the maps for their cosmological implications, the galaxy is masked out. In these figures you can do that in your mind’s eye by excluding regions between latitudes plus and minus 20° as indicated by the dashed lines on the top map. Northward and southward of this region, the maps show a seemingly random array of hot and cold regions of various irregular shapes and sizes.

The pattern of hot and cold regions is the CMB anisotropy. It is measured with high precision and is the same as determined with two completely independent satellites. As discussed in the previous section, these maps give a picture of the edge of the observable universe from 400,000 years after the Big Bang. Although the universe went through decou-

---

15 The central horizontal lines in the top and bottom of Figure 10 correspond to the line marked as the galactic plane, or GP, in Figure 1 or to the central horizontal line in Figure 2.

16 These maps have had a pattern called the “CMB dipole” removed. This pattern results from the velocity of the satellite with respect to preferred reference for the universe described in Section 9.
pling throughout its entire volume at this time, we can think of this radiation as coming to
us from a surrounding surface because that is the region from which the CMB we now mea-
sure originated. The region is sometimes called the “decoupling surface” to remind us that
the light we detect decoupled from the primordial plasma there. In Figure 7 the decoupling
surface is the outer-most shell.

What do the hot and cold regions show us? They are a map of the gravitational landscape
of the universe just 400,000 years after the Big Bang. In Section 12 we gave a simple one
dimension picture for how mass clumps to form structure. In the universe that of course
happens in three dimensions. When the mass clumps, the strength of gravity in that region
of space is stronger than in others. For example, if the Earth were the same size but more
massive, we would weigh more because the force of gravity would be stronger. Similarly, the
more mass that clumps into a fixed volume, the greater the strength of gravity. We call this
variation in the strength of gravity throughout space the gravitational landscape. In turn,
the gravitational landscape produces the temperature differences in the map as we describe
next.

It is sometimes easier to think of the clumping on a two-dimensional slice through space.
In two dimensions we can imagine a section of land with hills and valleys of all sorts of differ-
ent widths and heights. The different heights of the land represents the different strengths of
gravity. Gravity is stronger in the valleys than on the hill tops. As the universe evolves prior
to decoupling, the dark matter clumps, and the valleys get deeper. The plasma of photons,
electrons, and nuclei tries to fall into the valleys but it is so energetic that it doesn’t clump.
Instead, it sloshes around somewhat like a jigger of water thrown into a shallow cereal dish.
Unlike water, the plasma is compressible. When it falls into a valley it compresses, heats up,
and then bounces back. The whole universe at this time is full of a compressing, rarefying,
bouncing, jiggling, oscillating plasma that is trying to collect in the valleys but can’t quite do
it. Then, in a relatively short time around 400,000 years the universe cools enough for atoms
to form, the CMB is set free, and the plasma state ends. This is the decoupling discussed
in the previous section. The CMB records the state of the universe at this time. To a good
approximation, the hot regions, the red in the maps, tell us the locations of the gravitational
valleys and the blue regions tell us the locations of the hills. The CMB gives us a snapshot
of the primordial gravitational landscape. The freed up atoms then go on and respond to
this landscape to form the cosmic structure discussed earlier.

This process was first spelled out in the 1970s by Jim Peebles and Jer Yu and in a related
paper by Rashid Sunyaev and Yakob Zel’dovich. The model was refined and augmented over
the decades but it is a testament to the universality of physics that such predictions can be
made and tested. The physics behind the hot and cold regions is more involved than we
described above, but the basic picture we presented is the dominant one that gives rise to
the features in the CMB maps.

Figure 11 shows a zoomed in picture of the region in the small grey box in the top map
of Figure 10. It is about eight full-moon diameters on a side. Although the hot and cold
regions are splotchy and irregularly shaped they do have a characteristic size. For example,
the image does not look like a pointillist painting made up of thousands of tiny patches of
color. Nor are the color patches so large that they take up half the image. The characteristic
patch size looks like roughly twice the diameter of the full moon.
Figure 10: Maps of the CMB anisotropy and galactic emission. Top: The Planck map at a wavelength of 0.2 cm. Radiation from the Milky Way is primarily between the dashed lines. Most of the signal above and below these lines is the CMB anisotropy although in a few places the Milky Way emission pokes through. The little square box on the left just above the top dashed line is centered on the North Star. A zoom in of the boxed region is shown in Figure 11. Bottom: The WMAP map at a wavelength of 0.5 cm. Even though this is a much different wavelength and from a completely different satellite, all the features are the same as the Planck image once you get away from the Milky Way. The temperature color scale runs from -300 millionths of a degree to +300 millionths of a degree. The “µ” sign means “millionth.” These are both Mollweide projections of the full sky. The analysis of these images is the foundation for this essay.
Figure 11: *Left:* A zoom in of a $4^\circ \times 4^\circ$ section of the Planck map in Figure 10 centered on the north celestial pole. Think of this the next time you look at the North Star. For scale, the white circle shows the size of the full moon. Of course, the full moon is not near the North Star. The characteristic sizes of the red and blue regions, in a rough sense, are about twice the size of the full moon. *Right:* The average of over 10,000 hot (or red) spots in the Planck map also shown as a $4^\circ \times 4^\circ$ image. The irregularities of individual spots average out.

Why is there a characteristic splotch size? Let’s go back to the gravitational landscape. The primordial plasma will be hottest when it is compressed the most. This happens when the plasma flows down the valley walls from all sides just once from the time between when it starts flowing, roughly 50,000 years after the Big Bang, and decoupling. The flow speed of the plasma is fixed by fundamental physics. The time over which it can flow is set by the expansion of the universe because after decoupling there is no longer a plasma and the CMB is set free. The flow speed multiplied by a time is a distance. Therefore there is a special size of valley that is especially effective at creating hot splotches. Similarly, the same special size of hill makes the cold splotches. The speed of the plasma and the optimal valley size may be computed theoretically to high accuracy using conventional physics. To be sure there are valleys of all sizes and depths but the CMB highlights a special size and this size may be computed in units of light-years.

We now return to the maps to make quantitative sense of them. Somehow we need to quantify the randomly spaced set of hot and cold regions of different temperatures and irregular sizes. Mathematically, these maps are two-dimensional sets of random numbers on a sphere. Over the decades a few methods have been developed to characterize such sets. We will look at two.

The first method is simple and powerful. We will focus on the hot regions but the same applies to the cold regions. We simply go through the map and everywhere there is a hot spot we extract a four degree by four degree section of map centered on the hot spot. (We’ll use the notation $4^\circ \times 4^\circ$.) We have to be a little careful to make sure we do not double count
but we can experiment and devise an algorithm for that. In Figure 11, I get that there are about a dozen hot spots. Over the galaxy free part of the sky, there are about 10,000 hot spots. We then take all those $4^\circ \times 4^\circ$ maps and average them together, or in other words stack them on top of each other.

Figure 11 shows the average hot spot map for Planck. This is an amazing picture. It is showing us that special valley size in enormous detail. By eye it looks to be between roughly two full moons across. A more detailed analysis gives the angular diameter as $1.193^\circ$. This along with the CMB temperature are the two most precisely measured numbers in cosmology. We will return to the cosmological implications of this measurements in section 15.

The second method for making sense of the maps is more involved but it displays the details of the spot more clearly. The method results in a plot called a “power spectrum” as shown in Figure 12. An earlier version appeared in a front page article in the New York Times in February 2003. The plot tells us the magnitude of the fluctuations in the map for different angular sizes. We already know from the above that the largest fluctuations in temperature will be for regions around a degree in size. This corresponds to the maximum in the plot near $1^\circ$.

Operationally here is how to make the plot. The following details are not important for other sections but hopefully they give more insight into Figure 12. Take a map, cut out the region contaminated by the Milky Way, and then cut the remainder of the map into circles say $8^\circ$ across (16 full moons). For each of those $8^\circ$ diameter disks, compute the average temperature. Of course, in an $8^\circ$ disk there will be lots of smaller hot and cold regions but they will average out. You thus end up with a set of average temperatures for all the $8^\circ$ disks. Some will be hotter than zero, others colder. We don’t so much care about the average of those temperatures; we just want to know how much they scatter around zero. One common way to determine this is to square the temperatures of the disks, because this makes them all positive, and then average those. This is called a “variance” and is why the $y$-axis of the plot has units of $(\mu K)^2$. Now, repeat the process for a list of say fifty disks ranging in size from $32^\circ$ in diameter all the way down to $1/8^\circ$ in diameter. Lastly, because a smaller disk will have all the variance of the next larger disk and then some, you have to go through and subtract entry 49 from entry 50, subtract entry 48 from entry 49, etc. You’ll end up with a new list that has the variance associated just with each disk’s angular size. You then plot the entry in the new list on the $y$-axis and the disk diameter on the $x$-axis and the resulting figure will resemble Figure 12. The actual procedure is more accurate and makes better use of all the data in the map but conceptually does the same thing.

In Figure 12 we see there is more going on than just the degree-scale fluctuations. The other ups and downs come from different ways the plasma oscillates and interacts with the gravitation landscape. For each data point in Figure 12 there is an uncertainty represented by the vertical “error bar.” The smooth line that goes through the data is the standard model of cosmology. You can now see why the CMB anisotropy is so powerful. The measurements are extremely accurate and highly constraining. Any potential theoretical model of the universe has to fit these data. If the model doesn’t fit, it is ruled out. If a model cannot make a prediction for this plot it is not a contender. You can also see why cosmologists are confident that we have the basic picture correct even though we don’t know all the elements of the model in depth. After describing how we measure the CMB in the next section, we
Figure 12: The power spectrum of the CMB anisotropy from WMAP and Planck. On the $y$-axis is the variance of the fluctuations and on the $x$-axis is the angular size. The angular scale markers are not evenly spaced and decrease by half with each tick. The maximum is near $1^\circ$ roughly corresponding to the diameter of the hot spot in Figure 11. The red line shows the best fit model based on the parameters in Section 17. One way to think about this plot is as a graphic equalizer for a fancy audio system. The map is the music. The base notes are on the left side and the treble notes on the right arranged similarly to a piano keyboard. The $y$-axis then corresponds to the loudness of each audio frequency or pitch. If the peak near $1^\circ$ corresponds to middle C at 261 Hz, the second peak would be at 635 Hz or just below E in the next octave higher, and the third peak would be at 963 Hz in this same octave but just below B. The plot shows us the music of the cosmos. From the positions and amplitudes of the peaks one can determine the composition of the universe.
discuss how the model relates to the curve.

Before moving on, let’s do a thought experiment to put our map of the CMB anisotropy in perspective. Imagine that you were instantaneously transported to the edge of the observable universe and looked back towards Earth. What would you see? Around you the galactic environment would look similar to the environment around us now. Recall that at any fixed age, the universe looks the same everywhere. Your Hubble patch would be different, you would see galaxies that we cannot see from Earth, but it would look the same on average. As you looked back towards Earth and our local group you would see a CMB hot region because we know matter had to clump near us to make the local group. But, you would not see any galaxies because the light from them would not have reached you.

So, where did the Big Bang happen? It happened everywhere at the same time. About 13.8 billion years ago there was a hot primordial plasma where we are today. Decoupling took place here as well, only the CMB photons that last scattered here have had 13.8 billion years to travel to the edge of the observable universe. There was a valley in the gravitational landscape right here so that the local group could form. The same processes took place everywhere in space at the same time, only some locations are in the bottoms of valleys, others at the top of hills, and most others somewhere in between. All the processes in the cosmic timeline take place right here at a time you would measure on your watch.

14 How Do We Measure the CMB?

The CMB has been measured with many different methods over the years. The promise of what we might learn has driven the development of multiple new technologies and impressive instruments for measuring both the absolute temperature, the 2.725 K, and the anisotropy. In this section we will focus primarily on the latter. In the late 1960s, the skies were scanned with single room-temperature detectors to look for temperature differences of roughly a thousandth of a Kelvin. We have now advanced to instruments with thousands of detectors cooled to 0.1 K that run around the clock and measure CMB temperature differences at the level of a millionth of a Kelvin or better. The experimental challenge for the state of the art is to measure these minute differences from position to position on the sky while observing with an instrument that is typically 300 Kelvin, almost a billion times hotter than the signal. The steady advances in techniques and technologies that enable us to do this have been profound and sustained.

The CMB shines over a broad range of wavelengths although it is strongest near 0.1 cm as shown in Figure 8. Between the wavelengths of 30 cm and 0.05 cm it is brighter than anything else in the sky if you look away from the galactic plane and observe from above the Earth’s atmosphere. Especially at the wavelengths shorter than 0.3 cm, water vapor in the atmosphere can make the measurements from low elevation sites difficult if not impossible. This has driven researchers to take their instruments to high and dry locations, for example White Mountain in California, the Chilean Andes, and the South Pole, or to fly them on balloons. However, the ultimate platform for measuring the CMB is a satellite.
The first satellite\(^\text{17}\) dedicated to measuring the CMB and the infrared emission was NASA’s COBE, which we discussed above. After that came the Wilkinson Microwave Anisotropy Probe (WMAP) and most recently Planck. We will focus on these last two as they have given us the best and most complete picture of the CMB anisotropy, and have done so in much different ways. They are shown in Figure 13.

The CMB was discovered by Arno Penzias and Robert Wilson at a wavelength of 7.4 cm. This is in the “microwave” band and hence the name CMB stuck even though most of the emission is close to 0.1 cm. Because 7.4 cm is a relatively long wavelength it is not so affected by the atmosphere. Other common devices in the microwave band include TV stations (Channels 2-83, with wavelengths 500 cm to 34 cm) and microwave ovens (12.2 cm). Modern TV satellite dishes operate near a wavelength of 1 cm. You can see in Figure 13 that WMAP looks like it has two back-to-back TV satellite dishes. This is no coincidence! It operates between the wavelengths of 1.3 cm and 0.3 cm in five different bands.

To get a better sense for how the measurement is done, we can think about an old-fashioned TV. Let’s say you had the type of antenna that attaches directly to your TV. If you then tune to Channel 83 and there is no broadcast there you will just get fuzz or noise on your TV screen. This fuzz comes from two sources. It is a combination of microwaves coming into the TV from the environment through the antenna plus noise from the electronics inside the TV. Consider the antenna component of the fuzz. The incident microwaves move the electrons in the antenna structure. In turn, those electrons tickle the input of the transistors in the TV receiver and the rest of the TV receiver amplifies and packages the signal so you can visualize it. The CMB enters the antenna just like a broadcast signal would but it looks like noise. To a crude approximation, about 1% of the total noise on the TV screen is from the CMB that enters through the antenna.

To measure the anisotropy you would point the TV antenna in a specific direction and record the amount of fuzz say by taking a picture of it. Without changing anything on the TV, you’d then point the antenna to a new location and again record the loudness of the fuzz. The difference in the amount of fuzz directly corresponds to the difference in temperature of the radiation coming into the antenna.

It is somewhat counter intuitive but it is not necessary to make your instrument colder than the CMB to measure it. The key is that the electrons in the detecting elements are free to respond to the CMB and there are a number of ways to accomplish this.

It is easy to imagine how to improve the measurement. You’d definitely want to get a less noisy TV receiver so that more of the noise came from the CMB. This could be accomplished by cooling the transistors to reduce their electrical noise. You could increase the signal by listening in multiple channels at once. You’d want to make sure that your antenna only accepted TV waves from directions close to where you were pointing it. You’d want to make a TV that worked at shorter wavelengths to listen at shorter wavelengths etc. In effect WMAP does all of this although with very fancy transistors and a lot of attention to the details.

WMAP’s other key feature is that it accepts radiation from two different directions simul-

---

\(^{17}\)Physicists in the Soviet Union mounted a CMB radiometer on the Relikt satellite that came close to detecting the CMB anisotropy.
Figure 13: The WMAP satellite is on the left with perspective (top) and side (bottom) views. The Planck satellite is on the right. For scale, the large reflector or “dish” on WMAP, the grayish disk, is 140 cm by 160 cm. For Planck it is 150 cm by 190 cm. Thus the satellites are comparable in size with overall heights about 300 cm. In the bottom figures the Sun is in the direction of the bottom of the page. The thermal shielding allows WMAP’s primary reflector to cool to 60 K and Planck’s to 40 K. For both satellites, the detecting elements are just below the primary reflectors. For WMAP, the detectors are passively cooled to 90 K by radiating the heat from the detectors with the large tan-colored ears. Planck uses an active system to cool the longer wavelength set of detectors to 20 K and the shorter wavelength ones to 0.1 K. Note that WMAP uses two back-to-back dishes to collect the radiation whereas Planck uses a single-sided input.
taneously and compares them. This is why there are back-to-back dishes. The instrument is not capable of measuring the absolute temperature, only a large collection of temperature differences. Then, sophisticated algorithms take all those difference measurements and combine them to produce a map of just the fluctuations as shown in Figure 10.

The Planck satellite, shown in Figure 13, takes a different approach. Here there is just one receiving dish and the idea is to spin the satellite, find the average over the spin, and look at temperature differences around that. In total there are 72 channels peering at the sky at any time as opposed to WMAP’s 20 and so there is quite a bit of redundancy. Planck has two instruments that measure between the wavelengths of 1 cm and 0.035 cm in nine different bands. The lower three bands are similar to WMAP’s but the upper six are different and use a different technology called bolometry.

A bolometer is a device that is a lot like a thermometer. It simply measures the amount of thermal energy dumped onto it. The key to using them is to isolate them so well that only the radiation you intend gets to them. The ones on Planck are cooled to 0.1 K above absolute zero. In one second they can measure a temperature difference smaller than one ten thousandth of a Kelvin. They are fantastically sensitive.

Both satellites observed from “the second Lagrange point” or L2. This is a location in the solar system that is about a million miles away from the Earth in the direction opposite the Sun. In 1772 Joseph-Louis Lagrange found there were five places in the solar system where the gravitational pull of the Earth and Sun balanced in just the right way to make an orbit of constant shape possible. Unlike most satellites, WMAP and Planck orbit the Sun and not the Earth.

From L2, the dishes of both satellites look generally away from the Sun, Earth, and Moon, or at least never get closer than about 100°. This is important because these bodies are hot compared to the tiny temperature we need to measure. Another important feature of L2 is that it is thermally stable. There are no day/night cycles. This stability is important for being able to measure the sky over and oven and then averaging the data together. WMAP observed for nine years; Planck for four years.

Although the function of the satellites is straight forward—they simply measure the radiation temperature of the sky—getting them to work to the limits they have achieved has required unprecedented control of systematic sources of error. The devil is in the details. For example, one has to be sure that a measurement made on one day can be directly compared to a measurement made two years hence at a level set by the fundamental noise characteristics of the instrument. By far the most computationally intensive part of the data analysis is in checking that one understands the instrument and the environmental impacts on it.

One of the recurring questions when measuring the CMB is: “how do you know that you’re really looking to the furthest reaches of light and not just at something in the galaxy or in the local group?” The primary way we have of determining this is by measuring the anisotropy at different wavelengths. Just like Planck’s equation precisely describes the amount of power per wavelength band for a blackbody, it also describes the same thing for the fluctuations. Both Planck and WMAP have multiple wavelength bands that allow us to clearly separate the “foreground emission” from the cosmic emission.

When looking at the Hubble Ultra Deep Field in Figure 4, one might wonder why CMB experiments don’t see all of those galaxies. There are three reasons: those galaxies emit
radiation at different wavelengths, there is a lot of space between the galaxies (most of the image is black), and the galaxies are small in angular extent. When we measure the sky with CMB telescopes with high angular resolution, we can see galaxies and clusters of galaxies and thereby determine their precise contribution to the maps in Figure 10.

Measuring the CMB is now “big science.” Up until the 1990s groups of two or three researchers could make a break-through measurement with inexpensive equipment. Now there are thousands of people in the field and the instruments cost many millions of dollars or hundreds of millions for a satellite. There is already a network of ground based telescopes that are in the process of mapping the CMB over half the sky with even greater precision than Planck or WMAP.

Up until this point we have focused on measurements of the cosmos and on how to think about them in, hopefully, a physically intuitive way. We now change gears and introduce the major elements of the standard cosmological model. This will require more advanced concepts from physics and the reader will have to take more on faith. The reward is that we will reach a description of the six cosmological parameters that characterize our universe and all the measurements so far made of its large-scale properties. We start by considering the geometry of the universe.

15 The Geometry of the Universe

With the maps in Figure 10 and the related ones at other wavelengths we can determine characteristics of the universe. One of the fundamental characteristics is its geometry. Geometry is the study of the relations between points, lines, angles, surfaces, etc. Let’s go back to thinking about space. We have noted that it can expand at different rates. It is also malleable and can be warped or curved. When light from a distant star travels to us on a path that goes near to the Sun it is deflected ever so slightly. This can be thought of as the Sun’s gravitational pull on the light. A better way to look at it is that the space around the Sun is curved and that the light from the distant star follows the easiest path on its way to us. Figure 14 shows one way to visualize this.

It is difficult to think about our three-dimensional space as being curved into a fourth spatial dimension. In the mid 1800s Georg Friedrich Bernhard Reimann showed that even without going to the next higher dimension we can tell if we live in a curved space. To understand his insight we will work with two-dimensional surfaces curved into our familiar three-dimensional space as shown in Figure 15. Imagine that you are an ant walking around on the two-dimension surface taking a trip between the three vertices of a triangle. Think of the surface as being very large, the ant as very small with negligible height, and that all motion is confined to the surface. If you walk the perimeter of any triangle on a flat piece of paper and sum up the interior angles you will get 180°. The piece of paper is said to have a “flat” geometry and the two-dimensional space is infinite so that there are no edges.\footnote{A Möbius strip is an example of a “flat” geometry that is not infinite. It is said to have a non-trivial \textit{topology}. In this essay we assume that universe is characterized by its geometry and not it topology.}

40
Figure 14: An example of curved two-dimensional space. The Sun is represented as the ball in the center of the image. It warps space as would a bowling ball on a large rubber sheet. Light’s path by the Sun is analogous to a quickly rolled small marble going past the bowling ball. The marble follows the contour of the two-dimensional rubber sheet and its trajectory is bent towards the bowling ball, away from a straight path. Similarly, in three dimensions a light ray going by the Sun follows the easiest path. Its trajectory is bent by the curvature of three-dimensional space or, equivalently, by the force of gravity. Image from: http://astronomyonline.org/images/ImagesFromPapers/CurvedLight.jpg
convention, we use the word “flat” even if make the triangles in three-dimensional space with all possible orientations as opposed to on a flat sheet of paper.

![Figure 15: Examples of possible geometries of two-dimensional space. On the left is the saddle-surface-like open geometry. Think of it going on forever. The red lines show a triangle whose interior angles sum to less than 180°. In the middle is a flat sheet-of-paper-like geometry. Think of it too as going on forever. Here the red lines show a triangle whose interior angles sum up to 180°. On the right is a spherical-surface-like geometry. This one is finite. Here the red lines show a triangle whose interior angles sum up to greater than 180°.

Let’s consider instead a spherical shell. This is an example of a finite and closed positively-curved space. The ant could walk a triangular path from the north pole to the equator, around the equator by a quarter of the circumference, and then back to the north pole. Now the ant would find the sum of the interior angles to be greater than 180° and for this particular path the sum would be 270°. If the ant shot a laser beam out in this space it would come back and hit his back side because in two dimensions the laser beam is confined to follow the surface of the sphere.

A horse saddle is an example of an open, negatively-curved space. Unlike the case for the spherical shell, the saddle surface, like the flat piece of paper, goes on forever. If the ant were to walk the perimeter of a triangle on the saddle surface and sum up the interior angles, he would find that they were less than 180°. If we think of the leather as space and if we tried to flatten the saddle onto a flat surface, we would have a bunch of folds left over. An open negatively-curved space is one in which there is more available space, more leather, the further away we go.

This same process of determining the overall geometry by measuring the sum of the interior angles of a triangle works for a three-dimensional space that is curved into a fourth spatial dimension. All we have to do to tell the geometry of our three dimensional space is to traverse a large triangular path and sum up the angles. The CMB gives us a way to do this as shown in Figure 16. The hot and cold spots form one side of the triangle. In effect we average all the spot sizes together. As we discussed in the previous section, we can compute to high accuracy the average physical size of a hot or cold spot in, say, light-years. With our maps we can measure the average angular size to high accuracy as shown in Figure 11 and

---

Predictions from cosmologies with non-trivial topologies can be tested with CMB maps. So far, there is no strong evidence for a non-trivial topology.
12, and depicted in Figure 16. We need to know one more bit of information to completely specify the triangle that has, say, a hotspot as the far side. Although it is not obvious, that piece of information is the Hubble constant because it links the physical size of the hot spot to the distance of the hot spot. The result is that the sum of the interior angles is $180^\circ$. To the limit of measurement, the geometry of the universe is “flat.”

![Figure 16: A depiction of measuring the size of the hot and cold spots in the CMB. We think of the CMB as a spherical surface at the edge of the observable that surrounds us. With our telescopes, we measure the pattern of hot and cold regions and discover that there is a characteristic angular size. This average hotspot size is shown in Figure 11 and corresponds to the peak of the power spectrum in Figure 12. From theory based on well-established physics, we can also compute the physical size of the spots in light-years. By combining the measured angle and the computed size with knowledge of the Hubble constant, we can determine the geometry of the universe.](image)

To summarize, the geometry of the universe is like the geometry many of us learned in grade school. It is the simplest one we can think of. It is what you would have expected if you had never heard of Einstein or Reimann. More importantly, the geometry has been determined by measurement and can be checked with other measurements of the universe at huge distances. In the standard model, we reverse the historical process we just went through, fix the geometry to be flat, and derive the Hubble constant.

16 Quantum Mechanics and the Seeds of Cosmic Structure Formation.

The very very earliest instants of the universe are still not well understood. The reason is that as yet there is no fundamental theory that combines gravity with the standard model of
particle physics. In place of this we have “effective theories” and paradigms that are deeply rooted in the physics we know and that can explain the observations. The best known of these is inflation. We touch on aspects of it now but keep in mind that this is still a very active area of theoretical research.

One of the great mysteries of the cosmos before the inflation model was invented was: “why are the properties of the universe in two opposite directions so similar?” To be concrete, let’s take the north celestial pole and south celestial pole as our opposite directions and consider how the CMB can have the same temperature in the two directions. According to the picture we have been developing, the light from both sides of the observable universe is just reaching us now. No information can travel faster than light and so there is no way that radiation from the north celestial polar direction could have passed us and gone on to affect what we see in the south celestial polar direction, and vice versa. Yet, they are nearly the same temperature, 2.725 K, and have the same properties.

In the inflation model, the space in the very early universe, before there were any particles, had an enormous energy density. Associated with this energy density was a pressure that made space at an unimaginably fast, exponential pace. At the start of the process imagine you have two regions, call them Alice and Bob, that are right next to each other and that share information. In inflation, the space between Alice and Bob is made so rapidly that they can no longer communicate with each other. Their apparent speed of separation is faster than the speed of light. They are separated beyond, perhaps many many times beyond, the distance over which they can affect each other.

Inflation takes place over an extremely short period of time, very roughly a billionth of a billionth of a billionth of a second. After inflation ends, the universe settles into a calmer pace of expansion. As the universe ages, the observable universe gets larger and larger because we can look further and further away. At some point, Alice and Bob come into view with Alice, say, in the north polar direction and Bob in the south. Now we have a mechanism for saying why opposite sides of the universe should look the same. They communicated with each other very very early on, became hugely separated during the inflation epoch, and they are now just coming into our observable universe.

There are many variants of inflation but the simplest model has two other features that are relevant to the CMB. The first is that the universe is geometrically flat to at least a part in 10,000. This corresponds to trying to tell if a meter stick is flat or bowed up at one end by 100 microns. As the inflation model was made prior to the observations, it gained a lot of credibility in that the data showed a flat geometry. The idea is that even if the earliest geometry were, say, positively curved, inflation would have expanded it so much that it would effectively be flat. For example, in two dimensions if you are on the surface of a sphere, such as the Earth, you can tell the surface is curved. But, if the radius were a billion billion times larger, it would be difficult to tell you were on a sphere. To the limits of measurement, our geometry is flat but we cannot rule out the possibility that it is just ever so slightly positively or negatively curved.

The second feature is that inflation incorporates a mechanism for generating the seeds for the formation of cosmic structure discussed in Section 12. The picture is that the seeds are quantum fluctuations in the primordial energy density that were stretched out to cosmic scales through the inflation of space. The fluctuations in the primordial field are now seen as
the gravitational landscape that produces the hot and cold spots in the CMB. This means that when we look at the CMB we are directly looking at a manifestation of quantum mechanics. The random distribution in space of the hot and cold patches is a result of our quantum origins. We usually associate quantum processes as taking place on atomic or sub-atomic scale. This is still true; it is just that inflation expands space so much that the quantum scale becomes the cosmic scale. This is truly mind blowing.

The expansion in inflation is similar in character to that for the cosmological constant discussed in Section 11, but in inflation the pressures are much much greater. It may be that the origin of the processes is related. We don’t know. Also, it may be that inflation is not the correct paradigm. It could be, for example, that the universe goes through cycles of expansion and that we are in just one of the cycles.

17 Pulling it all Together with the Standard Model of Cosmology

While explaining the CMB has guided our path to an understanding of the universe, there are many other ways to study it. Cosmology is a broad field. The physics brought to bear includes everything from the General Theory of Relativity, to thermodynamics, to elementary particle theory. The observations are made in nearly every wavelength regime accessible to measurement and with state-of-the-art particle detectors. The observations come from near by and the furthest reaches of space. All of this evidence and theory is encompassed in the surprisingly simple standard model. Before summarizing the model, we touch on two major frontiers we have not explicitly mentioned.

The most time-honored approach to cosmology is through observations of galaxies. As we saw, this was how Hubble pointed out that the universe was expanding. In addition to telescopes like the Hubble Space Telescope that can peer deeply and with high resolution in a given direction, there are others that measure the properties of millions of galaxies over more than a third of the sky. Probably the best known is the Sloan Digital Sky Survey. From this and related efforts we now have maps of the three-dimensional distribution of galaxies throughout much of the observable universe. We see in detail how galaxies clump. We can see how light from distant galaxies is bent on its way to us by the curved space near the intervening galaxies. By averaging over large volumes we can even see that there is a characteristic size for the clumping of galaxies that corresponds to the average CMB hot spots and cold spots sizes in Figure 11. To emphasize, the signature of the physical processes that produced the hot and cold spots in the CMB is also detected in the distribution of galaxies.

Quite independently of the galaxy and CMB observations, cosmologists have worked out the nuclear physics of the first three minutes of the universe. The study is call Big Bang Nucleosynthesis. The inputs to the calculation are the CMB temperature and nuclear interaction rates as measured in laboratories. The outputs are the abundances of the lightest elements: hydrogen, helium, lithium, and beryllium The main predictions of the calculation are the overall cosmic density of atoms, that the atoms in the universe are primarily hydrogen (75% by mass) and helium (25% by mass) with only trace amounts of the other two elements. The calculation shows that elements heavier than beryllium could not have been formed in
the early universe. For the predicted nuclear abundance to match observations, there must have been about two billion photons per proton. These photons are of course the CMB. The measurements of cosmic abundances of these elements is in excellent agreement with what one would infer from the CMB with one exception, lithium. One finds less than is predicted. It is likely that the early stars eat the lithium but the mismatch between expectations and measurement may be telling us that there is an element of the calculation or model we are missing.

We now summarize the cosmological model. It has six parameters that we describe below. The particular values we give result from fitting the red curve in Figure 12 to the CMB data. The values don’t shift much, and the uncertainties improve, when additional data sets, such as the distribution of galaxies, are combined with the CMB. We give the specific symbols for the parameters as they are often encountered in the scientific literature.

The model stipulates that the universe is geometrically flat.

The first three parameters tell us about the contents of the universe. They are specified as fractions of the total like the components in a typical pie chart.

1. Atoms account for about 5% of the universe. In the CMB anisotropy spectrum, Figure 12, it turns out that the ratio of the height of the first to the second peak gives a measure of the density of atomic nuclei in the early universe. It is not obvious that this should be the case and was only understood after we knew how to compute the curve in the figure. The value from the CMB anisotropy agrees with the value from Big Bang Nucleosynthesis. The fact that the stuff of which we are made accounts for just 5% of the net cosmic energy density gives a new perspective on our place in the universe. We specify this fraction with the Greek letter “Omega” and say $\Omega_{\text{atoms}} = 0.05$.

2. Dark Matter accounts for 25% of the universe. In the CMB anisotropy spectrum, it turns out that the ratio of the height of the first to the third peak gives a measure of the dark matter density. This too is not obvious. The third peak is the evidence that dark matter existed in the early universe. The value from the CMB anisotropy agrees with the value from observations of the motions of stars and galaxies discussed in section 10. So, the stuff of which we are made accounts for just 1/6 the total mass in the universe. We write $\Omega_{\text{DM}} = 0.25$.

3. The cosmological constant accounts for 70% of the universe. We don’t know what it is but we have directly measured its presence through the cosmic acceleration. In the CMB, we determine it from the angular size of first peak. The value from the supernovae observations agrees with the value from the CMB. We write $\Omega_{\Lambda} = 0.70$.

There are of course other components, for example the current energy density of the CMB radiation itself and the mass density of neutrinos. We know they are there but they

---

19The energy density corresponding to the cosmological constant is converted to a mass density by dividing by the speed of light squared according to Einstein’s famous equation $E = mc^2$. 

46
are not significant enough that, with the current level of precision, they need to included in the overall budget for the energy density.

The following is the most astrophysical of the parameters. It captures our rather scant knowledge of the entire complex process of the formation and subsequent explosion of the first stars and the formation of the first galaxies discussed in Section 12. The intense light from these early stars and galaxies broke apart the hydrogen into its constituent protons and electrons reionizing the universe. With the current level of precision, we need just one parameter to account for what no doubt we will find to be a rich process.

4. In the process of reionization about 5-8% of the CMB photons were rescattered. The symbol for the process is \( \tau \) and is called the ‘optical depth.” We say \( \tau = 0.05 - 0.08 \). With the temperature anisotropy data alone it cannot be determined. We also need to measure the polarization of the CMB, a topic we have not discussed. Light reflected off the hood of your car is polarized and polaroid sunglasses block the reflected glare. Similarly, the electrons freed up in reionization scatter and polarize the CMB. If you could look at the CMB with polarized “sunglasses” it would look slightly different. In Figure 12 reionization causes an overall suppression of the spectrum with just slightly more suppression at the largest angular scales. The optical depth is the least well known of the cosmological parameters.

The next two parameters characterize the seeds of the fluctuations that gave rise to all the structure in the universe. The concepts underlying these are involved and beyond the scope of this essay but we include them for completeness. These seeds led to both the CMB anisotropy spectrum and to the fluctuations in the total matter in spheres of 25 million light-years in diameter that were discussed in Section 3. Recall that volumes of space 25 million light-years across give a fair sample of the cosmic constituents. These primordial fluctuations are described by the “primordial power spectrum.” It is similar in character to the CMB anisotropy power spectrum (Figure 12) but instead of describing the decoupling surface it describes fluctuations in three-dimensional space.

5. The amplitude of the primordial power spectrum is encoded in the formidable symbol \( \Delta^2_R \). If we had a complete model of the universe that began with the quantum fluctuations and predicted, say, the fluctuations in matter in spheres of 25 million light years diameter, we could relate \( \Delta^2_R \) to the rest of physics and its value would be known. Unfortunately, while we have a very successful framework, we do not yet know all the connections and so require it as a parameter.

6. The final parameter, called the “scalar spectral index” or \( n_s \), is the most difficult to understand but also our best window into the birth of the universe. Like \( \Delta^2_R \), it tells us about the primordial fluctuations. But, instead of specifying the overall amplitude, it tells us how the amplitude changes with angular scale. We observe that the primordial fluctuations, the “seeds,” were ever so slightly larger in amplitude at large angular scales than they were at smaller ones.
As the process of cosmic structure formation was being understood, the scalar spectral index was argued to be unity, $n_s = 1$, on general grounds. It was called the Harrison-Peebles-Zel’dovich spectrum after its authors. Then, in the early 1980s, it was realized by Viatcheslav Mukhanov and Gennady Chibisov that this quantity could be computed from quantum principles operating as the universe was being born. We now know this index is different than unity by about 5%, that is $n_s = 0.95$. This is the evidence that all the structure in the universe arose from quantum mechanics operating at a time when the universe was so compact and energetic that no known particles yet existed.

With these six parameters we can compute the properties and spectrum (the red line in Figure 12) not only for the CMB but for any cosmological measurement. We can compute the age of the universe. The single most constraining observation is the CMB anisotropy, but the model is consistent with all measurements. In short, no matter how one looks at the cosmos, with galaxy surveys, through exploding stars, through the abundance of the light elements, through the velocity of galaxies, or through the CMB, one needs only the six parameters given above and the physical processes we have described in the preceding sections to describe the observations. This is a truly remarkable state of affairs.

18 Frontiers

The standard model of cosmology is so successful that it is now a foundation from which we can look for departures. Through more precise measurements of the CMB we will learn, for example, the total mass of neutrinos. We might find that there are remnant gravitational waves from the birth of the universe that now pervade the universe. We could perhaps find that the General Theory of Relativity is not complete. Perhaps the universe isn’t quite geometrically flat. Perhaps the fluctuations have a slightly different form and spectrum than we now measure. Perhaps a new particle in the early universe will reveal itself. To be able to tell any of these we need more precise data. We expand on four frontier areas that are especially active.

The neutrino mass. We have mentioned neutrinos a few times already. These are very light neutral particles that are produced in nuclear reactions. They interact negligible with matter. They fly through the Earth with no trouble. From nuclear reactions in the Sun’s core alone there are about a hundred billion neutrinos per square centimeter per second going through us. From the nuclear reactions in the early universe there are a comparable number. We are sieves to these particles.

Until very recently they were thought to be massless. We now know they have to be more massive than one ten millionth the electrons mass, but less than ten times that. Because there are so many in the universe, about 300 per cubic centimeter, they affect how cosmic structure grows and in turn we can detect this with the CMB. At one point it was thought that these neutrinos might be the dark matter but we now know this cannot be the case.

Gravitational waves. In many variants of the standard model there is a background of gravitational waves produced in the early universe. They are another form of the quantum fluctuations. In general these waves are a distortion of space and time that propagate across
the universe at the speed of light. For example, if a wave were aimed at a 100 cm by 100 cm plate, then in one half cycle it would shrink the width and expand the height. A half cycle later it would shrink the height and expand the width. If the change in height was 1 cm we would say the strain is one part in a hundred or 1%. The LIGO detector on Earth detected gravitational waves from a pair of in-spiraling and coalescing black holes that were about 1.2 billion light-years away. The strain they measured was a part in 1 followed by 21 zeros. This is equivalent to detecting a change in distance between us and Proxima Centauri, 4.3 light-years away, with the precision of the width of a human hair. This is a staggeringly precise measurement.

The Big Bang might produce similar waves but with wavelengths ranging in size from about 1 % up to the size of the observable universe. Since the wavelengths are so large, the distortions produced by the waves appear stationary to us. Some current models predict the strain should be about a part in 100,000. This is a far greater strain than detected by LIGO and corresponds to measuring the height of a human to the width of a human hair. The cosmological gravitational waves are most easily identified through a characteristic pattern in the polarization of the CMB called a “primordial B mode.”

A detection of a primordial B mode would be very exciting. It would provide a new and deep connection between the quantum regime of the very early universe and gravity. It would have impact far beyond cosmology.

Structure formation and basic physics. It is one thing to specify the contents of the universe. It is quite another to be able to understand how those ingredients combine and work together over billions of years to produce the universe we observe today. By carefully measuring how mass assembles over the ages we can test to see if the General Theory of Relativity is correct and we can see if the cosmological constant is indeed constant with time. This research is taking place on many fronts.

The temperature spectrum. We noted earlier that if the source of radiation is a blackbody, all you need to do is specify its temperature and you know the intensity at all wavelengths. To the limits of measurement we know the CMB is a blackbody, that is it described by the Planck function. However, if the source is not a blackbody then the effective temperature depends on the wavelength. This could be the case if there was a large injection of energy during the cosmic evolution, say from the decay of some particle, of if the universe evolved in such a manner that the radiation did not have time to come into equilibrium with the particles. There are a number of known effects, such as reionization, that should alter the temperature spectrum at levels a little more than a factor of ten below current limits. This is too small to detect with the current experimental methods but satellite instruments are being designed to search for this and other features.

19 Endnote

It is a remarkable that humankind has arrived at the standard model of cosmology. We cosmologists feel fortunate to have been alive in the decades when the explosion of knowledge took place. It is well within the memory of most in the field of a time when we did not know the geometry of the universe, its contents, or its age. As the data have become more and
more precise, whole classes of cosmological models have been shown to be wrong. As we have emphasized, it is precise measurements that are the foundation of the model. The dramatic advance in cosmology has come through the ability to compare models to measurements. We learn so much because the physics is well understood, and the early universe, as it turns out, is so simple.

We now have this wonderfully predictive model but even within that context there is a lot we still do not know. We are possibly on the trail of the dark matter but the cosmological constant is an even greater mystery. As far as we know there is no “need” for it. Although we talk about “making space” we do not know what space is. The clues may well be all around us but we haven’t thought of them in the right way.

We still do not know what happened at the very earliest times. We do not know the physics. And even though we learn more and more about how the universe began, we of course do not know why it began or even if that is a relevant question. We do not know if there are multiple universes or if we are in just one of an endless series of cycles. We do not know why our universe is predominantly matter as opposed to a combination of matter and anti-matter. We do not know why we have one dimension of time and three of space other than the fact that it is consistent with our existence.

The cosmos have captured the imagination of humans since recorded history. Although the recent advances are dramatic, the quest for ever deeper knowledge on both theoretical and experimental fronts continues. For those observing the cosmos the most exciting thing is to find something new, or to find out that one of the elements of the standard model needs to be looked at in a new light. There is a huge about left to be learned from the CMB and we will likely be measuring it for years to come.
A Appendix A: The Electromagnetic Spectrum

Figure A shows the electromagnetic spectrum over a wide range of wavelengths. The units on the $x$-axis change from cm on the left to microns on the right to connect with the text. Of course, $0.1\text{ cm} = 1\text{ mm}=1000\text{ microns}$. Note that the wavelength gets smaller going to the right which means that the energy of a photon increases in going from left to the right.

Channel 83 on your TV, which is not in general use, has a wavelength of $34\text{ cm}$. Microwave ovens operate at $12.2\text{ cm}$. These are indicated as lines because most of the energy is concentrated near one wavelength. The CMB is a blackbody emitter that peaks near $0.1\text{ cm}$ but emits power over a large range of wavelengths. This is the same spectrum as shown in Figure 8 but now in a broader context. The next mark shows the wavelength for Figure 2. The next blackbody spectrum is for a room temperature blackbody (300 K). IR cameras measure this thermal emission. The spectrum for the Sun peaks around a wavelength of $0.5\text{ microns}$. The colors correspond to the colors of visible light our eyes detect. At slightly shorter wavelengths is the UV spectrum. UV B radiation is at $0.3\text{ microns}$. You can see that the Sun is still quite intense there but we can’t see the UV “light.” Along the top of the plot are the designations for the “Microwave” wavelength band, “Far infrared” band, and “Mid” and “Near” infrared bands.
Appendix B: Expanding Space

“Expanding space” is a controversial phrase. We use it simply as an intuitive description of the change in the scale of the universe with time. We take guidance from Einstein: “In that sense one can say, according to Friedman\textsuperscript{20}, that the theory demands an expansion of space.” and “It is indeed an exacting requirement to have to ascribe physical reality to space in general, and especially to empty space.”

The coordinate system we use to measure the location of objects in the universe is unambiguously expanding. At the same time, for most of the age of the universe, there is nothing that pushes galaxies apart that the phrase “expanding space” might bring to mind. Gravity is only attractive. The evolution of the universe during this time, over regions larger than shown in Figures 5 and 6, can be described by giving the galaxies initial velocities and computing how they interact under the force of gravity.

However, for the past 4 billion years, since \( \Lambda \)-matter equality, a new force has come to dominate the universe that does indeed push galaxies apart. That force is quantified with the cosmological constant. Its action can be described as “making space.” Similarly, if inflation is the correct model for the early universe, it too can be described as “making space” but at an exponential rate over a very brief time. During inflation there is a force that pushes particles apart that is much stronger than gravity. The source of this force is an effective cosmological constant that is much much larger than the one we currently observe.

Here is another example of “making space.” If the universe were described by a closed geometry, corresponding to the righthand image in Figure 15, the volume of the universe would be finite and would change with time. Space would indeed be created.

Understanding the nature of space—the nature of the vacuum—is at the forefront of physics. We do not understand the vacuum or “space” at a deep level. For some situations, we are almost forced to think of space as expanding, for others the expansion of space may lull us into thinking there are forces that don’t exist. Nevertheless, we find the concept of an expanding space useful for envisioning many aspects of the universe.

\textsuperscript{20}Friedmann derived the equations that describe cosmology from general relativity. The quotes are from “Relativity” by Einstein, Crown Publishers, 1961.
## Appendix C: Significant Events in the Cosmic Timeline

<table>
<thead>
<tr>
<th>Age</th>
<th>Compactness Scale factor</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Minuscule!</td>
<td>Our definition of the “Big Bang.” §5</td>
</tr>
<tr>
<td>$1.4 \times 10^{-14}$ sec</td>
<td>$2.2 \times 10^{-17}$</td>
<td>Typical energy of a photon equals the particle interaction energy at the LHC. §9.</td>
</tr>
<tr>
<td>0.000025 sec</td>
<td>$1 \times 10^{-12}$</td>
<td>Quark-gluon plasma as seen at RHIC. §9</td>
</tr>
<tr>
<td>3 min</td>
<td>$3 \times 10^{-9}$</td>
<td>The nuclei of H, He, Li, and Be formed. The temperature was 1 billion K. §9,12</td>
</tr>
<tr>
<td>1 year</td>
<td>$1 \times 10^{-6}$</td>
<td>“Matter-radiation equality.” The dominant form of energy density changes from radiation to matter and cosmic structure can start to grow. §12</td>
</tr>
<tr>
<td>51,000 yrs</td>
<td>0.00029</td>
<td>Appendix D</td>
</tr>
<tr>
<td>400,000 yrs</td>
<td>0.001</td>
<td>“Decoupling.” Hydrogen atoms form and the CMB is free to roam the universe. Some call this time “recombination.” §12. §9,12,13</td>
</tr>
<tr>
<td>1 million yrs</td>
<td>0.0017</td>
<td>First objects form. §8, 12</td>
</tr>
<tr>
<td>200 million yrs</td>
<td>0.05</td>
<td>Most distant object yet identified.</td>
</tr>
<tr>
<td>370 million</td>
<td>0.078</td>
<td>Most distant objects in the Hubble Ultra Deep Field. §8,12</td>
</tr>
<tr>
<td>0.4-0.7 billion yrs</td>
<td>0.08-0.12</td>
<td>“Reionization.” The universe was reionized by the first stars and the free electrons scatter 5%-8% of the CMB photons. §12</td>
</tr>
<tr>
<td>0.5-1 billion yrs</td>
<td>0.1-0.15</td>
<td>Universe is twice as compact. §5, 8,12</td>
</tr>
<tr>
<td>5.9 billion yrs</td>
<td>0.5</td>
<td>Time when the Earth and Moon appeared. §5</td>
</tr>
<tr>
<td>9.3 billion yrs</td>
<td>0.71</td>
<td>“A-matter equality.” The dominant form of effective energy density changes from matter to dark energy.</td>
</tr>
<tr>
<td>10 billion yrs</td>
<td>0.75</td>
<td>Dinosaurs roamed the Earth. §5</td>
</tr>
<tr>
<td>13.7 billion yrs</td>
<td>0.993</td>
<td>We live in a LCDM universe.</td>
</tr>
<tr>
<td>13.8 billion yrs</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

For the extremely small numbers, we have had to introduce scientific notation in which the exponent tells where to place the decimal point. For example, $1 \times 10^2 = 100$ and $1 \times 10^{-2} = 0.01$. The compactness is the number by which one should multiply the scale of the current universe to determine how much closer objects were in the past.
D Appendix D: Size and Age of the Observable Universe

Figure D shows the size versus age of the observable universe. In going from left to right, the vertical dashed blue lines show when cosmic structure started to grow (Section 9), when the CMB decoupled from the primordial plasma (Section 12), and “distance” to the furthest identifiable object.

Here we give a more detailed picture of thinking about distances in the universe. When we see a distant object, often the first question that jumps to most of our minds is “how far away is it?” For the universe, we have to be especially careful in specifying when we want to know how far away it is. As light propagates to us from a distant object, the universe expands. By the time we receive the light, the universe has expanded. For example, the “light travel distance” back to the Big Bang is 13.8 billion light years. However, over that 13.8 billion years the universe has expanded a huge amount so its current “size,” called the “comoving distance” in scientific literature or more popularly the “observable universe,” is 46 billion light years.

The natural way to think about the universe is in terms of its compactness or “scale.” We ask its age and size in regards to when it was 10 or 100 or a billion times more compact. The reason for this is that the fundamental physical properties—the temperature, the
densities, the rate of expansion, etc.—depend on the compactness. Then, from the history of compactness, we deduce the age and size. For example, from the left hand side of the figure we read off that when the universe was one million times more compact, it was 1 year old, and the size of the observable universe was a couple of million light years. We can tell the expansion rate was incredibly fast at this time. In its first year the universe expanded a factor of over a million in size. At this time the CMB was one million times hotter because temperature is directly proportional to compactness.

The current furthest identifiable object is a galaxy called EGSY8p7. The light from it started on its way to us when the universe was about ten time more compact. The light we now observe was emitted when the universe was 0.6 billion years old and so has been traveling to us for $13.8 - 0.6 = 13.2$ billion years. In Figure 7 this is in the purple band that the Hubble Ultra Deep Field can reach. The question of “how far away is it” really isn’t really the right question to ask because the universe has expanded so much since it emitted the light we just now observe.

We could have extended the left hand side of the plot much further. Some earlier times and associated events are given in Appendix C.