

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](https://ocw.mit.edu).

**PROFESSOR:** OK, in that case, let's take off. There's a fair amount I'd like to do before the end of the term. First, let me quickly review what we talked about last time. We talked about the actual supernovae data, which gives us brightness as a function of redshift for very distant objects and produced the first discovery that the model is not fit very well by the standard called dark matter model, which would be this lower line.

But it fits much better by this lambda CDM, a model which involves a significant fraction-- 0.76 was used here-- and significant fraction of vacuum energy along with cold, dark matter and baryons. And that was a model that fit the data well. And this, along with similar data from another group of astronomers, was a big bombshell of 1998, showing that the universe appears to not be slowing down under the influence of gravity but rather to be accelerating due to some kind of repulsive, presumably, gravitational force.

We talked about the overall evidence for this idea of an accelerating universe. It certainly began with the supernovae data that we just talked about. The basic fact that characterizes that data is that the most distant supernovae are dimmer than you'd expect by 20% to 30%. And people did cook up other possible explanations for what might have made supernovae at certain distances look dimmer. None of those really held up very well.

But, in addition, several other important pieces of evidence came in to support this idea of an accelerating universe. Most important, more precise measurements of the cosmic background radiation anisotropies came in. And this pattern of anisotropies can be fit to a theoretical model, which includes all of the parameters of cosmology, basically, and turns out to give a very precise setting for essentially all

the parameters of cosmology. They now really all have their most precise values coming out of these CMB measurements.

And the CMB measurements gave a value of  $\Omega$  [? vac ?], which is very close to what we'll get from the supernovae, which makes it all look very convincing. And furthermore, the cosmic background radiation shows that  $\Omega$  total is equal to 1 to about 1/2% accuracy, which is very hard to account for if one doesn't assume that there's a very significant amount of dark energy. Because there just does not appear to be nearly enough of anything else to make  $\Omega$  equal to 1.

And finally, we pointed out that this vacuum energy also improves the age calculations. Without vacuum energy, we tend to define that the age of the universe as calculated from the Big Bang theory always ended up being a little younger than the ages of the oldest stars, which didn't make sense. But with the vacuum energy, that changes the cosmological calculation of the age producing older ages.

So with vacuum energy of the sort that we think exists, we get ages like 13.7 or 13.8 billion years. And that's completely consistent with what we think about the ages of the oldest stars. So everything fits together.

So by now, I would say that with these three arguments together, essentially, everybody is convinced that this acceleration is real. I do know a few people who aren't convinced, but they're oddballs. Most of us are convinced.

And the simplest explanation for this dark energy is simply vacuum energy. And every measurement that's been made so far is consistent with the idea of vacuum energy. There is still an alternative possibility which is called quintessence, which would be a very slowly evolving scalar field.

And it would show up, because you would see some evolution. And so far nobody has seen any evolution of the amount of dark energy in the universe. So that's basically where things stand as far as the observations of acceleration of dark energy. Any questions about that?

OK, next, we went on to talk about the physics of vacuum energy or a cosmological

constant. A cosmological constant and vacuum energy are really synonymous. And they're related to each other by the energy entering the vacuum being equal to this expression, where  $\lambda$  is what Einstein originally called the cosmological constant and what we still call the cosmological constant.

We discussed the fact that there are basically three contributions in a quantum field theory to the energy of a vacuum. We do not expect it to be zero, because there are these complicated contributions. There are, first of all, the quantum fluctuations of the photon and other Bosonic fields, Bosonic fields meaning particles that do not obey the Pauli exclusion principle.

And that gives us a positive contribution to the energy, which is, in fact, divergent. It diverges because every standing wave contributes. And there's no lower bound to the wavelength of a standing wave. So by considering shorter and shorter wavelengths, one gets larger and larger contributions to this vacuum energy. And in the quantum field theory, it's just unbounded.

Similarly, there are quantum fluctuations to other fields like the electron field which is a Fermionic field, a field that describes a particle that obeys the Pauli exclusion principle. And those fields behave somewhat differently. Like the photon, the electron is viewed as the quantum excitation of this field. And that turns out to be by far, basically, the only way we know to describe relativistic particles in a totally consistent way.

In this case, again, the contribution to the vacuum energy is divergent. But in this case, it's negative and divergent, allowing possibilities of some kind of cancellation, but no reason that we know of why they should cancel each other. They seem to just be totally different objects.

And, finally, there are some fields which have nonzero values in the vacuum. And, in particular, the Higgs field of the standard model is believed to have a nonzero value even in the vacuum. So this is the basic story.

We commented that if we cut off these infinities by saying that we don't understand

things at very, very short wavelengths, at least one plausible cut off would be the Planck scale, which is the scale associated with where we think quantum gravity becomes important. And if we cut off at this Planck scale, the energies become finite but still too large compared to what we observe by more than 120 orders of magnitude. And on the homework set that's due next Monday, you'll be calculating this number for yourself. It's a little bit more than 120 orders of magnitude.

So it's a colossal failure indicating that we really don't understand what controls the value of this vacuum energy. And I think I mentioned last time, and I'll mention it a little more explicitly by writing it on the transparency this time, that the situation is so desperate in that we've had so much trouble trying to find any way of explaining why the vacuum energy should be so small that it has become quite popular to accept the possibility, at least, that the vacuum energy is determined by what is called the anthropic selection principal or anthropic selection effect. And Steve Weinberg was actually one of the first people who advocated this point of view. I'm sort of a recent convert to taking this point of view seriously.

But the idea is that there might be more than one possible type of vacuum. And, in fact, string theory comes in here in an important way. String theory seems to really predict that there's a colossal number of different types of vacuum, perhaps 10 to the 500 different types of vacuum or more. And each one would have its own vacuum energy.

So with that many, some of them would have a, by coincidence, near perfect cancellation between the positive and negative contributions producing a net vacuum energy that could be very, very small. But it would be a tiny fraction of all of the possible vacua, a fraction like 10 to the minus 120, since we have 120 orders of magnitude mismatch of these ranges. So you would still have to press yourself to figure out what would be the explanation why we should be living in such an atypical vacuum.

And the proposed answer is that it's anthropically selected, where anthropic means having to do with life. Whereas, the claim is made that life only evolves in vacua

which have incredibly small vacuum energies. Because if the vacuum energy is much larger, if it's positive, it blows the universe apart before structures can form. And if it's negative, it implodes the universe before there's time for structures to form.

So a long-lived universe requires a very small vacuum energy density. And the claim is that those are the only kinds of universes that support life. So we're here because it's the only kind of universe in which life can exist is the claim. Yes?

**AUDIENCE:** So, different types of energies, obviously, affect the acceleration rate and stuff of the universe. But do they also affect, in any way, the fundamental forces, or would those be the same in all of the cases?

**PROFESSOR:** OK, the question is, would the different kinds of vacuum affect the kinds of fundamental forces that exist besides the force of the cosmological constant on the acceleration of the universe itself? The answer is, yeah, it would affect really everything. These different vacua would be very different from each other. They would each have their own version of what we call the standard model of particle physics.

And that's because the standard model of particle physics would be viewed as what happens when you have small perturbations about our particular type of vacuum. And with different types of vacuum you get different time types of small perturbations about those vacua. So the physics really could be completely different in all the different vacua that string theory suggests exist.

So the story here, basically, is a big mystery. Not everybody accepts these anthropic ideas. They are talked about. At almost any cosmology conference, there will be some session where people talk about these things. They are widely discussed but by no means completely agreed upon. And it's very much an open question, what it is that explains the very small vacuum energy density that we observe.

OK, moving on, in the last lecture I also gave a quick historical overview of the

interactions between Einstein and Friedmann, which I found rather interesting. And just a quick summary here, in 1922 June 29, to be precise, Alexander Friedmann's first paper about the Friedmann equations and the dynamical model of the universe were received at [INAUDIBLE]. Einstein learned about it and immediately decided that it had to be wrong and fired off a refutation claiming that Friedmann had gotten his equations wrong. And if he had gotten them right, he would have discovered that  $\rho \dot{}$ , the rate of change of the energy density, had to be zero and that there was nothing but the static solution allowed.

Einstein then met a friend of Friedmann's Yuri Krutkov at a retirement lecture by Lawrence in Leiden the following spring. And Krutkov convinced Einstein that he was wrong about this calculation. Einstein had also received a letter from Friedmann, which he probably didn't read until this time, but the letter was apparently also convincing. So Einstein did finally retract. And at the end of May 1923, his retraction was received at *Zeitschrift fur Physik*.

And another interesting fact about that is that the original handwritten draft of that retraction still exists. And it had the curious comment, which was crossed out, where Einstein suggested that the Friedmann solutions could be modified by the phrase, "a physical significance can hardly be ascribed to them." But at the last minute, apparently, Einstein decided he didn't really have a very good foundation for that statement and crossed it out.

So I like the story, first of all, because it illustrates that we're not the only people who make mistakes. Even great physicists like Einstein make really silly mistakes. It really was just a dumb calculational error. And it also, I think, points out how important it is not to allow yourself to be caught in the grip of some firm idea that you cease to question, which apparently is what happened to Einstein with his belief that the universe was static.

He was so sure that the universe was static that he very quickly looked at Friedmann's paper and reached the incorrect conclusion that Friedmann had gotten his calculus wrong. In fact, it was Einstein who got it wrong. So that summarizes the

last lecture. Any further questions?

OK, in that case, I think I am done with that. Yeah, that comes later. OK, what I want to do next is to talk about two problems associated with the conventional cosmology that we've been learning about and, in particular, I mean cosmology without inflation, which we have not learned about yet. So I am talking about the cosmology that we've learned about so far.

So there are a total of three that I want to discuss problems associated with conventional cosmology which serve as motivations for the inflationary modification that, I think, you'll be learning about next time from Scott Hughes. But today I want to talk about the problems. So the first of 3 is sometimes called the homogeneity, or the horizon problem.

And this is the problem that the universe is observed to be incredibly uniform. And this uniformity shows up most clearly in the cosmic microwave background radiation, where astronomers have now made very sensitive measurements of the temperature as a function of angle in the sky. And it's found that that radiation is uniform to one part in 100,000, part in  $10^{-5}$ . Now, the CMB is essentially a snapshot of what the universe looked like at about 370,000 years after the Big Bang at the time that we call  $t_{\text{sub } d}$ , the time of decoupling. Yes?

**AUDIENCE:** This measurement, the  $10^{-5}$ , it's not a limit that we've reached measurement technique-wise? That's what it actually is, [INAUDIBLE]?

**PROFESSOR:** Yes, I was going to mention that. We actually do see fluctuations at the level of one part in  $10^{-5}$ . So it's not just a limit. It is an actual observation.

And what we interpret is that the photons that we're seeing in the CMB have been flying along on straight lines since the time of decoupling. And therefore, what they show us really is an image of what the universe look like at the time of decoupling. And that image is an image of the universe which is almost a perfectly smooth mass density and a perfectly smooth temperature-- it really is just radiation-- but tiny ripples superimposed on top of that uniformity where the ripples have an amplitude

of order of  $10^{-5}$ .

And those ripples are important, because we think that those are the origin of all structure in the universe. The universe is gravitationally unstable where there's a positive ripple making the mass density slightly higher than average. That creates a slightly stronger than average gravitational field pulling in extra matter, creating a still stronger gravitational field. And the process cascades until you ultimately have galaxies and clusters of galaxies and all the complexity in the universe. But it starts from these seeds, these minor fluctuations at the level of one part in  $10^5$ .

But for now we want to discuss is simply the question of how did we get so uniform. We'll talk about how the non-uniformities arise later in the context of inflation. The basic picture is that we are someplace. I'll put us here in a little diagram.

We are receiving photons, say, from opposite directions in the sky. Those little arrows represent the incoming patterns of two different CMB photons coming from opposite directions. And what I'm interested in doing to understand the situation with regard to this uniformity is I'm interested in tracing these photons back to where they originated at time  $t_{\text{sub } d}$ .

And I want to do that on both sides. But, of course, it's symmetric. So I only need to do one calculation. And what I want to know is how far apart were these two points. Because I want to explore the question of whether or not this uniformity in temperature could just be mundane.

If you let any object sit for a long time, it will approach a uniform temperature. That's why pizza gets cold when you take it out of the oven. So could that be responsible for this uniformity?

And what we'll see is that it cannot. Because these two points are just too far apart for them to come in to thermal equilibrium by ordinary thermal equilibrium processes in the context of the conventional big bang theory. So we want to calculate how far apart these points were at the time of emission.

So what do we know? We know that the temperature at the time of decoupling was



about 3,000 Kelvin, which is really where we started with our discussion of decoupling. We did not do the statistical mechanics associated with this statement. But for a given density, you can calculate at what temperature hydrogen ionizes.

And for the density that we expect for the early universe, that's the temperature at which the ionization occurs. So that's where decoupling occurs. That's where it becomes neutral as the universe expands.

We also know that during this period,  $aT$ , the scale factor times the temperature is about equal to a constant, which follows as a consequence of conservation of entropy, the idea that the universe is near thermal equilibrium. So the entropy does not change. Then we can calculate the  $z$  for decoupling, because it would just be the ratio of the temperatures. It's defined by the ratio of the scale factors. This defines what you mean by  $1 + z$  decoupling.

But if  $aT$  is about equal to a constant, we can relate this to the temperatures inversely. So the temperature of decoupling goes in the numerator. And the temperature today goes into the denominator. And, numerically, that's about 1,100. So the  $z$  of the cosmic background radiation is about 1,100, vastly larger than the red shifts associated with observations of galaxies or supernovae.

From the  $z$ , we can calculate the physical distance today of these two locations, because this calculation we already did. So I'm going to call  $l_p$  the physical distance between us and the source of this radiation. And its value today-- I'm starting with this formula simply because we already derived it on a homework set-- it's  $2c h \text{ naught inverse times } 1 \text{ minus } 1 \text{ over the square root of } 1 \text{ plus } z$ . And this calculation was done for a flat matter dominated universe, flat matter dominated.

Of course, that's only an approximation, because we know our real universe was matter dominated at the start of this period. But it did not remain matter dominated through to the present at about 5,000 or 6,000 years ago. We switched to a situation where the dark energy is actually larger than the non-relativistic matter. So we're ignoring that effect, which means we're only going to get an approximation here. But it will still be easily good enough to make the point.

For a  $z$  this large, this factor is a small correction. I think this ends up being 0.97, or something like that, very close to 1, which means what we're getting is very close to  $2c h_0^{-1}$ , which is the actual horizon. The horizon corresponds to  $z$  equals infinity. If you think about it, that's what you expect the horizon to be. It corresponds to infinite red shift. And you don't see anything beyond that.

So if we take the best value we have for  $h_0$ , which I'm taking from the Planck satellite, 57.3 kilometers per second per megaparsec, and put that and the value for  $z$  into this formula, we get  $l_{p \text{ of } t_0}$  of 28.2 billion light years times  $10^9$  light year. So it's of course larger than  $ct$  as we expect. It's basically  $3ct$  for a matter dominated universe. And  $3ct$  is the same as  $2c h_0^{-1}$ .

Now, what we want to know, though, is how far away was this when the emission occurred, not the present distance. We looked at the present distance simply because we had a formula for it from our homework set. But we know how to extrapolate that backwards,  $l_{\text{sub } t \text{ at time } t_{\text{sub } d}}$ . Distances that are fixed in co-moving space, which these are, are just stretched with the scale factor. So this will just be the scale factor at the time of decoupling divided by the scale factor today times the present distance.

And this is, again, given by this ratio of temperatures. So it's 1 over 1,100, the inverse of what we had over there. So the separation at this early time is just 1,100 times smaller than the separation today. And that can be evaluated numerically. And it gives us 2.56 times  $10^7$  light years, so 25 million light years.

Now, the point is that that's significantly larger than the horizon distance at that time. And remember, the horizon distance is the maximum possible distance that anything can travel limited by the speed of light from the time of the big bang up to any given point in cosmic history. So the horizon at time  $t_{\text{sub } d}$  is just given by the simple formula that the physical value of the horizon distance,  $l_{\text{sub } h \text{ phys}}$ ,  $l_{\text{sub } h \text{ phys}}$ , at time  $t_{\text{sub } d}$  is just equal to, for a matter dominated universe,  $3c$  times  $t_{\text{sub } d}$ .

And that we can evaluate, given what we have. And it's about 1.1 times 10 to the sixth light years, which is significantly less than 2.56 times 10 to the seven light years. And, in fact, the ratio of the two, given these numbers, is that  $\frac{t_{sub\ p}}{t_{sub\ d}}$  over  $\frac{d}{h}$  is also of  $\frac{t_{sub\ d}}{d}$  is about equal to 23, just doing the arithmetic.

And that means if we go back to this picture, these two points of emission were separated from each other by about 46 horizon distances. And that's enough to imply that there's no way that this point could have known anything whatever about what was going on at this point. Yet somehow they knew to emit these two photons at the same time at the same temperature. And that's the mystery.

One can get around this mystery if one simply assumes that the singularity that created all of this produced a perfectly homogeneous universe from the very start. Since we don't understand that singularity, we're allowed to attribute anything we want to it. So in particular, you can attribute perfect homogeneity to the singularity.

But that's not really an explanation. That's an assumption. So if one wants to be able to explain this uniformity, then one simply cannot do it in the context of conventional cosmology. There's just no way that causality, the limit of the speed of light, allows this point to know anything about what's going on at that point. Yes?

**AUDIENCE:** How could a singularity not be uniform? Because if it had non-uniform [INAUDIBLE], then not be singular?

**PROFESSOR:** OK, the question is how can a singularity not be uniform? The answer is, yes, singularities can not be uniform. And I think the way one can show that is a little hard. But you have to imagine a non-uniform thing collapsing. And then it would just be the time reverse, everything popping out of the singularity.

So you can ask, does a non-uniform thing collapse to a singularity? And the answer to that question is not obvious and really was debated for a long time. But there were theorems proven by Hawking and Penrose that indeed not only do the homogeneous solutions that we look at collapse but in homogeneous solutions also collapse to singularities. So a singularity does not have to be uniform.

OK, so this is the story of the horizon problem. And as I said, you can get around it if you're willing to just assume that the singularity was homogeneous. But if you want to have a dynamical explanation for how the uniformity of the universe was established, then you need some model other than this conventional cosmological model that we've been discussing. And inflation will be such a twist which will allow a solution to this problem.

OK, so if there are no questions, no further questions, we'll go on to the next problem I want to discuss, which is of a similar nature in that you can get around it by making strong assumptions about the initial singularity. But if one wants, again, something you can put your hands on, rather than just an assumption about a singularity, then inflation will do the job. But you cannot solve the problem in the context of a conventional big bang theory, because the mechanics of the conventional big bang theory are simply well-defined.

So what I want to talk here is what is called the flatness problem, where flatness is in the sense of  $\Omega$  very near 1. And this is basically the problem of why is  $\Omega$  today somewhere near 1? So  $\Omega_{naught}$  is the present value of  $\Omega$ , why is it about equal to 1?

Now, what do we know first of all about it being about equal to 1? The best value from the Planck group, this famous Planck satellite that I've been quoting a lot of numbers from-- and I think in all cases, I've been quoting numbers that they've established combining their own data with some other pieces of data. So it's not quite the satellite alone. Although, they do give numbers for the satellite alone which are just a little bit less precise.

But the best number they give for  $\Omega_{naught}$  is minus 0.0010 plus or minus 0.0065. Oops, I didn't put enough zeroes there. So it's 0.0065 is the error. So the error is just a little bit more than a half of a percent. And as you see, consistent with-- I'm sorry, I meant this to be 1.

Hold on. This is  $\Omega_{naught} - 1$  that I'm writing a formula for. So  $\Omega_{naught}$  is very near 1 up to that accuracy.

What I want to emphasize in terms of this flatness problem is that you don't need to know that Omega naught is very, very close to 1 today, which we now do know. But even back when inflation was first invented around 1980, in circa 1980 we certainly didn't know that Omega was so incredibly close to 1. But we did know that Omega was somewhere in the range of about 0.1 and 0.2, which is not nearly as close to 1 as what we know now, but still close to 1.

I'll argue that the flatness problem exists for these numbers almost as strongly as it exists for those numbers. Differ, but this is still a very, very strong argument that even a number like this is amazingly close to 1 considering what you should expect. Now, what underlies this is the expectations, how close should we expect Omega to be to 1?

And the important underlying piece of dynamics that controls this is the fact that Omega equals 1 is an unstable equilibrium point. That means it's like a pencil balancing on its tip. If Omega is exactly equal to 1, that means you have a flat universe. And an exactly flat universe will remain an exactly flat universe forever. So if Omega is exactly equal to 1, it will remain exactly equal to 1 forever.

But if Omega in the early universe were just a tiny bit bigger than 1-- and we're about to calculate this, but I'll first qualitatively describe the result-- it would rise and would rapidly reach infinity, which is what it reaches if you have a closed universe when a closed universe reaches its maximum size. So Omega becomes infinity and then the universe recollapses. So if Omega were bigger than 1, it would rapidly approach infinity. If Omega in the early universe were just a little bit less than 1, it would rapidly trail off towards 0 and not stay 1 for any length of time.

So the only way to get Omega near 1 today is like having a pencil that's almost straight up after standing there for 1 billion years. It'd have to have started out incredibly close to being straight up. It has to have started out incredibly close to Omega equals 1. And we're going to calculate how close. So that's the set-up.

So the question we want to ask is how close did Omega have to be to 1 in the early

universe to be in either one of these allowed ranges today. And for the early universe, I'm going to take  $t$  equals one second as my time at which I'll do these calculations. And, historically, that's where this problem was first discussed by Dicke and Peebles back in 1979.

And the reason why one second was chosen by them, and why it's a sensible time for us to talk about as well, is that one second is the beginning of the processes of nucleosynthesis, which you've read about in Weinberg and in Ryden, and provides a real test of our understanding of cosmology at those times. So we could say that we have real empirical evidence in the statement that the predictions of the chemical elements work. We could say that we have real empirical evidence that our cosmological model works back to one second after the Big Bang.

So we're going to choose one second for the time at which we're going to calculate what  $\Omega$  must've been then for it to be an allowed range today. How close must  $\Omega$  have been to 1 at  $t$  equals 1 second? Question mark.

OK, now, to do this calculation, you don't need to know anything that you don't already know. It really follows as a consequence of the Friedmann equation and how matter and temperature and so on behave with time during radiation in matter dominated eras. So we're going to start with just the plain old first order Friedmann equation,  $h^2$  is equal to  $\frac{8\pi}{3} \rho - \frac{kc^2}{a^2}$ , which you have seen many, many times already in this course.

We can combine that with other equations that you've also seen many times. The critical density is just the value of the density when  $k$  equals 0. So you just solve this equation for  $\rho$ . And you get  $\frac{3h^2}{8\pi G}$ . This defines the critical density. It's that density which makes the universe flat,  $k$  equals 0.

And then our standard definition is that  $\Omega$  is just defined to be the actual mass density divided by the critical mass density. And  $\Omega$  will be the quantity we're trying to trace. And we're also going to make use of the fact that during the era that we're talking about,  $\rho$  is essentially equal to a constant.

It does change a little bit when electron and positron pairs freeze out. It changes by a factor of something like  $4/11$  to the  $1/3$  power or something like that. But that power will be of order one for our purposes.

But I guess this is a good reason why I should put a squiggle here instead of an equal sign as an approximate equality, but easily good enough for our purposes, meaning the corrections of order one. We're going to see the problem is much, much bigger than order one. So a correction of order one doesn't matter.

Now, I'm going to start by using the Planck satellite limits. And at the end, I'll just make a comment about the circa 1980 situation. But if we look at the Planck limits-- I'm sorry. Since I'm going to write an equation for a peculiar looking quantity, I should motivate the peculiar looking quantity first.

It turns out to be useful for these purposes. And this purpose means we're trying to track how  $\Omega$  changes with time. It turns out to be useful to reshuffle the Friedmann. And it is just an algebraic reshuffling of the Friedmann equation and the definitions that we have here.

We can rewrite the Friedmann equation as  $\Omega - 1$  over  $\Omega$  is equal to a quantity called  $a$  times the temperature squared over  $\rho$ . Now, the temperature didn't even occur in the original equation. So things might look a little bit suspicious.

I haven't told you what  $a$  is yet.  $a$  is  $3k c^2$  over  $8\pi g a^2 t^2$ . So when you put the  $a$  into this equation, the  $t$  squares cancel. So the equation doesn't really have any temperature dependence.

But I factored things this way, because we know that  $a$  is approximately a constant. And that means that this capital  $a$ , which is just other things which are definitely constant times,  $a^2 t^2$  in the denominator, this  $a$  is approximately a constant. And you'll have to check me at home that this is exactly equivalent to the original Friedmann equation, no approximations whatever, just substitutions of  $\Omega$  and the definition of  $\rho_{sub c}$ .

So the nice thing about this is that we can read off the time dependence of the right-

hand side as long as we know the time dependence in the temperature and the time dependence of the energy density. And we do for matter dominated and radiation dominated eras. So this, essentially, solves the problem for us. And now it's really just a question of looking at the numerics that follow as a consequence of that equation.

And this quantity, we're really interested in just  $\Omega - 1$ . The Friedmann equation gave us the extra complication of an  $\Omega$  in the denominator. But in the end, we're going to be interested in cases where  $\Omega$  is very, very close to 1. So the  $\Omega$  in the denominator we could just set equal to one. And it's the  $\Omega - 1$  in the numerator that controls the value of the left-hand side.

So if we look at these Planck limits, we could ask how big can that be? And it's biggest if the error occurs on the negative side. So it contributes to this small mean value which is slightly negative.

And it gives you 0.0075 for  $\Omega - 1$ . And then if you put that in the numerator and the same thing in the denominator, you get something like 0.0076. But I'm just going to use the bound that  $\Omega - 1$  over  $\Omega$  is less than 0.01.

But the more accurate thing would be 0.076. But, again, we're not really interested in small factors here. And this is a one sigma error. So the actual error could be larger than this, but not too much larger than this.

So I'm going to divide the time interval between one second and now up into two integrals. From one second to about 50,000 years, the universe was radiation dominated. We figured out earlier that the matter to radiation equality happens at about 50,000 years. I think we may have gotten 47,000 years or something like that when we calculated it.

So for  $t$  equals 1 second to-- I'm sorry, I'm going to do it the other order. I'm going to start with the present and work backwards. So for  $t$  equals 50,000 years to the present, the universe is matter dominated.



And the next thing is that we know how matter dominated universe's behave. We don't need to recalculate it. We know that the scale factor for a matter dominated flat universe goes like  $t$  to the  $2/3$  power, I should have a proportionality sign here.  $a$  of  $t$  is proportional to  $t$  to  $2/3$ .

And it's fair to assume flat, because we're always going to be talking about universes that are nearly flat and becoming more and more flat as we go backwards, as we'll see. And again, this isn't an approximate calculation. One could do it more accurately if one wanted to. But there's really no need to, because the result will be so extreme.

The temperature behaves like one over the scale factor. And that will be true for both the matter dominated and a radiation dominated universe. And the energy density will be proportional to one over the scale factor cubed.

And then if we put those together and use the formula on the other blackboard and ask how  $\Omega_{-1}$  over  $\Omega$  behaves, it's proportional to the temperature squared divided by the energy density. The temperature goes like  $1/a$ . So temperature squared goes like  $1/a^2$ .

But  $\rho$  in the denominator goes like  $1/a^3$ . So you have  $1/a^2$  divided by  $1/a^3$ . And that means it just goes like  $a$ , the scale factor itself. So  $\Omega_{-1}$  over  $\Omega$  is proportional to  $a$ . And that means it's proportional to  $t$  to the  $2/3$ .

So that allows us to write down an equation, since we want to relate everything to the value of  $\Omega_{-1}$  over  $\Omega$  today, we can write  $\Omega_{-1}$  over  $\Omega$  at 50,000 years is about equal to the ratio of the 2 times the 50,000 years and today, which is 13.8 billion years, to the  $2/3$  power since  $\Omega_{-1}$  grows like  $t$  to the  $2/3$ . I should maybe have pointed out here, this telling us that  $\Omega_{-1}$  grows with time. That's the important feature. It grows  $t$  to the  $2/3$ . So the value at 50,000 years is this ratio to the  $2/3$  power times  $\Omega_{-1}$  over  $\Omega$  today, which I can indicate just by putting subscript zeros on my  $\Omega$ s. And that makes it today.

And I've written this as a fraction less than one. This says that  $\Omega_{-1}$  over  $\Omega$  was smaller than it is now by this ratio to the  $2/3$  power, which follows from the fact that  $\Omega_{-1}$  over  $\Omega$  grows like  $t$  to the  $2/3$ . OK, we're now halfway there. And the other half is similar, so it will go quickly.

We now want to go from 50,000 years to one second using the fact that during that era the universe was radiation dominated. So for  $t$  equals 1 second to 50,000 years, the universe is radiation dominated. And that implies that the scale factor is proportional to  $t$  to the  $1/2$ . The temperature is, again, proportional to 1 over the scale factor. That's just conservation of entropy. And the energy density goes one over the scale factor to the fourth.

So, again, we go back to this formula and do the corresponding arithmetic. Temperature goes like 1 over  $a$ . Temperature squared goes like 1 over  $a$  squared. That's our numerator.

This time, in the denominator, we have  $\rho$ , which goes like one over  $a$  to the fourth. So we have 1 over  $a$  squared divided by 1 over  $a$  to the fourth. And that means it goes like  $a$  squared.

So we get  $\Omega_{-1}$  over  $\Omega$  is proportional to  $a$  squared. And since  $a$  goes like the square root of  $t$ ,  $a$  squared goes like  $t$ . So during the radiation dominated era this diverges even a little faster.

**PROFESSOR:** It goes like  $t$ , rather than like  $t$  to the  $2/3$ , which is a slightly slower growth. And using this fact, we can put it all together now and say that  $\Omega_{-1}$  over  $\Omega$  at 1 second is about equal to 1 second over 50,000 years to the first power-- this is going like the first power of  $t$ -- times the value of  $\Omega_{-1}$  over  $\Omega$  at 50,000 years. And  $\Omega$  at 50,000 years, we can put in that equality and relate everything to the present value.

And when you do that, putting it all together, you ultimately find that  $\Omega_{-1}$  in magnitude at  $t$  equals 1 second is less than about 10 to the minus 18. This is just

putting together these inequalities and using the Planck value for the present value, the Planck inequality. So then  $10^{-18}$  is a powerfully small number. What we're saying is that to be in the allowed range today, at one second after the Big Bang,  $\Omega$  have to have been equal to 1 in the context of this conventional cosmology to an accuracy of 18 decimal places.

And the reason that's a problem is that we don't know any physics whatever that forces  $\Omega$  to be equal to 1. Yet, somehow  $\Omega$  apparently has chosen to be 1 to an accuracy of 18 decimal places. And I mention that the argument wasn't that different in 1980.

In 1980, we only knew this instead of that. And that meant that instead of having  $10^{-18}$  on the right-hand side here, we would have had  $10^{-15}$  differing by three orders of magnitude. So instead of getting  $10^{-18}$  here, we would have gotten  $10^{-15}$ .

And  $10^{-15}$  is, I guess, a lot bigger than  $10^{-18}$  by a factor of 1,000. But still, it's an incredibly small number. And the argument really sounds exactly the same. The question is, how did  $\Omega - 1$  get to be so incredibly small? What mechanism was there?

Now, like the horizon problem, you can get around it by attributing your ignorance to the singularity. You can say the universe started out with  $\Omega$  exactly equal to 1 or  $\Omega$  equal to 1 to some extraordinary accuracy. But that's not really an explanation. It really is just a hope for an explanation.

And the point is that inflation, which you'll be learning about in the next few lectures, provides an actual explanation. It provides a mechanism that drives the early universe towards  $\Omega = 1$ , thereby explaining why the early universe had a value of  $\Omega$  so incredibly close to 1. So that's what we're going to be learning shortly.

But at the present time, the takeaway message is simply that for  $\Omega$  to be in the allowed range today it had to start out unbelievably close to 1 at, for example,  $t$

equals 1 second. And within conventional cosmology, there's no explanation for why Omega so close to 1 was in any way preferred. Any questions about that? Yes?

**AUDIENCE:** Is there any heuristic argument that omega [INAUDIBLE] universe has total energy zero? So isn't that, at least, appealing?

**PROFESSOR:** OK the question is, isn't it maybe appealing that Omega should equal 1 because Omega equals 1 is a flat universe, which has 0 total energy? I guess, the point is that any closed universe also has zero total energy. So I don't think Omega equals 1 is so special.

And furthermore, if you look at the spectrum of possible values of Omega, it can be positive-- I'm sorry, not with Omega. Let me look at the curvature itself, little k. Little k can be positive, in which case, you have a closed universe.

It can be negative, in which case, you have an open universe. And only for the one special case of k equals 0, which really is one number in the whole real line of possible numbers, do you get exactly flat. So I think from that point of view flat looks highly special and not at all plausible as what you'd get if you just grabbed something out of a grab bag.

But, ultimately, I think there's no way of knowing for sure. Whether or not Omega equals 1 coming out of the singularity is plausible really depends on knowing something about the singularity, which we don't. So you're free to speculate. But the nice thing about inflation is that you don't need to speculate. Inflation really does provide a physical mechanism that we can understand that drives Omega to be 1 exactly like what we see.

Any other questions? OK, in that case, what I'd like to do is to move on to problem number three, which is the magnetic monopole problem, which unfortunately requires some background to understand. And we don't have too much time. So I'm going to go through things rather quickly.

This magnetic monopole problem is different from the other two in the first two problems I discussed are just problems of basic classical cosmology. The magnetic

monopole problem only arises if we believe that physics at very high energies is described what are called grand unified theories, which then imply that these magnetic monopoles exist and allow us a root for estimating how many of them would have been produced. And the point is that if we assume that grand unified theories are the right description of physics at very high energies, then we conclude that far too many magnetic monopoles would be produced if we had just the standard cosmology that we've been talking about without inflation.

So that's going to be the thrust of the argument. And it will all go away if you decide you don't believe in grand unified theories, which you're allowed to. But there is some evidence for grand unified theories. And I'll talk about that a little bit.

Now, I'm not going to have time to completely describe grand unified theories. But I will try to tell you enough odd facts about grand unified theories. So there will be kind of a consistent picture that will hang together, even though there's no claim that I can completely teach you grand unified theories in the next 10 minutes and then talk about the production of magnetic monopoles and those theories in the next five minutes. But that will be sort of the goal.

So to start with, I mentioned that there's something called the standard model of particle physics, which is enormously successful. It's been developed really since the 1970s and has not changed too much since maybe 1975 or so. We have, since 1975, learned that neutrinos have masses. And those can be incorporated into the standard model. And that's a recent addition.

And, I guess, in 1975 I'm not sure if we knew all three generations that we now know. But the matter, the fundamental particles fall into three generations, these particles of a different type. And we'll talk about them later.

But these are the quarks. These are the spin-1/2 particles, these three columns on the left. On the top, we have the quarks, up, down, charm, strange, top, and bottom. There are six different flavors of quarks.

Each quark, by the way, comes in three different colors. The different colors are

absolutely identical to each other. There's a perfect symmetry among colors. There's no perfect symmetry here. Each of these quarks is a little bit different from the others. Although, there are approximate symmetries.

And related to each family of quarks is a family of leptons, particles that do not undergo strong interactions in the electron-like particles and neutrinos. This row is the neutrinos. There's an electron neutrino, a muon neutrino, and a tau neutrino, like we've already said. And there's an electron, a muon, and a tau, which I guess we've also already said.

So the particles on the left are all of the spin-1/2 particles that exist in the standard model of particle physics. And then on the right, we have the Boson particles, the particles of integer spin, starting with the photon on the top. Under that in this list-- there's no particular order in here really-- is the gluon which is the particle that's like the photon but the particle which describes the strong interactions, which are somewhat more complicated and electromagnetism, but still described by spin-1 particles just like the photon.

And then two other spin-1 particles, the  $Z^0$  and the  $W^+$  and  $W^-$ , which are a neutral particle and a charged particle, which are the carriers of the so-called weak interactions. The weak interactions being the only non-gravitational interactions that neutrinos undergo. And the weak interactions are responsible for certain particle decays.

For example, a neutron can decay into a proton giving off also an electron-- producing a proton, yeah-- charge has to be conserved, proton is positive. So it's produced with an electron and then an anti-electron neutrino to balance the so-called electron lepton number. And that's a weak direction. Essentially, anything that involves neutrinos is going to be weak interaction.

So these are the characters. And there's a set of interactions that go with this set of characters. So we have here a complete model of how elementary particles interact. And the model has been totally successful. It actually gives predictions that are consistent with every reliable experiment that has been done since the mid-1970s

up to the present. So it's made particle physics a bit dull since we have a theory that seems to predict everything. But it's also a magnificent achievement that we have such a theory.

Now, in spite of the fact that this theory is so unbelievably successful I don't think I know anybody who really regards this as a candidate even or the ultimate theory of nature. And the reason for that is maybe twofold, first is that it does not incorporate gravity, it only incorporates particle interactions. And we know that gravity exists and has to be put in somehow. And there doesn't seem to be any simple way of putting gravity into this theory.

And, secondly-- maybe there's three reasons-- second, it does not include any good candidates for the dark matter that we know exists in cosmology. And third, excuse me-- and this is given a lot of importance, even though it's an aesthetic argument-- this model has something like 28 free parameters, quantities that you just have to go out and measure before you can use the model to make predictions.

And a large number of free parameters is associated, by theoretical physicists, with ugliness. So this is considered a very ugly model. And we have no real way of knowing, but almost all theoretical physicists believe that the correct theory of nature is going to be simpler and involve many fewer, maybe none at all, free knobs that can be turned to produce different kinds of physics.

OK, what I want to talk about next, leading up to grand unified theories, is the notion of gauge theories. And, yes?

**AUDIENCE:** I'm sorry, question real quick from the chart. I basically heard the explanation that the reason for the short range of the weak force was the massive mediator that is the cause of exponential field decay. But if the [INAUDIBLE] is massless, how do we explain that to [INAUDIBLE]?

**PROFESSOR:** Right, OK, the question is for those of you who couldn't hear it is that the short range of the weak interactions, although I didn't talk about it, is usually explained and is explained by the fact that the z and w naught Bosons are very heavy. And

heavy particles have a short range. But the strong interactions seem to also have a short range. And yet, the gluon is effectively massless.

That's related to a complicated issue which goes by the name of confinement. Although the gluon is massless, it's confined. And confined means that it cannot exist as a free particle. In some sense, the strong interactions do have a long range in that if you took a meson, which is made out of a quark and an anti-quark, in principle, if you pulled it apart, there'd be a string of gluon force between the quark and the anti-quark. And that would produce a constant force no matter how far apart you pulled them.

And the only thing that intervenes, and it is important that it does intervene, is that if you pulled them too far apart it would become energetically favorable for a quark anti-quark pair to materialize in the middle. And then instead of having a quark here and an anti-quark here and a string between them, you would have a quark here and an anti-quark there and a string between them, and an anti-quark here and-- I'm sorry, I guess it's a quark here and an anti-quark there and a string between those. And then they would just fly apart. So the string can break by producing quark anti-quark pairs. But the string can never just end in the middle of nowhere.

And that's the essence of confinement. And it's due to the peculiar interactions that these gluons are believed to obey. So the gluons behave in a way which is somewhat uncharacteristic of particles. Except at very short distances, they behave very much like ordinary particles. But at larger distances, these effects of confinement play a very significant role. Any other questions?

OK, onward, I want talk about gauge theories, because gauge theories have a lot to do with how one gets into grand unified theories from the standard model. And, basically, a gauge theory is a theory in which the fundamental fields are not themselves reality. But rather there's a set of transformations that the fields can undergo which take you from one description to an equivalent description of the same reality.

So there's not a one to one mapping between the fields and reality. There are many



different field configurations that correspond to the same reality. And that's basically what characterizes what we call a gauge theory. And you do know one example of a gauge theory. And that's  $e$  and  $m$ .

If  $e$  and  $m$  is expressed in terms of the potentials  $\Phi$  and  $A$ , you can write  $e$  in terms of the potential that way and  $b$  as the curl of  $A$ , you could put  $\Phi$  and  $A$  together and make a four-vector if you want to do things in a Lorentz covariant way. And the important point, though, whether you put them together or not, is that you can always define a gauge transformation depending on some arbitrary function  $\Lambda$ , which is a function of position and time. I didn't write in the arguments, but  $\Lambda$  is just an arbitrary function of position and time.

And for any  $\Lambda$  you can replace  $\psi$  by  $\psi'$ , given by this line, and  $a$  by  $a'$ , given by that line. And if you go back and compute  $e$  and  $b$ , you'll find that they'll be unchanged. And therefore, the physics is not changed, because the physics really is all contained in  $e$  and  $b$ .

So this gauge transformation is a transformation on the fields of the theory-- it can be written covariantly this way-- which leaves the physics invariant. And it turns out that all the important field theories that we know of are gauge theories. And that's why it's worth mentioning here.

Now, for  $e$  and  $m$ , the gauge parameter is just this function  $\Lambda$ , which is a function of position and time. And an important issue is what happens when you combine gauge transformations, because the succession of two transformations had better also be a symmetry transformation. So it's worth understanding that group structure.

And for the case of  $e$  and  $m$ , these  $\Lambda$ s just add if we make successive transformations. And that means the group is Abelian. It's commutative. But that's not always the case.

Let's see, where am I going? OK, next slide actually comes somewhat later. Let me go back to the blackboard.

It turns out that the important generalization of gauge theories is the generalization from Abelian gauge theories to non-Abelian ones, which was done originally by Yang and Mills in 1954, I think. And when it was first proposed, nobody knew what to do with it. But, ultimately, these non-Abelian gauge theories became the standard model of particle physics.

And in non-Abelian gauge theories the parameter that describes the gauge transformation is a group element, not just something that gets added. And group elements multiply, according to the procedures of some group. And in particular, the standard model is built out of three groups. And the fact that there are three groups not one is just an example of this ugliness that I mentioned and is responsible for the fact that there's some significant number of parameters even if there were no other complications.

So the standard model is based on three gauge groups, SU3, SU2, and U1. And it won't really be too important for us what exactly these groups are. Let me just mention quickly, SU3 is a group of 3 by 3 matrices which are unitary in the sense that  $u^\dagger u = 1$  and special in the sense that they have determinant 1.

And the set of all 3 by 3 matrices have those properties form a group. And that group is SU3. SU2 is the same thing but replace the 3 in all those sentences by 2s. U1 is just the group of phases. That is the group of real numbers that could be written as either the  $e^{i\phi}$ .

So it's just a complex number of magnitude 1 with just a phase. And you can multiply those and they form a group. And the standard model contains these three groups. And the three groups all act independently, which means that if you know about group products, one can say that the full group is the product group.

And that just means that a full description of a group element is really just a set of an element of SU3, and an element of SU2, and an element of U1. And if you put together three group elements in each group and put them together with commas, that becomes a group element of the group SU3 cross SU2 cross U1. And that is

the gauge group of the standard model of particle physics.

OK, now grand unified theories, a grand unified theory is based on the idea that this set of three groups can all be embedded in a single simple group. Now, simple actually has a mathematical group theory meaning. But it also, for our purposes, just means simple, which is good enough for our brush through of these arguments.

And, for example-- and the example is shown a little bit in the lecture notes that will be posted shortly-- an example of a grand unified theory, and indeed the first grand unified theory that was invented, is based on the full gauge group SU5, which is just a group of 5 by 5 matrices which are unitary and have determinant 1. And there's an easy way to embed SU3 and SU2 and U1 into SU5. And that's the way that was used to construct this grand unified theory.

One can take a 5 by 5 matrix-- so this is a 5 by 5 matrix-- and one can simply take the upper 3 by 3 block and put an SU3 matrix there. And one can take the lower 2 by 2 block and put an SU2 matrix there.

And then the U1 piece-- there's supposed to be a U1 left over-- the U1 piece can be obtained by giving an overall phase to this and an overall phase to that in such a way that the product of the five phases is 0. So the determinant has not changed.

So one can put an  $e^{i2\Phi}$  there and a factor of  $e^{-i3\Phi}$  there for any  $\Phi$ . And then this  $\Phi$  becomes the description of the U1 piece of this construction. So we can take an arbitrary SU3 matrix, and arbitrary SU2 matrix, and an arbitrary U1 value expressed by  $\Phi$  and put them together to make an SU5 matrix.

And if you think about it, the SU3 piece will commute with the SU2 piece and with the U1 piece. These three pieces will all commute with each other, if you think about how multiplication works with this construction. So it does exactly what we want. It decomposes SU5. So it has a subgroup of SU3 cross SU2 across U1. And that's how the simplest grand unified theory works.

OK, now, there are important things that need to be said, but we're out of time. So I

guess what we need to do is to withhold from the next problem set, the magnetic monopole problem. Maybe I was a bit over-ambitious to put it on the problem set. So I'll send an email announcing that.

But the one problem on the problem set for next week about grand unified theories will be withheld. And Scott Hughes will pick up this discussion next Tuesday. So I will see all of you-- gee willikers, if you come to my office hour, I'll see you then. But otherwise, I may not see you until the quiz. So have a good Thanksgiving and good luck on the quiz.