

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.

PROFESSOR: OK if there are no questions, we will get back to physics. What I want to do today, as it suggests on the slide, is to finish the kinematics of homogeneous expansion that we were talking about last time. And the one topic in that category that we have not discussed yet is the cosmological redshift So we'll begin by going over that. And then we'll begin to go on to the next topic altogether, which is the dynamics of homogeneous expansion -- how do we understand how gravity affects the expansion of the universe? So that will be the main subject of today's lecture, once we've finished up the issue of the cosmological redshift shift.

Let me remind you that at end of the last lecture, we were talking about the synchronization of clocks, and the coordinate system that we'll be using to describe the homogeneously expanding model of the universe. Remember, we are introducing spatial coordinates that grow with the universe, so that we're going to be assuming the fact, literally, that the universe is perfectly homogeneous and isotropic, which means that all objects will be literally addressed, relative to this coordinate system. If we're talking about the real universe, then there would be some motion relative to this coordinate system, because the universe is not exactly homogeneous. But we're going to be working for now with the approximation that our model universe is exactly homogeneous, which means that all matter is completely at rest, relative to this expanding coordinate system.

And now we want to talk about how to define time, or to review what we said last time when we talked about how to define time. What we will imagine is that in every location in the universe at rest, relative to the matter, is a clock. And each clock ticks off time, and all those clocks will be acceptable as a clock which measures the time at relevant positions-- time is measured locally-- but we still have to talk about synchronizing those clocks.

And what we said last time is that we can synchronize the clocks as long as there's some cosmic phenomena that can be seen everywhere, which has some time evolution. And we gave two examples-- one is the evolution of the Hubble expansion rate, which can be measured locally, and everybody can agree to set their clocks to midnight when the Hubble expansion rate has a certain value. And another cosmic variable is the temperature of the cosmic microwave background radiation.

So, everybody in this model universe will agree that we'll set the clocks to midnight when the temperature of the cosmic background radiation goes to 5 degrees, or any specified number. So as long as there's a phenomena of that sort, which there is in our universe, it's possible to synchronize these clocks in a unique way. And the important thing to realize is that once they're synchronized at one time, they will remain synchronized as a consequence of our assumption of homogeneity. That is, if everybody agrees that the cosmic background radiation has a temperature of 10 degrees at midnight, if everybody waits for 15 minutes after midnight, everybody should see the same fall in temperature during that time interval, otherwise it would be a violation of this hypothesis of perfect homogeneity. Yes, question.

AUDIENCE: Is it verified that temperature is invariant for all observers-- all Lorentz observers?

PROFESSOR: OK the question is, is temperature invariant for all observers? And the question even included all Lorentz observers. It's not really invariant to different Lorentz observers. We're talking about a privileged class of observers, all of whom are at rest, relative to the average matter. If you move through the cosmic background radiation, then you don't see uniform thermal distribution any more. Rather what you see is radiation that's hotter in the forward direction and colder in the backward direction. And we in fact, as I think I have mentioned here, see that effect in our real universe. We're apparently moving relative to the cosmic background radiation, at about 1/1000th of the speed of light. So it's not invariant with respect to motion.

There's the additional question, though is, is it the same everywhere in the visible universe? As far as we can tell, it is. There is some direct measurement of that, that

we'll probably talk about later in the course, by looking at certain spectral lines in distant galaxies. One can effectively measure the temperature of the cosmic microwave background radiation in some distant galaxies. This line cannot be seen in all galaxies, and the extent that it's been measured in degrees. So certainly in our model, we're going to assume complete homogeneity, so everything's the same everywhere, and there is strong evidence for that homogeneity. Although it's not exact, but there's strong evidence for approximate homogeneity in the real universe. Yes?

AUDIENCE: If you were really close to the black hole [INAUDIBLE].

PROFESSOR: OK. The question is, suppose we're a little bit more careful, and talk about the fact that some people might be living near black holes, and other people are not. Will that affect the synchronization of clocks for the people who are living near black holes? The answer is sure, it will. We can only synchronize clocks cosmically if we assume that the universe is absolutely homogeneous. As soon as you introduce inhomogeneities like black holes, or even just stars like the sun, they create small perturbations, which then make it really impossible to expect clocks to stay in sync with each other. So as soon as you have concentrations of mass, then the fact that what we're talking about now is an approximation becomes real. But those deviations are small. The deviations coming from the sun are only on the order of a part in a million or so. So, to a very good approximation, the universe obeys what we're describing, although if you went very close to the surface of one of these super-massive black holes in the centers of galaxies, or something, you would in fact find they had a very significant effect on your clocks. Any other questions?

OK. Let me move on now. The next topic, as I have warned you, is the cosmological redshift. Now in the first lecture beyond the overview, which I guess was a combination of the second and third lectures in the course, we talked about the Doppler shift for sound waves, and we talked about the relativistic Doppler shift for light waves-- that was all in the context of special relativity. Now what we're going to face is the fact that cosmology is not really governed entirely by special relativity,

although special relativity still holds locally in our cosmology. But special relativity does not include the effects of gravity, and on a global scale, the effects of gravity are very important for cosmology, and therefore special relativity by itself is not enough to understand many properties of the universe, including the cosmological redshift. It turns out though, that there's a way of describing the cosmological redshift which will make it sound even simpler than special relativity. And I'll describe that first, and then afterwards, we'll talk a little bit about how this very simple-looking derivation jives with the special relativity derivation, which must also be correct, at least locally.

OK. So, the question we want to ask ourselves, is suppose we look at a distant galaxy, and light is emitted from that galaxy. How will the frequency of that light shift between the frequency it had when it was emitted, and the frequency that we would measure as we received the light. So to draw the situation on the blackboard, let's introduce a coordinate system, x . And this will be our comoving coordinate system. X is measured in notches. We'll put ourselves at the origin-- there is us. And we'll put our galaxy out here someplace-- there is the distant galaxy that we will be observing. The galaxy will be at some particular coordinate, which I will call $l_{sub\ c}$, c for coordinate distance, so $l_{sub\ c}$ is the coordinate distance to the galaxy. And then the physical distance-- is what we've been calling $l_{sub\ p}$, p for physical, which depends on time, because there's Hubble expansion. So $l_{sub\ p}$ of t , as we've said a number of times already, is a of t times $l_{sub\ c}$. The scale factor, which depends on time, times the coordinate distance, which does not depend on time. So everything just expands with the scale factor a of t .

So this describes the situation, and now what we want to ask ourselves, is suppose a wave is being emitted by the galaxy-- and we'll be trying to track the distance between wave crests, which determines what the wavelength is. Since we'll only be interested in wave crests, we will talk in language where we just imagine there's a pulse at each crest, and what happens in between doesn't matter for what we're talking about. So we want to track successive pulses emitted by the galaxy.

Now the important feature of our system is that we have argued that we know how

to track light waves through this kind of coordinate system. If x is our cosmic coordinate, dx/dt , the coordinate velocity of light, is just equal to the ordinary velocity of light, c , but rescaled by the scale factor. And the scale factor here is playing the role of converting meters to notches. So c is measured in meters per second. By dividing by a of t , we get the speed in notches per second, which is what we want, because x is measured, not in meters, but in notches. A notch being the arbitrary coordinate-- the arbitrary unit that we adopt to describe our comoving coordinate system.

Now the important feature of this equation, for our current purpose, is that the speed of light, as we're going to follow these light pulses through our coordinate system, depends on time, but it does not depend on x . Our universe is homogeneous, so all points x are the same. So two pulses will travel at the same speed at the same time, no matter where they are. And that's all we really need, to understand the fact that if one pulse leaves our galaxy and is coming towards us-- I should do that with my right hand, because the second pulse is going to be my other hand-- as that second pulse follows it, the second pulse, at any given time-- even though the speed will change with time, but at any given time-- the second pulse will be traveling at the same speed as the first pulse.

And that means that it'll look something like this. The speed might change with time, but as long as they both travel at exactly the same speed at any given time, they will stay exactly the same distance apart in our comoving coordinate system. Δx , the x distance between the two pulses, will not change with time. And if the coordinate distance does not change with time-- the physical distance is always the scale factor times the coordinate distance-- it means that the physical wavelength of the light pulse will simply be stretched with the scale factor, which means you'll be stretched with the expansion of the universe, in exactly the same way as any other distance in this model universe will be stretched as the universe expands. So that's the key idea, and it's very simple, and those words really say it all.

Δx equals constant implies Δl physical is proportional to a of t , and that implies that the wavelength of the light, as a function of t , is proportional to a of t .

Wavelength is actually what I was calling Δl physical, the distance between these two pulses, where each pulse represents a crest of the wave. And λ is the standard letter of the wavelength.

Now the wavelength is related to the period of a wave simply by the relationship that λ is equal to c times Δt . Wavelength is just the distance the wave travels in one period. So if λ is proportional to Δt , so is the time interval, Δt , the period of the wave, going to be proportional to Δt . So we have been defining the redshift in terms of the period. So Δt observed over Δt at the source is equal to λ observed over λ at the source. λ and Δt are proportional to each other. And-- let me finish and I'll get to you, OK?

AUDIENCE: Yes.

PROFESSOR: This then, the ratio of the lengths, is just the amount by which universe has stretched over that time. So just the ratio of the scale factors at the two times. So this is equal to just Δt of the time of observation, which I'll call $t_{\text{sub } o}$, over Δt of the time of the source, $t_{\text{sub } s}$. So this is the scale factor at source, and the numerator here is the scale factor at observation. And this ratio of times, or ratio of wavelengths, or ratio of scale factors, is defined to be $1 + z$, as we have always done. The ratio of the time intervals we had defined originally as $1 + z$, we'll keep that definition, and that defines the redshift shift, z . Question now? Yes.

AUDIENCE: Is that definition of λ , does that have anything to do with the Lorentz invariant? Like, it just kind of struck me as the first term?

PROFESSOR: Not sure what you mean? What-- Lorentz invariant what?

AUDIENCE: Like the $c \Delta \tau$ squared equals $c \Delta t$ --

PROFESSOR: Oh. Well, the Δt could be put into that formula, but that's formula could measure any Δt .

AUDIENCE: Yeah

PROFESSOR: So of course Lorentz is a special case, but any Δt would be a special case of

that formula, so I don't think there's a lot to say about it being a special case.

AUDIENCE: All right, cool.

PROFESSOR: Any other questions? Yes?

AUDIENCE: Is this like fundamentally different? Or is it similar [INAUDIBLE]?

PROFESSOR: [INAUDIBLE] I was going to come to that. That's the question of how the cosmological redshift relates to the special relativity redshift that we derived earlier, and I'm coming to that immediately. Good question, we're getting there. Any other questions, though, before I go there? In my point of view, that's the next topic. OK.

OK, so let me move on to exactly that question. How does this relate to what we already said about the redshift? This answer-- I would like to quantify things and say that it differs in two ways from the calculation that we've done previously. And the first is-- the reason why it's important to us-- is that this actually takes into account, effects which were not taken into account by our earlier calculation. In particular, even though we derived this by a very simple kinematic argument, which didn't seem to involve much math at all, it actually is incredibly strong, in that it encompasses not only special relativity, but also general relativity. It includes all the effects of gravity. If you think about what gravity might do to what we're talking about, gravity doesn't change the fact that the speed of light is going to be c over a of t . That really is just a unit conversion, combined with the fundamental physics assumption that the speed of light is always measured at c , relative to any observer.

So when we put in gravity, this relationship continues to hold-- that was really all we used to drive this-- so gravity is not going to affect the answer. If you think about special relativity, is there something left out? Everything I said here, Newton would have understood perfectly. I didn't have to mention time dilation, which was crucial to our special relativity calculation of the redshift shift.

Did I make a mistake? Is there some place where time dilation should come in here? The answer, really, is no, if you think about it. We had two clocks involved in

our system, a clock on the galaxy, and a clock at us, which we used to measure the period of emission, and the period of reception, but those clocks are each at rest, relative to matter in the region-- even though they're moving with respect to each other-- so by definition, they do measure cosmic time. Cosmic time is a very peculiar kind of time, it's not the time in any inertial frame. These clocks are moving with respect to each other, so if you were defining inertial frame time, their clocks could never be synchronized and would never agree with each other.

But in this concept of cosmic time, they do agree with each other, by construction. And since each clock is at rest, relative to its local matter, it measures this t that we're talking about, this cosmic time variable. And when the pulse arrives at us, when we measure Δt on our clock, that's exactly the quantity that, in the end, we want to talk about-- $\Delta t_{\text{observer}}$. The quantity measured on our clock, which is a clock which also measures cosmic time. So there's no place for any time dilation to enter. It's not that we forgot it, it's not there. It's not part of this calculation.

So this result, as simple as it looks, actually fully encompasses the effects of both special relativity and gravity. Now let me just mention, it's not obvious how gravity came in here. I'm telling you it satisfies-- includes all the effects of gravity. Where is gravity hidden? Let me throw that out as a question. How does gravity affect this calculation, even though I didn't have to mention gravity when I described the calculation? Yeah, in back.

AUDIENCE: The scale factor?

PROFESSOR: That's right, the scale factor. We have not yet talked about how a of t evolves. And the evolution of the a of t will explicitly involve the effects of gravity. And that's why this result depends on gravity, even though we didn't need to use gravity, or say anything about gravity to get the results. So this is the first difference. This calculation includes the effect of gravity. Which is through a of t . Now, because this calculation seems to include everything that the first calculation included and more, you'd expect to be more complicated, but it's less complicated. Could we have saved ourselves a lot of time last week by just giving this calculation, and deriving

the other answer from it. The answer is, not easily, it would not have saved time, one can't, in principle, do it that way.

But the other important difference between these two calculations is the variables that you're using to express your answer. Once you ask a question, if you ask the question vaguely, there could be many different answers to that question, depending on what variables are used to express the answer. So what we're doing here is we're expressing the redshift z for objects which are in fact at rest in the comoving coordinate system. The special relativity calculation-- I think I'm going to need another blackboard. The special relativity calculation, on the other hand gives z as a function of the velocity, as measured in an inertial coordinate system. So the answers are just being expressed in terms of totally different things, and the answer is so simple here because a of t already incorporates a lot of information, and we've just taken advantage of that to be able to give a very simple answer in terms of a of t , without yet saying how we're going to calculate a of t . Yes.

AUDIENCE: [INAUDIBLE] two questions. One is about that constant time.

PROFESSOR: Yes.

AUDIENCE: How is that different than the Newton or Galilean idea of absolute time?

PROFESSOR: OK. The question was how does the notion of cosmic time differ from Newton's or Galileo's notion of absolute time? And the answer is perhaps not much. Operationally, I think it is pretty much the same, but the real point is that Newton and Galileo did not know anything about relative effects like time dilation. So for them, it was just obvious that all clocks ran at the same speed, and time was naturally universal-- naturally absolute. In this case, we're aware of the fact that moving clocks run at different speeds. So if we were to take these clocks between us and the galaxy, and transport one to the other, depending on what path we used to transport them on, in the end, they would probably not agree with each other. So, we're setting up a definition of what we're going to define locally as time, recognizing that what time means here, versus what time means there, is a consequence of our assumptions about how we define things. It is not given

automatically by the fact that all clocks will run at the same speed. Follow up?

AUDIENCE: Yeah. An addition, this is slightly different. So, in the special relativity calculations, [INAUDIBLE] z could be [INAUDIBLE]

PROFESSOR: Absolutely.

AUDIENCE: So here we're only seeing a red shift, but we would obtain a blue shift if we allowed a or t to be decreasing, right?

PROFESSOR: That's right. If the universe contracted, we would get a blue shift.

AUDIENCE: [INAUDIBLE]

PROFESSOR: That's right. It would correspond to the special relativity case. I was going to say a few words about the correspondence, but I'll answer questions first. Yes.

AUDIENCE: I'd kind of like to add on to that question regarding the causal time.

PROFESSOR: Yes.

AUDIENCE: Isn't the fact that you've scaled the speed of light, that's what takes care of this discrepancy between the clocks themselves?

PROFESSOR: The question is, does the fact that we've rescaled the speed of light take care of the discrepancy of times? Well, partially, but it doesn't say anything about what moving clocks will do. If you had a clock moving through this universe, you would have to calculate a time dilation for that clock, just as in any other case.

AUDIENCE: What about the two end points, say, of the path. Is that why you're scaling the speed of light?

PROFESSOR: Not really. The scaling of the speed of light really comes about through the scaling of space. This in fact is just the scale factor that we scale space. Time is measured locally on every clock, and we don't think of it as being rescaled. The speed of light looks different, simply because a notch is changing with time. And that formula tells you how to convert meters per second, which will always be the same to the speed

of light, to notches per second, which will change as the size of a notch changes

AUDIENCE: Right,yeah. OK. I understand.

PROFESSOR: Yes.

AUDIENCE: Further in the line of questioning about cosmological time-- so we expect that us and that other galaxy have simultaneous clocks relative to the cosmic time, and also we expect our own clocks to be simultaneous with our cosmological clocks, I assume. So if we-- Is that true?

PROFESSOR: That's right, yes. Our own clock is just an example of one of the clocks sitting on the place called us, and all clocks sitting on that place will behave the same way. And they define the local definition of cosmic time.

AUDIENCE: So if we take those clocks and move very slowly across to the other galaxy, in cosmological-- in comoving coordinates, we wouldn't expect there to be any time dilation, in the respect that clocks stay simultaneous. Safe to say, that we would think it would be simultaneous with us the whole time, until we got to the other galaxy. And then it would still be simultaneous. But, they're moving at a speed relative to us, so we wouldn't expect [INAUDIBLE].

PROFESSOR: Right. OK. You raise a good question, which I would have to think about the answer. If we brought-- if we carried our clock very slowly to this galaxy, and the limit was infinitely slow, would it agree when it got there? Let me think about that, and answer it next time. I'm not altogether sure. Any other questions? OK. I want to say something about the relationship between these two calculations.

What would happen if we tried to actually compare the answers that we got for the relativistic Doppler shift, and for this answer, for the cosmological redshift. There's really only one case where it would be legitimate to compare them. Since the calculation we just did was supposed to include the effects of gravity, and special relativity calculation does not include the effects of gravity, the only way we should be able to compare them, and see that they agree, would be the case where gravity is negligible.

And one can talk about a cosmological model where gravity is negligible, there's nothing inconsistent about that. If gravity were negligible, what would we expect for the behavior of a of t for this [INAUDIBLE] question. I hear a constant. Constant is certainly a possibility, but it's not the only possibility, so try to think a little harder, and ask if there are other possibilities. Yes.

AUDIENCE: I'm sorry, could you rephrase the question?

PROFESSOR: Rephrase the question. The question is, if gravity were negligible, what would we expect for the behavior of the scale factor a of t ? And so far, it's been suggested that it could be a constant, and that's true, but that's not the most general answer. Yes.

AUDIENCE: It could be negative.

PROFESSOR: Could be negative? I don't know what would mean, actually.

AUDIENCE: What do you--

PROFESSOR: It would mean the universe was inside out.

AUDIENCE: Oh.

PROFESSOR: It would really just mean that you've reversed your coordinates. I don't think it would have any significance.

AUDIENCE: Oh, the expansion would actually be a contraction?

PROFESSOR: Oh, well it could decrease with time, that's not the same as being negative.

AUDIENCE: Oh, I'm sorry

PROFESSOR: It could always increase or decrease with time, whether gravity is present or not. For our universe it's increasing with time, but one could imagine a contracting universe. Yes, Aviv.

AUDIENCE: Linear?

PROFESSOR: Linear. That's right. If there's no gravity, a of t should be a constant times t . The constant could be zero, and then a of t is-- and maybe I should say it should be a constant plus a constant times t , and then in a special case it could just be a constant. But it should vary linearly with time. And that simply means that all velocities are constant. If all velocities are constant, then a of t is varied linearly with time, so that the distance -- the famous relationship is the distance of a of t times t sub c . If this distance were growing linearly with time, it could just be a constant velocity, which is certainly allowed in the absence of gravity. It would mean that a of t was growing linearly in time. So that would be the special case of absence of gravity, a of t growing linearly in time. And one can always set the constant that would be added to the linear to be zero, just by choosing zero of time to be the time at which a of t is zero. So, in the absence of gravity, one can say that a of t should just be proportional to t .

So for that special case, these two calculations should really agree. And it will be, I'm pretty sure, an extra-credit homework problem coming up soon, in which you'll get a chance to calculate that. It's not easy, which is why it will be an extra-credit problem, probably, not required problem, because it involves understanding the relationship between these two coordinate systems. The special relativity answer is given in inertial coordinate system which, when gravity is present, doesn't exist at all. In the presence of gravity, there is no global inertial coordinate system. But without its action, there is. But it's related to this coordinate system, where everything's expanding in a complicated way, because of the various time dilations and Lorentz contractions associated with the motions that are taking place in our expanding universe.

So what you'll need to do is to figure out the relationship between these two coordinate systems. And when you do, and actually compare the answers, is you find that they actually do agree exactly. This is all perfectly consistent with special relativity, but the special case where there is no gravity. OK. Ready to leave cosmological redshift altogether, unless there are any further questions? OK. In that

case, Onward to the next major topic.

We've now finished what I wanted to say about the kinematics of homogeneously expanding universes, and now we're ready to talk about the dynamics. What happens when we try to think about what gravity is going to do to this universe, to be able to calculate how a of t is going to vary with time. That will be the only goal, to understand the behavior of a of t .

Now this problem, in a way, goes back to Isaac Newton. And I might just give a little aside here, and mention that one of the fun things about cosmology, actually, is that if one looks back at the history of cosmology, many great physicists have made great blunders in trying to analyze cosmological questions. And in the discussion today, we'll be discussing one of Newton's blunders. And to me, it's very consoling to know that even physicists as great as Newton can make stupid mistakes. And he actually did make a stupid mistake, in terms of analyzing the cosmological effect of his own theory of gravity.

At issue was Newton's view of the universe, and Newton, like everybody, really, until Hubble, believed that the universe was static. He imagined the universe as a static distribution of stars scattered through space. And early in his career, from what I understand of the history, he assumed that this distribution of stars was finite, and an infinite background space. But he realized at some point that if you had a finite distribution of mass, in otherwise empty space, that everything would attract everything else, with his one over r squared force of gravity-- which he knew about, he invented it-- and the result would be everything would collapse to a point. So he decided that would not work, but he was still sure everything was static. Because everything looked static, stars don't seem to move very much. So he asked what he could change, and decided that instead of assuming that the stars made up a finite distribution, he could assume that they were an infinite distribution, sharing all of space. And he reasoned-- and this is really where the fallacy showed up-- but he reasoned that if the stars filled the infinite space that, even though they would all be tugging on each other through the force of gravity, they wouldn't know which way to go. And since they wouldn't know which way to go, because they'd be tugged in all

directions, they would stand still. So he believed that an infinite, uniform, distribution of mass would be stable-- that there'd be no gravitational forces resulting from the masses in this infinite distribution.

And I have some quotes here, which I think are kind of cute, so I'll show them to you. Newton had a long discussion about these issues with Richard Bentley, the theologian. And we get to read about it, because all these letters have been preserved. In fact, I'm told that the original letters are actually still in existence at Trinity College in Cambridge University. And you can find them on the web even, I'll give you a web reference for the text of these letters, and they're in books and various places.

So let me read this to you. I think it's a cute quotation. "As to your first query"-- by the way, I think we don't have the letters that Richard Bentley sent to Newton, only the responses. But Newton fortunately responded in a way that made the questions pretty clear, so it's not an important problem in understanding what's going on. Newton says, "It seems to me that if the matter of our sun and planets and all the matter of the universe were evenly scattered throughout all the heavens, and every particle had an innate gravity toward all the rest, and the whole space throughout which this matter was scattered was but finite, the matter on the outside space would, by its gravity, tend toward all the matter on the inside"-- this is a finite universe he's talking about -- and he says, "that by its consequence, everything would fall down into the middle of the whole space, and there compose one great spherical mass." So, there he's describing how it would not work if you had a finite collection of matter.

But, he says, "If the matter was evenly disposed throughout an infinite space, it could never convene into one mass, but some of it would convene into one mass, and some into another, to as to make an infinite number of great masses, scattered at great distances from one to another, throughout all that infinite space." So he thought there'd be local coagulation, which of course is what we see in our real world. We see stars that have formed, and now we know about galaxies, which Newton had no way of knowing about. That's the kind of coagulation process that

he's discussing. And he-- oops, sorry. "And thus might the sun and the fixed stars be formed, supposing the matter were of a lucid nature."

That's a cute phrase. I can tell you what it means, it may not be obvious. But at this point, nobody had any idea what the sun was made out of, and why the sun was different from the earth. In fact nobody really had much of a real idea what the earth was made out of either, here. Chemistry wasn't really invented yet. So the assumption was that there were two kinds of matter, lucid matter and opaque matter. Where lucid matter is the stuff the sun is made out of, and the stars, that glows, and is fundamentally different in some way, that was of course not understood at all, from opaque matter, which is what the Earth is made out of. You can't see through it, and it doesn't, obviously, glow. So here, when he's talking about matter forming the stars and the sun, he says if the matter was lucid, if it was the kind of matter that glows.

Going on-- so far, what he said sounds pretty good. Going on, he goes on now to talk more about this lucid versus opaque business. And I think it's cute. I don't know where exactly it's going, but it shows something about Newton's personality, which one might not have known otherwise. "But how the matter should divide itself"-- I should also warn you, all of this is one sentence. If you think sometimes my sentences sound convoluted, just think how lucky you are that you don't have Newton here as your lecturer. This is just impossible. So, "But how the matter should divide itself into two sorts"-- how we'd have lucid and opaque matter in the right places-- "and that part of it which is to comprise a shining body should fall down into one mass, and make a sun, and the rest, of which is fit to compose an opaque body, should coalesce not into one great body, like shiny matter, but into many little ones"-- somehow he's forgotten about the stars here, when he's talking about the sun and the planets, many planets and one sun.

So he says that, "how the opaque matter should fall instead into many little masses"-- and then he talks about other possibilities. It's wonderful the way he lists all the possibilities. "Or," he says, "if the sun were at first an opaque body, like the planets, or if the planets were lucid bodies like the sun, how he alone"-- he being

the sun, if you track everything back, "how the sun alone should be changed into a shiny body, while all the"-- lost track -- "where all they"-- of the planets-- "continue to be opaque, or"-- he's considering all possibilities-- "or they all be changed to opaque ones, while he,"-- the sun-- "remains unchanged as a lucid one."

He does not know how to explain all that, is what he's saying. Bottom line of the sentence is, I don't know, I don't have a clue. And he says, "I don't think it's explicable by mere natural causes, but am forced," Newton says, "to ascribe it to the council and contrivance of a voluntary agent." So the theory of intelligence design, as well the theory of gravity, actually both go back to Newton, it turns out. Newton was a very religious person, and in certain aspects of physics, he was happy to ascribe to a voluntary agent, as he calls it. I have some references here, and I'll be posting this so you'll be able to read those references and type them in, if you want.

Now Newton decided that you could not have a finite distribution, because it would collapse. If you had an infinite distribution, he thought it would be stable, but he apparently had heard different arguments to that same conclusion. And one argument that you might give for saying that the infinite distribution would be stable would be the argument that if you look at the force on any one particle, there is an infinite force pulling it to the right-- my right, your left-- and an infinite force pulling it to my left, your right, and since they're both infinite, they would cancel each other. Newton did not accept that argument. He was sophisticated enough to realize that infinity minus infinity isn't necessarily zero. And he has a bit of a tirade on that, that I thought was worth quoting.

And this is a second letter to the same Richard Bentley. I guess it was Bentley who made this argument, and Newton rejected it. Infinity minus infinity, Newton realized, is ambiguous. It's not something that we should necessarily think of zero. "But you argue in the next paragraph of your letter that every particle of the matter in the infinite space has an infinite quantity of matter on all sides, and by consequence, an infinite attraction every way, and therefore must rest in equilibrio, because all infinities are equal:"-- he's summarizing Richard Bentley's argument-- "yet you suspect a parologism"-- that means logical error, I think-- "in this argument: and I

can see the paralogism lies in the position that all infinities are equal. The generality of mankind consider infinities no other ways than indefinitely"-- and in this sentence they said all infinities are equal-- "though they would speak more truly if they should say that they are neither equal nor unequal, nor have any certain difference or proportion, one to another."

So he realizes that the ratio of infinity could be anything, and infinity minus infinity could be anything, all of which is consistent with our modern view of how to do the mathematics. "In this sense, therefore, no conclusions can be drawn from them about the equality, proportions or differences of things, and they that attempt to do so usually fall into paralogisms." He goes on, now I just have one more Newton quote-- I like Newton quotes--

I have one more Newton quote, again from the same series of letters. These are all from 1692 and 1693, I believe, where he gives an example-- I think this follows the quotes of the previous slide immediately-- where he gives an example of a false argument that you get into-- and apparently it's an argument that he had heard from other people-- if you think all infinities are equal. What he says is, "So when men"-- he doesn't say who men are, and I don't really know the history. He may referring to some particular philosophers at the time-- "when men argue that the infinite divisibility of magnitude by saying that an inch may be divided into an infinite number of parts, the sum of those parts would be an inch-- and a foot can be divided into an infinite number of parts, the sum of those parts must be a foot-- and therefore, since all infinities are equal, these sums must be equal."

Understand the argument here. He's saying that if you divide an inch into an infinite number of parts-- this is all you've been given as a foil. He's not claiming the argument is right, he's claiming it's wrong-- that argument is that if you divide an inch into an infinite number of parts, you get an infinite number of points, if you put them together, you get an inch. If you divide a foot into an infinite number of parts, you get an infinite number of points, and if you put them together, you should get a foot. But they're both an infinite number of points in the description. So if you think all infinities are equal, the infinite number of points that make an inch should be the

same as the infinite number of points that make a foot, therefore a foot should equal an inch, obviously. Right. Not right, he know.

So he says that the falseness of the conclusion shows an error in the premises, and the error lies in the position that all infinities are equal. So Newton has given us a very nice example of how you can convince yourself that you get into logical paradoxes if you pretend that all infinities are equal. But, this does not change the fact that Newton was still convinced that an infinite distribution of mass would be stable. The argument that convinced him was not the infinity on each side, but rather the symmetry. Newton's argument, the one he believed, was that if you look at any point in this infinite distribution, if you look around that point, all directions would look exactly the same, with matter extending off to infinity, and therefore there'd be no direction that the force should point on any given particle. And if there's no direction in that force at the point, it must be zero. That was the argument Newton believed.

OK. What I want to do now is to talk about this in a little bit more detail, and try to understand how modern folks would look at the argument. And by the way, I might just add a little bit more about the history first. Newton's argument, as far as I know, was not questioned by anybody for hundreds of years, until the time of Albert Einstein. Albert Einstein, in trying to describe cosmology using his new theory of general relativity, was the first person, as far as I know, to realize that even if you had an infinite distribution of mass, it would collapse-- and we'll talk about why. And Einstein did realize that the same thing would happen with Newtonian physics, it's not really a special feature of general relativity, it just somehow historically took the invention of general relativity to cause people to rethink these ideas and realize that Newton had been wrong. So, what's going on.

The difficulty in trying to analyze things the way in which Newton did is that Newton was thinking of gravity, in the language that he first proposed it, as a force at a distance. If you have two objects in space, the distance r apart, they will exert a force on each other proportional to one over r squared. Since the time of Newton, other ways of describing Newtonian gravity itself have been invented, which make it

much more clear what's going on. The difficulty in using Newton's method-- we'll talk about in more detail in a few minutes-- but it's simply that we try to add up all of these one over r squared forces, you get divergent sums that you have to figure out how to interpret. But to understand that Newton couldn't possibly have been right, the easiest thing to do is to look at other formulations of Newton's gravity. And I'll describe two of them, both of which will probably have some familiarity to you.

The first one I'm quite sure will. And I'm going to describe it by analogy with Coulomb's law, because 802 goes a little further with Coulomb's law than any course you are likely to have taken has gone with gravity. But Coulomb's law is really the same as the force law of gravity. So Coulomb's law says that any charged particle will create an electric field, which is the charge divided by the distance squared times the unit vector pointing radially outward. That's Coulomb's law. People can-- sometimes there's constants in here, depending on what units you measure q in, but that won't be important for us. So I'm going to assume we're using this where that constant is one.

You know that Coulomb's law can be reformulated in terms of what we call Gauss's law. If Coulomb's law is true, you can make a definite statement about what happens when you integrate the flux of the electric field over any surface. It's proportionate to the total amount of charge inside. So Coulomb's law implies Gauss's law, which says that the integral over any closed surface of $\mathbf{E} \cdot d\mathbf{a}$ is equal to 4π times the total enclosed charge. q encloses the total amount of charge inside that volume. And what constants appear depends on what constants appear here, which depends on what units you're using, but these equations are consistent. Those are the correct constants, if you measure charge in a way which makes the electric field be given by that simple formula.

OK, so I'm going to assume you know this, that you learned it in 802 or elsewhere. If this is true, then, since this is the same inverse square law, if we write down Newton's law of gravity, almost as Newton would have written it, we can express it as the acceleration of gravity at a given distance from an object. So we could write Newton's law of gravity by saying the acceleration of gravity is equal to minus

Newton's constant times the mass of the object, the analog of the charge up there, divided by r^2 times \hat{r} .

Again, it's the inverse r^2 law, and the point radiating outward is just like Coulomb's law, except for the constant out front. The constant actually has the opposite sign, which is important for some issues, but not for what we're saying now. The important point is that this can also be recast as a Gauss's law, and it's called Gauss's law of gravity. And the only thing that differs is a constant out front, so it's a trivial transformation. The integral over any closed surface of the gravitational acceleration vector, $\mathbf{g} \cdot d\mathbf{a}$ is equal to minus $4\pi G$ times the total mass enclosed. The only difference is the minus sign, and the factor of G , which follow from the difference of the minus sign and the factor of G in the formula on the left. OK, does everybody believe that?

OK, now let's think about this homogeneous distribution of mass that Newton was trying to think about. Newton's claim was that you could have a homogeneous distribution of mass filling all of the infinite space, and that would be static, that is, there would be no acceleration. No acceleration means Newton is claiming in this language that \mathbf{g} could be zero everywhere. But if you look at this formula, if \mathbf{g} is zero everywhere, then the integral of \mathbf{g} over any surface is going to zero, and therefore the total mass enclosed had better be zero. But if we have a uniform distribution of mass, the total mass enclosed will certainly not be zero for anything with non-zero volume. So clearly this assertion that the system would be static was in direct contradiction with the Gauss's law formulation of Newton's law of gravity.

Just for the fun of it, I'll give you another similar argument using another more modern formulation of Newtonian gravity. Another way of formulating Newtonian gravity, which you may or may not have seen-- and if you haven't seen it, don't understand what I'm saying, don't worry about it, it's not that important. But for those of you who have seen it, I'll give you this argument. Another way of formulating Newtonian gravity is to introduce the gravitational potential. So I'm going to use the letter ϕ for the gravitational potential. I'll tell you in a second how that relates to gravity-- well, I guess I'll tell you now. It's related to the gravitational

acceleration by g is equal to minus the gradient of ϕ , and gradient of ϕ is something that you probably all learned in 802, but I'll write down the formula anyway. It's equal to \hat{i} , a unit vector in the x direction, times the derivative of ϕ with respect to x , plus \hat{j} , a unit vector in the y direction, times the partial of ϕ with respect to y , plus \hat{k} times the partial of ϕ with respect to z . And once one defines this gravitational potential, one can write down the differential form of the Gauss's law, which becomes what's called Laplace's equation. And it says the $\nabla^2 \phi$ is equal to 4π times Newton's constant times ρ , where ρ is the mass density.

And this is called Laplace's equation, and if you're given the mass density, it allows you to find the gravitational potential, and then you can take its gradient, and that determines what g is. And it's equivalent to the other formulations of gravity. But it gives us another test of Newton's claim that you could have a homogeneous distribution of matter, and no gravitational forces. If there are no gravitational forces, then g would have to be zero, as we said a minute ago, and this formulation of g is zero, that implies the gradient of ϕ is zero.

If we look at the formula for the gradient, it's a vector. For the vector to be zero, each of the three components has to be zero, and therefore the derivative of ϕ with respect to x has to vanish, the derivative of ϕ with respect to y has to vanish, the derivative of ϕ with respect to z has to vanish, that means ϕ has to be constant everywhere, it has no derivative with respect to any spacial coordinate. So if g vanishes, the gradient of ϕ vanishes, and ϕ is a constant throughout space. And if ϕ is a constant throughout space, now we can look at this formula-- and I forgot to write down the definition of ∇^2 . $\nabla^2 \phi$ is defined to be the second derivative of ϕ with respect to x squared, plus the second derivative of ϕ with respect to y squared, plus the second derivative of ϕ with respect to z squared. So if ϕ is a constant everywhere, as it would have to be if there were no gravitational forces, then one can see immediately from this equation that $\nabla^2 \phi$ would have to be zero, and one can see from this equation that ρ would have to be zero, there would have to be no mass density. But Newton wanted to have a non-zero mass density, the matter of the universe spread out uniformly

over an infinite space. So this is another demonstration that Newton's argument was inconsistent. Yes.

AUDIENCE: I'm sorry, what does phi represent?

PROFESSOR: Phi is really defined by these equations, it's defined, really, by this equation. The name is that it's the gravitational potential.

AUDIENCE: Potential.

PROFESSOR: And its physical meaning is simply that it gives you another way of writing what g is.

AUDIENCE: Yeah.

PROFESSOR: Any other questions? OK, so the conclusion seems to be that Newton has not gotten the right answer, here, but we still have to analyze Newton's argument a little bit more carefully, to see exactly where he went wrong. So, the next thing I want to talk about is the ambiguity associated with trying to add up the Newtonian gravitational forces, as Newton was thinking, for an infinite universe. I mentioned that the real problem with Newton's calculation is that the sum he was calculating actually diverges, and you have to be more careful about trying to calculate it in a reliable way.

So to make this clear, I want to begin by giving an example of this general notion of integrals that give ambiguous values. And I want to define just a couple of mathematical terms. I want to consider just-- again, starting talking about general functions, and when integrals are well defined and when they're not. I want to imagine that we just have some arbitrary function f of x where x would not just be one variable.

We'll generalize this to three dimensions, which is the case that we'll be interested in, but we'll start by talking in terms of one variable. If we have a function f of x , we can discuss what I'll call I_1 , which is the integral, from minus infinity to infinity, of f of x dx . This is exactly the kind of integral that you're thinking of when we wanted to-- thinking about adding up all the gravitational forces acting on a given

body. Now I want to consider the case where I_1 is finite.

I'm sorry. I need to first define more carefully what I mean by I_1 . OK, to even define what you mean by this minus v to infinity, you should say something a little bit more precise. So we could define I_1 a little bit more precisely, and I'll call this I_1 prime, for clarity. This will really just be a clearer way of describing what one probably meant when one wrote the first line. We can define the integral from minus infinity to infinity as the limit, as some quantity L goes to infinity, of the integral from minus L to L of f of x dx . So this says to do the integral from minus L to L , and if we assume f of x is itself finite, this is always finite. I will assume f of x itself is finite, we'll only worry about the convergence of the integral. So for any given L , this is a number, then you can ask, does this number approach a limit as L goes to infinity? And if it does, you say that's the value of this integral. That just defines what we mean by the integral from minus infinity to infinity.

I want to now consider the case where that exists. So consider the case where I_1 prime is-- I'll write is less than infinity, meaning it has some finite value. The limit as L goes to infinity exists. But now, I want to also consider-- and I'll move on to the next blackboard-- to consider this-- consider an integral that I'll call I_2 , for future reference, which is just defined to be the integral from minus infinity to infinity. Defined as the same kind of limit that we used here, but I won't rewrite it. I'll just assume that the integral from minus infinity to infinity means that limit. But I want to consider the integral from minus infinity to infinity of the absolute value of f of x dx .

And now I want to introduce some terminology. If I_2 is less than infinity, if it converges, then I_1 is called absolutely convergent. So absolutely convergent means that it would converge, even if you had absolute value signs. Conversely, this I_2 is divergent-- and I'll just write that as I_2 equals infinity, if that limit does not exist, if it's a divergent integral. But remember, we assumed I_1 did exist, so I_1 still converges, but it's called conditionally convergent. So if an integral converges, but the integral of the absolute value of that same function does not converge, that's the case that's called conditional convergence.

And the moral of the story, that I'll be beginning to tell you now, is that conditionally convergent integrals are very dangerous. What makes them dangerous is that they're not really well defined. You can get any value you want by adding up the integrand in different orders. As long as you stick to a particular order, which is how we define the symbol, you will get a unique answer, but if, for example, you just shift your origin, you can get a different answer, which is something you don't usually expect. You usually think of just integrating over the whole real line, it doesn't matter what you took to be the center of the line. So things become much less well defined when one is discussing conditionally convergent integrals.

And before we get to the particular integral that we're really interested in, which is trying to add up the gravitational forces of an infinite distribution of matter, which I'll get to, I'm going to give you an example of a very simple function that just illustrates this ambiguity, that the integral converges, but is not absolutely convergent. You can get any answer you want by adding it in different orders-- adding up the pieces of the integral in different orders. So let me consider an example-- and this is just to illustrate the ambiguity-- the example I'll consider will be a function f of x , which is defined to equal plus 1 if x is greater than zero, and minus one if x is less than zero. And I have neglected to specify what happens if x is exactly equal to zero, but when you integrate, that doesn't matter. A single point never matters. So you could measure it's anything you want at x equals zero, it won't change anything you're going to be saying.

Let me draw a graph of this. f of x versus x . I'll put plus 1 there, and minus 1 there. The function is plus 1. Maybe I have a little bit of colored chalk here to draw the function. The function is plus 1 for all positive values of x , and minus one for all negative values of x . And there's the function. And if we integrate it symmetrically, following this definition of what we mean by integrating from minus infinity to infinity, we do get a perfect cancellation. When you integrate from minus L to L , we get zero, because you get a perfect cancellation between the negative parts and the positive parts. And then if you take the limit as L goes to infinity, the limit of zero is zero. There's not really any ambiguity to that statement.

So in the order specified, this has unique integral, which is zero. But, it depends on how you've chosen to add things up. In particular, if you just change your origin, and integrate starting moving outward from the new origin, you'll get a different answer, and that's what I want to illustrate next. Suppose-- suppose we consider the limit as L goes to infinity, we'll pick the limit the same way, but instead of integrating from minus L to L , we can integrate from a minus L to a plus L of f of x dx . Now this is really the same integral, we've just basically changed our origin by integrating from a outwards. In the special case a equals zero, it's exactly the same as what we did before, but if a is non-zero, it means that our integral is centered about x equals a , instead of centered about x equals zero.

So we can draw that on the blackboard. If we let a be over here, our integral will go from a minus L , and that will be to the left of distance L , you will extend to a plus L , which will be to the right by distance L . The integral defined by the equation on the blackboard at the left will correspond to that region of integration. And the specification is that we should do that interval first, and then take the limit as L goes to infinity, and see what we get.

It's easy to see what we will get. Once L is bigger than a , you can see that the answer won't change any more, as we make L bigger. As you make L bigger, we will always be adding a certain amount of minus 1 on the left, and certain amount-- the same amount of plus 1 on the right, and they will cancel each other once L is bigger than a . And we don't care about small L , because we're only interested in taking the limit of large L , but we should look at what happens when L equals a , and then from any bigger value of L will give us exactly the same number. And when L equals a , the integral will go from 0 up to $2a$ -- a plus L which is a equals L , so that's $2a$. So the integral will be only on the positive side, and we'll have a length of $2a$, and that means the integral will be $2a$, because we're just integrating one from 0 to $2a$. And that will be what we get for any bigger value of L also, because as we increase L , as I said, we just get a cancellation between adding more plus 1 on the right, and adding more minus 1 on the left. So this limit has a perfectly well defined value, which is $2a$.

And a is just where we chose to start integrating, so a could be anything. We could choose a to be anything we want if we're free to integrate in any order. So we can get any answer we want, if we're free to integrate in any order, to add up the pieces of this integral in the order that we choose. And that is a fundamental ambiguity of conditionally convergent integrals. And what we'll see is that trying to add up the force on a particle in an infinite mass distribution is exactly this kind of conditionally convergent integral. And that's why you get any answer you want, and it doesn't really mean anything unless you do things very carefully.

OK. Let's move on. We only have a few minutes left, which I guess means I will set up this calculation, but not quite get the answer, and we'll continue next time. I actually have some diagrams here on my slides. What I want to do now, is calculate the force on some particle in an infinite mass distribution, and show you that I can get different answers, depending on what order I add things up. I will add things up in a definite order at each stage, so I will get a definite answer at each stage, though I'll get different answers, depending on what ordering I choose.

So, we're going to start by trying to calculate-- and the only thing [? of interest, ?] actually, in calculating the gravitational force on some point, p in an infinite distribution of mass. Mass fills the slide, and everything, out to infinity. And we're going to add up that mass in contributions that are specified.

And for our first calculation, we're going to add up the forces for masses that are defined in concentric shells, where we're going to take the innermost shell first, then the second shell, then the third shell, going outward from the center. In that case, it's easy to see that the force on p calculated in that order of integration is 0, because every shell has p exactly at the center, and by symmetry, it has to cancel exactly. In fact, we know-- and we'll use this fact shortly-- that the gravitational field of a shell, inside the shell, is exactly zero-- Newton figured this out-- and outside the shell, the gravitational field of a shell looks exactly the same as the gravitational field of a point mass located at the center of the shell with the same total mass. So we're going to be using those facts. And clearly those facts indicate that, for this case, the answer is 0. P equals 0.

Now we're going to consider a more complicated case-- going too far, here, don't want to tell you the answer yet-- this more complicated case, we're going to still calculate the force at the point p , but we're going to choose concentric spherical shells which are centered around a different point, q . So q just defines the shells that we're going to use for adding things up, and we're still going to add up all the shells out to infinity, so we're going to be adding up the force on p due to the entire infinite mass distribution, but we'll be taking those contributions in a different order, because we're going to be ordering it according to shells that are all centered on q , starting with the innermost, and then the second, and then the third, and so on. Now in this case, we can first talk about the contribution of the shaded region, which are all the shells around q which have radii which are less than the distance to p . For all of these shells, p lies outside the shell. And therefore all of those shells act just like a point mass, with the same total mass concentrated at q , the center of all those spheres. So the mass that's in the shaded region will give a contribution to the force at p , which is just equal to the force of the mass given by the same total mass the point q , located at q .

On the other hand, all the shells outside will be shells for which p is inside. p is no longer at the center of those shells, but Newton figured out, and I'll assume that we all believe, it doesn't matter. Inside the spherical shell, the gravitational force is zero anywhere, no matter how close you are to the boundaries. It just cancels out perfectly. As you get closer to one boundary, you might think you'd be pulled toward that boundary, but-- let me just tell you what's happening here-- as you get closer to one boundary, it is true that the force pulling you towards the particular particles at the boundary get stronger, because it's 1 over r squared, but as you get close to this boundary, there's more mass on this side, because all the mass except for a little sliver is on the other side. And those two effects cancel out exactly.

So the force on a particle inside a shell is exactly zero, as you can prove very easily by the way, from the Gauss's law of formulation of gravity. And therefore, the outer shells give no contribution. So we've completely calculated now the force at p is just equal to the force due to the shaded mass. It's just given by that simple formula, it's

g times the total mass, divided by b squared, that would be it's distance between q and p. And it's non-zero. So you get 0 or non-zero answer, depending on what ordering you chose for adding up the pieces of the mass that are going to make up this infinite distribution. And furthermore, this answer could be anything you want, because I could let b be anything I want. And this answer depends on b, and becomes arbitrarily large in magnitude as b gets bigger. The mass grows like b cubed. It might look like it falls with b, but actually it grows with b. And we could get a point in any direction, by choosing q on any side we want of p, so we can get, really, any answer we want by using this particular way of adding up the masses. Yes.

AUDIENCE: Well, although we can get any answer we want, every answer [INAUDIBLE]

PROFESSOR: Every answer, say again?

AUDIENCE: Like every single one of those answers corresponds to a setup. I mean like the g equals 0, [INAUDIBLE]

PROFESSOR: Well the reason it's a problem is that these shells don't really exist. We're just thinking about these shells. The shells only determine what order we are going to use for adding up the different contributions. The matter is just uniformly distributed and there's no shells present. The shells are purely a mental construct, which should not affect the answer. This is not part of the physical system at all. The shells only reflect the order that we have used to add up the masses.

So we'll stop there. If anybody has questions, we can talk after class, and we can talk more about the question at the beginning of the next class, but class is over for now.