

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](http://ocw.mit.edu).

**PROFESSOR:**

I really stated everything about discrete coding as clearly as I could in the notes. I stated it again as clearly as I could in class. I stated it again as clearly as I could in making up problems that would illustrate the ideas. If I talked about it again it would just be totally repetitive. So, at this point, if you want to understand things better, you gotta come up with specific questions and I will be delighted to deal with them. So we want to go on--. Oh, there's one other thing I wanted to talk about. We're not having a new problem set out today. I don't think most of you would concentrate on it very well. I'll tell you what the problems are, which will be due on October 14th. I'll pass it out later. It's problems 1 through 7, and the end of lectures 8 to 10, and one other one all the way at the end, problem 26. So, 1, 2, 3, 4, 5, 6, 7 and 26. So you can get started on them whenever you choose and I'll pass out a traditional problem set form next time.

Last time we started to talk about the difference between Riemann and Lebesgue integration. Most people tell me this is further into mathematics than I should go. If you agree with them after we spend a week or two on this, please let me know and I won't torture future students with it. My sense is that in the things we're going to be dealing with for most of the rest of the term, knowing a little bit extra about mathematics is going to save you an awful lot of time worrying about trivial little things that you shouldn't be worrying about. In other words, the great mathematicians of the 19th century who developed -- yeah, the 19th century, but partly the 20th century. These mathematicians were really engineers at heart. The 20th century and the 21st century, mathematics has very much become very separated from physics and applications.

But in the 19th century and in the early 20th century, mathematicians and physicists were almost the same animals. You could scratch one and you'd find another one

there. They were very much interested in dealing with real things. They were like the very best of mathematicians everywhere and the very best of engineers everywhere. They really wanted to make life simpler instead of making life more complicated. One way that many people express it is that mathematicians are lazy, and because they're lazy, they don't want to go through a lot of work, and therefore, they feel driven to simplify things. There's an awful lot of truth in that. What we're going to learn here I think will, in fact, simplify what you know about Fourier series and Fourier integrals a great deal. Engineers typically don't worry about those things.

An awful lot of engineers, and unfortunately, even those who write books often state theorems and leave out the last clause of the theorem, and the last clause of most of those theorems the only way to make them true is to add the clause or not, as the case may be, at the end of it, which makes what they state absolutely meaningless, because anything you can add or not, as the case may be, and it becomes true at that point. The whole question becomes well, what are those cases under which it's true. That's what we're going to deal with a little bit here. We're going to say just enough so you start to understand these major things that cause problems, and I hope you will get to the point where you don't have to worry about them after that. Well, the first thing, we started talking about it last time, we were talking about the difference between Riemann integration and Lebesgue integration.

Riemann was a great mathematician, but he came before all of these mathematicians who were following Lebesgue, and who started to deal with all of the problems with this classical integration, which started to fall apart throughout most of the 19th century. What Riemann said, well, split up the axis into equal intervals, then approximate the function within each interval, add up all of those approximate values, and then let this quantization over the time axis become finer and finer and finer. If you're lucky, you will come to a limit. You can sort of see when you get to a limit and when you don't get to a limit. If the function is smooth enough and you break it up finely enough, you're going to get a very good approximation, and if you break it up more and more finely, the approximation gets better and

better.

If the function is very wild, if it jumps around wildly, and we'll look at examples of that later, then this doesn't work. We'll see why, in fact, this approach does work.

Lebesgue said, no, instead of quantizing along this axis, quantize on this axis. So he said, OK, start with a zero, quantize into epsilon, 2 epsilon, 3 epsilon and so forth, and everybody when they talk about epsilon they're thinking about something small. They're also thinking about making it smaller and smaller and smaller and hoping that something nice happens.

So then what he said is after you quantize this axis, start to look at how much of the function lies in each one of those little windows. I've drawn that out here,  $\mu_2$  is the amount of the function that lies between 2 epsilon and 3 epsilon. Now the function lies between 2 epsilon and 3 epsilon starting at this point, which I've labeled  $t_1$ , going up to this point, which I've labeled  $t_2$ , on over here. It's not in this interval until we get back to  $t_3$ . It stays in this interval until  $t_4$ . Lebesgue said, OK, let's say that the function is between 2 epsilon and 3 epsilon over a region of size,  $t_2$  minus  $t_1$ , which is this size interval, plus  $t_4$  minus  $t_3$ , which is this size interval. Instead of saying size, he said size gets too confusing so I'll call it measure instead. That was the beginning of measure theory, essentially. Well, in fact other people talked about measure before Lebesgue. There are a lot of famous mathematicians who were all involved in doing this. Anyway, that's the basic idea of what measure is concerned with.

Now, for this curve here, it's a nice smooth curve, and you can almost see intuitively that you're going to get the same thing looking at it this way or looking at it this way. In fact, you do. Anyway, what he finally wound up with is after saying what the measure was on each one of those little slices, how much of the function lay in each one of those intervals, he would just add them all up. He would add up how much of the function was in each slice, he would multiply how much of the function was in a slice by how high the slice was up, and then he'd get an answer. One difference between what he did and what Riemann did was that he always got a lower bound in doing it this way. If he was dealing with a non-negative function, his

approximation was always a little less than what the function was, because anything that lay between 2 epsilon and 3 epsilon, he would approximate it as 2 epsilon, which is a little less than numbers between 2 epsilon and 3 epsilon. So this is a lower bound, whereas this is whatever it happens to be -- however you decide to approximate the function there, and there are lots of ways of doing it.

Well, I won't prove any of these things, I just want to point them out so that when you get frustrated with this, you can always rely on this. Which says that whenever the Riemann integral exists, in other words, whenever the integral that you're used to exists, mainly, whenever it has meaning, Lebesgue integral gives you the same value. In other words, you haven't lost anything by going from Riemann integration to Lebesgue integration. You can only gain, you can't lose. The familiar rules for calculating Riemann integrals also apply for Lebesgue integrals. You remember what all those rules are, you probably know all of them better even than this fundamental definition of an integral, which is split up the function into tiny little increments, because throughout all of the courses that you've taken, learning about integration, learning about differentiation, learning about all of these things, what you've done for the most part is to go through exercises using these various rules.

So, you know how to integrate lots of traditional functions. You memorized what the integral of many of them is. You had many ways of combining them to find out what the integral of many other functions is. If you program a computer to calculate these integrals, a computer can do it both ways. It can either use all the rules to find out what the value of an integral is, or it can chop things up finely and find out that way. As I said before, if you think that being able to calculate integrals is what engineering is about, think again. I told you before you could be replaced by a digital computer. It's worse than that. You could be replaced by your handheld Palm, whatever it's called. You can be replaced by anything. After a while we're going to wear little things embedded in our body that will tell us when we're sick and things like this. You can be replaced by those things even. So you really want to learn more than just that.

For some very weird functions, the Lebesgue integral exists, but the Riemann

integral doesn't exist. Why do we want to worry about weird functions? I want to tell you two reasons for it -- I told you about it last time. One thing is an awful lot of communication theory is concerned with going back and forth between the time domain and the frequency domain. When you talk about things which are straightforward in the time domain, often they become very, very weird in the frequency domain and vice versa. I'm going to give you a beautiful example of that probably in the next class. You can look at it -- it is in the appendix. You'll see this absolutely weird function which has a perfectly well-defined Fourier transform. Therefore, the inverse Fourier transform of this nice looking thing is absolutely weird.

The other reason is even more important. We have to deal with random processes. We have to deal with noise functions, which are continuously varying functions of time. We have to deal with what we transmit, which is continuously varying functions of time. Just like when we were dealing with sources, we said if we want to compress a source it's not enough to think about what one source sequence is. We have to think about it probabilistically. Namely, we have to model these functions as sample values of what we call a stochastic process or a random process. In other words, when we start talking about that, these functions which already look complicated just become sample points in this much bigger space where we're dealing with random processes.

Now, when you're dealing with sample points of these random processes, weirdness just crops up as a necessary part of all of this. You can't define a random process which consists of only non-weird things. If you do you get something that doesn't work very well, it doesn't do much for you. People do that all the time, but you run out of steam with it very, very quickly. So for both reasons we have to be able to deal with weird things. The most ideal thing is to be able to deal with weird things and get rid of them without even thinking about them. What's the nice thing about Lebesgue theory? It let's you get rid of all the weirdness without even thinking about it. But in order to understand it to start with so we get rid of all that weird stuff, we have to understand just a little bit about what it's about to start with. So that's where we're going. Well, in this picture -- let me put the picture back up again -- I

sort of sluffed over something. For a nice, well-behaved function, it's easy to find out what the measure of a function is. Namely, what the measure is of the set of values where you're in some tiny little interval, and that's straightforward, it was just the sum of a bunch of intervals. But now we come to the clincher, which is we want to be able to deal with weird functions too, and for weird functions this measure, the set of values of time in which the function is in some tiny little slice here, the measure of that is going to become rather complicated. Therefore, we have to understand how to find the measure of some pretty weird sets.

So, Lebesgue went on and said OK, I'll define how to find measure of things. It's all very intuitive. If you have any real numbers,  $a$  and  $b$  -- if I have two real numbers I might as well say one of them was less than or equal to the other one, so  $a$  is less than or equal to  $b$ . I want to include minus infinity and plus infinity here. When we say the set of real numbers, we don't include minus infinity and plus infinity. When we talk about the extended set of real numbers, we do include minus infinity and plus infinity. Here for the most part when we're dealing with measure theory, we really want to include minus infinity and plus infinity also because they make things easier to talk about.

So what Lebesgue said is for any interval the set of points which lie between  $a$  and  $b$ , and he started out with the open interval -- namely the set of points not including  $a$  and not including  $b$ , but including all the real numbers in between -- the measure of that is what we said before, and none of you objected to it. The measure of an interval ought to be the size of the interval, because after all, all Lebesgue was doing was taking size and putting another name on it because we all thought of it as being size. So the measure of the interval  $ab$  is  $b$  minus  $a$ . Now as engineers we all know that it doesn't really make any difference whether you include  $a$  or you don't include  $a$  when we're trying to find the size of an interval.

So, he said OK, the size of that interval is  $b$  minus  $a$ , whether or not you include either of the end points. Then he went on to say something else, the measure of a countable union of disjoint intervals is the sum of the measure of each interval. We already said when we were trying to figure out what the Lebesgue integral was, that

if you had several intervals, the measure of the entire interval, namely the measure of the union of those intervals ought to be the sum of the measure of each interval. So all he's adding here is let's go on and go to the limit and talk about a countably infinite number of intervals and use the same definition.

Now there's one nice thing about that and what is it? If we take the sum of a countable set of things, each of those things is non-negative. So what happens when we take the sum of a bunch of non-negative numbers? There are only two possibilities when you take the sum of even a countable set of non-negative numbers and what are they?

**AUDIENCE:** Either to the limit or it goes to infinity?

**PROFESSOR:** It goes to a limit or it goes to infinity, yes. Lebesgue said well, let's say if it goes to infinity that's a limit also. So, in other words, there are only two possibilities -- the sum is finite, you can find the sum, and the sum is infinite. Nothing else can happen. So that's nice.

Then Lebesgue said one other thing. A bunch of mathematicians trying to deal with this did different things when they were trying to deal with very small sets. Some of them said well these very small sets aren't measurable, others, and part of Lebesgue's genius was that he said if you can take a set of points and you can put them inside another set of points, which doesn't amount to anything, then this smaller set of points should have a measure which is less than or equal to the bigger set of points. So that that says that any subset of something that has zero measure also ought to have zero measure. Now when we look at some examples you'll see that that's not quite as obvious as it seems. But anyway, he said that.

An even nicer way to put that is when you're talking about weird intervals, try to cover that weird set with a bunch of intervals. If you can cover it with a bunch of intervals in different ways and the measure of the bunch of intervals, the countable set of intervals, is you can make it arbitrarily small, then we also say that this set has measure zero. So any time you have a set  $s$  and you can cover it with intervals which have arbitrarily small measure, namely we can make the measure of those

intervals as small as we want to make them, we say  $S$  has zero measure. Now that's going to be very nice, because it let's us get rid of an awful lot of things. Because any time we have some set which has zero measure in this sense, when we look back at what we did with the Lebesgue integral, it says good, forget about it. If it has zero measure it doesn't come into the integral at all. That's exactly the thing we want. We'd like to get rid of all that stuff. So if we're going to get rid of it we have these sets which have measure zero, they don't amount to anything, and we're not going to worry about them.

Let's do an example. It's a famous example. What about the set of rationals which lie between zero and 1? Well, rational numbers are numbers where there's an integer numerator and an integer denominator, and people have shown that there are numbers which are approximated by rational numbers but they're not rational numbers. You can take that set of rational numbers and you can order them. The way I've ordered them here, you can't order them in terms of how big they are. You can't start with the smallest rational number which is greater than zero because there isn't any such thing. Whatever rational number you find which is greater than zero, I can take a half of it, that's a rational number, and that also is greater than zero and it's smaller than your number. If you don't like me getting the better of you, you can then take half of that and come up with a smaller number than I could find, and we can go back and forth on this as long as we want.

So the way we have to order these is a little more subtle than that. Here we're going to order them in terms of first the size of the denominator and next, the size of the numerator. So there's only one fraction in this interval -- oh, I screwed that up. I really wanted to look at the set of rational in the open interval between zero and 1, because I don't want zero in it, and I don't want 1 in it, but that's a triviality. You can put zero and 1 in or not. So, anyway, we'll leave zero and 1 out. So I start out with  $1/2$  -- that's the only rational number strictly between zero and 1 with a denominator of 2. Then look at the numbers which have a denominator of 3, so I have  $1/3$  and  $2/3$ . I look then at the rational numbers which have a denominator of 4. I have  $1/4$ , I have to leave out  $2/4$ , because that's the same as  $1/2$  which I've already counted,



so I have  $3/4$ , Then I go on to  $1/5, 2/5, 3/5, 4/5$  and so forth. In doing this I have actually counted what the integers are. Namely, whenever you can take a set and put it into correspondence with the integers, you're showing that it's countable. I've even labeled what the counting is. A sub 1 is  $1/2$ , A sub 2 is  $1/3$  and so forth.

Now, I want to stick this set inside of a set of intervals, and that's pretty easy. I stick A sub  $i$  inside of the closed interval, which is A sub  $i$  on one side and A sub  $i$  on the other side. So I'm sticking it inside of an interval which has zero measure. Then what I'm going to do is I'm going to add up all of those things. Whenever you add up an infinite number of things, you really have to add up a finite number of them and then look at what happens when you go to the limit. So I add up all of these zeroes, and when I add up  $n$  of them I get zero. I continue to add zeroes, I continue to have zero, and the limit is zero. Now here's where the mathematicians have pulled a very clever swindle on you, because what they're saying is that infinity times zero is zero. But they're saying it in a very precise way. But anyway, since we said it in a very precise way, infinity times zero here is zero, the way we've said it. But somehow that's not very satisfying. I mean it really looks like we've cheated.

So let's go on and do this in a different way. What we're going to do now is for each of these rational numbers, we're going to put a little hat around them, a little rectangular hat around them. Namely a little interval which includes A sub  $i$ , and which also goes  $\delta$  over 2 that way,  $\delta$  over 2 this way, and also multiplies that by -- my computer knew I was going to talk about computers replacing people, so it made some mistakes to get even with me. So this is  $\delta$  times 2 to the minus  $i$  minus 1, and  $\delta$  times 2 to the minus  $i$  minus 1. In other words, you take  $1/2$  and you put a pretty big interval around it. You take  $1/3$ , you put a smaller interval around it.  $2/3$  you put a smaller interval around that and so forth. Well, these intervals here are going to overlap.

But anyway, the union of two overlapping open intervals is an open interval which has a smaller measure than with two of them. In other words, if you take this, the measure of this and add it to the measure of this, the union of this and this is just this. We don't have to be very sophisticated to see that this length is less than this

length plus length, because I've double counted things here. So anyway, when I do this I add up the measure of all of these things and I get delta, and I then let delta get as small as I want to. Again, I find that the measure of the rationals is zero.

Now if you don't like this you should be very happy with yourselves, because I've struggled with this for years and it's not intuitive, because with these intervals I'm putting in here, no matter how small I make that interval, there's an infinite number of rational numbers which are in that interval. In other words, the thing we're trying to do is to separate the interval  $0,1$  into some union of intervals which don't amount to anything but which include all of the rational numbers. Somehow this argument is a little bit bogus because no matter what number I look at between zero and 1, there are rational numbers arbitrarily close to it. In other words, what's going on here is strictly a matter of which order we take limits in, and that's what makes the argument subtle.

But anyway, that is a perfectly sound mathematical argument. You can't get around it. It's why people objected to what Lebesgue was doing for a long time, because it wasn't intuitive to them either. It was intuitive to Lebesgue, and finally it's become intuitive to everyone, but not really intuitive. It's just that mathematicians have heard it so many times that they believe it. I mean one of the problems with any society is that if you tell people things often enough they start to believe them. Unfortunately, that's true in mathematics, too. Fortunately in mathematics we have proofs of things, so that when somebody is telling you something again and again which is false, there are always people who will look at it and say no, that's not true. Whereas in other cases, not necessarily.

There are also uncountable sets with measure zero. For those of you who are already sort of overwhelmed by this, why don't you go to sleep for three minutes, it'll only take three minutes to talk about this, but it's kind of cute. In the set I want to talk about something closely related to what people call the Cantor set, but it's a little bit simpler than that. So what I'd like to do, and this is already familiar to you. I can take numbers between zero and 1 and I can represent them in a binary expansion. I can also represent them in a ternary expansion. I can also represent

them in a decimal expansion, which is what you've been doing since you were three years old or five years older or whenever you started doing this. Well, ternary is simpler than decimal, so you could have done this a year before you started to deal with decimal expansions.

So I want to look at all of the ternary expansions. Each real number corresponds to a ternary expansion, which is an infinite sequence of numbers each of which are zero, 1 or 2. Now, what I'm going to do is I'm going to remove all of the sequences which contain any 1's in them at all. Now it's not immediately clear what that's going to do, but think of it this way. If I first look at the sequences, I'm going to remove all the sequences which start with 1. So the sequences which start with 1 and have anything else after them, that's really the interval that starts at  $1/3$  and ends at  $2/3$ , because that's really what starting with 1 means. This is what we talked about when we talked about approximating binary numbers also, if you remember, is the way we proved the Kraft inequality. It was the same idea.

The sequences which have a 1 in the second position, when we remove them we're removing the interval from  $1/9$  to  $2/9$ . We're also removing the interval from  $7/9$  to  $8/9$ . We're also removing the  $4/9$  to  $5/9$  interval, but we removed that before. So we wind up with something which is now -- we've taken out this, we've now taken out this, and we've taken out this. So the only thing left is this and this and this and this, right? And we've removed everything else. We keep on doing this. Well, each time we remove one of these, all of the numbers that's -- when I do this  $n$  times, the first time I go through this process, I removed  $1/3$  of the numbers, I'm left with  $2/3$  of the interval. When I remove everything that has a 1 in the second position, I'm down with  $2/3$  squared. When I remove everything which has a 1 in the third position, I'm down to  $2/3$  cubed. When I keep on doing this forever, what happens to  $2/3$  to the  $n$ ? Well,  $2/3$  to the  $n$  goes to zero. In other words, I have removed an interval, I have removed a set, the measure 1, and therefore I'm left with a set of measure zero. You can see this happening. I mean you only have diddly left here and I keep cutting away at it.

So less and less gets left. But what we now have is all sequences, all infinite

sequences of zeroes and 2's. So I'm left with all binary sequences except instead of binary sequences with zeros and 1's, I now have binary sequences with zeros and 2's. How many binary sequences are there when I continue forever? Well, you know they're an uncountable number, because if I take all the numbers between zero and 1, I represent them in binary zeroes and 1's, I have an uncountable number of them. Well, because I have to have an uncountable number because we already showed that any countable set doesn't amount to anything. Countable sets are diddly. Countable sets all just go away. So, anything which gets left has to be uncountable. Again, people had to worry about this for a long time. But anyway, this gives you an uncountable set which has measure zero.

So, back to measurable functions. I'm going to get off of mathematics relatively soon, but we need at least this much to figure out what's going on here. We say that a function is measurable. Before we were only talking about sets of numbers being measurable. We had to talk about sets of numbers being measurable because we were interested in the question of what's the set of times for which a function lies between  $2\epsilon$  and  $3\epsilon$ , for example. What we said is we can say a great deal about that because we can not only add up a bunch of intervals, we can also add up a countable bunch of intervals, and we can also get rid of anything which is negligible. So, a function is measurable if the set of  $t$ , such that  $u$  of  $t$  lies between these two points is measurable for each interval. In other words, if no matter how I split up this interval, if no matter what slice I look at, the set of times over which the function lies in there is measurable. That's what a measurable function is.

Everybody understand what I just said? Let me try to say it once more. A function is measurable if for every two values, say  $3\epsilon$  and  $2\epsilon$ , if the set of values  $t$  for which the function lies between  $2\epsilon$  and  $3\epsilon$ , if that set is measurable. In other words, that's the set we were talking about before which went from here to here, and which went from here to there. In this case for this very simple function, that's just the sum of two intervals. If I make the function wiggle a great deal more, it's the sum of a lot more intervals. So, we say the function is measurable if all of the sets are measurable. Now, what I'm going to do is when I'm trying to define this integral, I'm going to have to go to smaller and smaller intervals.

Let's start out with epsilon, 2 epsilon, 3 epsilon and so forth. Let's look at a non-negative function--. Yeah?

**AUDIENCE:** Maybe I missed something, but could you tell me [UNINTELLIGIBLE] definition for a set, if measurable.

**PROFESSOR:** What's the definition for a set is measurable. I didn't really say, and that's good. I gave you a bunch of conditions under which a set is measurable, and if I have enough conditions for which it's measurable then I don't have to worry about--. I said that it is measurable under all of these conditions. I'm saying I don't have to worry about the rest of them because these are enough conditions to talk about everything I want to talk about.

**AUDIENCE:** [UNINTELLIGIBLE].

**PROFESSOR:** I will define my measure as all of these things. Unfortunately, you need a little bit more, and if you want to get more you better take a course in real variables and measure theory. Good.

So, if I want to now make this epsilon smaller, what I'm going to do is do it in a particular way. I'm going to start out partitioning this into intervals of size epsilon. Then I'm going to partition it into intervals of size epsilon over 2. When I partition it into intervals of size epsilon over 2, I'm adding a bunch of extra things. This thing gets added because when I'm looking at the interval between epsilon and 3 epsilon over 2, the function is in this interval here, over this [UNINTELLIGIBLE]. It's in this interval over this whole thing. I'm representing it by this value down here. Now when I have this tinier interval, I see that this function is really in this interval from here to there also, and therefore, instead of representing the function over this interval by epsilon, I'm representing it by 3 epsilon over 2. In other words, as I add these extra quantization levels, I can never lose anything, I only gain things. So I gain all of these cross-hatched regions when I do this, which says that when I add up all these things in the integral, every time I decrease epsilon by 2, the integral that I've got, the approximation to the integral increases.

Now what happens when you take a sum of a set of numbers which are increasing? Well, they're increasing, the result that you get when you add them all up is increasing also, and therefore, as I go from  $\epsilon$  to  $\epsilon/2$  to  $\epsilon/4$  and so forth, I keep climbing up. Conclusion, I either get to a finite number or I get to infinity -- only two possibilities. Which says that if I'm looking at non-negative functions, if I'm only looking at real functions which have non-negative values, the Lebesgue integral for a measurable function always exists, if I include infinite limits as well as finite limits.

Now, if you think back to what you learned about integration, and I hope you at least learned enough about it that you remember there are a lot of very nasty conditions about when integrals exist and when they don't exist. Here that's all gone away. This is a beautifully simple statement. You take the integral of a non-negative function, if it's measurable, and there are only two possibilities. The integral of some finite number or the integral is infinite. It's never undefined, it's always defined. I think that's neat. A few people don't think it's neat, too bad. I guess when I was first studying this, I didn't think it was neat either because it was too complicated. So you have an excuse. If you think about it for a while and you understand it and you don't think it's neat, then I think you have a real problem.

So now -- I did this. I'm getting too many slides out here. Here we go. Here's something new. Hardly looks new. Let's look at a function now, just defined on the interval zero to 1. Suppose that  $h$  of  $t$  is equal to 1 for each rational number and at zero for each irrational number. In other words, this is a function which looks absolutely wild. It just goes up to here and it's 1 or zero. It's 1 at this dense set of points, which we've already said doesn't amount to anything, and it's zero everywhere else. Now you put that into the Riemann integral, and the Riemann integral goes crazy, because no matter how small I make this interval, there are an infinite number of rational numbers in that interval, and therefore, the Riemann integral can never even get started.

For the Lebesgue integral, on the other hand, look at what happens now. We have a bunch of points which are sitting at 1, we have a bunch of points which are sitting

at zero. The only thing we have to do is evaluate the measure of the set of points which are up in some tiny interval up in here. What's this measure of the set of  $t$ 's corresponding to the rational numbers? Well, you already said that was zero. Now, that's why Lebesgue integration works. Any countable set, and in fact, any of these uncountable sets that measure zero get lost in here because you're combining them all together and you say they don't contribute to the integral at all. That's why Lebesgue integration is so simple. You can forget about all that stuff.

When we looked last time at the Fourier series for a square wave, you remember we found that everything behaved very nicely, except where the square wave had a discontinuity, the Fourier series converged to the mid-point, and that was kind of awkward. Well, the mid-points where the function is discontinuous don't amount to anything, because in that case there were just two of them, there were only two points. If they had measure zero it just washes away. You all felt intuitively when you saw that example, that those points were not important. You felt that this was mathematical carping. Well, Lebesgue felt it was mathematical carping too, but he went one step further and he said here's a way of getting rid of all of that and not having to worry about it anymore. So you've now gotten to the point where you don't have to worry about any of this stuff anymore.

Now let's go a little bit further. We're almost at the end of this. If I take a function which maps the real numbers into the real numbers, in other words, it's a function which you can draw on the line. You take time going from minus infinity to plus infinity, you define what this function is at each time. That's what I'm talking about here, a function which you can draw. The functions magnitude of  $u$  of  $t$ , and the function magnitude of  $u$  of  $t$  squared are both non-negative. Now I'm not going to prove this, but it turns out that the magnitude and the magnitude squared are both measurable functions if  $u$  of  $t$  is a measurable function.

In fact, from now on we're just going to assume that everything we deal with is measurable, every function is measurable. I challenge any of you without looking it up in a book to find an example of a non-measurable function. I challenge any of you to find an example of a non-measurable set. I challenge any of you to

understand the definition of a non-measurable set if you look it up in a book. You've heard about things like the axiom of choice and things like that, which are very fishy kinds of things -- that's all involved in finding non-measurable function. So, any function that you think about is going to be measurable. I hate people who say things like that, but it's the only way to get around this because I don't want to give you any examples of that because they're awful.

Since magnitude of  $u$  of  $t$  and magnitude of  $u$  of  $t$  squared are measurable and they're non-negative, their integrals exist. Their integrals exist and are either a finite number or they're infinite. So, we define  $L^1$  functions, and we'll be dealing with  $L^1$  functions and  $L^2$  functions all the way through the course,  $u$  of  $t$  is an  $L^1$  function if it's measurable, and if this integral is less than infinity. That's all there is to it.  $u^2$  is  $L^2$  if it's measurable and the integral of  $u$  of  $t$  squared is less than infinity. I could have said that at the beginning, but now you see that it makes a lot more sense than it did before because we know that if  $u$  of  $t$  is measurable, this integral exists -- it's either a finite number or infinity. The  $L^1$  functions are those particular functions where it's finite and not infinite. Same thing here. The  $L^2$  functions are those where this is finite. This is really the energy of the function, but now we can measure the energy of even weird functions which are zero on the irrationals and one on the rationals. Even things which are zero on the non-cantor set points and 1 on the cantor set points, still it all works. So those define the set  $L^1$  and  $L^2$ .

Now, a complex function  $u$  of  $t$ , which maps  $r$  into  $c$ , why does a complex function map  $r$  into  $c$ ? What's the  $r$  doing there? Think of any old complex function you can think of.  $e$  to the  $i, 2\pi t$ . That's something that wiggles around, the sinusoid.  $t$  is a real number. So that function  $e$  to the  $i 2\pi t$  is mapping real numbers into complex numbers. That's what we mean by something which maps the real numbers into complex numbers. We always call these complex functions. I mean mathematicians would say yeah, a function could be anything. But you know, when most of us think of a function, we're thinking of mapping a real variable into something else, and when we're thinking of mapping a real variable into something else, we're usually thinking of mapping it into real numbers or mapping it into complex numbers, and



because we want to deal with these complex sinusoids, we have to include complex numbers also.

So a complex function is measurable by definition if the real part and the imaginary part are each measurable. We already know when a real function is measurable. Namely, a real function is measurable if each of these slices are measurable. So now we know when a complex function is measurable. We already said that all of the complex functions you can think of and all the ones we'll ever deal with are all measurable. So  $L^1$  and  $L^2$  are defined in the same way when we're dealing with complex functions. Namely, just whether this integral is less than infinity and this integral is less than infinity. Since these functions -- this is a real function from real into real, this is a real function from real into real. So those are well-defined.

What's the relationship between  $L^1$  functions and  $L^2$  functions? Can a function be  $L^1$  and not  $L^2$ ? Can it be  $L^2$  and not  $L^1$ ? Yeah, it can be both, unfortunately. All possibilities exist. You can have functions which are neither  $L^1$  nor  $L^2$ , functions that are  $L^1$  but not  $L^2$ , functions that are  $L^2$  but not  $L^1$ , and functions that are both  $L^1$  and  $L^2$ . Those are the truly nice functions that we like to deal with. But there's one nice that you can say, and that follows from a simple argument here. If  $u$  of  $t$  is less than or equal to 1, if the magnitude of  $u$  of  $t$  is less than or equal to 1 --. Let's start out by looking at  $u$  of  $t$  being greater than or equal to 1. If  $u$  of  $t$  is greater than or equal to 1, then  $u$  squared of  $t$  is even bigger. You see the thing that happened is when  $u$  of  $t$  becomes bigger than  $t$ ,  $u$  squared of  $t$  becomes even more bigger. When  $u$  of  $t$  is less than 1,  $u$  squared of  $t$  is less than  $u$  of  $t$ . But if  $u$  of  $t$  is less than or equal to 1, it's less than 1.

So in all cases, for all  $t$ ,  $u$  of  $t$ , magnitude is less than or equal to  $u$  of  $t$  squared plus 1. So that takes into account both cases. It's a bound. So if I'm looking at functions which only exist between over some limited time interval, and I take the integral from minus  $t$  over 2 to the  $t$  over 2 of the magnitude of  $u$  of  $t$ , I get something which is less than or equal to the integral of  $u$  squared of  $t$  plus the integral of 1, and the integral of 1 over this finite limit is just  $t$ . This says that if a function is  $L^2$  and the function only exists over a finite interval, then the function is  $L^1$  also. So as long as

I'm dealing with Fourier series, as long as I'm dealing with finite duration functions,  $L_2$  means  $L_1$ . All of the nice things that you get with  $L_1$  functions apply to  $L_2$  functions also, and there are a lot of nice things that happen for  $L_1$  functions. There are a lot of nice things that happen for  $L_2$  functions. You take the union of what happens for  $L_1$  and for  $L_2$ , and that's beautiful. You can say anything then. Can't calculate anything, of course, but we all said, we leave that to computers.

Let's go back Fourier series now, let's go back to the real world. Any old function we have  $u$  of  $t$ , the magnitude of  $u$  of  $t$  and the magnitude of  $u$  of  $t$  times either the  $2\pi$   $\text{fft}$ , for any old  $f$ , this thing has magnitude 1, right, a complex exponential. Real  $f$ , real  $t$ . This just has magnitude 1. And therefore, this magnitude is equal to this magnitude. This says that if the function  $u$  of  $t$  is  $L_1$ , then the function  $u$  of  $t$  times either the  $2\pi$   $\text{IFT}$  is also  $L_1$ , which says if we can integrate one we can integrate the other. In other words, the integral of  $u$  of  $t$ , either the  $2\pi$   $\text{ift}$ , the magnitude of the  $fdt$  is going to be less than infinity. Since we're taking the magnitude, it's either finite or it's infinite, and since  $u$  of  $t$  in magnitude when we integrate it is less than infinity, this thing is less than infinity also.

Now, this is a complex number in here. So you can break it up into a real part and an imaginary part, and if this whole thing, if the magnitude is less than infinity, then the magnitude of the real part is finite. If you take the real part over the region where this is positive and the region where it's negative, you still get non-negative numbers. In other words, if we're taking the integral of something which has positive values -- I should have written this out in more detail, it's not enough to--. I'm taking the integral of something which--. This is  $u$  of  $t$ . If I can find another color. Let me draw magnitude of  $u$  of  $t$  on top of this. This thing here is magnitude of  $u$  of  $t$ . What?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** I'm making it real for the time being because I'm just looking at the real part. In other words, what I'm looking is this quantity here. Later you can imagine doing the same thing out on complex numbers, OK? So what I'm saying is if I just look at the real part of  $u$  of  $t$  -- call this real part of  $u$  of  $t$ , if you like. If I know that this is finite,

this is non-negative, I know that the positive part of this function has a finite integral, I know that the negative part of it has a finite integral. In other words, the thing which makes integration messy is you sometimes have the positive part being infinite, the negative part being infinite also, and the two of them cancel out somehow to give you something finite.

When you're dealing with the magnitude of  $u$  of  $t$ , if the magnitude of  $u$  of  $t$  has an infinite integral, then that messy thing can't happen. It says that the positive part has a finite integral, the negative part has a finite integral also. If you take the imaginary part, visualize that out this way, the positive part of the imaginary part has a finite integral, the negative part of the imaginary part has a finite integral also. It says if the magnitude of  $u$  of  $t$ , that integral always exists, and if it's finite, then these positive parts of the real, the positive part of the imaginary part, all of those are finite, and it says the integral itself is finite. Which says that this integral here has to be finite. Namely, the positive part of the real part, the negative part of the real part, the positive part of the imaginary part, the negative part of the imaginary part, all four have to be finite, just because this quantity here is finite. Now, if  $u$  of  $2$  is  $L_2$  and also time limited, it's  $L_1$ , and the same conclusion follows. So this integral always exists if it's over a finite interval.

So at this point we're really ready to go back to this theorem about Fourier series that we stated last time and which was a little bit mysterious at that point. In fact, at this point we've already proven part of it. I wanted to prove it because I wanted you to know that not everything in measure theory is difficult. An awful lot of these things, after you know just a very small number of the ideas, are very, very simple. Now, you will think this is not simple because you haven't had time to think about the ten slides that have gone before. But if you go back and you look at them again, if you read the notes, you will see that, in fact, it all is pretty simple. What this says is if  $u$  of  $t$ , a complex function real into  $\mathbb{C}$ , but time limited, suppose it's an  $L_2$  function, then it's also  $L_1$  over that interval minus  $t$  over  $2$  to plus  $t$  over  $2$ . Then for each  $k$  and  $z$ , then for each integer  $k$ , this integral here, this function here, is now an  $L_1$  function, and therefore, this integral exists, is finite. You divide by  $t$  is still finite. So that Fourier coefficient has to exist and it has to exist as a finite value.

Now, you look at Reimann integration, and you look at the theorems about Reimann integration, and if they're stated by somebody who was stating theorems, the conditions are monstrous. This is not monstrous. It says all you need is measurability and  $L^1$ , which says it doesn't go up to infinity. That's enough to say that every one of these Fourier coefficients has to exist. It might be hard to integrate it, but you know it has to exist. Now the next thing it says is that -- this is more complicated. What we would like to say and what we tried to say before and what you like to say with the Fourier series is that  $u$  of  $t$  is equal to this, where you sum from minus infinity to plus infinity. We saw that we can't say that. We saw that we can't it for functions which have step discontinuities, because whenever you have a step discontinuity, the Fourier series converges to the mid-point of that discontinuity. If you were unfortunate enough to try to make life simple and define the function without defining it at the step discontinuity as the mid-point, then the Fourier series would not be equal to  $u$  of  $t$ . But what this says is that if you take the difference between  $u$  of  $t$  and a finite expansion, and then you look at the energy in that difference, it says that the energy and the difference goes to zero. Now that's far more important than having this integral be equal to that, because frankly, we don't care a fig for whether this is equal to this at every  $t$  or not. What we care about is when we add more and more terms onto this Fourier series.

I mean in engineering we're always approximating things. We have to approximate things. We talk about functions  $u$  of  $t$ , but our functions  $u$  of  $t$  are just models of things anyway, and we want those models to really converge to something, which means that as we take more and more terms in the Fourier series, we get something which comes closer to  $u$  of  $t$  and it comes closer in energy terms. Remember when we take this, when we approximate it in this way by a finite Fourier series, and we then quantize coefficients, and then we go back to the function, we have lost all the coefficients and all of these terms for the very high frequencies. We've dropped them off. What this says is if we take more and more of them and then we quantize and we go back to a function  $v$  of  $t$  it says as we add more and more Fourier coefficients and quantize carefully, we can come closer and closer to

the function that we started with. You don't get that by talking about point to point equality, because as soon as we quantize you lose all of that anyway. You don't have a quality anywhere anymore, and the only thing you can hope for is a small mean square error.

So, this looks more complicated than what you would like, but what I'm trying to tell you is that this is far more important than what you would like. What you would like is not important at all. I can give you examples of things where this converges to this everywhere, but in fact, no matter how many terms you take, the energy difference between these two things is very large. Those are the ugly things. That says you never get a good approximation, even though looking at things point-wise everything looks great. But what you're really interested in is these mean square approximations and what this Fourier series thing says is that if you deal with measurable functions then this converges just to that in this very nice energy sense, and energy is what we're interested in.

The final part of this -- I mean I talked about this a little bit last time -- sometimes instead of starting out with a function and approximating it with the Fourier series, you want to start out with the coefficients and find the function. In fact, when we start talking about modulation, that's exactly what we're going to be doing because we're always going to be starting out with these digital sequences, we're going to be finding functions from the digital sequences. This final part of the theorem says yes, you can do that and it works. It says that given any set of coefficients where the coefficients have finite energy, in other words, where the sum is less than infinity, then there's an L2 function,  $u$  of  $t$ , which satisfies the above in this limiting sense. Now this is the hardest thing of the theorem to prove. It looks obvious but it's not. But anyway, it's there. It says these approximations are rock solid and you can now forget about all of this measure theoretic stuff, and you can just live with the fact that all of these in terms of L2 approximations, work. I mean again, it doesn't look beautiful at this point because you haven't seen it for long enough. It really is beautiful. It's beautiful stuff. When you think about it long enough, for 40 years is how long I've been thinking about it, it becomes more and more beautiful every year. So, if you live long enough it'll be beautiful.

Any time you talk about this type of convergence, we call it convergence in the mean, and then notation we will use is l.i.m., which looks like limit but which really stands for limit in the mean. We will write that complicated thing on the last slide, which I said is not really that complicated, but we will write this as this. In other words, when we write limit in the mean, we mean that the difference between these two sides in energy sense goes to zero as  $k$  gets large. That's what this limit means. It just means the statement that we talked about before. So the Fourier series, the theorem really says you can find all of the Fourier coefficients in this way, they all exist, they all exist exactly, they're all finite. The function then exists as this limit in the mean, which is the thing that we're always interested in. So, every  $L^2$  function defined over a finite interval, and every  $L^1$  function defined over a finite interval, all of these have a Fourier series which satisfies these two relationships.

Now, what we're going to do with all of this is we're going to segment an arbitrary function, an arbitrary  $L^2$  function over the entire time range, and we're going to segment it into pieces, all of width  $t$ . Then we're going to expand each of those segments into a Fourier series. That's what people do any time they compress voice that I said several times. When you compress voice -- for some reason or other everybody does it in 20 millisecond increments. You chop up the voice into 20 millisecond increments. You then, at least, conceptually look at those 20 millisecond increments, think of them in terms of the Fourier series, you expand them in the Fourier series. So for each increment in time we get a different Fourier series, and we use those Fourier series to approximate the function.

So, each one of the increments now is going to be represented in this way here. Since we're only looking at the function now over the interval around time,  $mt$ , we look at it this way, we calculate the coefficients, again, in terms of this rectangular function spaced off by  $m$ . This is saying the same thing that we said before, just moving it from  $\frac{-t}{2}$  to  $\frac{t}{2}$  to  $\frac{t}{2} - mt$ . Here's  $\frac{-t}{2}$ . The  $\frac{t}{2}$ . Next we're looking at  $\frac{t}{2}$  to  $\frac{3t}{2}$ . Next we're looking at  $\frac{5t}{2}$ . This is  $m$  equals zero. This is  $m$  equals 1. This is  $m$  equals 2. This notation here you should get used to it, it will be confusing for a while. All it means is that, and it works.

So, in fact, you can take all these terms, you can break them up this way. So what you've really done is you've broken  $u$  of  $t$  into a double sum expansion. These exponentials limited in time, that entire set of functions are talking to each other. A function living in this interval of time and a function living in this interval of time have to be orthogonal, because I multiply this by this and I get zero at each time. In one interval these exponentials are orthogonal to each other -- we've pointed that out before. So we have a doubly exponential, we have a double sum of orthogonal functions, and what we're saying is that any old  $L_2$  function at all we can break up into this kind of sum here. This is a very complicated way of saying what we think of physically anyway. It says take a big long function, segment it into intervals of length  $t$ , break up each interval of length  $t$  into a Fourier series. That's all that's saying. So it's saying what is obvious.

We're going to find a number of such orthogonal expansions which work for arbitrary  $L_2$  functions. As I said, it's a conceptual basis for voice compress algorithms. Even more, next time we're going to go into the Fourier integral. You think of the Fourier integral as being the right way to go from time to frequency, but, in fact, it's not really the right way to go from time to frequency. When we think of voice or you think of the wave form of a symphony, for example, what going on there? Over every little interval of time you hear various frequencies, right? In fact, if you make  $t$  equal to the timing of the music, the idea becomes very, very clean because at each time somebody changes a note. So you go from one frequency to another frequency, so it's a rather clean way of looking at this. Our notion of frequency that we think of intuitively is much more closely associated with the idea of frequencies is changing in time than it is of frequencies being constant in time. When we look at a Fourier integral, everything is frozen in time. As soon as you take the Fourier integral, you've glopped everything from time minus infinity to time plus infinity all together. Here when we look at these expansions, we're looking at these frequencies changing.

There's an unfortunate part about frequencies changing, and that is frequencies, unfortunately, live over the entire time interval. These truncated frequencies work

very nicely but they don't quite correspond to non-truncated time intervals, but it still does match our intuition and this is a useful way to think about functions that change in time. If you believe what I've just said, why do people ever worry about the Fourier integral at all? Well, you see the problem is this quantity  $t$  here that I'm segmenting over is unnatural. If I look at voice, there's no  $t$  that you can take which is a natural quantity. The 20 milliseconds is just something arbitrary that somebody did once and it worked. Too many other engineers in the field said, well that works, I'll pick the same thing instead of trying to think through what a better number would be. So they all use the same  $t$ . It doesn't correspond to anything physically. So, in fact, people do try the same thing.

Let me just say what a Fourier transform is and we'll talk about it most of next time because that's the next thing we have to deal with. Something you're all familiar with are these Fourier transform relationships. There's the function of time, there's the function of frequency. You can go from one to the other, mapping. Well, if you know this you can find this. If you know this, you can find this. It looks a little bit like the Fourier series, and usually when people learn about the Fourier integral they start with a Fourier series and start letting  $t$  become large, which doesn't quite work. But anyway, that's what we're going to do. If the function is well-behaved, the first integral exists and the second exists. What does well-behaved mean? It means what we usually mean by well-behaved, it means it works. So again, this theorem here is another example, or not, as the case may be. But we'll make this clearer next time.