

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: --And go on with lectures 8 to 10. First I want to briefly review what we said about measurable functions. Again, I encourage you if you hate this material and you think it's only for mathematicians, please let me know. I don't know whether it's appropriate to cover it in this class either. I think it probably is because you need to know a little more mathematics when you deal with these Fourier transforms and Fourier series. When you're dealing with communication theory, then you need to know for signal processing and things like that. When you get into learning a little more, my sense at this point is it's much easier to learn a good deal more than it is to learn just a little bit more, because a little bit more you're always faced with all of these questions of what does this really mean. It turns out that if you put just a little bit of measure theory into your thinking, it makes all these problems very much simpler.

So you remember we talked about what a measurable set was last time. What a measurable set is, is a set which you can essentially break up into a countable union of intervals, and something which is bounded by a countable set of intervals which have arbitrarily small measure. So you can take all of these sets of zero measure -- countable sets, cantor sets, all of these things, and you can just throw all of those out from consideration. That's what makes measure theory useful and interesting and what simplifies things. So we then said that a function is measurable if each of these sets here -- if the set of times for which a is less than or equal to u of t is less than or equal to b -- these things are these level sets here. It's a set of values, t , along this axis in which the function is nailed between any two of these points here. For a function to be measurable, all of these sets in here have to be measurable.

We haven't given you any examples of sets which are not measurable, and

therefore, we haven't given you any examples of functions which are not measurable. It's very, very difficult to construct such functions, and since it's so difficult to construct them, it's easier to just say that all of the functions we can think of are measurable. That includes an awful lot more functions than you ever deal with in engineering. If you ever make a serious mistake in your engineering work because of thinking that a function is measurable when it's not measurable, I will give you \$1,000. I make that promise that you because I don't think it's ever happened in history, I don't think it ever will happen in history, and that's just how sure I am of it. Unless you try to deliberately make such a mistake in order to collect \$1,000. Of course, that's not fair.

Then we said the way we define this approximation to the integral, namely, we always defined it from underneath, and therefore, if we started to take these intervals here with epsilon and split them in half, the thing that happens is you always get extra components put in. So that as you make the scaling finer and finer, what happens is the approximation into the integral gets larger and larger. Because of that, if you're dealing with a non-negative function, the only two things that can happen, as you go to the limit of finer and finer scaling, is one, you come to a finite limit, and two, you come to an infinite limit. Nothing else can happen. You see that's remarkably simple as mathematics go. This is dealing with all of these functions you can ever think of, and those are the only two things that can happen here.

So then we went on to say that a function as L^1 , if it's measurable, and if the integral of its magnitude is less than infinity, and so far we're still talking about real functions. If u of t is L^1 , then you can take the integral of u of t and you can split it up into two things. You can split it up into the set of times over which u of t is positive, and the set of times over which u of t is negative. The set of times over which u of t is equal to zero doesn't contribute to the integral at all so you can forget about that. This is well-defined as the function is measurable. What's now happening is that if the integral of this magnitude is less than infinity, then both this has to be less than infinity, and this has to be less than infinity, which says that so long as you're dealing with L^1 functions, you never have to worry about this nasty

situation where the negative part of the function is infinite, the positive part of the function is infinite, and therefore, the difference, plus infinity, minus infinity doesn't make any sense at all. And instead of having an integral which is infinite or finite, you have a function which might just be undefined completely. Well this says that if the function is L^1 , you can't ever have any problem with integrals of this sort thing being undefined. So, and in fact, u of t is always integral and has a finite value in this case.

Now, we say that a complex function is measurable if both the real part is measurable and the imaginary part is measurable. Therefore, you don't have to worry about complex functions as really being any more than twice as complicated as real functions, and conceptually they aren't any more complicated at all because you just treat them as two separate functions, and everything you know about real functions you now know about complex functions. Now, as far as Fourier theory is concerned -- Fourier series, Fourier integrals, discrete time, Fourier transforms -- all of these things, if u of t is an L^1 function, then u of t times e to the $2\pi i f t$ has to be L^1 , and this has to be true for all possible values of f . Why is that? Well, it's just that the absolute value of u of t for any given t , is exactly the same as the absolute value of u of t times e to the $2\pi i f t$. Namely, this is a quantity whose magnitude is always 1, and therefore, this quantity, this magnitude is equal to this magnitude. Therefore, when you integrate this magnitude, you get the same thing as when you integrate this magnitude. So, there's no problem here.

So what that says, using this idea, too, of positive and negative parts, it says that the integral of u of t times e to the $2\pi i f t$ dt has to exist for all real f . That covers Fourier series and Fourier integral. We haven't talked about the Fourier integral yet, but it says that just by defining u of t to be L^1 , you avoid all of these problems of when these Fourier integrals exist and when they don't exist. They always exist.

So let's talk about the Fourier transform then since we're backing into this slowly anyway. What you've learned in probably two or three classes by now is that the Fourier transform is defined in this way. Namely, the transform, a function of frequency, \hat{u} of f , we use hats here instead of capital letters because we like to

use capital letters for random variables. So that this transform here is the integral from minus infinity to infinity of the original function, we start with $u(t)$ times $e^{-i2\pi ft}$ dt. Now, what's the nice part about this? The nice part about this is that if $u(t)$ is L^1 , then in fact this Fourier transform has to exist everywhere.

Namely, what you've learned before in all of these classes says essentially that if these functions are well-behaved then these transforms exist. What you've learned about well-behaved is completely circular, namely, a function is well-behaved if the transform exists, and the transform exists if the function is well-behaved. You have no idea of what kinds of functions the Fourier transform exists and what kinds of functions it doesn't exist.

There's another thing buried in here, in this transform relationship. Namely, the first thing is we're trying to define the transform function of frequency in terms of the function of time. Well and good. Then we try to define a function of time in terms of a function of frequency. What's hidden here is if you start with $u(t)$, you then get a function of frequency, and this sort of implicitly says that you get back again the same thing you started with. That, in fact, is the most ticklish part of all of this. So that if we start with an L^1 function $u(t)$, we wind up with a Fourier transform, $\hat{u}(f)$. Then $\hat{u}(f)$ goes in here. We hope we might get back the sum transform here.

The nasty thing, and one of the reasons we're going to get away from L^1 in just a minute, is that if a function is L^1 , its Fourier transform is not necessarily L^1 . What that means is that you have to learn all of this stuff about L^1 functions, and then as soon as you take the Fourier transform, bingo, it's all gone up in a lot of smoke, and you have to start all over again saying something about what properties the transform might have. But anyway, it's nice to start with because when $u(t)$ is L^1 , we know that this function actually exists. It actually exists as a complex number. It exists as a complex number for every possible real f here. Namely, there aren't any if, and's, but's or maybe's here, there's nothing like L^2 convergence that we were talking about a little bit before. This just exists, period.

There's something more here, that if this is L^1 , this not only exists but it's also a

continuous function, and those don't prove that. If you've taken a course in analysis and you know what a complex function is and you're quite patient, you can sit down and actually show this yourselves, but it's a bit of a pain. So for a well-behaved function, the first integral exists for all f , the second exists for all t , and results in the original u of t . But then more specifically, if u of t is L^1 , the first integral exists for all f and it's a continuous function. It's also a bounded function, as we'll see in a little bit. If \hat{u} of f is L^1 , the second integral exists for all t , and u of t is continuous.

Therefore, if you assume at the output at the onset of things that both this is L^1 and this is L^1 , then everything is also continuous and you have a very nice theory, which doesn't apply to too many of the things that you're interested in. And I'll explain why in a little bit.

Anyway, for these well-behaved functions, we have all of these relationships that I'm sure you've learned in whatever linear systems course you've taken. Since you should know all of these things, I just want to briefly talk about them I mean this linearity idea is something you would just use without thinking about it, I think. In other words, if you'd never learned this and you were trying to work out a problem you would just use it anyway, because anything respectable has to have -- well, anything respectable, again, means anything which has this property. This conjugate property you can derive that easily from the Fourier transform relationships also, if you have a well-behaved function. This quantity here, this duality, is particularly interesting because it isn't really duality, it's something called hermitian duality. You start out with this formula here to go to there and you use almost the same formula to get back again. The only difference is instead of a minus 2π ift , you have a plus 2π ift . In other words, this is the conjugate of this, which is why this is called hermitian duality. But aside from that, everything you learn about the Fourier transform, you also automatically know about the inverse Fourier transform for these well-behaved functions. Otherwise you don't know whether the other one exists or not, and we'll certainly get into that.

So, the duality is expressed this way, the Fourier transform of, blah. If you take a function, u of t , and regard that as a Fourier transform, then -- I always have trouble saying this. If you take a function, u of t , and then you regard that as a function of

frequency -- OK, that's this -- and then you regard it as a function of minus frequency, namely, you substitute minus f for t in whatever time function you start with. You start with a time function u of t , you substitute minus f for t , which gives you a function of frequency. The inverse Fourier transform of that is what you get by taking the Fourier transform of u of t , and then substituting t for f in it. It's much harder to say it than to do. This time shift, you've all seen that. If you shift a function in time, the only thing that happens is you get this rotating term in it. Same thing for a frequency shift. You want to have an interesting exercise, take time shift plus duality and derive the scaling and frequency from it -- it has to work, and of course, it does.

Scaling -- there's this relationship here. This one is always a funny one. It's a little strange because when you scale here, it's not too surprising that when you take a function and you squash it down that the Fourier transform gets squashed upwards. Because in a sense what you're doing when you squash a function down you're making everything happen faster than it did before, which means that all the frequencies in it are going to get higher than they are before. But also when you squash it down, the amount of energy in the function is going to go down.

One of the most important properties that we're going to find and what you ought to know already is that when you take the energy in a function, you get the same answer as you get when you take the energy in the Fourier transform. Namely, you integrate u of f squared over frequency and you get the same as if you integrate u of t squared over time. That's an important check that you use on all sorts of things. The thing that happens now then when you're scaling is when you scale a function, u of t , you bring it down and you spread it out when you bring it down. The frequency function goes up so the energy in the time function goes down, the energy in the frequency function goes up, and you need something in order to keep the energy relationship working properly. This is what you need. Actually, if you derive this, the t just falls out naturally. So we get the same thing if we do scaling and frequency. I don't think I put that down but it's the same relationship.

There's differentiation. Differentiation we won't talk about or use it a whole lot. All of

these things turn out to be remarkably robust. When you're dealing with L1 functions or L2 functions and you scale them or you shift them or do any of those things, if they're L1, they're still L1 after you're through, if they're L2, they're L2 after you're through. If you differentiate a function, all those properties change. You can't be sure of anything anymore. There's convolution, which I'm sure you've derived many times, and which one of the exercises derives again.

There's correlation, and what this sort of relationship says is taking products at the frequency domain is the same as going through this convolution relationship and the time domain. Of course, there's a dual relation to that, which you don't use very often but it still exists. Correlation is you actually get correlation by using one of the conjugate properties on the convolution and I'm sure you've all seen that. That's something you should all have been using and familiar with for a long time.

Two special cases of the Fourier transform is that u of zero is what happens when you take the Fourier transform and you evaluate it at t equals zero. You get u of zero is just the integral of u hat of f . u hat of zero is just the integral of u of t . What do you use these for? Well the thing I use them for is this half the time when you're working out a problem it's obvious by inspection what this integral is or it's obvious by inspection what this integral is, and by doing that you can check whether you have all the constants right in the transform that you've taken. I don't know anybody who can take Fourier transforms without getting at least one constant wrong at least half the time they do it. I probably get one constant wrong about three-quarters of the time that I do it, and I'm sure I'll do it here in class a number of times. I hope you find it, but it's one of these things we all do. This is one of the best ways of finding out what you've done and going back and checking it.

Parseval's Theorem, Parseval's Theorem is really just this convolution equation which we're applying at τ equals zero. You take the convolution, you apply it at τ equals zero and what do you get? You get the integral of u of t times some other conjugate -- the conjugate of the other function is equal to the integral of u hat of f times the complex conjugate of v of f . Much more important than this is what happens if v happens to be the same as u , and that gives you the energy equation

here, which is what I was talking about. It says the integral in a function you can find it two ways, either by looking at it in time or by looking at it in frequency. I urge you to always think about doing this whenever you're working problems, because often the Fourier integral it's very easy to find the integral and the and the time function is very difficult. A good example of this is sinc functions and rectangular functions. Anybody can take a rectangular function, square it and integrate it. It takes a good deal of skill if you don't use this relationship to take a sinc function, to square it and to integrate it. You can do it if you're skillful at integration, you might regard it as a challenge, but after you get done you realize that you've really wasted a lot of time because this is the right way of doing it here.

Now, as I mentioned before, it's starting to look like Fourier series and Fourier integrals are much nicer when you have L1 functions, and they are, but L1 functions are not terribly useful as far as most communication functions go. In other words, not enough functions are L1 to provide suitable models for the communication systems that we want to look at. In fact, most of the models that we're going to look at, the functions that we're dealing with are L2 functions. One of the reasons for this is this sinc function, sine x over x function is not L1. A sinc function just goes down as 1 over t, and since it goes down as 1 over t, you take the absolute value of it and you integrate 1 over t and what do you get when you integrate it from minus infinity to infinity. A function that's 1 over t, well, if you really want to go through the trouble of integrating it, you can integrate it over limits and you get limits where you have to evaluate the limits on a logarithmic function. When you get all done with that you get an infinite value. And you can see this. You could take 1 over t and you just look at it, as you go further and further out it just gets bigger and bigger without limit. So, sinc t is not L1. Sinc function is a function we'd like to use.

Any function with a discontinuity can't be the Fourier transform of any L1 function. In other words, we said that if you take the Fourier transform of an L1 function, one of the nice things about it is you get a continuous function. One of the nasty things about it is you get a continuous function. Since you get a continuous function it says that any time you want to deal with transforms which are not continuous, you can't be talking about time functions which are L1. One of the frequency functions we

want to look at a great deal is a frequency function corresponding to a band-limited function. When you take a band-limited function you just chop it off at some frequency, and usually when you chop it off, you chop it off and get a discontinuity. When you chop it off and get a discontinuity, bingo, the time function you're dealing with cannot be L1. It has to dribble away much too slowly as time goes to infinity. That's an extraordinarily important thing to remember. Any time you get a function which is discontinuous in the frequency domain, the function cannot go to zero any faster in a time domain than $1/t$ and vice versa in the frequency domain.

L1 functions sometimes have infinite energy. In other words, $\text{sinc } t$ is not L1 -- well, that's not a good example because that's not L1, and it also has infinite energy, but you can just as easily find functions which drop off a little more slowly than $\text{sinc } t$, and which have infinite energy because they--. Excuse me. Sometimes you have functions which go off to infinity too fast as you approach time equals zero, things which are a little bit like impulses but not really. Impulses are awful and we'll talk about them in a minute, because they don't have finite energy as we said before. We have functions which just slowly go off to infinity and they are L2 but they aren't L1. Why do we care about those weird functions? We care about them, as I said before, because we would like to be able to make statements which are simple which we can believe in. In other words, you don't want to go through a course like this with a whole bunch of things that you have to leave. It's very nice to have some things that you really believe, and whether you believe them or not, it's nice to have theorems about them. Even if you don't believe the theorems, at least you have theorems so you can fool other people about it, if nothing else.

Well, it turns out that L2 functions are really the right class to look at here. Oh, I think I left out one of the most important things here. Maybe I put it down here. No, probably not. One of the reasons we want to deal with L2 functions is if you're dealing with compression, for example, and you take a function, if the function has infinite energy in it, then one of the things that happens is that any time you expand it into any kind of orthonormal expansion or orthogonal expansion, which we'll talk

about later, you have coefficients, which have infinite energy. In other words, they have infinite values, or the sum squared of the coefficients is equal to infinity. When we try to compress those we're going to find that no matter how we do it our mean square error is going to be infinite. Yes, we will talk about that later when we get to talking about expansions.

So for all those reasons L1 isn't the right thing, L2 is the right thing. A function going from the reals into the complexes, in other words, a complex valued function is L2 if it's measurable, and if this integral is less than infinity, in other words, if that has finite energy. Primarily it means it has finite energy because all the functions you can think of are measurable. So it really says you're dealing with functions that have finite energy here.

So, let's go on to Fourier transforms then. Interesting simple theorem. I think I stated this last time, also. If a function is L2 and its time limited, it's also L1. So we've already found that if functions are L1, they have Fourier transforms that exist. The reason for this is if you square u of t , take the magnitude squared of u of t for any given t , it has to be less than or equal to the sum of u of t plus 1. In fact, it has to be less than that. Why? How would you prove this if you had to prove it? Well, you say gee, this is two separate terms here. Why don't I look at two separate cases. The two separate cases are first, suppose u of t itself as a magnitude less than 1. If u of t has a magnitude less than 1 and you square it, you get something even smaller. So, any time u of t has a magnitude less than 1, u squared of t is less than or equal to u of t . Blah blah blah blah blah blah blah. If u of t -- what did I do here? No wonder I couldn't explain this to you. Let's try it that way and see if it works. If you can prove something, turn it around and see if you can prove it then.

Now, two cases. First one, let's suppose that u of t is less than or equal to 1. Well then, u of t is less than or equal to 1. And this is positive so this is less than or equal to that. Let's look at the other case. u of t is greater than or equal to 1, magnitude. Well then, it's less than or equal to u of t magnitude squared. So either way this is less than or equal to that. What that says is the integral over any finite limits of u of t dt is less than or equal to the integral of this. Well, the integral of this splits up into

the integral magnitude u of t squared dt plus the integral of 1. Now, the integral of 1 over any finite limits is just b minus a . That's where the finite limits come in. Finite limits say you don't have to worry about this term because it's finite. So that says at any time you have an L^2 function over a finite range, that function is also L^1 over that finite range. Which says that any time you take a Fourier transform of an L^2 function, which is only non-zero over a finite range, bingo, it's L^1 and you get all these nice properties. It has to exist, it has to be continuous, it has to be bounded, and all of that neat stuff.

So for any L^2 function u of t , what I'm going to try to do now, and I'm just copying what a guy by the name of Plancherel did a long time ago. The thing that Plancherel did was he said how do I know when a Fourier transform exists or not. I would like to make it exist for L^2 functions, how do I do it? Well, he said OK, the thing I'm going to do is to take the function u of t and I'm first going to truncate it. In fact, if you think in terms of Riemann integration and things like that, any time you take an integral from minus infinity to plus infinity, what do you mean by it? You mean the limit as you integrate the function over finite limits and then you let the limits go to infinity. So all we're doing is the same trick here. So we're going to take u of t , we're going to truncate it to some minus a to plus a over some finite range, no matter how big a happens to be. We're going to call the function b sub a of t , u of t truncated to these limits. In other words, u of t times a rectangular function evaluated at t over $2a$. Now, can all of you look at this function and see that it just means truncate from minus a to plus a ? No. Well, you should learn to do this. One of the ways to do it is to say OK, the rectangular function is defined as having the value 1 between minus $1/2$ and plus $1/2$ and it's zero everywhere else. I think I said that before in class, didn't I? Certainly it's in the notes. Rectangle $f(t)$ equals 1 for t less than or equal to $1/2$, zero else.

So with this definition you just evaluate what happens when t is equal to minus a , you get rectangle of minus a over $2a$, which is minus $1/2$. When t is equal to a , you're up to the other limit, so this function is 1 for t between minus a and plus a and zero everywhere else. Please get used to using this and become a little facile at sorting out what it means because it's a very handy way to avoid writing awkward

things like this.

So v_a of t by what we've said is both L^2 and its L^1 . We started out with a function which is L^2 and we truncated. Then according to this theorem here, this function v_a of t is now time limited. It's also L^2 , and therefore, by the theorem it's also L^1 . Therefore, it's continuous and you can take the Fourier transform of it -- \hat{v}_a of f exists for all f and it's continuous. So this function is just the Fourier transform that you get when you truncate the function, which is what you would think of as a way to find the Fourier transform anyway. I mean if this is not a reasonable approximation to the Fourier transform a function, you haven't modeled the function very well. Because when a gets extraordinarily large, if there's anything of significance that happens before year ten to the minus 6 or which happens after year ten to the plus 6, and you're dealing with electronic speeds, your models don't make any sense. So for anything of any interest, these functions here are going to start approximating u of t , and therefore we hope this will start approximating the Fourier transform of u of t . Who can make that more precise for me?

What happens when u of t has finite energy? If it has finite energy it means that the integral of u of t magnitude squared over the infinite interval is finite. So you start integrating it over bigger and bigger minus a to plus a . What happens is as the minus a to plus a gets bigger and bigger and bigger, you're including more and more of the function, so that the integral of u of t squared over that bigger and bigger interval has to be getting closer and closer to the overall integral of u of t . Which says that the energy in u of t minus the a of t has to get very, very small as a gets large. That's one of the reasons why we like to deal with finite energy functions. By definition they cannot have a appreciable energy outside of very large limits. How large those large limits have to be depends on the function, but if you make them large enough you will always get negligible energy outside of those limits. So then we can take the Fourier transform of this function within those limits and we get something which we hope is going to be a reasonable approximation of the Fourier transform of u of t .

That's what Plancherel said. Plancherel said if we have an L^2 function, u of t , then

there is an L^2 function \hat{u} of f , which is really the Fourier transform of u of t . Some people call this the Plancherel transform of u of t and say that indeed Plancherel was the one that invented Fourier transforms or Plancherel transforms for L^2 functions. That's probably giving him a little too much credit, and Fourier somewhat less than due credit. But it was a neat theorem. What he said is that there is a function \hat{u} of f , which we'll call the Fourier transform, which has the property that when you take the integral of the difference between \hat{u} of f and the transform of b sub a of t , when you take the integral of this dt -- in other words, when you evaluate the energy in the difference between \hat{u} of f and v sub a of f , that goes to zero. Well this isn't a big deal. In other words, this is plausible, since this integral has to go to zero for an L^2 function, that's what we just said, then therefore, using the energy relation, this also has to go to zero.

So is this another example where Plancherel just came along at the right time and he said something totally trivial and became famous because of it? I mean as I've urged all of you, work on problems that other people haven't worked on. If you're lucky, you will do something trivial and become famous. As another piece of philosophy, you become far more famous for doing something trivial than for doing something difficult, because everybody remembers something trivial. And if you do something difficult nobody even understands it.

But no, it wasn't that because there's something hidden here. He says a function like this exists. In other words, the problem is these functions get closer and closer to something. They get closer and closer to each other as n gets bigger and bigger. You can show that because you have a handle on these functions. Whether they get closer and closer to a real bonafide function or not is another question. Back when you studied arithmetic, if you were in any kind of advanced class studying arithmetic, you studied the rational numbers and the real numbers you remember. And you remember the problem of what happens if you take a sequence of rational numbers which is approaching a limit. There's a big problem there because when you take a sequence of rational numbers that approaches a limit, the limit might not be rational. In other words, when you take sequences of things you can get out of

the domain of the things you're working with.

Now, we can't get out of the domain of being L_2 , but we might get out of domain of measurable functions, we might get out of the domain of functions at all. We can have all sorts of strange things happen. The nice thing here, which was really a theorem by [? Resenage ?] a long time ago. It says that when you take cosine sequences of L_2 functions, they converge to an L_2 function. So that's really what's involved in here. So maybe this should be called the [? Resenage ?] transform, I don't know. But anyway, whatever this says, the theorem says, the first part of Plancherel's theorem says that this function exists and you get a handle on it by taking this transform, making a bigger and bigger, and it says it will converge to something in this energy sense. Bingo, when you're all done this goes to zero. We're going to denote this function as a limit and a mean of the Fourier transform. In other words, we do have a Fourier transform in the same sense that we had a Fourier series before. We didn't know whether the Fourier series would converge at every point, but we knew that it converged at enough points, namely, almost everywhere, everywhere but on a set of measure zero. It converges so that, in fact, you get this kind of relationship.

Now, do you have to worry about that? No. Again, this is one of these very nice things that says there is a Fourier transform, you don't have to worry about what goes on at these oddball sets where the function has discontinuities and things like that. You can forget all of that. You can be as careless as you've ever been, and now you know that it all works out mathematically, so long as you stick to L_2 functions. So, sticking to L_2 functions says you can be a careless engineer, you can use lousy mathematics and you'll always get the right answer. So, it's nice for engineers, it's nice for me. I don't like to be careful all the time. I like to be careful once and then solve that and go on.

Well, because of time frequency duality, you can do exactly the same thing in the frequency domain. So, you start out defining some b , which is bigger than zero which is arbitrarily large, you define a finite bandwidth approximation as \hat{u}_b of f is \hat{u} of f . We now know that \hat{u}_b of f exists and it's an L_2 function. \hat{u}_b of f

times this rectangular function, that's f over $2b$. In other words, it's \hat{u} of f truncated to a big bandwidth minus b to plus b . Since w sub b of f is L^1 , as well as L^2 , this always exists. So long as you deal with a finite bandwidth, this quantity exists. It exists for all t and r . It's continuous. The second part of Plancherel's theorem then says that the limit as b goes to infinity of the integral of u of t minus this truncated function, magnitude squared the energy in that, goes to zero. This now is a little different than what we did before. It's easier in the sense that we don't have to worry about the existence of this function because we started out with this to start with. It's a little harder because we know that a function exists, but we don't know that it's u of t . So, in fact, poor old Plancherel had to do something other than just this very simple argument that says all the energy works out right. He had to also show that you really wind up with the right function when you get all through with it. But again, this is the same kind of energy convergence that we had before. Yeah?

AUDIENCE: Could you discuss non-uniqueness? Clearly, [INAUDIBLE] \hat{u} of f to satisfy Plancherel 1 and Plancherel 2.

PROFESSOR: Yeah, in fact, any two functions which are L^2 equivalent. But you see the nice thing is when you take this finite bandwidth approximation there's only one. It's only when you get to the limit that all of this mess occurs. If you take these different possible functions, \hat{u} of f , which just differ in these negligible sets of measure zero, those don't affect this integral. Sets of measure zero don't affect integrals at all. So the mathematicians deal with L^2 theory by talking about equivalence classes of functions. I find it hard to think about equivalence classes of functions and partitioning the set of all functions into a bunch of equivalence classes. So I just sort of remember in the back of my mind that there are all these functions which differ in a strange way. We'll talk about that more when we get to the sampling theorem later today, because there it happens to be important. Here it's not really important, here we don't have to worry about it. Anyway, we can always get back to the u of t that we started with in this way.

Now, this says that all L^2 functions have Fourier transforms in this very nice sense.

In other words, at this point you don't have to worry about continuity, you don't have to worry about how fast things drop off, you don't have to worry about anything. So long as you have finite energy functions, this beautiful result always holds true. There always is a Fourier transform, it always has this nice property that it has the same energy as the function you started with. The only nasty thing, as Dave pointed out, is that, in fact, it might not be a unique function, but it's close enough to unique. It's unique in an engineering sense.

The other thing is that L2 wave forms don't include some of your favorite wave forms. They don't include constants, they don't include sine waves, and they don't include Dirac impulse functions. All of them have infinite energy. I pointed out in class, spent quite a bit of time explaining why an impulse function had infinite energy by looking at it as a very narrow pulse of width ϵ and a height $1/\epsilon$, and showing that the energy in that is $1/\epsilon$, and as ϵ goes to zero and the pulse gets narrower and narrower, bingo, the energy goes to infinity.

Constants are the same way, they extend on and on forever. Therefore, they have infinite energy, except if the constant happens to be zero. Sine waves are the same way, they dribble on forever. So the question is are these good models of reality? The answer is they're good for some things and they're very bad for other things. The point in this course is that if you're looking for wave forms that are good models for either the kinds of functions that we're going to quantize, namely, source wave forms, or if you're looking for the kinds of things that we're going to transmit on channels, these are very lousy functions. They don't make any sense in a communication context. But anyway, where did these things come from?

Constants and sine waves result from refusing to explicitly model when very long-lasting functions terminate. In other words, if you're looking at a carrier function in a communication's system, $\sin(2\pi f \text{ carrier times } t)$, it just keeps on wiggling around forever. Since you want to talk about that over the complete time of interest, you don't want to say what the time of interest is, you don't want to admit to your employer that this thing is going to stop working after one month because you want

to let him think that he's going to make a profit off of this forever, and you don't want to commit to putting it into use in one month when you know it's going to get delayed for a whole year. So you want to think of this as something which is permanent. You don't want to answer the question at what time does it start and at what time does it end, because for many of the questions you ask, you can just regard it as going on forever.

You have the same thing with impulses. Impulses are always models of short pulses. If you put these short pulses through a filter, the only thing which is of interest in them is what their integral is. And since the only thing of interest is their integral, you call it an impulse and you don't worry about just how narrow it is, except that it has infinite energy and, therefore, whenever you start to deal with a situation in which energy is important, these become lousy models. So we can't use these when we're talking about L2 functions. That's the price we pay for dealing with L2 functions. But it's a small price because for almost all the things we'll be dealing with, it's the energy of the functions that are really important. So as communication wave forms, infinite energy wave forms make mean square error quantization results meaningless. In other words, when you sample these infinite energy wave forms you get results that don't make any sense, and they make those channel results meaningless. Therefore, from now on, whether I remember to say it or not, everything we deal with, unless we're looking at counter examples to something, is going to be an L2 function.

Let's go on. I'm starting to feel like I'm back in our signals and systems course, because at this point I'm defining my third different kind of transform. Fortunately, this is the last transform we will have to talk about, so we're all done with this. The other nice thing is that the discrete time Fourier transform happens to be just the time frequency dual of the Fourier series. So that whether you've ever studied the dtft or not, you already know everything there is to know about it, because the only things there are to know about it are the results about Fourier series.

So the theorem is really the same theorem that we had for Fourier series. Assume that you have a function of frequency, \hat{u} of f -- before we had a function of time,

now we have a function of frequency. Suppose it's defined over the interval $-\omega$ to $+\omega$. In other words, a way we often say that a function is truncated is to say it goes from some interval into c . This is a complex function which is non-zero only for f in this finite bandwidth range. We want to assume that this is L^2 and thus, we know it's also L^1 . Then we're going to take the Fourier coefficients. Before we thought of the Fourier coefficients as corresponding to what goes on at different frequencies, now we're going to regard them as time quantities, and we'll see exactly why later. So we'll define these as the Fourier coefficients of this function. So they're $\frac{1}{2\omega}$ times this integral here. You remember before when we dealt with the Fourier series, we went from $-\frac{t}{2}$ to $+\frac{t}{2}$. Now we're going from $-\omega$ to $+\omega$. Why? It's just convention, there's no real reason. So that what's happening here is that this is a Fourier series formula for a coefficient where we're substituting ω for $\frac{t}{2}$. We're putting a plus sign in the exponent instead of a minus sign. In other words, we're doing this hermitian duality bit. What's the other difference? We're interchanging time and frequency. Aside from that it's exactly the same theorem that we established -- well, that we stated before.

So we know that this quantity here, since \hat{f} is L^1 , this is finite and it exists -- that's just a finite complex number and nothing more -- for all integer k . Also, the convergence result when we go back, this is the formula we try to use to go back to the function we started with. It's just a finite approximation to the Fourier series. We're saying that as k gets larger and larger, this finite approximation approaches the function in energy sense. So this is exactly what you should mean by a Fourier series anyway. It's exactly what you should mean by a Fourier transform anyway. As you go to the limit with more and more terms you get something which is equal to what you started with, except on the set of measure zero. In other words, it converges everywhere where it matters. It converges to something in the sense that the energy is the same. I said that in such a way that makes it a little simpler than it really is.

You can't always say that this converges in any nice way. Next time I'm going to show you a truly awful function which we'll use in the Fourier series instead of \hat{f} , which is time limited and which is just incredibly messy and it'll show you why you

have to be a little bit careful about stating these results. But you don't have to worry about it most of the time, because this theorem is still true. It's just that you have to be a little careful about how to interpret it because you don't know whether this is going to reach a limit or not. All you know is that this will be true, this energy difference goes to zero.

Also, the energy in the function is equal to the energy in the coefficients. This was the thing that we found so useful with the Fourier series. It's why we can play this game that we play with mean square quantization error of taking a function and then turning it into a sequence of samples, trying to quantize the samples for minimum mean square error and associate the mean square error in the samples with the mean square error on the function. You can't do that with anything that I know of other than mean square error. If you want to deal with other kinds of quantization errors, you have a real problem going from coefficients to functions.

And finally, for any set of numbers u_k , if the sum is less than infinity, in other words, if you're dealing with a sequence of finite energy, there always is such a frequency function. Many people when they use the discrete time Fourier transform think of starting with the sequence and taking a sequence and saying well it's nice to think of this sequence in the frequency domain, and then they say a function f exists such that this is true, or they just say that this is equal to that without worrying about the convergence at all, which is more common. But since we've already gone through all of this for the Fourier series, we might as well say it right here also. So there's really nothing different here. But now the question is why do these u_k 's -- I mean why do we think of those as time coefficients? I mean what's really going on in this discrete time Fourier transform. At this point it just looks like a lot of mathematics and it's hard to interpret what any of these things mean. Well, the next thing I want to do is to go into the sampling theorem. The sampling theorem, in fact, is going to interpret for you what this discrete time Fourier transform is, because the sampling theorem and the discrete time Fourier transform are just intimately related, they're hand and glove with each other, and that's the next thing we want to do. But first we have to re-write this a little bit. We're going to say that this frequency

function is the limit in the mean of this -- this rectangular function is what we use just to make sure we're only talking about frequency between minus w and plus w . The limit in the mean, there's a little notational trick that we use so that we can think of this as just a limit instead of thinking of it as this crazy thing that we just derive, which is really not so crazy. That means we can talk about this transform without always rubbing our noses in all of this mess here. It just means that once in awhile we go back and think what does this really mean. It means convergence in energy rather than convergence point-wise, because we might not have convergence point-wise.

So we're going to write this also as the limit in the mean of the sum over k of u of k . We're going to glop all of this together. This is just some function of k and a frequency. So we're going to call that ϕ sub k of f at some wave form, and the wave form is this. What happens if you look at the relationship between ϕ of k and ϕ of k' ? Namely, if you look at two different functions. These two functions are orthogonal to each other for the same reason that the functions that the sinusoid you used in the Fourier series are orthogonal. Namely, you take this function, you multiply it by $e^{-2\pi i k' f / 2w}$ times this rectangular function. You integrate from minus w to w and what do you get? You're just integrating a sinusoid -- the whole thing is one big sinusoid -- over one period of that sinusoid or multiple periods of the sinusoid. Actually, $k - k'$ periods of the sinusoid. And when you integrate a sinusoid over a period, what do you get? You get zero. So it says that these functions are all orthogonal to each other. So, presto, we have another orthogonal expansion just like the Fourier series gave us an orthogonal expansion. And in fact, it's the same orthogonal expansion.

Now the next thing to observe is that we have done the Fourier transform and we've also done the discrete time Fourier transform. In both of them we're dealing with some frequency function. Now we're dealing with some frequency function which is limited to minus w to plus w , but we have two expansions for it. We have the Fourier transform, so we can go to a function u of t , and we also have this discrete time Fourier transform. So u of t is equal to this Fourier transform here. Again, I should write limit in the mean here, but then I think about and I say do I have to write limit in

the mean? No. I don't need a limit in the mean here because I'm integrating this over finite limits. Since I'm taking a function over finite limits, \hat{u} of f is over these limits, is an L^1 function, therefore, this integral exists. This function is a continuous function. There aren't any sets of measure zero involved here. This is one specific function which is always the same. You know what it is exactly at every point. At every point t , this converges.

So then what we're going to do, what the sampling theorem does is it relates this to what you get with a dtft. So the sampling theorem says let \hat{u} of f be an L^2 function which goes from minus w to c , and let u of t be this, namely, that, which we now know exists and is continuous. Define capital T as 1 over $2w$. You don't have to do that if you don't want to, but it's a little easier to think in terms of some increment of time, T , here. Then u of t is continuous, L^2 and bounded. It's bounded by u of t less than or equal to this. Why is that? It doesn't make any sense as I stated it. Now it makes sense. OK, its magnitude is bounded. Its magnitude is bounded because if you take u of t magnitude, it's equal to the magnitude of this which is less than or equal to the integral of the magnitude of this, which is equal to the integral of the magnitude of just \hat{u} of f , which is what we have here. So all of that works nicely.

So, u of t is a nice, well-defined function. Then the other part of it is that u of t is equal to the sum of its values at these sample points times sinc of t minus kt over t . Now you've probably seen this sampling theorem before. How many people haven't seen this before? I mean aside from the question of trying to do it -- do it in a way that makes sense. OK, you've all seen it. So good. What it's saying is you can represent a function in terms of just knowing what its samples are. Or you can take the function, you can sample it, and when you sample it if you put these little sinc hats around all the samples, you get back to the function again. So you take all the samples, you then put these sincs around them, add them all up, and bingo, you got the function.

Let's see why that's true. Here's the sinc function here. The important thing about $\text{sinc } t$, $\frac{\sin \pi t}{\pi t}$, which you can see by just looking at the sine function, is it

has the value 1 when t is equal to zero. I mean to get that value 1, you really have to go through a limiting operation here to think of $\sin \pi t$ when t is very small as being approximately equal to πt . When you divide πt by πt you get 1, so that's its value there and value around there. At every other sample point, namely, at t equals 1, \sin of πt is zero. At t equals 2, \sin of πt is equal to zero. So the sinc function is 1 at zero and is zero at every other integer point. Now to see why it's true and to understand what the dtft is all about, note that we have said that a frequency function, \hat{u} of f , can be expressed as the sum over k -- and I should use a limit in the mean here but I'm not -- of u_k times this transform relationship here. Well, these are these functions that we talked about awhile ago. They're these functions. They're the sinusoids, periodic sinusoids in k truncated in frequency.

So we know that \hat{u} of f is equal to that dtft expansion. If I take the inverse Fourier transform of that, I can take the inverse Fourier transform of all these functions and I'll get u of t is equal to the sum over k , of $u_k p_k$ of t . I'm being careless about the mathematics here. I've been careful about it all along. The notes does it carefully, particularly in the appendix. I'm not going to worry about that here. If I take the function p_k of f , which is this truncated sinusoid, and I take the inverse transform of that -- take the transform of this, I get this. Can you see that just by inspection? If you were really hot on these things, if you just finished taking 6.003, you could probably see that by inspection. If you remember all of those relationships that we went through before, you can see it by inspection. The Fourier transform of a sinc function is a sinc function. This exponential here when you go into the time domain corresponds to a time shift, so that gives rise to this time shift here. The 1 over t is just one of these constants you have to keep straight, and which I would do just by integrating the things to see what I get.

Finally, u of kt , if I look at this, is just 1 over t times u of k , because these functions here are these sinc functions, which are zero everywhere but on their own point. So if I look at u of kt , it's a sum over k of c_k of kt , and c_k of kt is only 1 when t is equal to k times capital T , and therefore, I get that point there. Therefore, u of kt is just 1 over t times u of k . That finishes the sampling theorem except for really

tracing through all of these things about convergence, but it also tells you what the discrete time Fourier transform is. Because the discrete time Fourier transform is just scaled samples of u of t . In other words, you start out with this frequency function, you take the inverse Fourier transform of it, you get a time function. You take the samples of that, you scale them, and those are the coefficients in the discrete time Fourier transform, which is what you use discrete time Fourier transforms for. You think of sampling of function, then you represent the function in terms of those samples and then you want to go into the frequency domain as a way of dealing with the properties of those samples, and all you're doing is just going into the Fourier transform of u of t , and all of that works out.

There's one bizarre thing here, and I'm going to talk about that more next time. That is that when you look at this time frequency limited function, \hat{u} of f , \hat{u} of f can be very badly behaved. You can have a frequency limited function which does all sorts of crazy things. Since it's frequency limited as inverse transform, it's beautifully behaved. It's just the sum of sinc functions -- it's bounded, it's continuous and everything else. When we go back into the frequency domain it is just as ugly as can be. So what we have in the sampling theorem, it comes out particularly clearly. Is that L1 functions in one domain are nice, continuous, beautiful functions which are not L1 in the other domain. So you sort of go from L1 to continuous.

Now when we're dealing with L2 functions, they're not continuous, they're not anything else, but you always go from L2 to L2. In other words, you can't get out of the L2 domain, and therefore, when you're dealing with L2 functions, all you have to worry about is L2 functions, because you always stay there no matter what. When we start talking about stochastic processes, we'll find out you always stay there also. In other words, you only have to know about one thing. We've seen here that to interpret what these Fourier transforms mean, it's nice to have a little idea about L1 functions also because when we think of going to the limit, when we go to the limit we get something which is badly behaved, and for these finite time and finite frequency approximations, we have things which are beautifully behaved.

I'm going to stop now so we can pass out the quizzes.