TheFourierTransformAndItsApplications-Lecture24

Instructor (Brad Osgood):That was quick. Looks like the crowd's a little thin out there. This looks more like the Wednesday before Thanksgiving than – instead of the Monday before the Wednesday before Thanksgiving. Okay.

That was the big fear that – when Stanford switched to giving the entire week off of Thanksgiving, they just sort of – people would take off earlier and earlier. And sure enough – all right.

Anyway, I want to continue with our discussion of linear systems – big day today, actually. It's really – the state of affairs, as you'll see when we wrap up the discussion, is really quite satisfactory and understanding linear systems and the structure of linear systems. And then we'll make further use of it next time when we talk about – more essentially about time invariant systems as a special case.

So let me remind you sort of where we finished up last time. This is linear systems part deux. And I was sort of making the comparison between the continuous case and the discreet case. We're gonna spend most of our time today talking about the continuous case, but one of the places where I made the connection or the analogy was in the discreet case, any linear system has to be given by multiplication by a matrix. All right?

So in the finite dimensional, I said – when I talk about discreet case, I'm sorta thinking about N-by-N matrices and discreet signals and so on. So in the discreet finite dimensional case, as opposed to the continuous infinite dimensional case where the inputs and outputs are functions – so here, they're just finite vectors. Any linear system is given by multiplication of the matrix by a matrix. All right? That's a fundamental fact that you will learn in linear algebra.

The matrices are different. You get a matrix by choosing a basis of the space of inputs, and then you express the matrix in terms of what happens to the bases. And I'm not gonna go through this because I'm assuming that you've seen this in linear algebra, although you may not have thought about it quite in these terms.

And what I wanna see is that an analogous result holds in the continuous case, and it's really quite striking.

So an analogous result holds in the continuous infinite dimensional case when the inputs and outputs are functions of a continuous variable instead of discreet functions finite – just a finite list of numbers – a vector – infinite dimensional case.

All right, now this gets into – to get a completely precise statement of this. And to see to what extent this really holds gets into pretty deep waters, all right? And that's not my intention to go there, but I do want you to sort of see what the point is.

All right. And what it amounts to is understanding a little bit better about the special case of a linear system is given by integration against the kernel. So we'll see a linear system, in this case – in the infinite dimensional continuous case – is given by integration against the kernel – against a kernel.

This was sort of a example of – one example of a linear system in the infinite dimensional continuous case K of XY, V of Y, DY. That's L operating on V is of that form, and that we talked about last time as an analogy to matrix multiplication where K is playing the role of the matrix. And what I wanna see is that this is not just a good example of a linear system, it's the only example of a linear system. That is any linear system essentially looks like this.

So to do that, I have to tell you what the K is, and I have to actually produce the K – the kernel – against what you integrate. So I have to produce a K for a general linear system. All right? In order to show something like this.

All right. Now to do this, I need a brief digression into the idea of cascading linear systems – following one linear system with another. So we need a little digression here, and I'm not gonna take this very far, I just wanna – in a particular case. So we need a digression on cascading or composing. Engineers tend to use the word cascading. Mathematicians tend to use the word composing linear systems.

The idea being that if you have an input into two linear systems, one linear system following another, say LM. This is V and this is W. Then if L and M are each linear, so is their cascade. If L and M are each linear, then M – L followed by M, W equals ML applied to V. First L then M – you can see by the way I've written it there in that diagrammatic way – is also linear. All right.

I'm not gonna check that. You have seen this before. It's not interesting to check. But it comes up often, all right? The idea of following one linear system with another, and the fact that linearity is preserved under a system like that.

Now, the only way – the only place I wanna use this is when one of the systems is – the first system is given by a integration against the kernel, and then you follow that with another system. So I wanna look at a special case. When – say L is given by integration against the kernel. So L of V of X is, say, the integral for minus infinity to infinity. Once again, K of XY, V of Y, DY. All right?

What happens if you follow that with another linear system L – M, excuse me. So what is M applied to L at V?

All right. That's not hard to see. And again, it's very – it's quite a satisfactory result. The result of this, actually, is applying M to K. So in fact, I'll give you the punch line, and then I'll tell you why it's true. M applied to L at V of X is given by the integral from minus infinity to infinity of M applied to K. Now, let me get a little bit more careful about this. Let me write it down, and I'll tell you what I mean.

I put a little subscript X here indicating M acts on functions of one variable, so to speak. It takes inputs to outputs, where the input is a function and the output is a function. So when I'm writing this, I'm meaning that M is operating on a function K of XY in the X variable, so to speak. All right?

So it's not a great notation, but it's the best I can do. I just want to indicate this. There's a little – something a little bit that has to be said here because K is a function of two variables. So you think of K of XY operate it – operate on it as if Y were fixed, and then operate it with respect to the X variable. And this is what happens when you operate on the system is given by integration against the kernel.

Now let me show you why that's – that's pretty nice, right? That's a nice result. And let me show you why that's true – again, without all the details.

All of these things, if you wanna give a very careful mathematical statement of them, require extra assumptions – require taking limits, knowing when you can take limits and so on. And again, this is beyond our concern. They say the rigger police are off duty.

But why does this work? Well, the idea is you approximate the integral by a sum and then apply ordinary linearity. So K of XY, D of Y, DY is approximately – the integral's approximately the sum. Let me not write the [inaudible]. Just sum over I, K of XYI, V of YI, delta YI. All right? The usual way you approximate an integral by the sums. So I'm leaving X here as the continuous variable and replacing Y by a discreet set of measurements Y1, Y2 – all of them. Whatever. Okay? Because I'm integrating with respect to Y.

Now, if you operate with M on the sum – and here you'll see why – what I mean by M operating in the X variable. So M operating on the sum K, XYI, VYI, delta YI – and it's linear. So a linear operator applied to a sum is the sum of M applied to the components in here. So this is the sum over I of M applied to all this inside – K of XYI, V of YI, delta YI. But again, as far as M is concerned – if M is operating in the X variable, then V evaluator YI and delta YI are constants as far as M is concerned. All right?

So here is where – and here's what I meant. Here's where you can see what I meant by M operating on the X variable. That is, again, because it's linear, the [inaudible] position is M of a constant times a function is a constant times M of the function. In this case, the constants are these V of YI and the delta YI. So this is the sum over I of – let me write it like this: M operating the X variable K of XYI times V of YI delta YI.

And now, think of this as the approximation going the other way. That is, this is the sum approximating the integral. We started off as the integral being approximated by the sum. Now think of this as the sum approximating an integral. So this is approximately the integral for minus infinity to infinity M operating the X variable K of XY, V of Y, DY. All right?

So once again, just to summarize – so if L looks like this: L of V of X is the integral from minus infinity to infinity K of XY, V of Y, DY, then M of L of V of X is the integral from minus infinity to infinity of M operating on K of XY, V of Y, DY. Okay?

This would take a lot of work to actually make this into a completely rigorous proof. And again, I don't wanna do that. Okay?

But it is this sort – and I wouldn't expect you to do it. But it is – but I would expect you to be able to believe the argument. And this is the sort of argument that you should be able to do more and more on your own to see what should the properties be, how should it work and so on to see that it's reasonable. All right? To see that something like this is – try to – if somebody gives you a statement like this, ask yourself why is it reasonable? And to give yourself an argument like that to justify it, I think, is a good thing to develop – to develop being able to do. All right?

Okay. Now, that's all I wanted to say about cascading. Cascading is a big – there are lots of examples. There are lots of different things you can do with it. But I'm only mentioning it toward the end of establishing this result – this really striking result that any linear system is integration against some kernel – can be given as integration against some kernel. So it's back to live action.

Back to the main plot – the main plot being that any linear system is given by integration against a kernel. All right. Now, I'm gonna play – I'm gonna bring back distributions into the picture. All right? And I'm actually gonna write distributions in terms of integration. So when I say integration against a kernel when we follow through using that language – and at the end, I'll say a little bit more generally how you would make this more precise in terms of pairings and distributions and all the rest of that stuff. But put that aside for the moment and just think about distributions and, particularly, delta functions in the way you probably first learned how to use them. That is to say if V of X is any function, I know I can write V of X as the integral from minus infinity to infinity delta X minus Y – it's like convolution with a delta function – V of Y, DY. All right? That expresses V in terms of a delta function. You know what this is, actually – is this is the continuous analogue of expressing a vector in terms of the sum of its components. All right? If you think of the discreet delta functions as just one in one slot and zeros every place else, then writing a vector as V1 times the first basis vector plus V2 times the second basis vector plus V3 times the third basis vector is exactly the continuous analogue of writing V as the integral against a shifted delta function here. It's exactly the same thing.

So again, if I were doing this – and I'll come back to this later. If I were doing this precisely with distributions, I wouldn't talk about integration. I'd talk about a pairing. But think of it in terms of integration now. All right? Because it's the way you're probably a little bit more familiar with, and it allows us to keep up the analogy with the finite case. All right. Well now, if L is any linear system – I'll call it L instead of M – I wanna apply L to V by applying L to that integral. All right? So if L is a linear system, then I find LV by applying L to the integral. All right? That is to say L of V of X is L applied to the integral from minus infinity to infinity V of Y delta – let me write the delta first, there –

delta X minus Y, V of Y, DY. And from what I just said, you can already view that as a linear system – as integration against a kernel, the kernel being the delta function. So what happens is L comes inside and gets applied to the delta function. That is – this is the integral from minus infinity to infinity of L of delta X minus Y. And if I wanna be a little careful here, it's like L applied in the X variable. All right? Delta X minus Y times V of Y, DY. And if you'll look at it, and if you believe in fairies – if you believe in magic, we're done. That is to say now set – say H or K – that K of XY to be L of delta X minus Y, L operating in the X variable. I won't keep writing that, but that's always what's going on here. All right. Then it's exactly – then the system is given exactly in the form of integration against a kernel.

Then we have L of V of X is the integral from minus infinity to infinity of K of XY, V of Y, DY. Amazing, actually. It's really quite striking. Any linear system is given by integration against a kernel. What is the kernel? The kernel is what the system does when you feed it a delta function. Now there's terminology that goes along with this that, again, I'm sure you have heard. Delta is called a respond – delta is called an impulse. When you feed an impulse into the system, the system responds. And so you call K of XY the impulse response. All right. That's the terminology that is always used here. It is how a system responds to feeding it an impulse. It's shifted impulse, not just a delta function of zero. But the delta function at Y – that is to say delta of X minus Y. The system L responds to the input – to an impulsive input, namely delta X minus Y. Okay? And I think I've already broken protocol here because the impulse response is always denoted by H instead of K. So sue me. All right. Before, I was calling things K instead of H. You always – you almost always, for some reason – history seems to dictate that you call the impulse response H of XY instead of K of XY. All right. Maybe I'll go back to that. Now, play it a little fast and loose with the facts here. All right? That is we were writing a delta function. We were writing how do you use a delta function. Well, you use it in terms of integration. We wrote sort of the V of X is equal to the integral from minus infinity to infinity delta X minus Y, V of Y, DY and so on. We applied L to not just an ordinary function, but to a distribution. Does that really all make sense and so on.

Now this whole method and – this method of writing – of working with linear systems was very well to engineers and was used very effectively in a lot of different applications before mathematicians decided to get their hands on it and try to make it rigorous. And again, it was Laurent Schwartz – the same person who founded the rigorous theory of distributions – who also found a rigorous treatment of just this result. All right? And it would probably be considered – I think it's probably fair to say it would be considered the deepest theorem – the hardest theorem in his whole theory of distributions. All right? And I won't write it out in detail, but I think it's worth noting here. And I also want to recall that Schwartz taught at the École Polytechnique in Paris. All right? That's France's leading engineering school, and so I'm sure – quite sure – I don't know the history of this really, and I'd be interested to know – that he was thoroughly familiar with the applications and thoroughly familiar with sort of how engineers viewed things. All right? So I wouldn't be surprised at all if he understood that that was in the context. But being a mathematician, and so socially deviant, he felt like he had to put this on a rigorous foundation, which he did. And it's usually called the Schwartz Kernel Theorem because

the deepest fact in the – I think the way I described this in the notes is the deepest fact in the entire theory of distributions is well known – sort of been used every day by every electrical engineer – namely this sort of – the system is given by integration against the impulse response. Schwartz would say it something like this: if M – if L is a linear operator on distributions – so it sends one distribution to another distribution.

And again, you have to make certain assumptions here on continuity or boundedness, but I think the assumptions are generally pretty mild. If L is a linear operator on distributions, then there is a unique – actually, part of the contribution was the unique kernel, K, which is another distribution, so that L of V is given by the pairing of K with V. Okay? He would say it something like this. And you have to understand the nature of the pairing and so on. All right? So there is a precise statement of what I've written down here using delta functions and integrating against delta functions and the whole theory of distributions – I'm not gonna go – I'm not gonna really say anything more about this other than this is – what we said so far can be justified. Although we'll use it in practice like I've written it down there, in fact, it does fit into a more general framework. And it's really nicely and fully established, but it's very – it's quite hard. All right? This is, I think, considered – I think it's fair to say this is probably considered the hardest theorem to prove it in its full generality in the whole theory of distributions. All right? And for us, again, in many cases, we consider the pairing to be given by integration. Now, we know it's not always given by integration. There's a pairing – that's how distributions are defined in terms of a pairing – but in many circumstances, certainly practically working with it, you often view that pairing as given in terms of integration. So that's where the connection is. All right? So this is called the Schwartz Kernel Theorem. I hope I'm spelling his name right. I can never quite remember – Kernel Theorem. Okay? It's a big deal. It's a big deal.

Now, let's do a couple of examples of this. Okay? And again, by the way, the kernel in Schwartz's theory works out to be L applied to the delta function or the delta distribution. All right? So that – it's not as though you sort of prove it in one case not very rigorously and then do something completely different in the case where you're trying to prove it rigorously. Again, in Schwartz's theory, K turns out to be L applied to the delta, but things have to be understood in terms – things have to be understood properly in terms of pairing and what you assume about the operator and so on. But let's look at some examples. All right. What is the impulse response for the Fourier Transform? View it as a linear system. There's a curveball for ya. We've been studying now for a couple of months. Now if you view the Fourier Transform as a linear system, what is its impulse response?

Now you can answer – actually answer this several different ways. All right? You can answer it based on the theorem, which is how I [inaudible], or you can answer it based on the properties of a Fourier Transform that we have seen, namely the Fourier Transform and the definition of the impulse response that we have seen. By definition, the impulse response is the system applied to a shifted delta function. All right? So the Fourier Transform – the impulse response of the Fourier Transform is how the Fourier Transform responds when you apply it to an impulse. So you have to know the Fourier Transform

delta X minus Y. And we actually know that. We figured this out. It is E to the minus two pi IXY. Okay?

And actually, that was a fairly straightforward calculation that we did on the basis of the definition – the more general definition of the Fourier Transform. However, let me also point out that you know the Fourier Transform is given by this formula. The Fourier Transform – I'll use variables X and Y instead of S and T – is integral from minus infinity to infinity. Even the minus two pi IXY, F of Y, DY. Okay? Now this exhibits the Fourier Transform as integration against a kernel. All right? The Fourier Transform is a linear system. When you write it like that – is realized as integration against a kernel. Here's the kernel. Okay? The kernel K of XY or H of XY – I'll call it H so – to keep my response – to keep my traditions up – is E to the minus two pi IXY. Schwartz's theorem says – and this is the aspect of Schwartz's theorem that I wanted to mention – is it says there is a unique kernel so that the linear operator is given by the pairing of the kernel with V. All right? And that kernel has to be the impulse response. That kernel has to be L applied to a delta function, and so the fact that the kernel – the fact that the Fourier Transform is given by integration against this kernel implies – and in fact, the kernel is unique – has to be the impulse response – actually implies that the Fourier Transform applied to delta X minus Y, which is how you get the impulse response, has to be given by this complex exponential. You get the reasoning there? All right? It's not circular – complicated way of getting the fact that the Fourier Transform over shifts the delta function as a complex exponential, but it's – it sort of shows how all these ideas are consistent with each other and kind of circle around – not circular reasoning, but how all the things I just circled around – let me just say it one more time.

You know by – so we know that any linear operator has to be given by integration against the impulse response. Let me just not call it kernel – the integration against the impulse response. All right? And furthermore, the impulse response is unique. So if you express – this can sometimes be useful. If you can express your linear operator as integration, then you have found the impulse response. All right? I'll say that one more time. If you – if by – if somebody gives it to you or you can figure it out how your linear system can be written as integration – as an integral of H of XY times the input, then that H of XY has to be the impulse response. All right? So in this case, the Fourier Transform is given by integration against a kernel. That kernel has to be the impulse response. That has to be the Fourier Transform applied to delta X minus Y. Cute. It's cute. I like these cute things.

Let me look at another example. Actually, let me give an example for you to think about. Let me go back briefly. I won't say anything more about it other than to raise the question. We start off by saying this is analogous to the discreet – this is the continuous analogue to the discreet case – the infinite dimensional analogue to the finite dimensional case. Well, what about the finite dimensional case? Finite dimensional case of the discreet finite dimensional discreet case. All right? That is to say if L is a linear operator, L of V plus VA times V where A is a matrix. All right? Any linear system – any finite dimensional linear system is given by multiplication by a matrix. What is the impulse response? All right. All the same words make sense, and all the same reasoning makes sense in the discreet case as it did in the continuous case. You have discreet delta

functions. Integration is replaced by sonation. The continuous variables were replaced by discreet variables. It makes sense to talk about – I'm not gonna give the definition for you. All right? I think I actually – it's written down in the notes, but just think about now yourself. All the definitions make sense. It would be good for you to actually reason this through. All right? Yourself in a quite room, sitting there in the dark. What is the impulse response?

Well, I'll give you the answer, but you should be sure that you understand the answer and think it through yourself. In this case, the impulse response is the system itself. This says the impulse response is A – is A. If a system is given by multiplication by A, then A is also the impulse response for this finite dimensional continuous case. So reason that through. Reason that through. That is to say ask yourself and answer to yourself: what do you mean by the impulse response? What is the analogy here? Does it really work the same way? And you'll find – I work with a discreet delta function instead of a continuous delta function, integration has been replaced by summation and so on. All the same words apply in pretty much the same form. It makes sense to talk about the impulse response for a discreet system. What is it? In this case, if it's given by matrix multiplication, then it's the matrix A.

Let's do another example. Let's do an example of the switch. All right? The switch was – one way of reviewing it is V is – L of V is equal to the rectangle function times V. Okay? You switch on for a second. You switch off after that. All right? So this is just a rectangle function center to the origin of duration one half. Okay? What is the impulse response? And let's check to see that it works, actually, also. Let's check to see they actually get the interval – get the system back by integrating against the impulse response. Okay?

So what is it? Well, L of delta X minus Y. I have to find the impulse response, and I have to compute. I have to see how the system responds to an impulse – to the shift in impulse. So that is equal to pi of X times delta X minus Y if I put my variables in there. Right? Now you will remember – and the sort of L operating in the X variable. Now you remember the property of multiplying a function times a delta function. It's the sampling property of the delta function. That is this is pi of Y times delta X minus Y. All right? That's the sampling property of the delta function. So it appears that H of XY – the impulse response to the switch system as a linear system – and it is a linear system because it's multiplication. It's direct proportion. All right?

So H of X minus Y is pi Y times delta X minus Y. And that's the impulse response. Now if you really want – now, if you're a little skittish about this, we can test it. All right? Does it really work? Does it really work meaning is the system really given by integration against the impulse response. So let's check this. Check this that is what if I get the integral from minus infinity to infinity of K – of H of XY, V of Y, DY. That's integrating X against the impulse response is the integral from minus infinity to infinity of pi Y, delta X minus Y, V of Y, DY. Now put the pi together with the V – that is the integral from minus infinity to infinity, delta X minus Y, pi of Y, V of Y, DY. And again, that's sort of the convolution property of the delta function. If I integrate delta against this function: pi of Y times V of Y, I know what I get. I get pi of X times V of X. Okay? And

that's the system. So it works. Was there ever any doubt? I ask you. I'm a professional. Okay?

There are many other examples. You'll have some examples in homework. There are other examples that are mentioned in the notes. I think this example is actually done in the notes. Let me look at a special case of integration against the kernel – a very important special case of convolution. All right. Everything that we've said before applies in great generality. Now, I wanna look at the special case where the system is given by convolution – special case of convolution. All right – where L of V is equal to H convolved with V. So H is a fixed function here. The input is V. The output is H convolved with V where H is fixed. All right? Which is, of course, so L of V of X is given by the integral from minus infinity to infinity H of X minus Y, V of Y, DY. All right? So now, if you want, we can appeal the Schwartz Kernel Theorem that says the system is given as a integration against a kernel. That kernel must be the impulse response. All right? This H of X minus Y must be the impulse response. So you conclude. There are other ways of seeing this, as well, but I wanna conclude it from that theorem. You conclude that L of delta X minus Y is equal to H of X minus Y.

Okay? That must be the impulse response. Now you can also check this directly. I'm gonna say that directly from the definition that this is so. All right? Again, I'm concluding it from this uniqueness statement in this very general Schwartz Kernel Theorem. Once again, if the linear system is given by pairing with a kernel or integration against a kernel, then that kernel must be the impulse response. All right? Now, there is a special property – and I alluded to this before, and I know you've seen this. There's a special property of convolution that makes it particularly important for linear systems. I mean, we've seen the importance of convolution in all sorts of different aspects – in many guises. All right?

For linear systems, it's the relationship between convolution and delay that turns out to be worth singling out. All right? There's a relationship – a simple relationship between convolution and delay or shift – same thing. All right? So let me remind you – all right – let me remind you of – I wanna write it in terms of this delay operator. All right? So tal sub A of V at X is just delay V by A – V of X minus A. Incidentally, you could ask – I may even ask you at some point – is tal sub A as the delay operator a linear operator? And what's its impulse response? I may ask you that. But now – don't worry. Never mind that right now. It's a delay operator. We've seen this before. And you showed on homework that the delay of the convolution is the convolution of the delay in words. Very nice to say that in words. You showed that the convolution of a delay – of a delayed signal is the delay of the convolution of a signal – of the convolution. That is to say, you showed that if I take H convolved with a delayed signal – this is – so I'm writing this without writing any variables here. But that's to say if I first delay V and then take the convolution, that's the same thing as taking the convolution and then delaying the whole thing.

This is tal sub A of H convolved with V. Okay? You showed that. I can first delay the signal and then convolve, or I can convolve and then delay the signal, and I get the same

result. Okay? Now, I just wanna reinterpret this in terms of linear systems. That's a standard result, and it's a very important result. I'm not gonna say anything different here. I'm just gonna give it, now, a different interpretation – reinterpret in terms of linear systems. All right? It says if the system is given by, again, convolution – here's the system: L of V is equal to H star V. So it's given by convolution. All right? It says if V is delayed – so let's say W is the output. W is equal to L of V – that's H star V – H convolved with V. All right?

So it says if I delay – a delay of the input – that is V going into V of T minus A – or X minus A, all right – causes an identical delay of the output. That is if V goes into W by L – that's H star V – all right? And if V gets delayed to V of T minus A – that's tal sub A applied to V – then the delayed output goes to the delayed input – the delayed input goes to the delayed output W of T minus A. This is H convolved with the delayed signal, but that's the same as – this is tal sub A of W. You like this? That's quite a diagram. That's the same as tal sub A of H star V. All right? Make sure you understand all the different arrows in this diagram and what's happening here. All right? V goes to W by convolution. That's what the system is doing. All right? If V is delayed by applying the delay operator, that means V goes to T minus A. All right? What happens to the output? Well, this is the fundamental relationship between delay and convolution. The convolution of a delay is the same as the delay of the convolution. All right? So if I delay V, and I convolve with H, that's applying L to the delayed signal. Convolving H with the delay of V is applying L to the delayed signal. All right? So this is L of tal sub A of V.

Applying H to the delayed signal is the same thing as applying the delay to the original – to L of the signal. All right? And that produces, therefore, a delay of the output. So delayed inputs go to delayed outputs by the same amount in the case of convolution. And because of that property, you say that the system is time invariant or – it's better to say shift invariant because you don't always think of the variables representing time. But nevertheless, it's almost always referred to as a time invariant system. And I know you guys have seen this before, probably in a signals and systems class. There is a mistake of assuming that every single – every system you see is time invariant because it is such an important class of systems, but it's not the only class of systems. It is the class of systems that is associated intimately with convolution. So again, what have we shown here? So you say that L is, as a system, time invariant or shift invariant. All right? So again, what that means is that if W is equal to L of V, then W at X minus A is equal to L applied to V of X minus A. Shifted inputs go to shifted outputs. That's the definition of time invariance. All right?

Now that's a general definition, and what we just showed was that convolution is a time invariant or shift invariant system. This definition can apply to any linear system. All right? You can ask the question of any linear system: is it time invariant? Is it shift invariant? In which case, you are asking is that statement true? If I delay the input, does that cause a corresponding delay in the output? All right? And what we just showed was that if the system is given by convolution, then that is the case. All right? So we just saw that if the system is given by convolution, then you have time invariance. All right?

Now, the remarkable thing is the converse is also true. I'm just tingling to think about it, but that's me. All right? The converse is also true. What does the converse say? The converse says if you have a time invariance system, then it must be given by convolution. I.e., if L is time invariant, then L is given by convolution. All right. Now, I'm gonna – and I'm gonna show this to you in just a second. This is another indication – one that we haven't seen earlier and maybe the last sort of one we'll see. The general property of convolution indicates why convolution is such an important – and in some case – in many ways such a natural operation. I mean, it was an odd thing to write down – when we first wrote down convolution – where the hell did that come from? And why would anybody think of writing anything like that down? But it's not so – from the point of view of linear systems, it's not so odd to ask it be time invariant. It's not so odd to ask that a system be such that, if you delay the input, you also get a delay of the output by the same amount. That's not – that's a reasonable thing to expect of a system. All right? If I run a program today, I get an answer. If I run a program tomorrow, I'm gonna get the same answer delayed by 24 hours – delayed by a certain amount of time, but I'm gonna get the same result, but I'm gonna get the same result delayed. All right?

If I flip a switch today and something happens, if I flip the switch tomorrow, the same thing happens, we hope. So the input is the same as the output but delayed. Actually flipping the switch may not be a good example because it's not time invariant. Never mind. But you know what I mean? You get the idea that it's not an unnatural thing to say. And the fact is, a time invariant system has to be given by convolution. Now why. All right. We know that it's given by integration against the kernel. All right? So we know that L of V of X is given by the integral from minus infinity to infinity of the impulse – of the impulse response integrated against the function L of delta X minus Y of V of Y times V of Y, DY. We know that for any linear system. All right? That's the impulse response. This is L applied to delta. Okay? So the question is what is that?

So what is L of delta X minus Y? It's gonna be some general thing, but it's gonna be a general thing of a very – of a special type. That is instead of the following, all right? So let me let L of delta X – the unshifted delta function – be H of X. All right? I'm gonna define H of X by that. Okay? Then L of delta X minus Y – think of it this way – is L applied to the shifted delta function. Delta X of Y applied to delta. Okay? But if it's a time invariant system, a shift in input goes to the shift in output. The shift in input goes to the shift in output. That is this is the same thing as tal sub Y of L applied to delta – perhaps I should have said that over here when I was talking about generalities. All right? To say that a shifted input – matter of fact, let me just say that. It says that if a shift – I mean, I sort of said this, but I didn't write it down. If I shift the input by A, that's the same thing as shifting the output by A. All right? So symbolically, I would write it like that. I'm sorry. I probably should have said that. All right?

So this says L applied to the shifted input is the same thing as shifting the output. So another way of saying that, actually, quite nice is that L commutes with shifts. That's the only way of saying – that's probably the quickest mathematical way of saying what a time invariant system is. It commutes with a delay operator. The linear system commutes with the delay operator. First delay then apply L as the same thing as first applying L and

then applying delay. All right. So back to live action, once again, if I shift the inputs, that's shifting tal by Y. That's the same thing as shifting the output. That is tal sub Y apply – of H of X. That is H of – we'll put an X in there. That is H of X minus Y. Okay? Now, it doesn't look like I did much there. That's where the assumption of time invariance comes in – right at this stage. Okay? It makes sense. This definition makes sense. All right? To define – I'm defining function H by how L responds to the impulse centered at zero. But if I know the system is time invariant, and if I know how the system responds to an impulse at zero, I know how it responds to any delayed impulse. It responds to the delayed impulse by delaying the output – delaying the response. Right?

That is L of the delayed signal is the delay of L of a signal – L of the impulse H of X minus Y. All right. So what does that say? You know it's given by this. So that says that L of V of X is equal to the integral from minus infinity to infinity of H of X minus Y, V of Y, DY because it's always given by the integral against the impulse response. Where H of X is equal to L of delta X and H of X minus Y is equal to L of delta X minus Y. The difference between a time invariant system and a general linear system is the impulse response depends only on the difference of the variables X and Y rather than on X and Y separately. All right? It depends only on the difference. And I mentioned this, actually, when we were first talking about linear systems and so on and delays. All right? The difference between a general linear system and a time invariant system is the impulse response is not a function of X and Y independently, but is rather a function of X minus Y. Now this is – so for example, you can see that the switch is not time invariant. All right. Let's go back to that example.

All right, back to the switch. V of W is equal to pi times V. All right? Or L of V is equal to the rectangle function times V. What was the impulse response? The impulse response was, if I remember right – what was it? Was it pi of Y times delta X minus Y. Right? Is that right? Remind me. Yeah. Okay. Is that a function only of X minus Y? No. All right? Not quite. It's not a function of X minus Y alone. It's pi of Y times delta X – this is a function only of X minus Y, but this isn't. All right? Not of the form function of X minus Y alone. Okay? The switch is not a time invariance system. This can be a little disappointing, right? The simplest linear system – the relation of direct proportion multiplication is actually not a time invariant system. It's only when you actually complicate sort of direct proportion and multiplication into this operation of convolution that you actually get a time invariance system. All right. Now, let's summarize. This is a very – I think it's a very – it's a beautiful state of affairs intellectually. I mean, it's a very satisfactory state of affairs. Okay? There are two things that we've really shown today. We've shown that any linear system is integration against a kernel.

As a matter of fact, let me just say integration – well, integration against a kernel – and the kernel's the impulse response – any linear system. All right? The special case of the time invariance system is also the integration against a kernel, but it's a special integration. It's convolution. The impulse response is a function of X minus Y – the difference of the variables, not the variables separately. Two, a system is time invariant if and only if it is given by convolution. I think that's just gorgeous. All right? I think intellectually, that's just very satisfying. Every electrical engineer knows this somehow.

Right? You're all taught this. Maybe not quite in the sweeping grandeur of general linear systems, but it's a very satisfactory state of affairs. You have – any linear system is given by integration against a kernel. That has to be if you really want to understand the theory of it. It gets quite complicated. But never mind. All right. There's the impulse response – linear systems given by integration against the impulse response. In the special important case of time invariance systems, the impulse response takes a special form. The linear system takes a special form. The integration takes a special form of convolution. And if nothing else, it's yet another indication of how fundamental an operation of convolution really is. Okay? Now – one second. We're going. We're about to go. We are about to get a week off. All right? The fact that convolution is entered into the picture now means that the Fourier Transform cannot be far behind because as soon as somebody says convolution to you, your trained behavior at this point should be to salivate and say, "I'll take the Fourier Transform. I know I should. I should do that. Yes. Yes. Yes. Yes. Yes. I'll take the Fourier Transform." And so we will. All right? Anytime convolution comes into the picture, the Fourier Transform cannot be far behind, but it will be a week delayed. All right? So I hope everybody has a wonderful holiday. I will see everybody after Thanksgiving.

[End of Audio]

Duration: 57 minutes