try to be exact.
So today we talk about signal processing.
As one of you already reviewed the chapter
and got rough ideas and steps, that's very good.
And as I mentioned earlier, the book chapters
were drafted, so some typos.
So it's very important that you read the chapter,
still pay attention to what I'm going to explain.
So I fixed the typos, so this is going to be more accurate.
And also, you read the book.
The book chapter, the derivations
may not be as clear as I would explain to you today.
So this is an example.
You can see the difference.
Reading textbook, why you still want
to come to the classroom.
So you compare what you learned from the book chapter
and what I will explain.
Then after the lecture, you do a little bit of review.
And you should get a pretty good understanding.
That's what I hope.
And again, we are on schedule.
So let me show you the schedule here.
And I think early March, I'm going to attend a seminar.
So you can watch the video recorded before.
And also, later March, I'm going to visit Japan
to give a lecture.
So again, we will be arranging one time point here.
And I will be returned.
But I will prepare examination two before my trip.
Anyway, I will be more specific as time goes by.
And so far, we have learned the linear system
and learned the Fourier analysis.
And then I designed a logo to give you
the very, very high level concrete idea.
That's a delta times e to the power i theta.
So this reminds us the Euler formula.
And then you have sine and cosine components.
Delta is imposed.
So the e to the power i theta bring you real
and imaginary part, cosine and sine.
So to have a very general representation power,
you really need to do shift and scaling.
So the delta function, you really need
a safe delta function all the places
and the scale accordingly.
And therefore, sinusoidal waves, you change frequency.
So you can delay it, you can comprise,
all look like a sinusoidal.
And also, amplitude can be scaled.
So with the scaling and the shifting operation,
given your very fundamental building blocks,
such as delta and the sine,
so you can put the components,
translated and the scaled components,
basic components or building blocks together.
You can represent an arbitrary continuous
or piecewise continuous functions.
So that's what we learned.
And then most specifically, you see general function
and then whether the period is a unit or not a unit.
If not a unit, it doesn't make a variable change.
And then you can use what we explained to you.
So you know periodic function with a period of one,

and then you can express that into Fourier expansion.
And therefore, arbitrary periodic function
over the interval from zero to capital T,
you do variable change.
So the S over capital T is a whole thing.
Okay?
We'll have the unit one.
So you can use what I derived for you before.
So you do the variable change
and the T equal to S over T.
For variable S, the period is from zero to capital T.
And what I explained to you,
I like to cite the interval from minus T,
capital T divided by two,
to positive capital T divided by two.
That is easier for us to see
how you can make a transition
from Fourier series to Fourier transform.
Basically, you just like T approaches infinity.
So the minus T over two and the plus T over two,
they approach negative and the positive infinity.
So you have Fourier integrals on here.
So these are very nice tools.
And for Fourier series,
and we say you can just use a complex form.
And that's very compact.
Only one term, you have complex frequency components.
And you have e to the power two pi i n t
or small t or big t.
So just shown here,
this is a complex harmonic or sinusoidal component.
And again, you should remember
one of the most beautiful formulas called the Euler formula.
And how you compute this Cn,
so you do this integral,
you do average over the interval zero to capital T.
So you have this as a basis function doing,
you are doing inner product
with this given arbitrary function.
This is inner product, the geometrical meaning
is the function projected onto a basis function.
So the function can be considered as a vector.
Just its dimensionality is very, very high,
infinitely many dimensions.
Okay, so you learned this.
Again, you learned convolution theorem,
and I said convolution theorem is so important.
And this kind of closely tied linear system
and the Fourier analysis together.
And when the linear system
is a safety invariant linear system,
you know the output is a convolution of input
and the system's impulse response function.
Okay, this is a convolution theorem.
And then you can perform Fourier analysis
upon the input, impulse, response, and the output.
Then we say if you do convolutions,
for example, in time domain,
then we say this operation is same thing
as a multiplication in Fourier domain.
Multiplication between the Fourier transform
of impulse response function
times Fourier transform of input function.
The product is nothing but a Fourier transform

of your output function.
Now you have two ways to compute
a safety invariant system, linear system output.
One way you just do convolution.
The other way you just perform Fourier analysis.
You do multiplication in Fourier domain
to get time domain response.
You perform inverse Fourier transform.
You just get it back.
And then I have one slide.
So how you prove this profound theorem
called convolution theorem, okay?
And then one slide I wouldn't reproduce here.
But I give you an example.
If you have two gate function, you do convolution.
Two gate function, you do convolution.
You just flip one, it's just a gate.
A rectangular function, doesn't matter.
You do translation, then for each translated position,
you do multiplication, add it together.
So you kind of have the area under curve.
So get maximum value when the two rectangular function
perfectly overlap together.
When you move away the value,
the area under curve linearly goes down.
So the two rectangular function,
you do convolution, you got this triangular function.
So this just visualize this process.
Very simple, but a good example.
So on one way, you know the theory.
On the other way, you have some representative examples
which are vividly remembered in your mind.
That's a very good thing.
So this is time domain convolution.
And according to convolution theorem,
you can do the convolution the other way.
We know that rectangular function
has a four-way transform, which is a synchro function.
Kind of sign something over something,
so this is called a synchro function.
So the convolution of rectangular function
is identical copy of a rectangular function.
Then in four-way domain, the synchro function
will be multiplied with itself.
So you got a squared synchro.
According to convolution theorem,
this synchro square will be, must be,
the four-way transform of this triangular function.
This is just a fact, just to help you remember,
see the linkage between linear system and four-way analysis
and visually see how the convolution theorem
works in this simple example.
I asked some questions beyond the scope of the syllabus.
They say why we have convolution theorem,
and we proved in the previous class.
And I asked why the convolution theorem is unique.
And I say this is not required,
but just for you to think about that.
So just say a few more comments here.
Basically, I pointed out in the last lecture,
you say for safe environment linear system,
if you give a sinusoidal as the input to the system,
what you have output must be also
sinusoidal at the same frequency.

This can be proved.
And I also say sinusoidal in, sinusoidal out.
Frequency now changed.
No new frequency ever generated.
And this only holds for sinusoidal function.
If you combine both facts,
right facts and blue facts together,
you will understand that you must have convolution theorem.
And also the convolution theorem
can only hold for four-way transform.
No other transform will have such a good property.
And let me explain more for this.
That's just, you can treat as a homework exercise, okay?
I say you could write something quite good
if you do not use convolution theorem
to show this right point and the blue point.
The right point basically say sinusoidal in,
you must have a sinusoidal out.
The blue point say if you have some functional form in,
the same functional form will be out.
Then the functional form must be sinusoidal, okay?
These are two different points.
And I mean to show these two points
without using what we showed last lecture.
You just derive the convolution theorem.
But these two points can be easily illustrated
with the convolution theorem.
So let me just show you, okay?
It's just a kind of, I think,
a good homework exercise.
So you have one dimensional
safety environment linear system.
So the output is output T
and according to what we learned in the linear system,
safety environment linear system class.
And the output must be impulse response
convolved with input, right?
Now we already know what's hold valid
is convolution theorem.
Then in frequency domain we say the Fourier spectrum
of the output function must be product of impulse,
the spectrum of impulse response function.
This is, so this is lowercase h.
Here you have capital H.
Okay, then the input Fourier counterpart is capital I,
is lowercase i, then I use the frequency component u.
So first line is just soil.
This is what we mean by convolution theorem.
That holds, okay?
So suppose we gave a sinusoidal input.
Sinusoidal input here is special case.
It's just the arbitrary sinusoidal components.
E to the power i two pi u zero t.
So you have a frequency u zero.
And the exponential components, you have sine part,
you have cosine part, you have sine part,
real and imaginary part.
So if you gave a system input as a set,
so what will be the Fourier expression
of this input?
And according to what we learned
in Fourier transform lecture.
So the Fourier spectrum of this sinusoidal,
complex sinusoidal component is a delta function.

So you got this delta function, okay?
So this is, again, this is the very simple
Fourier transform exercise.
So according to frequency expression,
and now you know input is delta u minus u zero.
So you just insert this expression here.
Then you can immediately infer in the frequency domain.
The output Fourier spectrum must be product
of this delta function, this delta function,
and this impulse, the Fourier spectrum
of impulse response, so H, capital H of u.
Capital H of u is also called the transfer function.
So these two things, transfer function
and this delta expression, put it together.
This is the Fourier spectrum of the output function.
It must be this way, right?
So according to the theorem.
And this has a time domain counterpart.
And it is quite similar to input,
but it's weighted by H of u zero.
So you see, I go through these steps.
I already show you.
If your input is a complex sinusoidal component,
output must be also a complex sinusoidal component.
The only difference is coefficient.
Coefficient doesn't matter, just scaling factor.
You scale it larger or smaller,
and then the scaling factor could be complex.
And even the face could change a little bit.
But the frequency is not changed.
So no new frequency could be created
by a shaped environment linear system.
So I utilize the convolution theorem
to show this quite easily.
But this can also be shown without using
convolution theorem, without using first line.
So that needs a little bit more exercise
if you are interested.
And on the other hand, let me show you.
If the above invariability holds for another function,
then another function must be sinusoidal function,
cannot be any other function,
unless the impulse response is delta function.
If the system response is delta function,
then anything you put into the system,
the system will do nothing, just pass the thing out.
So remember, we learned in the linear system theory.
So if the system response is delta function,
if your camera has an impulse response as a delta function,
anything you take a picture,
you have a perfectly captured scene.
So nothing changed.
This is just an idealized system.
So if the impulse response is not a delta function,
that means H of T is not a delta function,
maybe scaled by a constant K.
So this is just say,
we consider a practical system is not idealized system.
Any camera will take a picture,
you cannot get all the details.
So the point response will be a blurred version.
You can never get a real delta response with a camera.
So we say if the system response is not delta,
so input delta, the output will be something more blurry.

It's not just the original delta.
So in this case, the Fourier domain expression
will not be a constant.
We know the delta function, you do Fourier transform,
and then you have a constant spectrum.
You just plug in delta into Fourier definition.
You will know delta function has a Fourier delta function,
delta of T has a Fourier spectrum of one,
all the way constant.
So if your impulse response is not delta,
the Fourier transform will not be constant,
not be constant K in this case.
Suppose we have an input I, I of T,
and it goes through the system.
In other words, we do convolution.
So the system response function convolved with input.
What's the output?
Output is still the same.
So you input I, and output is still I,
but scaled by a constant alpha.
So that means the shape of function is I of T,
scale somehow, but the shape is still the same.
So I say the shape invariability holds
for such a linear system.
Then in frequency domain,
it just put similar the first line
according to convolution theorem.
So this part really means alpha times capital I,
equal to capital H times the capital I, right?
So this convolution becomes a multiplication.
Let's see I of U, I of U are the same.
So that means, this part means H of U
must be equal to alpha, okay?
If you have a constant in the frequency domain,
in the time domain,
the function must be alpha times delta of T.
And we already decided that H is not supposed
to be a delta function.
So this is in contradiction to what we assumed.
So you know this is just what I claimed,
and you cannot have this general invariability.
Only two sinusoidal functions,
system will give you sinusoidal functions.
And the change is this weighting factor.
So each frequency component is multiplied
by a number at that corresponding frequency.
So convolution in time domain
is equivalent to multiplication in frequency domain.
This is a way to see it clearly.
Otherwise just a statement, a bunch of formula,
but this helps you understand what's going on, okay?
Some logic behind the theorem.
And also I mentioned this possible identity.
And I briefly mentioned you have energy conservation
meaning behind.
Also, if you consider F of T or F hat of S
as vectors in infinity many dimensional space.
And this integral and the right hand counterpart are nothing,
but they're all about the total length of the vector.
Because Fourier transformation by definition
is also normal transformation.
So the total length of the function,
the function total length of the vector
or normal of the function will be kept the same.

So this identity holds for any inner product space.
So this is just a geometrical explanation.
And for the energy conservation part
and encourage you anything you learn properly,
you may figure out some simple, very simple examples.
So you have a very simple example.
Sine of T equal to two sine hundred T.
So this is a sinusoidal wave.
And two can be amplitude of voltage.
So this sinusoidal wave goes through resistance.
What's the average power over interval minus capital T
divided by two and the two positive half capital T.
So this is average power.
The current squared times resistance.
Just say resistance is one.
So this is average power going through the unit register.
This is time domain computation.
And also you perform Fourier analysis.
So this can be decomposed
into these two frequency components.
You can do so using Euler formula.
Do this expression.
Now we claim the coefficient in the Fourier space.
This is C one, C minus one.
And you get a square together.
You just compute the components in the Fourier domain.
You add it together.
That's a total power.
Just according to individual harmonic components
of the wave form.
So you just add it together, got two.
You see frequency computation
and the time domain computation, they agree each other.
So it's just two different ways.
You can think of the X of T
has a different multiple sinusoidal components.
So add it together, it can be quite arbitrary form.
But this computation will remain.
You either do it in time domain.
You do this way.
What is our word?
The wave form of X of T
is you do time domain computation using this formula.
You get power consumed in the unit register.
Or you just perform Fourier analysis.
You get a bunch of Fourier components.
You may have complex components.
Frequency 100, frequency 99, frequency one million.
What's our word?
You add all those C coefficients squared together.
That's a total power according to frequency components.
And they should be the same.
Just review these a little bit
so you'll get a better understanding.
The delta function turned out to be very important.
And I mentioned if the system impels function
is delta function,
any input will be convolved with delta function.
What's the outcome?
Outcome is identically the same as the input.
Again, the camera example, ideal camera.
Every detail is captured.
Delta captured as delta, no blurry.
So this is the physical meaning.

But a mathematical formula is the bottom line
that says this delta function,
you can represent a continuous function as such.
So the continuous function,
really you cut it into small pieces.
These pieces is delta function
and at the corresponding location.
But at that location,
the amplitude of the magnitude of the delta function
is weighted by original function,
functional value at that point.
So tau equal to t.
At that point, you weight the delta function.
You basically say the function is really viewed
as a many, many small delta function.
They just collectively represent the original function.
So this is the discrete, the pixel perspective,
particle perspective of the continuous function.
The convolution theorem is something so important
and that we will utilize extensively.
So you need to be familiar with it.
This is another example.
So time domain, you have a functional form
and for simplicity, you have a gate function.
Or you can sync any function.
Then you have, say, two delta function here.
You do convolution.
And each delta function,
delta convolved with an input.
It's input itself,
just centered around the delta function.
So I gave you this example.
Say, when the delta function is shown here,
you do convolution.
You get the original function,
another copy of the convolution operation.
So this delta function is centered at the system origin.
So the function, the input function is reproduced.
No change.
So if the delta function is sifted a little bit,
just this delta function,
you do convolution of this sifted delta function
with original function.
So it's still original function,
but the convolved result will be centered
at the location of this left delta function.
You got this copy.
And then you have this right delta function
and you do convolution of this input
with this right delta function.
You got this copy.
So this is something you can easily verify.
And wherever you put the delta function,
single delta function,
you do convolution with the arbitrary continuous function.
You reproduce that function,
but just around the location of your sifted delta function.
And if you have many, many delta functions
and these delta functions are added together,
collectively doing convolution
with your input function here, f of xt,
you get many, many copies of f of x,
then just at corresponding locations
of the delta functions here.

So this is just a convolution property.
So convolution with delta is like a copying machine.
And you just keep copying input signals.
And you have many, many delta functions.
You have many, many copying machines.
So you copy whatever you input to this system
at locations where the copying machines are.
So this first line is just a time domain explanation.
Time domain explanation.
And according to convolution theorem,
okay, you see this is the gate function.
Gate function has a Fourier transform like this.
Okay, so single function.
Here is two delta functions at a symmetric location.
So each represent a complex component.
One is a negative complex index.
The other is the real complex,
the positive complex index.
So two things added together,
you get a real function which is a cosine function.
So this time domain,
pile of delta function has a Fourier transformation here.
Okay.
Then this convolution equivalent to multiplication
of these two Fourier spectrums.
So you do multiplication of this sinc function
with this sinusoidal function.
You have a modulated sinc function shown here.
So the bottom line represents
a frequency domain computation.
And then the theorem claims that
this time domain result and the frequency domain result,
they are pile of forward and inverse Fourier transform.
In other words, you do Fourier transformation
of this twin rectangular function,
you will get a modulated sinc function.
So so much is warming up for sampling theorem
and also kind of review what we learned.
So what we have learned so far,
Fourier series, Fourier theorem and Fourier transform
all talking about how can we represent
a continuous function using sinusoidal components.
The continuous function or piecewise continuous function
could be periodic, could be non-periodic,
but they are all, basically they are all continuous.
But now we are living in digital age.
So you just have a digital signals all the time.
So you understand very well, you need a computer
and the computer can only take
and output digital signals.
And graphically you see, you have computer
and to get a physical measurement into the system,
you need the first half of sensors,
all kinds of sensors.
The sensors give you basically analog signal,
like temperature, like dose, like distance,
like physical quantities.
Mostly many of them are continuous signals.
Then you need half a transducer.
You transform the physical signals
into electrical signals and they still kind of noisy,
noisy continuous signal.
And then you can do a little bit of conditioning.
You try to smooth it, do some local averaging,

you get a signal like this.
Then the analog to digital converter
really take this continuous signal into digital counterpart.
Digital counterpart can be coded in binary format,
zero and ones, and to see it better.
You see this is a continuous signal.
The digital computer cannot put a digital,
cannot put a continuous signal into computer memory.
So you have to represent the continuous signal
at a multiple sampling point.
So say at this point, what is a functional value?
And then at the next one, what's a value?
And the value itself cannot be continuous either.
You have to make it discretized.
So how many basic unit, how many minimum measurable unit?
So here you have three, here you got five,
maybe not exactly five, just you got some,
you got to do some approximation.
That is called quantization error.
You have a discretization along time axis.
Then at each sample point, and then you do quantization.
So you got just all the integer numbers.
So three, five, six, and so on.
All these integer numbers can be put in binary format.
The integer numbers, we are familiar,
we use decimal numbers, you have the 10 digit,
zero until nine.
And this 10 digit format can be converted
into binary format.
So you use zero, one, like one is just the one.
Zero is zero.
But the two becomes, so just the one, zero.
Then the three become one, one.
So you'll notice in the programming code.
So this is the step we need to convert continuous function
into discretized quantized function, the signal.
So we can save the data into computer.
So analog to digital conversion.
So the most important step is discretization.
So you do sampling, it's a point you've got these values,
okay, just so and here.
And also, I said, and this amplitude needed to go through
another step called quantization.
So you cannot just represent a pi, absolutely accurate.
You have to represent into some approximation,
do some approximation, say pi equal to 1.3.14 something,
you cannot go on for all time.
But for our purpose, we just think this quantization
is accurate enough.
And you can use binary code to precisely,
quite precisely represent the amplitude.
So that part is ignored in our analysis.
We just think the amplitude is represented
rather accurately.
But this discretization issue
must be explicitly handled in the next.
So this is the picture,
how you convert analog signal to digital signal.
And again, remember, two things in the conversion
from analog to digital.
One is discretization,
that's just the discrete sampling shown here.
The other is for each sample,
you need to do quantization.

But we ignore that part,
just think quantization is just so accurate.
We use very, very small unit,
like 10 to the power minus 20 something,
it's so small, just ignore that.
So we just need to deal with things like this.
So now we know the real thing,
the physical thing is this continuous function.
Now for your convenience,
you want to utilize computer.
Just for your convenience,
you do this sampling of discretization.
And the problem is,
and you do this approximation,
and if any information already lost in this process,
and after all, the continuous function
is what you need to do, to deal with.
But you just do second best,
you use a discrete version.
So any problem in this conversion,
how can we interpret the signal processing results
based on sampled data,
and then relay that back to continuous real world.
So that's a very important question.
You see, this is a continuous wave, okay?
This is the arbitrary function, it's a continuous.
And you can sample this wave with a small sampling interval
and at a frequency four hertz,
the signal has a frequency four hertz,
but your sampling rate is much higher.
So you got all these samples.
So heuristically you see, okay,
this sampled curve pretty much represented
the original curve.
No major information lost, as I expected.
Okay, it's just good.
This is a well sampled signal.
However, if you sample the signal,
it's not frequently enough.
Like in this case, you do equi-spatial sampling.
You got all these right points.
So the real thing is a blue oscillation,
but you didn't sample densely enough.
And the result of the knowledge of the blue curve,
and you just look at it, you sample the result,
you tend to interpret,
there is a low frequency sinusoidal wave.
That is what I'm sampling, but that is clearly wrong.
So this is an example for under-sampled situation.
So you have continuous signal.
You do sample, you want the,
this create a version,
fairly, faithfully represented the continuous case.
And you could do so very well
if you sample the original function densely enough.
Densely enough is a key word.
What do you mean densely enough?
And here is not densely enough
because you just imagine what is a curve.
You tend to think this is a result,
but this bump and all these small peaks will be missed.
So this small change may be very important
for the process of your interest.
So the missing information, in this case,

due to the sampling rate, the sampling frequency,
is not high enough.
So higher frequency components got lost.
So that's a problem, aliasing problem.
So in spatial domain, when you do sampling,
you are doing multiplication.
So you have a train of impulse function here.
You do multiplication.
You do the multiplication point-wise.
You got a sample of the signal.
This is a frequency domain picture.
And what is, this is a spatial domain picture.
So what is a frequency domain picture?
Frequency domain picture is like this, okay.
You have a train of delta function in spatial domain.
And in frequency domain,
you will also have a train of delta function
and a period in the frequency domain
and a period in the spatial or time domain.
So this period and this period, they are closely related.
So the two period, you time them together,
you got the number one.
So one is reciprocal of the other.
So in frequency domain, you do convolution, okay.
The convolution, you really just each delta component
in the frequency domain, you do convolution.
These components, you really get a copy
of this blue Fourier spectrum.
So you got a copy here.
And the central one, you got a central copy.
This happened to be the original one.
And this right one, you got a right copy.
The three copy together.
And this is a linear system, three copy,
really they try to accumulate together.
But because the Fourier spectrum spread so wide
with respect to the period of the delta train,
so then these spectral components,
they got mixed together.
So this is aliasing.
You got two spectrum mixed, shown here,
just like algebraically, you know x plus y
equal to certain number.
You have no way to tell what is x, what is y, okay.
So you just do not know how to interpret
this Fourier spectrum.
So if you just get the spectrum, this part out,
the original function will no longer be the same
as the truth, so they will have aliasing effect.
So this is just the idea in frequency of Fourier domain.
So one way to help with the problem,
and in the time or spatial domain,
you weight the function with something like a Gaussian,
so you do this multiplication.
So you got this input signal, look a little bit smoother.
And this way, so you will get a spectrum
not that widely spread.
So you just make the spectrum a little bit narrower.
Then you do the time domain sampling,
which is equivalent to frequency domain convolution.
Then you will have multiple copy.
So this is one copy, this is second copy,
this is central copy, this is first right copy,
this is first left copy.

Suddenly you have many copies,
it's just an infinity line, a train of delta functions.
So in this case, you can just think about low pass filter.
You just take this one out,
then you can recover the original function.
So ideal sampling filter, just here,
just put multiplication, just put a gate function,
cover this part.
You do multiplication, only this part left.
And the rest, just zero out, because gate function,
you have zero values beyond the support of the gate function.
So this is ideal thing to do in Fourier space.
In Fourier space, you do multiplication,
that is equivalent to you do convolution
in time or spatial domain.
And the gate function has a gate function in Fourier domain,
has a Fourier, inverse Fourier transform,
which is a sinker function.
So we learned before.
So sinker function in spatial or temporal domain
has an infinity line ranging, so keep going on.
So it's very long, not so convenient.
So if you use, in frequency domain,
you use this sinker function in time domain,
it's very easy.
The time domain, you just got this rectangular function.
So that means in time domain, you do from a discrete point,
you do local averaging, you do just moving average,
so you can just recover the continuous function.
So this is easy, but this is not ideal,
because the ideal filter is a gate function.
So compromise will be Gaussian function.
So you do Gaussian multiplication
in Fourier or frequency space,
then in time or spatial domain,
temporal, spatial or temporal domain,
and then you do convolution with Gaussian.
So Gaussian function has a good form,
and just something, good compromise.
You know the ideal filter is this gate filter.
So you just take this out.
And then you use Gaussian as an ideal filter,
as a filter, it's not ideal, but that's something easier.
So in both domain, you kind of have finite support.
The Gaussian function, the functional value goes,
decays exponentially fast, so very quickly,
you do not have the infinity ranging.
So that's just a heuristic explanation what's going on.
So if you haven't reviewed this before,
so this gave you some idea,
but it may not be crystal clear.
And the next part of the lecture,
I will get more quantitative,
so kind of repeat what I show you.
Those pictures are kind of cartoon,
but from the next part,
and you will see exactly the formula.
I will give you more rigorous derivation,
but I'll let you have 10 minutes, then come back, okay?
Now we are ready to continue.
So in the first part, I gave you some cartoon-like pictures.
Now we are getting more mathematical.
So the delta function is so important
for digital signal processing.

And we particularly interested in a series
of delta functions, or training of impulses,
or sometimes, we just call it a comb,
because it's just like a comb you use
to just sort out your hires.
So this is a fundamental relationship.
Fourier series, Fourier transformation,
and also the training of impulses
in different textbook, they all represent differently.
I kind of like frequency domain symbol.
I use u as a frequency.
Here, they use f, and for delta function,
training of delta function, they use lowercase s of t,
so there's just many delta functions.
They are evenly spread out,
with the period equal to t, so here.
And if you perform Fourier transformation,
and then you also got similar thing,
it's just another train of delta function,
but the period is one over t, here is t is one over t.
So it's same thing, but their periods
are related reciprocally, and not only that,
and also there is a scaling factor.
This is easy to remember.
They see n over t, and it's one over t,
so this is just to show you
this very important relationship.
And to get a Fourier transformation
from delta function to Fourier x present is not hard.
So if you insert a delta function
into definition of Fourier transform,
you immediately see the delta function
is easy to integrate, and according to the definition
of the generalized function.
So the delta of t has a Fourier spectrum one,
but if you have delta of t minus N capital T,
this is something you need the Fourier safety
or translation theorem, you can compute.
And then you can go from functional space
to Fourier domain, you can go the other way around,
because the Fourier transform and the inverse transform,
they look very similar.
The kernel, all e to the power two pi N U T,
but one is a minus sign, the other is a positive sign.
Minus and positive doesn't make a whole lot of difference.
So if delta is in frequency domain,
you perform inverse Fourier transform,
you can show this corresponds to e to the power
j two pi f zero, this f zero is safe.
Okay, t, so some people use i, and some use j,
and i use i, here you use j,
but basically this is a relationship.
And if you want to underline,
the sampling is done in frequency domain
with the step delta t, because I have delta t
small increment, you just sample signal.
So you can put delta t here,
so this capital T becomes delta t.
And in the frequency domain,
and the period will be one over delta t.
And this is getting, fact is not one over t,
is one over delta t.
So the two versions kind of get you familiar with this.
I want to underline, this is a Fourier pile.

So this is a train of impulse or delta functions,
and that corresponds to another train of delta functions.
So the cone, and that is a mirror in the Fourier space,
is a very important tool for our analysis.
How do you understand this?
Certainly this is a way, just a wave of hand,
you can derive this way.
And in the textbook, and I first gave a relatively long
and hard to, difficult to understand argument,
but more rigorous.
I also gave some easy way, and roughly or loosely,
we can just say something on this page.
I call it much simpler, but in precise way to read.
So this unit period in time domain,
a train of delta function.
The Fourier transform is also a train of delta function
with unit period.
So you have same thing, both hand side.
And if you have period is not a unit,
it's general capital T, then you have one over t.
Here is one over t, like what I showed you
in the previous slides.
Okay, so you have this.
How you prove this loosely?
So you see, this is a delta function.
This is a train of delta functions.
So this is a periodic function.
The period is capital T, indicated by this capital T.
A periodic function can be expanded as a Fourier series.
And the first or second slides,
at the beginning of this lecture,
we know this periodic function can be expanded this way.
And the cn is computed through the averaging over capital T.
Remember, I told you that on second slide, but anyway.
So you can do this integral to find coefficient.
And this s sub t is delta function.
You put a delta function,
and the delta function basically only take a value
at system origin t equal to zero.
So this whole integral becomes one.
So the value is one over t.
The cn is one over t.
You put this back.
So you have this expression shown like this.
So this can be viewed as a frequency representation.
And the period here is one over t,
because you have t in the bottom part.
And the amplitude is weighted by one over t, you got this.
So this is nothing but this claim.
So whenever you have a train of impulse function,
this period capital T, or what so ever you call it,
P in one domain.
And in the other domain, you will have counterpart.
And the period will be reciprocal
of the period in the original domain.
So this is very important to remember.
And also don't forget this weighting factor.
So this is derived this way.
And if you are mathematically curious enough,
you read other pages I have.
And just give your understanding.
And such correspondence makes sense in some fancy space
and in the sense of inner product in terms of measurement.
So it's kind of complicated.

But an easy way is to lose in the imprecise way.
Just remember you have this tool.
But Regulus theory is beyond the scope of this.
Lecture.
And again, you see, if we think in the time domain,
you have the signal as sub capital T,
then lowercase t is this one.
In the Fourier domain, you have one over t.
So if in the Fourier domain,
you think this period is capital T,
then in the time domain,
the sampling distance is a small delta t.
So just this counterpart relationship.
And with such a tool,
now we can explain to you
and what is the sampling problem
and what is the sampling theorem.
If you understand the sampling theorem
after this lecture,
that will be a great achievement.
So first look at the first column.
So this is a continuous function.
And our goal or motivation
is to convert this into digital signal
because we need to sample the signal
and put it into computer.
So what we will do is just at regular time intervals.
And then you just do the measurement.
And then the sampling point is evenly distributed
because beforehand, you don't know,
you really don't know where you should sample
more densely, where you should sample more sparsely.
And without prior knowledge,
it's just to do even distribution.
So I sample the signal using the right,
you're using the train of right impulses.
So you got this one.
This is a multiplication process.
This is the yellow multiplication.
You do this way.
This is a mathematical model.
What you are doing,
how you deal with continuous signal.
You just take a value like your thermometer.
You keep taking value every hour.
Then you got a discretized sampled data shown here.
So this is what you really do as an engineer.
You use some sensors, transducers,
to interface with continuous analog physical process.
As a result, you put a computer,
all these discretized signals or digital signals.
So you want to precise the signals for different purposes.
But this is very heuristic.
Now we have the Fourier analysis as a powerful tool.
Now let's see what's happened,
what was happening when we do sampling in the first column.
So second column, now look at the second column.
So this continuous function
would have a continuous Fourier spectrum.
Okay, the Fourier spectrum has a significant spread.
So from minus W to positive capital W.
So this is a range over which
the Fourier components are significant.
So you cannot ignore anything.

But I want to say, if time domain signal
is of finite support, finite support I mean,
over this finite interval, all values are zero, okay?
So if you have a finitely supported time domain signal,
and its Fourier spectrum must have infinite support.
That means non-zero values will not be restricted
within a finite interval.
It'll keep going, all spread out.
But the nice thing is that the Fourier spectrum will,
the amplitude will decay quickly,
particularly when this function is smooth,
smooth in first order derivative or second order derivative.
So the Fourier spectrum will go, will decay rather quickly.
So beyond a certain point, in this case,
beyond W, the value just is so small,
you can just safely ignore just the simple explanation.
So you have this dual pair, both finitely supported,
meaning that this function is only,
time function is only significant over this interval.
And the spectral components are only significant
over minus W, all the way to capital W.
And beyond that, just think, no problem, okay?
And you do multiplication between this time function
and these right pulse sequences, right pulse sequence,
many delta functions.
So this right train of impulses,
you do Fourier transform, as I explained,
you have another train of delta functions,
also right, all the right impulses.
And the period is, this is very dense in the time domain.
But in the Fourier domain,
these delta functions are widely spread.
And the period is one over delta t,
the smaller delta t, the larger capital P.
So you've got all these copying machines
in the Fourier domain.
So the multiplication here corresponds
to convolution in the Fourier domain.
So each copying machine, I mean, each delta function,
will bring a copy around the machine.
So the central one, we got a central copy.
This one got this copy.
So you have infinity many delta functions in time domain.
And then you will correspondingly
have infinity many delta functions in frequency domain.
And the adjacent ones are P distance apart.
So you got all these multiple copies
of same Fourier spectrum.
So this is what's happening graphically.
So now the question you may curious enough,
you want to know how densely we should sample the signal.
So we can recover the original signal
from its discrete copy, discrete version.
From this, I want to recover the continuous function
perfectly.
So if I can do that, then I know,
although for use of computer,
I converted the continuous function
into discrete digital signals,
nothing has been lost.
So that's good to know.
So I think just graphically,
if you just make the sample,
the sample, the data densely enough,

delta T small enough, then P is larger enough,
you have many copies here.
But these copies do not mess up.
So each copy from any copy,
you can recover original signal.
So in this case, if you perform a Fourier analysis
of this discrete type of signal,
you got all these things.
Then we can purposely keep the frequency components
between minus W and the capital W.
Then from this Fourier component,
we can recover the original function
through inverse Fourier transform.
So this is the idea.
So how densely is dense enough?
So recall, we have a Fourier analysis.
So this is the formula.
And the Fourier series expansion
and the formulas for coefficient, okay?
We know the components.
If we just take a finite term,
the highest frequency is capital N times two pi,
two pi capital N.
That's the highest frequency.
Lowest frequency is one N equal to one.
So lowest frequency is two pi.
It's just two pi.
So the period is one.
So this function over a unit interval
will be able to represent a periodic function.
And if you do normalization
and extend the period from one to capital T,
and then you can really periodically
express function with period capital P.
So this is just a Fourier expression capability.
Then you say, how many samples I should do
to get the original function completely recovered?
So first, let's just see, you have original function,
and then you can put a bunch of samples there.
The first question I want to ask you,
how can you estimate the DC?
The DC is computed, this is our region operation
for continuous function over the period, say unit period.
And when you just discretize the function
over the interval zero to one,
you get many, many samples.
So one way heuristically you think,
you can estimate the direct DC component.
Here you do continuous function,
you do integral for continuous function.
Then once you sample the data,
you add all the samples together,
then you do average.
So this is DC component,
this is just a constant component,
can be easily recovered that way, okay?
Then we don't worry about DC,
you just sample the original signal,
you do average, you can get DC fixed,
the constant term fixed.
Then you have, say, you have two capital N terms
for cosine and sine components, okay?
So how many unknowns you have?
You have two times capital N unknowns, okay?

So each sample, at a time point, say T1,
T1, you do measurement, this is one sample point.
This is a linear equation.
So you have one equation, it contains two N unknowns.
Two N unknowns, so heuristically you think
you would need at least two N measurement,
so you can uniquely solve A and BN, right?
So the number of sample for this periodic function
over zero and one should be at least two capital N.
That means this periodic function over zero and one,
over the interval zero one,
you need to just put two N or more sampling point.
Then you set up the linear equation,
you can solve for A and BN, all together,
A and BN come together, you have two N point.
And I told you, you don't need to worry about constant term.
So this is one observation.
And also sine and cosine term, you say,
you have sine term, cosine term, this is A and this is BN.
Sine and cosine term can be combined
into a single sine term.
So if you like, you click this,
you will see how to do this conversion.
Sum of sine and cosine becomes a single sine function.
But in this case, the amplitude R becomes unknown.
So you have two unknown, A and B.
Here you have one unknown R,
but don't forget you have another unknown,
which is a phase factor, this is alpha.
So you still have two unknown.
So heuristic analysis is showing like this, okay?
You have this summation, sine component and cosine component
added together, N terms, you have two N unknowns,
amplitude and phase, just based on this formula.
And all you just based on the original formula,
you have two unknown, A and BN,
all together you have two N, two capital N unknown.
So for the highest frequency, so for each cycle,
you want to just get two samples out of that cycle.
So you can fix the two unknowns, amplitude and phase.
And then you have all together capital N components,
you need, I told you you need at least two,
capital N samples.
So sample sampling frequency wise,
so over one period, one period,
you need to get two over capital N,
or here you call the V, two V.
So that's just the sampling frequency.
Next is the sampling frequency.
So the double the highest frequency.
So highest frequency in the original approximation,
what's the highest frequency?
It's two pi capital N is the highest frequency.
What's your sampling frequency?
Your sampling frequency should be four pi capital N.
So highest frequency is two pi capital N.
And I say sampling frequency should be greater
than twice of that.
So we've got a four pi capital N.
And the argument is to solve,
to have an argument is to have enough equations
to solve for all the unknowns.
So this is a heuristic analysis.
Okay, a little bit more mathematical,

but still not a regular enough.
Then this slice is really the most important slice.
I hope you can follow this slice.
And what's in the draft book has typos.
So I fixed them, but I haven't uploaded it yet.
So for this part, please follow this slice
and follow my explanation in the rest of the lecture.
So do not be bothered by the typo in my book draft.
Anyway, so this is a derivation of sampling theorem.
But just make sure, once you have discretized samples,
how you can recover the continuous function
from the sample data.
Assume the sampling rate is equal or beyond
so-called negative sampling rate.
This sampling rate is satisfied,
so no aliasing artifacts.
So now let me explain to you.
So if you just look at these derivations step by step,
a chance is that you just follow them.
Then you see, okay, from discretized samples,
all these sample points,
and then you do convolution with a single kernel,
you just recover the original function.
This is the result, and this is the y.
And now let me explain each of the steps
and why we have that.
So just follow with me.
This becomes more mathematical now.
So first let me explain why we say f of t
equal to inverse Fourier transform of all these things.
So let me explain to this underlined right-hand side,
this way.
So remember you see this.
When the sampling rate is high enough,
delta t is small enough,
then the capital P is large enough,
you have multiple copies of original Fourier spectrum,
multiple copies.
But each of them is sufficiently far away from neighbors.
So you don't have aliasing problem.
So they all spread out.
So this is Fourier spectrum of original function.
And then you do convolution with this train
of delta function in Fourier domain.
So the period is capital P.
So this is capital P.
So each copy is separated from neighboring copy
by this period.
Let's say from here to here is capital P.
So you do convolution.
So this part really give you all these things, okay?
Then to recover the original function,
what do you need to do?
If you understand this part, you have many copies.
This thing, okay, I put a gate function here.
So gate function multiply with this train
of Fourier spectrum.
Only the central one is kept intact.
Other copies will be zero out.
So that is great.
So this is the gate function, okay?
Only keep the central copy.
And the rest things, other copies, all zero out.
So the central copy is saved.

Then you perform inverse Fourier transform.
So you keep the central copy.
This is the central copy.
You perform inverse Fourier transform.
You get this back, okay?
This is, this line says, okay?
You follow me, so now you understand this part, great.
If you understand this part, now I will rely
on convolution theorem to explain the next two steps.
So you see here, this is a Fourier transform, okay?
And this is a Fourier transform
convolved with another Fourier transform.
So after all, this is a Fourier transform
because the variable is a frequency variable.
So this is a Fourier transform.
So this Fourier transform times another Fourier spectrum
according to convolution theorem.
You see, you have a multiplication here.
Then you have convolution in the time domain.
So this is a Fourier transform.
And the time domain thing is this one
subject to Fourier, inverse Fourier transform.
This time domain thing.
And the convolved, this whole thing is a spectrum.
And this whole thing is subject to
inverse Fourier transform is a time domain thing.
So multiplication in Fourier domain
is equivalent to convolution in time domain.
So this f of t is equal to this time signal
convolved with this time signal
by the convolution theorem, okay?
The second line, okay?
And keep going, keep going.
So the first part just copied onto the third line.
And this part you see the convolution, okay?
Is a convolution in frequency domain.
Is a convolution in frequency domain
is a multiplication in time domain.
So this Fourier spectrum will do
inverse Fourier transform.
Is a time domain, time domain function.
Likewise, you do inverse Fourier transform
is a time domain function.
So the convolution becomes multiplication here.
Got multiplication, okay?
Now we got this line, the third line.
So we use convolution theorem.
So convolution theorem is not limited to one domain.
You do convolution in time domain.
Then you have multiplication in frequency domain.
If you do convolution in frequency domain,
you'll have multiplication in time domain.
It can be proved similarly.
So we use the convolution theorem twice.
So we are now at the third line, okay?
Now we are ready to go back to the time domain.
Let me say, see you have convolution here.
So first we bring the first part to time domain.
So frequency domain, that's a gate function.
Gate function is a period P.
Why we have a period P?
Because we want to keep the central copy.
We single out the central copy.
We argued before.

But this gate function in frequency domain
with this period capital P in time domain
is a single function.
It's not a simple single function.
It's a weighted version of single function.
The single function pi capital PT over pi capital PT.
This whole thing is really single PT.
And I cannot type single PT.
Any of you know how to use MATLAB to type single
in MATLAB editor?
Let me know.
I couldn't make it.
So anyway, so just got this single part.
But this single function is weighted by capital P.
So this underline the part here.
Why you have this one?
The pi mean this gate function has a time domain.
Now you can just use a definition
of inverse Fourier transform.
You perform inverse transform
upon rectangular sub capital P.
You will get this one.
I gave you even better example.
It's a 2D rectangular function.
Rectangular function along x axis
with side total lines or side lines capital X.
And along y direction, you have a total side lines y.
Then you can perform the Fourier transformation.
So you just see the x part, the y part.
You got this x factor out.
And in the main part, you have the single function.
This is time domain computation.
Likewise, you can perform if the gate function,
this rectangular function, is not in spatial domain.
Rather, it's in UV, in frequency domain.
It's similar thing.
So just to show you, you can do similar thing
to show in one dimensional case.
So this gate function along your axis,
you can get time domain sinker.
So just example show you how to do it.
This is not exactly same.
This is 2D example first place.
Second, the gate function, the rectangular function,
is in time domain.
But what I want to explain is gate function
in frequency domain, you perform inverse Fourier transform.
Here, you perform forward Fourier transform.
But they are very similar.
Similar enough, you can figure out.
So I explained why you got this part.
Now we go line by line.
We are here.
And now we say, look at the second part.
So this is Fourier transformation of original f of t.
Then you perform inverse Fourier transform.
You did the first forward transform.
You do inverse transform.
That's easy.
You just bring back the function f of t.
You perform Fourier transform.
You perform inverse transform.
You just return original thing.
You just have something.

You rotate, say, 45 degrees this way.
Then you rotate it back.
You just original function.
You got this part.
This is easy.
From this part to here is easy.
And here, you have a frequency domain train
of impulse functions with period of capital P.
Inverse transform, we say, is in time domain,
period becomes 1 over capital P, the reciprocal relationship.
The frequency component, u, becomes t.
And don't forget this weighting factor, 1 over p.
So from here to here, it's just we utilize what I explained
to you before, comb and it's a mirror in Fourier space.
Remember here, I showed you this part.
You see you have t.
You have 1 over t, and here, 1 over t.
Now I call 1 over t is p.
OK, this 1 over t becomes p.
Then the t becomes 1 over p.
This p can be moved to this part.
This line is exactly what I just utilized
on the previous slides.
You see in frequency domain, you have
the train of delta functions with period of capital P.
And you perform inverse Fourier transform.
And in time domain, you got a train of delta functions
with period of 1 over capital P. And this whole train
is weighted by 1 over p.
See here, 1 over p is 1 over p.
So it's just I explained that part.
You see from here to here, what's the difference?
So this just put the original definition, this sigma.
So many, many, many, by definition,
just a train of delta function with period of 1 over p.
But where is this 1 over p?
It's missing here, right?
This is copied here.
Just think, this sigma, all these delta functions,
is I of t times I of t.
That's OK here.
But where is 1 over p?
This 1 over p is canceled out with this additional p here,
p, 1 over p.
So just cancel out.
So this is missing.
You only got this single function here.
So now you have this single function.
Then you have this I of t continuous function.
And then multiplied with all these delta functions.
So this multiplication is the sampling process.
And then you show it before, this sampling process.
This is a continuous function.
You have a train of delta function.
And then you do multiplication, you sample, you got this one.
So now I explained this part.
So this is your sample version of continuous time function.
This sample version.
This sample version is convolved
with this synchro function.
The synchro function is something
like sinusoidal oscillation.
But this gradually reduced amplitude.
So you do the convolution.

So these discrete values convolve.
These discrete values convolve with this kernel.
This sinusoidal kernel, you do the convolution
with all these delta functions.
So you basically, each delta function, as I told you,
each delta function behaves like a copying machine.
Each delta function convolved with this synchro function.
You got a copy of synchro function.
And then you have many, many digital samples.
You have many, many copied, scaled synchro functions.
Because this is not only delta function.
It has a weighting factor of f of t.
Really, f of t you sampled at a location k over capital P.
The k is a dummy index.
You can call it a k, call it an n.
Doesn't matter.
So you basically think you have a bunch of discrete.
Let me just think of this.
It may be better.
You have these sampled signals.
And along these sampled signals, you
make a copy of the sinusoidal, another sinusoidal,
the synchro functions.
So each of them, really, you've got a synchro function.
All these weighted synchro functions are added together.
So this discrete signal.
But each discrete signal is converted to a synchro profile.
And all these synchro profiles, just linearly added together,
magically, you will have this original,
original continuous function, f of t, recovered.
So you see there's a copying of our convolution
with delta function.
And you showed this before.
So it's just reproduced, this slide reproduced here.
You have functional form.
You have a delta function.
You do convolution.
You just reproduce here.
Do convolution with another delta function,
you have the copy here.
But now, for the sampling theorem,
and the functional form is this synchro function.
So you have multiple synchro functions.
It's not just two, multiple of them together.
And also, all these delta functions
are not equal amplitude.
So the delta function amplitude is weighted
by f of k over capital P. That's a digital value
at the sampling point.
So not all these amplitudes are equal.
So you have many, many synchro functions
with different weighting factor.
All of them added together.
You have this continuous function recovered.
This is the main conclusion of a sampling theorem.
From discretized sample data through this formula,
you really go back to the original continuous function.
So just go step by step.
You see how we make the magic happen.
So just one slide, this slide, you go through.
You got a good understanding.
And then you can safely sample a continuous signal.
As long as the sampling rate is good,
you can always recover the original signal.

And under something of band-limitedness,
you go double the maximum bandwidth, maximum frequency.
So this is a piece of classic signal processing theory.
And then the latest theory, it just
go beyond the necklace of the sampling rate.
You don't need a sample that densely.
And you can still recover the signal very well.
Just a few words.
You see the green button.
I do not expect that you really understand thoroughly.
So instead of assuming the signal is band-limited,
and you do sampling, double the maximum bandwidth,
just the necklace of the rate.
Go beyond the necklace of the rate.
And instead of assuming that model,
you assume the signal is sparse in a way
like you perform some transformation,
like a wavelet transformation.
Or you just try to compute a signal difference.
You add all the absolute value together.
Basically, you have a measure in some domain, some space.
You want the signal to be very sparse, like images.
You perform a wavelet transform.
Then the effective coefficients are not many.
So you want to recover a signal among all possible solutions.
If you don't have enough sample to determine the signal,
you will have many, many solutions.
All of them can explain your measured data.
And among all these candidates, you
find those who give you sparsity,
means the function is smooth, or the wavelet transform
has a very compact support.
So you have some other kinds of assumptions.
Then you can also do a very good job for signal recovery
from discrete sparse sample to recover continuous signal.
So this is a big picture.
And what we explained, pretty much this part.
So through this part, you know continuous function
can be sampled, turned into discretized version.
From discretized signal, we can use
negative sampling rate, or higher rate.
And just use the syndrome sampling kernel
according to sampling theorem.
And from the discretized signal, we
can recover continuous signal.
So we can discretize signal, and then we can go back.
So this is nice to know.
And this is only part of the story.
And we want to do convolution, Fourier analysis,
filtering, not only in time domain.
And the filtering oftentimes should
be done in frequency domain.
In the last lecture, I showed you
how to use the Fourier transform to do denoising,
edge detection, and so on.
So you want to work in frequency domain.
But now, the frequency domain thing,
you have a continuous spectrum.
And you do sampling.
You make the problem even more substantial.
And you have a multiple copy of a continuous spectrum.
So you have all continuous thing.
And I told you digital computer have
to take discretized signals.

So we need to discretize Fourier spectrum.
How should we x-price approximate a continuous
spectrum into digital signals?
So we need to sample Fourier spectrum
and put it into computer.
In other words, if you want to perform good digital analysis
in time or spatial domain, temporal, spatial temporal
and also frequency domain, you need
to have discretized signal in both spaces.
So you need to have something like this.
So this part will be covered on Friday.
So please review.
Depending on this week, I'm rather busy.
Several NIH review meetings.
So I couldn't update my book chapters.
And I hope to update a little later.
But I suggest you review the signal processing lecture
based on this slide and the derivations step by step.
So do not be confused by typo in the book chapter.
But you can review the next lecture
called discrete Fourier transform according
to the book chapter if you like.
And this is your homework.
So so much for today.
We have an office hour from 3 o'clock.
And if you couldn't make me, and you have any suggestions
about the pace, the material, which part you feel challenged,
which part you like, you don't like, you can also email me.
So just interact.
And the next lecture pretty much finish Fourier analysis.
OK.