Okay, let's start.
So I chatted with some of you and it seems so far so good.
The best format, you read the chapter and if you like you can watch the video we recorded last year and at least you read the chapter.
Then you attend the class, they will do homework and review again and that will give you a very good understanding, I hope.
And today we talk about Fourier series and it's slightly harder than convolution.
So step forward.
Very important and I try to explain the key ideas.
Not like some class material or textbook that gives you a formula, then you do Fourier analysis and then you can still do computation but you really don't know what's going on.
And what I hope to offer in this class, so we start from higher dimensional space and thinking about geometry and this higher dimensional space, angulation, distance, inner product and so on, putting together will give you a very good understanding.
So called Fourier series, Fourier transform are nothing but some transform and then you can have a clear geometrical picture.
Whenever you have geometrical picture, you can visualize things.
So you have a good understanding, you have a good memory and you know what you are doing.
Otherwise just some program lines, you just keep running, you really don't know what's going on.
That's not good.
So you really need to know what's happening.
And the last lecture we mentioned safety environment linear system.
We discussed about convolution in continuous and discrete cases.
And then we also mentioned some equality called Cauchy-Schwarz equality.
A little confused perhaps to some of you.
And I mentioned inner product and motivated by convolution and cross correlation and those things you see a common format like here.
So you have two vectors A and B, a concept called inner product.
You basically have two vectors and then you pair the corresponding elements together and then each pair of the corresponding element are multiplied together, you get a partial product.
You match many things together, you get many parts of product.
Like the example, hands-on example, you have a vector A, vector B, you do this pairwise matching.
Do all the partial product added together.
That is called inner product.
When you do cross correlation or convolution, you need to do inner product multiple times and each time give you a data point.
And in the radar signal processing, when the matching filter and the underlying signal like an echo from an airplane, they match together.
The cross correlation gives you maximum output.
So you can use that to do signal detection.
And the linear system perspective tells us you can do correlation and the correlation really will flip over, you do this matching a bunch of inner products.
That will give you response of safety environment linear system.
That's very important.
So this is something I want to look deeper in this class to prepare you for Fourier series.
The inner product has multiple nice properties and the most important one is this one.
So you have two vectors and each vector has a total length.
Okay.
Now you have two vectors and the inner product is nothing is the length of vector A times the length of vector B, then weighted by angle between these two vectors.
So that's called the cosine theta.
So the claim can be shown to be true in two dimensional space, three dimensional space, and this is very important.

If the ambulation is such that they perpendicular to each other, then this cosine theta will give you the value zero.
So your inner product has a minimum value.
Or you can have a maximum value when cosine theta is equal to one.
So you got this case.
And in the special case, A and B are the same, so you got A times B, the length of vector A times the length of vector B.
In this case, the vector B is A.
So you got the length squared and you do square root operation.
So you got a total length of A.
So this inner product can give you distance, total length, and also angulation.
So that is the essential feature of a two, three, or n-dimensional space.
The claim, and I kind of hinted you last lecture, so this pairing, partial product summation is something similar.
So the analogy of biological analogy would be DNA sequence.
You're matching up together, so this summation overall effect is very important.
So let me just review and also kind of summarize why we need inner product.
There are really multiple reasons.
So one is from the perspective of safety environment linear system.
And remember this slide, you go step by step.
Whenever you are given a safety environment linear system, and also you are given impulse response.
So once you know impulse response and you know it's a safety environment linear system, then you are OK.
Any arbitrary input, and you can do convolution.
You can find the system output.
So the process is shown here.
So looking at the convolution, you see this slide, you need to flip and save.
But whenever you finish flipping and saving, to find the particular value of the convolution, it's an inner product matching together.
OK, so similar argument I mentioned, the radar signal detects it.
And that is to define cross correlation and just do the matching filtering.
And again, the fundamental operation is inner product.
And this can be viewed from a geometrical point of view.
So this part is from the book chapter I drafted.
And then you can read it.
But let me explain to you.
So suppose you have a vector A with an element ai, lowercase ai.
Maybe you just think in two-dimensional space to be easier.
OK, you have this vector capital A.
So vector capital A is a vector.
But the paper of the vector is a point in the space.
OK, just think two-dimensional thing.
So vector A is a point.
Let's think about another vector, vector B.
This is arbitrarily another point.
Suppose this vector B is not zero.
OK, then this non-zero vector B and the zero vector is the system origin.
So these two points will determine a line.
OK, so this line can be described as this.
OK, let me get a pointer.
OK, this line can be described by this expression.
So basically, the vector B has multiple components.
And in two-dimensional case, you have B1, B2.
Then you just scale each component with equal proportion signified by the constant k.
k can change any real value.
So keep changing k so you really can trace any point along the line specified by system origin, zero vector, and the vector B.
The vector B is also a single point.
OK, the paper of vector B.
So you keep changing k.
You got just arbitrary.
So this is a vector B.

This is arbitrary line.
So this is system origin.
OK, system origin.
So this is a vector B.
You keep changing k so you can go anywhere.
Then we also have vector A here.
OK, the question is, what's the minimum distance from this vector or point A to this line? OK, if you find this minimum distance, this is a vector A, this is a vector B.
Then this total distance D can be geometrically interpreted as a projection of vector A onto this line, determined by vector B and the system origin O, or zero vector.
So that's a geometrical idea.
And then we can compute.
We can compute this minimum distance.
Equivalently, we can compute this projection, vector A projected onto the line determined by vector B.
So we just formulate what's the distance of this point A to arbitrary point on this line specified by vector B.
So the squared distance is nothing but component A minus arbitrary components along the B-I axis.
So this is the kBi, because this is a general formula.
Arbitrary components corresponding to the index i is kBi.
k can be any real number.
So the component-wise difference or subtraction is squared.
You add all these things together.
It's nothing just compute Euclidean distance.
OK, remember high school formula, x1 minus x2 squared plus y1 minus y2 squared.
Then square root.
That gives you distance from point A, x1 by 1, to point B, x2 by 2.
So this formula is still good.
We are talking about two-dimensional space.
The same idea applies to three-dimensional space.
And this distance computation, element-wise, squared, then summed together, square root.
And we can safely extend it to n-dimensional space.
So this is the objective function.
You want the distance.
You want to have this distance smallest.
Then you just calculate the first derivative of this objective function.
You set it to zero.
You solve for a particular k.
This particular A will specify this point.
And this point A to this particular point on the line specified by vector B will give you minimum distance.
Once you know this minimum distance corresponding to k, this is k, then you know this k.
So this total distance B can be computed.
So this is from system origin O.
So this is y.
This is two-dimensional case.
Same thing for higher-dimensional case.
So this is k.
You can compute.
Once you know k, the distance from origin to this projected point will be the coefficient k times the total length of this B vector.
So you got this one.
So you just got this x present.
So this is to say you can find this projection.
The projection is nothing but this result, inner product divided by the length of vector B.
So this is how inner product is related to projection.
So this is something we need to know, the geometrical picture.
So once you know that, you can understand a lot of things easier.

The earlier slides show inner product equal to the length of vector A times the length of vector B weighted by cosine theta.
Here it says this projected distance, you can think projected distance is total length and weighted by this cosine angle, this angle, either in two-dimensional space or higher-dimensional space.
You can define the angle in this meaningful way.
So whenever you have this line, you have this vector A, and you think there is an angle, and cosine this angle times this total length A, so total length A cosine angle, so that will be this total distance.
So that's just geometrical meaning.
So the inner product is nothing geometrically, but the length of vector A times the length of vector B weighted by cosine theta.
So in other words, you think this is the definition of cosine theta in higher-dimensional space.
So look at the Cauchy-Schwarz equality, and I'm not sure any of you didn't read the proof, but now you should read the proof.
So just algebraically, we can prove this relationship.
So this is inner product, but it's just squared.
And on the right-hand side of inequality, so you have the length squared for vector A, and vector B's length is squared.
So if you do square root on both sides, so on the left-hand side you have inner product.
So inner product is always less than or equal to product of the length of vector A and the length of vector B.
So this is geometrical meaning.
So this is the equality proved algebraically, and we can prove it in an easy way, so geometrically.
So you think about this.
Again, the same inequality.
So inner product is nothing but the length of vector A times the length of vector B and the times cosine theta.
The cosine theta normally is less than 1, so the right-hand side is always generally greater.
But when the angle theta is 0, then it takes equal side, so just a geometrical picture.
So when A and B has an angle between them, 0, then in that case A and B will be similar.
In what sense I say similar? Because in this case we can say when it will take equality.
So when the components BI divided by AI is constant, so component-wise they all have the same ratio, so that means one vector is scaled worse than the other vector.
They look similar.
So I sign the radar signal and let it go and hit the airplane and come back.
And depending on distance and the material of the airplane, so my signal will be attenuated, will be delayed.
So the amplitude when the echo back will be different, but shape wise is the same.
So I will just try to do cross correlation.
So if the signal, my expected echo signal is convolved with the whole noisy background signal,
whenever I see a maximum response and I think, okay, I would infer there is a plane, that's my echo.
So that's the idea, matching filtering.
So the equality will take equal sign when the two involved vectors are different only by a scaling vector.
So that's just the idea, simply explained in two-dimensional space.
And you need to use your imagination and just extend the argument, the explanation from 2D to 3D.
That's not too hard, and if you like you can geometrically show the same thing.
And the n-dimensional space, you need to do a similar thing.
And you just have two vectors, V and W, capital V with the components Vi, i may be equal to 1, 2, 3, depending on over a finite range.
Likewise for vector W, capital W with the components, lowercase wi, and the

tapes of these two vectors are specified by all these VW components.
The distance between them, as I argued earlier, can be computed.
It can be computed element-wise, you find the difference squared, you add them together.
This is how you compute distance.
And again, in two-dimensional case, this is nothing but a piezo-green theorem.
You extend it to three-dimensional space, so you think you have a cube, and you have A, B, C, so what is the longest diagonal will be A squared plus B squared plus C squared square root.
And what is the distance in our n-dimensional space? This is the distance.
So the V vector can be projected onto W vector, or W vector can be projected onto V vector, the line determined by V vector.
So this is the argument that runs through.
You minimize the distance, you find the point, then you can find the distance from that projected point to the system origin.
So we talked about this.
Then using the inner product, so when you solve this, minimize this objective function.
Inner product, the mathematical formula, the definition for inner product will come out naturally, so that's why you have this definition called inner product.
So inner product is nothing but vector lines times vector lines, there's two vector lines, times multiply them together.
Then you weight the product with this cosine theta, this cosine theta, always between 0 and 1.
So that naturally defines the angle and orthogonality when the angle is 90 degrees.
So you have this inner product evaluated to be 0.
So this is the idea.
So for n-dimensional space and this two-dimensional space, you can have two basis vectors, say along X or along Y.
When you have the three-dimensional case, you are very familiar with this picture.
So arbitrary vector A, the A can be this right vector, or you can think A is just a point, the tape of the vector.
And you can represent this vector A in terms of its three components, X component, Y component, Z component.
As I mentioned, if you know this vector A projected onto X axis, so you got the value projected result, XA, you have YA, ZA.
So you have the vector A, it's equivalent for you to have three components, XA, YA, ZA.
And how you can find XA, YA, ZA? You just project vector A onto corresponding axis.
But if you know the three components equivalently, you can build your original vector A.
So this is idea in three-dimensional case.
And in n-dimensional case, you can still do similar orthogonal representation.
And in three-dimensional space, you have three natural basis vectors, say 1, 0, 0.
Then the second one should be 0, 1, 0.
It's some typo, I need to correct.
And for the z-directional unit vector, 0, 0, 1, the arbitrary components, X, Y, Z, you can just give certain weight for the unit vector along X, along Y, along Z.
Then how you compute the coefficient? As I argued here, the coefficient is inner product of the vector to be represented.
It's vector A projected onto vector EX.
EX is a natural basis vector along X axis.
And likewise, you can do other things.
And for n-dimensional space, you have this general orthogonal representation.
So general vector W, just linear combination of all these natural basis vectors.
And the general way I call it, vector lowercase e with subscript n.
n equal to 1, 2, 3 until capital N.
And what's the coefficient? The coefficient is computed in terms of inner product.

Because inner product means projection.
So in that case, the vector W is projected onto natural basis vector EN, vectorized EN.
So this is inner product.
So you do the inner product.
And this EN has a unit length.
So you don't need to do normalization.
So the W projected onto the natural basis vector EN, that's just the coefficient.
And you compute all the coefficients for all the natural unit basis vectors.
And then you do linear combination.
So this is what you are doing.
So you get the original representation.
So just represent the function in different ways.
And this is one way.
And I say different ways, I indicate that you are not restricted to represent this vector, vector W, using the natural basis vectors.
And the basis, or basis, the set of basis vectors is not unique.
And for example, in two-dimensional space, and you can think that the basis vectors are two.
So one along x-axis, the other along y-axis.
So along x-axis, so you have 1, 0.
Along y, you have 0, 1.
So this is a set of basis vectors.
Or you simply call it a basis.
That's good.
You have a black base.
And you just rotate the coordinate by a certain angle, 5.
And I can just select my basis vectors along this right-axis, along this y-axis.
That's my basis orthogonal unit vectors.
And I can represent any point, this arbitrary point, either in terms of black basis, or I can represent the same point in terms of right basis.
So it doesn't matter.
So the two transforms is related by this rotational, rotation matrix.
And you read this part, and this has a good thing.
You know just that you can convert back and forth.
Similarly, you can do three-dimensional cases.
And then you can rotate with respect to each of the three orthogonal axes.
So you have many ways to form basis.
And you just do any way convenient.
And for general n-dimensional cases, we can have vectors of n elements.
So we call this space vector space.
So generally, it's just a kind of summary slice.
So you have two arbitrary n-dimensional vectors, v and w.
So inner product, we see this form.
Now we feel comfortable.
This form is well motivated by linear system theory, by signal detection, by geometrical argument.
I would think geometrical argument is most fundamental.
And I like think geometrical pictures.
I like visualize things.
So the geometrical argument is most important.
So when v and w are set the same, you can just do the inner product to retrieve the squared lines of the vector v.
So you have concept of distance.
And then you have the concept of angulation between vector v and vector w.
And I argued about that.
And then we can have a particular natural basis for n-dimensional space.
OK, just e1, e2 along the way until e n, 0, 0.
Last element is 1.
So all these are orthogonal.
You perform inner product among them.
As long as the two natural basis vectors are not the same, what would be their inner product?

They are orthogonal.
So inner product will be 0.
Inner product being 0 indicates that they are independent, linearly independent.
The arbitrary vector, you need n components because they are independent.
And you can arbitrarily specify the contribution from each dimension.
So add all things together doing the linear combination.
You recover the original vector v.
And the coefficient is computed in terms of inner product.
So the vector v projected onto the case, natural basis vector, then you have the case contribution, the coefficient.
All these coefficients combined vector-wise, then you can get the original vector.
So this is the concept, all geometrical concept.
So vector has many points, n-dimensional vector, like 100-dimensional vector.
You have 100 elements.
So we can have a geometrical picture.
So each vector, we think each vector is not a bunch of numbers.
We think each vector is just a single point.
And this is true.
So in two-dimensional space, you can have a unit circle.
So each point is just a tape of the vector with unit lines.
And all these put together, you trace a unit circle.
And in three-dimensional space, you have a unit sphere.
And any point on the sphere is a tape of the vector, three-dimensional vector.
The length of the vector is 1.
Likewise, for n-dimensional sphere, you just ask the length of the vector, n-dimensional vector, to be 1 is a unit sphere.
Or to be r, that will be a sphere with radius r.
So each point, each vector is a point.
And furthermore, we say function.
Normally, you think function is not a point.
I have y equal to f of x.
x is a point that keeps moving on.
Then I need to trace a nice picture.
That's a curve of function f of x.
And we say this is a curve, right? A curve.
As a function of x, x trace just a range.
So any point x corresponding to a value y.
So that determines a point in two-dimensional space.
But we say any function, if you just discretize a function, that becomes a series of bar charts, a series of impulses.
So it can be a vector, because you have, say, vector element 1, element 2, and so on.
So if you think the function can be discretized, so function is not so much different from vector.
So vector can be viewed as a single point.
Then we say the function can also be viewed as a single point.
It's not, say, a single point in two-dimensional space.
Rather, it's a single point in higher-dimensional space.
Say, if you approximate this function using a vector, six elements, then this function is a single point in six-dimensional space.
So you can make six hundred, six million, six billion, so you get higher and higher accuracy.
So from an engineering perspective, vector and function are basically the same.
And a vector can be viewed as a single point, and a function can be viewed the same way.
And for vector, you can compute inner product for the reasons I explained.
Then for function, you can also compute inner product.
Then this summation becomes integral.
I forgot equal sign.
I will fix after the lecture.
So this inner product defined here is just a summation of partial product.
And here is an integral of product continuously, and it runs through the whole axis, all through a finite interval.

But anyway, I just tried to say vector space and functional space, they are closely related.
And a vector space normally is a finite dimensional.
Functional space is an infinite dimensionality.
So that's the difference.
But mathematical form is kind of the same.
As long as you are happy, summation and integral, they are essentially the same under certain conditions.
So we can express arbitrary vectors using basis vectors.
Then we can also similarly express an arbitrary continuous function using certain basis functions.
And for example, what we learned the delta function is such a base.
And I copied this formula from your convolution lecture.
And we showed f of x can be written as an integral.
This is a convolution.
And the delta function and the f function itself.
So the magical meaning of this right-hand side is nothing but to say the original function, f of x.
And it can be viewed as many, many slices.
And these slices can be approximated with a delta function.
This delta function has the total height determined by this original function.
You just put all kinds of delta functions at all possible places.
So there are many, many locations.
So you have many, many basis functions called delta of x minus a.
Just like originally you have lowercase e, you have e1, e2, en.
Only one place, the element is 1.
The rest element is 0.
Here only at position t equal to tau, you have a value delta.
You have impulse here.
Or if you like, you think that's just a rectangular bar there.
There's total amplitude.
So that's the same idea.
So we can represent a function as a sum of impulses.
So just all these.
Each impulse, the unit delta function translated at a corresponding location forms a basis function.
So you have many of these basis functions.
Then you can compute weighting factors for all these basis functions respectively.
Then add them together.
You just recover the original function f of t.
This is one-dimensional case.
If you think function is multidimensional function, then you can extend delta function from 1D to multidimensional case.
OK.
So this is very nice.
And when we talk about digital pictures, we really take the same point of view.
And we think the picture can be decomposed into many pixels.
Each pixel corresponding to an impulse, stronger or weaker.
That's the weighting factor.
So the delta function will be obtained when the pixel approaches 0 in size.
So that's one very good idea.
So arbitrary function can be picture or signal, one-dimensional signal, two-dimensional picture, or three-dimensional image volume can be decomposed into pixels or voxels.
And that can be represented as basis functions in terms of unit pixel discretized delta function, or in limiting case, continuous delta function.
So this is one view.
Another view, we think about sinusoidal functions.
The sinusoidal functions are rather common, just like the motion.
One common trajectory is straight line.
The other one is just circular, like the moon orbiting around the Earth.
So when you have uniform speed, circular motion, and component-wise, you naturally recover sine and cosine functions.

So I'm trying to say sine and cosine functions are very natural.
So you have sine, you have cosine, and for cosine, this is an even function.
For sine is an odd function.
Sine and cosine, they are actually even or odd.
It depends on the face.
So it's not an essential thing, because a coordinate system is set up by an observer, by a researcher.
We can set up a coordinate system this way, then the green function is even.
I can set up a system here, then the green function becomes odd.
So this just depends on.
And the interesting thing is that, remember earlier slides, I said function as a sum of impulses.
And this slide is important.
I say function can also be viewed as a sum of many waves.
All these waves are sinusoidal functions.
They may be different in amplitude, in frequency, in phase, but all these things add together.
And you can recover an arbitrary function.
So that's just another point of view.
This is the same function.
You can either represent the function as a summation of impulses, or you can express the function as a number of sinusoidal waves together.
So this is a demo.
So just let the animation play out.
You see it.
Right function is the target function.
Then we can use many blue sinusoidal waves, add them together to approximate this right function.
And for each frequency, you have different amplitudes.
Frequency versus amplitude.
That's our Fourier spectrum.
So this is an idea.
You perform one-dimensional Fourier analysis.
You represent an arbitrary function as a sum of various waves.
And then you can do so for two-dimensional pictures, the Albert Einstein.
So you perform Fourier analysis.
We will learn later.
Basically, this arbitrary picture is decomposed into many 2D wave fields.
Some waves travel horizontally.
Some waves travel diagonally or vertically.
And these are different frequencies.
So you have all these two-dimensional sinusoidal wave components.
And you add them together.
Certainly, for any given picture, the amplitude and the phase would be different.
You add them together, you can recover the original two-dimensional pictures.
The picture can be arbitrary.
So the building blocks are very smooth and regular, all sinusoidal things.
As long as you make the sinusoidal components right, amplitude wise, frequency wise, phase wise, you make all these right, you can recover the original picture.
So this is a very interesting alternative way to represent a function.
So in physics, we know anything has dual characteristics, particle and wave.
And in information science, we see similar things.
Particle wise, you have the pixel or continuous impulse, discrete impulse.
So you can have pixel-wise representation.
So basically, you decompose the function or picture into summation of many small elements, pixel or many delta functions.
And also, we can represent functions or pictures as many kinds of waves.
So that's just the point of view of wave characteristics.
In this case, we use the sinusoidal functions.
That is very reasonable.
You know Maxwell equation, you solve equation, and then you get the sinusoidal solutions when you deal with just, say, free space, free space, and EM field

propagation.
So that's a very natural solution.
And the information science and the classic material or physical science, they are really linked together.
Here, all we have been saying.
So we can represent a function in different ways.
The basis function can be all kinds of delta functions or all kinds of sinusoidal wave functions.
You want to represent the same thing from a different point of view, so you have a deeper understanding.
And then the representation is nothing but you think the function is a vector in finite or infinite functional space.
Then you have a set of basis vectors or basis sinusoidal functions.
And the function can be projected onto individual elements of the basis.
Then you have a linear combination.
You recover the original function.
And the basis is not unique.
So I give you two examples.
All delta functions, you save the delta functions to all possible locations.
So all these safety words together form a basis, delta basis.
And all sinusoidal functions, you have a sinusoidal functional form, and you can stretch them with different frequency.
You can certainly save them a little bit.
So all these things together form a sinusoidal basis.
So they are just different sets of basis elements.
And they can do the same thing, but from a different point of view.
So for the higher dimensional space, orthogonal representation, these are good for you to have a geometrical picture.
Then the next part of the lecture, we talk about Fourier series and the periodic function.
And once you understand the also normal basis, and then you represent arbitrary function in terms of a given also normal basis.
And it will be clear why you can represent the function f of t as an integral convolution with delta function.
And then next you see you represent the function as a Fourier series.
So we have 10 minutes rest, then we really talk about the Fourier series.
Previously, we learned to use delta functions, linearly combine different kinds of delta functions to represent the function f of t.
So we covered the original function f of t, discrete or continuous.
Now I just suggested that we could do similar thing with wave or sinusoidal functions.
Pretty much like in physical science, and you have particle view, you have wave view, depends on situations that you can treat in particle view or you treat the situation from a wave perspective.
So this is something very similar to what we are going to do for imaging science and engineering.
And the first that we talk about a particular type of periodic function, a function called the periodic function.
So arbitrary function can be periodic or non-periodic.
And we just now we think function is a periodic.
And then later we will extend the discussion to non-periodic function.
This is convenient.
So for sinusoidal representation of a periodic function, so we need to use so-called Fourier series.
So this is just a one-dimensional example of what we mean by sinusoidal function.
This is a two-dimensional thing, the sinusoidal function.
You see the two-dimensional pattern is repeated horizontally and vertically, and you can have periodic functions in any dimensionality.
So the question is how we can represent arbitrary sinusoidal function in terms of summation of sinusoidal waves.
Because a function is a sinusoidal, so we need to make sure we use sinusoidal periodic function, add them together, sinusoidal functions, add them together to perfectly recover

any copy, say this copy.
If we have this part represented, and the other part will be represented accordingly.
So that's just our task.
And what we will use is the so-called Fourier series without loss of generality.
We consider the function is defined over a unit interval, zero, one.
We just consider the function will repeat itself.
So from interval one to two, two to three, every subsequent interval you will see the same thing as you would over interval zero, one.
Likewise, you have negative part, and just say interval minus one, zero, minus two, minus one.
So you just have multiple copies.
So we focus on expressing of the function over this unit period, this unit interval.
And later we will extend the work to arbitrary intervals, say from A to B, or from zero to T, capital T, or from minus capital T divided by two to positive capital T divided by two.
So we can extend the Fourier series from unit interval zero, one to arbitrary interval from A to B.
So let's focus on representation of a function on unit interval.
So the function can be defined, any function can be defined over this interval.
We are particularly interested in real value functions defined over this interval, yet the function is square integrable.
That means the function say F of T, you square the function, you do the integral, you make sure this integral from zero to one will give you a definite value.
If you do integral, give you a value infinity, you cannot deal with that.
So we deal with meaningful things, and normally meaningful things in engineering practice.
That means you have finite, definite results.
Infinity is not a number, it's something called divergence.
So you just cannot deal with divergence in your daily life.
And the idea is that we focus on this interval, and the functional elements are all square integrable, so we denote the space as L2, and O1, that means the function is well behaved, so it does not give us a divergence problem.
And as I mentioned, it's a continuous function, and it can be discretized into small segments, and each segment is a discrete delta function.
In the limiting case, you have just the many, many small discrete delta functions, that becomes a real dirac delta function, safety, the scale, so that is the kind of basis we explained to you in the convolution lecture.
And a while ago, I just mentioned, also normal basis is not unique.
In addition to the safety, the scale of the delta functions, and we can represent the original function as a summation of many sinusoidal waves.
And particularly, I claim, so you use the sinusoidal waves, cosine and sine, cosine and sine are essentially the same, you just shift it a little differently.
Anyway, so sinusoidal function.
In the special case, you have a sinusoidal function, it varies so slow, infinitely slow, that becomes a constant.
So constant is a special sinusoidal function.
So if I just put a constant, I put just a cosine component, sine component, and with different frequency components, with n, say n equal to 1, you have lowest frequency.
You can keep increasing n, so the oscillation rate becomes higher and higher.
So you have all these things, I put, why I put 1 instead of 2? Because we are talking about natural unit lines.
So you do inner product 1, which is the sine of, and the length is 1.
The cosine square root 2, cosine 2 pi nt, you do inner product, this element, this second element, which is sine of, over interval from 0 to 1.
It gives you the length of that base function, gives you unit length, so we make sure each of the elements, any two distinct elements, you do inner product, it gives you value 0.
So we need to do inner product in terms of integral, not summation, because we are talking about continuous functions.

Continuous functions live in infinitely dimensional space.
And in infinitely dimensional space, you can imagine you have infinitely many basis functions, so each of this list is such a function, so they form a basis, sinusoidal basis.
So this function can be represented as a sum of all these sinusoidal things added together.
So what we don't know is coefficient, a0, an, bn, you don't know coefficient, but you know all these basis functions.
So you just need to compute all these coefficients, then you can find a representation of f of t in terms of cosine and sine components plus dc components.
So that is Fourier series, it's nothing but also normal representation of original function.
I say also normal presentation, I just claimed this line forms a basis, and this line really just convinces you, indeed that line forms a basis.
So you do inner product 1 and 1 itself, so your inner product defined in terms of integral from 0 to 1, you can easily verify this is 1.
And the inner product 1 with sine or cosine components, that will give you 0, the sinusoidal component, you have positive cycle and negative cycle weighted by 1, that doesn't matter, that will give you 0.
And the cosine and sine components, they do not play together well, so any sine components, you do inner product with cosine components, it gives you 0.
So cosine and cosine, just the frequency difference, when m and n are not the same, the inner product is 0, but when the m and n is the same, it's the only case you have non-zero value 1 over 2, and likewise for sine components, you have similar things.
So that's square 1 over 2, you think when you do the normalization, you come with square root of 2.
So the idea here is to show 1 cosine 2 pi nt and sine 2 pi nt, so you have a long list of these members.
Any two of them, if they are different, and you do the inner product, that will give you 0.
And if the same element do the inner product with itself, it will give you unit value.
So that shows this is indeed also normal basis.
And also normal basis can be used to represent the original function.
The coefficient is computed in terms of inner product.
That's the inner product of the arbitrary vector.
In this case, it's an arbitrary function, and projected onto a given sinusoidal function.
So pretty much like what I explained here, two-dimensional, three-dimensional space, how you represent the vector.
You project the vector onto x, y, z, respectively.
Here you project the arbitrary function, f of t, onto 1, onto cosine 2 pi nt, sine 2 pi nt, and could be 1, 2, 3, anything.
You just do all these things, you can find all these components.
And the sinusoidal components, they are very natural.
They do nice oscillation.
They have another name called harmonic component.
They look so nice.
And because n equal to 1, sine or cosine will do this full circle over 0 and 1.
When n equal to 2, you just shorten the period by half.
So you just double the rate of oscillation.
And you can do triple the frequency.
If you make n equal to 3, you can keep doing so.
But whenever, whatsoever you do, you have a series of discrete numbers or integers, n equal to 1, 2, 3.
And so you always start from 0 and from 0, not only value-wise, also change rate.
You see, when you deal with periodic function, you have one cycle.
So the ending point of a given cycle is the starting point of the next cycle.
So if you don't think about DC components, the ending point and the starting point really should seamlessly link together.
Not only in terms of functional value, also in terms of rate of change.

So the function is continuously changed.
This sinusoidal component will change at the same rate.
So if you look at this, you'll move this way.
You look back, it's this part.
It's just symmetric with respect to the starting location.
So this is to say the Fourier series over unit interval.
So f of xt defined over the unit interval, 0 and 1, can be expressed as three terms.
The first is DC components.
This is a simple thing.
This is a background thing.
On top of the background, you have a cosine part.
You have sine part.
A cosine part is an even part, is an even function.
The sine part is an odd part.
So you have three parts together.
And as we show here, so once you have representation,
we'll approximate it over interval 0 and 1.
And the function is sinusoidal, and the cosine is the same thing.
It's a periodic thing.
So if you have the waveform here, right waveform, next cycle, you will have the same thing.
And the right situation will be the same as what we see on these slides.
For right components, green components, then the blue components, purple things, all the same.
That is to say, if in this particular interval, you see you're playing iPhone,
in the ivory period, you'll have many, many copies.
They're all playing iPhone, so the same thing, the periodic thing.
So this is common sinusoidal, sorry, arbitrary function can be naturally represented
in terms of even part and the odd part.
The DC component is both even and odd, only if DC is 0.
Otherwise, it's an even part, it's a constant, so you think it's an even function.
So anyway, so you have this representation.
More generally, and this is just a small trick, any function can be always,
even any function, f of x, you can always make an even function this way.
You can make an odd function in this way.
And the even function and the odd function, so constructed, added together,
will form, will recover, return to the original function.
So that is to say, any function can always be decomposed into even part and the odd part.
And this is just consistent to what I told you here.
Now you want to know, and we have this Fourier series, how should we find the coefficient?
You already know the building block, and you need to have the coefficient
to weight the building block, then add it together.
That will be the particular function you want to represent.
How do you find the coefficient?
So given what I explained a lot about higher dimensional space, orthogonal representation,
inner product to find the coefficient, so all these things we just learned apply here.
So how to find the Fourier series coefficient becomes straightforward.
You just do inner product.
For example, you see here, you want to find, just for example, you have b, bn, n equal to 0, 1, 2, 3.
Just for example, how do you find the coefficient b3?
That's clear.
I do inner product of this function.
This f of t is given, so you know this.
You do inner product with sine 2 pi 3t.
Sine 2 pi 3t is a very specific function.
You know that, and you know f of t, you do inner product on this part.

So this can be done.
And you see this sine 2 pi 3t, you do inner product this part.
You should also do the same inner product operation on the right-hand side.
Then inner product for the whole thing is equal to inner product with this first term plus inner product with second term with third term.
And the inner product sine 2 pi 3t with first term, because also normal property, will give you 0.
And here will give you 0.
Here you have many numbers, b1, b2, b3, the courage bounding to sine 2 pi, n equal to 1, 2, 3, and so on.
From 0, 0, 1, 2, 3, and so on.
So only one term will not give you non-zero value.
That is 1, n equal to 3.
So when you do inner product with this whole series, only one term, single out, will not be eliminated, will not vanish.
That is b3.
So you do inner product, you got the value for b3.
So you see, 1, n equal to 3, you do inner product.
2 pi 3t with I of t given.
So you know I of t.
You compute this value.
This value, whatever you find, will happen to be the value b3.
Likewise, b0, b1, b2, b4, b5 can be computed similarly.
And similarly, a n can be computed.
This is really the merit of also normal basis.
So you pick up a single element, and you do inner product.
And also normal presentation will be such that the inner product with a particular element in the basis will give you courage bounding coefficient.
And again, you recall what you learned.
The three-dimensional vector, how you find the three components, you do inner product.
With respect to ex, that's a unit vector along x-axis.
With ey, ez, the same thing.
Just the basis, not natural unit vector.
Rather, sinusoidal components.
The dimensionality of the space is not three, not two.
It's infinitely many.
But the geometrical picture is exactly the same.
So Fourier series, the complete summary.
So arbitrary, not really arbitrary.
Many functions, like continuous functions, can be represented as summation of sinusoidal components.
And this is just the overall structure.
It's not specific.
To find Fourier series for a specific f of t.
And you need to specify a0, a n, b n.
So a0, a n, b n can be computed by inner product operations as explained before.
So this is the whole summary.
And it gives you a function.
See, I can arbitrarily draw a function.
Parabolic function, or just something like triangular function.
You can always go through the exercise to compute a0, a n, b n.
Then plug a0, a n, b n there.
And do the right-hand side summation.
So that means all these sinusoidal functions are added together.
And then magically, you just find that this summation gives you an original function.
So that is just another way to represent f of t.
So given a target f of t, and we spend a lot of time to represent f of t in terms of delta function,
in terms of sinusoidal function.
And this is our collective experience.
Many scientists and engineers, whenever you deal with anything,
you have more than one way to represent, to understand, to treat the task.

And you normally do a good job.
So now we have at least two ways to represent a signal.
The delta type expansion, or pixel-wise point of view.
And now we're talking about the sinusoidal expression.
So that is from a wave perspective.
So this is the main thing about Fourier expansion.
So the rest of the part, I will go a step further.
So you look at this slide, everything is clear.
All real functions, real coefficients.
Now I give you some trouble.
So I want to represent a real form of Fourier series into a complex form.
So we involve a complex number.
So complex number means more complex.
The imaginary number is something weird.
Why we bother go from such nice, clear, real form?
Why bother use a complex form?
Let me tell you one reason.
See, there's a Fourier series.
You have a direct term.
You have a cosine part.
You have a sine part.
You have three parts.
If you use a complex representation,
under this formulation, three parts just become one part.
You can unify sine and cosine using a complex exponential function.
So the expression all of a sudden comprises from this long to one third.
Then the derivation, everything becomes easy.
That's just a practical reason.
There are other reasons that you will appreciate as time goes by.
That gives us convenience.
Let me just underline.
So just trust me.
We want to move from real form to complex form.
But the essential idea already revealed to you in the real case.
The complex representation, and you know this.
So you have imaginary number, unit i.
The definition, the first step, is really weird to me.
So you know anything squared will be normally what we learned in high school.
Anything squared will be zero or positive.
But if something squared to be negative, that's contradictory.
But let's just think.
There is a number.
We imagine there is a number called i.
And when it is squared, it gives us minus one.
So we accept this as a fact.
Then we think, we construct this number and this imaginary thing.
Based on this, we just apply, just extend normal operations.
Like addition, subtraction, multiplication.
In all the reasonable ways, as reasonable as we could think.
So we just follow this logic.
We want to say, if we define i squared equal to minus one,
then we extend the algebraic operations as best as we could.
And we hope all the operations would not end up with contradiction.
Then we have a new mathematical system called complex analysis.
So in a way, mathematics is like a complex analysis.
It's just our intelligent construction.
It's a creation.
But anyway, it works very well.
At least for Fourier expression, we can use complex form
to greatly reduce the workload of mathematical derivation.
So just to review this, complex number
has a real part and an imaginary part.
So you think this imaginary thing, the unit,
is a special thing orthogonal to the real axis.

You have an imaginary axis.
So the complex number is a point on a complex plane
defined as x plus iy.
The conjugate has just changed the sign of the imaginary part.
So that will make a complex conjugate of a complex number
symmetric with respect to the real axis.
And the multiplication defined naturally,
so we want to have similar rules, same rules applied
to multiplication in the real domain.
So for complex number, you can talk about the real part.
You can talk about the imaginary part.
In terms of conjugate, imaginary and the real part
can be expressed easily.
And for inner product, we can think
about the conjugate of inner product.
That will be conjugate of corresponding element
subject to inner product.
So conjugate can be this top bar or this star sign.
So just notations.
And for complex number, you can have Cartesian expression,
x plus iy or yj.
ij both commonly used for imaginary unit.
And in the draft, I prefer to use i.
Anyway, that's just a notation thing.
Don't worry about that.
So the polar Cartesian rectangular representation,
they are interchangeable.
Then something a little tricky about inner product.
I keep saying inner product, just element-wise, point-wise,
matching partial product.
Then add together or integrate the multiplication together.
But in complex case, the inner product
is not just two vectors you match up together.
And one of the vectors need to take a conjugate first.
You do conjugating operation first.
Then you do pair-wise matching.
Add them together.
Why bother about this conjugate operation?
So I gave two arguments in the textbook draft
so you can read.
The first argument, so I put a green button.
Just to let you know, you do not necessarily
remember these things for examination.
But it's just nice for you to know.
So you have two complex vectors.
Simple case, two-dimensional, complex dimensional.
So x, you have first element and second element.
Each is a complex number.
You have a second vector, also two-dimensional.
You have this and this.
And we can naturally define lines of the vector.
This is really a four-dimensional thing.
Although you have a two-dimensional complex vector,
this is really four-dimensional degree of freedom.
So the lines is a1 squared plus b1 squared plus a2 squared
plus b2 squared.
Likewise, you can compute lines of y, lines of y vector.
Then we think if these x and y are orthogonal to each other,
then you can add them together.
So x plus y, you form a new vector z.
So this is the new vector z.
You add these together.
You have the new vector z.
So the new vector z, according to the rule

we defined in the complex domain, is this one.
Then the total lines of this new vector z
and the similar argument will be this one.
Because we say if x and y are orthogonal,
then the lines of z should be this one.
And we want to have the property,
So the lines of z squared should be equal to
lines of vector x squared plus lines of vector y squared.
We want to have this.
Then you just do a little computation.
That means to have this nice property,
piece of green theorem in this case,
you need to have this part to be zero.
Otherwise, you don't have this nice property.
So this suggests that you do the inner product.
You want to ask the inner product to be zero
for the orthogonality to hold.
So that suggests that you really need to compute
inner product in this way.
So x vector, inner product, this is pairwise,
element-wise matching in the production,
really need inner product with vector y after conjugation.
So that's why we define in complex domain
the two vectors.
You perform inner product operation.
It's not simple-minded x dot y.
Rather, it should be x dot y conjugated.
That's y star.
So a simpler argument would say you have a complex vector.
And then you can do inner product with itself
to compute the squared lines.
And then you want that number to be real.
So you better do conjugation.
When you do conjugation, you do multiplication.
The phase angle cancel out.
So you have a real number.
So just understand the definition of inner product
in the complex space.
So the inner product is important to induce
geometrical structures into new space.
So inner product give you distance,
give you angulation.
And for the geometrical intuition or ideas
applied to complex space,
inner product need to be defined that way.
OK.
And once we have the complex number domain and operations,
you can extend the definition of exponential function
into complex domain.
You got this one.
OK.
And you can define these infinite series.
And you need to make sure the definition is meaningful
and that means you need to make sure
such infinite summation converge to finite number.
And this can be shown using result test.
Using result test, you can prove this converge.
And when z is x, this is just the returns
is degraded to the original real exponential function.
But when z is arbitrary, z is equal to x plus iy
is an extended definition.
It's meaningful.
We show it converge.
So this is something we can say this is a function.

If it just diverged, the definition wouldn't make sense.
OK.
We argue that this makes sense.
And particularly, we write z equal to i theta.
Then we just do a little bit of arrangement.
We see when z equal to i theta, and this infinite series
can be separated into two parts.
One happens to be cosine theta.
That's just the original Taylor expansion for cosine theta.
And i factorizes out.
You have this sine theta.
And this is so-called Euler's formula.
With Euler's formula, you do some simple steps.
And you see the cosine can be expressed
as two terms of complex exponential function.
Likewise, sine can be expressed in terms of e to the power i
theta.
So this is nice, because once we know
how to represent sine and cosine in terms of e to the power
i theta, then we are ready to represent f of t
in terms of e to the power i theta.
Here theta is something 2 pi nt.
Because cosine 2 pi nt is just the whole thing.
The theta can be expressed as right-hand side, likewise
for sine.
So this is just the same thing.
This is theta.
Here is theta.
The next slide here is 2 pi nt plus rho theta.
So you have 2 pi nt here.
So all secret is here from real to complex form.
So this is a real form, so you have sine and cosine.
Now we know either sine or cosine
can be expressed in terms of e to the power i 2 pi nt.
Or you have a minus sign.
It doesn't matter.
So you can just express sine and cosine in terms of e
to the power i 2 pi nt.
And this is exponential basis function as claimed here.
So all these things, n, m, can be positive, negative.
And all these things can be summarized as the last line.
You do inner product, usually give you 0.
Otherwise, m and n is equal.
So this is in complex space.
The basis in complex space.
So once you have basis in complex space,
then you can represent f of t in terms of these basis functions.
So just to reinforce your memory,
this is the complex basis.
And why you do inner product?
You do inner product.
Remember, you have a conjugate here.
Conjugate means you change the phase angle
from positive to negative.
You keep doing this operation step by step.
You show the result is delta n, m.
When n equal to m, you've got 1.
Otherwise, it's 0.
It's just the same thing.
It's the basis.
Any two elements, different elements,
their inner product will be 0.
And the inner product element itself is 1.
So this is good for our functional representation.

So we now represent function f of t as a summation
of many, many, many, many complex sinusoidal terms.
So each term has a different complex frequency.
i, 2 pi, np.
n could be 0, 1, 2, 3.
Could be minus 1, minus 2, and so on.
So it really goes from minus infinity to infinity.
So if you believe in Euler formula,
then the sine and cosine can be expressed
in terms of complex harmonic components.
All of them form a complex basis.
Now, how you compute cn?
Again, our old trick, we do inner product.
And only given n, so for given n,
so you do inner product.
Only that given n, the coefficients corresponding
to that given n come out.
You got this number.
Or other than the inner product for the corresponding term
will be 0.
So you don't have it.
So this is a nice separation.
So this way you got cn.
And cn certainly is related to an and bn and a0 as such.
So you just see the relationship.
And if f of t is real, then cn and c minus n
are related by the conjugate operation.
You can just check for the mathematical detail.
But the basic idea, the fundamental idea,
all summarized here, Fourier series
can be visualized as a representation
of the original function in infinite dimensional space.
And if you work in real space, the basis
has infinitely many numbers.
So the general numbers is cosine 2 pi mt,
or sine 2 pi nt, could be any mn,
or all these things together, plus constant 1.
So all these things form also normal basis.
They weighted together can represent f t.
Not so much different from a three-dimensional,
two-dimensional case.
You have arbitrary vector.
Then you can represent that vector
in terms of unit vectors along x, y, z.
And how you find the co-efficient,
you do inner product.
Here you do inner product as well.
In complex domain, you have complex harmonic element.
So this is just e to the power 2 pi mt,
e to the power 2 pi nt.
And you really need to make a different type,
just the change of the class.
So all these, you just need to stress your imagination.
So many, many unit vectors.
And then you just imagine it's just
a high, infinite dimensional space.
And what do you do when you perform Fourier analysis?
It's nothing just decompose your arbitrary function,
or if you like, arbitrary vector into components.
You have infinitely many components.
So what's the component contribution?
You do inner product to find co-efficient by projection.
That's all you do.
Then you can see a specific example.

You can review the example after the class.
And this particular example, square wave,
and you just go through the computation
and all the mathematical steps.
So you can get a finite summation fairly close
to the original term.
And in the middle, so you have discontinuous point.
The summation will give you average.
If in the place where function is continuous,
you can approximate the function very well.
In small terms, you use the better accuracy.
But close to the discontinuous point,
you always have this so-called Gibbs oscillation.
So all these are details.
As the number goes larger and larger,
so I will be very small.
But the discontinuity is harder to be
fit using continuous sinusoidal waves.
So you have the Gibbs effect.
When the period is not united,
there are one paragraph in your textbook
to discuss how do you do extension.
So you can just change the variable
using a scaling factor.
You just map an arbitrary interval
into the interval 0, 1.
Then you can use what we just learned.
And then you change the variable back.
You will get this arbitrary period.
Function over arbitrary period expressed as a Fourier series.
One dimensional case can be extended
to two-dimensional, three-dimensional cases
in similar steps.
So this is your homework.
So much for today.
And this is slightly more complicated than convolution.
So please review.
And if you understand well,
then you will have an easier time
for Fourier transform,
which will be explained in the next lecture.