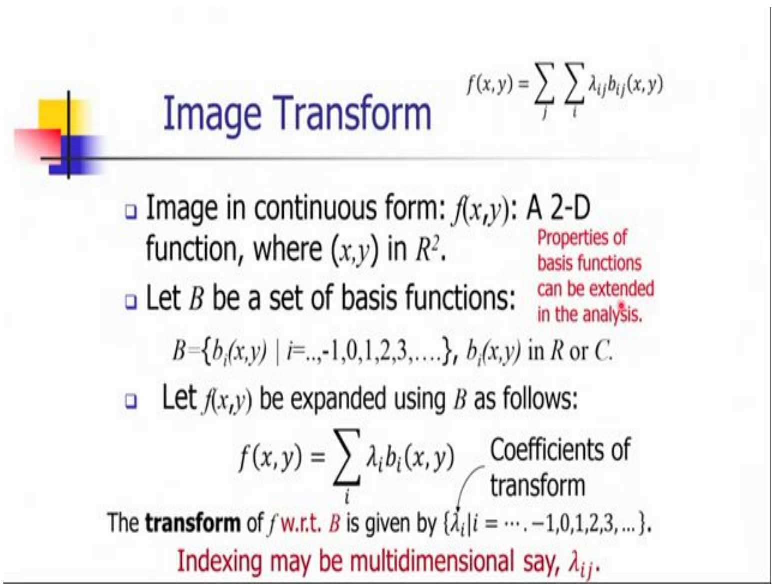
Computer Vision

Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur

Lecture - 03
Image Transform Part - I

In this lecture I will introduce Image Transforms.

(Refer Slide Time: 00:21)



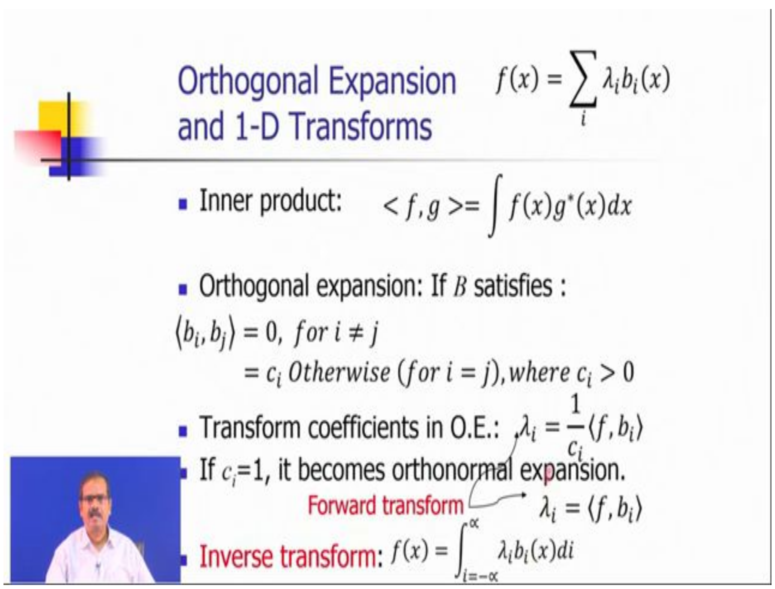
Let us consider image as a continuous function; it is a two-dimensional function and a point is in the two-dimensional real space and let us consider set of basis functions which are also a two-dimensional functions, we can represent it as a set where each function is given by say

$b_i(x,y)$. you should note that this could be the functional value, could be either in real or in the complex domain. We can represent, we can expand the two dimensional function f(x,y) using B as a linear combination of this basis functions as it is given here in this form that $\lambda_i \times b_i(x, y)$ and where i is the indexes of the basis functions as given in the set.

So, the transform of f with respect to B is given by the set of set of this coefficients which are λ_i these are called coefficients of transform and we can represent the function as the linear combination of this basis functions where the coefficients of the linear combinations are listed here. So, you can see that this is an alternative description of the image instead of representing the image by the functional form of f(x, y); I can simply represent it as a list of coefficients or even these coefficients can be a function of indexes.

So, indexing maybe multidimensional, for example for a two-dimensional function indexing we can use two such indexes to denote a coefficient and in that case we can say that this function can be expanded in the form of a linear combination of two dimensional functions and we can have this double summations in this representation. So, one of the advantage of image transform is that this properties of basis functions this can be extended in the analysis.

(Refer Slide Time: 02:37)



Let us consider a particular type of property which is very useful when this basis functions they have this property. This property is called orthogonality property and if I expand the function in terms of this orthogonal basis functions then we call that expansion as orthogonal expansion. We will consider our discussion following up the discussion on this image transform, we will restrict our discussion to one-dimension first and later on we will see that now these properties can be easily extended to two dimension. So, we will understand the one-dimensional transform initially.

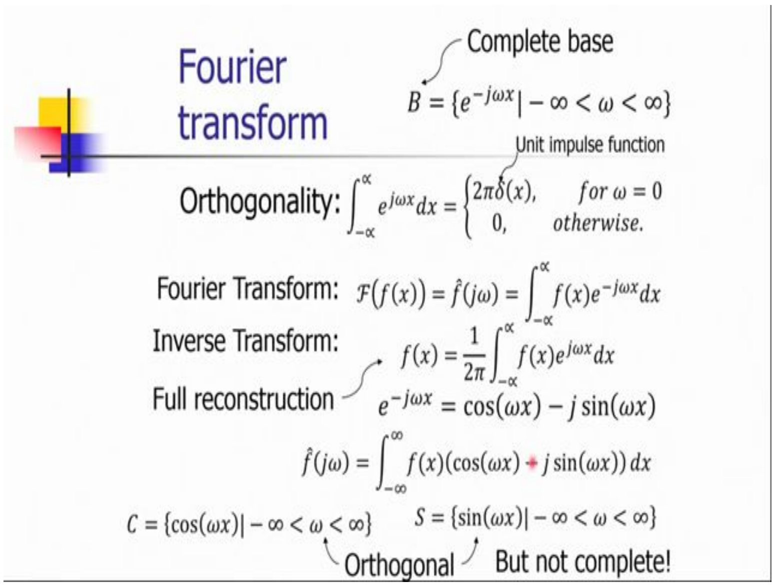
So, one of the thing that we would like to define here this operation inner product. So, inner product is a binary operation where two operands are two functions you can see that this function f(x) and the another function g(x). So, it is the product of this two functions and integral of this product values at every point in the space x. Now of occurs in this product there is a there is there is something we should note that it is not a simple product, it is a product of function f(x) with the complex conjugate of g(x) and that is we are considering here. If both f(x) and g(x) functional values are in the real domain then complex conjugate itself will be the same functional value, so then we can write it as f(x) g(x) dx.

So, orthogonal expansion it is possible when the basis functions this satisfies certain property, that means the set of basis functions it satisfies this particular property of a orthogonality. What we can see that inner product of any two different basis functions that should be equal to 0, whereas inner product of the same basis function will have a non zero value and which is a positive value. If this is true for any pair of basis functions in the set B, then we say that this basis functions they are all orthogonal in that set.

So, transform coefficients in orthogonal expansion that could be easily computed by exploiting this property, that is one of the usefulness of this particular property and you can say that simply if I take the inner product of the function with a basis function, then we can and also divide it by c_i then we can get the corresponding λ_i's. And if c_i is equal to 1 then it becomes orthonormal expansion then we can simply write λ equal to inner product of function and b_i. So, this is what of since we are expressing the functions in terms of only transform coefficients, so we call this operation is a Forward transform operation.

So now, function instead of being represented by f(x), now the functions function is represented by these λ's and the reverse transform or inverse transform would be to compute the function from this coefficients back and because of the orthogonal property and also orthonormal property, we can simply write it in this form. That means, simply it is a; it is nothing but the linear combination of those basis functions and since it is a continuous domain. So, we are taking the integrations over all the index values otherwise in a discrete domain we can write it in this form.

(Refer Slide Time: 06:29)



So, one of the special case of this orthogonal expansion is Fourier transform and in this case you can see that set of basis functions is given by in the following form:

$B = e^{j\omega x} | -\infty < \omega < \infty$ (eq. 1)

it is a complex sinusoid which is the member of the set, it is given in the above form and the completeness (Refer Time: 06:53) of this basis function is : any orthogonal set which is a subset of any orthogonal basis set will also remain orthogonal, but using the linear combination of that subset, you will not get the complete reconstruction of the function.

The basis set which keeps the complete reconstruction that is called the complete base. So, in the Fourier transform, in fact, the set what is defined here it is a complete base because, it can give me back this function as a linear combination of this function. You can see that actually this (eq. 1) is a infinite set, though individually every sinusoid can be distinguished here.

So, the orthogonality property is reflected by this particular relationships where you can see the that □(x) is the unit impulse function whose area is equal to 1 centering at ω equal to 0 and otherwise □(x) value would be 0 in everywhere. So, it is an unit impulse function and this particular property gives you the orthogonal property of the Fourier transform this base.

So, Fourier transform can be defined in the following form which is the forward transform .

$$F(f(x)) = \hat{f}(j\omega) = \int_{-\infty}^{\infty} f(x)\, e^{-j\omega x}$$

as you see in the above that it is an inner product of f(x) and also the complete base. So, the base is $e^{j\omega x}$,. So, you take the complex conjugate of the base which is $e^{-j\omega x}$. And if I take the inverse transform then once again this is considered as the inverse transform;

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(j\omega) e^{j\omega x}\, dx$$

So, $\hat{f}(j\omega)$ is the corresponding coefficients and the linear combination of this basis function will give you the corresponding inverse transform.

So, it gives you the full reconstruction because it is a complete base as I mentioned. One of the interesting fact that can be noticed in the following complex sinusoid

$$e^{-j\omega x} = cos(\omega x) - j sin(\omega x)$$

it can be recomposed into two real and imaginary parts, real part consists of $cos(\omega x)$ and imaginary part consists of say $-sin(\omega x)$ in this case. So, this forward transform can be expressed using the following expression itself.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(j\omega)(cos(\omega x) - jsin(\omega x))dx$$

So, it has one transform component which consists of real part, another transform component which consists of imaginary part and in the real part we are using the basis functions as $cos(\omega x)$ where as in the imaginary part we will be using the basis functions at $sin(\omega x)$ or $-sin(\omega x)$ whatever be your interpretation.

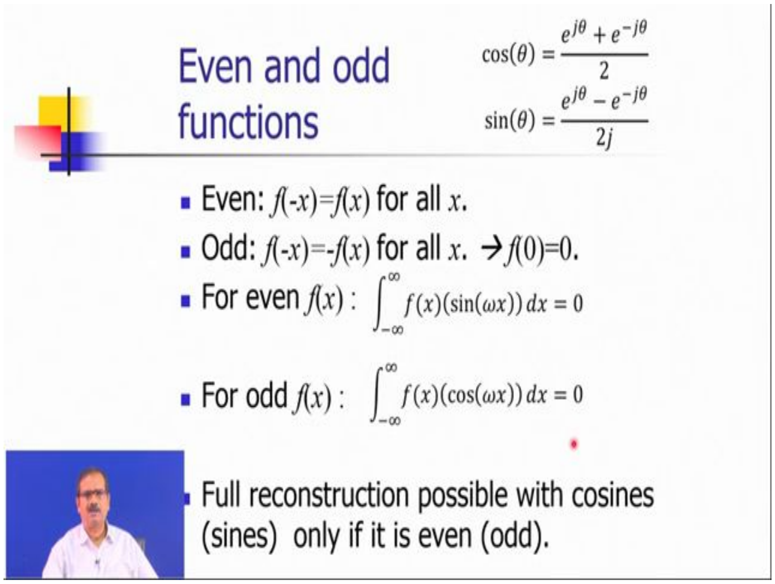
So, we can consider $cos(\omega x)$ as the set of basis functions and this is also orthogonal,

$$C = cos(\omega x) | -\infty < \omega < \infty$$

$$S = sin(\omega x) | -\infty < \omega < \infty$$

The above trigonometric functions are also orthogonal we know and $sin(\omega x)$ is also orthogonal. But the thing is that as I mentioned that they will not form the complete base. So, we can use only $cos(\omega x)$. If I use $f(j\omega)$ $cos(\omega x)$ the real part of $cos(\omega x)$ then that will not give you back the full function. Similarly if I use the corresponding imaginary part of the transform and use the $sin(\omega x)$ that will not also give me back the full function.

(Refer Slide Time: 10:55)



So, it is not a complete base but there are certain functions where actually if you use only cosine functions or sine functions, you can reconstruct it fully, so these functions are called even and odd functions.

There is a property like for an even functions, it should satisfy the following property:

$f(-x) = f(x)$ for all x

i.e, it should be symmetric around the origin or around at $x = 0$ at both ends say $f(-x)$ should be equal to $f(x)$ for all x

$f(-x) = -f(x)$ for all x

Whereas for odd it should be antisymmetric. i.e, $f(-x)$ equal to $-f(x)$ for all x and

$f(0) = 0$

naturally at the $x = 0$, the value has to be equal to 0 for this definition.

So, a function could be even, it could be odd or it could be neither, when they are belong when they have this property, then you can expand them using only cosine or only sine functions. So, let us see so even $f(x)$ we can have this is a property which is satisfied because, in that case if I take the integrations while taking the product with $sin(\omega x)$ that would be 0. I.e,

$$\int_{-\infty}^{\infty} \hat{f}(j\omega)(sin(\omega x))dx = 0$$

So, all sinusoidal terms would be 0; so that is why only cosine terms will remain and your transform coefficients can be sufficient to prescrib by only cosine transformations.

$$\int_{-\infty}^{\infty} \hat{f}(j\omega)(cos(\omega x))dx = 0$$

Similarly for odd $f(x)$ the above property is true and that is why using only sine sinusoidal basis you can reconstruct it.

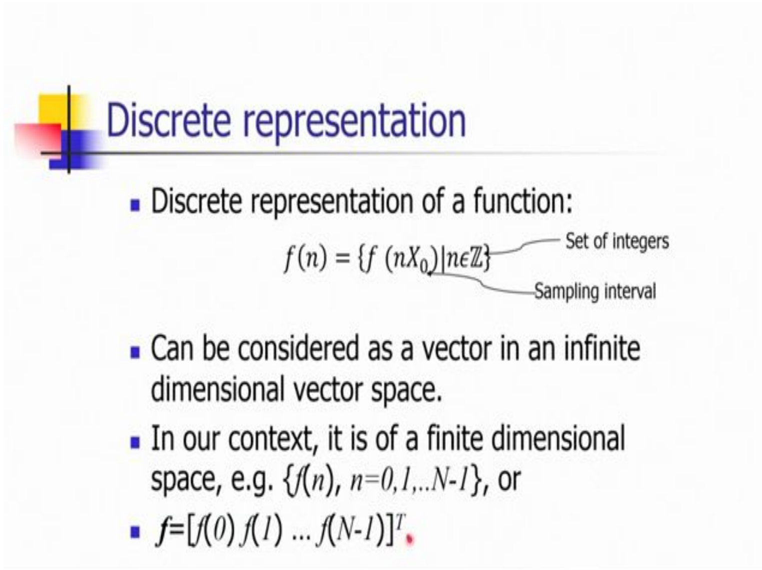
Now this could be easily derived if I consider this relationships of $cos(\theta)$ and $sin(\theta)$ in terms of the complex exponential quantities as below.

$$cos(\theta) = \frac{e^{j\theta}+e^{-j\theta}}{2}$$

$$sin(\theta) = \frac{e^{j\theta}-e^{-j\theta}}{2j}$$

So, full reconstruction is possible with cosines when the function is even and with sines when the function is odd.

(Refer Slide Time: 13:03)



Now, let us consider the discrete representation. So, a discrete representation of a function can be made in the following form :

$f(n) = \{f(nX_0)|n \in Z\}$

that the function needs to be sampled at periodic interval and it will provide a sequence of functional values where each sequence position is an integer set and sampling interval which is associated with this particular definition.

So, it can be also considered as a vector in an infinite dimensional vector space, but in our consideration since we will be always using images of finite dimension or the signals of finite dimension, you are representing them in the computation in the discrete domain.

So, there we will be having only a finite dimensional vector.

$f(n),\ n = 0, 1, ..., N-1$

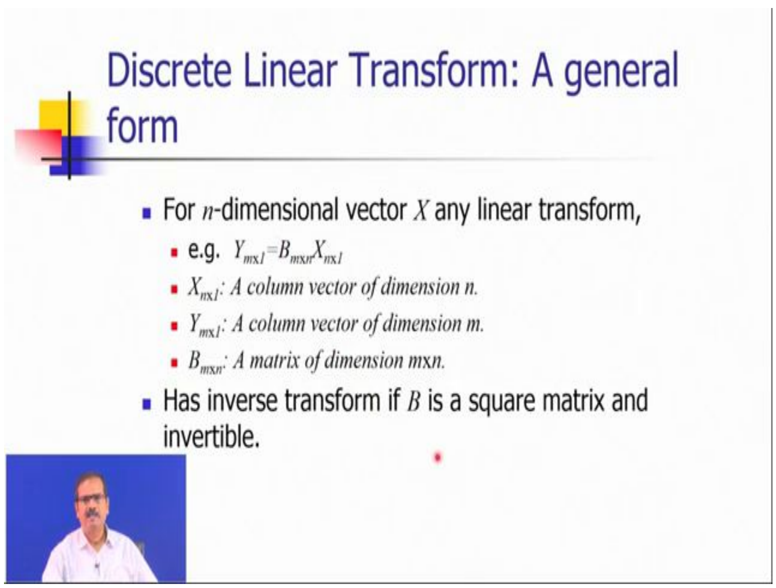
For example we can represent a function from within certain interval from say 0 to $N-1$ *as shown above.*

You should note here the sampling interval is implicitly represented in this form. So, even without sampling interval we have a representation of the function and when you are trying to interpret the function with the physical terms, physical quantities in the functional space then only the sampling interval has to be used.

So, let us consider that we are representing a function in this case with a finite dimensional vector and say it is an n dimensional vector in this case; it is a column vector representation. So, that is why the transpose operation has been used as shown below.

$$f = [f(0)f(1)...f(N-1)]^T$$

(Refer Slide Time: 15:05)



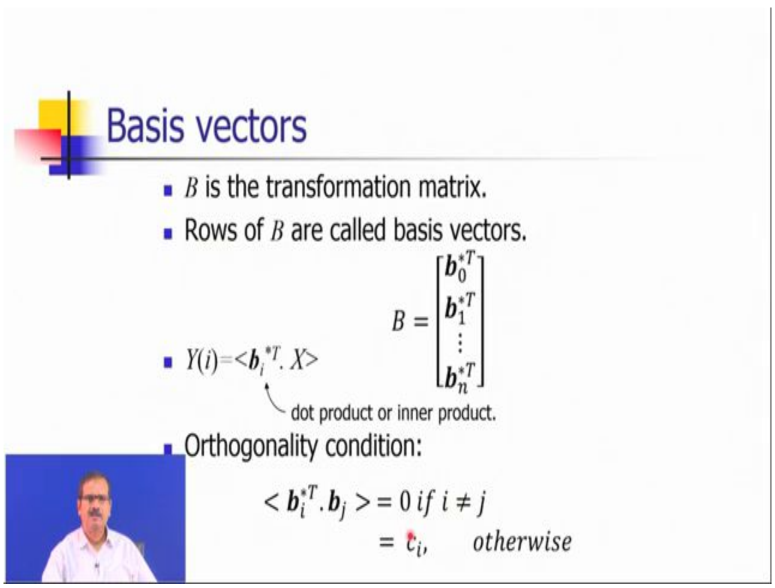
So, then how do you define a discrete linear transform? It is very simple because you know that whenever we perform any matrix multiplication in the following form:

$$Y_{m\times 1} = B_{m\times n} X_{n\times 1}$$

, say you have a column vector of n dimensional column vector and let there may be a matrix of $m \times n$ dimension and if I multiply them then you will get another vector of $m \times 1$ dimension, that means m dimensional vectors.

So, this is a transformation of this column vector into another column vector of a different dimension. we call it a linear transform or as it is discrete since we were using the discrete presentation, let us call it as discrete linear transform. And this transform has inverse when this matrix which is called say transformation matrix B is a square matrix and also invertible.

(Refer Slide Time: 16:11)



So, one of the interesting facts about this transformation matrix that we can note that we can consider rows of these transformation matrix B, they form the basis vectors, this is the analogy with respect to the basis functions. Because we will see instead of inner product between two functions we are having here, inner product of two vectors which is equivalent to the dot product of two vectors.

So, let us consider this say this is the representation of the transformation matrix and these are the row vectors which is indicated by the transpose operations and you can say that this is a row vector you are considering, these are the complex conjugate operations by keeping it consistent with representation what we have for the inner product.

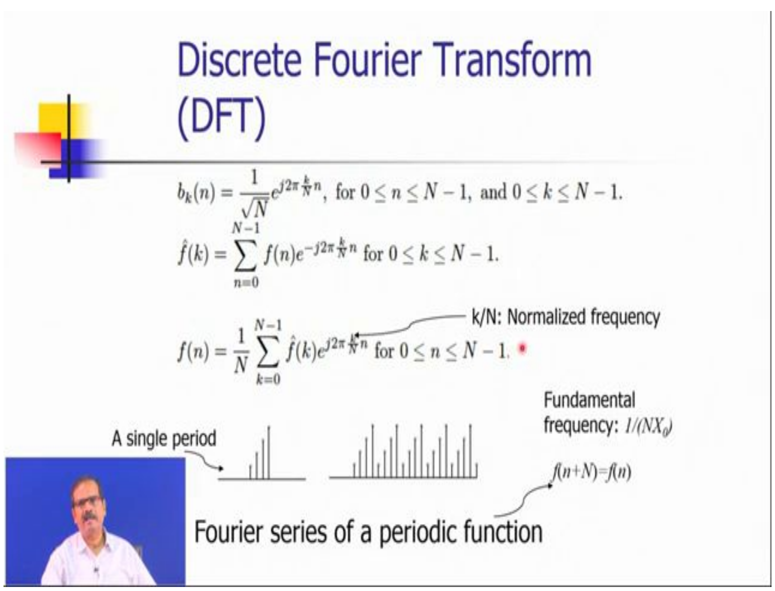
So, when we perform this dot product or inner product between two vectors, then you get the corresponding element. So, i^{th} basis vector provides you the i^{th} element of the transformed vector which is Y here.

$$< b_i^{*T} . b_j >= 0 \; if \; i \neq j$$

$$= c_{i,} \text{ otherwise}$$

So, the orthogonality condition in this case is reflected in the above form such that if you take any pair of two basis vector, then their inner product should be equal to 0 if they are different otherwise they should have a non zero value c_i.

(Refer Slide Time: 17:43)



So, with this form we can consider that it will have the similar representations. You can consider that a function as a linear combination of all those basis vectors. So, if I look at the discrete Fourier transform expressions, the basis vectors for the discrete Fourier transforms are represented in this form. So, it is $\frac{1}{\sqrt{N}}\ e^{j2\pi\frac{k}{N}n}$ you should note here this is defined for n functional points.

$$b_k(n) = \frac{1}{\sqrt{N}}e^{j2\pi\frac{k}{N}n} \text{ for } 0 \le n \le N-1, \text{ and } 0 \le k \le N-1$$

So, $b_k(n)$ is a small n^{th} element which means you can form a basis vector from by computing it $n = 0$ to $n = N-1$ at it at each integer value of n within these interval. So, that gives me a vector that is a k^{th} vector and there are n such vectors where the k indexes k is indexes vary from 0 to $N-1$.

$$\hat{f}(k) = \sum_{n=0}^{N-1} f(n)e^{j2\pi\frac{k}{N}n} \text{ for } 0 \le k \le N-1$$

So, forward transform or discrete Fourier transform of a discrete sequence $f(n)$ which is a finite sequence of length n can be expressed in the above form.

$$f(n) = \frac{1}{N}\sum_{n=0}^{N-1} \hat{f}(k)e^{j2\pi\frac{k}{N}n} \text{ for } 0 \le n \le N-1$$

You can find out that this is nothing but the inner product of $f(n)$ *and* $e^{j2\pi\frac{k}{N}n}$ .only thing is that instead of putting $\frac{1}{\sqrt{N}}$ in this expression ,we have taken care of the square root operation during the inverse transform by multiplying $\frac{1}{N}$ as shown above .This is a simple operation .

So, we kept this division normalization operation. We removed that operation in forward transform and included it in during that reconstruction. So, actually the value what we will get that would be proportional and it does not matter at this stage. But when you reconstruct, it will again recover the same value because, you are taken care of that particular normalization during reconstruction operation.

So, you can see that it is a linear combination of the corresponding of basis vector in the reconstruction also and the coefficients which is given by $\hat{f}(k)$ are the coefficients from discrete Fourier transform. We can also observe from the above expression that discrete Fourier transform is nothing but Fourier series of a periodic function. So, let us consider a finite sequence. So, in this case for simplicity let me take only four functional values and so, which means my value of N is 4 here and this is a functions for which I will be doing discrete Fourier transform.

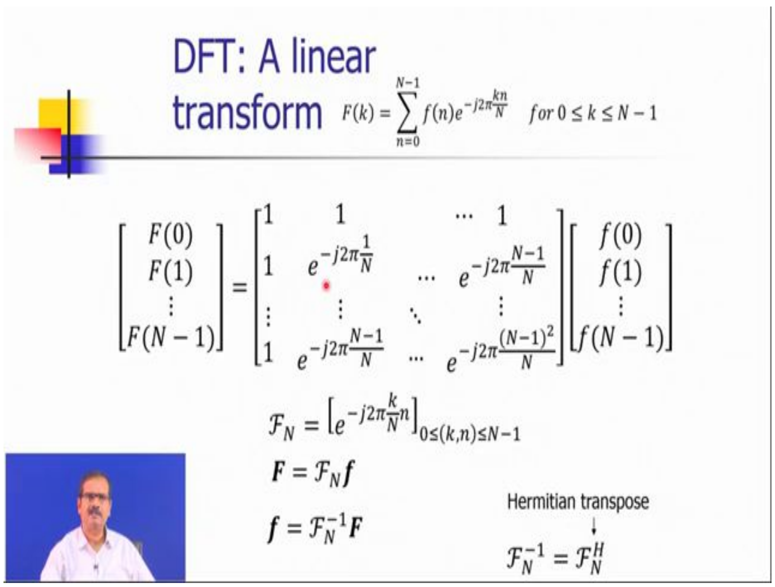
So, what I will consider because there is no definitions outside this interval, I can use any definition as per my convenience any other functional values and perform a transformation and after inverse transformation once again I will keep my observation window within the interval from 0 to 3 in this case. So, a periodic extension of this signal could be in this form that means, it is repeated so that you know in a periodic function this property needs to be a satisfied a periodic function with a period capital N should be $f(n+N)=f(n)$. So, that is satisfied as you can see here, here the value of N is equal to 4.

So, you will see after every fourth sample again it is repeating the same value. So, it becomes a periodic signal and as you know any periodic signal can be expressed as a linear combination of sinusoid functions and you can perform Fourier series ,that is ,what you are doing here in this case. After that, while in performing inverse Fourier transform , you are only performing for these four sample points. So, and just to note that how it is related with the actual physical sampling interval which is say X_0 here.

So, the fundamental frequency would be determined by the length of the period of this signal which is NX_0 So, fundamental frequency is $\frac{1}{NX_0}$.

And you can see that harmonics is represented by harmonic $\frac{k}{N}$ actual implicitly there is $\frac{k}{N} X_0$ that is the physical frequency. So, we call $k \times N$ as the normalized frequency in this representation.

(Refer Slide Time: 22:47)



So, discrete Fourier transform that can be also expressed in terms of a linear transforms which means we can express a discrete Fourier transform as a matrix multiplication of a column vector where the column vector is given by the functional values a finite dimensional vector as we have considered earlier and if I multiply these matrix then we will get the transform matrix and you can see that these elements they are obtained from this corresponding basis vectors definition.

So, we can represent this matrix in a shorter form where each k and N^{th} element is represented by the corresponding expression as shown below

$$\mathcal{F}_N = [e^{-j2\pi\frac{k}{N}n}]_{0\le(k,n)\le N-1}$$

and if the value of k and n they range from 0 to $N-1$ which will give me an $m \times n$ matrix. So, in a forward transform what we are doing; We are simply multiplying this transformation matrix with a column vector f which is representing the particular column vector $[f(0)\, f(1)\, ... f(N)]^T$.

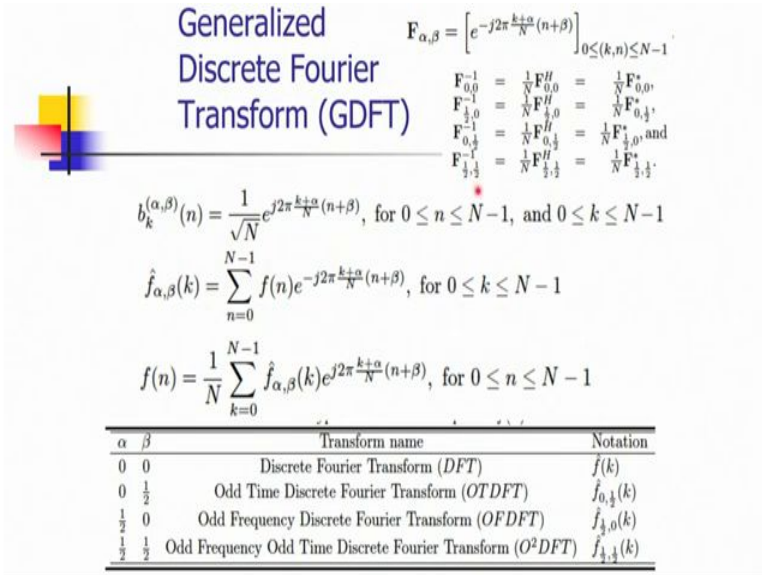
Then, we get the output transform matrix. So, the ($N \times N$) transform matrix transform column vector which is representing the column vector $[F(0)\, F(1) .... F(N-1)]^T$ which is actually the coefficients of discrete Fourier transform. $F = \mathfrak{F}_N f$.

and inverse transform will be naturally if I take the inverse of $F(N)$ and multiply with F then we will get back once again column vector. $f = (\mathfrak{F}_N)^{-1} F$

Incidentally because of the orthogonality property and also orthonormal property of these functions. So, you can show that the inverse of the discrete Fourier transform matrix is nothing but its Hermitian transpose of the corresponding matrix. $(\mathfrak{F}_N)^{-1} = (\mathfrak{F}_N)^H$

Hermitian transpose is the transpose of this matrix and also you should have to perform complex conjugate operation that would give you the Hermitian matrix.

(Refer Slide Time: 24:53)



Now, discrete there could be other kinds of know say it off orthogonal basis vectors and some of them could be derived or extended from discrete Fourier transform representation. So, we call it generalized discrete Fourier transform.

If you can observe that in the discrete Fourier transform ,basis vectors are generated by sampling the complex sinusoid within an interval between 0 to $N-1$ and then we have sampled at regular interval .Now if I make a phase shift there in that interval. So, there itself we can have a variation instead of we can give a phase shift of $\boldsymbol{\beta}$ and also while defining basis vector we have considered harmonics and we have generated harmonics at regular interval.

$$b_k^{\alpha,\beta}(n) = \frac{1}{\sqrt{N}}e^{j2\pi\frac{k+\alpha}{N}(n+\beta)} \text{ for } 0 \leq n \leq N-1, \text{ and } 0 \leq k \leq N-1$$

So, if I give also a shift in the frequency space ,then also you can have a different basis vector. Now the above representation will generate also n basis vectors of N dimensions and that would be once again an orthogonal basis vector which is a square matrix which could be invertible.

So, it could be easily used for once again for making a transform. So, this is the generalized discrete Fourier transform. We can use this basis vectors and we can get this expressions for discrete Fourier transforms. Similarly we can get back the function by applying the inverse Fourier transform it is the same similar form what we did for the case of discrete Fourier transform and the corresponding transformation matrix can be expressed in this form, here the elements as you can see it retains the same similar expressions only there are parameters α and β which is giving you a different set of transformation matrix.

There are some popular transformation matrix as you can see a special value of α at zero and β that itself will give you the discrete Fourier transform what we have discussed earlier. If I took $\alpha=0$ and set $\beta = \frac{1}{2}$, we call that transform as Odd Time Discrete Fourier Transform or OTDFT and you can represent the transform in this form; similarly $\alpha = \frac{1}{2}$ and $\beta = 0$ would be Odd Frequency Discrete Fourier Transform and if both are half; Odd Frequency Odd Time Discrete Fourier Transform.

There are different properties which I am not discussing here; just for your example we have given this particular thing and they have their inverse transform in this case also. Those are related and I have shown you in this particular grade.

So, there are different relationships of the inverse transform. I think let us stop here at this point and we can start from this point in the next lecture, where will see that though it is not possible in the continuous domain to have cosine transform and sine transform for every kind of function but in discrete domain for any finite dimensional sequence you can define cosine transforms and sine transforms.

So, for that we will be using this generalized discrete Fourier transform. So, thank you for listening this talk and we will move over to the next lecture.

Thank you.