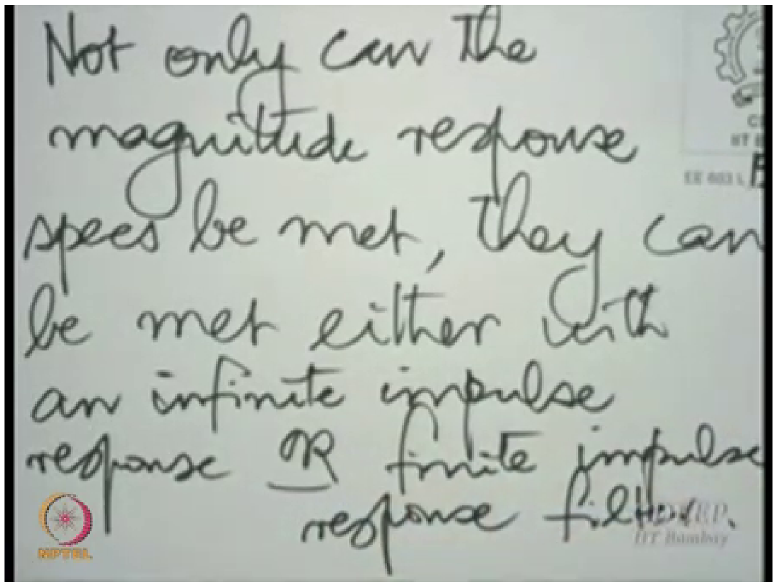
Digital Signal Processing and Its Applications
Professor Vikram M. Gadre
Department of Electrical Engineering
Indian Institute of Technology, Bombay
Lecture No. 19b
IIR and FIR filters; Approach to Filter Design

(Refer Slide Time: 0:29)



So, coming back to this, you see not only can we meet the magnitude response specifications with, you know, any δ_1, δ_2 and ω_p minus or transition band but not only can they be met, not only can the magnitude response specs be met, they can be met either with an infinite impulse response system or a finite impulse response system.

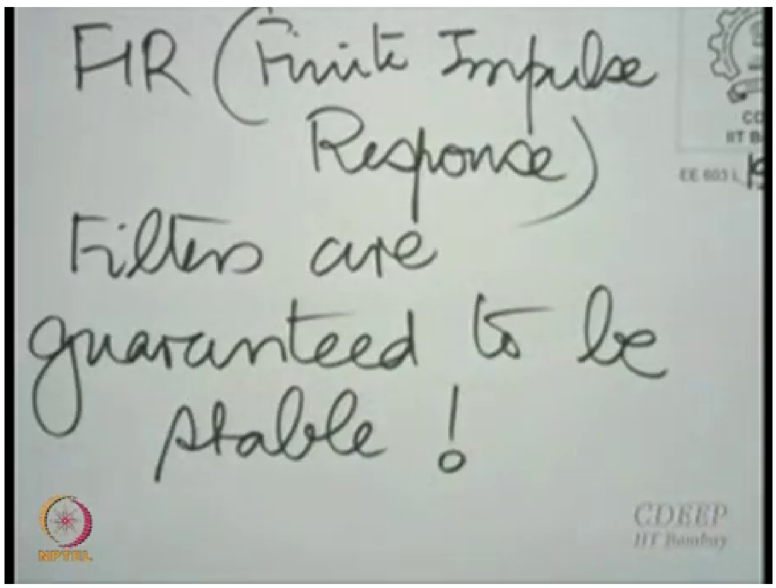
So, what is the infinite impulse response filter? Infinite impulse response filter is a filter whose impulse response has an infinite number of non-zero samples. No doubt they all occur on the non-negative side of n because you want the system to be causal but there are an infinite number of non-zero samples.

A finite impulse response system on the other hand has only a finite number of non-zero samples again on the non-negative side of n. Now you see we may spend a minute here in asking why we should look at both these possibilities. Why should you want to meet them either with an IIR or infinite impulse response filter or with an FIR or finite impulse response system? Why not just be happy with FIR?

You see the moment the system is finite impulse response many things are going to happen automatically; at least we will see a couple of them at the moment. The first thing that is going to happen automatically is that, the system is bound to be cause, it is bound to be stable because you can always, you see.

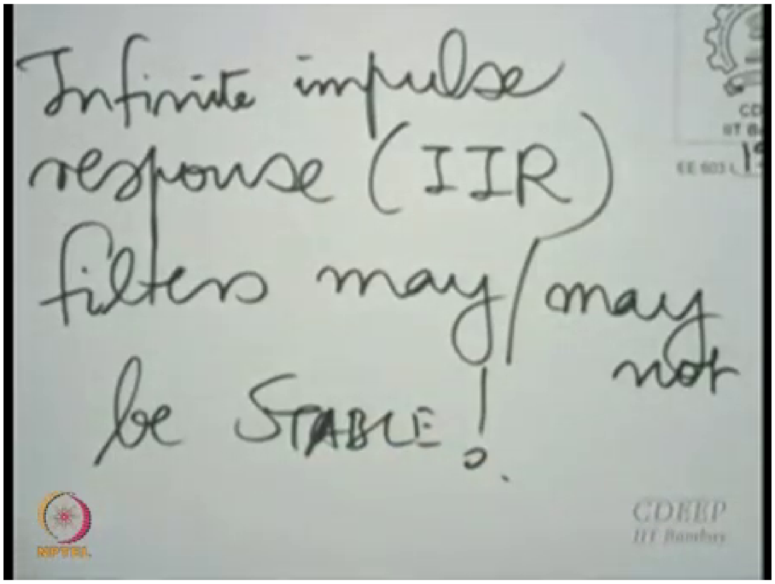
If you have a finite number of non-zero samples in the impulse response if each of these samples is finite in value finite that is all that you want and that of course you know is a reasonable assumption to make certainly. Then the absolute sum of the impulse response has to be finite.

(Refer Slide Time: 02:45)



So, stability is guaranteed. So, far systems are guaranteed FIR filters, FIR systems are guaranteed to be stable. That is very easy to see. That is not true with infinite impulse response systems.

(Refer Slide Time: 03:09)



Infinite impulse response systems may become unstable, may or may not be stable. So, why at all should we look at IIR filters as an option? Well, there are two reasons for it. One is you may call it historical or you may call it convenient. You see, you have a plethora of literature on analog filter design available to you, in fact, at the time when digital signal processing was evolving as a subject that could be said around the late 60s, the 70s and the 80s.

You see the 70s of the previous century and the 80s is the time when the subject of digital signal processing evolved to the maximum, filter design at least. The subject continues to evolve today, but filter design, the fundamentals of discrete time filter design, evolved to a maximum in the 70s and the 80s and of course, even now there are variants still coming in.

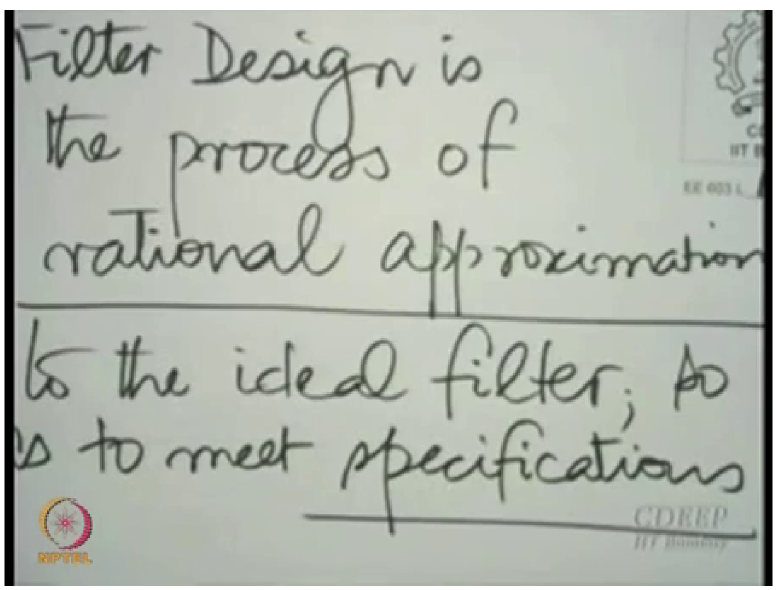
But during that period the initial approach was to take advantage of what was known with analog filter design. Now the beauty is analog filters can never have a finite impulse response. That is something interesting. Analog filters can never have a finite impulse response and therefore if you wish to take advantage of analog filter design; the easiest way to do it was to allow for infinite impulse response discrete time filters.

That was one important reasons why IIR filters are still of value. But that is not, you see that is essentially bound by time. One could have always argued that in time one should have got over that limitation and come up and done away with IIR. No, that is not. You see we will see later that IIR filters have advantages too.

For example, the computational demand the resource demand in IIR filters is very often less, much less, than it is for FIR filters to meet the same specifications. In fact, in the assignment on filter design that you will do in this course you will actually verify this. You will compare the resource requirement for the IIR filter that you design or the IIR filters that you design and the FIR filters that you design.

And then you would see why IIR filters still hold promise even though they could become unstable with numerical effects and so on. Anyway what we are now going to do is therefore to look at these two approaches to filter design. So, essentially what is filter design really? Filter design is an effort to make a rational approximation to the ideal filter right. So, let us define what filter design is.

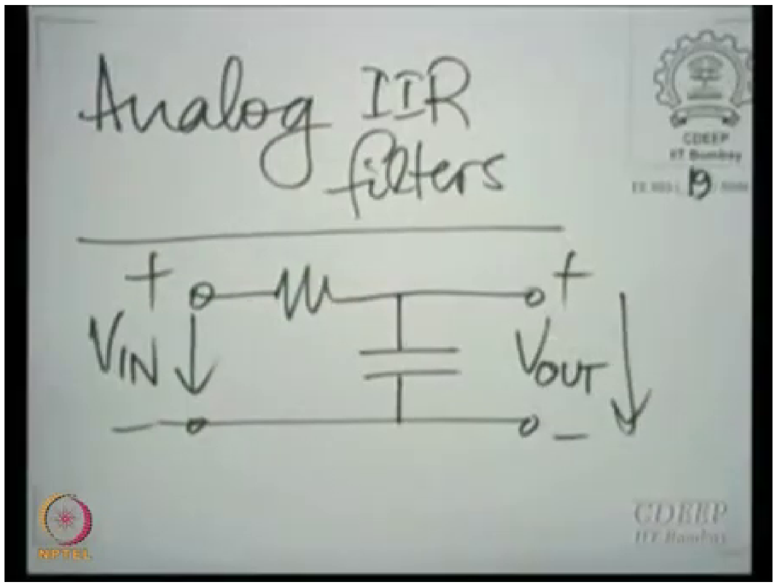
(Refer Slide Time: 06:18)



Filter design is the process of rational approximation to the ideal filter. Of course, the rational approximation meets certain specifications and what we are going to deal with is only magnitude specifications not phase specifications. In fact we will see later that as far as phase goes again FIR filters core over IIR filters. We shall see that too later.

But even so we do not do away with IIR filters in fact so much so just to proceed in the same way as a subject evolve we shall first look at how we could design IIR filters and then we shall see how to design FIR filters. Now what we are going to do is to take an approach to IIR filter design, where we take full advantage of the IIR system function in the analog domain. What kind of a system function do we have for IIR filters in the analog domain?

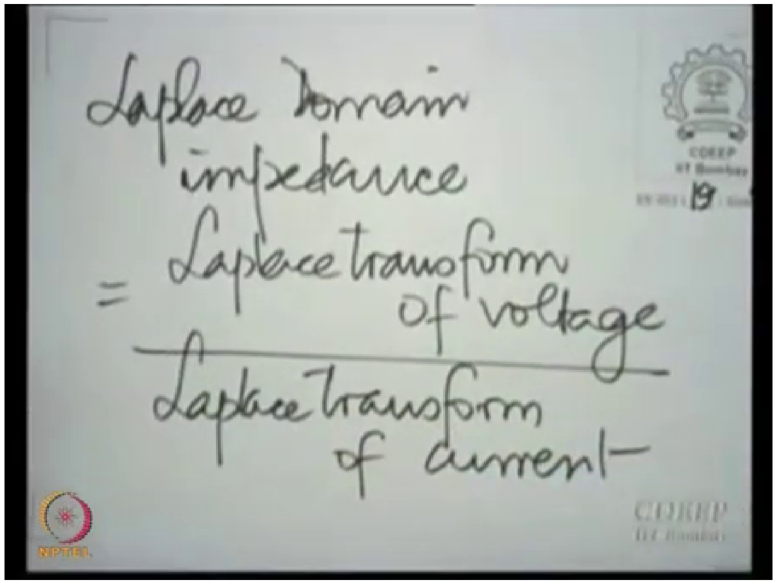
(Refer Slide Time: 07:43)



Let us take an example. So, you see let us take the very simple example of a resistance and a capacitance placed in series and you have the input applied here and the output voltage taken here. Now we know how to deal with this filter. What we do first is to write down what we call a Laplace domain circuit for this filter.

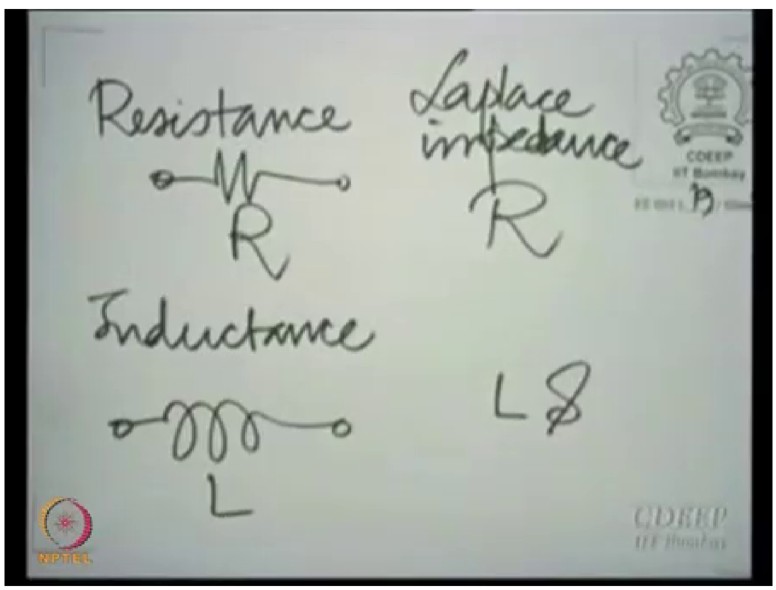
So, what we essentially do is we write down the ratio of the Laplace transform of the voltage to the current for each of the elements there we call it the Laplace domain impedance.

(Refer Slide Time: 08:34)



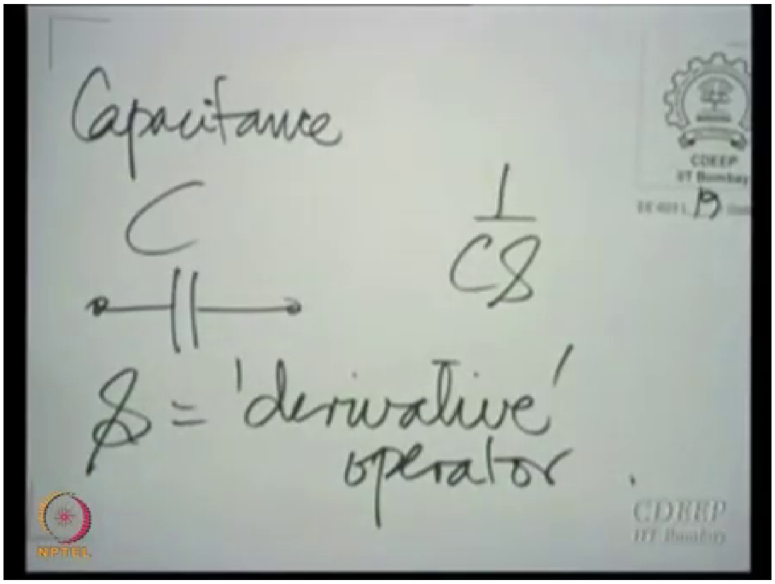
So, the Laplace domain impedance for each element is the Laplace transform of voltage by the Laplace transform of current in that particular element.

(Refer Slide Time: 09:11)



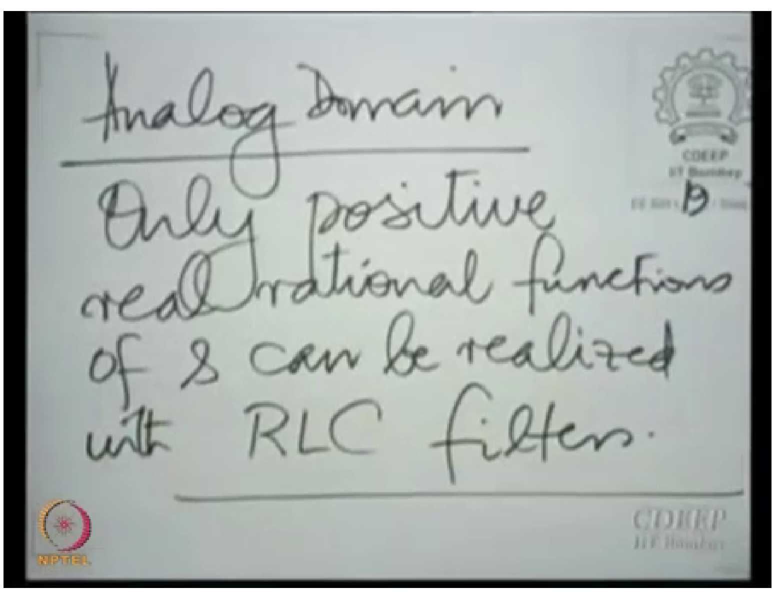
So, for a resistance, for example, of value R the Laplace domain impedance is just R. For an inductance of value L, the Laplace impedance is L times s. And remember s is the Laplace variable and there is a physical significance associated with s actually. Remember in an inductance the voltage is L times the derivative of the current. So, the process of taking the derivative is captured by the s. So, in a way the Laplace variable s captures a derivative operator.

(Refer Slide Time: 10:06)



So, there we are. In a capacitance of value C, where the Laplace domain impedance is $\frac{1}{Cs}$ and in all these cases s essentially captures the derivative operator and therefore one approach that we could take to move from an analog filter to the corresponding discrete time filter is to make a change on s.

(Refer Slide Time: 11:33)

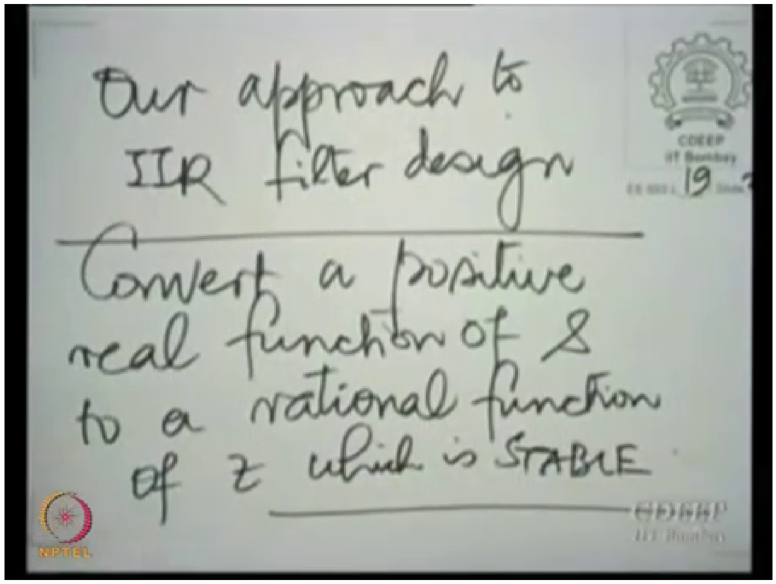


What is it that we want? We want a rational function in the analog domain. Remember again it is only rational analog filters, rational functions of s. That to a certain class of rational functions of s, that are realizable. At least if you want to realize them in terms of just resistances, inductances and capacitances then you read the rational function to be what are called positive real.

So, it is only positive real rational analog functions in s that are realizable with a combination of inductance, resistance and capacitance. Let us make a note of that. Now positive realness is something that we cannot quite deal with entirely in this course. I assume that some exposure may have been given to you in a course on network theory.

If you have had one but if you if you have not do not worry too much about it. Anyway let us make a remark. For the analog domain only positive real rational functions of s can be realized with RLC filters and therefore the approach that we would take is to convert a positive real function of s into a rational function of z.

(Refer Slide Time: 12:23)

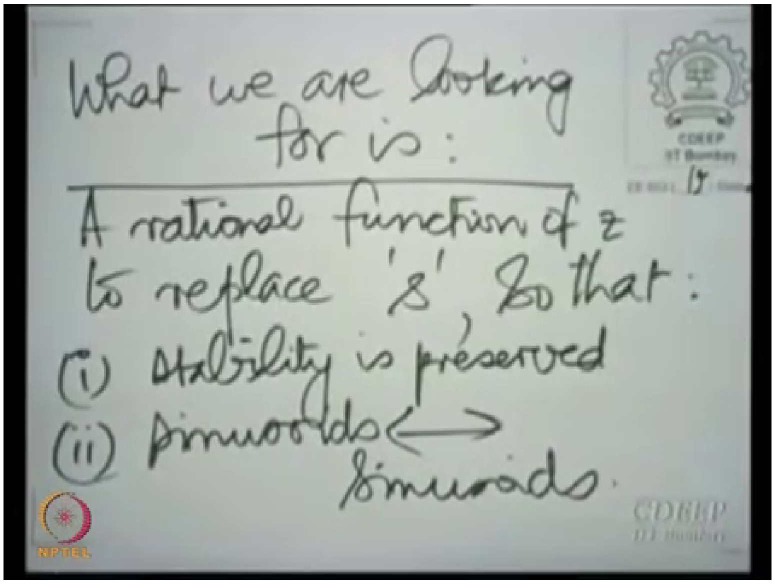


So, what is the first approach that we are going to take? The first approach to IIR filter design and in fact the main approach that we are going to take, convert a positive real function of s to a rational function of z, which is stable. Now here I make a remark. When the function is rational if you wish that it have a frequency response it has no choice but to be stable that is very easy to see.

If you want that it has a frequency response and that it be rational you want the unit circle to be included in the region of convergence. If the unit circle is included in the region of convergence the rational function has no choice but to be stable. In fact we saw in a couple of lectures before that the necessary and sufficient condition for stability as far as rational functions go is that the unit circle be in the region of convergence.

So, the moment you ask for a frequency response from a rational function you have no choice but to insist that it be stable, that means that when we go from the analog domain as a function in s to a rational function of z then you must maintain the stability that is there in the in the analog filter right. So, essentially it is a stability preserving.

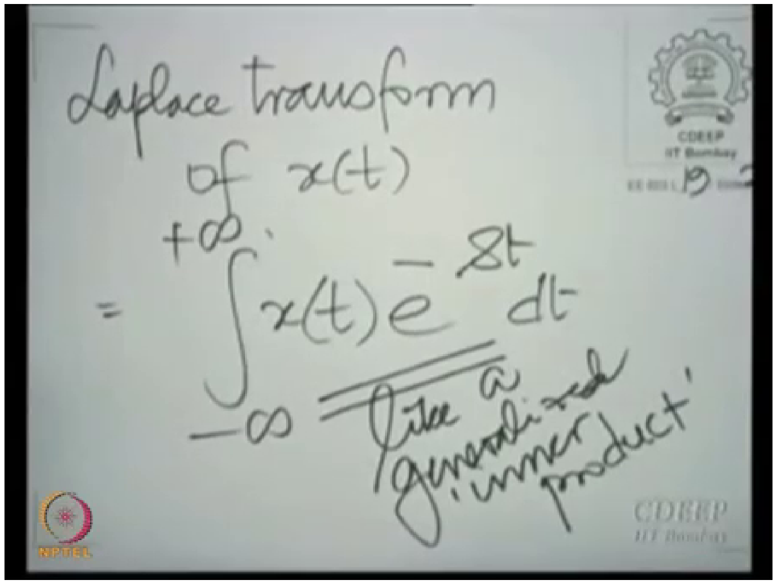
(Refer Slide Time: 14:04)



So, stability, so we want what we want? what we are looking for is a rational function of z to replace s so that the following happens. Point number one stability is preserved sine waves go to sine waves, sinusoids go to sinusoids. What does that mean?

What do you mean by sinusoids going to sinusoids? Well where do sinusoids come from in the analog domain? Sinusoids if you look at the Laplace variable s. It is when you substitute $s = j\Omega$ that you get sinusoidal response.

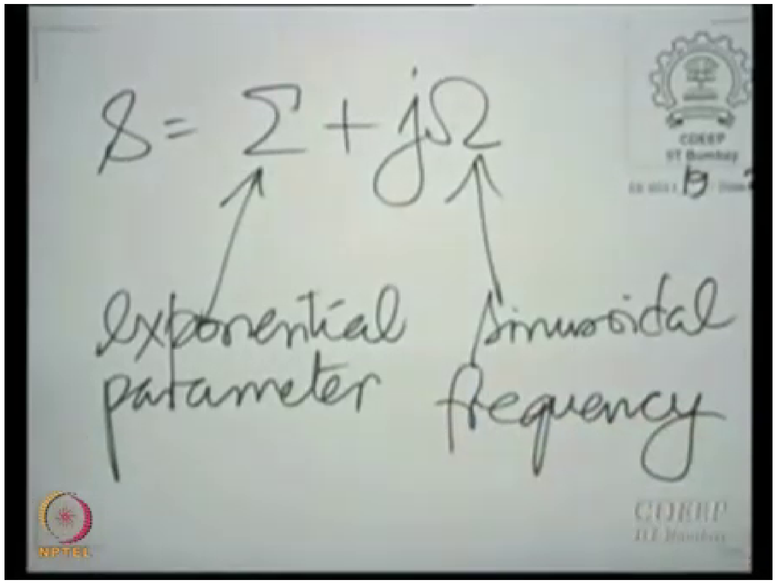
(Refer Slide Time: 15:20)



In fact if you just recapitulate what the Laplace transform is all about? The Laplace transform of a function is essentially $x(t) = \int_{-\infty}^{\infty} x(t)\, e^{-st} dt$. And essentially this is the generalization of the idea of a dot product, is like a generalized dot product or inner product, like we have seen before.

In some sense it is like finding the component of $x(t)$ along not a sinusoid but a sinusoid multiplied by an exponential and the real part of s gives you the exponential and the imaginary part of s gives you the frequency of the sinusoid. Let us make a note of that.

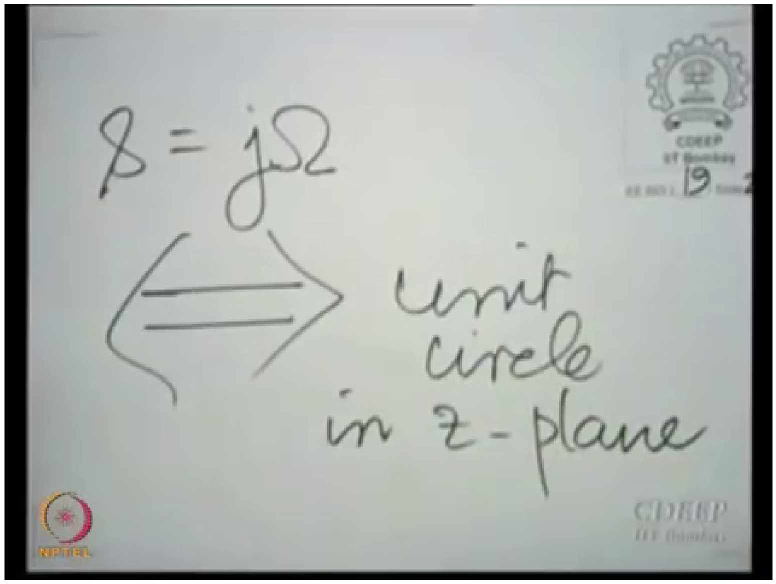
(Refer Slide Time: 16:23)



What we are saying is that if $s = \sum + j\Omega$ then this tells us the exponential part of the function exponential parameter and this tells us the sinusoidal frequency. Of course angular frequency if you please that is another way to interpret this just as you could interpret the Laplace transform as essentially being the derivative or derivative operator.

You could think of s as representing an exponential part and a sinusoidal part, the real part is the exponential part and omega the imaginary part of the sinusoidal part. So, if the exponential part is absent meaning if sigma is 0 then you have only sinusoids and that of course you see to be the Fourier transform in continuous time.

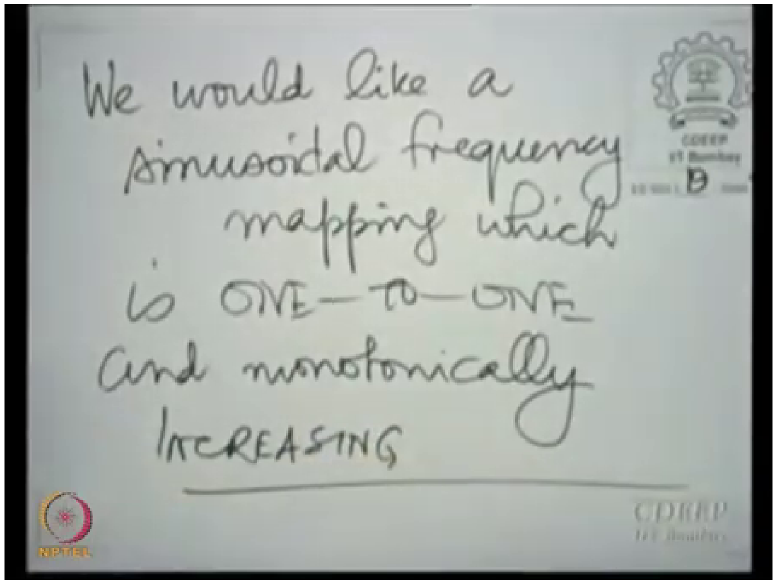
(Refer Slide Time: 17:23)



So, putting $s = j\Omega$ what we are saying in effect is that $s = j\Omega$ should correspond to the unit circle in the z plane. That is what we mean by sinusoids going to sinusoids. You see what we want is that once you have a frequency response that you have built in the analog domain you want to be able to carry that frequency response albeit with some changes or with some distortion of the axis into the discrete time domain.

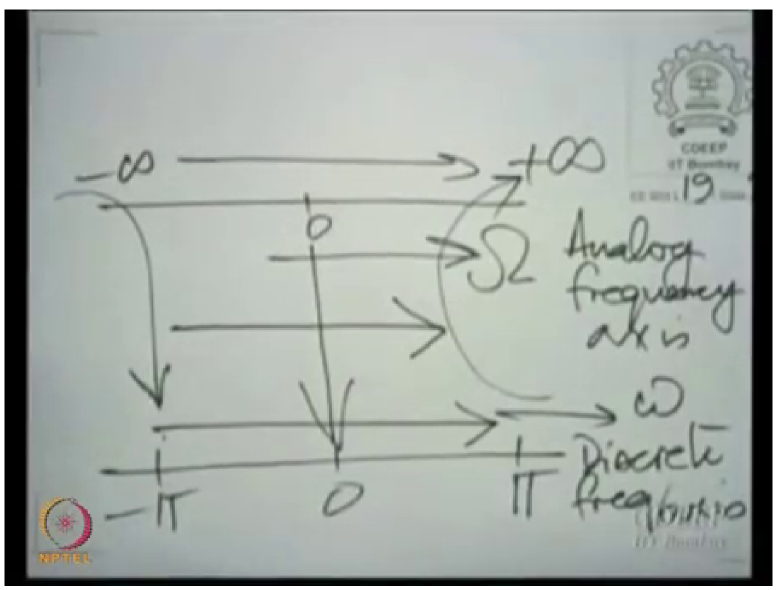
Now there is a third thing that you also want you see. You want to make your life easy. You do not want an analog high pass filter to become a discrete time low pass filter. If you want to design a discrete time low pass filter then you would like to start with an analog low pass filter. So, you would like the nature of the filter to be carried as it is from analog to discrete time. For that it is not adequate that sinusoids go to sinusoids. Not only must sinusoids go to sinusoids but there must be a one-to-one and increasing relationship between the sinusoids.

(Refer Slide Time: 18:36)



What I mean by that is we would like a sinusoidal frequency mapping of the following nature which is one to one and monotonically increasing.

(Refer Slide Time: 19:17)



Let us explain what we mean by this. You see as you traverse the analog frequency axis from minus infinity towards plus infinity you would like to traverse the discrete frequency axis from $-\pi$ to π . So, you would like $-\infty$ to come to $-\pi$, you would like plus π to be essentially a mapping of plus $\infty$to map to 0 and there to be an increasing character in between.

So, as you move from $-\infty$ to ∞ on the analog frequency axis you would like to move monotonically from $-\pi$ to π on the discrete frequency axis this is what you would want. If you do not have this the nature of the filter can change. So, high pass filter could become a low pass filter.

For example, suppose you reverse the frequency traversal and you are going in the other direction. Suppose in going from 0 to ∞ you are going from π to 0. You are reversing the nature of the filter. You do not want that. Obviously a mapping which takes you from an infinite axis to a finite axis must be non-linear and must also distort the frequency axis. We shall see more about this in the next lecture.