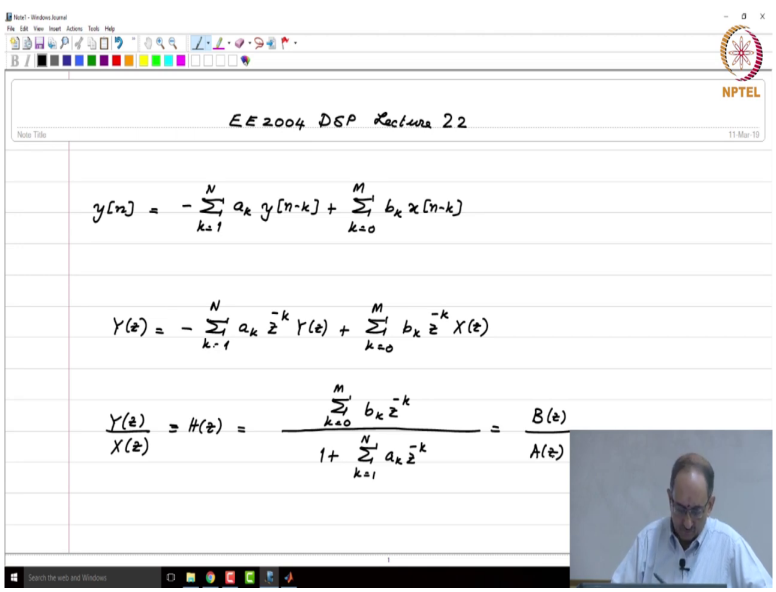
Digital Signal Processing

Prof. C.S. Ramalingam
Department Electrical Engineering
Indian Institute of Technology, Madras

Lecture 49:
Magnitude Response (1)
Geometric interpretation of the frequency response

All right let us continue our discussion on frequency response. We just got started with the definition of frequency response. We made the statement that if the region of convergence contains the unit circle, you can evaluate the Z-transform along the unit circle. And we also made the observation that if the region of convergence contains the unit circle, then the impulse response is absolutely summable and we are interested in the class of causal and stable systems.

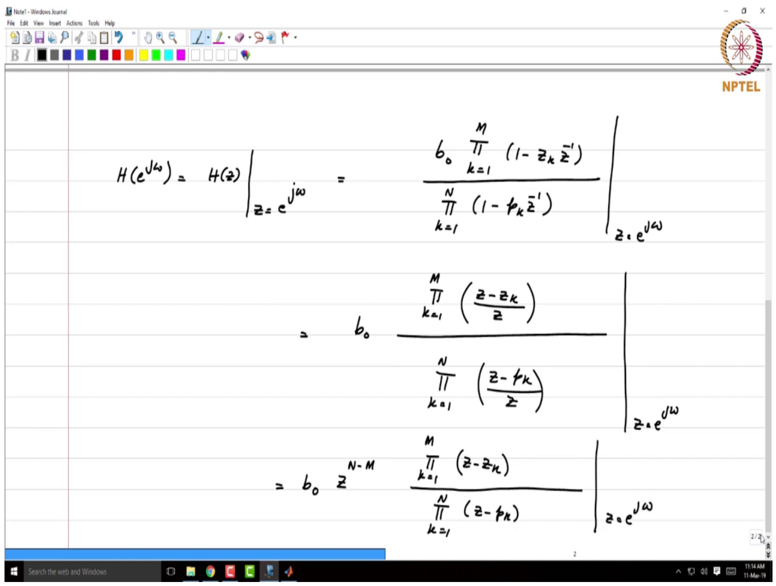
(Refer Slide Time: 00:51)



And further. we will focus our discussion on the class of systems whose input-output relationship is given by linear constant coefficient difference equation. We are focusing on this class, because that is the only class that can be realized in practice. So, $y[n] = -\sum_{k=1}^{N} a_k y[n-k] + \sum_{k=0}^{M} b_k x[n-k]$.

So, this is the basic input-output relationship that belongs to the class of LCCDE and now what we will do is, we will look at the system's transfer function. We have already looked at the system transfer function for this class, we will just go through the steps quickly and recall the result. So, this is $Y(z) = -\sum_{k=1}^{N} a_k z^{-k} Y(z) + \sum_{k=0}^{M} b_k z^{-k} X(z)$ and hence $\frac{Y(z)}{X(z)}$, this we denote as $H(z)$ which is the system transfer function and this turns out to be, $1+\sum_{k=1}^{N} a_k z^{-k}$ is the denominator polynomial and the numerator of course, is $\sum_{k=0}^{M} b_k z^{-k}$ and then we have seen that this is written in the form $\dfrac{B(z)}{A(z)}$.

(Refer Slide Time: 03:10)



So, this is the transfer function corresponding to this class and as we have made this point before this is the class of rational transfer function and the frequency response is $H(z)$ evaluated at $z = e^{j\omega}$. And then you are going to look the frequency response of the above system, it is advantageous to write this in terms of the products of its poles and zeros.

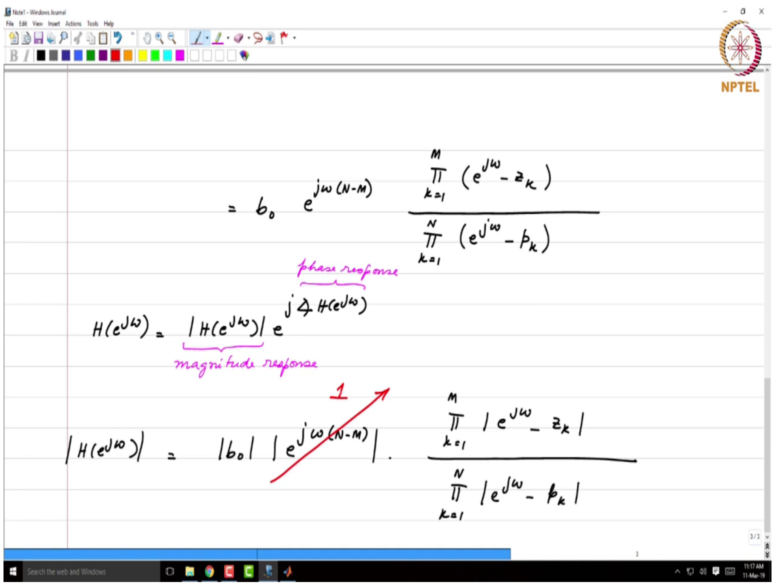
That is we will factor the numerator polynomial and the denominator polynomial. Therefore, the denominator becomes $\prod_{k=1}^{N}\left(1-p_k z^{-1}\right)$. All I have done this I have now expressed $A(z)$ in terms of its roots and the numerator is $\prod_{k=1}^{M}\left(1-z_k z^{-1}\right)$, where z_k are the zeroes and p_k are the poles.

This is not quite complete. To complete it, something more is needed which is all I have done, this I have written the numerator polynomial $B(z)$ in its product of its roots. I am sorry, this should be the $\prod$ product, ok. Now what is missing here? This is not quite what the previous expression is. To make it exactly equal, what is needed is b_0, very good. b_0 is needed, because without b_0 this if you look at this the z^0 term, the coefficient will be 1. Whereas, the constant term for this polynomial is b_0.

So, to make this equal, you need to have b_0 here. And now we are going to evaluate this along the unit circle. Before we do that, let us do some further simplifications here. So, this is b_0. So, this is the $\prod_{k=1}^{M}\left(\dfrac{z-z_k}{z}\right)$. Similarly, in the denominator, you have $\prod_{k=1}^{N}\left(\dfrac{z-p_k}{z}\right)$ of course, you are going to evaluate this along the unit circle.

Now if you look at this $\prod$ product, each term in the $\prod$ product has a denominator z. So, when you multiply this out, when you are done with, you will have z^M in the denominator. Similarly, here, you have z^N in the denominator of the denominator. Therefore, if you simplify, this will give rise to $b_0 z^{N-M} \dfrac{\prod_{k=1}^{M}(z-z_k)}{\prod_{k=1}^{N}(z-p_k)}$, all of this being evaluated along the unit circle.

(Refer Slide Time: 07:13)



So, this is b_0, now we will substitute $z = e^{j\omega}$. So, this becomes $e^{j\omega(N-M)}$ and this of course, is $b_0 e^{j\omega(N-M)} \dfrac{\prod_{k=1}^{M}\left(e^{j\omega}-z_k\right)}{\prod_{k=1}^{N}\left(e^{j\omega}-p_k\right)}$. So, this is the frequency transfer function and the frequency transfer function, it is useful to think of it being expressed in this form. So, this is $|H(e^{j\omega})|e^{j\angle H(e^{j\omega})}$ and no different from what you had seen in the continuous-time case.

So, this is the magnitude response and this of course is the phase response and we will study each of these parts separately. We will first focus on magnitude response and then on phase response and hence if you look at the magnitude response, this now becomes $|b_0||e^{j\omega(N-M)}|(\ )$. Magnitude of the product is product of the magnitude so, this becomes $\displaystyle\prod_{k=1}^{M}|e^{j\omega}-z_k|$ and in the denominator, you have a similar term $\displaystyle\prod_{k=1}^{N}|e^{j\omega}-p_k|$ and $|e^{j\omega(N-M)}|$ of course, is 1 and if you look at this term whose magnitude is 1.

That factor is coming from z^{N-M} and depending on the values of $N-M$, z^{N-M} will be either a trivial pole or a trivial zero. If $N > M$, this will be a trivial zero, otherwise will be a trivial pole of order magnitude of $N-M$. And hence you see that, this factor which is contributed either by trivial poles or trivial zeros as the case may be, they do not contribute to the magnitude response and this is precisely the reason why also zeros at the origin are called trivial poles are trivial zeros, because they do not contribute to the magnitude frequency response. However, later we will see that they will indeed contribute to the phase response.

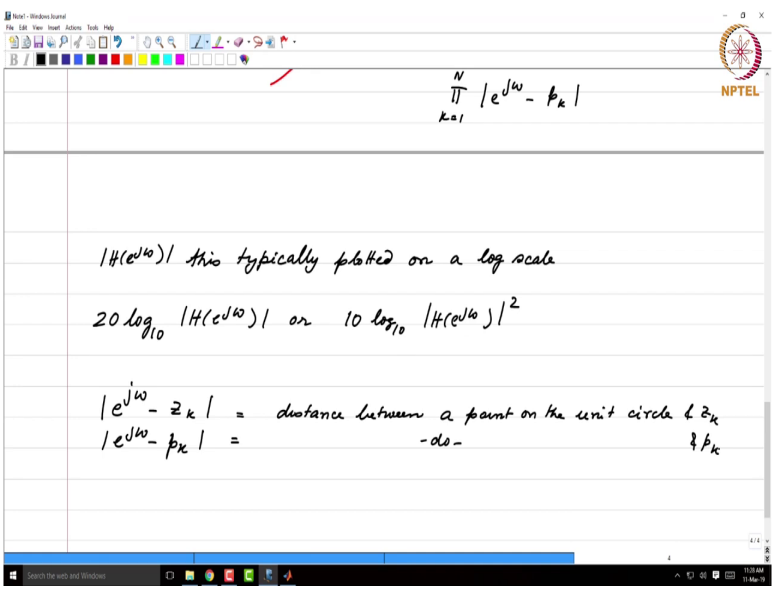
You cannot have a term like z^{N-M} and not contribute to both magnitude and phase. So, it does not contribute to the magnitude, but it will certainly contribute to the phase. And this is just the overall gain term and when we plot the magnitude frequency response, will either plot, you can look at this as either being the magnitude transfer function or the magnitude square transfer function.

But typically, when you are looking at frequency response, how do you plot them? You plot versus frequency. Do you plot the magnitude or magnitude squared or do you plot something else? You plot the log magnitude and log magnitude, typically the scale is what is called as the dB scale. So, why is it that you are going to plot the frequency response and log magnitude, why do not you see this plotted on the linear scale?

So, typically you are, you are told that if you have a large range to cover, to see that you will use a log scale. But for example, if this were a filter and if this were a passive filter, then the maximum magnitude gain will be? For a passive filter, the magnitude gain maximum will be 1 and the minimum gain will be 0 and 0 to 1 is not a large range and yet you will find frequency response being plotted on a log scale.

No, I think when you are saying it approximated straight and you think of bode plot and that is log log. Here I am talking about the x-axis frequency being still linear, but the y-axis being log magnitude and typically it will be $20\log_{10}$ or $10\log_{10}$ magnitude squared. But, given that you have 0 to 1 not being a large range, why is it that you still want to plot this on a log scale?

(Refer Slide Time: 13:15)



Sorry, so this is typically plotted on a log scale and you have you either plot $20\log_{10}|H(e^{j\omega})|$ or $10\log_{10}|H(e^{j\omega})|^2$, same thing. Yes, somebody was saying something.

Student: (Refer Time: 13:45).

So, in the discrete-time case, your frequency response goes from π to π and if the frequency response if the impulse response is real valued, you need to restrict yourself only to the range 0 to π, that also is a finite interval. So, that is not a large range, ok. So, you are now looking at a later stage, but if I did not want to do any of that, if I just wanted to plot the spectrum, why would I plot it on a log scale rather than the usual linear scale?

So, all of this is coming later. So, if we want to look at this in terms of its contribution to each pole or zero, things like that. The reason why you are going to plot this on a log scale is if you are going to look at say the low pass filter and if you are looking at the stop band and if filter A has attenuation of say 30 dB and filter B has attenuation of say 60 dB, on a linear scale if you plot these two frequency responses, assume that both these filters have roughly the same pass band characteristics and the transition band let us say they are similar, but in the stop band is where the main difference is.

So, one is say 30 dB rejection, the other is 60 dB. And we plot this on a linear scale, in the pass band, the gain will roughly be 1 and in the stop band, you will not be able to see the difference between these two responses if you plot it on a linear scale. If you plot this on a linear scale, if the pass band gain is 1 and the stop band will look roughly 0, but you will not be able to see the difference between these two filters. These are three orders of magnitude difference and you will not be able to see this; 10 to the -3 versus 10 to the -6. You will not be able to see the difference, if you plot this on a linear scale.

Difference in the stop band of two filters that have rejections of 30 and 60, no way will you be able to see this on a linear scale. To see the difference, you need to see this plotted on a log scale for you to make out the difference. And just to give you a feel for what typical rejection is needed for a filter that is practical if you want to implement this filter and you want to get rejections that are meaningful, you have a sense of how much of rejection you should have in the stop band for it to be considered adequate or at least minimum acceptable.

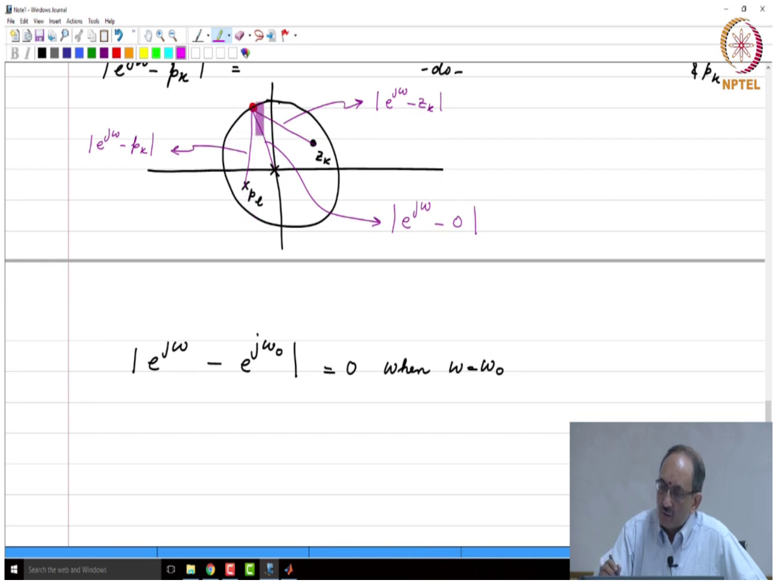
So, if I have a low pass filter, pass band gain is 1 which translates to 0 dB. Stop band gain, what kind of attenuation do you think might be needed in terms of dB for it to be acceptable? 60, no this is a reasonable guess, it is really large. So, you need at least 40 dB. Any filter that does not even give you 40 dB of rejection in most cases is not adequate. So, you need at least 40 dB rejection in the stop band for you to get adequate performance.

So, now, let us go back to this and focus on terms like this. So, this magnitude of b_0 once you consider the log plot, all this is going to do is its going to curve up or down, because this is going to jump out as a an added term and any change in magnitude of b naught will merely will shift the curve up or down. So, we will focus on this and we have already seen what this can be interpreted as geometrically.

So, magnitude of $e^{j\omega} - z_k$, geometrically when you look at the magnitude of 2 numbers, complex numbers a and b, magnitude of $a - b$ is the is the distance between the two points. So, in this case, it is now the distance between a point on the unit circle, because $e^{j\omega}$ is magnitude is 1 and as you vary ω, you are moving along the unit circle.

Therefore, one of the points is on the unit circle, the other point is z_k which is zero. So, distance between point on the unit circle and z_k, z_k happens to be the zero here. Similarly, $e^{j\omega} - p_k$ is similar, except here it is now p_k.

(Refer Slide Time: 19:34)



Let us explore this more. So, I have here z_k and then I may have another pole here and there could also be a trivial pole or a trivial zero. Here I am shown so, z_k should be marked with a filled circle because that is a zero, a cross is a symbol used for the pole and this is right now, I have shown a trivial pole, this could as well have been a trivial zero. Now here is a point on the unit circle.

So, this is the point on the unit circle therefore, what is happening here is $e^{j\omega} - z_k$, its magnitude is nothing but this line segment. Similarly, $e^{j\omega} - p_k$ is this therefore, this is $|e^{j\omega} - z_k|$. This is the $|e^{j\omega} - p_k|$ and this is nothing, but $|e^{j\omega} - 0|$ and hence as ω varies, the frequency response will vary.

So, you will start up with $\omega = 0$, which means you are on this point on the unit circle and then you are going to move along the unit circle, you will hit π and from π again you will hit 2π. If this were the pole-zero plot of a system with real valued impulse response, the poles and zeros will occur in complex conjugate pairs and the magnitude response will be even and hence you can limit yourself to looking at the magnitude response from 0 to π.

So, this is one part. The other part is the geometric interpretation of this frequency response. Therefore, you see that this frequency response can be interpreted as wherever you are on the unit circle, you find the distance from that point on the unit circle to each and every zero and then you take the product of all such distances. Similarly for the same point on the unit circle, find the distances to each and every pole and you take the product of all such distances and the ratio of these two numbers is the magnitude frequency response at that particular frequency and this point on the unit circle will vary from 0 to 2π.

Once you have this geometric interpretation in place; for example, when you are here on the unit circle, you are closest to this zero which means this distance will be very small. Therefore, the numerator which is the product of all such distances, the overall product will get pulled down, because it is going to get multiplied by this distance which is very small and hence the overall numerator will come down in value and therefore, you can expect that the frequency response will exhibit a dip when you are near a zero.

Similarly, suppose you are near a pole which means say you are here on the unit circle, then this particular distance will be very small. So, what is happening is the numerator is now going to happen to the denominator and hence the product of all the distances will get pulled down, because it is getting multiplied by this small factor and hence the denominator will get small and the overall transfer function will have a large value.

Therefore, this interpretation gives you a feel that zeros produce dips in the frequency response whereas, poles provide peaks in the frequency response. Therefore, zeros provide attenuation, poles provide gain. So, this is an intuition that you can develop from the geometric interpretation of this algebraic formula. Carrying this further, suppose this zero were on the unit circle, then you have this kind of situation; algebraically you have $e^{j\omega}$, here this ω denotes a general point on the unit circle and since the zero is on the unit circle, it can be represented as $e^{j\omega_0}$.

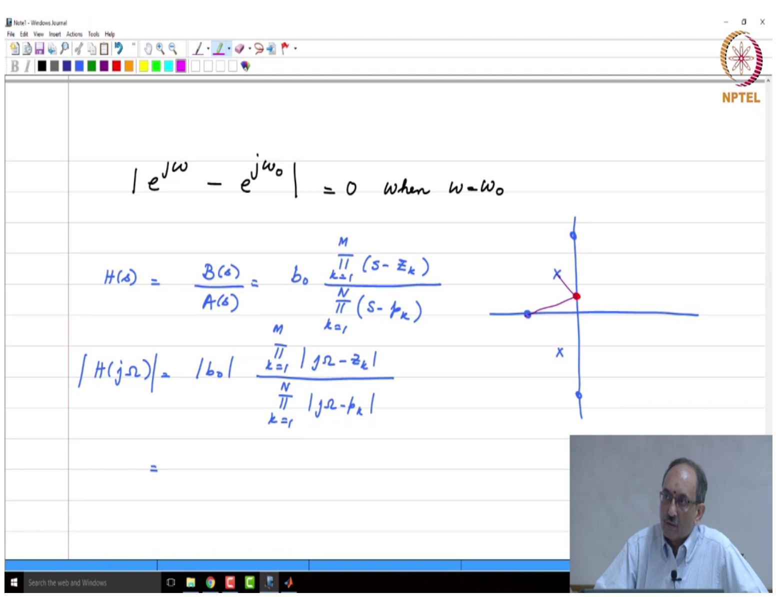
And now as ω varies, when $\omega = \omega_0$, this equals 0. Geometrically also this can be inferred. So, when you when a zero is on the unit circle, as you approach the zero, then you reach a point where you are sitting on top of the zero, because you are on the unit circle; the zero is on the unit circle. Therefore, that particular distance is 0 at that particular frequency.

And hence the whole product goes to 0 and hence the frequency response goes to 0 at $\omega = \omega_0$. Note that because we are looking at causal and stable systems, a pole cannot be on the unit circle. If the pole is on the unit circle, the region of convergence does not include the unit circle and hence you cannot quite evaluate the frequency response.

Therefore, while the zero can be on the unit circle and provide you 0 gain, you cannot have a pole on the unit circle to provide infinite gain. But pole can be as close to the unit circle as possible and it can provide very large gain. So, these are the general geometric intuition you can draw from this algebraic expression and the point about poles are zeros at the origin is, wherever you are on the unit circle, you are at a constant distance of 1 and hence they do not contribute to the magnitude response.

Actually, what we have seen is absolutely no different from what is happening in the continuous-time case. In the continuous-time case, you had $H(s)$ and this was of the form $\dfrac{B(s)}{A(s)}$ and your frequency response was $H(j\Omega)$ and what you did was, you evaluated the Laplace transform by replacing s by $j\Omega$.

(Refer Slide Time: 27:24)



So, this is nothing, but $\frac{B(j\Omega)}{A(j\Omega)}$ and this in turn can be written as some general gain by $b_0 \frac{\prod_{k=1}^{M}(s-z_k)}{\prod_{k=1}^{N}(s-p_k)}$. And you can think of your magnitude response exactly similar to what was done in the discrete-time case. So, this is now, you are going to replace s by $j\Omega$ therefore, this is $|b_0| \frac{\prod_{k=1}^{M}|j\Omega-z_k|}{\prod_{k=1}^{N}|j\Omega-p_k|}$.

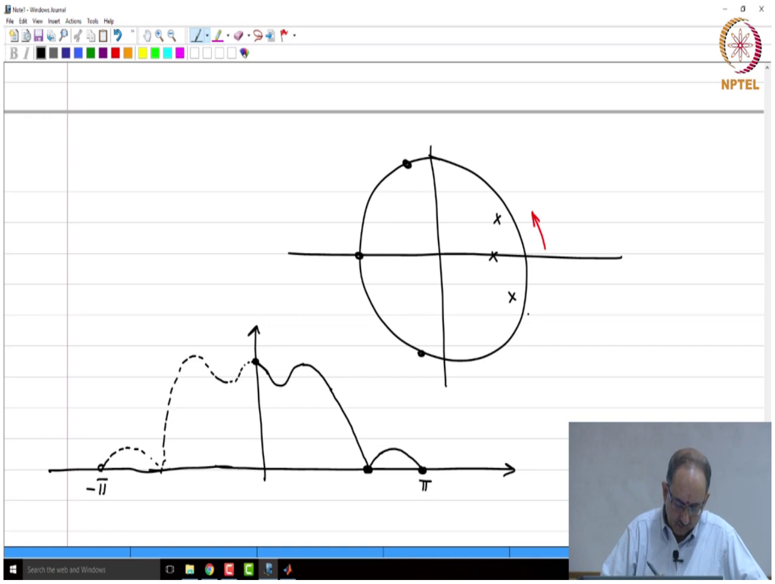
And this has exactly similar geometric interpretation as what we had just now discussed, only that rather than moving along the unit circle, you are moving along the $j\Omega$ axis. Therefore, if you have a pole here and say a zero here and a zero here and a zero here, if this way the frequency response, you are somewhere along the $j\Omega$ axis and the interpretation of the magnitude response is wherever you are on the $j\Omega$ axis, find out the distances to all the zeros, take the product and you will get one number.

Similarly, for the same point, find the distances to all the poles, take the product and you will get one number. The ratio of these two gives the magnitude frequency response except for this gain factor. If you ignore this gain factor, the shape of the frequency response is governed by the behavior in relationship to the poles and zeros. Exactly similar to what was happening in the earlier case; if there was zero on the $j\Omega$ axis, when you are going along the $j\Omega$ axis and you land on top of this zero, the frequency response will go to 0. If the zero were not on the $j\Omega$ axis, but close to it when you go past that zero, the frequency response will take a dip.

Similarly, if the pole were close to the $j\Omega$ axis, when you go past that pole on the $j\Omega$ axis, you will get a gain; you will get a peak. Therefore, poles and zeros shape the frequency response, be it in the continuous-time case, Laplace $H(s)$ or be it in the discrete-time case $H(z)$. So, this is the general intuition. Is there a counterpart to the trivial pole in the continuous-time case, finite s plane, this is a good guess by the way.

The answer is there is no counterpart to the trivial pole case. There it was at trivial pole or a trivial zero which was at a constant distance when you moved along the unit circle. The origin always was at a constant distance, is there a point here which is always at a constant distance as you move along the j omega axis? No therefore, there is no counterpart to the trivial pole or trivial zero case that was there in the discrete-time system. Such a concept does not exist in the continuous-time case. Now let us get further feel for this.

(Refer Slide Time: 32:07)



Once you have this geometric interpretation intuition with you, you can kind of guess the shape of the frequency response, rough plot. Suppose I had a pole here and here and suppose I had zeros like this. So, this is the pole-zero plot of a system, now what I want to do is I want to get a feel for the frequency response. So, what I am going to do is I am going to start off at this point, at $\omega = 0$. I will have a certain frequency response, then what is happening is, I am going to go along the unit circle like this.

When I am here and then I move away from $\omega = 0$, what I am doing is I am moving away from this pole. If I move away from the pole, the gain should, will it increase or decrease? It will decrease therefore; the gain will start to decrease and keep going along. When I keep going along, I encounter this pole which means when I approach this pole the gain should again start to increase.

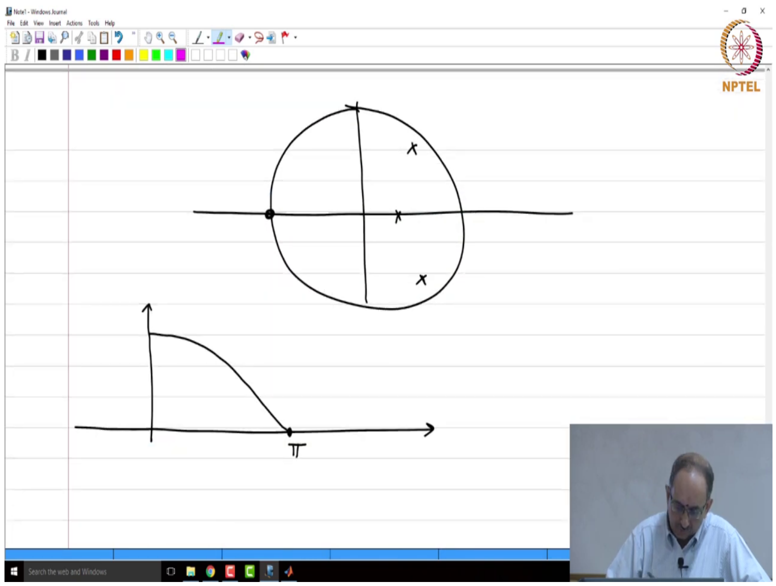
Therefore, the gain which is at this point started to decrease and then again it starts to increase when I am roughly in this neighborhood. And then what I do is I keep going along the unit circle which means I am going away from this pole, when I am going away from this pole the gain will start to become smaller, because I am going away from this pole. Therefore, what is going to happen is the gain starts to fall and then when I keep going along, I am now going to approach this 0.

This zero happens to be on the unit circle, therefore, what is going to happen here is response keeps falling down and when I am exactly on top of this zero, I hit a 0 in the frequency response. Then I go past this zero. When I go past this zero, the gain will increase and I keep going and when I hit $\omega = \pi$, the gain will again have to go to 0.

Therefore, the response will be like this and you can complete the frequency response. So, this is π so, I can you can complete the frequency response for the $-\pi$ to π part also. Let me so, this is $-\pi$ and π. One thing that I have assumed here is each pole is close enough such that when you are near that pole, it causes an individual bump in the frequency response. If you did not have this information, you cannot draw this plot.

The other thing that I am implicitly assumed in is all the heights are exactly the same. Since this is drawn by hand, it is not precisely the same, but the intent is all these peaks are exactly the same point even the dips, they have exactly the same dip value. So, this is what an equiripple response and there are filters that have this kind of frequency response that is equiripple, that is these ripples have the same height. Unless I mentioned that to you, you cannot draw this kind of plot. This is what is called an elliptic filter, all right.

(Refer Slide Time: 37:04)



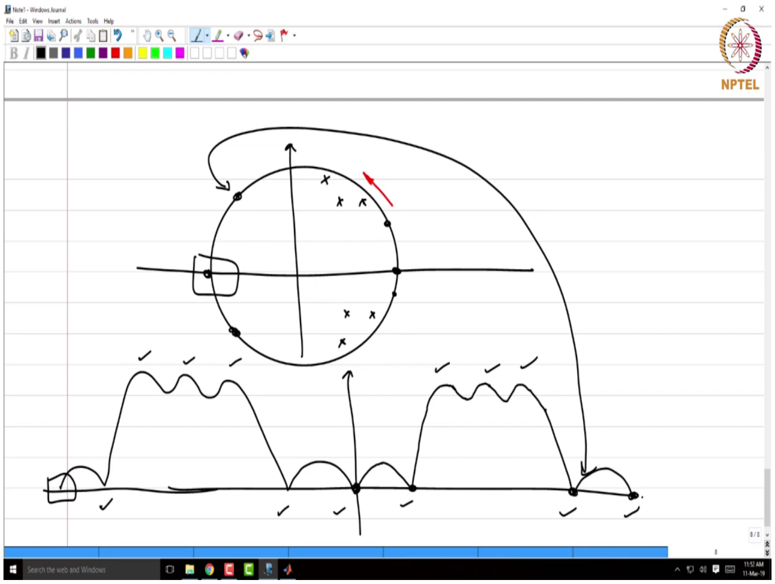
There is another kind of filter which is called as the Butterworth and let me draw the pole-zero plot of a third order Butterworth filter. And the point to note here is, these poles if you look at this plot compared to this these are clearly further inside and it so happens that if you plot the frequency response of this filter, the response will be something like this.

The placing of the poles is such that, when you start from here, the point at which you have the maximum gain, when you go along the unit circle you are going to move away from this, but this pole is further inside such that even when you approach it, it is not close enough to cause individual bump in the frequency response. So, the response monotonically decreases and falls down to π.

Therefore, this intuitive sketch based on this geometric interpretation this rough sketch, implicitly I have assumed that the poles are close enough to cause an individual bump and to know whether the pole-zero plot corresponds to something like this, that is the frequency response corresponding to this is this or is it this, you will not know unless you feed these numbers into say MATLAB and get the frequency response, all right.

But if I tell you that this is the pole-zero plot and if I also tell you each pole is close enough to cause individual bump and further if I tell you that the response is equiripple, then you should be able to come up with a plot like this. And this is a low pass filter, right this is a low pass filter, because it passes frequencies up to a certain point and then beyond a certain point it causes attenuation, all right.

(Refer Slide Time: 39:33)



And just to quickly continue on that low pass filter example, suppose I had a pole-zero plot something like this. Suppose I had something like this, then I had zeros wherever, then if I also tell you that this is the poles are close enough to cause an individual bump and if I also tell you that the response is equiripple, then you can easily come up with a response something like this.

So, clearly when you are at zero, the response is 0 and then when you move along, the response begins to rise and then when you hit this zero, the response will be precisely 0 and then when you go past this zero when you are in this region where the poles are, the response will begin to rise. And then when you are near this pole, you will have a bump and then you will go past this pole, the response will dip and then we are going to be close to this. So, this pole will cause this bump.

Similarly, the third pole will cause this bump and then when you go past this pole here, then the response will start to come down. And when you hit this zero so, this zero on the unit circle corresponds to this and then the next zero corresponds to this and this is a band pass filter. So, from the pole zero plot given information such as each pole is close enough to cause a bump and the response is equiripple given this pole-zero plot, you can easily come up with this. And there are 6 poles and there are 6 bumps; three from 0 to *pi* and similarly.

So, you have 1, 2, 3, 4, 5, 6; 6 poles, 6 bumps. And there are 6 zeros right; 1, 2, 3, 4, 5, 6, you do not want to count this twice, because this point is the same as this. So, this kind of gives you a very quick intuitive feel for the frequency response based on the pole-zero positions.