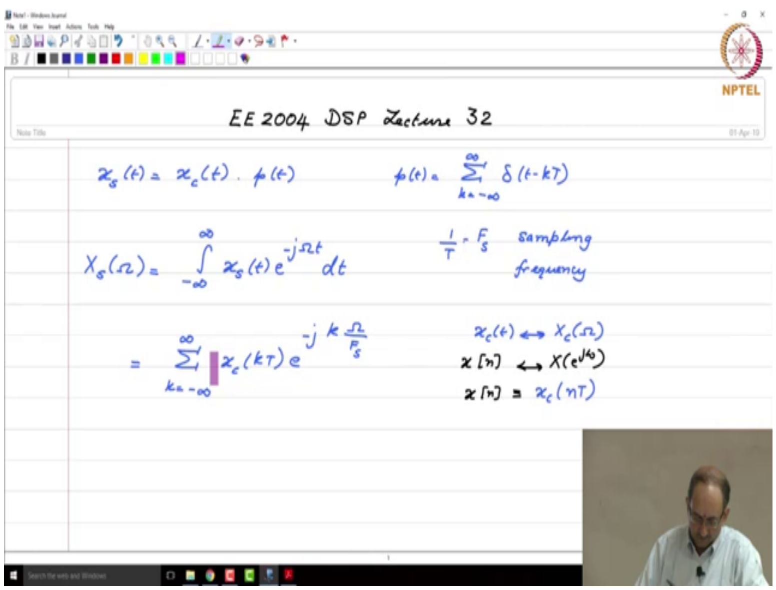
Digital Signal Processing

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

Lecture 69:
Sampling (2)
-Relationship between $X_s(\Omega)$ and $X_c(\Omega)$
-Aliasing
-Relationship between $X_s(\Omega)$ and $X(e^{j\omega})$

So, we got started with Sampling and the idea of sampling we started with impulse train sampling that is a theoretical construct.

(Refer Slide Time: 00:36)



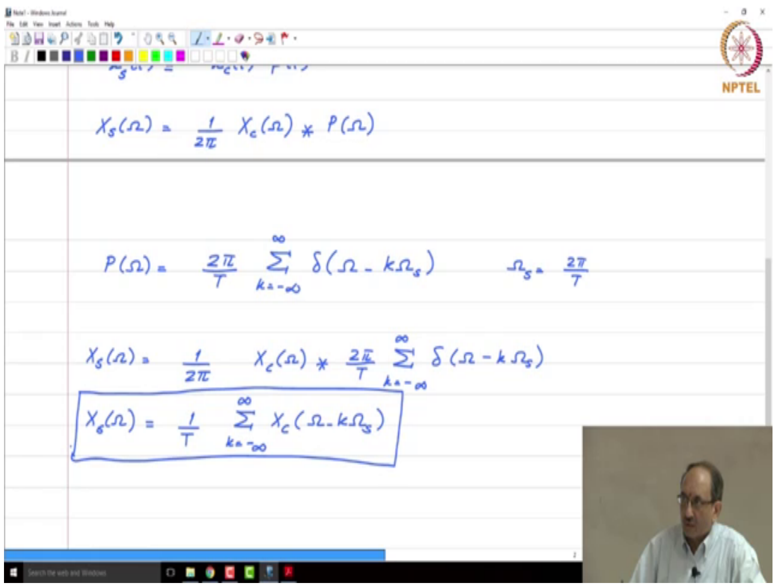
Therefore, given $x_c(t)$, we multiply by $p(t)$ which is an impulse train and this gives us the sampled signal and then we were looking at the continuous time Fourier transform of the impulse train sampled signal. Therefore, this is $\int_{-\infty}^{\infty} x_s(t)e^{-j\Omega t}dt$, and $p(t)$ of course is the impulse train.

So, this is $\sum_{k=-\infty}^{\infty} \delta(t - kT)$ is the impulse train; T is a sampling period, $1/T$ is called as F_s and this is the sampling frequency. And this we saw was nothing but, $\sum_{k=-\infty}^{\infty} x_c(kT)e^{-jk(\Omega/F_s)}$. So, this was the expression that was derived towards the end of the last lecture.

Remember, our goal is to relate the continuous time Fourier transform of $x_c(t)$. $x_c(t)$ has continuous time Fourier transform $X_c(\Omega)$, and the other sequence that we are interested in is $x[n]$, which has discrete time Fourier transform $X(e^{j\omega})$. And $x[n]$ is nothing but $x_c(nT)$. These are the values of the signal at the sampling instance. And our ultimate goal is to relate $X_c(\Omega)$ and $X(e^{j\omega})$.

So, we have come this far, but as you can see from this expression, on the right hand side, you have $x_c(kT)$ so, we do not have anything that relates or that brings in $X_c(\Omega)$. So, for this, we look at exactly the same spectrum, but from a slightly different view point.

(Refer Slide Time: 03:56)



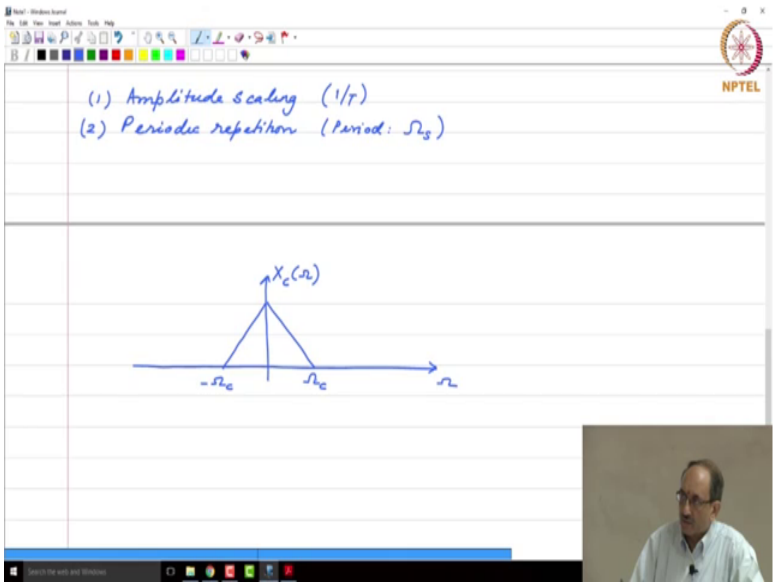
And the viewpoint we are going to look at is this, so we have $x_s(t)$ and this as before is $x_c(t)p(t)$ which is the impulse train. And, we also know that if you multiply in the time domain, you convolve in the frequency domain. Therefore, from the multiplication in the time domain relating to convolution in the frequency domain, the expression now becomes $X_c(\Omega) * P(\Omega)$.

So, $X_c(\Omega)$ is the Fourier transform of $x_c(t)$ and $P(\Omega)$ is the Fourier transform of the impulse train and if you recall the Fourier transform of an impulse train is also an impulse train. The main difference is the spacing. If the spacing is T in the time domain, the spacing will be inversely proportional in the frequency domain. Therefore, the Fourier transform of this is $\frac{2\pi}{T}\sum_{k=-\infty}^{\infty}\delta(\Omega - k\Omega_s)$, where T is a sampling interval. Delta, it is again going to be an impulse train, but it is going to be spaced Ω_s apart in the frequency domain, where $\Omega_s = \frac{2\pi}{T}$.

So, this is all from what you have learnt in the previous course and hence $X_s(\Omega) = \frac{1}{2\pi}X_c(\Omega) * \frac{2\pi}{T}\sum_{k=-\infty}^{\infty}\delta(\Omega - k\Omega_s)$. All I have done is, have just replaced the corresponding expressions for $X_c(\Omega)$ and $P(\Omega)$. So, the 2π cancels, so this becomes $1/T$ and then when you convolve something with an impulse train, it periodically replicates whatever you are convolving with. Therefore, this becomes $X_s(\Omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty}X_c(\Omega - k\Omega_s)$.

So, this is the formula we will focus on for now, draw some inferences before moving on to connecting $X_c(\Omega)$ and $X(e^{j\omega})$. Therefore, what this tells you is, this tells you that the impulse train sampled signal spectrum is a periodic repetition of the original signal spectrum. The original signal is $x_c(t)$ its spectrum is $X_c(\Omega)$ and now this is getting periodically repeated and the periodicity in the frequency domain is Ω_s.

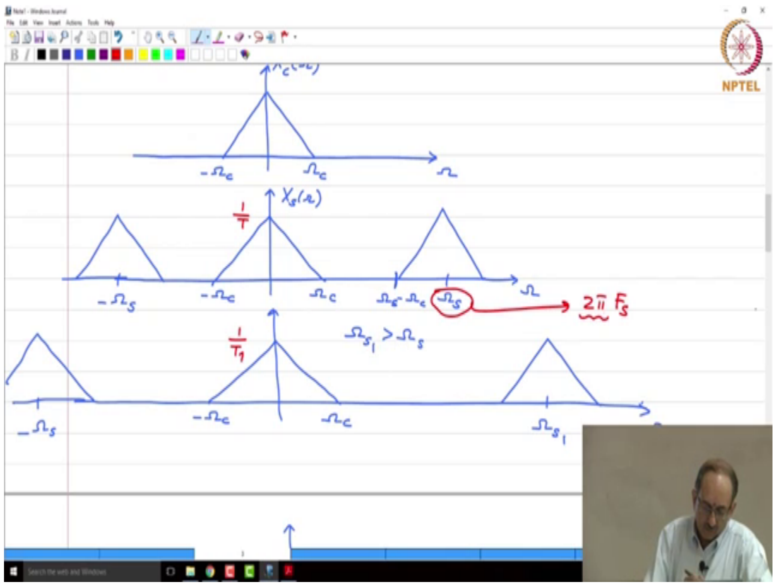
(Refer Slide Time: 07:35)



So, periodic repetition is one thing, the other important thing to note is not only is there periodic repetition, but there is also amplitude scaling and this amplitude scaling is $1/T$. And hence you have amplitude scaling and the scale factor is $1/T$. The other thing you have is periodic repetition, and the period of course, is Ω_s where $\Omega_s = 2\pi/T$. And remember, this can also be written as $2\pi F_s$ because $1/T = F_s$ and F_s has units of Hertz.

So, again this is all recap, this is all going over what you have learnt in sampling earlier and the picture that is normally associated with this is along these lines. So, you have $X_c(\Omega)$ and this is Ω and you have function something like this and this is between $-\Omega_c$ to $+\Omega_c$ so, this is the Fourier transform of the given $x_c(t)$.

(Refer Slide Time: 08:58)



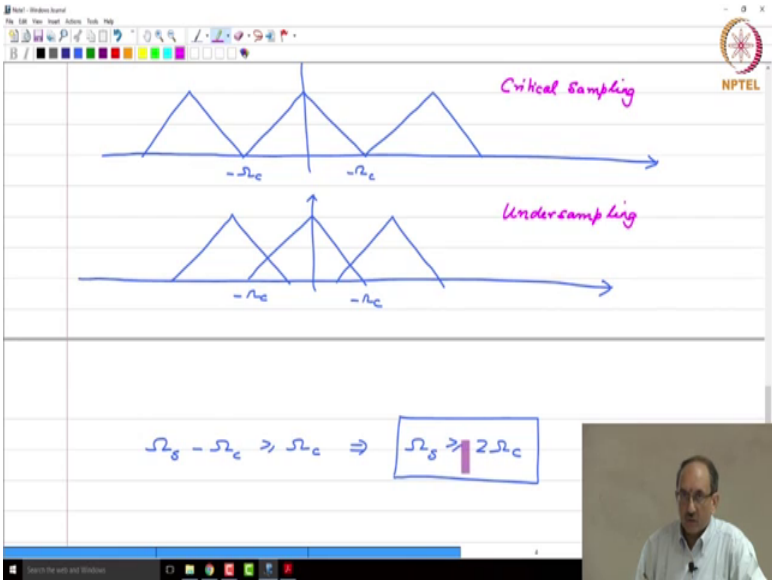
And, then what you have is $X_s(\Omega)$ which is the periodic repetition of $X_c(\Omega)$, the periodic repetition being Ω_s. So, I am just showing a couple of copies, the fundamental and the copy to the left and copy to the right. So, this is Ω_c, this is $-\Omega_c$.

The other important thing that you have to also keep in mind is the scale factor of $1/T$. And hence, if you now have another sampling frequency and suppose this sampling frequency is Ω_{s_1}. And Ω_{s_1}, suppose if it is larger than Ω_s, then the periodic repetition will occur later. So, this will be minus Ω_{s_1} and this of course, will be $1/T_1$.

So, the amplitude scale factor goes up, the repetitions occur later. Again these are not to scale I have shown them by the same height, but what this really meant is the heights are also different. And just to complete the picture for one more case. So, this is $\Omega_{s_1} > \Omega_s$ and immediately, you can also guess that if you now have a sampling frequency that is smaller say Ω_{s_2}, then the repetitions will occur sooner and this will be scaled by $1/T_2$.

So, all these X_c are, Ω which is in radians per second. So, you have this signal, its sampled version, periodic repetition will be there, amplitude scale factor will be there and the two plots below show the behavior for two different sampling frequencies; one that is larger than the first, the other that is smaller.

(Refer Slide Time: 11:52)



And hence, immediately this brings about this picture. So, the repetitions will occur earlier or later depending upon whether the sampling frequency is lower or higher. And, immediately you can see that if the sampling frequency is just about what is considered minimum, the repetitions will just abut each other. So, this is Ω_c, this is $-\Omega_c$. So, this is when the sampling frequency is what is called critical sampling frequency, that is, they just touch each other. And the other case of course is under sampling. So, you have $-\Omega_c$ to $+\Omega_c$, you have not sampled it enough. So, that the repetitions occur so much sooner that the overlap and this is what is called under sampling.

Again, in these last two plots, I have not shown the corresponding scale factor, scale factor is implied, here I am focusing on the overlap part. So, from this picture, you see that this point would be $\Omega_s - \Omega_c$ and you want this point to be at best equal to Ω_c or greater, only then will this repetition occur later. Therefore, if you do not want to overlap, then it is easy to see that you need $\Omega_s - \Omega_c$ to be greater than or equal to Ω_c. So, this implies that $\Omega_s \geq 2\Omega_c$.

And the other thing that is implied from the picture that we have drawn is, clearly we have assumed the signal to me band limited, that is why this frequency response is between $-\Omega_c$ and $+\Omega_c$ and there are no other components beyond this frequency range. So, assuming you have a band limited signal and as I mentioned earlier this signal is low pass, so what we are actually discussing is the low pass sampling theorem. You should be aware that there is also a corresponding band pass sampling theorem which you are not going to look at. So, for the low pass case, signals assumed to be band limited and in such cases, if you do not want overlap, this is what has to be satisfied. And,

Student: Sir.

Yes, question.

Student: So, I think we care about low pass band limited (Refer Time: 15:40) band limited add some and also higher frequency, (Refer Time: 15:45) repetition (Refer Time: 15:47).

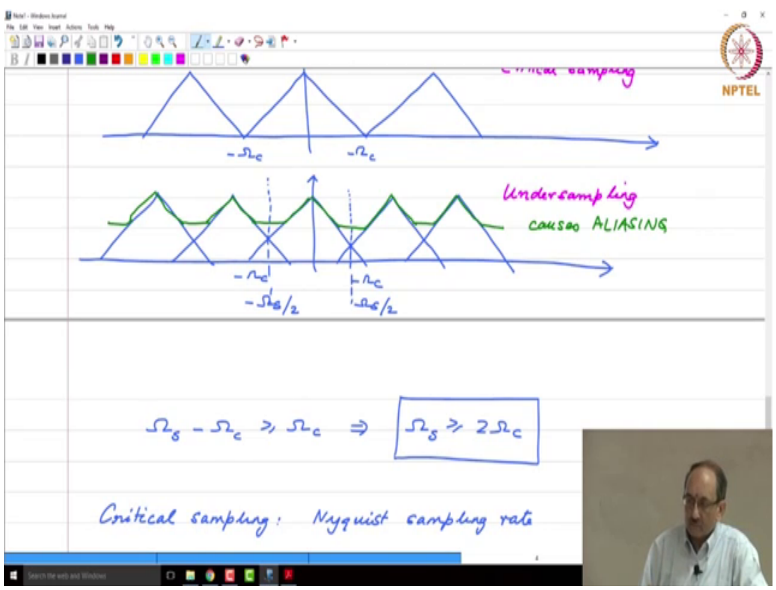
Yes.

Student: So.

So, what will happen in when the signal is band pass, there also as you rightly say, there will be repetitions and then you can come up with a version of sampling theorem such that the sampling frequency required is not twice the highest frequency; that is the point. The point is here what you are inferring is, you are inferring that the, for distortion to not happen, you need to sample at least at twice the highest frequency.

Even the signal where band pass, if you treat it as a low pass signal with the highest frequency being whatever the frequency is there is a spectrum, then you would end up sampling it at a far higher frequency than needed. Whereas, if you look at the band pass sampling theorem, you will infer that you do not need to sample it at twice the highest frequency but something much lower, that is the inference that will be drawn. And this critical sampling frequency is called the Nyquist frequency.

(Refer Slide Time: 17:02)

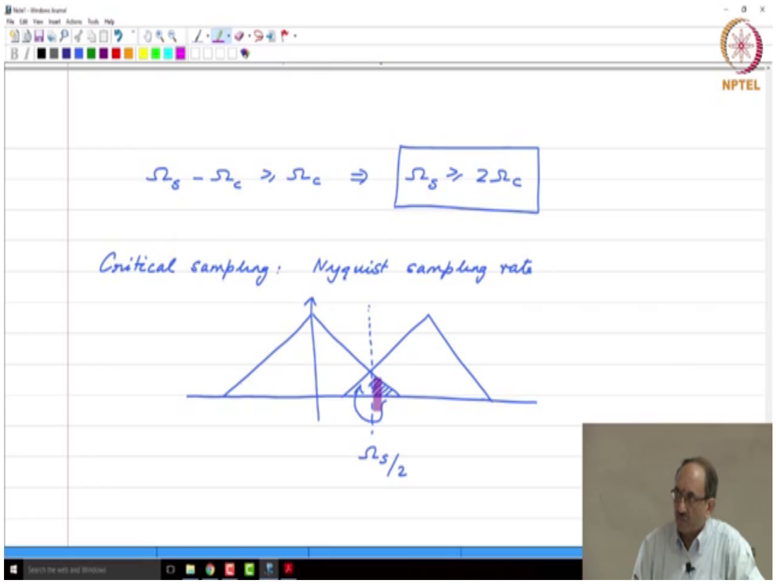


Therefore, this is critical sampling, and when you satisfy that theorem such that it just about these two repetitions abut each other but do not overlap, that particular sampling frequency is called the Nyquist sampling rate. And then this also you must have heard. So, these two overlap and it should be clear that whenever you have something that is periodic, then you need to consider the information only in the range 0 to the period or between minus half the period to plus half the period, anything outside this is the same.

And hence, since this is periodic with period Ω_s, either you need to consider the information in the interval 0 to Ω_s or between $-\Omega_s/2$ to $+\Omega_s/2$. Because everything else outside this is the same, you do not get any new information. And hence, if Ω_s were the sampling interval, all you need to do is you need to consider the information between $-\Omega_s/2$ to $+\Omega_s/2$, that is all.

And that is this dashed line and when these two things overlap, you produce distortion and this is what is called as aliasing. And hence, the spectrum that you have is really this. So, this is the aliased spectrum, you cause distortion and this is the distortion that is caused is called aliasing. And just to focus on what aliasing is, I am just drawing the positive part of the spectrum.

(Refer Slide Time: 19:32)

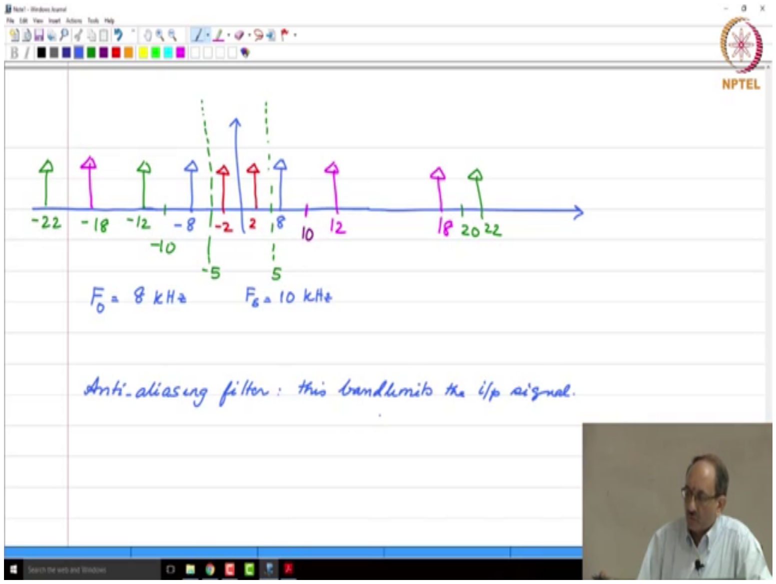


Remember, this is $\Omega_s/2$ and only the information between $-\Omega_s/2$ to $+\Omega_s/2$ needs to be considered. And because of this, what is happening here is, this portion folds over and appear as a low frequency. So, this is actually a higher frequency component in the signal originally, because you under sampled, it appears as a lower frequency component in the under sampled signal. So, a high frequency component masquerading as low frequency one is called aliasing.

And we will see a simple sinusoidal example where this is brought out. So, this is for the general case where I have assumed the signal spectrum to be like this and showing how high frequency component folds over and acts like a low frequency one. And this, because you can think of this as being folded over and appearing in this, this also is called as the folding frequency. Another name that you will find in the literature is, this $\Omega_s/2$ is also called as the folding frequency.

So, this can easily be illustrated, this aliasing which is really a higher frequency component appearing as if it is a low frequency component because of under sampling easily can be seen when you consider a sinusoid.

(Refer Slide Time: 21:23)



So, let us consider a pure infinite duration sinusoid. Therefore, its frequency is, transform is two impulses. So, I am assuming that $F_0 = 8$ kilo Hertz. So, this is the frequency of the sinusoid. If I did not want aliasing, I need to sample at least at twice the highest frequency. Therefore, I need to sample it at 16 kilo Hertz are beyond if I did not want aliasing therefore, this is -8 and $+8$ so, this is F and this is now in kilo Hertz.

So, now in this example, I am showing the frequency axis in terms of Hertz which is nothing but $\Omega/2\pi$ for ease of calculation and reference. So, this is my underlying sinusoid. Now, what I am going to do is, I am going to sample it at 10 kilo Hertz which means everything in the spectrum will be periodic and the periodicity will be F_s. In the Ω domain, it will be periodic with Ω_s, in the F_s domain, it will be periodic in F_s. Therefore, everything will repeat after 10 kilo Hertz.

Therefore, what will happen is, if I now look at periodic repetition, -8 will occur at $-8+10$, right therefore, -8 will now occur at 2, because $-8+10 = 2$. Then, 8 also will repeat later with a periodicity of 10 kilo Hertz therefore, $8 + 10 = 18$ and hence let me draw this with a different color so, I will have 18 here.

So, this is two repetitions on the right so, things will saw also will repeat on the left. Therefore, 8 will appear at $8 - 10$; you saw $8 + 10$ here, $8 - 10$ will be -2. And, -8 will also occur at -18, but also note that, 2 will again repeat with a period of 10 therefore, $2 + 10 = 12$, right. Therefore, so these are some of the repetitions here, remember what will happen at 0 will also happen at 10.

Therefore, so this is 10 kilo Hertz; so, what will happen at 0 will also happen at 10 therefore, you have this picture. And remember, whenever the sampling frequency is F_s, you need to worry about the spectrum only between $-F_s/2$ to $+F_s/2$ therefore, you need to only worry about the spectrum between -5 and $+5$, this clearly not to scale.

And whatever is happening around 0 will happen at multiples of 10 and -10. Therefore, now you have, when you concentrate between -5 kilo Hertz and $+5$ kilo Hertz, the original signal which was 8 kilo Hertz because you are now under sampled has now become 2 kilo Hertz signal.

And what is happening around 0 is happening around 10 therefore, around 0 you have $+2$ and -2, similarly around 10 you have 8 and 12 so, these are the other copies. And similarly, what is happening around 10 will happen around 20 and hence you will have 18 and 22. So, these are the further copies that are present and hence you will have something like this, you will have −8 and −12 which is centered around −10, and similarly you will have −18 and −22.

Therefore, a 2 kilo Hertz signal is now is what appears in the spectrum, and this is a result of at 8 kilo Hertz signal being sampled at 10 which is inadequate, and hence the underlying spectrum corresponds to the spectrum of a 2 kilo Hertz signal. And hence, a high frequency signal namely 8 kilo Hertz masquerading as a low frequency signal, namely 2 kilo Hertz happens because of under sampling, sampling it at 10 kilo Hertz rather than the required minimum of 16 or beyond. So, this is all the spectrum is.

So, when you have on a given underlying spectrum $X_c(\Omega)$, if you want to sample it at Ω_s or F_s, all you need to do is you need to repeat that spectrum with a periodicity of Ω_s or F_ss, equivalently F_s. And then focus on the interval between $-F_s/2$ to $+F_s/2$, that is all. And here is an example of aliasing, 8 kilo Hertz behaves as if it were a 2 kilo Hertz signal.

The other thing to remember in this context is, all practical signals are time limited. If a signal is time limited, what can you say about, in the frequency domain will be band unlimited. Therefore, nor time limited signal can be band limited, no band limited signal can be time limited. Now, the implication of this is that, since all practical signals are time limited, their frequency content in the Fourier representation will be infinite duration.

And hence you cannot sample them without causing aliasing, no matter what sampling frequency use because it is a band unlimited signal, you will always cause aliasing. One way to avoid aliasing is to use what is called an anti-aliasing filter. So this, what it does is it band limits the input signal. Once your band limited this input signal and if you know sampled, it will not cause aliasing.

But, if you look at this, because you have put this band limiting or aliasing filter at the front end, you are now going to lose some high frequency component, because you have band limited the signal before sampling. So, you have lost high frequency component, so this is on one side. Now, you have to compare this with a signal, if you do not do this band limiting filter, if you do not introduce this and band limit the signal, if you sample, you are going to cause aliasing because the signal is band unlimited to begin with.

So, no matter what you do, you are going to distort the given signal. The question is, which is better? And the answer to this is this, anti-aliasing filter at the front end is better, because in that case, you only lose the high frequency component; whereas, if you cause aliasing, not only are you going to lose the high frequency information beyond $F_s/2$, but that high frequency information is going to come back as lower frequency component in the signal and cause distortion.

And in practice, it has been found that this aliasing is perceptually very annoying, and hence all practically ADC's have a band limiting filter in the front end. So, any ADC chip that you use will always have a front end anti-aliasing filter. And, it has been found that this is less objectionable than aliasing occurring, because the high frequency component falls back and appears as a low frequency component.

Fortunately, all practical signals have a spectrum that falls off with increasing frequency, and hence if you put this band limiting filter and remove beyond a certain point, the information lost will not be too much of an issue. Perceptually, it may not matter much, you choose the band limiting filter high enough so that all the relevant information is captured. For example, in telephony, in the first days of digital telephony, your voice signal was band limited to, you know what the sampling frequency was, and hence what the?

Student: (Refer Time: 31:52).

No, I am talking about early days of digital telephony. So, what you are talking about 44.1 is what you use, it is one of the standards for music and other things. But telephony for voice, do you know what the numbers were, what the sampling frequency was?

Student: (Refer Time: 32:15).

Not clear (Refer Time: 32:17).

Student: (Refer Time: 32:18).

No, this is too high.

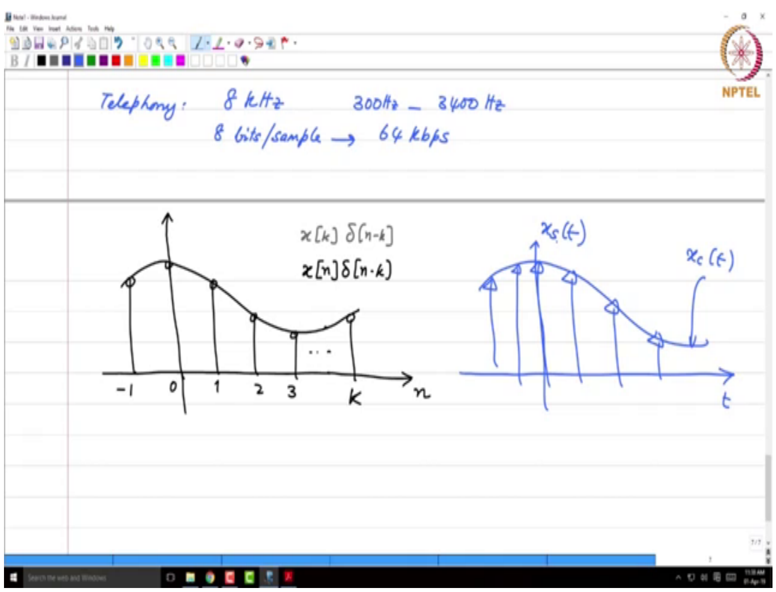
Student: (Refer Time: 32:21) 31, (Refer Time: 32:22).

32, no.

Student: (Refer Time: 32:25).

No; so, 8 kilo Hertz was the sampling frequency.

(Refer Slide Time: 32:35)



And, so in, the sampling frequency was 8 kilo Hertz and the front end filter was actually between 300 Hertz to 3400 Hertz. So, this was the front end filter that was used and then they were using 8 bits per sample, and this gave rise to 64 kilo bits per second telephony rate, sampling rate.

Again, 8 bits per sample is actually not adequate, you really need something like 12 bits; 12 or 13 bits. But they got around using 8 bits and still not sounding too bad, because they used what is called A-Law and μ-Law compression. Non-uniform quantization, that is, lower amplitude signals were quantized finer and higher amplitude signals were quantized coarser. Therefore, the 12 bits per second which is what was needed, if you used uniform quantization, that was overcome using just 8 bits by non-linear quantization. And μ-Law and A-Law are the compression standards used in the US and the other one is in Europe.

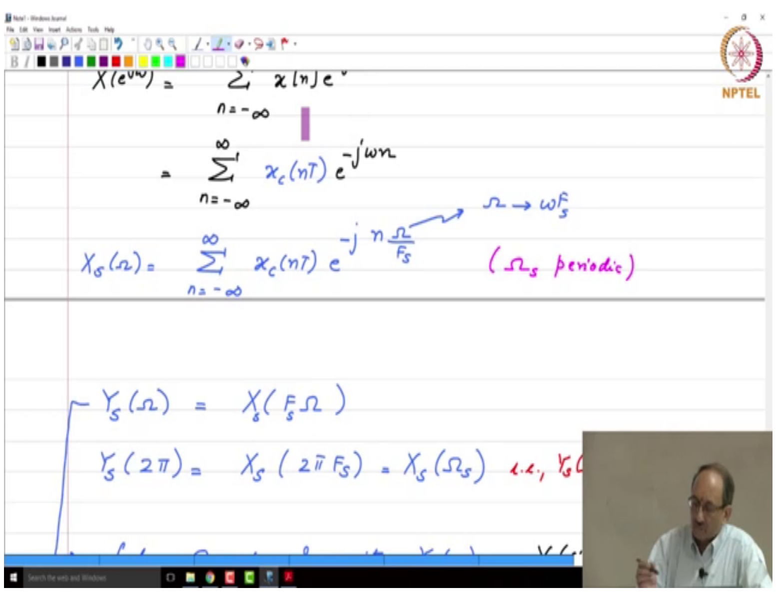
Therefore, you just needed 8 bits per sample and this gave 64 kilobits per second and the anti-aliasing filter was 300 to 3.4 kilo Hertz. And this is the reason why you will have trouble listening or distinguishing between s and F, because the distinguishing features of these sounds lay mainly in the higher frequency part of the spectrum. Because of this anti-aliasing filter that got knocked off, that is why people will say s as a son, F as a $father$. When you try to say some letters as abbreviations, you will have trouble at the other end unless you clarify what it is and that is why s as a son, F as a $father$ is needed, because those zones cannot be distinguished clearly at the other end, and that is because of this anti-aliasing filter that is cutting out high frequency components.

Because now, what we do is, we do not use just plain sampling, we use speech coding. You sample it at a certain rate, we take 30 milliseconds of speech, sample it and then use a source filter model and then transmit those parameters. So, the advantages you get, higher quality speech or you get a lower bit rate speech, but the price paid is more processing power; all this involves DSP processing. Anyway coming back to sampling.

So, anti aliasing filter is an integral part of all ADCs, that is because, you want to avoid distortion due to aliasing. And as I had mentioned all practical signals the spectrum decays fast enough, so that you can put a reasonable band limiting filter and still not perceptually feel the difference between the filtered and the unfiltered signal. Because the energy is so little beyond a certain frequency putting, this band limiting filter is not a big deal.

Now, let us come back, so, this was a digression for aliasing and what it does and how to mitigate the effects of this aliasing by putting this anti aliasing filter. Now, let us come back to the problem at hand. Remember, our goal is to connect samples of this signal. So, this is k and this is $x[n]\delta[n-k]$, so, this is your discrete time sequence. So, this is $0, 1, 2, 3$ and so on, and this in turn really becomes $x[k]\delta[n-k]$ because of the sifting property. And, what we really want is we are trying to relate the DTFT of this sequence and the spectrum of the impulse train sampled signal.

(Refer Slide Time: 37:43)



So, this is $x_s(t)$, this of course is $x_c(t)$ and the impulse train is the impulse train corresponding to $x_s(t)$. Now, we want to relate these two, that is, we have $X(e^{j\omega})$. So, this is nothing, but $\sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$, and this is nothing but $\sum_{n=-\infty}^{\infty} x_c(nT)e^{-j\omega n}$.

So, this is one part of the problem, the other thing that we have just now derived is this $X_s(\Omega)$, and $X_s(\Omega)$ we showed that this was nothing but $\sum_{n=-\infty}^{\infty} x_c(nT)e^{-jn(\Omega/F_s)}$, this is what we derived first. And what we had derived using this approach was to show the periodic repetition, that amplitude scaling by $1/T$ and then effects of aliasing. We will come back to this, this periodic repetition picture is very important. So, the periodic repetition picture was not present in this expression, but if you look at these two, you can now start to see the connection here.

Let us take $X_s(\Omega)$, this is actually, this is Ω_s periodic. We got the picture about the periodicity of X_s by the second approach; we know that this is Ω_s periodic. Now, let us consider $Y_s(\Omega)$, and $Y_s(\Omega) = X_s(F_s\Omega)$. So, what we are doing is, we are forming a new function $Y_s(\Omega)$ which is a scaled version of $X_s(\Omega)$, that is what we are doing here. And typically, F_s is greater than 1. Therefore, one interpretation of this is $Y_s(\Omega)$ is a compressed version of $X_s(\Omega)$, because $x(2t)$ is a compressed version by a factor of 2. And, hence to start to understand this, since typically F_s is greater than 1, you can think of $Y_s(\Omega)$ as $X_s(F_s\Omega)$.

So, this is compression by a factor of F_s, that is what is happening here. what about $Y_s(2\pi)$? $Y_s(2\pi)$ is X_s of; wherever Ω is, you replace $\Omega/2\pi$. Therefore, this becomes $2\pi F_s$. And this is nothing, but $X_s(\Omega_s)$. Therefore, what can you say about the periodicity of $Y_s(\Omega)$? $Y_s(\Omega)$ is 2π periodic. So, $Y_s(\Omega)$ is 2π periodic.

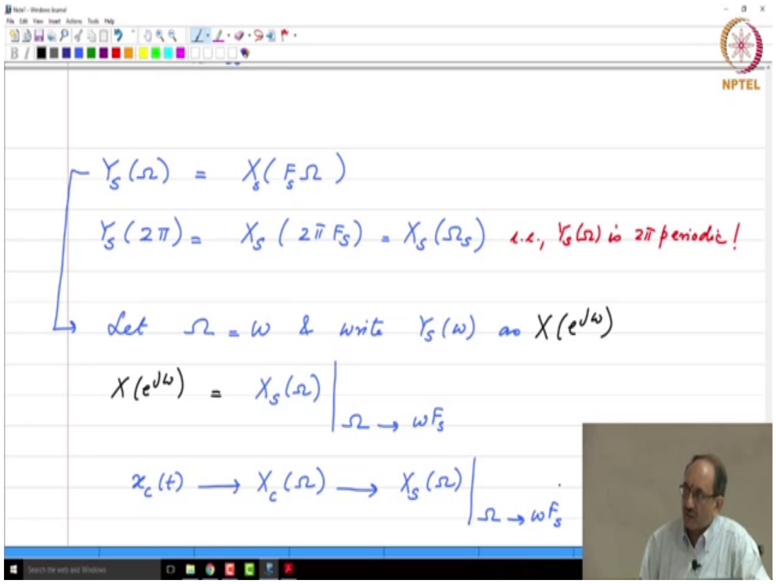
Now, what we will do is, we will replace Ω by ω, that is, in $Y_s(\Omega)$ which is 2π periodic, we will relabel the independent axis Ω by ω and remember, now we have relabeled the independent axis from Ω to ω and then write $Y_s(\Omega)$ as $X(e^{j\omega})$. So, basically, what is happening is, let us go back to this picture. So, $X_s(\Omega)$ is the impulse train sample signal spectrum. It is periodic with period Ω_s, that is one thing, of course the other thing is the amplitude scale factor.

So, all you need to do to relate this to the underlying sequences DTFT is, scale the frequency axis by F_s, that is all. And why does that make sense? That makes sense, because this is $2\pi F_s$. If you now scale this by F_s, this point will now get mapped to 2π, and the DTFT is nothing but 2π periodic. That is all, as simple as that.

And, hence to get $X(e^{j\omega})$, all you need to do is, you need to take $X_s(\Omega)$ and then replace Ω by ωF_s. Go back here, take this Ω and replace Ω by ωF_s. So, now, take Ω and then replace Ω by $\omega times F_s$, if you did that, this F_s will get cancelled. So, $-jn(\Omega/F_s)$ will become $-j\omega n$ which is exactly this and $x_c(nT)$ is nothing but $x[n]$, that is all as simple as that.

So, basically you are taking the impulse train sampled signal which is Ω_s periodic and compressing the frequency axis by F_s, so that the Ω_s periodic function becomes 2π periodic which is what the DTFT is.

(Refer Slide Time: 45:39)



So, now what we will do is, we will start off with $x_c(t)$, then we know what its Fourier transform is which is $X_c(\Omega)$, then what we will do is, we will form $X_s(\Omega)$. So, two things will happen when you go from $X_c(\Omega)$ to $X_s(\Omega)$; first thing is periodic repetition with period Ω_s, the second thing is amplitude scaling. Then, what we will do is, we will take $X_s(\Omega)$ and then replace Ω by ωF_s and then we will see that this will give us $X(e^{j\omega})$.

And we will do this by taking specific examples for which we know the CTFT, and the corresponding sequence's DTFT also will be known. What we do is, we will start off from the CTFT and derive the DTFT going through these steps for well known functions. And then, we will see that, we will get back the same expression that was derived when we encountered the sequence's DTFT. So, now you will be able to see the connection much closer.