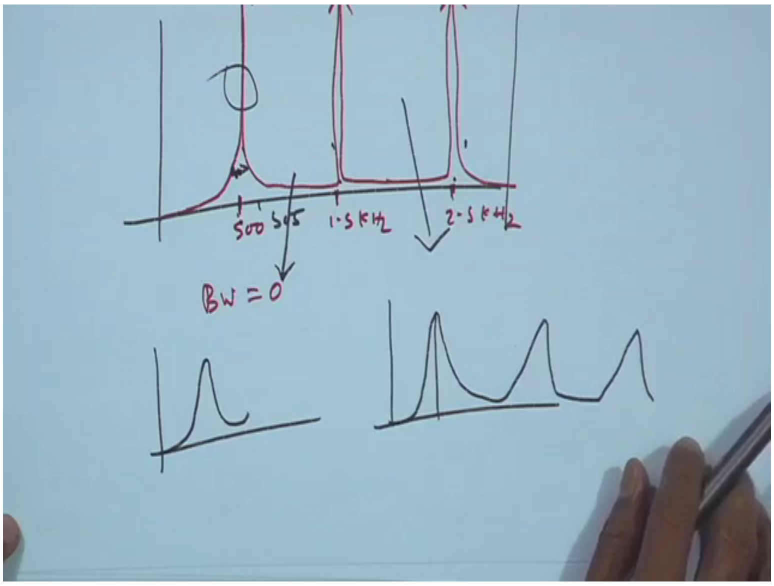
Digital Speech Processing
Prof. S. K. Das Mandal
Centre for Educational Technology
Indian Institute of Technology, Kharagpur

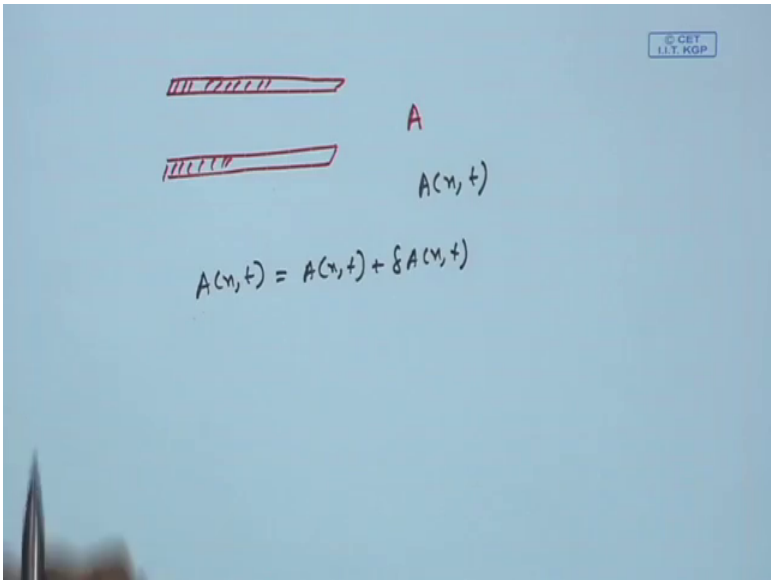
Lecture - 11
Uniform Tube Modeling Of Speech Processing Part – III

So now, we have derived that transfer function of lossless tube, without considering any loss infinite power at 500 hertz, 1.5 kilo hertz, 2.5 kilo hertz.

(Refer Slide Time: 00:26)

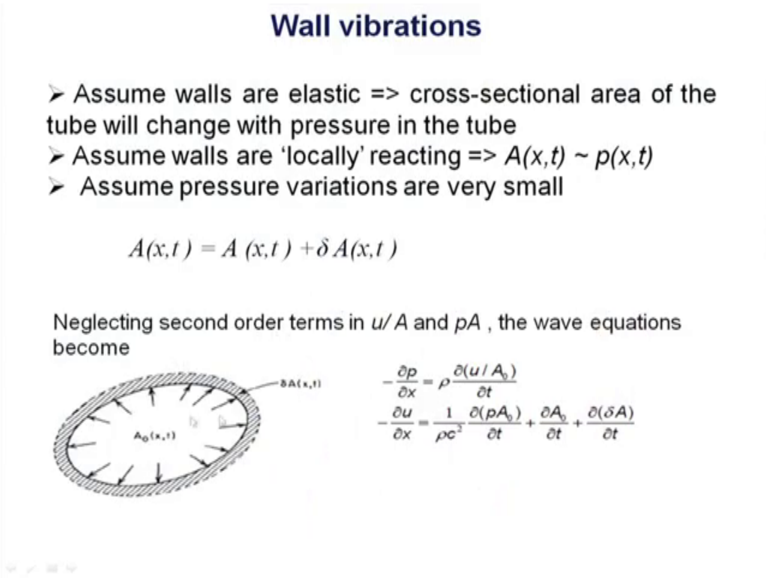


(Refer Slide Time: 00:37)



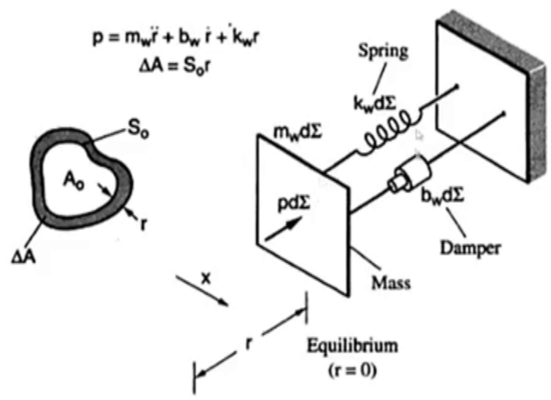
Now, if I consider that the tube the wall which I supposed to be rigid. If this wall is not rigid. This wall can flexible, this wall can modify if pressure is high. If there is a pressure wall can change. So, what is happening instead of cross sectional area a is fixed. Now a is also function of x and t a is also function of x and t. So, anytime the cross sectional area A x t is nothing but a average area or you can say the A x t plus delta, change of A x t due to the pressure at that position at that time delta A.

(Refer Slide Time: 01:32)



So, if you see the slides if you see the slides the slides is there. So, inside it is a 0 t; I can say it is a 0 t fixed, and then it is a delta A x t is expansion. So now, if I instead of u A and rho A the wave equation I have modified because the wave equation instead of fixed A, A is modified, then if I put this one a 0 x t into del A x t for a 0 x t, then the modified equation will be like this. So, I am not going details modified equation mathematical things then it will be take lot of time. So, I this is not objective to derive the details can be available in the papers also mathematical details. So, there is a complex equation will come out if I see the complex. So, I can say this tube is made of a muscles, I can say muscle is nothing but a spring mass action.

(Refer Slide Time: 02:31)



So, it is nothing but a spring mass mechanical oscillator, if I consider the second order equation of the spring mass mechanical oscillator, then this will be like this.

(Refer Slide Time: 02:43)

The differential equation relationship between area perturbation $\delta A(x,t)$ and the pressure variation, $p(x,t)$

$$m_w \frac{d^2(\delta A)}{dt^2} + b_w \frac{d(\delta A)}{dt} + k_w(\delta A) = p(x,t) \quad (1)$$

$m_w(x)$ = mass/unit length of the vocal tract wall

$b_w(x)$ = damping/unit length of the vocal tract wall

$k_w(x)$ = stiffness/unit length of the vocal tract wall

$$-\frac{\partial p}{\partial x} = \rho \frac{\partial (u/A_0)}{\partial t}$$

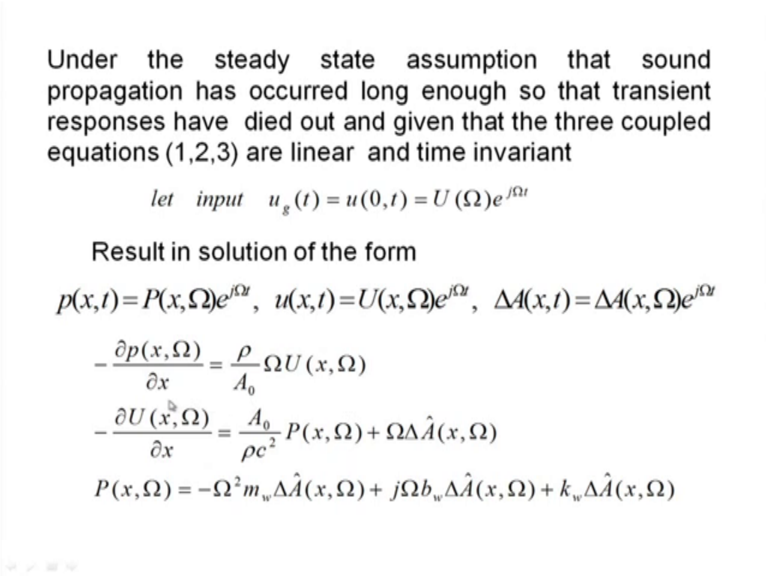
$$-\frac{\partial u}{\partial x} = \frac{1}{\rho c^2}\frac{\partial (pA_0)}{\partial t} + \frac{\partial A_0}{\partial t} + \frac{\partial (\delta A)}{\partial t}$$

$\Rightarrow$

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A_0}\frac{\partial u}{\partial t} \quad (2)$$

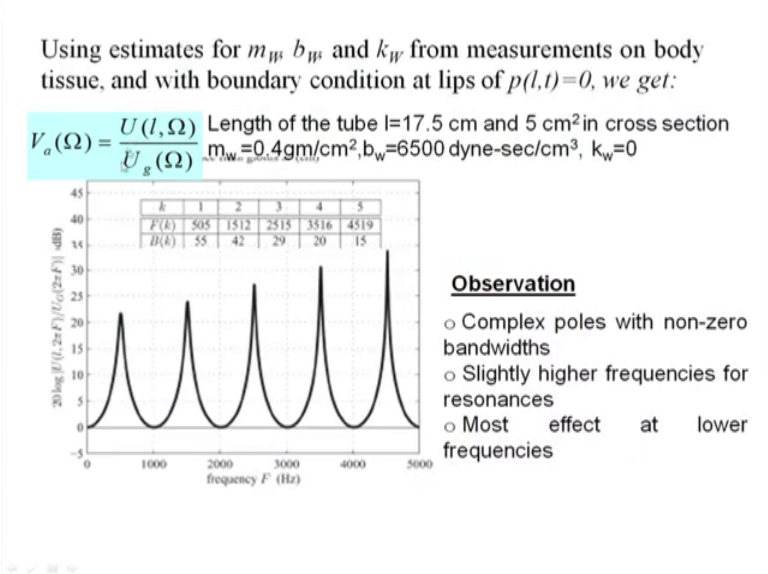
$$-\frac{\partial u}{\partial x} = \frac{A_0}{\rho c^2}\frac{\partial p}{\partial t} + \frac{\partial \Delta A}{\partial t} \quad (3)$$

(Refer Slide Time: 02:48)



Then again if I; 1, 2, 3, if I consider and then the wave equation instead of to wave equation with this 3 equation will derive and from that 3 equation I get an analytical solution of what will happen in the frequency response which is v a omega u l omega divided by u g omega.

(Refer Slide Time: 03:00)



Now, they said if the length of the tube is 17.5 centimeter and 5 centimeter square is the cross sectional area, and m omega is 0.5 gram per centimeter square b omega is 6 5 dyne; that means, m omega is nothing but a mass per unit length b omega damping per unit length and stiffness per unit length of the vocal tract wall.

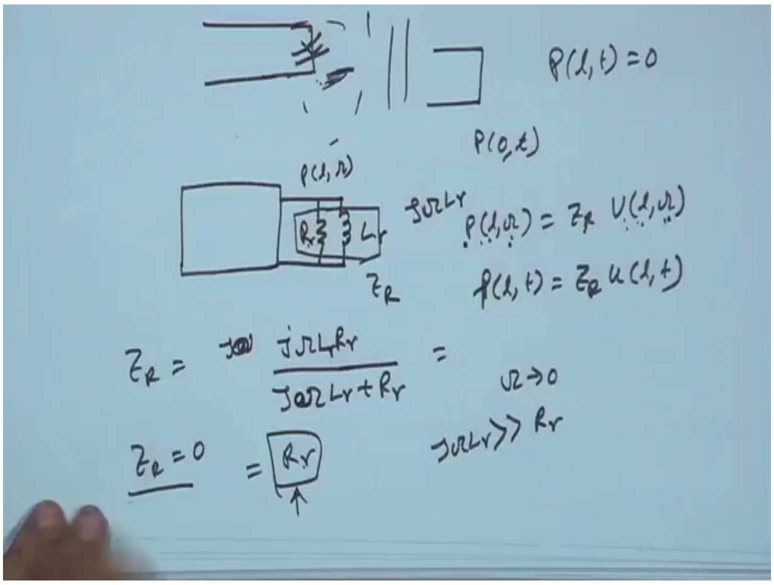
If I consider that value, then it is said it was found in the frequency response earlier it was infinite energy 500 hertz 1.5 kilo hertz 2.5 kilo hertz like that. It is said that complex pole with nonzero bandwidth. So, here it is 0 bandwidth, here it is 0 bandwidth, bandwidth is 0. In tube lossless condition. Now if I introduce the loss what will happen the instead of 0 bandwidth some bandwidth will be generated. So, the poles are complex with nonzero bandwidth. Slightly higher frequency the frequency will be formant instead of 500s hertz it maybe 500 and 5 hertz. Slightly formant equation is shifted towards the higher frequency side, most affected in the lower band.

Lower frequency will be the most. So, lower frequency bandwidth will be increase more compared to the higher frequency. So, most affected is the lower band. Now if I consider So I said that earlier I am not considered the friction loss viscous loss and thermal loss, now if I consider the all kinds of friction loss thermal conduction on the wall viscosity all kind of losses, then I found it increase the bandwidth of the complex pole, and decrease the resonance frequency slightly. So, ultimately I can say that yes, if I consider the losses and not rigid to wall. So, instead of a infinite 0 bandwidth infinite impulse at every resonant frequency it becomes a finite bandwidth with a slightly shifting higher direction of the formant.

So, formant is shifted slightly higher direction. So, instead of this figure ultimately I will get some bandwidth kind of formant, but if you see more or less it will be on an around of 500 hertz slightly shifted. So, I can easily say normally if it is tube is totally open, first formant is 500 hertz, second formant is 1.5 kilo hertz, third formant is 2.5 kilo hertz, third formant is 2.5 kilo hertz, fourth formant is 3.5 kilo hertz fifth formant is 4.5 kilo hertz. So now, if you see if a signal is band limited with 4 kilo hertz, then I can say I only can get up to 4th formant.

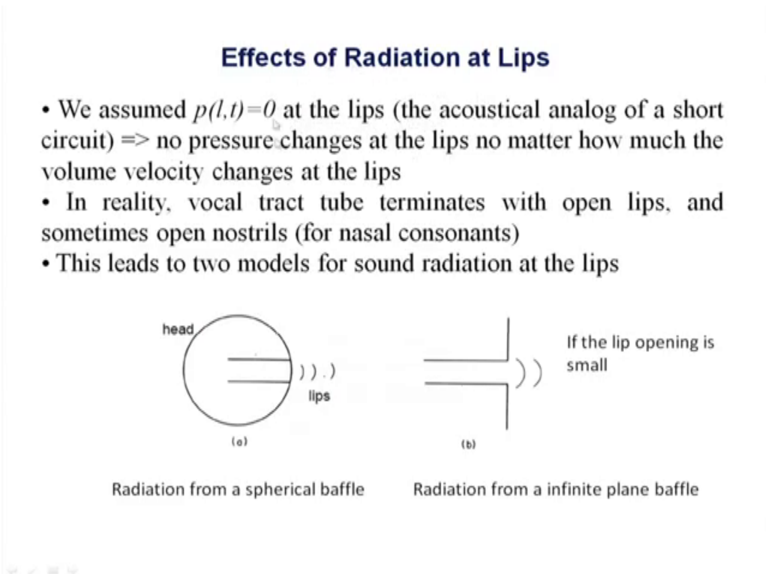
I cannot get the fifth formant, because fifth one 4.5 kilohertz. So, I will not get it. So, if I cut the signal in here I will get only 3 formant if I cut the signal in there I can get the 3 or 4 formant. So, depends on the sampling frequency format will come I will come this will discuss later on also. So, I am not detail discuss about the this kind of losses. Now second difficulty is the effect of radiation at lips. So, what I said, once the air is radiated from the mouth. So, I can say at the opening of the tube the acoustic wave is radiated.

(Refer Slide Time: 07:00)



So, the radiation this radiation why it is radiated in the atmospheric air; so, what will be the radiation losses? What kind of effect I will get due to this radiation? So now, if you see the assume that P l t is equal to 0 at lips; this is the we have assumed, the acoustical analogue this short circuit the output couple this thing.

(Refer Slide Time: 07:25)



If I say transmission line, I said the output is totally short circuit. P is equal to 0 t means the voltage is equal to 0. So, output is short circuit. Here I said there is a no load, no acoustic load. So, it is loaded 0 output is short circuit it is ideal condition, but it may not be the ideal condition. So, what will be the effect of air load on the lip? Or you can say the radiation effect on the acoustics wave transmission?

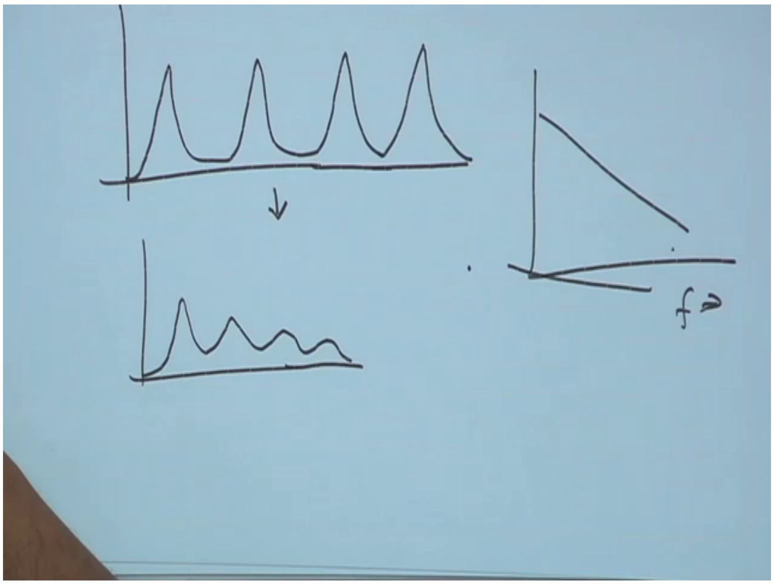
So, then I if I put the microphone then I can say what kind of signal I am expecting from the mouth. Now if I consider that how do you do that. So, I can say let us this is the tube and this is the mouth. Here it is radiated, So whole load is the atmospheric load is there in the lip. So, let us the load is nothing but a inductive and resistive load, L r and R r. Acoustics mass inductive load and resistive load. Now it is there, then I can say here it is p l omega is nothing but a p l omega is nothing but a let us this constitute this load constituted z r. So, z r into u l omega.

So, I have not writing small u keep it capital U l omega means length omega means, instead of time it is frequency it is length and frequency capital P frequency response p l omega is nothing but a z r into u l omega. Or I can say that p l omega p l t is nothing but a small p l t is nothing but a z r into u l t. Now what is z r? Z r is nothing but these 2 are in parallel. So, j omega you can say the j omega L R r divided by j omega omega L r plus R r. Because inductance j omega L r parallel with r. So, j omega L R r r divided by j omega L r plus R r. Now if you see, if omega is very omega tends to 0.

At very low frequency z r is equal to 0; that means, the acrostics load in the radium radiation load in here is not there it is totally short circuit which is the idle condition we said, p l t is equal to 0. So, the frequency response whatever we have get due to the considering the loss and that things that will be remain constant remain same; that means, that implies or; that means, that low frequency components are not affected by the radiation load. Because at low frequency radiation load does not effect the frequency response of the acoustical tube. So, low frequency are less affected. Now if I say j omega or omega L r is much, much greater than R r at high frequency. Much, much greater than R r.

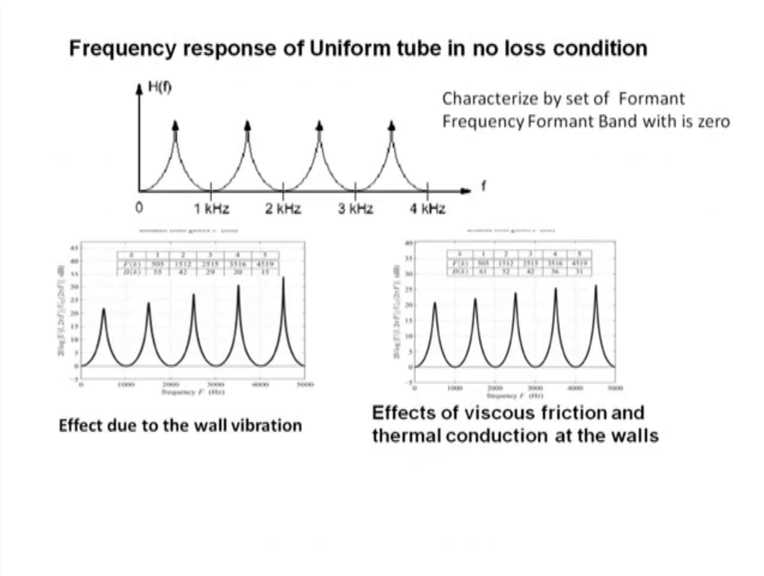
Then I can say it is nothing it is nothing but a R r j omega l j omega l cancel. So, it is nothing but a R r. So, load is totally resistive. So, radiation load is resistive means there is a loss radiation loss. So, I can say high frequency are affected due to the radiation loss ok.

(Refer Slide Time: 11:55)



So, I can say the frequency response which I will get after the considering the losses with the bandwidth. If it is passed through the radiation loss, then low frequency will not affected much more, but high frequency will be lost. So, I will explain it in here.

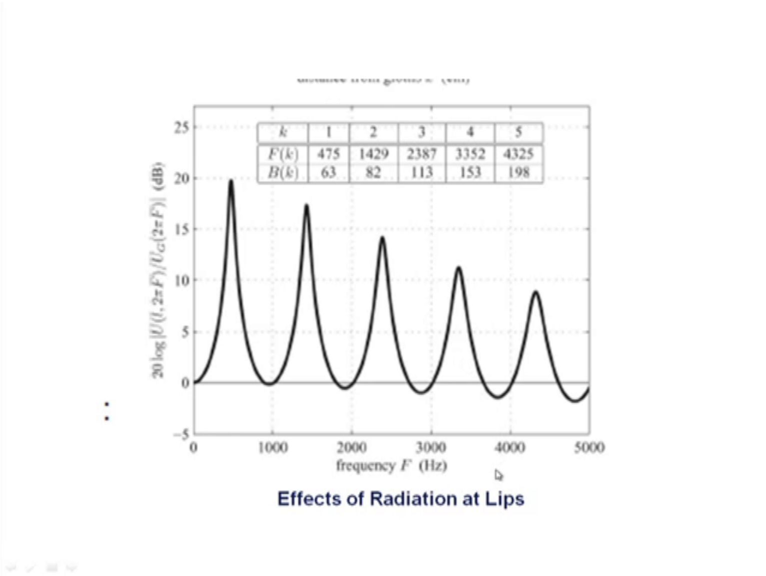
(Refer Slide Time: 12:20)



If you see this is the figure with the no loss condition. At 500 hertz infinite energy 0 bandwidth.

If I consider the wall vibration then I say the bandwidth is increases slightly amplitude increases in high frequency and band to bandwidth is introduced. And bandwidth in the low frequency are much more, then if I say the viscous loss friction law loss consider then again the bandwidth is increase the bandwidth and slightly decrease the formant frequency, it slightly increase the formant frequency. So, increase decrease if I cancel out I can say that formant frequency is and around what about 500 hertz I have getting here, but bandwidth is here, here is 0 bandwidth.

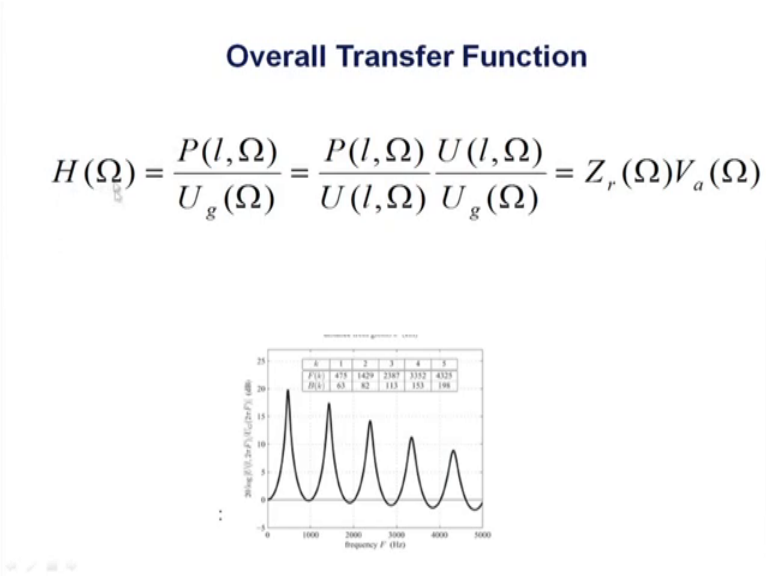
(Refer Slide Time: 13:06)



Now, if I consider the radiation loss. Then if you see high frequency are resistive radiation. So, this circuit this will be suppressed.

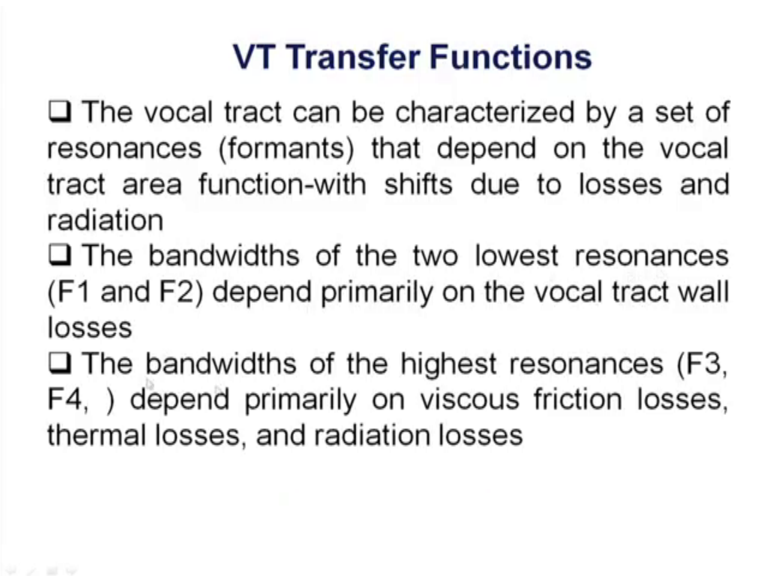
So, radiation loss due to the radiation loss high frequency components amplitude will be decrease. Now if you see that also here also then in mathematics h omega transfer function of the tube is nothing But a p l omega divided by u g omega which is p l omega u l omega.

(Refer Slide Time: 13:29)



So, it is nothing but a z r omega will be multiply with the v omega z r omega is 1 or no effect if omega is very low. If z r omega is sorry if the omega is low radiation effect is not negligible, if it is z r omega is omega is very high then there is a lot of attenuation in the high frequency. So, frequency response of the z r omega is nothing but like this. This stepper high if the frequency increases loss is increases. So, that is why if I see any speech signal.

(Refer Slide Time: 14:15)

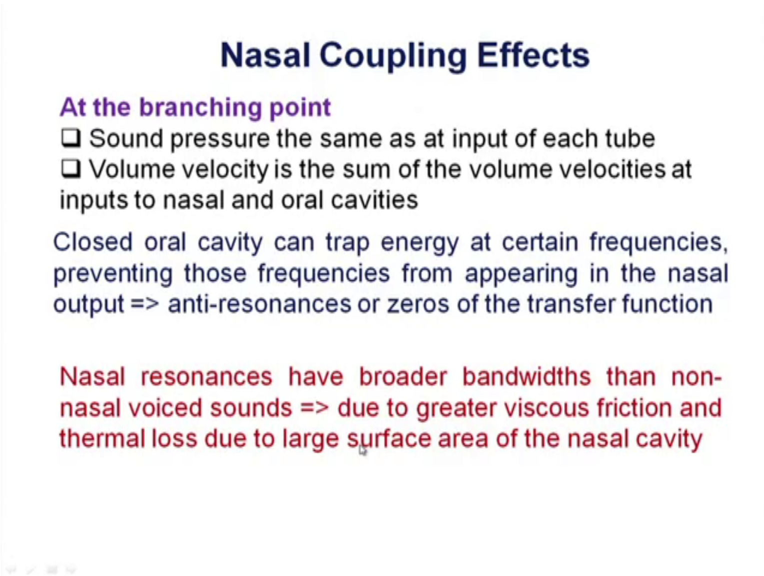


If you see the speech signal it will look like that high frequency or amplitude are less it is due to the radiation loss.

High frequency amplitude are very less. So, this is the frequency response of the uniform tube if I consider whole tube vocal tract in a single tube, then I say this is the frequency response this is a transfer function of the single tube model, where the frequency response is that formant frequency are 500 hertz 1.5 kilo hertz 2.5 kilo hertz, but the response is high frequency are attenuated due to the radiation loss. And it is not 0 bandwidth because of the if I consider the losses inside the tube.

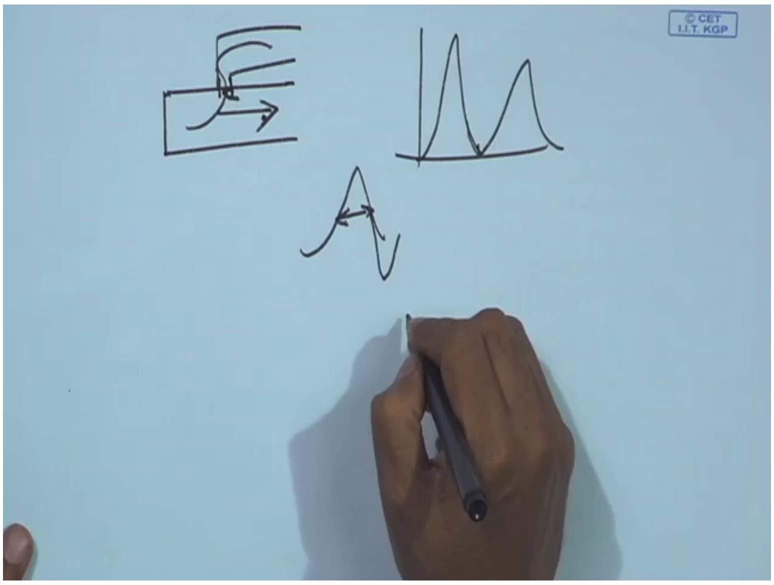
Now So, vocal tract can be characterized by a set of resonance that depends on the vocal tract area function with shifts due to the loss and radiation. The bandwidth of the 2 lowest resonance depend primarily on the vocal tract wall losses f 2 and f 2. And the bandwidth of the highest resonance depends the highest frequency resonance depend primarily on the viscous friction loss thermal loss and radiation loss. Because if you see I said the bandwidth is most affected frequency or low frequency in case of wall variation in the wall. In case of viscous law high frequency bandwidth are increase introduced.

(Refer Slide Time: 15:43)



Now, next one is nasal coupling effect. This is oral one cavity I have done this is the one cavity I have done.

(Refer Slide Time: 15:48)



Now, if you see you experience, that this is a cavity and it produces sound now if I put a hole in here sound will change.

If you see the flute there is a lot of hole in the top of the flute and by pressuring the closing the hole I can change the sound if you see that flute or saxophone just closing the hole means I am closing the hole in the tube. So, if I put a hole here you know the frequency response is change the 2 frequency response will be change. Now in case of our case in vocal tract also there is a nasal coupling in here. So, once the velam is open. So, nasal coupling is coupled with the oral tract give structure is changed frequency response will be change. So, sound pressure the same as at input for each tube the volume velocity is the sum of the volume velocity at the input to nasal and oral cavity. So, closed oral cavity can trap energy at certain frequency preventing those frequency from appearing in the nasal output.

So, if you think about what I producing a nasal consonant or nasal sound if my oral cavity is totally closed. Then the air is coming through the nasal. Now if you oral cavity is not totally closed then the nasal and oral both cavity are, So if I consider the nasal case sound is coming in this path then sound pressure will be trapped by these oral cavity. So, that trap will produce some anti resonance frequency. So, nasal resonance have border bandwidth loss is much more. So, bandwidth. So, in case of the formant band width will be border in case of nasal sound, and the anti resonance will be there.

So, I can say deep if I say this, this point will be much more larger point anti resonance will be introduced. So, bandwidth will be larger also. So, compared to the oral sound, nasal sound bandwidth is larger. Now let us I stop here because next class I will start about the sound another topic which is called how the excitation will excited the vocal tract, then will derive the 2 boundary condition which is at the lip at the glottis. After derived considering the 2 boundary condition will derive the total transfer function in digital domain and then try to implement is uniform tube model. Then you go for the multi tube modeling of the human vocal tract.

Thank you.