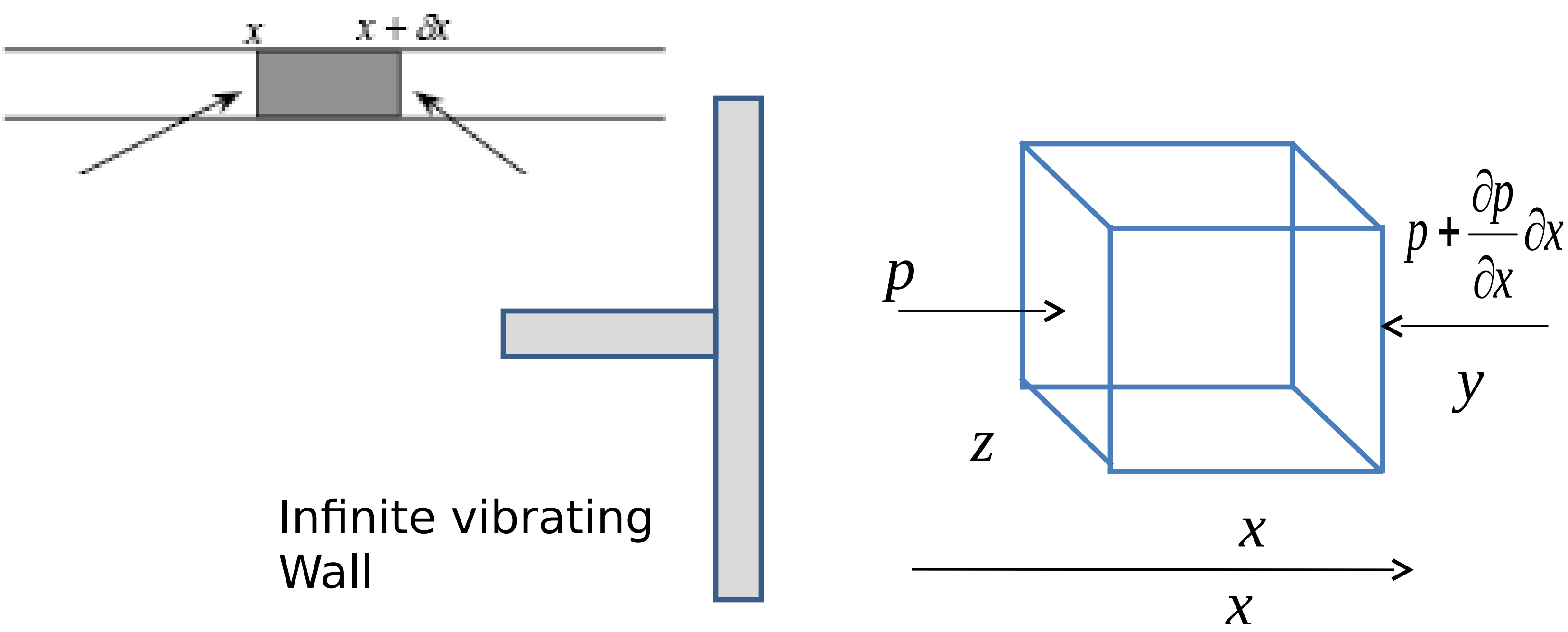
Uniform Tube Modeling of Speech Production

Wave equation



- Pressure $p(x, t)$ is a function of t and x and total pressure is

$$p(x + \Delta x, t) \approx p(x, t) + \frac{\partial p}{\partial x}\Delta x$$

- Particle velocity is $v(x, t)$
- Density of air is ρ and also function of x and t

To derive wave equation three law of physics is used

1. Newton’s Second law of motion → this law predicts that a constant applied force produces a constant acceleration
2. The Gas law from thermodynamics →relates pressure volume and temperature under the adiabatic condition
3. Conservation of mass

Assumption

1. The medium is homogeneous
2. The pressure change across a small distance can be linearized
3. There is no friction of air particle
4. The air particle velocity is small
5. Sound is adiabatic

$$p(x+\Delta x,t)=p(x,t)+\frac{\partial p}{\partial x}\Delta x$$

Net force on the chunk is

$$F=A[p-(p+\frac{\partial p}{\partial x}\Delta x)]=-A\frac{\partial p}{\partial x}\Delta x=(-\frac{\partial p}{\partial x}\Delta x)\Delta y\Delta z$$

Assumed density in the cube is constant $m=\rho\Delta x\Delta y\Delta z$

From Newton's Second Law of Motion $F=ma$

Acceleration of the cube air is $a=\frac{dv}{dt}$

$$-\frac{\partial p}{\partial x}=\rho\frac{dv}{dt} \qquad .....1$$

$\frac{dv}{dt}$ Is total derivative and *v* is a function of *x* and *t*

So true acceleration is $$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}$$

$$-\frac{\partial p}{\partial x} = \rho\frac{dv}{dt}$$

$$-\frac{\partial p}{\partial x} = \rho[\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}]$$

This is nonlinear equation in the variable *v* because the particle velocity *v* is multiply with $\frac{\partial v}{\partial x}$

$v\frac{\partial v}{\partial x}$ Is small relative to $\rho\frac{\partial v}{\partial t}$

$$-\frac{\partial p}{\partial x} = \rho\frac{\partial v}{\partial t}$$

Gas Law

We can express the pressure in terms of the density:

$$pV = nRT$$
$$= \frac{\rho V}{M} RT$$
$$\Rightarrow \quad p = \rho \times \frac{RT}{M}$$

n = moles of air = molecules $\div (6 \times 10^{23})$
R = gas constant = 8.314 J / (K · mol)
T = Temperature (° K)
ρ = density (≈ 1.225 kg / m^3)
M = molecular weight of air = 0.029 kg / mol
γ = specific heat ratio of air = 1.4

We define $\quad c^2 = \frac{\gamma RT}{M} \approx (340 \text{ m/s})^2 \quad \Rightarrow p\gamma = \rho c^2$

Adiabatic Gas Law: For pressure changes too rapid for heat conduction to occur (e.g. sound vibrations):

$$\frac{d}{dt}\left(pV^{\gamma}\right) = 0 \quad \Rightarrow \quad V^{\gamma}\frac{\partial p}{\partial t} + p\gamma V^{\gamma-1}\frac{\partial V}{\partial t} = 0$$

$$V = A * \Delta x \qquad \frac{\partial V}{\partial t} = A * \partial v = A * \partial x \frac{\partial v}{\partial x} = V \frac{\partial v}{\partial x}$$

$$-\frac{\partial p}{\partial t} = \rho c^2 \frac{\partial v}{\partial x} \quad .....(2)$$

$$-\frac{\partial p}{\partial x} = \rho \frac{\partial v}{\partial t}$$

$$-\frac{\partial p}{\partial t} = \rho c^2 \frac{\partial v}{\partial x}$$

$$-\frac{\partial p^2}{\partial x^2} = \rho \frac{\partial^2 v}{\partial x \partial t}$$

$$-\frac{\partial p^2}{\partial t^2} = \rho c^2 \frac{\partial^2 v}{\partial x \partial t} \quad \Rightarrow \quad \frac{\partial p^2}{\partial x^2} = \frac{1}{c^2} \frac{\partial p^2}{\partial t^2}$$

$$\frac{\partial v^2}{\partial x^2} = \frac{1}{c^2} \frac{\partial v^2}{\partial t^2}$$

Volume Velocity $u(x,t)$ define as the rate of flow of air particle perpendicularly through a specific area.

$$u(x,t) = Av(x,t)$$

u(x,t) is the volume velocity

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A}\frac{\partial u}{\partial t}$$

$$-\frac{\partial u}{\partial x} = \frac{A}{\rho c^2}\frac{\partial p}{\partial t}$$

Acoustic

$$-\frac{\partial v}{\partial x} = L\frac{\partial i}{\partial t}$$

$$-\frac{\partial i}{\partial x} = C\frac{\partial v}{\partial t}$$

Electrical

$p(x,t)$ Acoustic pressure

$u(x,t)$ Acoustic volume velocity

$\frac{\rho}{A}$ Acoustic inductance

$\frac{A}{\rho c^2}$ Acoustic Capacitance

$v(x,t)$ is the electrical voltage

$i(x,t)$ is the electrical current

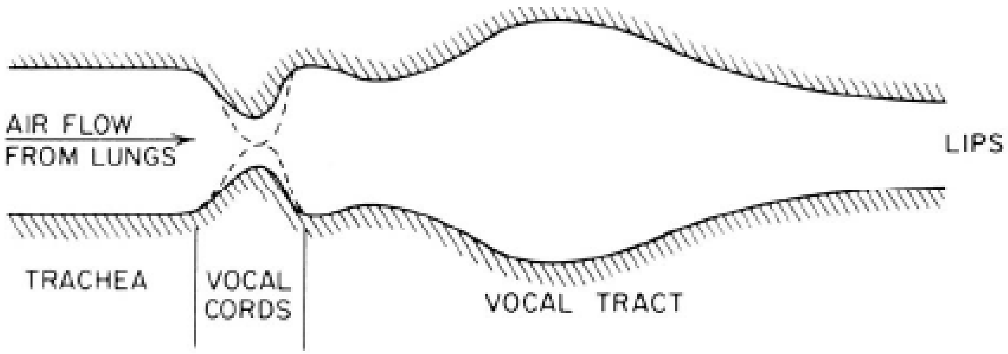
L electrical inductance

C electrical Capacitance

Lecture-6

Uniform Tube Modeling of Speech Production

Sound in the Vocal Tract



- Issues in creating a detailed physical model
  - time varying acoustic system
  - losses due to heat conduction and friction in the walls.
  - radiation of sound at the lips and nostrils
  - softness of the walls
  - nasal coupling
  - excitation of sound in the vocal tract

Use of basic physics to formulate ***air flow equations for vocal tract***

- Need to make ***simplifying assumptions about vocal tract shape*** and energy losses to solve air flow equations
- Some complicating factors:

➢ **time variation of the vocal tract shape *(*** we will look mainly at fixed shapes)

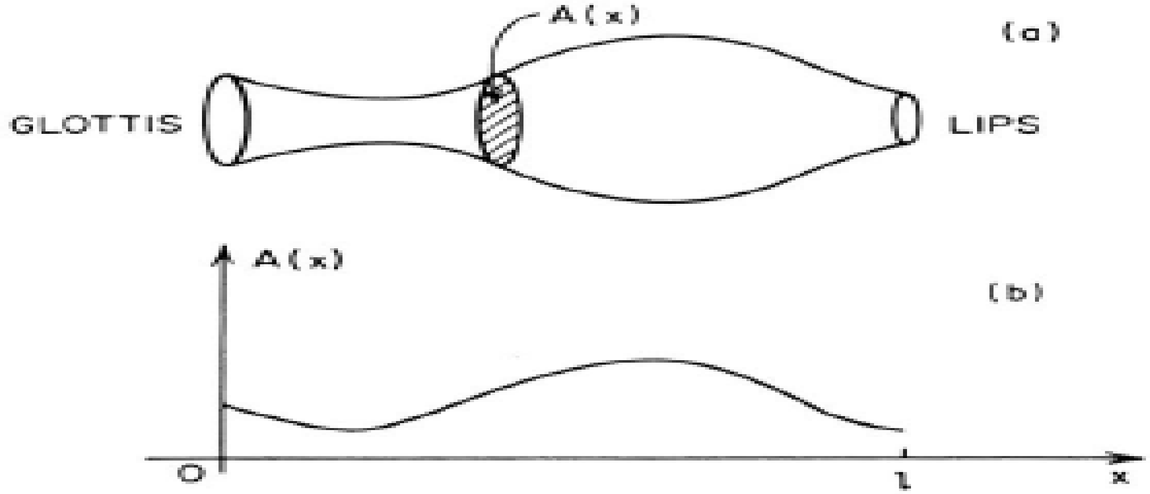
➢ **losses in flow at vocal tract walls *(***we will first assume no loss, then a simple model of loss)

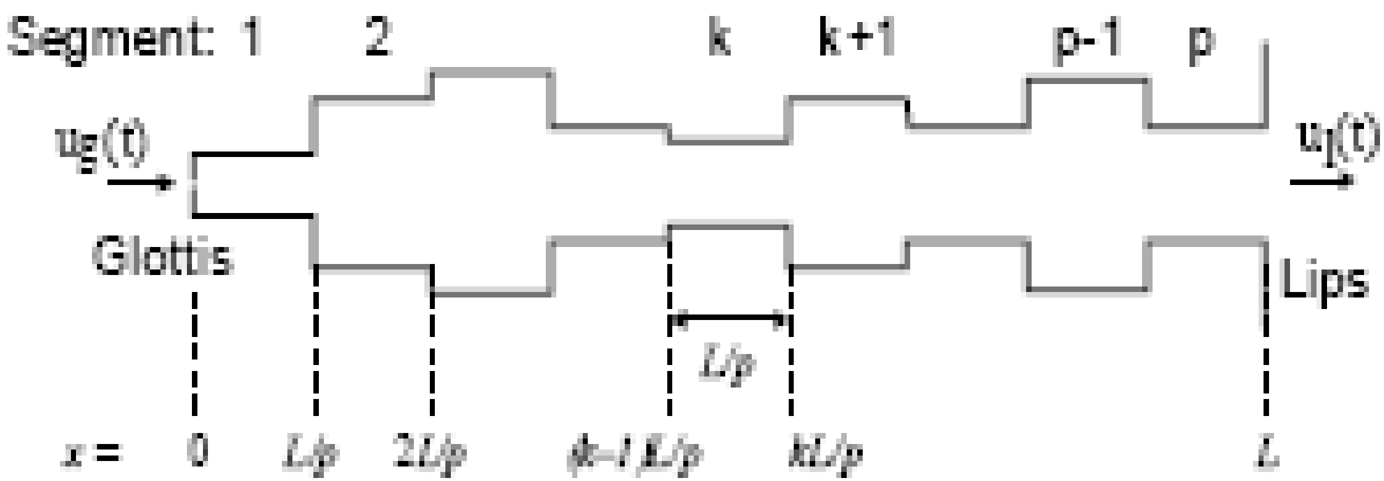
➢ **softness of vocal tract walls *(***leads to sound absorption issues***)***

➢ **radiation of sound at lips *(***need to model how radiation occurs***)***

➢ **nasal coupling *(***complicates the tube models as it leads to multi-tube solutions)

➢ **excitation of sound in the vocal tract *(***need to worry about vocal source coupling to vocal tract as well as source-system interactions)


A(x)
(a)
GLOTTIS
LIPS
A(x)
(b)
0
l
x


Segment: 1
2
k
k+1
p-1
p
ug(t)
ul(t)
Glottis
Lips
L/p
x =
0
L/p
2L/p
(k-1)L/p
kL/p
L

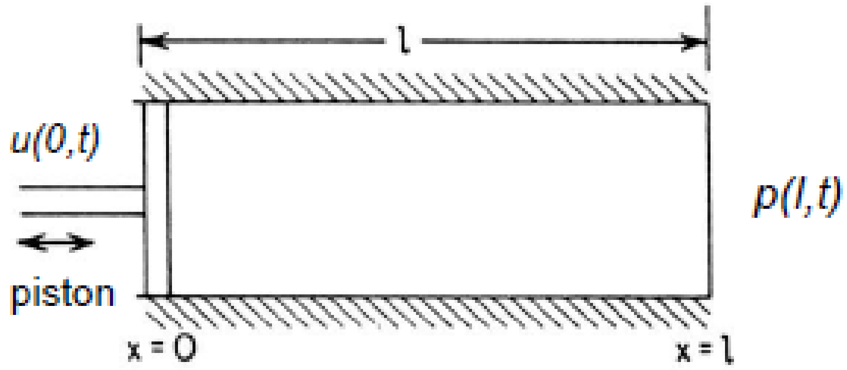
Effects of Losses in VT

- Several types of losses to be considered

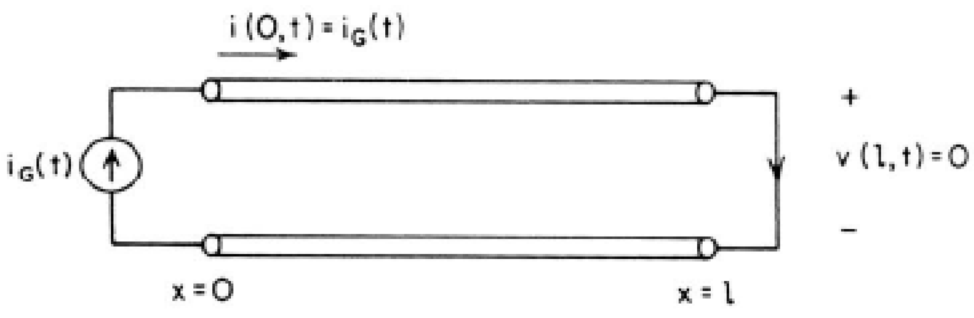
- Vibration of the tube walls
- Viscous friction at the walls of the tube
- Heat conduction through the walls of the tube

loss will change the frequency response of the tube

Uniform Lossless Tube



Assume uniform lossless tube => $A(x,t)=A$



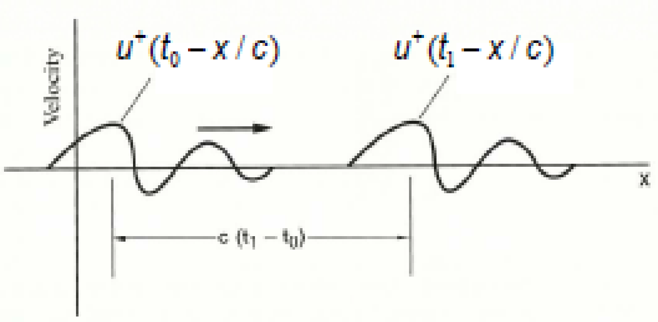
$$-\frac{\partial p}{\partial x} = \frac{\rho}{A}\frac{\partial u}{\partial t}$$

Traveling Wave Solution

$$-\frac{\partial u}{\partial x} = \frac{A}{\rho c^2}\frac{\partial p}{\partial t}$$

$$u(x,t) = u^+(t - x/c) - u^-(t + x/c)$$

$$p(x,t) = \frac{\rho c}{A}[u^+(t - x/c) + u^-(t + x/c)]$$



$u^+(t - x/c)$ *Wave traveling forward*

$u^-(t + x/c)$ *Wave traveling backward*

Two boundary conditions:

(a) at the glottis gives: $u(0,t) = U_G(\Omega)e^{j\Omega t}$

(b) at the lips gives: $p(l,t) = 0$

Since the differential equations are linear with constant coefficients, the solutions must be of the form where k^+ and k^- represent the amplitude of forward and backward wave

$$u^+(t - x/c) = k^+ e^{j\Omega(t - x/c)} \qquad u^-(t + x/c) = k^- e^{j\Omega(t + x/c)}$$

Traveling Wave Solution

- solve for K^+ and K^-

$$u(0,t) = U_G(\Omega)e^{j\Omega t} = K^+e^{j\Omega t} - K^-e^{j\Omega t}$$

$$p(\ell,t) = 0 = \frac{\rho c}{A}\left[K^+e^{j\Omega(t-\ell/c)} + K^-e^{j\Omega(t+\ell/c)}\right]$$

$$K^+ = U_G(\Omega)\frac{e^{2j\Omega\ell/c}}{1+e^{2j\Omega\ell/c}}; \qquad K^- = -\frac{U_G(\Omega)}{1+e^{2j\Omega\ell/c}}$$

- solve for $u(x,t)$ and $p(x,t)$

$$u(x,t) = U_G(\Omega)e^{j\Omega t}\left[\frac{e^{j\Omega(2\ell-x)/c} + e^{j\Omega x/c}}{1+e^{2j\Omega\ell/c}}\right]$$

$$p(x,t) = \frac{\rho c}{A}U_G(\Omega)e^{j\Omega t}\left[\frac{e^{j\Omega(2\ell-x)/c} - e^{j\Omega x/c}}{1+e^{2j\Omega\ell/c}}\right]$$

$$p(x,t) = jZ_0 \frac{\sin(\Omega(\ell - x)/c)}{\cos(\Omega\ell/c)} U_g(\Omega) e^{j\Omega t}$$ Where $Z_0 = \frac{\rho c}{A}$ 1

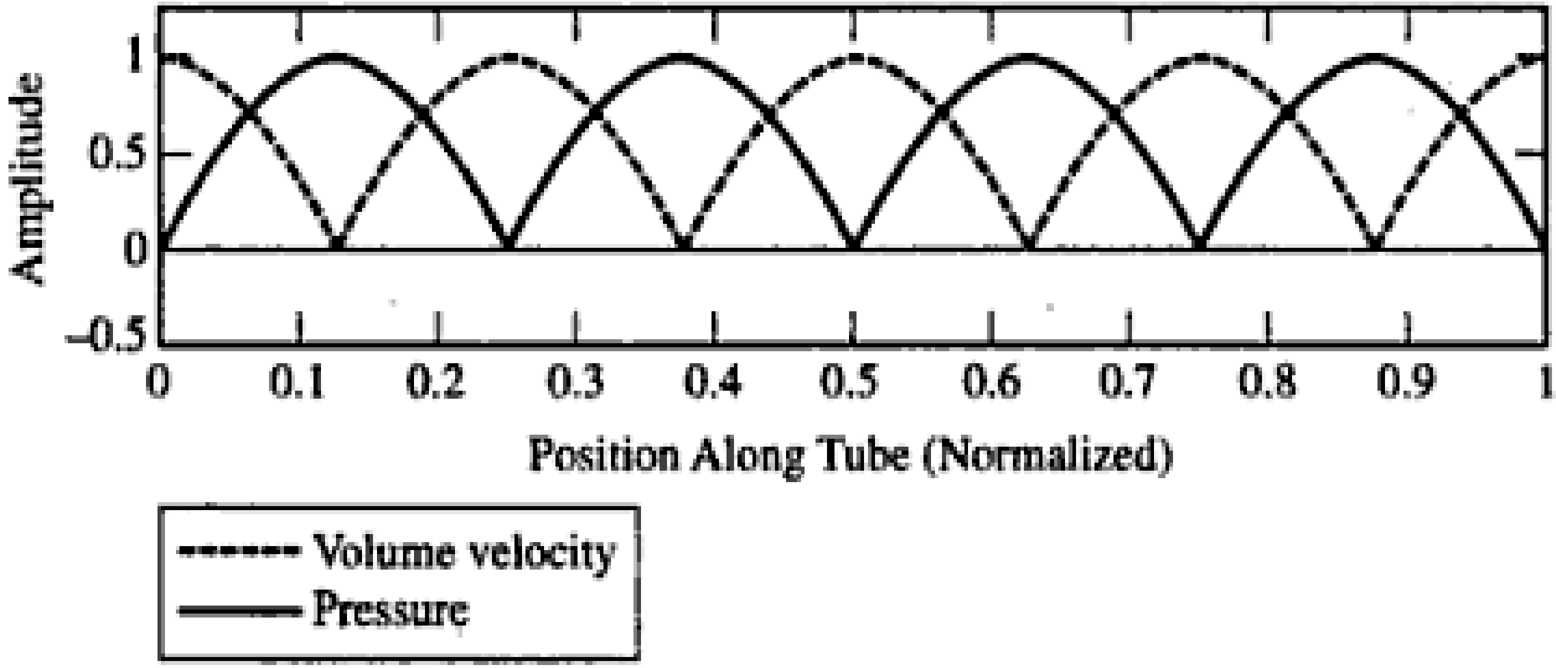
$$u(x,t) = \frac{\cos(\Omega(\ell - x)/c)}{\cos(\Omega\ell/c)} U_g(\Omega) e^{j\Omega t}$$ 2

Acoustic impedance $Z_A(\Omega) = j\frac{\rho c}{A}\tan[\Omega(l - x)/c]$

If Δx is very small then from Taylor series expansion we get

$$Z_A(\Omega) \approx j\frac{\rho}{A}\Delta x\Omega$$

$\frac{\rho \Delta x}{A}$ →Can be thought as an acoustic mass



$$p(x,t) = jZ_0 \frac{\sin(\Omega(\ell - x)/c)}{\cos(\Omega\ell/c)} U_g(\Omega) e^{j\Omega t}$$

Envelope $$\frac{\rho c}{A} \frac{\sin[\Omega(l - x/c)]}{\cos[\Omega l/c]}$$

$$u(x,t) = \frac{\cos(\Omega(\ell - x)/c)}{\cos(\Omega\ell/c)} U_g(\Omega) e^{j\Omega t}$$

Envelope $$\frac{\cos[\Omega(l - x/c)]}{\cos[\Omega l/c]}$$

From equation (1) and (2)

$$\mathrm{Re}[u(x,t)] = \frac{\cos[\Omega(l - x/c)]}{\cos[\Omega l/c]} U_g(\Omega)\cos(\Omega t)$$

$$\mathrm{Re}[p(x,t)] = \frac{\rho c}{A} \frac{\cos[\Omega(l - x/c) + \pi/2]}{\cos[\Omega l/c]} U_g(\Omega)\cos(\Omega t + \pi/2)$$

$$u(x,t) = \frac{\cos[\Omega(l - x/c)]}{\cos[\Omega l/c]} U_g(\Omega) e^{j\Omega t}$$

At $x=l$

$$u(l,t) = \frac{\cos[\Omega(l - l/c)]}{\cos[\Omega l/c]} U_g(\Omega) e^{j\Omega t} = \frac{1}{\cos[\Omega l/c]} U_g(\Omega) e^{j\Omega t}$$

$$U(l,t) = \frac{1}{\cos[\Omega l / c]} U_g(\Omega) \quad for \quad complex \quad input \quad U_g(\Omega) e^{j\Omega t}$$

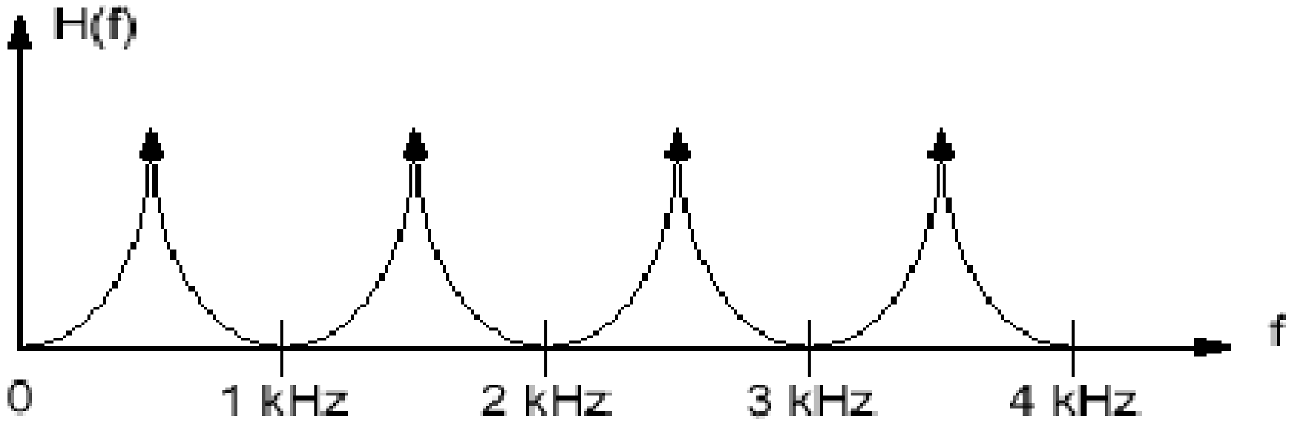
$$\frac{U(l,t)}{U_g(\Omega)} = \frac{1}{\cos[\Omega l / c]} = V_a(\Omega)$$

Frequency response of Uniform tube

$$\Omega = 2\pi f ; c = 35000 cm / \sec; l = 17.5 cm$$

$$\frac{\Omega l}{c} = \frac{2\pi f l}{c}$$

$$\cos\left(\frac{2\pi f l}{c}\right) = 0 \quad when \quad \frac{2\pi f l}{c} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2} ..$$

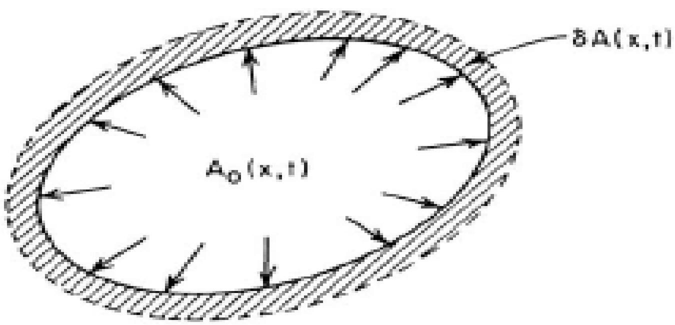


Wall vibrations

- ➢ Assume walls are elastic => cross-sectional area of the tube will change with pressure in the tube
- ➢ Assume walls are 'locally' reacting => *A(x,t) ~ p(x,t)*
- ➢ Assume pressure variations are very small

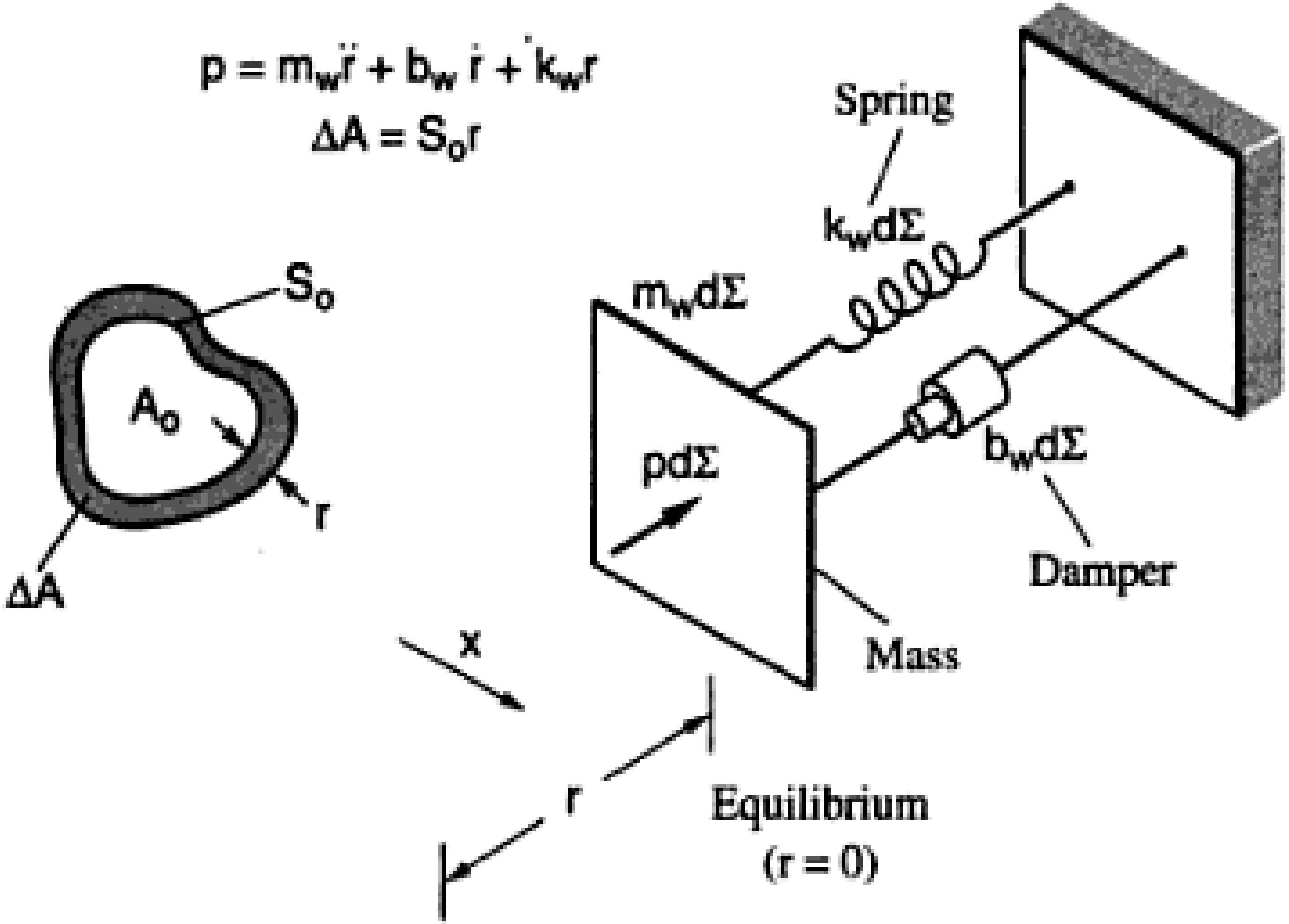
$$A(x,t) = A(x,t) + \delta A(x,t)$$

Neglecting second order terms in *u/ A* and *pA* , the wave equations become



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

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



$p = m_w\ddot{r} + b_w\dot{r} + k_w r$
$\Delta A = S_o r$
Spring
$k_w d\Sigma$
S_o
$m_w d\Sigma$
A_o
$p d\Sigma$
$b_w d\Sigma$
r
ΔA
Damper
Mass
x
r
Equilibrium
(r = 0)

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

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

$m_w(x) =$ 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

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

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

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A_0}\frac{\partial u}{\partial t} \qquad 2$$

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

Under the steady state assumption that sound propagation has occurred long enough so that transient responses have died out and given that the three coupled equations (1,2,3) are linear and time invariant

$$let \quad input \quad u_g(t) = u(0,t) = U(\Omega)e^{j\Omega t}$$

Result in solution of the form

$$p(x,t) = P(x,\Omega)e^{j\Omega t}, \quad u(x,t) = U(x,\Omega)e^{j\Omega t}, \quad \Delta A(x,t) = \Delta A(x,\Omega)e^{j\Omega t}$$

$$-\frac{\partial p(x,\Omega)}{\partial x} = \frac{\rho}{A_0}\Omega U(x,\Omega)$$

$$-\frac{\partial U(x,\Omega)}{\partial x} = \frac{A_0}{\rho c^2}P(x,\Omega) + \Omega\Delta\hat{A}(x,\Omega)$$

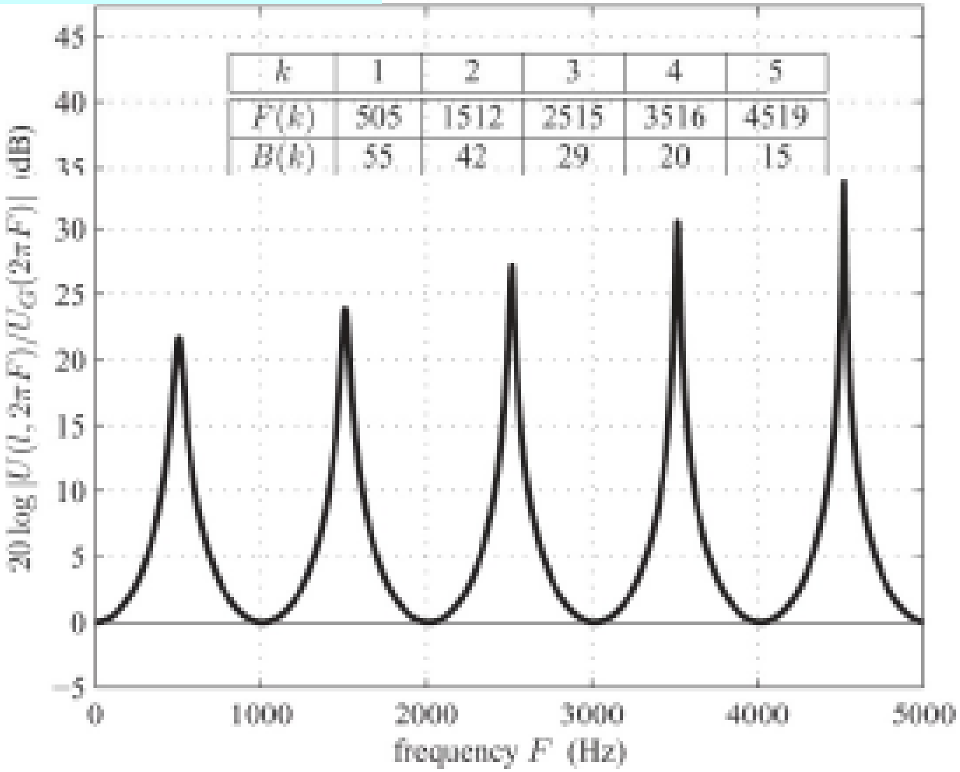
$$P(x,\Omega) = -\Omega^2 m_w \Delta\hat{A}(x,\Omega) + j\Omega b_w \Delta\hat{A}(x,\Omega) + k_w \Delta\hat{A}(x,\Omega)$$

Using estimates for m_W, b_W, and k_W from measurements on body tissue, and with boundary condition at lips of *p(l,t)=0, we get:*

$$V_a(\Omega) = \frac{U(l,\Omega)}{U_g(\Omega)}$$

Length of the tube l=17.5 cm and 5 cm^2 in cross section

m_w =0.4gm/cm^2,b_w=6500 dyne-sec/cm^3, k_w=0

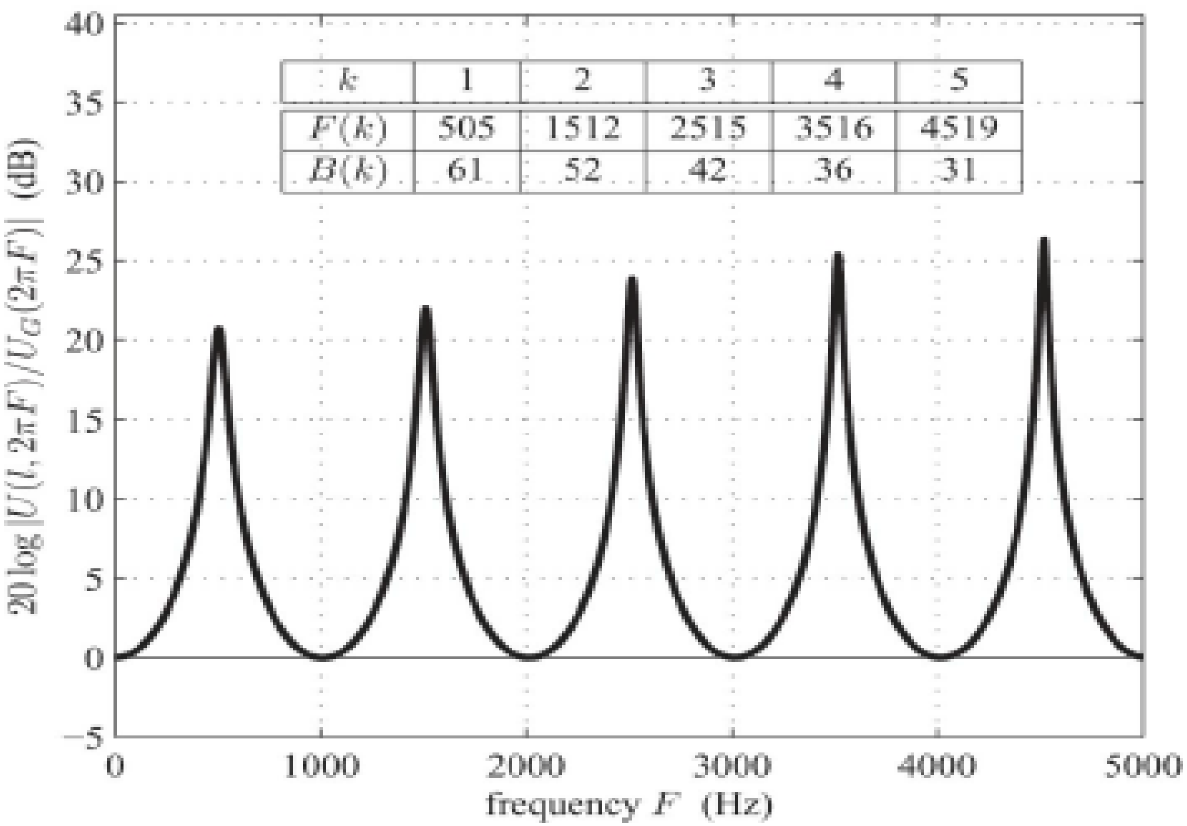


Observation

- Complex poles with non-zero bandwidths
- Slightly higher frequencies for resonances
- Most effect at lower frequencies

Similarly account for effects of viscous friction and thermal conduction at the walls

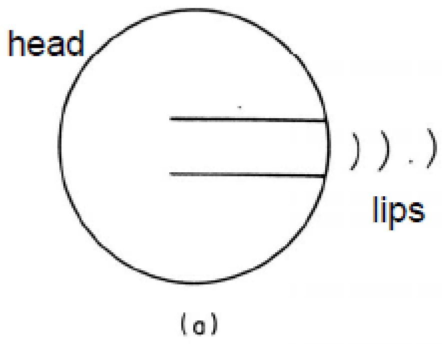
- Increases bandwidth of complex poles
- Decreases resonance frequency (slightly)

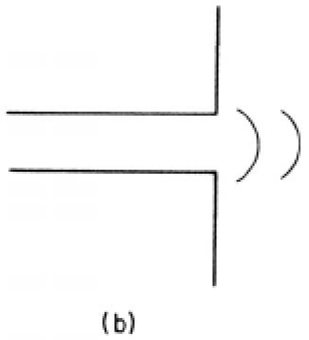


k	1	2	3	4	5
$F(k)$	505	1512	2515	3516	4519
$B(k)$	61	52	42	36	31

Effects of Radiation at Lips

- We assumed $p(l,t)=0$ at the lips (the acoustical analog of a short circuit) => no pressure changes at the lips no matter how much the volume velocity changes at the lips
- In reality, vocal tract tube terminates with open lips, and sometimes open nostrils (for nasal consonants)
- This leads to two models for sound radiation at the lips

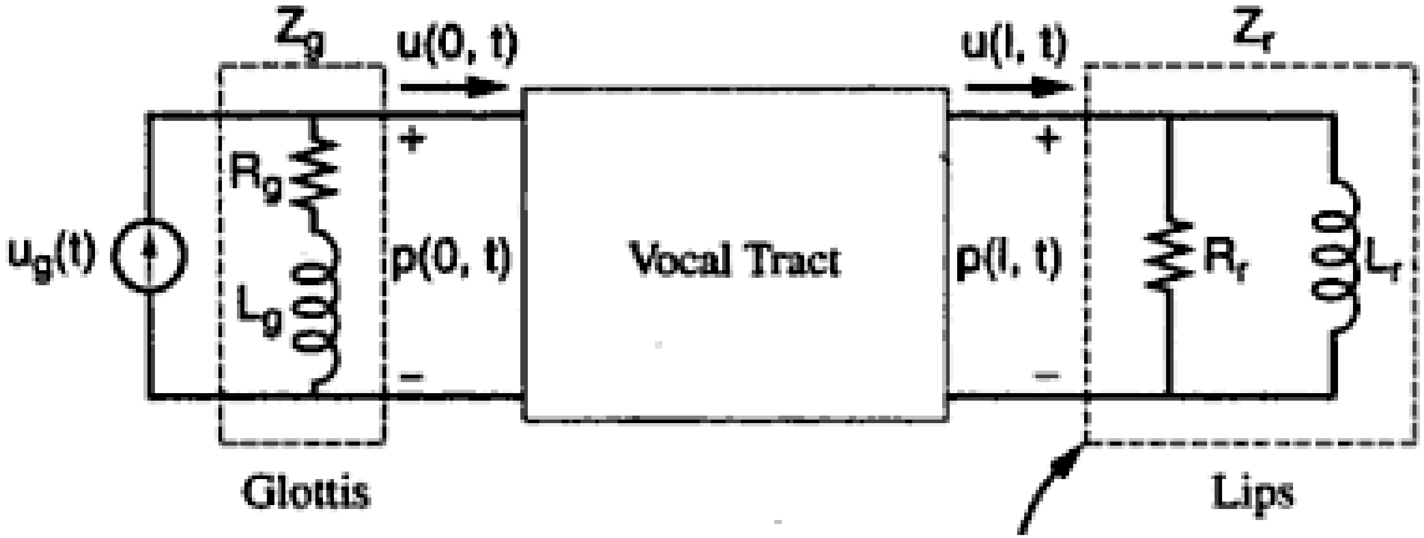




If the lip opening is small

Radiation from a spherical baffle

Radiation from a infinite plane baffle



This 'radiation load' is the equivalent of a parallel connection of a radiation resistance and a radiation inductance

$$P(l,\Omega) = Z_L(\Omega)U(l,\Omega)$$

Where $$Z_L(\Omega) = \frac{j\Omega L_r R_r}{R_r + j\Omega L_r}$$

For infinite baffle, Flanagan has given the value

$$R_r = \frac{128}{9\pi^2} \qquad L_r = \frac{8a}{3\pi c}$$

a is the radius of opening and c is the velocity of sound

$$Z_L(\Omega) = \frac{j\Omega L_r R_r}{R_r + j\Omega L_r}$$

1. At low frequencies, $Z_L(\Omega) \approx 0$

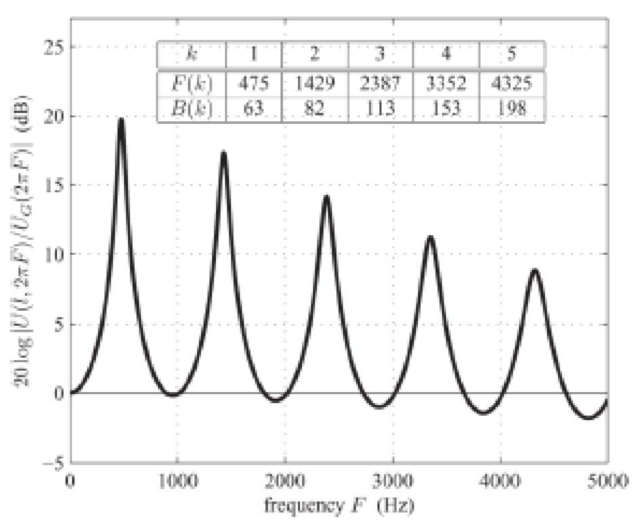
short circuit termination which is the old solution

2. At mid-range frequencies $Z_L(\Omega) \approx j\Omega L_r$ $\quad R_r >> \Omega L_r$

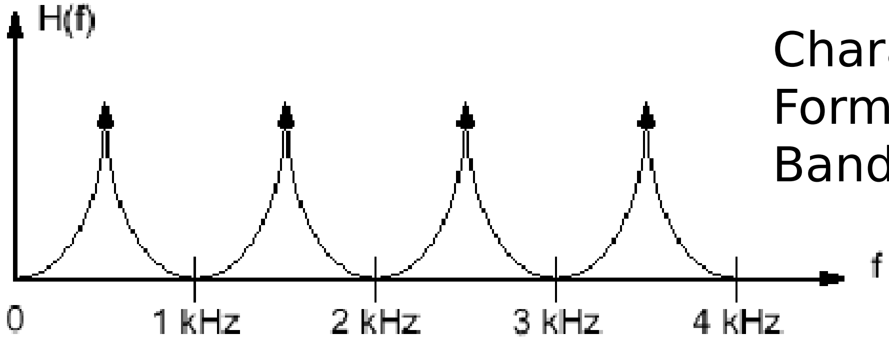
3. At higher frequencies $Z_L(\Omega) \approx R_r$ $\quad R_r << \Omega L_r$

Overall Transfer Function

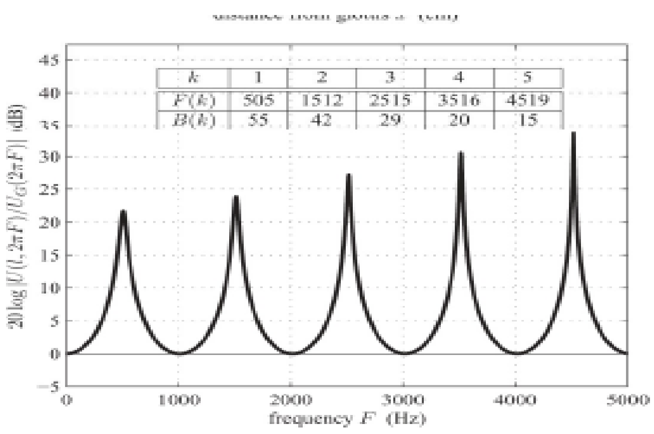
$$H(\Omega) = \frac{P(l,\Omega)}{U_g(\Omega)} = \frac{P(l,\Omega)}{U(l,\Omega)} \frac{U(l,\Omega)}{U_g(\Omega)} = Z_r(\Omega) V_a(\Omega)$$



Frequency response of Uniform tube in no loss condition

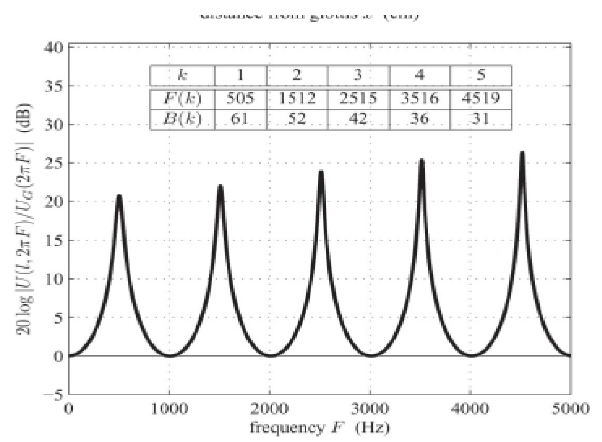


Characterize by set of Formant Frequency Formant Band with is zero



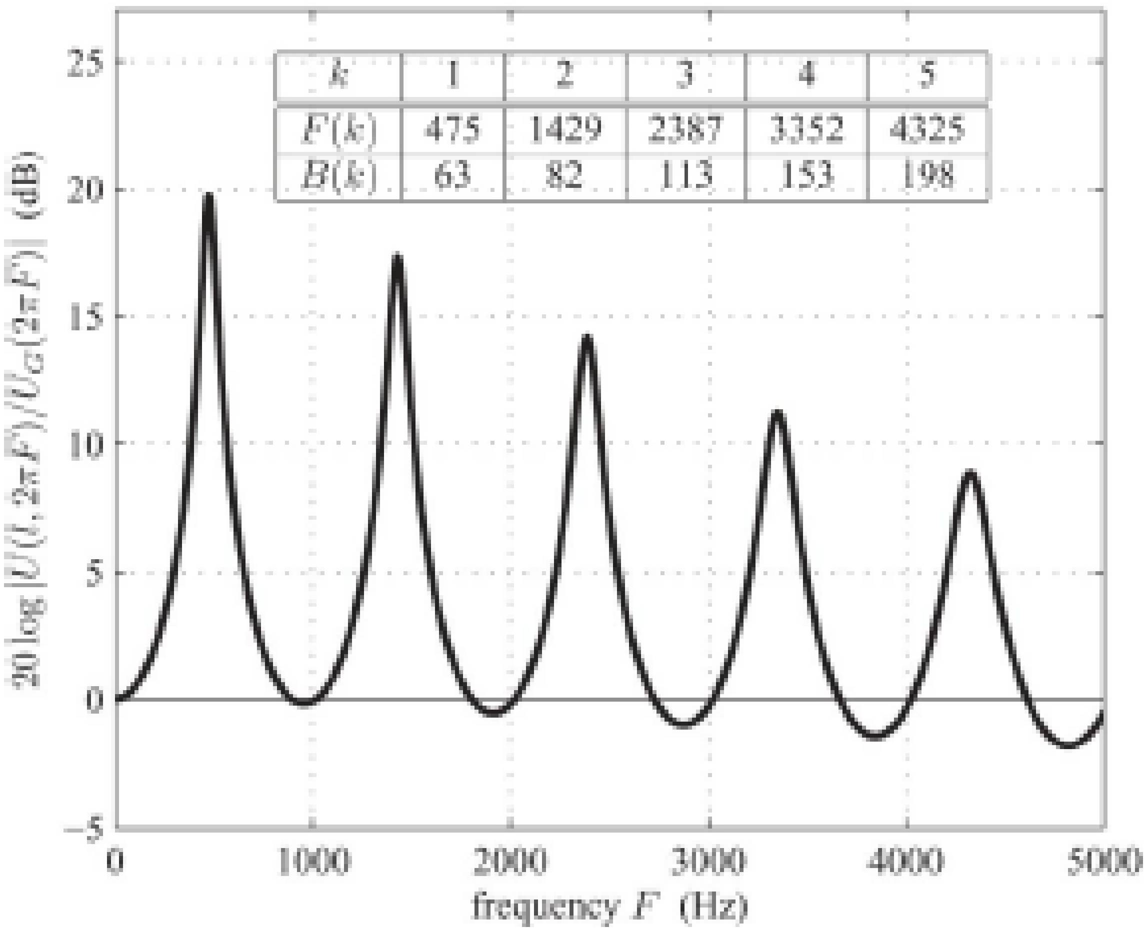
k	1	2	3	4	5
$F(k)$	505	1512	2515	3516	4519
$B(k)$	55	42	29	20	15

Effect due to the wall vibration



k	1	2	3	4	5
$F(k)$	505	1512	2515	3516	4519
$B(k)$	61	52	42	36	31

Effects of viscous friction and thermal conduction at the walls



Effects of Radiation at Lips

VT Transfer Functions

- The vocal tract can be characterized by a set of resonances (formants) that depend on the vocal tract area function-with shifts due to losses and radiation
- The bandwidths of the two lowest resonances (F1 and F2) depend primarily on the vocal tract wall losses
- The bandwidths of the highest resonances (F3, F4,) depend primarily on viscous friction losses, thermal losses, and radiation losses

Nasal Coupling Effects

At the branching point

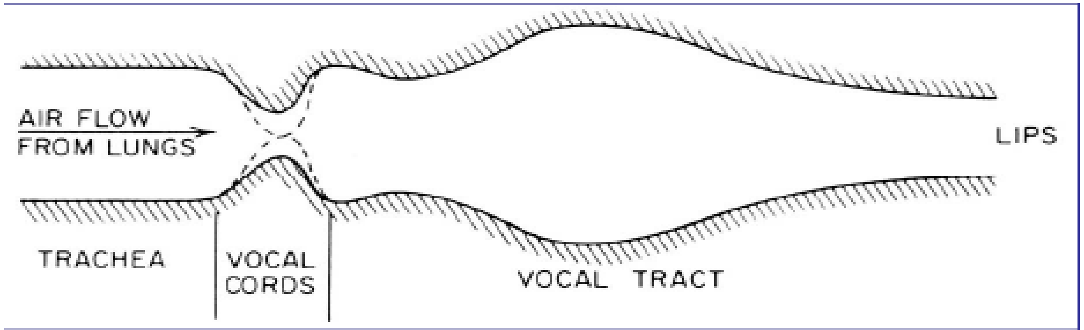
- ❑ Sound pressure the same as at input of each tube
- ❑ Volume velocity is the sum of the volume velocities at inputs to nasal and oral cavities

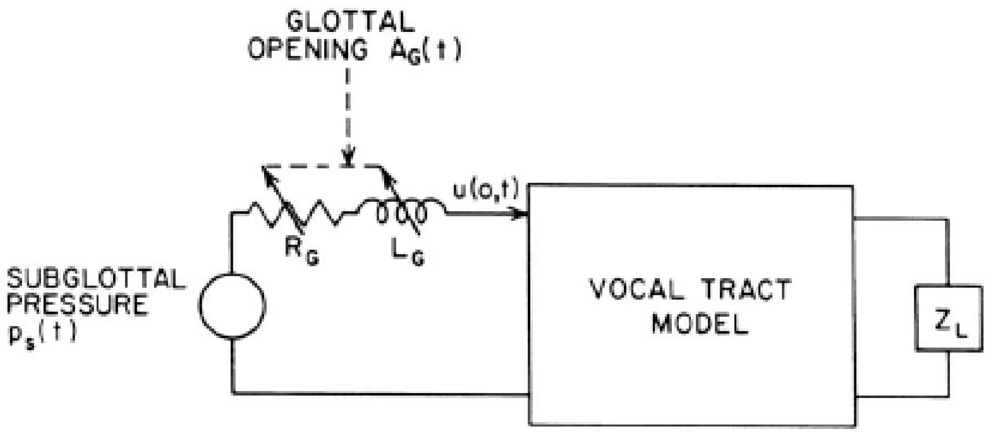
Closed oral cavity can trap energy at certain frequencies, preventing those frequencies from appearing in the nasal output => anti-resonances or zeros of the transfer function

Nasal resonances have broader bandwidths than non-nasal voiced sounds => due to greater viscous friction and thermal loss due to large surface area of the nasal cavity

Sound Excitation

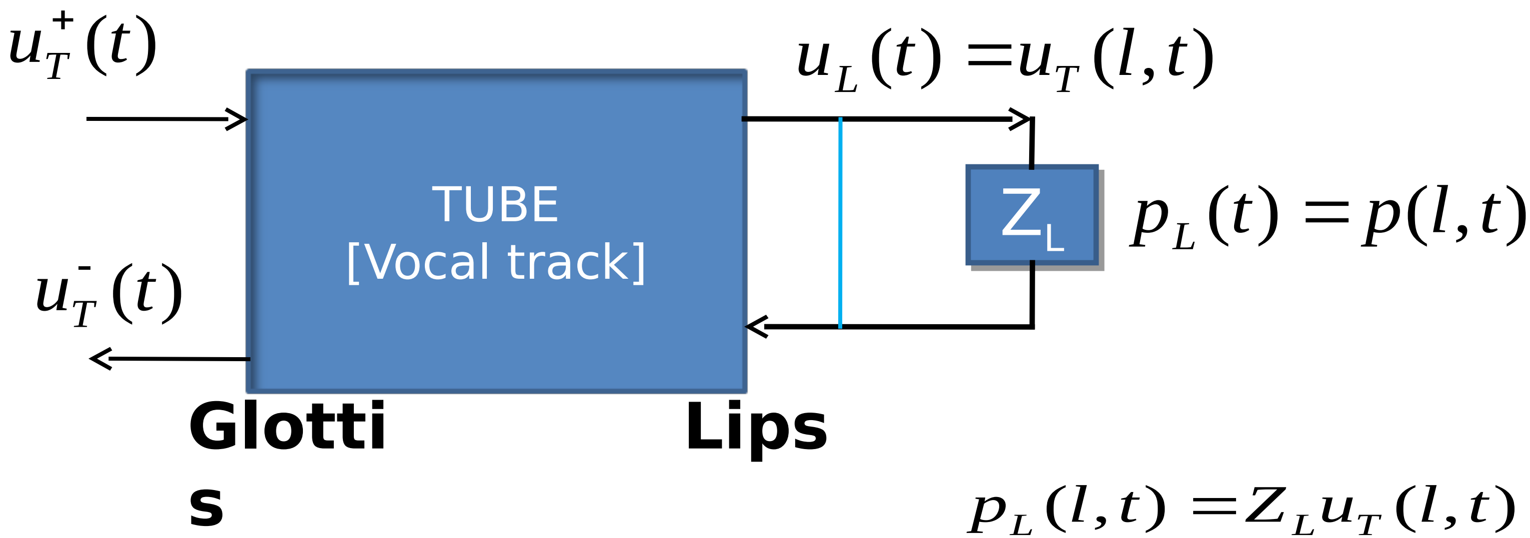
1. Air flow from lungs is modulated by vocal cord vibration, resulting in a quasi-periodic pulse-like source
2. Air flow from lungs becomes turbulent as air passes through a constriction in the vocal tract, resulting in a noise-like source
3. Air flow builds up pressure behind a point of total closure in the vocal tract => the rapid release of this pressure, by removing the constriction, causes a transigent excitation (pop like sound)





- Vocal tract acts as a load on the vocal cord oscillator
- Time varying glottal resistance and inductance-both functions of *1/AG(t)* => when *AG(t)=0 (total closure), impedance* is infinite and volume velocity is zero

Boundary condition (lips)

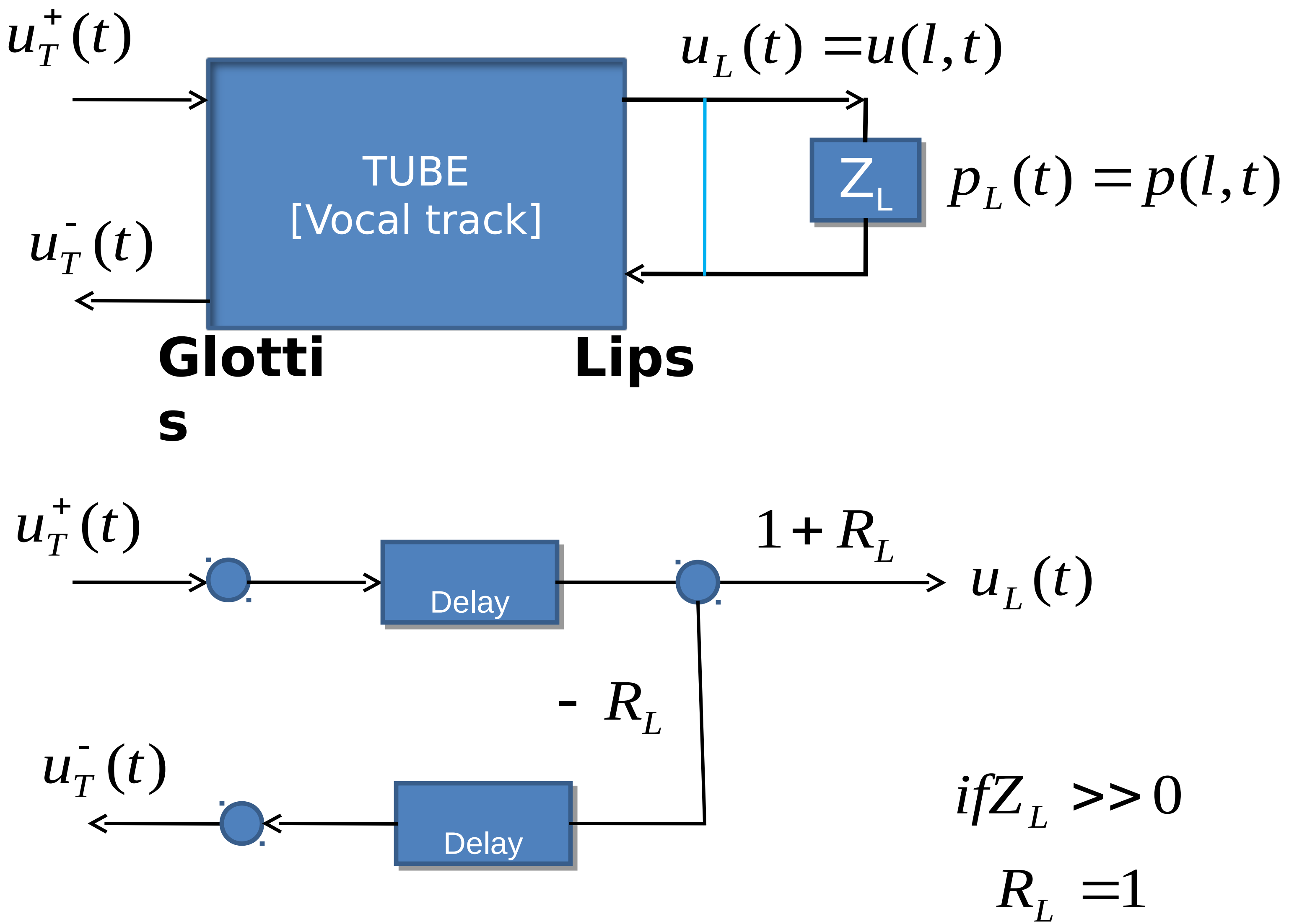


$$p_L(l,t) = Z_L u_T(l,t)$$

$$Z_T[u_T^+(t-l/c) + u_T^-(t+l/c)] = Z_L[u_T^+(t-l/c) - u_T^-(t+l/c)]$$

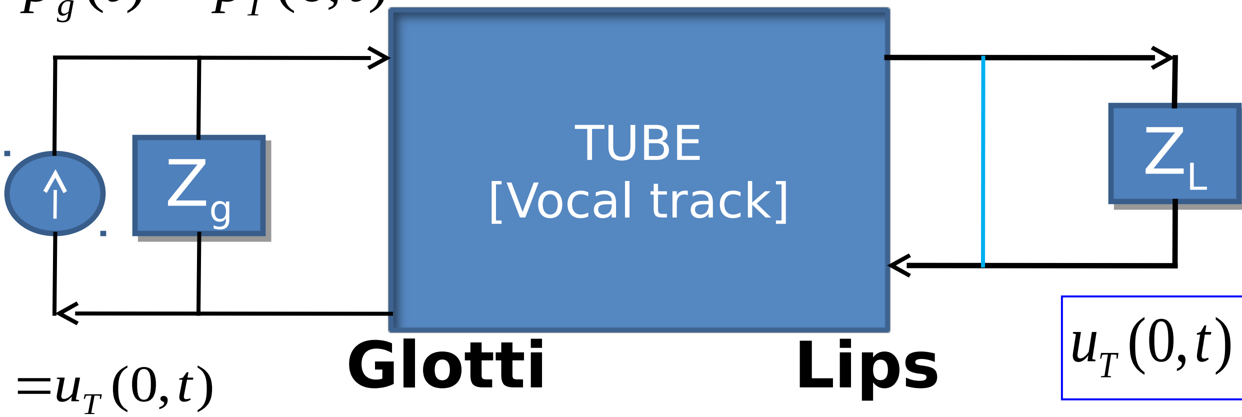
$$u_T^-(t+l/c)] = -\frac{Z_T - Z_L}{Z_T + Z_L} u_T^+(t-l/c)]$$

$$\boxed{u_L(t) = 1 + \frac{Z_T - Z_L}{Z_T + Z_L} u_T^+(t-l/c) = [1+R_L]u_T^+(t-l/c)}$$


$u_T^+(t)$
TUBE
[Vocal track]
$u_L(t) = u(l,t)$
Z_L
$p_L(t) = p(l,t)$
$u_T^-(t)$
Glottis
Lips
$u_T^+(t)$
Delay
$1 + R_L$
$u_L(t)$
$-R_L$
$u_T^-(t)$
Delay
$if Z_L >> 0$
$R_L = 1$

Boundary condition (glottis)

$$p_g(t) = p_T(0,t)$$



$$p_L(t) = p(l,t)$$

$$u_g(t) = u_T(0,t)$$

$$\boxed{u_T(0,t) = u_g(t) = p_g(t)/Z_g}$$

$$\boxed{u_T(0,t) = u_g(t) - \frac{Z_T}{Z_g}[u_T^+(t) + u_T^-(t)]}$$

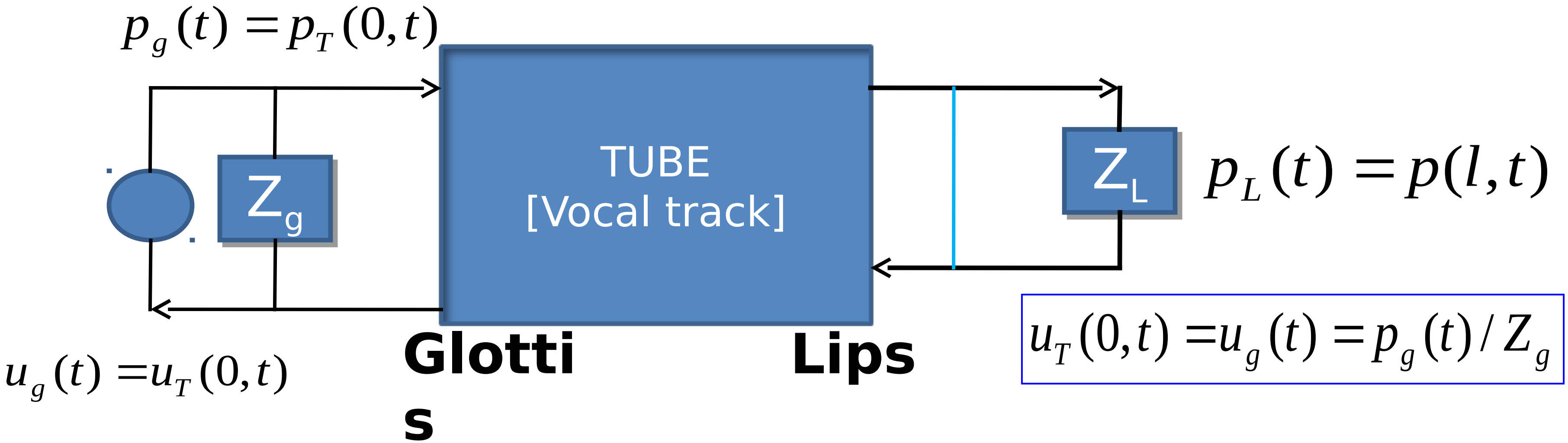
$$\boxed{u_T^+(t) + u_T^-(t) = u_g(t) - \frac{Z_T}{Z_g}[u_T^+(t) + u_T^-(t)]} \Rightarrow \boxed{u_T{}^+(t) = \frac{u_g(t)}{1 + \frac{Z_T}{Z_g}} + \frac{1 - Z_T/Z_g}{1 + Z_T/Z_g} u_T^-(t)}$$

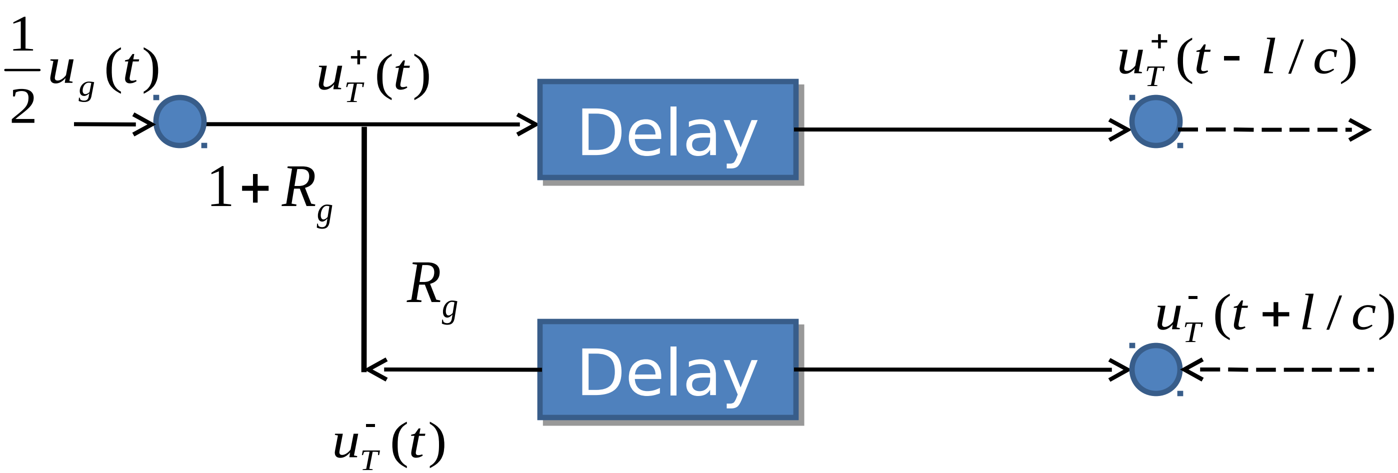
$$\Downarrow$$

$$\boxed{u_T{}^+(t) = \frac{1 + R_g}{2} u_g(t) + R_g u_T^-(t)}$$

$$R_g = \frac{Z_g - Z_T}{Z_g + Z_T}$$

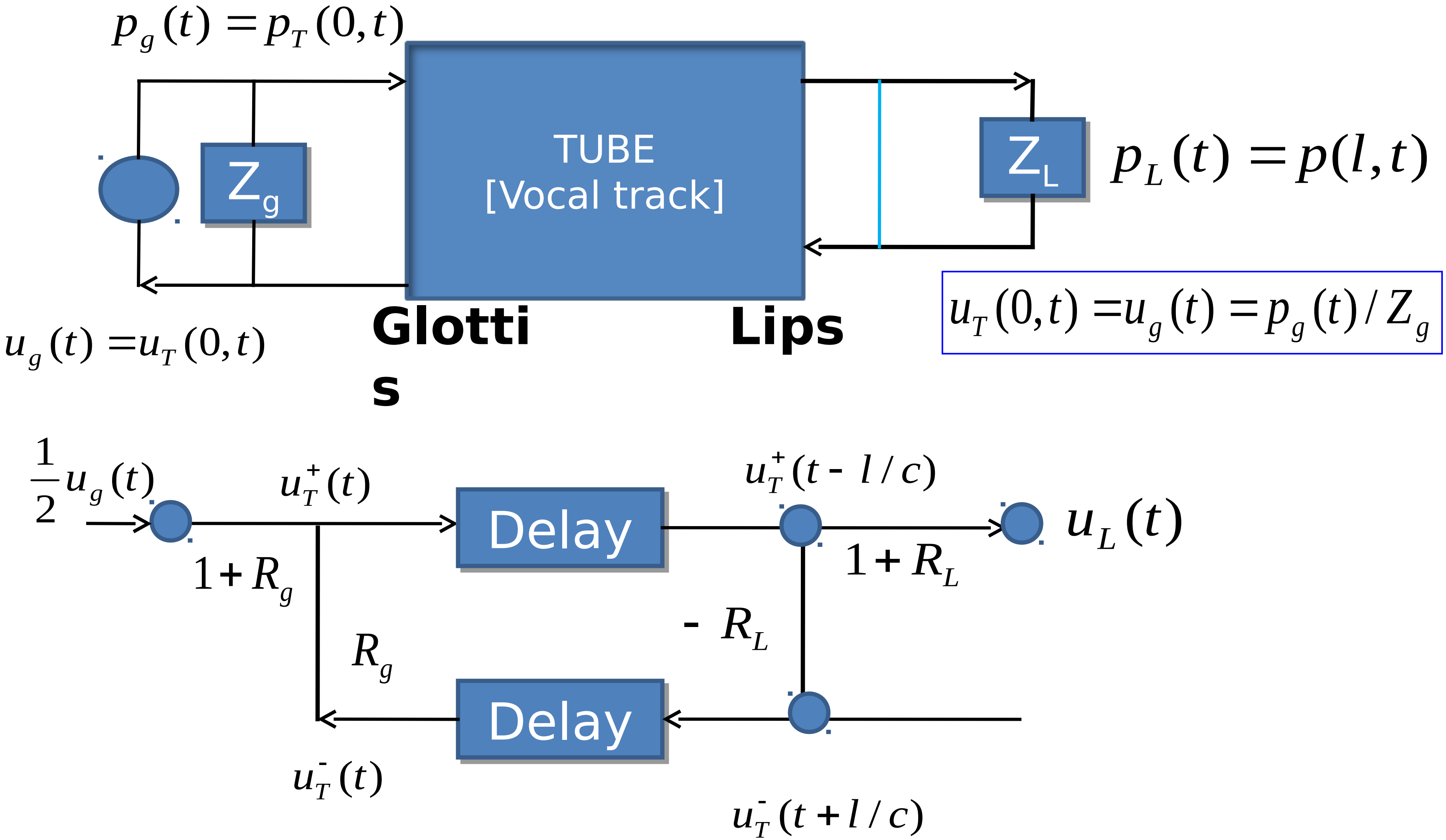
Boundary condition (glottis)



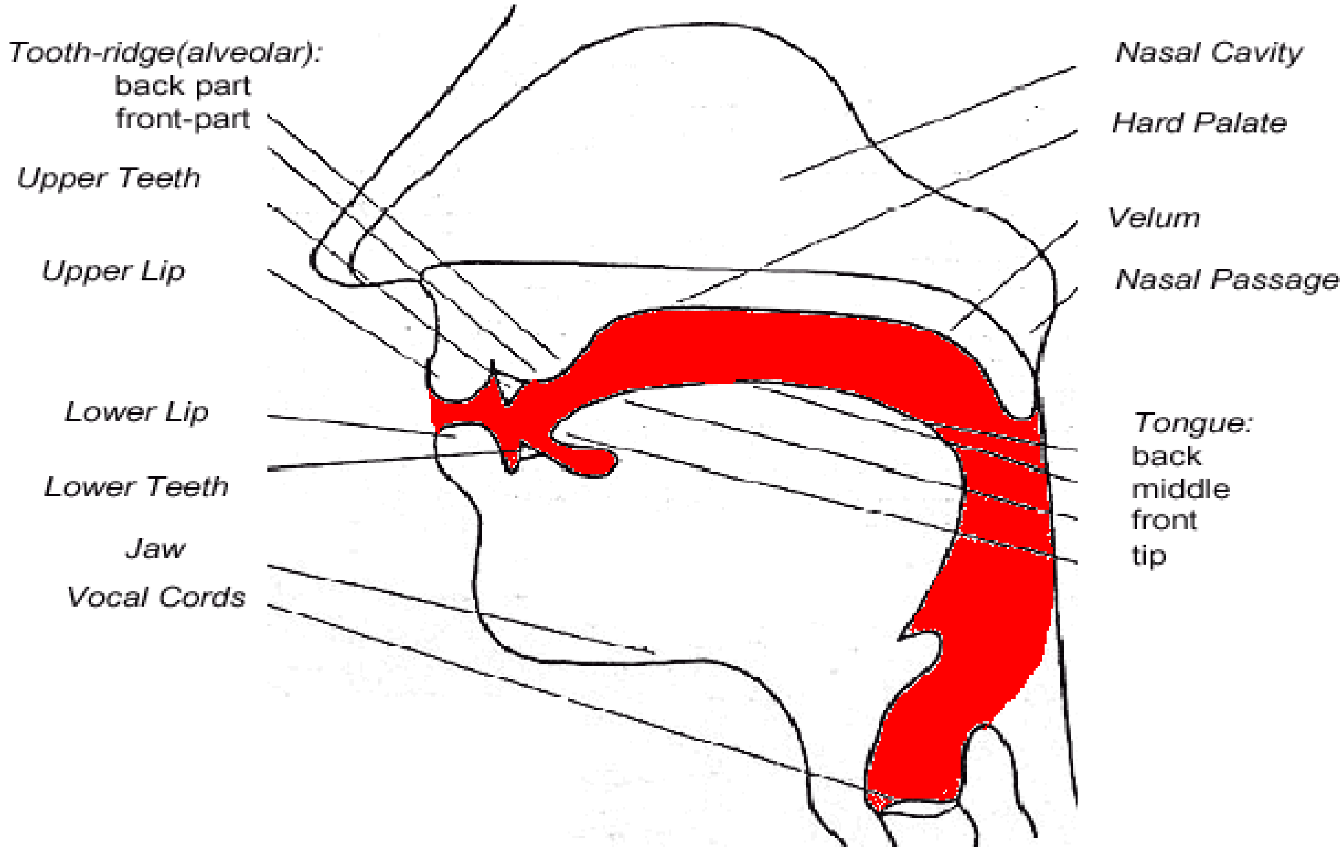


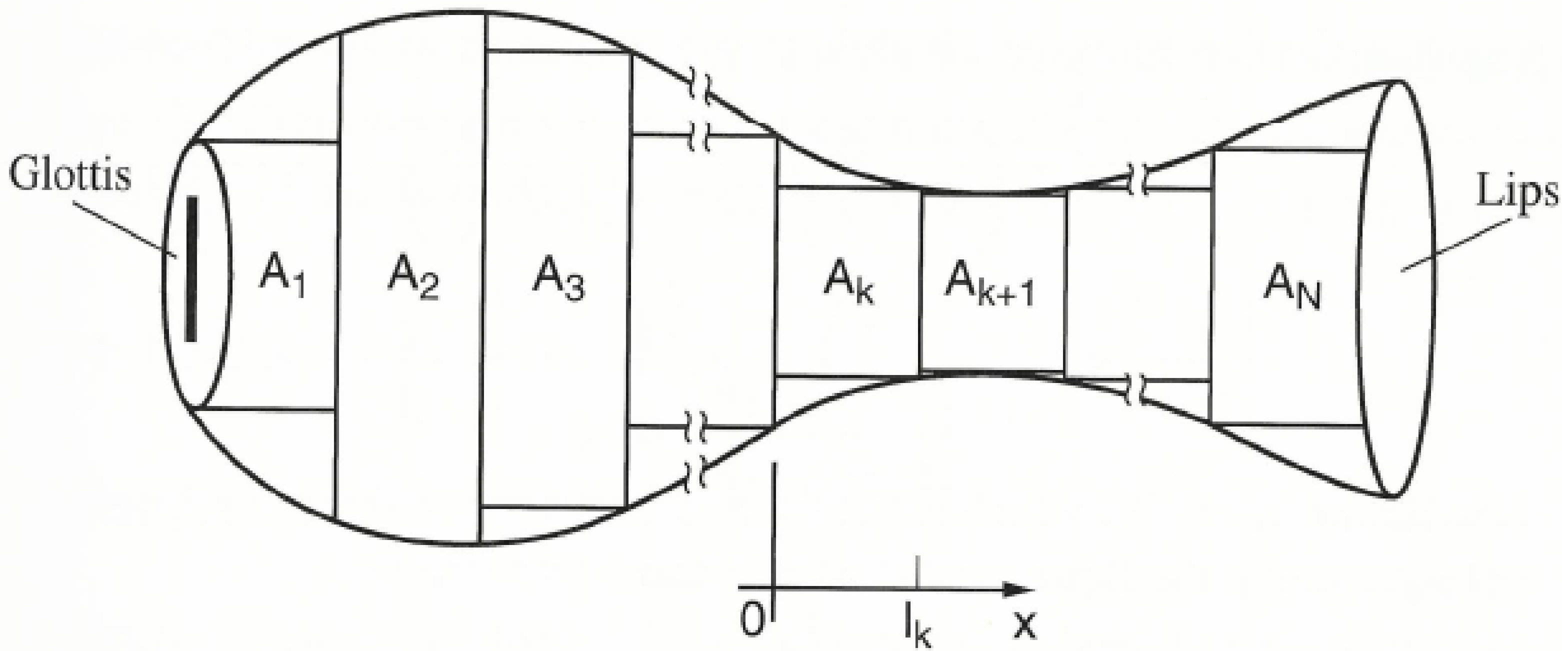
If $Z_g >> 0$ *then*
$R_g = 1$

Boundary condition (glottis)

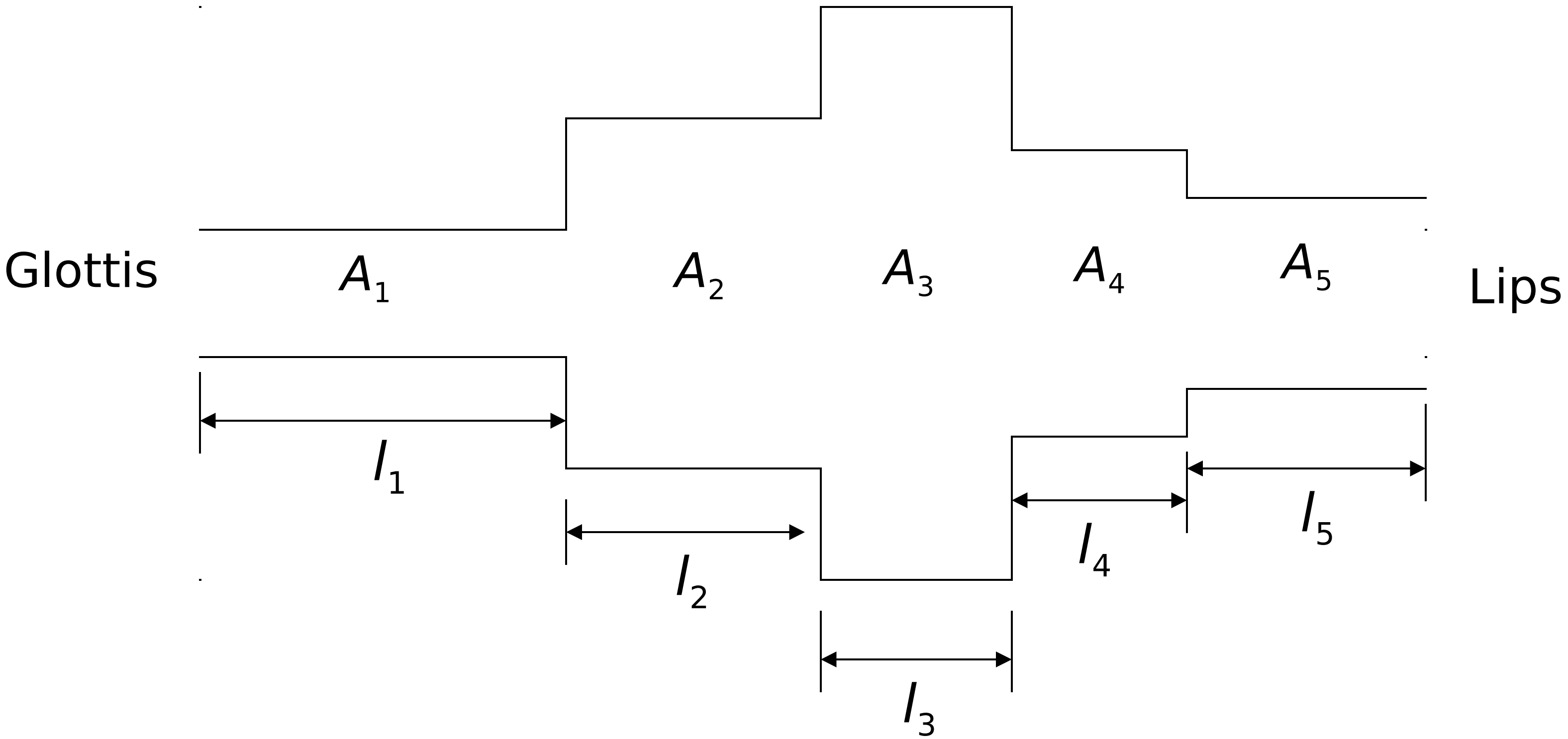


Speech Production System

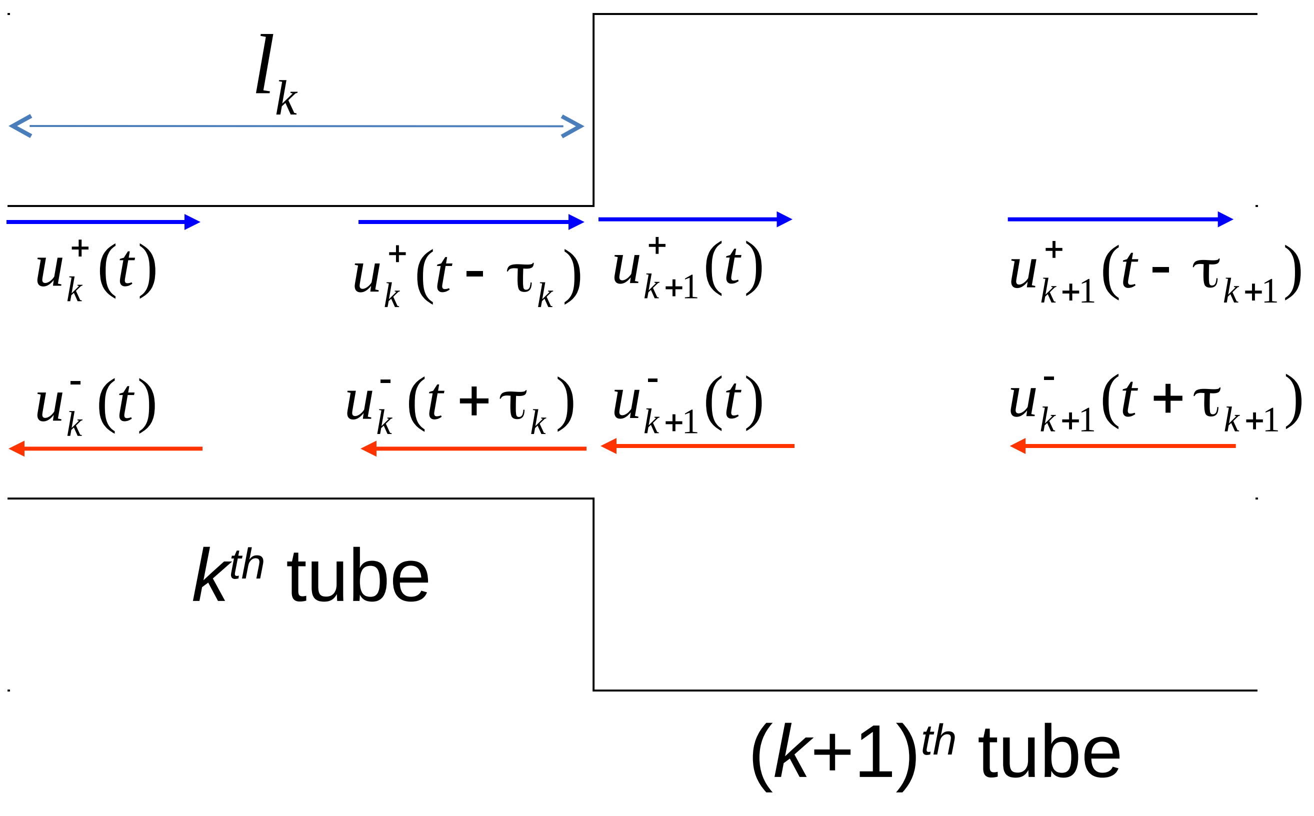



Glottis
A_1
A_2
A_3
A_k
A_{k+1}
A_N
Lips
0
l_k
x

Concatenated Tubes



Wave Propagation at the Junction



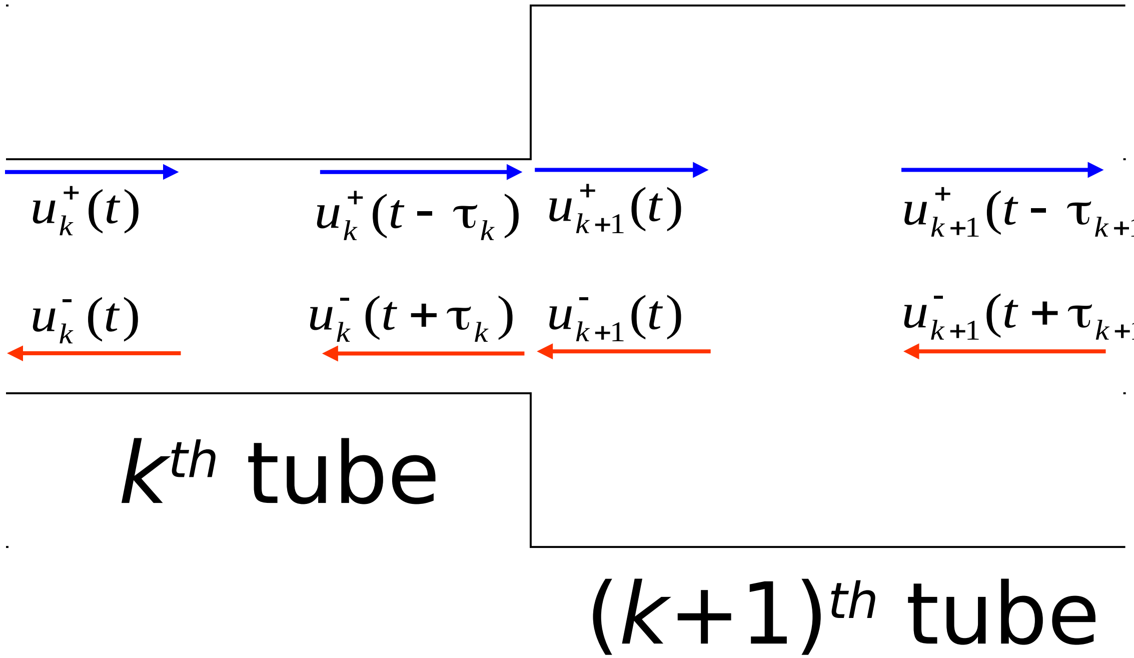
Wave Propagation at the Junction

$$u_k(x,t) = u_k^+(t-\tau_k) - u_k^-(t+\tau_k)$$

$$p_k(x,t) = Z_k[u_k^+(t-\tau_k) + u_k^-(t+\tau_k)]$$

$$u_{k+1}(x,t) = u_{k+1}^+(t) - u_{k+1}^-(t)$$

$$p_{k+1}(x,t) = Z_{k+1}[u_{k+1}^+(t) + u_{k+1}^-(t)]$$



$$Z_k = \frac{\rho c}{A_k}$$

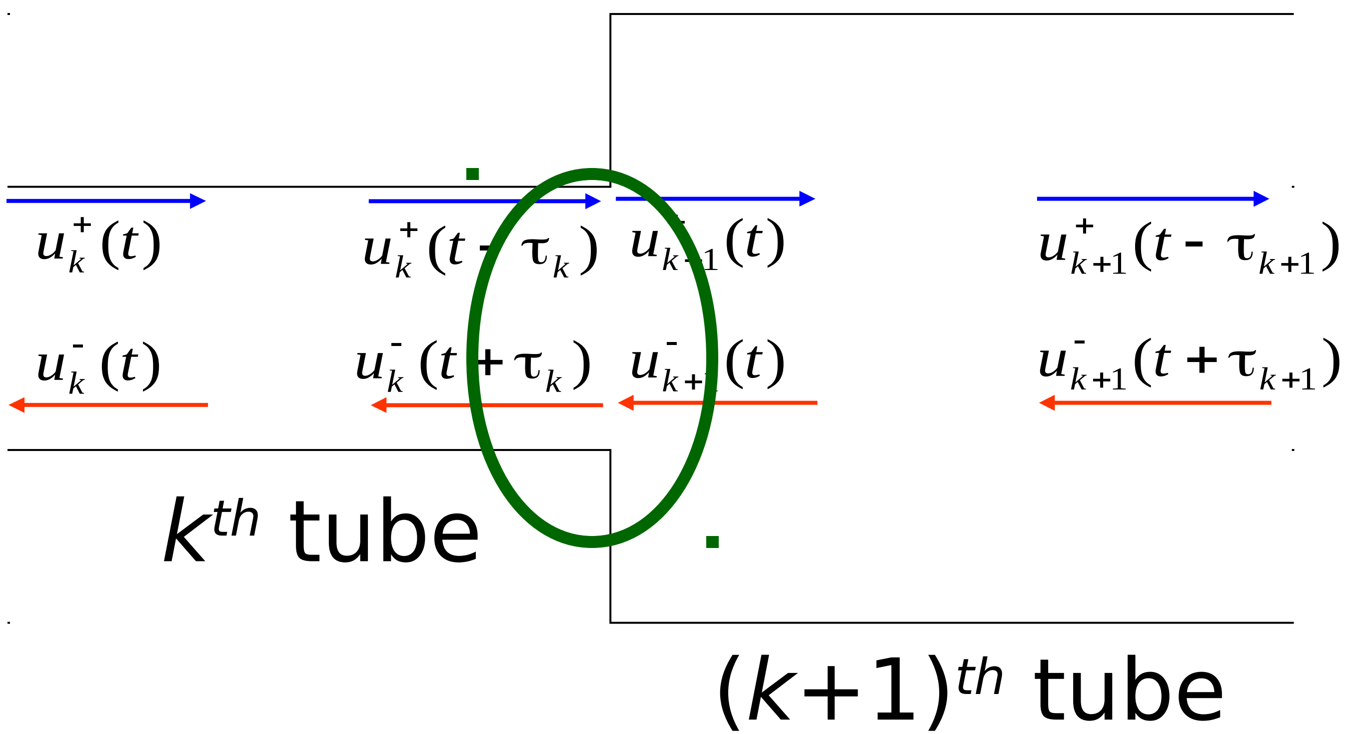
$$Z_{k+1} = \frac{\rho c}{A_{k+1}}$$

$$u_k(x,t) = u_k^+(t - \tau_k) - u_k^-(t + \tau_k)$$

$$u_{k+1}(x,t) = u_{k+1}^+(t) - u_{k+1}^-(t)$$

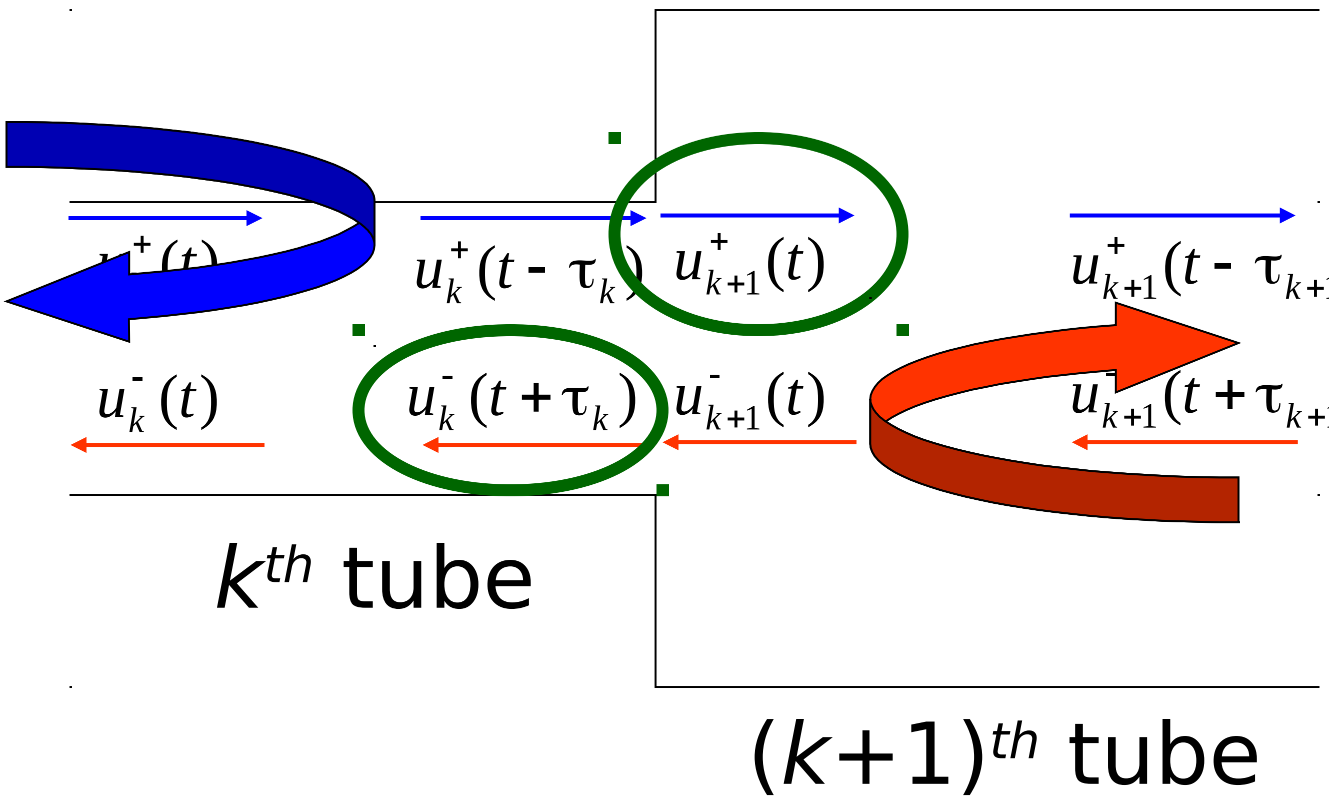
$$p_k(x,t) = Z_k[u_k^+(t - \tau_k) + u_k^-(t + \tau_k)]$$

$$p_{k+1}(x,t) = Z_{k+1}[u_{k+1}^+(t) + u_{k+1}^-(t)]$$



$$u_k^+(t-\tau_k) - u_k^-(t+\tau_k) = u_{k+1}^+(t) - u_{k+1}^-(t)$$

$$Z_k[u_k^+(t-\tau_k) + u_k^-(t+\tau_k)] = Z_{k+1}[u_{k+1}^+(t) + u_{k+1}^-(t)]$$



$$\boxed{u_k^+(t-\tau_k) - u_k^-(t+\tau_k) = u_{k+1}^+(t) - u_{k+1}^-(t)}$$

$$\boxed{Z_k[u_k^+(t-\tau_k) + u_k^-(t+\tau_k)] = Z_{k+1}[u_{k+1}^+(t) + u_{k+1}^-(t)]}$$

$$u_k^+(t-\tau_k) - u_k^-(t+\tau_k) = u_{k+1}^+(t) - u_{k+1}^-(t)$$

$$\frac{Z_k}{Z_{k+1}}[u_k^+(t-\tau_k) + u_k^-(t+\tau_k)] = u_{k+1}^+(t) + u_{k+1}^-(t)$$

$$u_k^-(t+\tau_k) + u_{k+1}^+(t) = u_k^+(t-\tau_k) + u_{k+1}^-(t)$$

$$-\frac{Z_k}{Z_{k+1}}u_k^-(t+\tau_k) + u_{k+1}^+(t) = \frac{Z_k}{Z_{k+1}}u_k^+(t-\tau_k) - u_{k+1}^-(t)$$

$$u_{k+1}^{+}(t) = \left[\frac{2Z_k}{Z_k + Z_{k+1}} \right] u_k^{+}(t - \tau_k) + \left[\frac{Z_k - Z_{k+1}}{Z_k + Z_{k+1}} \right] u_{k+1}^{-}(t)$$

$$u_k^{-}(t + \tau_k) = \left[\frac{-Z_k + Z_{k+1}}{Z_k + Z_{k+1}} \right] u_k^{+}(t - \tau_k) + \left[\frac{2Z_{k+1}}{Z_k + Z_{k+1}} \right] u_{k+1}^{-}(t)$$

Define $r_k = \dfrac{Z_k - Z_{k+1}}{Z_k + Z_{k+1}}$

$$u_{k+1}^{+}(t) = \left[\frac{2Z_k}{Z_k + Z_{k+1}}\right] u_k^{+}(t - \tau_k) + \left[\frac{Z_k - Z_{k+1}}{Z_k + Z_{k+1}}\right] u_{k+1}^{-}(t)$$

$$u_k^{-}(t + \tau_k) = \left[\frac{-Z_k + Z_{k+1}}{Z_k + Z_{k+1}}\right] u_k^{+}(t - \tau_k) + \left[\frac{2Z_{k+1}}{Z_k + Z_{k+1}}\right] u_{k+1}^{-}(t)$$

$$u_{k+1}^{+}(t) = (1 + r_k) u_k^{+}(t - \tau_k) + r_k u_{k+1}^{-}(t)$$

$$u_k^{-}(t + \tau_k) = -r_k u_k^{+}(t - \tau_k) + (1 - r_k) u_{k+1}^{-}(t)$$

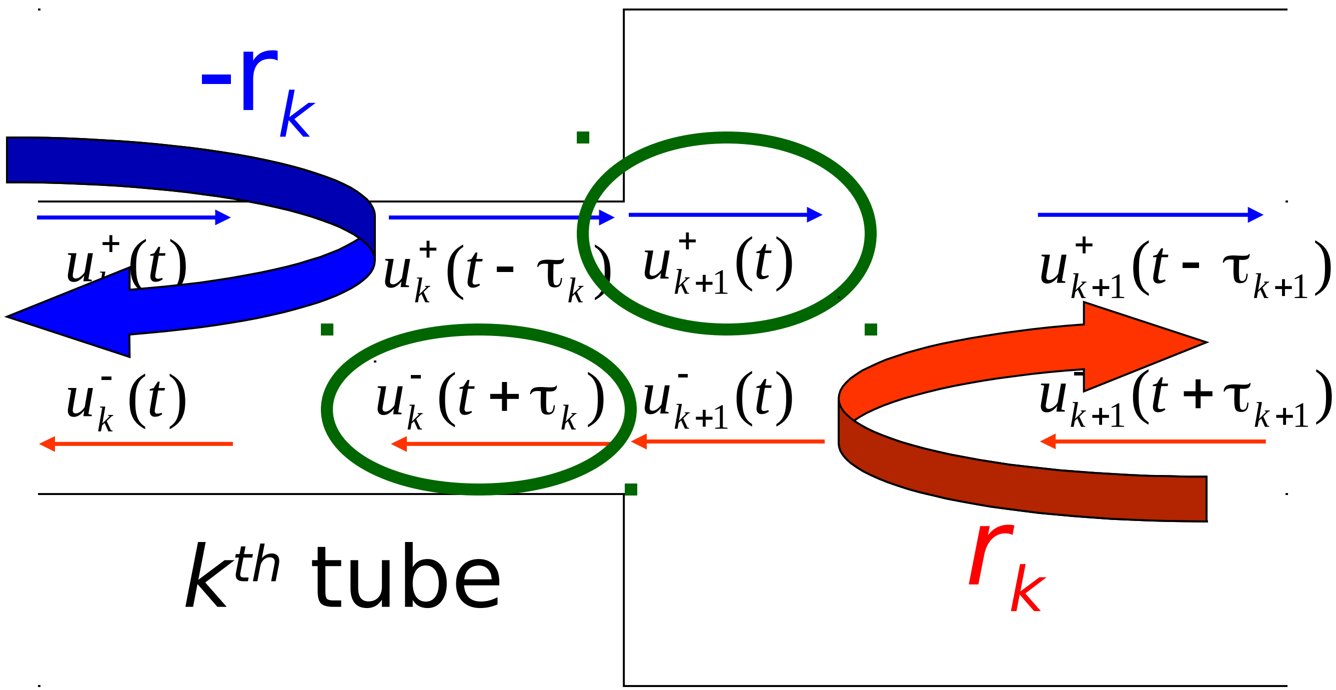
$$Z_k = \frac{\rho c}{A_k}$$

$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k}$$

$$Z_{k+1} = \frac{\rho c}{A_{k+1}}$$

$$u^{+}_{k+1}(t) = (1+r_k)\,u^{+}_{k}(t-\tau_k) + r_k\,u^{-}_{k+1}(t)$$

$$u^{-}_{k}(t+\tau_k) = -r_k\,u^{+}_{k}(t-\tau_k) + (1-r_k)\,u^{-}_{k+1}(t)$$



$(k+1)^{th}$ tube

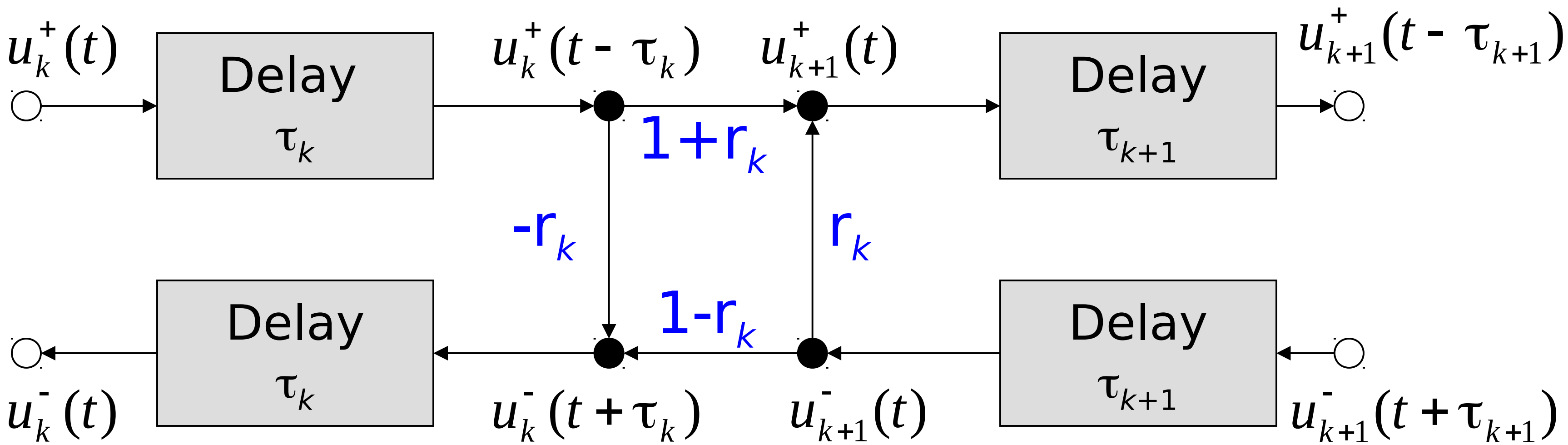
$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k}$$

$r_k > 0$: $A_{k+1} > A_k$

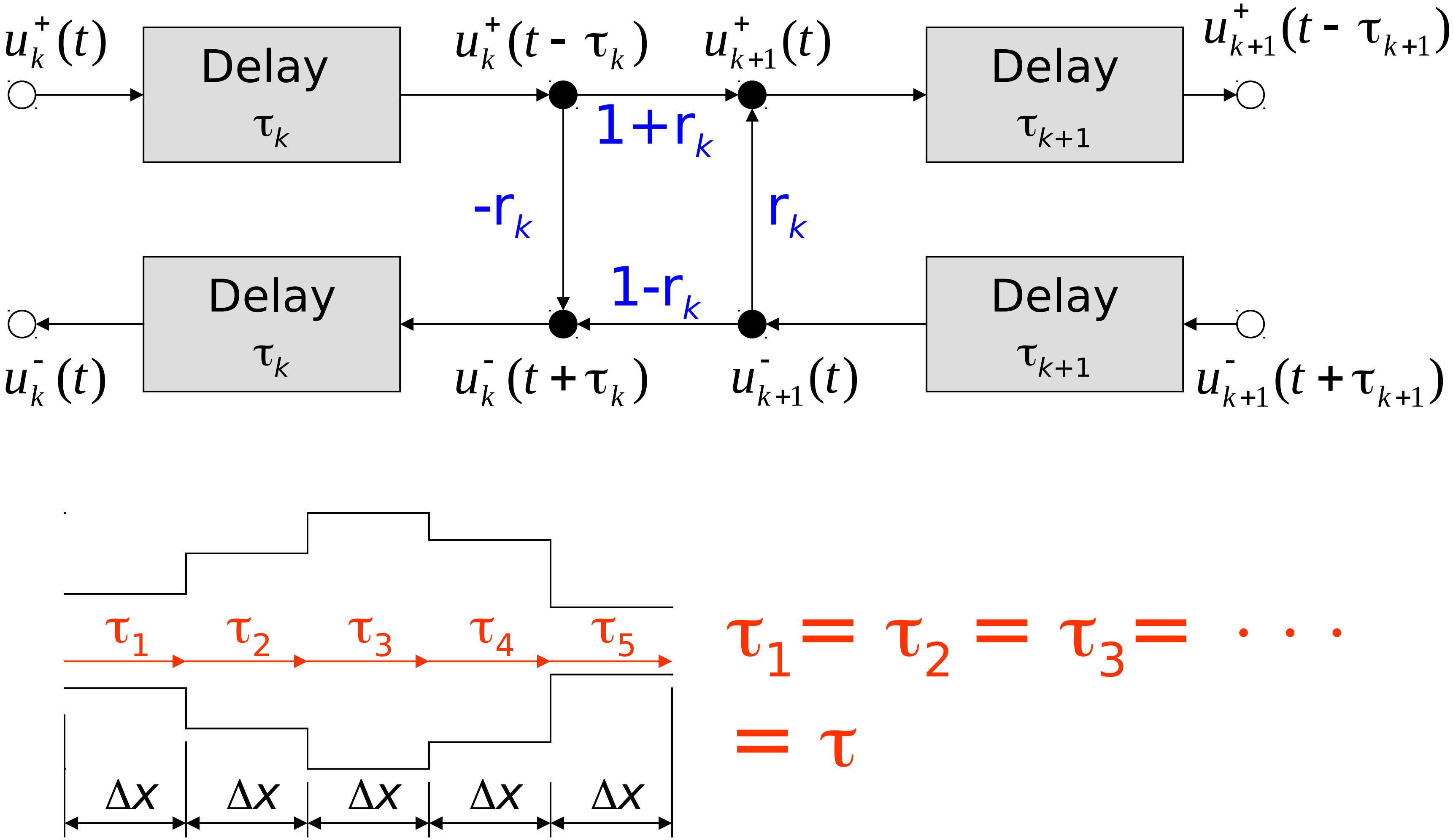
$r_k < 0$: $A_{k+1} < A_k$

$$u_{k+1}^{+}(t) = (1 + r_k) u_k^{+}(t - \tau_k) + r_k u_{k+1}^{-}(t)$$

$$u_k^{-}(t + \tau_k) = -r_k u_k^{+}(t - \tau_k) + (1 - r_k) u_{k+1}^{-}(t)$$



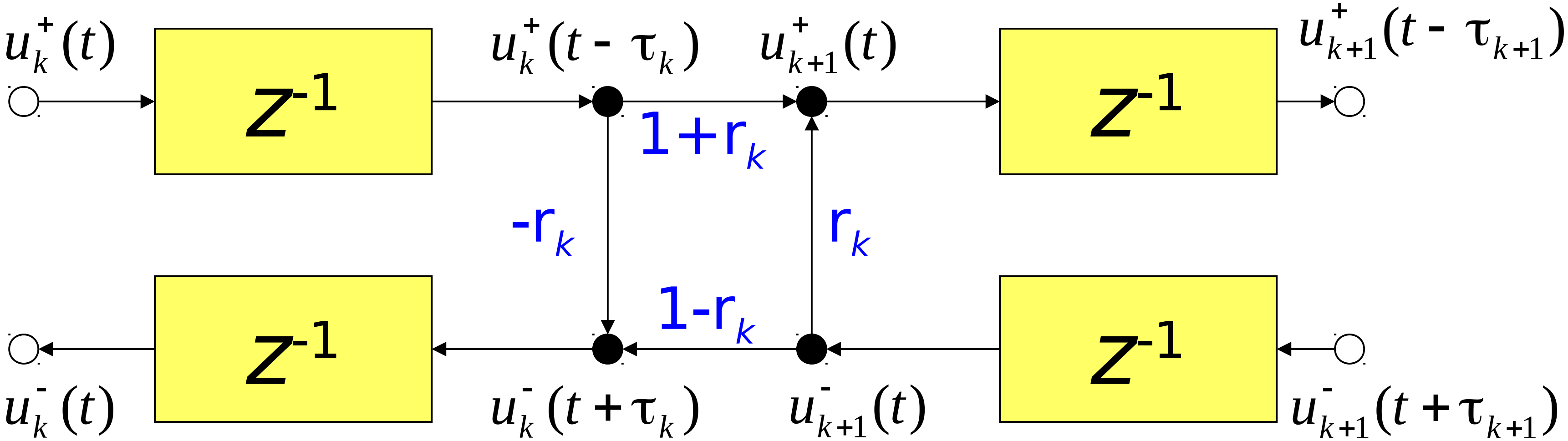
Equal Spaced Tubes

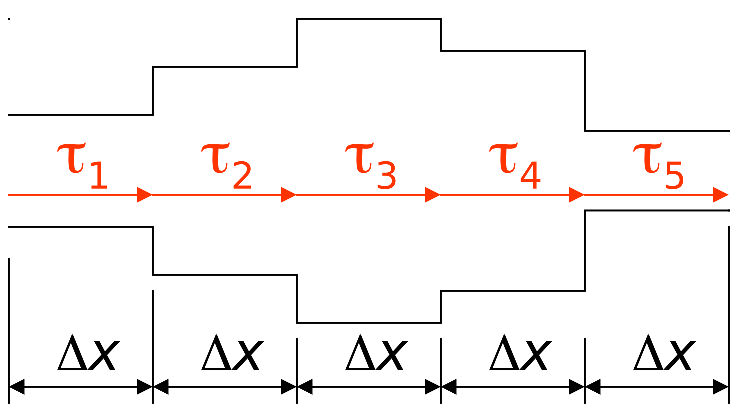


$$\tau_1 = \tau_2 = \tau_3 = \cdots = \tau$$

Equal Spaced Tubes

Let sampling period $T_1 = \tau$

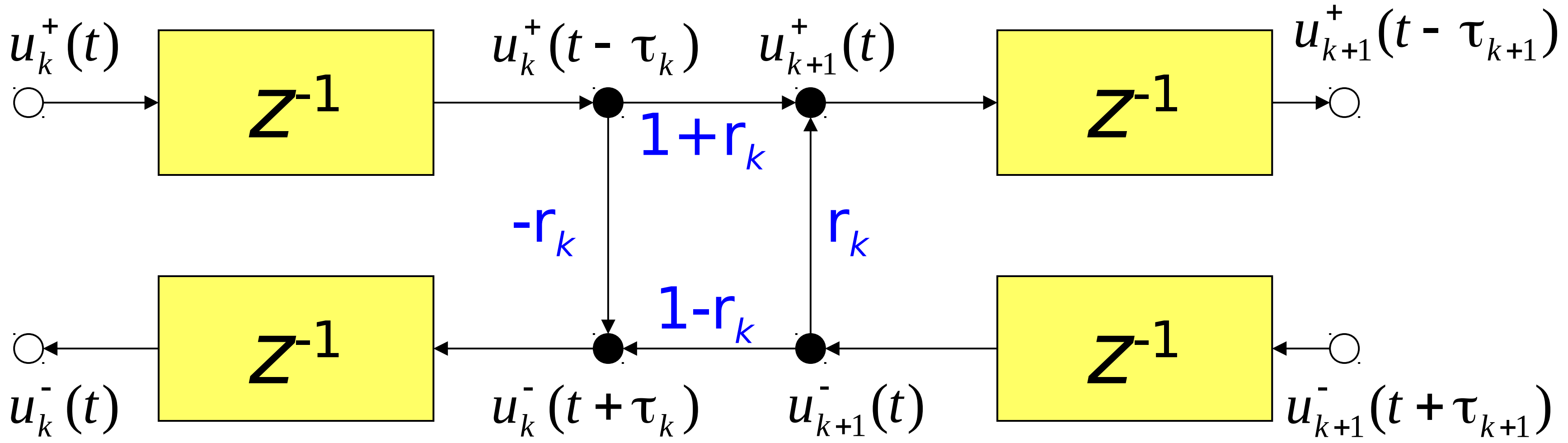


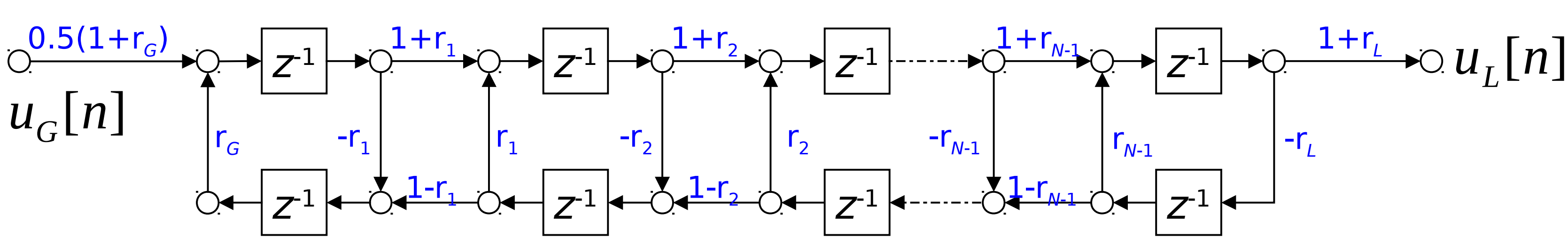


$$\tau_1 = \tau_2 = \tau_3 = \cdots = \tau$$

Vocal Tract Model

Let sampling period $T_1 = \tau$





$$v_a(t) = b_0\delta(t - N\tau) + \sum_{k=1}^{\infty} b_k\delta(t - N\tau - k2\tau)$$

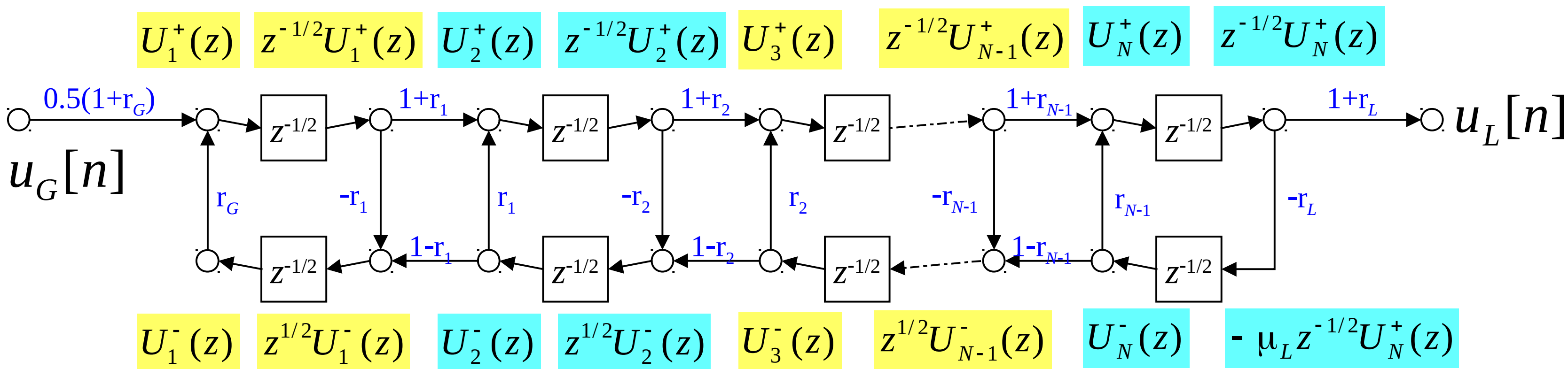
where 2τ is the round trip delay

earliest arrival is at time Nτ

and the following arrival occur at multiple of 2τ

Impulse Response

$$V(z) = \frac{U_L(z)}{U_G(z)} = ?$$



$$U_{k+1}^+(z) = (1+r_k)z^{-1/2}U_k^+(z) + r_k U_{k+1}^-(z)$$

$$U_k^+(z) = \frac{z^{1/2}}{1+r_k}U_{k+1}^+(z) - \frac{r_k z^{1/2}}{1+r_k}U_{k+1}^-(z)$$

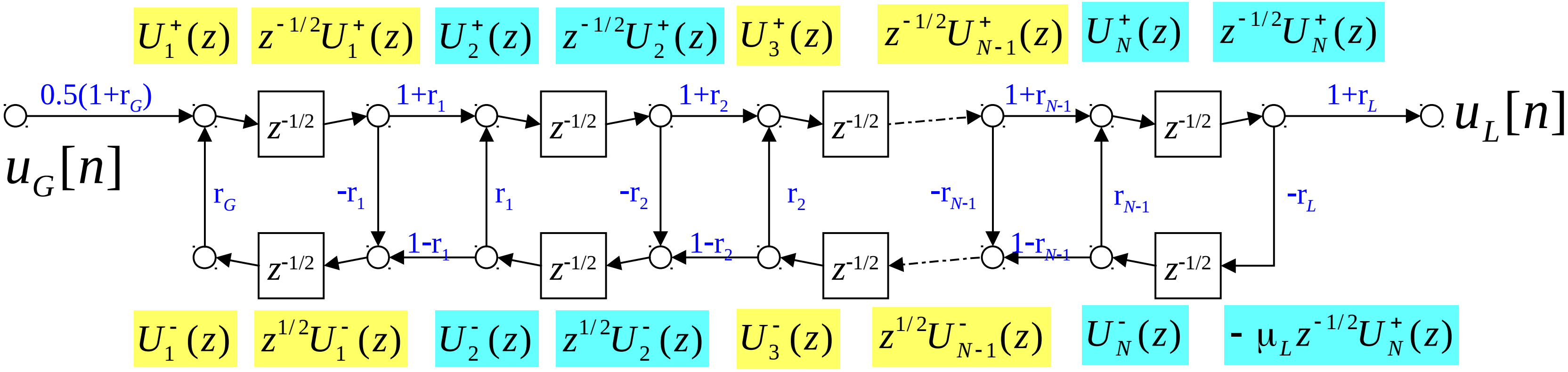
$$z^{1/2}U_k^-(z) = -r_k z^{-1/2}U_k^+(z) + (1-r_k)U_{k+1}^-(z)$$

$$U_k^-(z) = \frac{-r_k z^{-1/2}}{1+r_k}U_{k+1}^+(z) + \frac{z^{-1/2}}{1+r_k}U_{k+1}^-(z)$$

Impulse Response

$$\boxed{U_k^+(z) = \frac{z^{1/2}}{1+r_k} U_{k+1}^+(z) - \frac{r_k z^{1/2}}{1+r_k} U_{k+1}^-(z)}$$

$$\boxed{U_k^-(z) = \frac{-r_k z^{-1/2}}{1+r_k} U_{k+1}^+(z) + \frac{z^{-1/2}}{1+r_k} U_{k+1}^-(z)}$$



Define $\mathbf{U}_k = \begin{bmatrix} U_k^+(z) \\ U_k^-(z) \end{bmatrix}$ $\Rightarrow$ $\mathbf{U}_k = R_k \mathbf{U}_{k+1}$

$$\boxed{R_k = \frac{z^{1/2}}{1+r_k}\hat{R}_k}$$

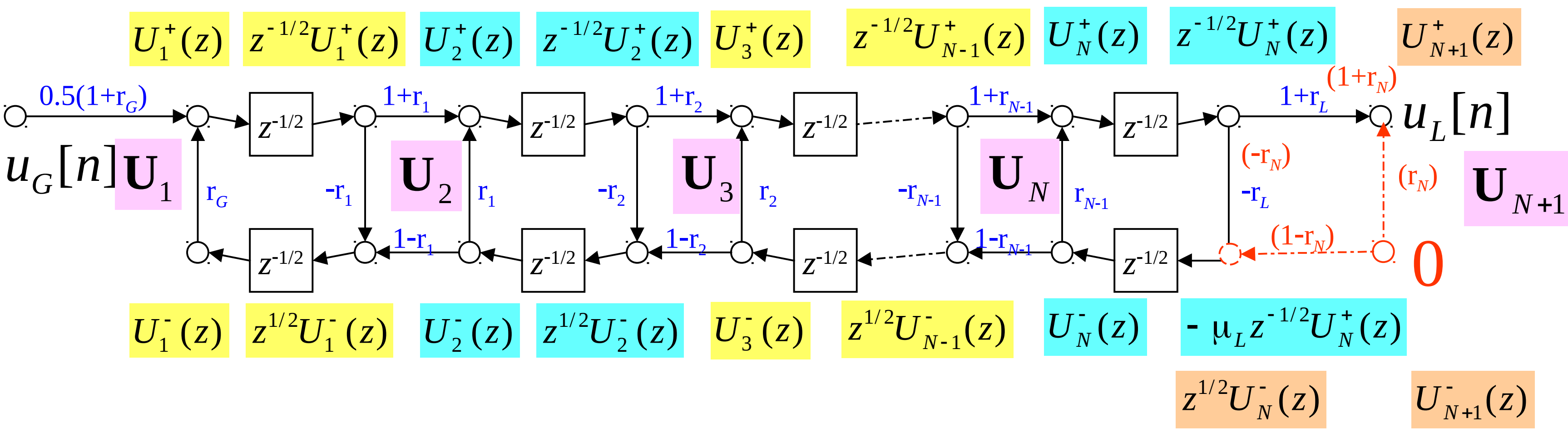
$$\mathbf{Q}_k = \begin{bmatrix} \dfrac{z^{1/2}}{1+r_k} & \dfrac{-r_k z^{1/2}}{1+r_k} \\ \dfrac{-r_k z^{-1/2}}{1+r_k} & \dfrac{z^{-1/2}}{1+r_k} \end{bmatrix} = \frac{z^{1/2}}{1+r_k}\begin{bmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{bmatrix} = \frac{z^{1/2}}{1+r_k}\hat{R}_k$$

$$\boxed{\hat{R}_k = \begin{bmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{bmatrix}}$$

Impulse Response

$$R_k = \frac{z^{1/2}}{1+r_k}\hat{R}_k$$

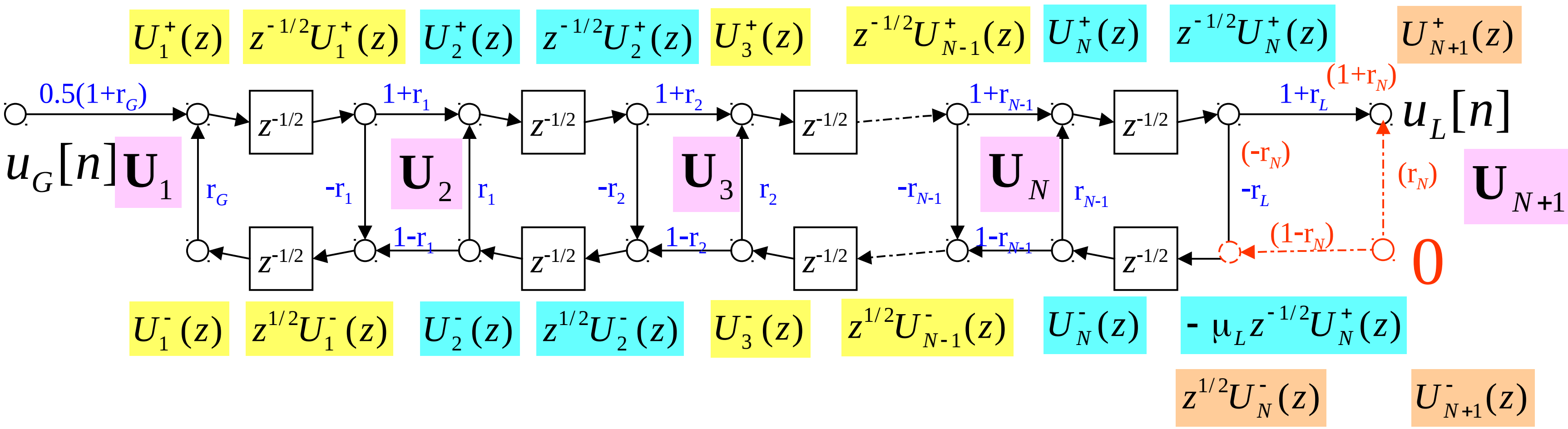
$$\hat{R}_k = \begin{bmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{bmatrix}$$



Define $\mathbf{U}_k = \begin{bmatrix} U_k^+(z) \\ U_k^-(z) \end{bmatrix} \Rightarrow \mathbf{U}_k = R_k \mathbf{U}_{k+1}$

Define $\mathbf{U}_{N+1} = \begin{bmatrix} U_L(z) \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} U_L(z) \Rightarrow \mathbf{U}_1 = R_1 R_2 \cdots R_N \mathbf{U}_{N+1}$

Impulse Response



$$U_1^+(z) = 0.5(1+r_G)U_G(z) + r_G U_1^-(z)$$

$$U_G(z) = \frac{2}{1+r_G}U_1^+(z) - \frac{2r_G}{1+r_G}U_1^-(z) \qquad = \frac{2}{1+r_G}\left[1, -r_G\right]\mathbf{U}_1$$

Impulse Response

$$R_k = \frac{z^{1/2}}{1+r_k}\hat{R}_k$$

$$\hat{R}_k = \begin{bmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{bmatrix}$$

$$U_G(z) = \left(\frac{2}{1+r_G}[1, -r_G]\left(R_1 R_2 \cdots R_N\right)\begin{bmatrix}1\\0\end{bmatrix}\right) U_L(z)$$

$$\frac{1}{V(z)} = \frac{U_G(z)}{U_L(z)} = \frac{2}{1+r_G}\left[1, -r_G\right]\left\{ z^{N/2}\prod_{k=1}^{N}\frac{1}{1+r_k}\right\}\left\{\prod_{k=1}^{N}\hat{R}_k\right\}\begin{bmatrix}1\\0\end{bmatrix}$$

Two-Tube Model

$$\boxed{R_k = \frac{z^{1/2}}{1+r_k}\hat{R}_k}$$

$$\boxed{\hat{R}_k = \begin{bmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{bmatrix}}$$

$$U_G(z) = \left\{\frac{2}{1+r_G}[1,-r_G]\right\} R_1 R_2 \cdots R_N \begin{bmatrix} 1 \\ 0 \end{bmatrix} U_L(z)$$

$$\boxed{\frac{1}{V(z)} = \frac{U_G(z)}{U_L(z)} = \frac{2}{1+r_G}\left[1,-r_G\right]\left\{z^{N/2}\prod_{k=1}^{N}\frac{1}{1+r_k}\right\}\left\{\prod_{k=1}^{N}\hat{R}_k\right\}\begin{bmatrix} 1 \\ 0 \end{bmatrix}}$$

$$\frac{1}{V(z)} = z\frac{2}{(1+r_G)(1+r_1)(1+r_2)}\left[1,-r_G\right]\begin{bmatrix} 1 & -r_1 \\ -r_1 z^{-1} & z^{-1} \end{bmatrix}\begin{bmatrix} 1 & -r_2 \\ -r_2 z^{-1} & z^{-1} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= z\frac{2}{(1+r_G)(1+r_1)(1+r_2)}\left[1,-r_G\right]\begin{bmatrix} 1 & -r_1 \\ -r_1 z^{-1} & z^{-1} \end{bmatrix}\begin{bmatrix} 1 \\ -r_2 z^{-1} \end{bmatrix}$$

$$= z\frac{2}{(1+r_G)(1+r_1)(1+r_2)}\left[1,-r_G\right]\begin{bmatrix} 1 + r_1 r_2 z^{-1} \\ -r_1 z^{-1} - r_2 z^{-2} \end{bmatrix} = z\frac{2(1+r_1 r_2 z^{-1} + r_1 r_G z^{-1} + r_2 r_G z^{-2})}{(1+r_G)(1+r_1)(1+r_2)}$$

Two-Tube Model

$$V(z) = \frac{0.5(1+r_g)(1+r_1)(1+r_2)z^{-1}}{1+(r_1 r_2 + r_1 r_G)z^{-1} + r_2 r_G z^{-2}}$$

o at origin

er (2 poles)

N-Tube Model

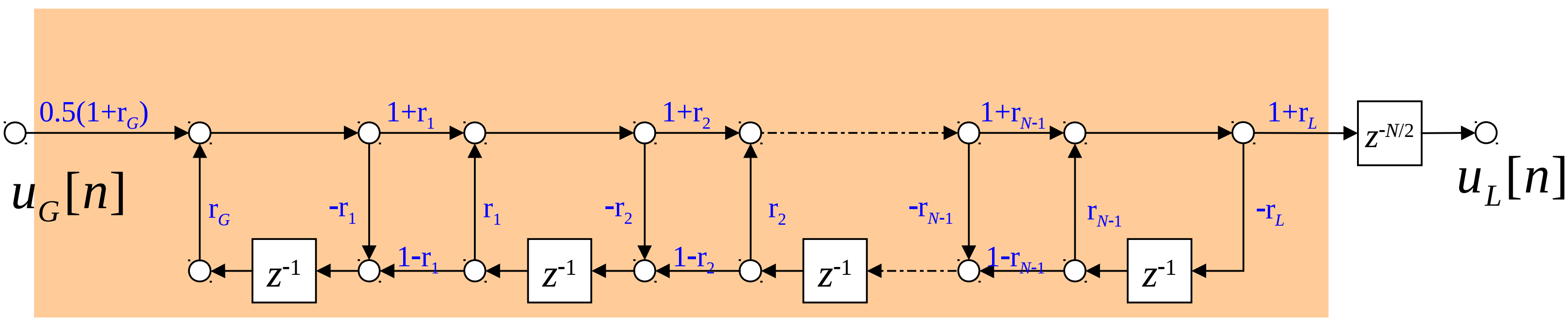
$$V(z) = \frac{0.5(1+r_g)(1+r_1)(1+r_2)z^{-1}}{1+(r_1r_2+r_1r_G)z^{-1}+r_2r_Gz^{-2}}$$

o at origin

er (2 poles)

$$\hat{V}(z) = z^{N/2}V(z) = \frac{G}{1+\sum_{k=1}^{N}\alpha_k z^{-k}} = \frac{G}{D(z)}$$

All-pole Model



Transfer Function of Lossless Tube Model

$$D(z) = 1 - \sum_{k=1}^{N} \alpha_k z^{-k}$$

- special case of $r_G = 1$ ($Z_G = \infty$)

$$D_0(z) = 1$$

$$D_k(z) = D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}), \qquad k = 1, 2, \ldots, N$$

$$D(z) = D_N(z)$$

- Examples:

$$D_1(z) = 1 + r_1 z^{-1} = D_0(z) + r_1 z^{-1} D_0(z^{-1}) = 1 + r_1 z^{-1}(1)$$

$$D_2(z) = 1 + r_1 z^{-1} + r_1 r_2 z^{-1} + r_2 z^{-2} = D_1(z) + r_2 z^{-2} D_1(z^{-1})$$

$$= 1 + r_1 z^{-1} + r_2 z^{-2}(1 + r_1 z) = 1 + r_1 z^{-1} + r_1 r_2 z^{-1} + r_2 z^{-2}$$

- choose $N = 10$ as a reasonable number of tubes for model

$r_N = 1 \Rightarrow A_{N+1} = \infty$ (infinite tube at lips)

$r_N = 0.714 \Rightarrow A_{N+1} = 28$ cm^2

Transfer Function of Lossless Tube Model

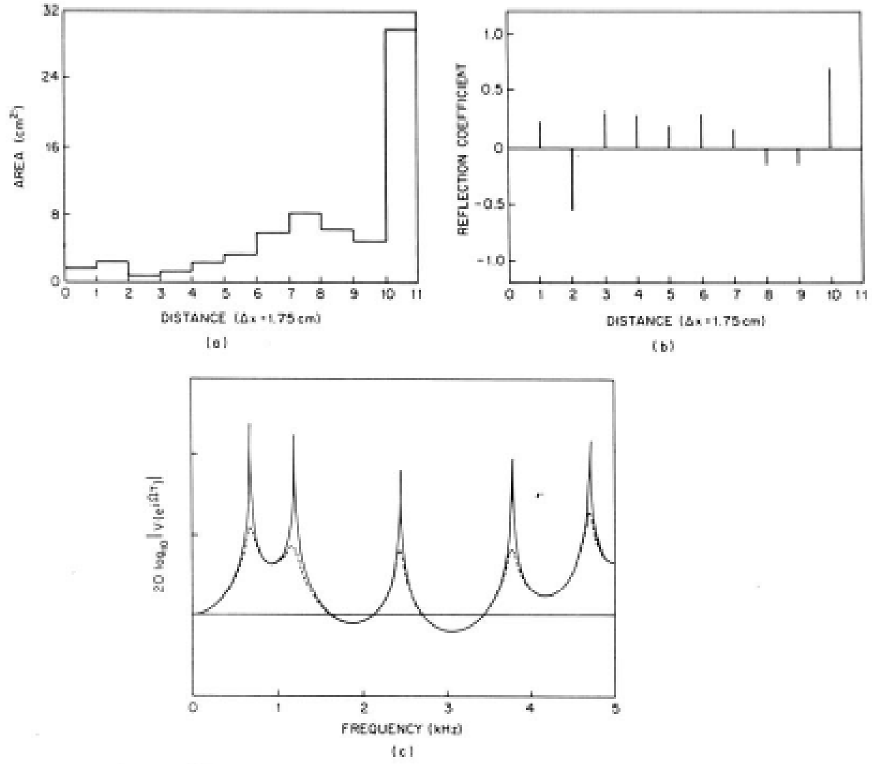


Fig. 3.43 (a) Area function for 10 section lossless tube terminated with reflectionless section of area 30 cm^2; (b) reflection coefficients for 10 section tube; (c) frequency response of 10 section tube; dotted curve corresponds to conditions of (b); solid curve corresponds to short-circuit termination. (Note area data of (a) estimated from data given by Fant [1] for the Russian vowel /a/.)

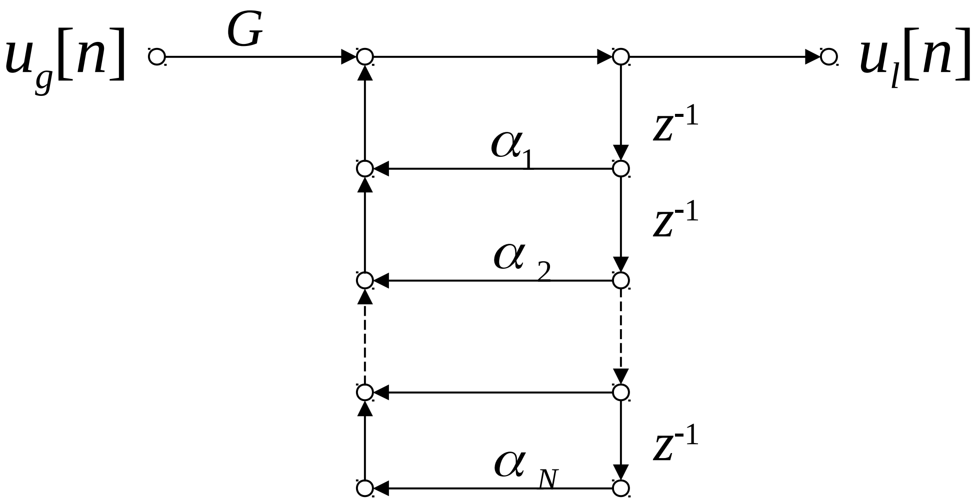
Let length of the vocal tract *l=17cm* and the velocity of sound *c=340m/s* find the number of section required to generate *5 kHz* bandwidth voiced signal

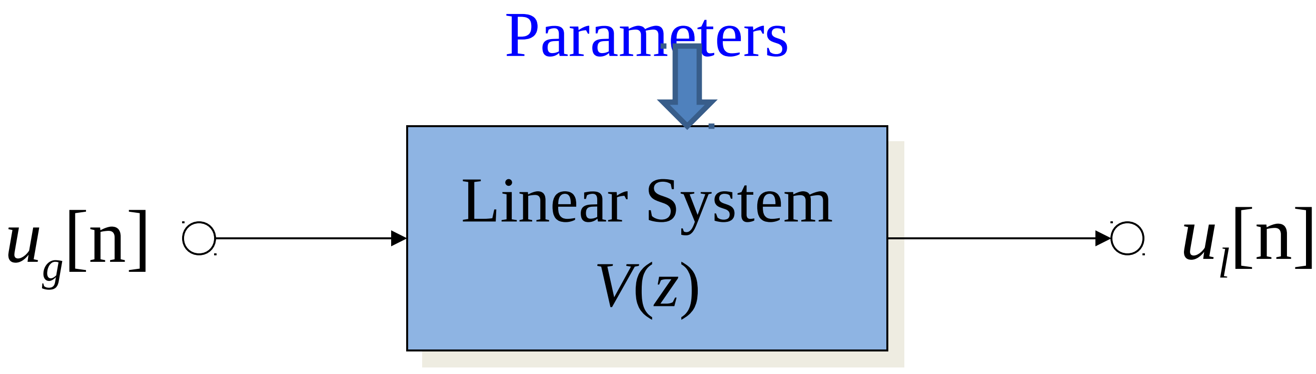
Lecture-9

Digital Models for Speech Signals

Direct Implementation

$$V(z) = \frac{G}{D(z)} \qquad D(z) = 1 + \sum_{k=1}^{N} \alpha_k z^{-k}$$



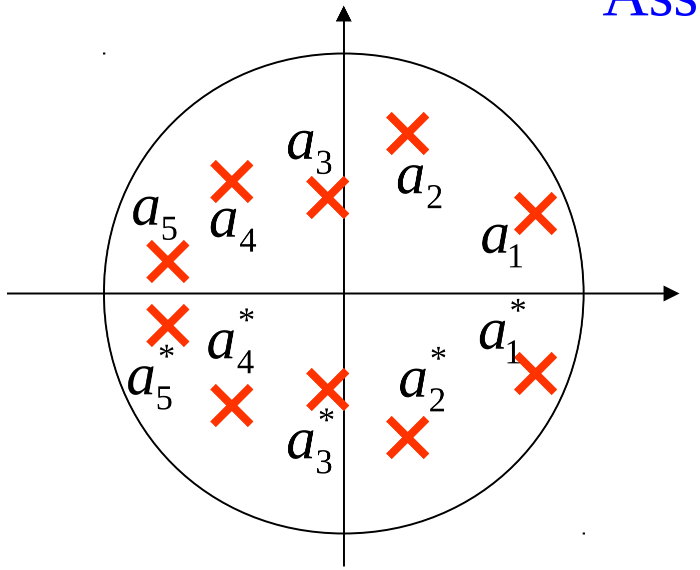


Poles of Vocal Tract Model

$$V(z) = \frac{G}{D(z)} \quad D(z) = 1 + \sum_{k=1}^{N} \alpha_k z^{-k}$$

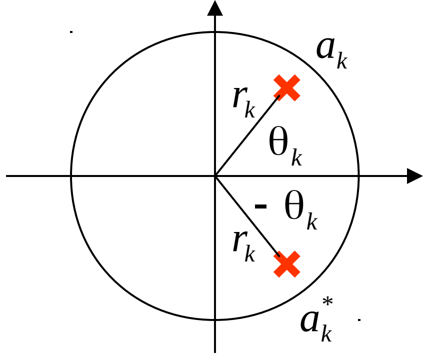
- The roots of *D(z)* will be either *real* of occur in *complex conjugate pairs*.

Assume all poles are in conjugated pairs.



$$V(z) = \frac{G}{\prod_{k=1}^{N/2} (1 - \alpha_k z^{-1})(1 - \alpha_k^* z^{-1})}$$

$$\alpha_k = r_k e^{j\theta_k} \quad \alpha_k^* = r_k e^{-j\theta_k}$$



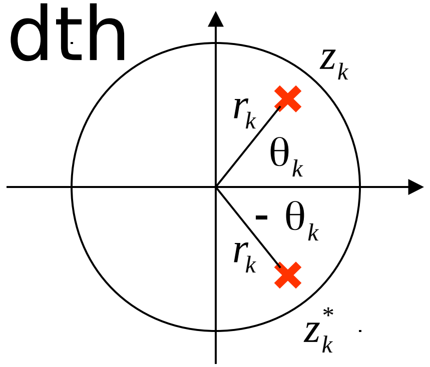
$$\text{Let } V_k(z) = \frac{1}{(1-\alpha_k z^{-1})(1-\alpha_k^* z^{-1})} = \frac{1}{1 - r_k e^{j\theta_k} z^{-1} - r_k e^{-j\theta_k} z^{-1} + r_k^2 z^{-2}}$$

$$= \frac{1}{1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2}}$$

$$\text{Let } r_k = e^{-b_k} \Rightarrow b_k = -\ln r_k$$

$$V_k(z) = \frac{1}{1 - 2e^{-b_k} \cos\theta_k z^{-1} + e^{-2b_k} z^{-2}}$$

Formant Frequency and Bandwidth

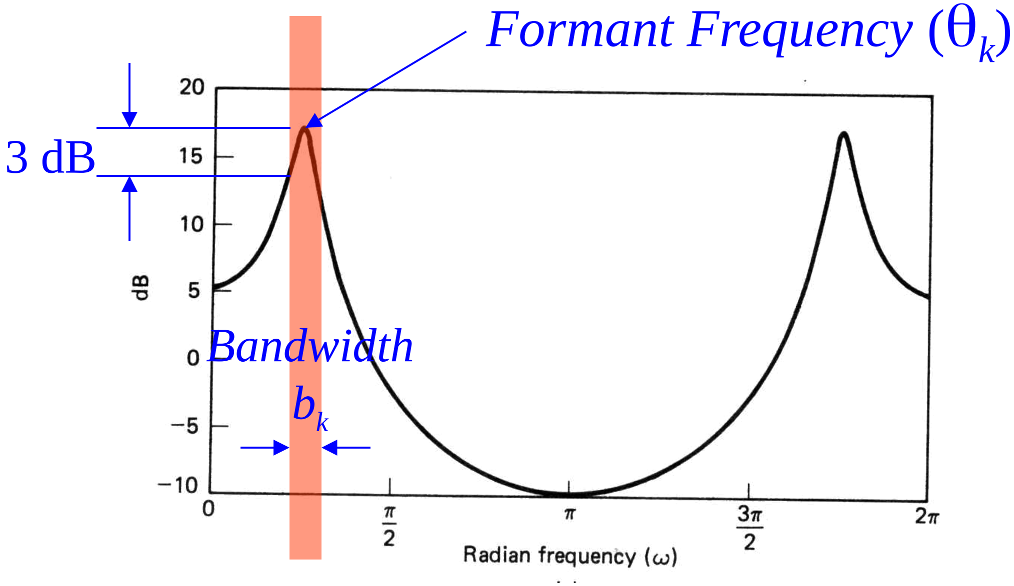


$$V(z) = \frac{G}{\prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

$r_k = 0.9$

$\theta_k = \pi / 4$

$b_k = 0.1053$

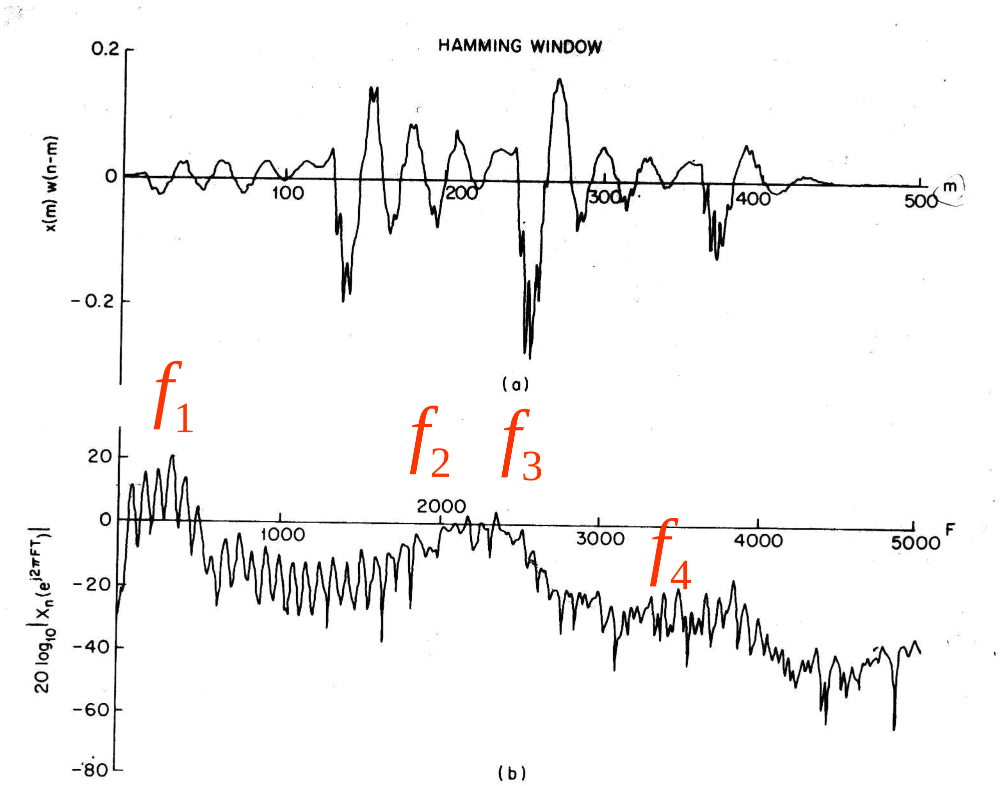


Let $r_k = e^{-b_k}$ $\Rightarrow$ $b_k = -\ln r_k$

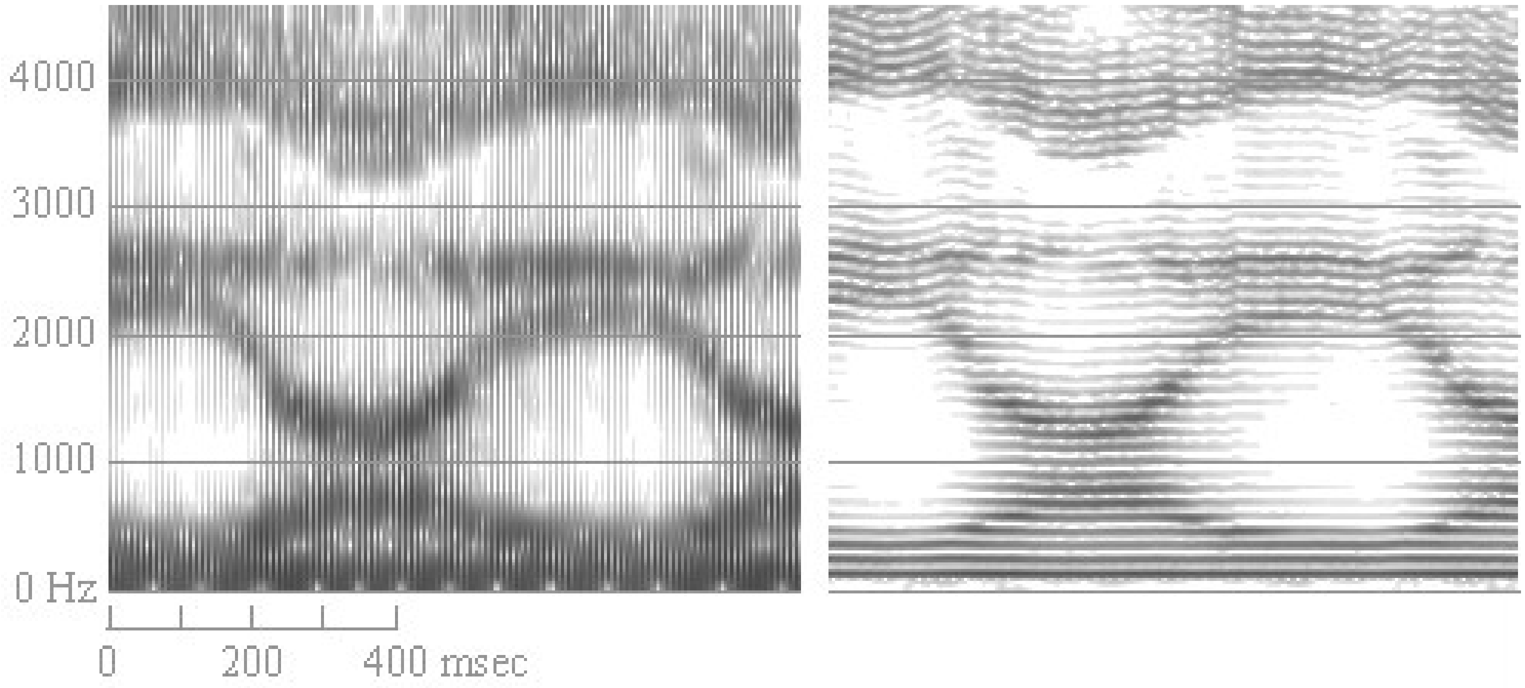
$$V_k(z) = \frac{1}{1 - 2e^{-b_k} \cos\theta_k z^{-1} + e^{-2b_k} z^{-2}}$$

Formant Frequency and Bandwidth

$$V(z) = \frac{G}{\prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

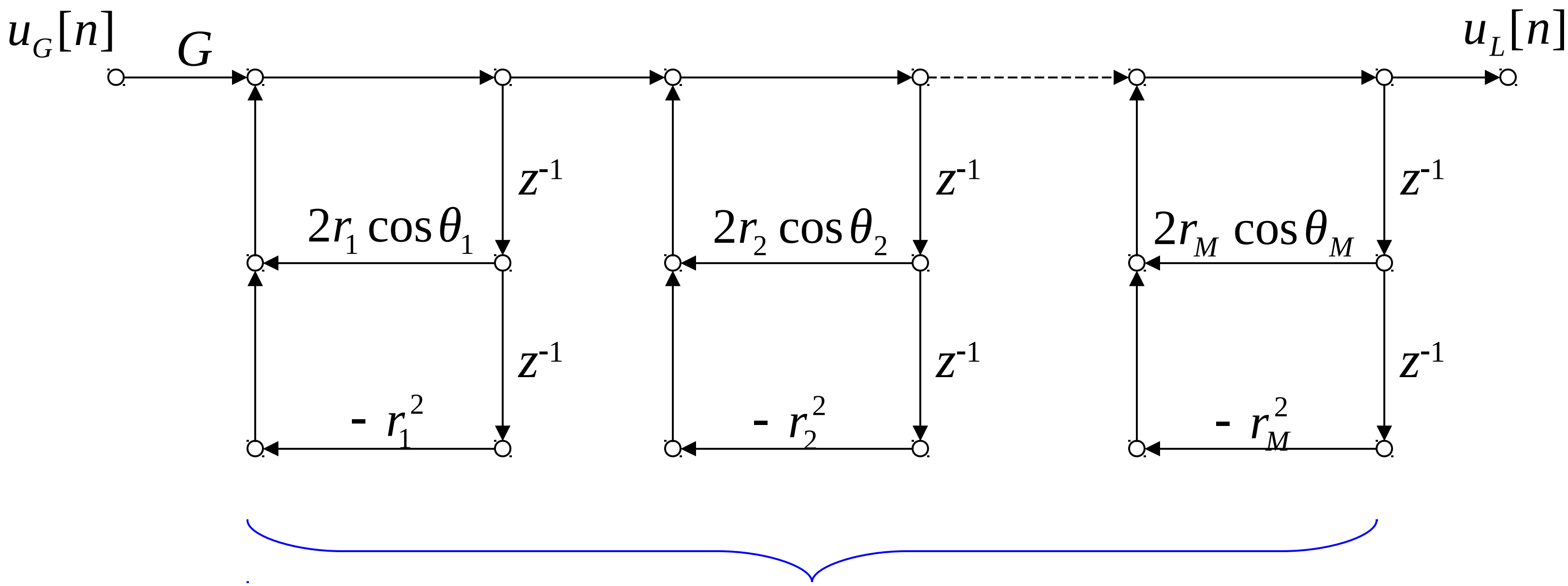


Formant Trajectory



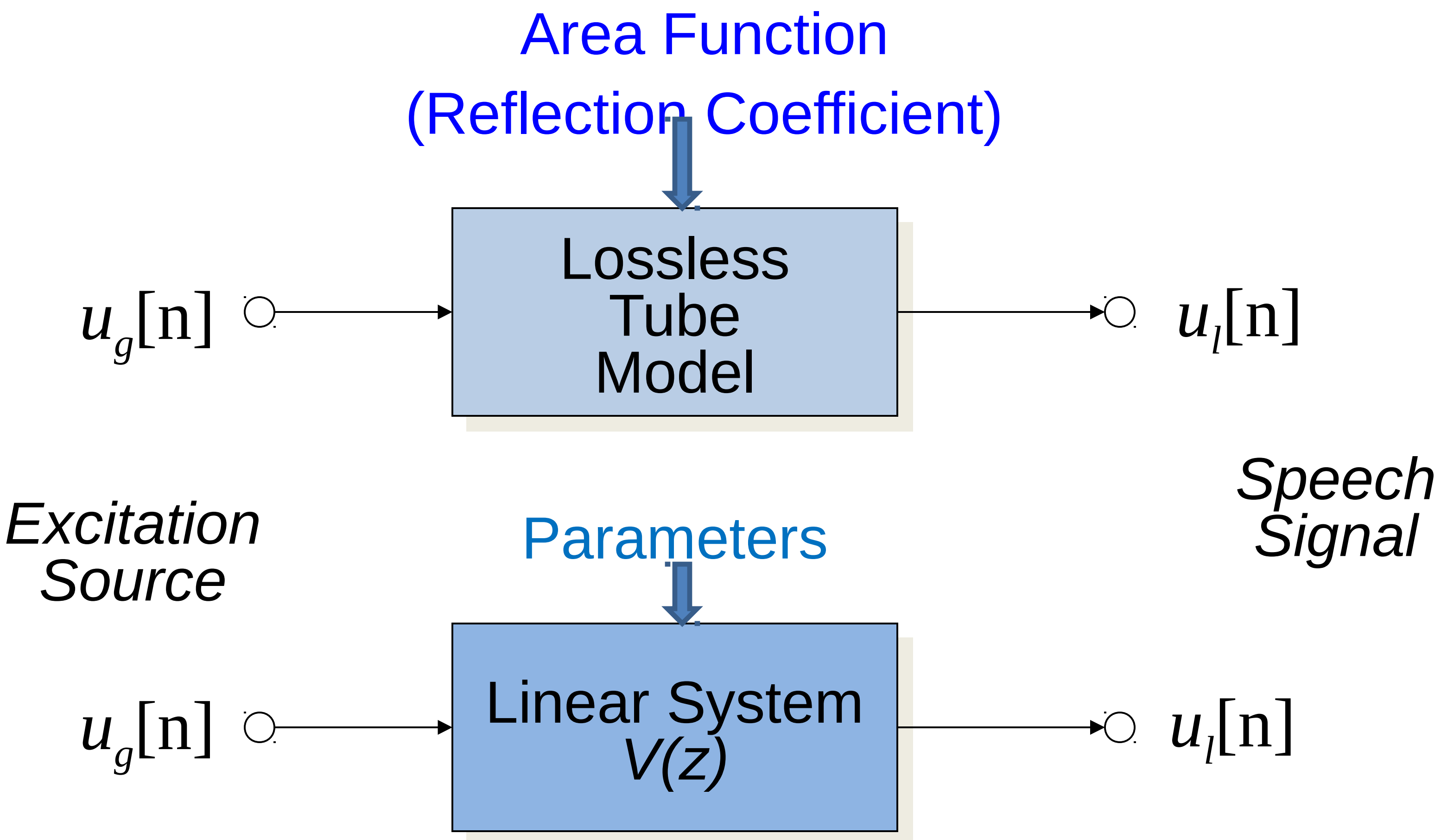
Cascade Implementation

$$V(z) = \frac{G}{\prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

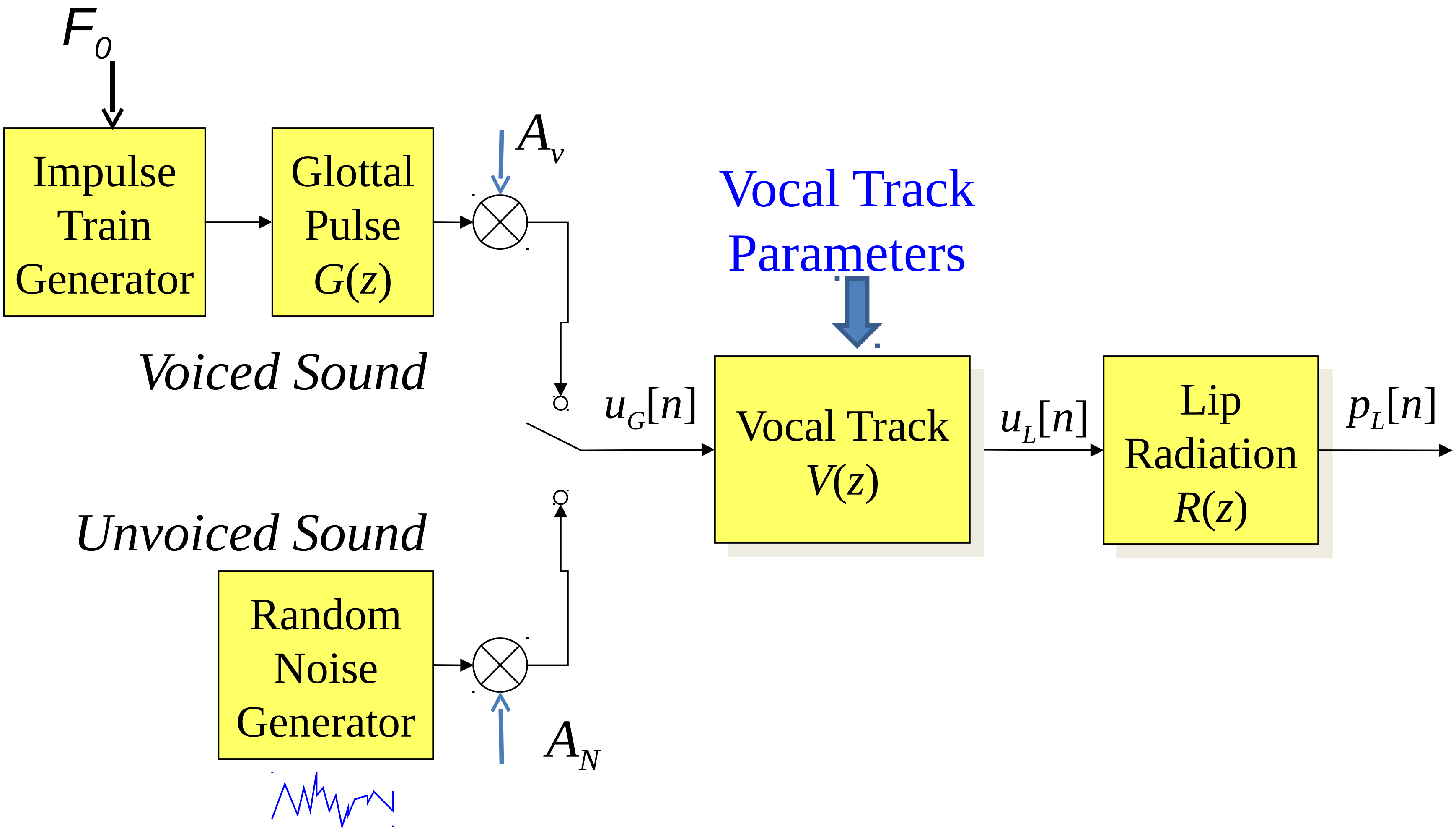


Each Stage represents one formant frequency and its corresponding bandwidth.

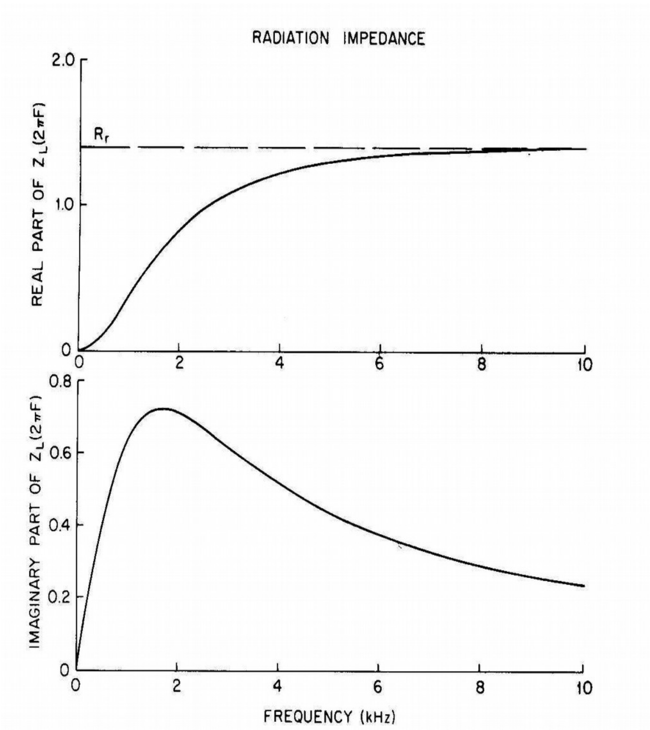
Digital Models for Speech Signals



Speech Production Model



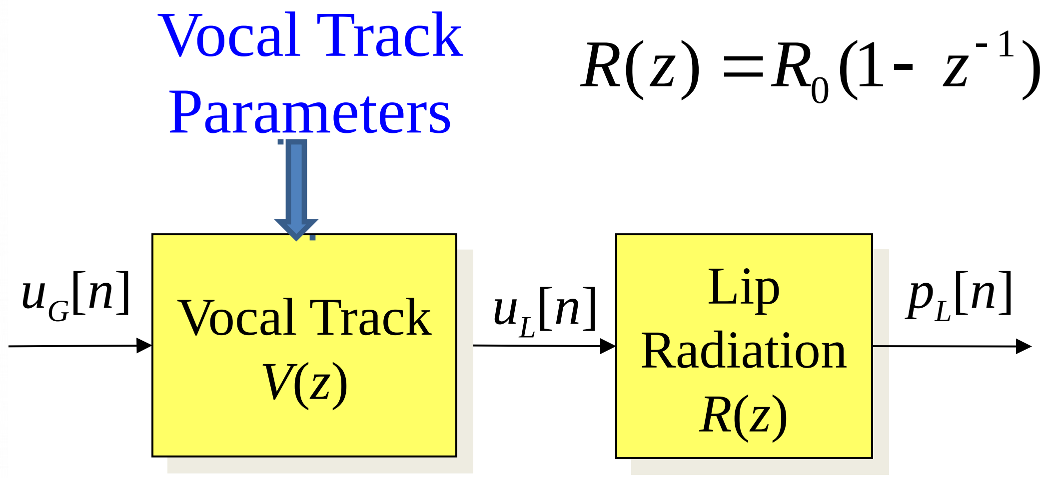
Radiation



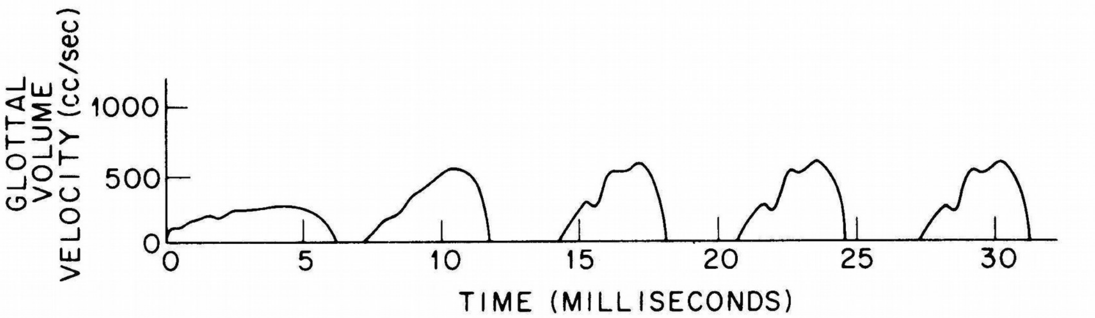
A High-Pass Filter

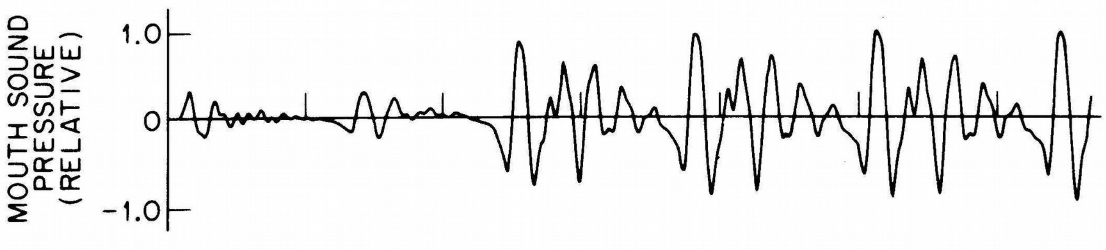
$$P_L(z) = R(z)U_L(z)$$

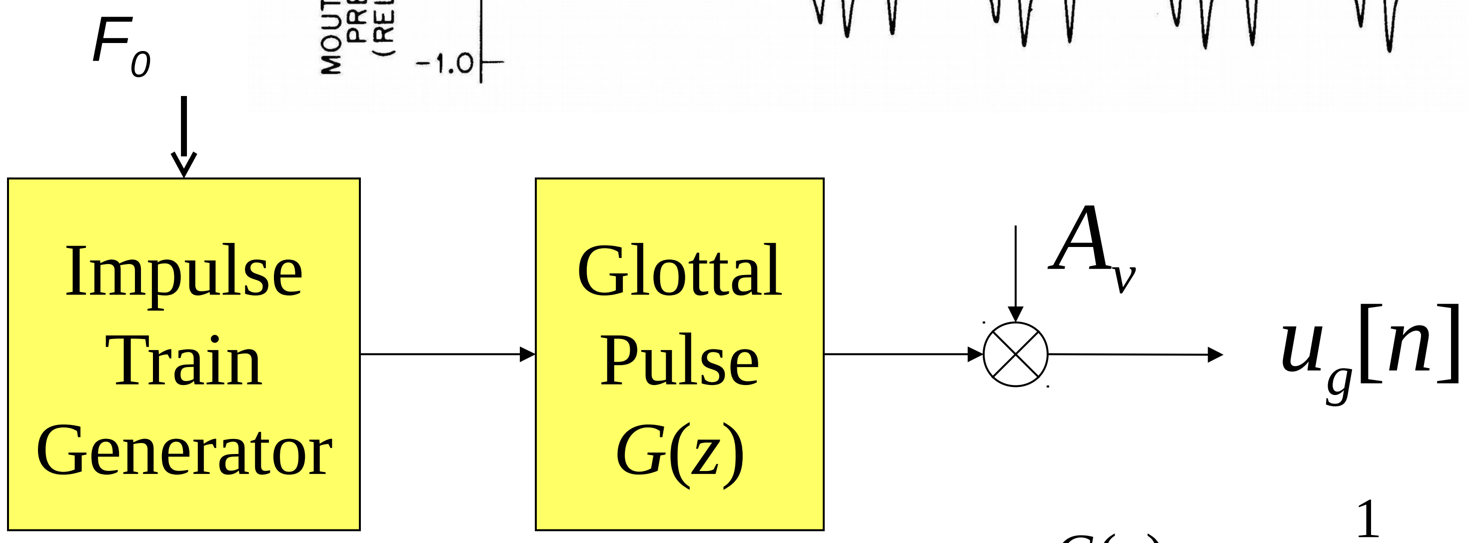
$$R(z) = R_0(1 - z^{-1})$$



Excitation





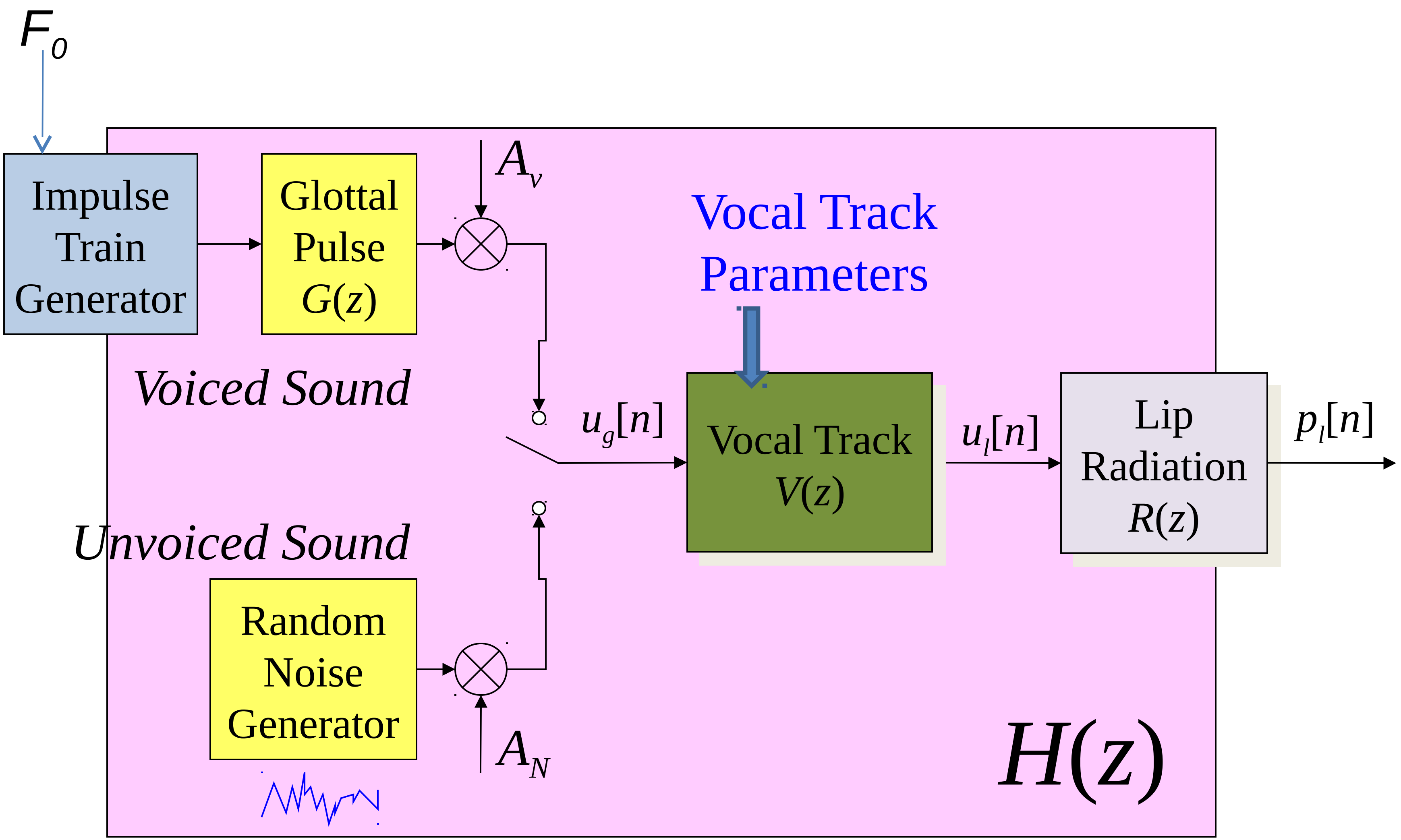


$$G(z) = \frac{1}{(1 - e^{-cT} z^{-1})^2}$$

$$e^{-cT} \approx 1$$

Speech Production Model

$$H(z) = G(z)V(z)R(z)$$



Speech Production Model

$$H(z) = G(z)V(z)R(z)$$

$$G(z) = \frac{1}{(1 - e^{-cT} z^{-1})^2}$$

$$V(z) = \frac{G}{\prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

$$R(z) = R_0 (1 - z^{-1})$$

$$H(z) = \frac{\sigma(1 - z^{-1})}{(1 - e^{-cT} z^{-1})^2 \prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

$$H(z) = \frac{\sigma}{(1 - e^{-cT} z^{-1})^1 \prod_{k=1}^{N/2} (1 - 2r_k \cos\theta_k z^{-1} + r_k^2 z^{-2})}$$

$$H(z) = \frac{\sigma}{1 + \sum_{k=1}^{p} a_k z^{-k}}$$