Digital Speech Processing— Lectures 5-6

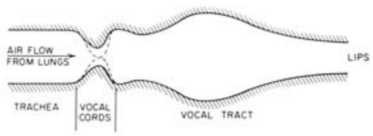
Sound Propagation in the Vocal Tract

Basics

- can use 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)

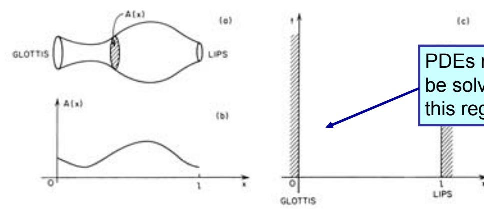
Bottom Line: simplify as much as possible and see what we can learn about the mechanics of sound propagation in the human vocal tract

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

Schematic Vocal Tract



PDEs must be solved in this region

Fig. 3.13 (a) Schematic vocal tract; (b) corresponding area function; (c) $x-t$ plane for solution of wave equation.

- simplified vocal tract area => **non-uniform tube** with time varying cross section
- ***plane wave propagation*** along the axis of the tube (this assumption valid for frequencies below about 4000 Hz)
- ***no losses at walls***

Sound Wave Propagation

- using the laws of conservation of mass, momentum and energy, it can be shown that sound wave propagation in a lossless tube satisfies the equations:

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

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

- where

$p = p(x,t) =$ sound pressure in the tube at position x and time t

$u = u(x,t) =$ volume velocity flow at position x and time t

$\rho =$ the density of air in the tube

$c =$ the velocity of sound

$A = A(x,t) =$ the 'area function' of the tube,

i.e., the cross-sectional area normal to the axis of the tube, as a function of the distance along the tube and as a function of time

Solutions to Wave Equation

- ***no closed form solutions*** exist for the propagation equations
  - need ***boundary conditions***, namely *u(0,t)* (the volume velocity flow at the glottis), and *p(l,t)*, (the sound pressure at the lips) to solve the equations numerically (by a process of iteration)
  - need ***complete specification of A(x,t),*** the vocal tract area function; for simplification purposes we will assume that there is no time variability in *A(x,t)* => the term related to the partial time derivative of *A* becomes 0
  - even with these simplifying assumptions, numerical solutions are very hard to compute

Consider simple cases and extrapolate results to more complicated cases

Uniform Lossless Tube

- Assume uniform lossless tube => *A(x,t)=A* (shape consistent with /UH/ vowel)

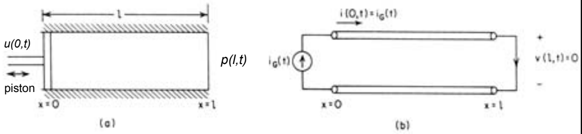


Fig. 3.14 (a) Uniform lossless tube with ideal terminations; (b) corresponding electrical transmission line analogy.

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A}\frac{\partial u}{\partial t} \qquad -\frac{\partial v}{\partial x} = L\frac{\partial i}{\partial t}$$

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

7

Acoustic-Electrical Analogs

Acoustic	Electrical
$p =$ pressure	$v =$ voltage
$u =$ volume velocity	$i =$ current
$\rho / A =$ acoustic inductance	$L =$ inductance
$A/(\rho c^2) =$ acoustic capacitance	$C =$ capacitance
uniform acoustic tube ⟺	lossless transmission line terminated in a short circuit, $v(l,t)=0$, at one end, excited by a current source $i(0,t)=i_G(t)$ at the other end

8

Traveling Wave Solution

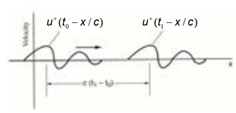
- assume traveling wave solution

$$u(x,t) = \left[u^+(t-x/c) - u^-(t+x/c)\right]$$

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

-- $u^+(t-x/c)$ wave travelling forward

-- $u^-(t+x/c)$ wave travelling backward



- boundary conditions at the glottis and at the lips gives:

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

$$p(\ell,t) = 0$$

- since the differential equations are linear with constant coefficients, the solutions must be of the form

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

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

9

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 $p(x,t)$

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

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

10

Traveling Wave Solution

- look at solution for $u(\ell,t)$

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

- giving for the transfer function of volume velocity

$$V_a(\Omega) = \frac{U(\ell,\Omega)}{U_G(\Omega)} = \frac{1}{\cos(\Omega\ell/c)}$$

11

Overall Transfer Function

- consider the volume velocity at the lips (*x=l*) as a function of the source (at the glottis)

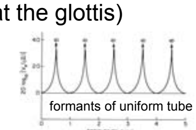
$$u(\ell,t) = U(\ell,\Omega)e^{j\Omega t}$$

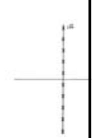
$$= \frac{1}{\cos(\Omega\ell/c)}U_G(\Omega)e^{j\Omega t}$$

$$\frac{U(\ell,\Omega)}{U_G(\Omega)} = V_a(\Omega) = \frac{1}{\cos(\Omega\ell/c)}$$

↑

Frequency response of uniform tube in terms of volume velocities





$\Omega = 2\pi f$; $c = 35{,}000$ cm/sec; ℓ=17.5 cm

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

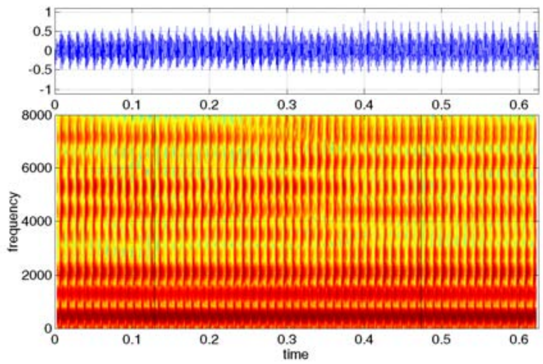
$\cos\left(\frac{2\pi f\ell}{c}\right) = 0$ when $\frac{2\pi f\ell}{c} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots$

i.e., when $f_n = \frac{c}{4\ell}\cdot(2n+1)$, $n = 0,1,2,\ldots$

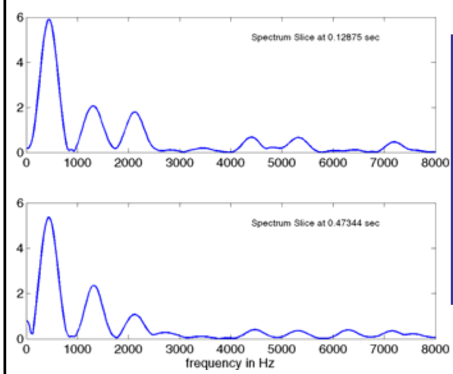
$f_0 = 500$Hz; $f_1 = 1500$ Hz; $f_2 = 2500$ Hz,...

12

Hardwalled Tube and Buzzer



Spectrum Slices



The formants are not at the frequencies 500, 1500, 2500, … Hz. What are some possible sources of error in this experiment?

14

Summary of Solution of Sound Propagation Equations in the Vocal Tract

Step 1--Basic sound wave propagation equations

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

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

Step 2--Boundary conditions

$u(0,t) = u_G(t) \Rightarrow$ sound source at glottis

$p(\ell,t) = 0 \Rightarrow$ no pressure at lips

Step 3--Simplifying assumption $\Rightarrow$ fixed area A

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

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

Step 4--Assume travelling waves form of solution

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

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

Step 5--Simplified boundary conditions

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

$$p(\ell,t) = 0$$

Step 6--Simplified forward and backward waves

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

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

Step 7--Determine K^+ and K^-, and solve for $u(x,t)$, $p(x,t)$

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

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

Step 8--Solve for $u(\ell,t)$

$$u(\ell,t) = U(\ell,\Omega)e^{j\Omega t}$$

Step 9--Solve for transfer function of volume velocity

$$V_a(\Omega) = \frac{U(\ell,\Omega)}{U_G(\Omega)} = \frac{1}{\cos(\Omega\ell/c)}$$

Step 10--Determine tube resonances

$\ell = 17.5$ cm; $c = 35{,}000$ cm/sec $\Rightarrow$

$$f_n = \frac{(2n+1)c}{4\ell} = 500, 1500, 2500, \ldots \text{Hz}$$

Step 11--Interpretation of traveling wave solution

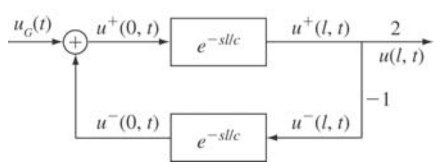
$$V_a(s) = \frac{U(\ell,s)}{U_G(s)} = \frac{2e^{-s\ell/c}}{1 + e^{-2s\ell/c}} = \frac{1}{\cosh(s\ell/c)}$$

15

Traveling Wave Solution

- using s-transform notation ($s = \sigma - j\Omega$) we get

$$V_a(s) = \frac{U(\ell,s)}{U_G(s)} = \frac{2e^{-s\ell/c}}{1 + e^{-2s\ell/c}} = \frac{1}{\cosh(s\ell/c)}$$



16

Frequency Domain Representation

- we can alternatively express $p(x,t)$ and $u(x,t)$ as

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

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

- where

$$Z_0 = \frac{\rho c}{A} = \text{characteristic acoustic impedance of tube}$$

17

Alternative Wave Equation Solution (avoids solution for forward and backward travelling waves)

- express $p(x,t)$ and $u(x,t)$ as complex transfer functions of the form:

$$p(x,t) = P(x,\Omega)e^{j\Omega t}$$

$$u(x,t) = U(x,\Omega)e^{j\Omega t}$$

- inserting these representations back into the wave equation gives:

$-\frac{dP}{dx} = ZU, \quad Z = j\Omega\rho/A =$ acoustic impedance per unit length

$-\frac{dU}{dx} = YP, \quad Y = j\Omega A/(\rho c^2) =$ acoustic admittance per unit length

Solution 1

- can show that solutions of wave equation have form:

$$P(x,\Omega) = Ae^{\gamma x} + Be^{-\gamma x}$$

$$U(x,\Omega) = Ce^{\gamma x} + De^{-\gamma x}$$

$$\gamma = \sqrt{ZY} = j\Omega/c$$

- by using boundary conditions, $P(\ell,\Omega) = 0$, $U(0,\Omega) = U_g(\Omega)$, can solve for A,B,C,D

18

Effects of Losses in VT

- several types of losses to be considered
  - viscous friction at the walls of the tube
  - heat conduction through the walls of the tube
  - vibration of the tube walls
- loss will change the frequency response of the tube
- consider first 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

$$A(x,t) = A_0(x,t) + \delta A(x,t)$$

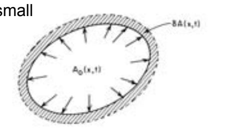


Fig. 3.16 Illustration of the effects of wall vibration.

Effects of Loss

- there is a differential equation relationship between area perturbation $\delta A(x,t)$ and the pressure variation, $p(x,t)$ of the form:

$$m_W \frac{d^2(\delta A)}{dt^2} + b_W \frac{d(\delta A)}{dt} + k_W(\delta A) = p(x,t) \quad \text{where}$$

$m_W(x) =$ 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

- neglecting second order terms in u/A and pA, the basic wave equations become

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

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

20

Losses in Frequency Domain

- consider a time-invariant constant area tube excited by a complex volume velocity source

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

- since the loss differential equation is linear and time-invariant, the form for *p(x,t)* and *u(x,t)* is:

$$p(x,t) = P(x,\Omega)e^{j\Omega t}$$

$$u(x,t) = U(x,\Omega)e^{j\Omega t} \qquad \delta A(x,t) = \delta A(x,\Omega)e^{j\Omega t}$$

- substituting into the wave equations yields the following:

$$-\frac{\partial P}{\partial x} = Z(x,\Omega)U, \qquad Z(x,\Omega) = j\Omega\frac{\rho}{A_0(x)}$$

$$-\frac{\partial U}{\partial x} = Y(x,\Omega)P + Y_W(x,\Omega)P, \qquad Y(x,\Omega) = j\Omega\frac{A_0(x)}{\rho c^2}$$

$$Y_W(x,\Omega) = \frac{1}{j\Omega m_W(x) + b_W(x) + \frac{k_W(x)}{j\Omega}}$$

Solution 2—same as Solution 1 with new term, Y_W, and with *Z* and *Y* terms being functions of *x*

21

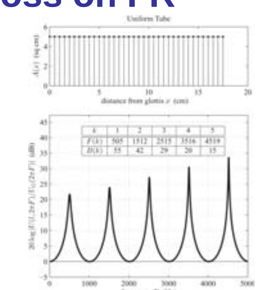
Effects of Loss on FR

- 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:

$$V_a(j\Omega) = \frac{U(\ell,\Omega)}{U_G(\Omega)}$$

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

- increases bandwidth of complex poles
- decreases resonance frequency (slightly)



- complex poles with non-zero bandwidths
- slightly higher frequencies for resonances
- most effect at lower frequencies

Friction and Thermal Conduction Losses

- Viscous friction can be accounted for in the frequency domain by including a real, frequency dependent term in the expression for the acoustic impedance, *Z,* of the form:

$$Z(x,\Omega) = \frac{S(x)}{[A_0(x)]^2}\sqrt{\lambda\rho\mu/2} + j\Omega\frac{\rho}{A_0(x)}$$

$S(x)$ is the circumference of the tube in cm

μ is the coefficient of friction (0.000186)

ρ is the density of air in the tube (0.00114 gm/cm^3)

- Heat conduction accounted for by adding a real frequency dependent term to the acoustic admittance, of the form:

$$Y(x,\Omega) = \frac{S(x)(\eta-1)}{\rho c^2}\sqrt{\frac{\lambda\Omega}{2c_p\rho}} + j\Omega\frac{A_0(x)}{\rho c^2}$$

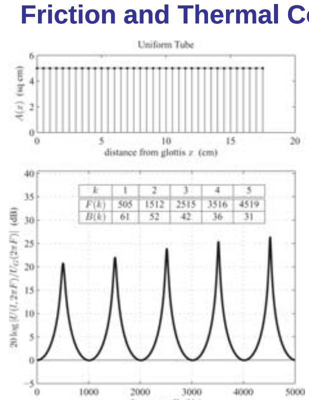
c_p is the specific heat at constant pressure (0.24)

η is the ratio of specifiec heat at constant pressure to that at constant volume (1.4)

λ is the coefficient of heat conduction (0.000055)

23

Friction and Thermal Conduction Losses



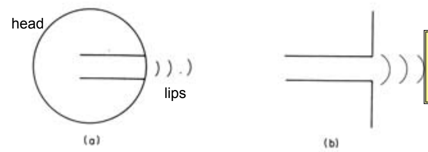
Main effect of friction and thermal conduction losses is that the formant bandwidths increase

- since friction and thermal losses increase with $\Omega^{1/2}$, the higher frequency resonances experience a greater broadening than the lower resonances
- the effects of friction and thermal loss are small compared to the effects of wall vibration for frequencies below 3-4 kHz

24

Effects of Radiation at Lips

- we have 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



model valid when lip opening is small compared to size of sphere

Fig. 3.19 (a) Radiation from a spherical baffle; (b) radiation from an infinite plane baffle.

Radiation at Lips

- using the infinite plane baffle model for radiation at the lips, can replace the boundary condition for a complex sinusoid input with the following:

$P(\ell,\Omega) = Z_L(\Omega)U(\ell,\Omega)$ where

$Z_L(\Omega) = \dfrac{j\Omega L_r R_r}{R_r + j\Omega L_r}$ -- 'radiation impedance' or 'radiation load' at lips

- this 'radiation load' is the equivalent of a parallel connection of a radiation resistance, R_r, and a radiation inductance, L_r. Suitable values for these components are:

$R_r = \dfrac{128}{9\pi^2}$, $L_r = \dfrac{8a}{3\pi c}$, where *a* is the radius of the opening and *c* is the velocity of sound

Behavior of Radiation Load

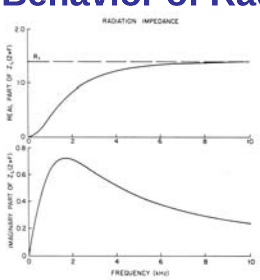


Fig. 3.20 Real and imaginary parts of the radiation impedance.

radiation losses most significant at higher frequencies

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

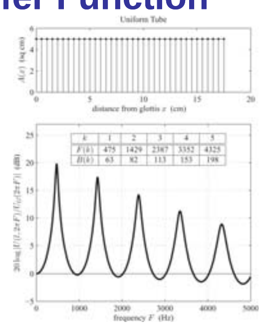
- at low frequencies, $Z_L(\Omega) \approx 0$ (short circuit termination) $\Rightarrow$ old solution
- at mid-range frequencies, $Z_L(\Omega) \approx j\Omega L_r$ (inductive load) $\Rightarrow R_r \gg \Omega L_r$
- at higher frequencies, $Z_L(\Omega) \approx R_r$ (resistive load) $\Rightarrow \Omega L_r \gg R_r$

Overall Transfer Function

- for the case of a uniform, time-invariant tube with yielding walls, friction and thermal losses, and radiation loss of an infinite plane baffle, can solve the wave equations for the transfer function:

$$V_a(j\Omega) = \frac{U(\ell,\Omega)}{U_G(\Omega)}$$

- assuming input at glottis of form: $U(0,t) = U_G(\Omega)e^{j\Omega t}$

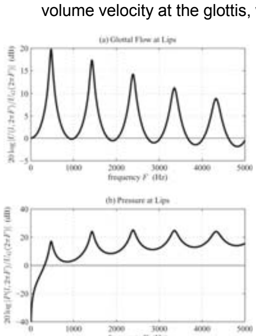


- higher bandwidths, lower resonance frequencies
  - first resonance is primarily determined by wall loss
  - higher resonance bandwidths are primarily determined by radiation losses

Vocal Tract Transfer Function

- look at transfer function of pressure at the lips and volume velocity at the glottis, which is of the form:

$$H_a(\Omega) = \frac{P(\ell,\Omega)}{U_G(\Omega)} = \frac{P(\ell,\Omega)}{U(\ell,\Omega)} \cdot \frac{U(\ell,\Omega)}{U_G(\Omega)} = Z_L(\Omega) \cdot V_a(\Omega)$$

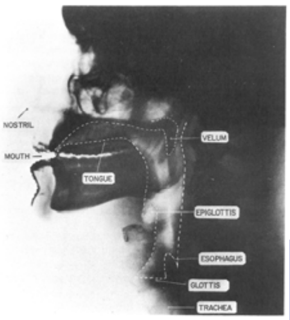


Notice:
- zero at *Ω=0*
- high frequency emphasis (compare with previous chart)

Vocal Tract Transfer Functions for Vowels

- using the frequency domain equations, can compute the frequency response functions for a set of area functions of the vocal tract for various vowel sounds, using all the loss mechanisms, assuming:
  - *A(x), 0≤x≤l* (glottis-to-lips) measured and known
  - steady state sounds (*dA/dt=0*)
  - measure $U(l,\Omega)/U_G(\Omega)$ for the vowels /AA/,/EH/,/IY/,/UW/

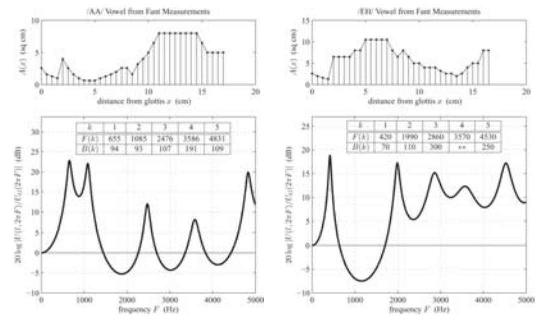
Area Function from X-Ray Photographs



Gunnar Fant, *Acoustic Theory of Speech Production*, Mouton, 1970

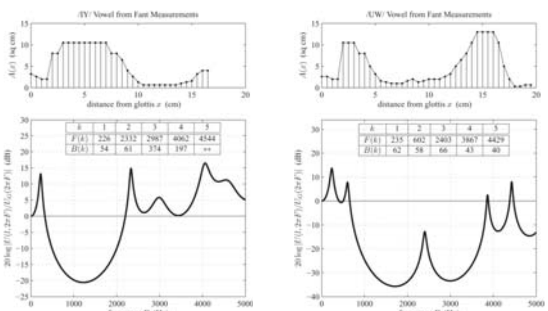
31

Area Functions and FR for Vowels /AA/ and /EH/



32

Area Functions and FR for Vowels /IY/ and /UW/



33

VT Transfer Functions

- the vocal tract tube 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, thermal losses, and radiation losses

34

Nasal Coupling Effects

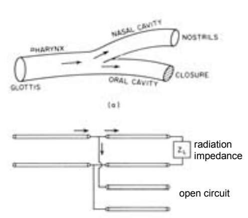


Fig. 3.27 (a) Tube model for production of nasals; (b) corresponding electrical analog.

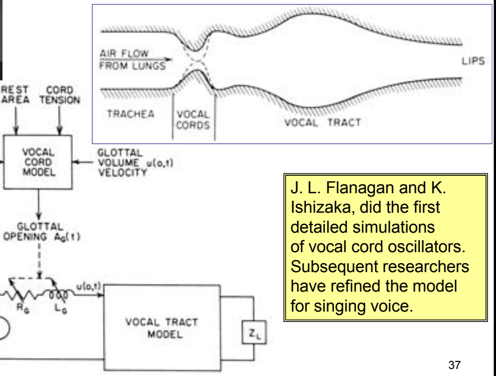
- at the branching point
  - sound pressure the same as at input of each tube
  - volume velocity is the sum of the volume velocities at inputs to nasal and oral cavities
- can solve flow equations numerically
  - results show resonances dependent on shape and length of the 3 tubes
- closed oral cavity 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 in VT

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 transient excitation (pop-like sound)

36

Vocal Cord Simulation



J. L. Flanagan and K. Ishizaka, did the first detailed simulations of vocal cord oscillators. Subsequent researchers have refined the model for singing voice.

37

Voiced Excitation in VT

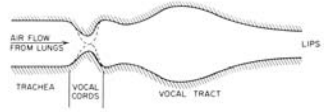


Fig. 3.28 Schematic representation of the vocal system.

- lung pressure is increased, causing air to flow out of the lungs and through the opening between the vocal cords (the glottis)
- according to Bernoulli's law, if the tension in the vocal cords is properly adjusted, the reduced pressure in the constriction allows the cords to come together, thereby constricting air flow (see dotted lines above)
- because of closure of the vocal cords, pressure increases behind the vocal cords and eventually reaches a level sufficient to force the vocal cords to open and allows air to flow through the glottis again

sustained Bernoulli oscillations => rate of opening and closing is controlled by air pressure in the lungs, tension and stiffness of the vocal cords, and area of the glottal opening; the vocal tract area at the glottis also effects the rate

Glottal Excitation Model

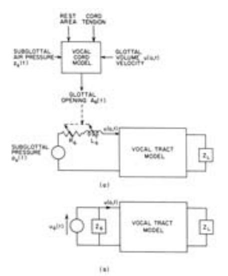


Fig. 3.29 (a) Diagram of vocal cord model. (b) approximate model for vocal cords.

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

build-up of glottal pulses

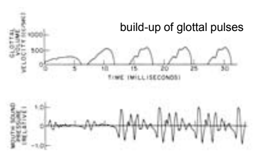
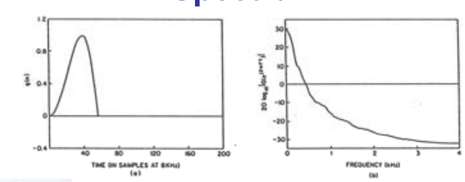


Fig. 3.30 Glottal volume velocity and sound pressure at the mouth for vowel /a/. (After Ishizaka and Flanagan [30].)

Rosenberg Glottal Pulse and Spectrum



$$
\begin{aligned}
g[n] &= 0.5[1-\cos(\pi n/N_1)] && 0 \le n \le N_1 \\
&= \cos[\pi(n-N_1)/(2N_2)] && N_1 \le n \le N_1+N_2 \\
&= 0 && \text{otherwise}
\end{aligned}
$$

Note the high frequency fall off due to the lowpass pulse shape

40

Other Excitation Sources

- voiceless excitation occurs at a constriction of the vocal tract when volume velocity exceeds a critical value (called the Reynolds number) => this can be modeled using a randomly time varying source at the point of constriction
- a combination of voiced and voiceless excitation is used for voiced fricatives
- a total closure of the tract is used for stop consonants

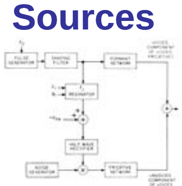


FIG. 2. Excitation network for voiced fricatives.

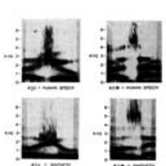
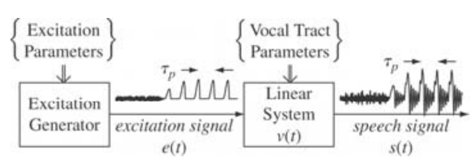


FIG. 3. Spectrographic examples of voiced fricatives.

41

Source-System Model



Pitch, Voiced/Unvoiced, Amplitude

Formants, Vocal Tract Area Functions, Articulatory Paramameters

42

Summary of Losses, Radiation and Boundary Condition Effects

- considered losses due to friction at walls, heat conduction through walls, vibration of walls
- losses introduce new terms into sound propagation equations
- effects of losses are increased bandwidth of complex poles (from 0 to a finite quantity) and changes in the regular spacing of the resonance (formant) frequencies of the tract
- radiation at lips adds a parallel resistance and inductance component and is most significant at higher frequencies
- nasal coupling adds components to solution which include anti-resonances (frequency response zeros)
- sound excitation models lead to simplified model with a distinct glottal pulse (for voiced speech) with strong high frequency drop-off in level
- the overall vocal tract is well modeled as a variable excitation generator exciting a time-varying linear system

43

Lossless Tube Models

- approximate *A(x)* by a series of lossless, constant cross sectional area, acoustic tubes of the form shown at the right
- as the number of tubes becomes larger (smaller approximation error for the vocal tract area function), the approximation error for modeling the vocal tract goes to zero

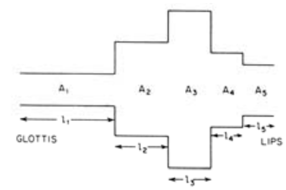
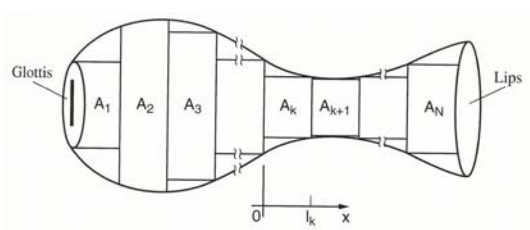


Fig. 3.32 Concatenation of 5 lossless acoustic tubes.

How do we use the lossless tube model to solve for various vocal tract transfer functions

Concatenated Tube Models



45

Wave Propagation in Lossless Tubes

- since each individual tube is lossless, can solve the basic wave equation for each individual tube, giving:

  - for k^{th} tube:

$$p_k(x,t) = \frac{\rho c}{A_k}\left[u_k^+(t - x/c) + u_k^-(t + x/c)\right], \quad 0 \le x \le \ell_k$$

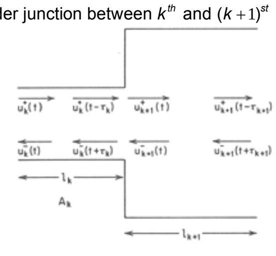
$$u_k(x,t) = u_k^+(t - x/c) - u_k^-(t + x/c), \quad 0 \le x \le \ell_k$$

- where x is the distance measured from the left-hand end of the k^{th} tube ($0 \le x \le \ell_k$) and $u_k^+(\ )$ and $u_k^-(\ )$ are positive-going and negative-going traveling waves in the k^{th} tube

46

Wave Propagation in Lossless Tubes

- boundary conditions at edges of adjacent tubes state that both pressure and volume velocity must be continuous in both time and space
- consider junction between k^{th} and $(k+1)^{st}$ tubes



47

Lossless Tube Junction

- at the junction between k^{th} and $(k+1)^{st}$ tubes, we get:

$$p_k(\ell_k, t) = p_{k+1}(0,t)$$
$$u_k(\ell_k, t) = u_{k+1}(0,t)$$

- substituting from the previous set of equations, we get:

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

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

- where $\tau_k = \ell_k / c$ is the time for a wave to travel the length of the k^{th} tube
- at the junction between tubes, part of the positive going wave is propagated to the right while part is reflected back to the left
- similarly part of the negative going wave is propagated to the left while part is reflected back to the right

48

Lossless Tube Model

- solve for $u^+_{k+1}(t)$ and $u^-_k(t+\tau_k)$ in terms of $u^-_{k+1}(t)$ and $u^+_k(t-\tau_k)$ to see how forward and reverse travelling waves propagate

$$u^+_{k+1}(t) = \left[\frac{2A_{k+1}}{A_{k+1}+A_k}\right]u^+_k(t-\tau_k) + \left[\frac{A_{k+1}-A_k}{A_{k+1}+A_k}\right]u^-_{k+1}(t)$$

$$u^-_k(t+\tau_k) = -\left[\frac{A_{k+1}-A_k}{A_{k+1}+A_k}\right]u^+_k(t-\tau_k) + \left[\frac{2A_k}{A_{k+1}+A_k}\right]u^-_{k+1}(t)$$

- the quantity

$$r_k = \left[\frac{A_{k+1}-A_k}{A_{k+1}+A_k}\right] = \text{amount of } u^-_{k+1}(t) \text{ that is reflected at the junction}$$

- r_k is called the 'reflection coefficient' for the k^{th} junction, with $-1 \le r_k \le 1$, and rewriting above equations as

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

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

49

Lossless Tube Model

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

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

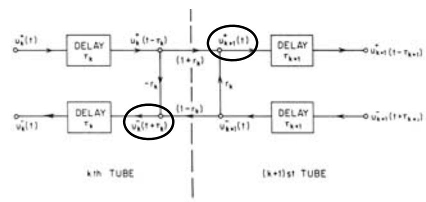


Fig. 3.34 Signal-flow representation of the junction between two lossless tubes.

50

Boundary Conditions

- for a 5-tube model there are 5 sets of forward and backward delays, 4 junctions (each characterized by a reflection coefficient), and a set of boundary conditions at the lips and glottis
- assume N sections (indexed 1 to N) starting at the glottis
- want to relate pressure $(p_N(l_N,t))$ and volume velocity $(u_N(l_N,t))$ at output of N^{th} tube to the radiated pressure and volume velocity (at the lips)

51

Boundary Conditions

- earlier we saw that at the lips we have

$$P_N(\ell_N,\Omega) = Z_L \cdot U_N(\ell_N,\Omega)$$

- giving the time domain relation

$$\frac{\rho c}{A_N}\left[u^+_N(t-\tau_N) + u^-_N(t+\tau_N)\right] = Z_L\left[u^+_N(t-\tau_N) - u^-_N(t+\tau_N)\right]$$

- solving for $u^-_N(t+\tau_N)$ we get

$$u^-_N(t+\tau_N) = -r_L u^+_N(t-\tau_N)$$

- where the reflection coefficient at the lips is

$$-1 \le r_L = \left[\frac{\rho c / A_N - Z_L}{\rho c / A_N + Z_L}\right] \le 1$$

52

Termination at Lips

$$u^-_N(t+\tau_N) = -r_L\, u^+_N(t-\tau_N)$$

- output volume velocity at the lips is

$$u_N(\ell_N,t) = u^+_N(t-\tau_N) - u^-_N(t+\tau_N)$$
$$= (1+r_L)u^+_N(t-\tau_N)$$

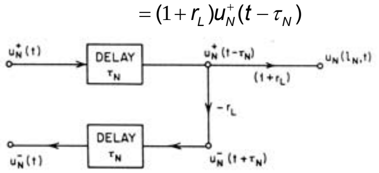


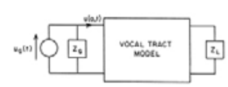
Fig. 3.35 Termination at lip end of a concatenation of lossless tubes. 53

Adding the Excitation Source

- assume that the excitation source and the vocal tract are linearly separable, then at the first tube we have the boundary relation:

$$U_1(0,\Omega) = U_G(\Omega) - P_1(0,\Omega)/Z_G$$

$$u^+_1(t) - u^-_1(t) = u_G(t) - \frac{\rho c}{A_1}\left[\frac{u^+_1(t) + u^-_1(t)}{Z_G}\right]$$



- solving for $u^+_1(t)$ we get:

$$u^+_1(t) = \frac{(1+r_G)}{2}u_G(t) + r_G u^-_1(t)$$

- where the reflection coefficient, r_G, is

$$-1 \le r_G = \left[\frac{Z_G - \dfrac{\rho c}{A_1}}{Z_G + \dfrac{\rho c}{A_1}}\right] \le 1$$

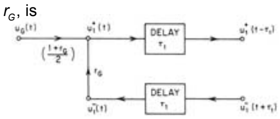


Fig. 3.36 Termination at glottal end of a concatenation of lossless tubes.

Lossless Two Tube Model

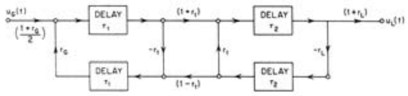


Fig. 3.37 Complete flow diagram of a two-tube model.

- volume velocity at lips is $u_L(t) = u_2(\ell_2, t)$
- transfer function from glottis to lips is

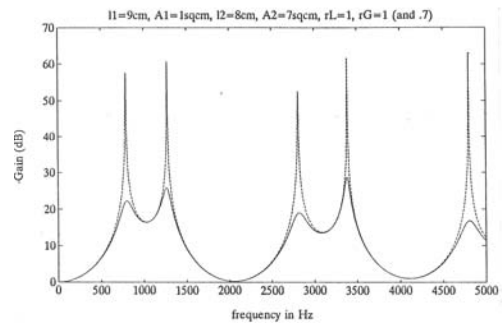
$$V_a(\Omega) = \frac{U_L(\Omega)}{U_G(\Omega)}$$

$$= \frac{0.5(1+r_G)(1+r_L)(1+r_1)e^{-j\Omega(\tau_1+\tau_2)}}{1+r_1r_Ge^{-j\Omega 2\tau_1}+r_1r_Le^{-j\Omega 2\tau_2}+r_Lr_Ge^{-j\Omega 2(\tau_1+\tau_2)}}$$

- note total delay of $(\tau_1 + \tau_2)$ is total propagation delay from glottis to lips

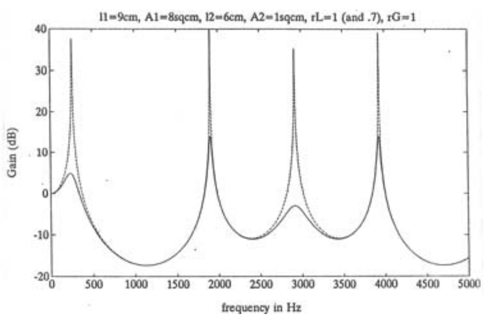
55

Two-Tube Model for Vowel /AA/



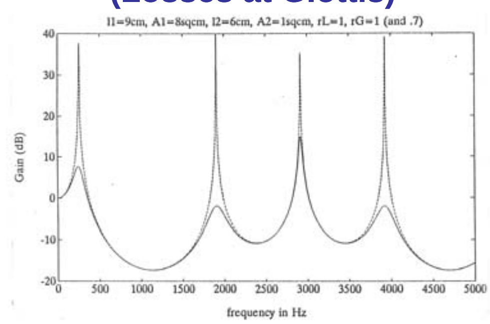
56

Two-Tube Model for Vowel /IY/ (Losses at Lips)



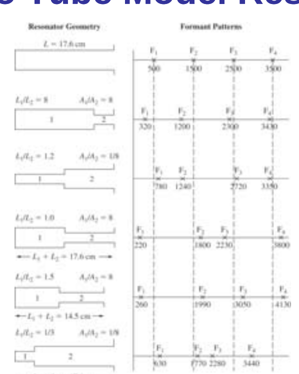
57

Two-Tube Model for Vowel /IY/ (Losses at Glottis)



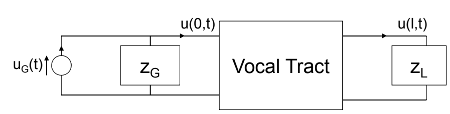
58

Two Tube Model Resonances



59

Summary of VT Model



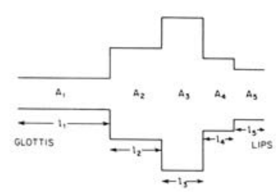
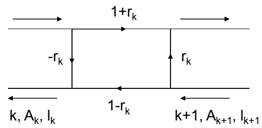


Fig. 3.32 Concatenation of 5 lossless acoustic tubes.

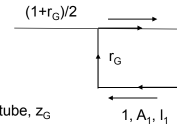
60

Lossless VT Sections

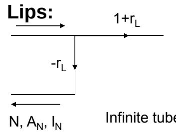


$r_k = \dfrac{A_{k+1} - A_k}{A_{k+1} + A_k}$; reflection coefficient

Glottis:



Lips:



$r_L = \dfrac{\dfrac{\rho c}{A_N} - z_L}{\dfrac{\rho c}{A_N} + z_L}$; reflection coefficient

$r_G = \dfrac{-\dfrac{\rho c}{A_1} + z_G}{\dfrac{\rho c}{A_1} + z_G}$; reflection coefficient

61

Relationship to Digital Filters

- observation that lossless tube model appears similar to digital filter implementations => consider a system of *N* lossless tubes, each of length *Δx=l/N* where *l* is the overall length of the vocal tract

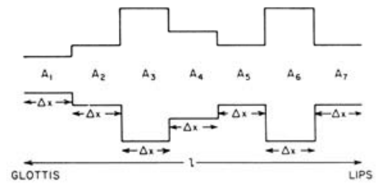


Fig. 3.38 Concatenation of ($N=7$) lossless tubes of equal length.

- all delays equal to $\tau = \Delta x/c$, the time to traverse the length of one tube

62

Relation to Digital Filters

- consider response to a unit impulse, $u_G(t) = \delta(t)$
- impulse propagates down the series of tubes, being partially reflected and partially propagated at the junctions; can show that the impulse response (the volume velocity at the lips) is of the form:

$$v_a(t) = \alpha_0 \delta(t - N\tau) + \sum_{k=1}^{\infty} \alpha_k \delta(t - N\tau - 2k\tau)$$

↑ initial propagation with delay N x Tau

↑ successive impulses with one or more reflections

- overall system function is:

$$V_a(s) = \sum_{k=0}^{\infty} \alpha_k e^{-s(N+2k)\tau} = e^{-sN\tau} \sum_{k=0}^{\infty} \alpha_k e^{-s2k\tau}$$

63

Relation to Digital Filters

$$V_a(s) = e^{-sN\tau} \sum_{k=0}^{\infty} \alpha_k e^{-s2k\tau}$$
$$= e^{-sN\tau} \hat{V}_a(s) \leftrightarrow \hat{v}_a(t) = v_a(t + N\tau)$$

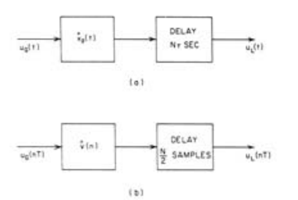


Fig. 3.39 (a) Block diagram representation of lossless acoustic tube model; (b) equivalent discrete-time system.

64

Relation to Digital Filters

- frequency response $\hat{V}_a(\Omega)$ is

$$\hat{V}_a(\Omega) = \sum_{k=0}^{\infty} \alpha_k e^{-j\Omega 2k\tau}$$

- can easily show that

$$\hat{V}_a(\Omega + \frac{2\pi}{2\tau}) = \hat{V}_a(\Omega) \quad => \text{ periodic in frequency}$$

- if input at glottis is bandlimited to frequencies below $\pi/(2\tau)$ we can sample the input at the glottis with period $T = 2\tau$ and filter the sampled signal with a digital filter whose IR is

$$\hat{v}[n] = \alpha_n \qquad n \geq 0$$
$$= 0 \qquad n < 0$$

- for $T = 2\tau$, the delay of $N\tau$ corresponds to a shift of $N/2$ samples giving the digital filter implementation of part b.

65

Signal Flow Graph Model

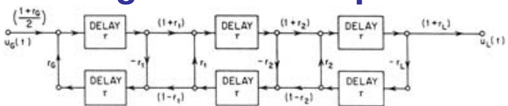
- z-transform of $\hat{v}(n)$ is simply $\hat{V}_a(s)$ with e^{sT} replaced by *z*, giving

$$\hat{V}(z) = \sum_{k=0}^{\infty} \alpha_k z^{-k}$$

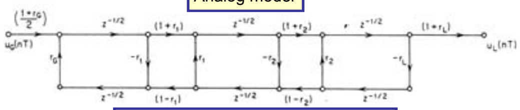
- signal flow graph for the discrete system can be obtained from the flow graph of the analog system by replacing each node variable in the analog system by the corresponding sequence of samples
- each τ second delay is replaced by a 1/2 sample delay since $\tau = T/2$
- section propagation delay is $z^{-1/2}$
- better digital signal flow graph model is of the form of a ladder with the delay elements only in the upper and lower paths; in the upper paths signals propagate to the right and in the lower paths they propagate to the left => can move delay elements in the lower paths to the upper paths; this only changes the overall delay from input to output, but has no practical significance

66

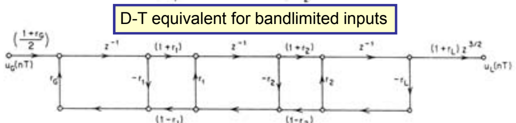
Signal Flow Graphs



Analog model



D-T equivalent for bandlimited inputs



D-T equivalent without half-sample delays

Transfer Function of Lossless Tube Model

- want to determine

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

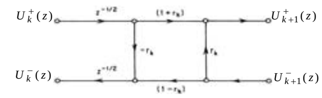
- at junctions we have the relations

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

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

- at lips use same formulation with fictitious $(N+1)st$ tube that is infinitely long (no negative going wave)=> $(N+1)st$ tube terminated in its characteristic impedance

$$U_{N+1}^{+}(z)=U_L(z)$$
$$U_{N+1}^{-}(z)=0$$
$$A_{N+1}=\frac{\rho c}{Z_L}, \qquad r_N=r_L$$



68

Transfer Function of Lossless Tube Model

- at the glottis

$$U_G(z)=\frac{2}{1+r_G}U_1^{+}(z)-\frac{2r_G}{1+r_G}U_1^{-}(z)$$

- putting it all together gives

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

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

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

69

Transfer Function of Lossless Tube Model

- consider a 2-section tube $(N=2)$

$$\frac{1}{V(z)}=\frac{2(1+r_1r_2z^{-1}+r_1r_Gz^{-1}+r_2r_Gz^{-2})z}{(1+r_G)(1+r_1)(1+r_2)}$$

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

- in general

$$V(z)=\frac{0.5(1+r_G)\left[\prod_{k=1}^{N}(1+r_k)\right]z^{-N/2}}{D(z)}$$

$$D(z)=\left[1,-r_G\right]\begin{bmatrix}1 & -r_1\\ -r_1z^{-1} & z^{-1}\end{bmatrix}\cdots\begin{bmatrix}1 & -r_N\\ -r_Nz^{-1} & z^{-1}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

70

Transfer Function of Lossless Tube Model

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

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

$$D_0(z)=1$$
$$D_k(z)=D_{k-1}(z)+r_kz^{-k}D_{k-1}(z^{-1}), \qquad k=1,2,\ldots,N$$
$$D(z)=D_N(z)$$

- Examples:

$$D_1(z)=1+r_1z^{-1}=D_0(z)+r_1z^{-1}D_0(z^{-1})=1+r_1z^{-1}(1)$$
$$D_2(z)=1+r_1z^{-1}+r_1r_2z^{-1}+r_2z^{-2}=D_1(z)+r_2z^{-2}D_1(z^{-1})$$
$$=1+r_1z^{-1}+r_2z^{-2}(1+r_1z)=1+r_1z^{-1}+r_1r_2z^{-1}+r_2z^{-2}$$

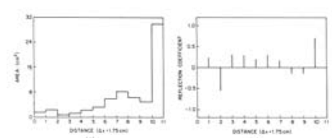
- choose $N=10$ as a reasonable number of tubes for model

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

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

71

Transfer Function of Lossless Tube Model



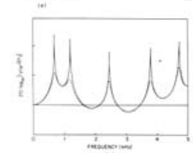


Fig. 3.43 (a) Area function for 10 section lossless tube terminated with reflectionless section of area 30 cm²; (b) reflection coefficients for 10 section tube; (c) frequency response of 10 section tube; dotted curve corresponds to conditions of (b); solid curve corresponds to short-circuit termination. (Note area data of (a) estimated from data given by Fant [1] for the Russian vowel /a/.)

72

Summary of Lossless Tube Models

1. The vocal tract area function, A, is now a function of x, $A(x)$

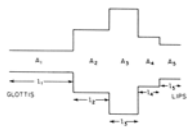


Fig. 3.32 Concatenation of 5 lossless acoustic tubes.

2. Solve the wave equation for the k^{th} tube:

$$p_k(x,t) = \frac{\rho c}{A}\left[u_k^+(t - x/c) + u_k^-(t + x/c)\right], \quad 0 \le x \le \ell_k$$

$$u_k(x,t) = \left[u_k^+(t - x/c) - u_k^-(t + x/c)\right], \quad 0 \le x \le \ell_k$$

Summary of Lossless Tube Models

3. add boundary conditions at the edges of adjacent tubes
   - both pressure and volume velocity must be continuous in both time and space at boundaries

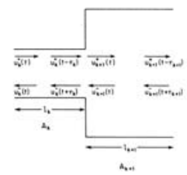


Fig. 3.33 Illustration of the junction between two lossless tubes.

1. $p_k(\ell_k, t) = p_{k+1}(0, t)$
2. $u_k(\ell_k, t) = u_{k+1}(0, t)$

Summary of Lossless Tube Models

4. at each junction:
   - part of the positive going wave is propagated to the right while part is reflected back to the left
   - part of the negative going wave is propagated to the left while part is reflected back to the right

5. $r_k = \dfrac{A_{k+1} - A_k}{A_{k+1} + A_k}$ = reflection coefficient for the k^{th} junction

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

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

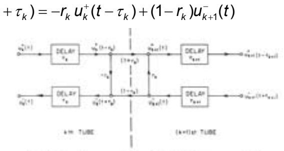


Fig. 3.34 Signal-flow representation of the junction between two lossless tubes.

Summary of Lossless Tube Models

6. for an N-tube model there are $(N-1)$ junctions with reflection coefficients
   - there are boundary conditions at the lips and glottis
7. need to relate $p_N(l_N, t)$ and $u_N(l_N, t)$ to pressure and volume velocity at the lips via:

$$r_L = \frac{\rho c / A_N - z_L}{\rho c / A_N + z_L}$$

where z_L is lip impedance, i.e.,

$$P_N(\ell_N, \Omega) = z_L U_N(\ell_N, \Omega)$$

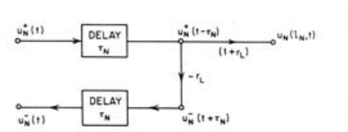


Fig. 3.35 Termination at lip end of a concatenation of lossless tubes.

Summary of Lossless Tube Models

8. boundary condition at the glottis:

$$r_G = \frac{z_G - \rho c / A_1}{z_G + \rho c / A_1}$$

where z_G is glottal impedance

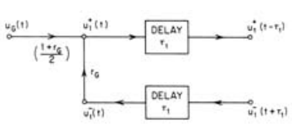


Fig. 3.36 Termination at glottal end of a concatenation of lossless tubes.

Summary of Lossless Tube Models

9. for the special case of a 2-tube model we get:

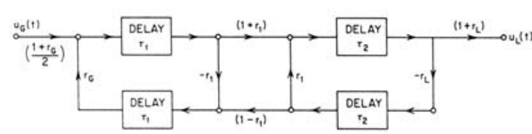


Fig. 3.37 Complete flow diagram of a two-tube model.

$$V_a(\Omega) = \frac{U_L(\Omega)}{U_G(\Omega)} = \frac{0.5(1 + r_G)(1 + r_1)(1 + r_L)e^{-j\Omega(\tau_1 + \tau_2)}}{1 + r_1 r_G e^{-j\Omega 2\tau_1} + r_1 r_L e^{-j\Omega 2\tau_2} + r_L r_G e^{-j\Omega 2(\tau_1 + \tau_2)}}$$

Digital Models for Speech

- sound is generated in 3 ways—each yielding a distinctive output
- vocal tract imposes resonance (and anti-resonance) structure so as to produce different speech sounds

=> this leads to a "terminal analog" speech model—correct at the terminal points, but without mimicing the physics of speech production

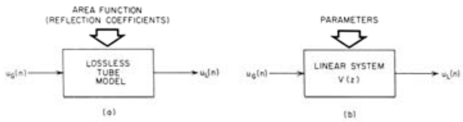


Fig. 3.44 (a) Block diagram representation of the lossless tube model; (b) terminal analog model.

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

- need to model time varying, $u_G(n)$

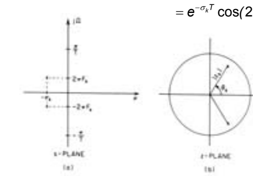
Digital Model Features

- poles of $V(z)$ correspond to vocal tract resonances
- missing zeros of $V(z)$ for nasals
  - modify $V(z)$
  - increase N to large number to model zeros
- roots of $V(z)$ are at:

$$s_k,\ s_k^* = -\sigma_k \pm j2\pi F_k \quad (s-plane)$$

$$z_k,\ z_k^* = e^{-\sigma_k T} e^{\pm j2\pi F_k T}$$

$$= e^{-\sigma_k T}\cos(2\pi F_k T) \pm j e^{-\sigma_k T}\sin(2\pi F_k T)$$

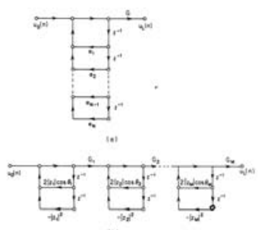


$$|z_k| = e^{-\sigma_k T},\ \theta_k = 2\pi F_k T$$

Fig. 3.45 (a) s-plane; and (b) z-plane representations of a vocal tract resonance.

Other Synthesis Implementations

- direct form difference equation
- cascade of second order systems



$$V(z) = \prod_{k=1}^{M} V_k(z)$$

$$V_k(z) = \frac{1 - 2|z_k|\cos(2\pi F_k T) + |z_k|^2}{1 - 2|z_k|\cos(2\pi F_k T)z^{-1} + |z_k|^2 z^{-2}}$$

Fig. 3.46 (a) Direct form implementation of all-pole transfer function; (b) cascade implementation of all-pole transfer function ($G_k = 1 - 2|z_k|\cos\theta_k + |z_k|^2$).

Radiation at Lips

$P_L(z) = R(z)U_L(z)$ -- high pass filtering

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

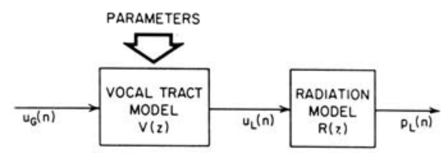


Fig. 3.47 Terminal analog model including radiation effects.

Excitation Model

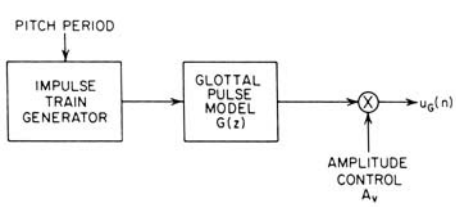
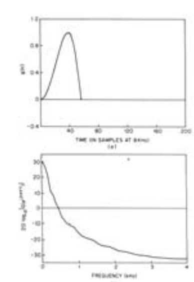


Fig. 3.48 Generation of the excitation signal for voiced speech.

Glottal Pulse Model

$$g[n] = 0.5[1 - \cos(\pi n / N_1)], \quad 0 \le n \le N_1$$
$$= \cos(\pi(n - N_1)/2N_2), \quad N_1 \le n \le N_1 + N_2$$
$$= 0 \quad \text{otherwise}$$

-lowpass filtering effect



General Synthesis Model

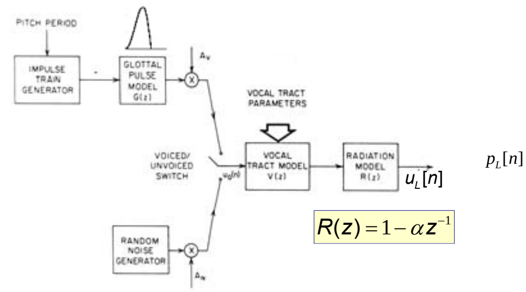
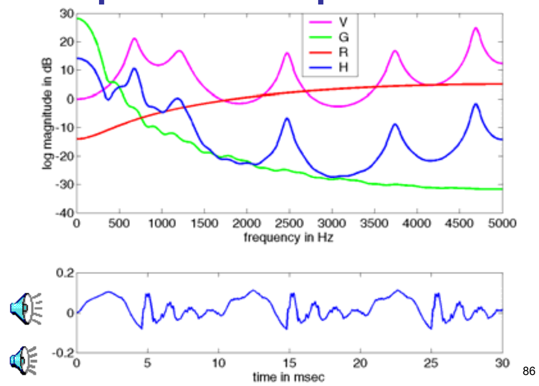


Fig. 3.50 General discrete-time model for speech production.

Components of Speech Model



Summary of Lecture

- Derived sound propagation equations for vocal tract
  - first considered uniform lossless tube
  - added simple models of loss
  - added model for radiation at lips
  - added source model at glottis
  - added nasal model for nasal tract
  - broadened the model to N-tube approximation—lossless case
  - looked at 2-tube models for simple vowels
  - examined range of digital speech production/synthesis models