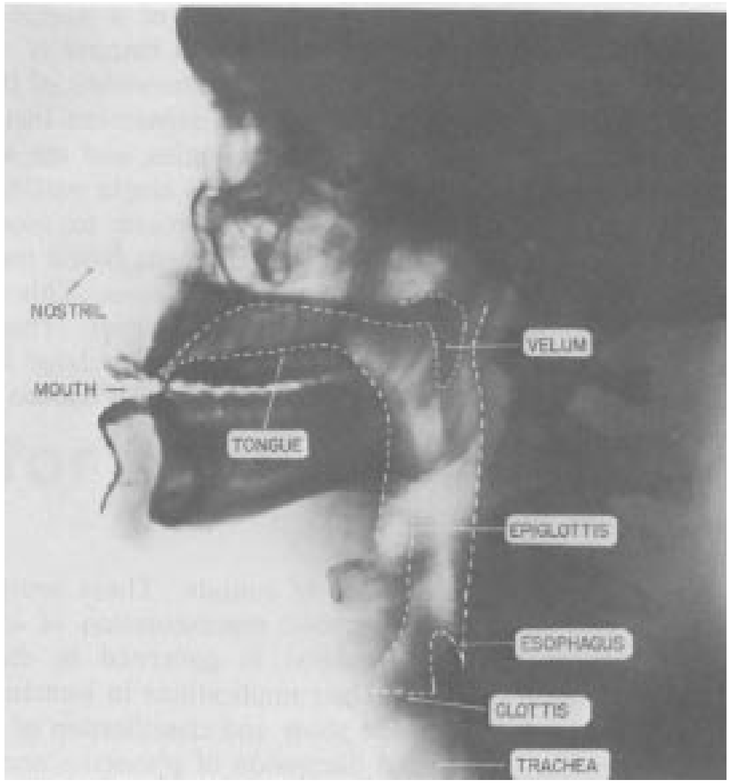
Sound Propagation



A detailed acoustic theory must consider the effects of the following:

- Time variation of the vocal tract shape
- Losses due to heat conduction and viscous friction at the vocal tract walls
- Softness of the vocal tract walls
- Radiation of sound at the lips
- Nasal coupling
- Excitation of sound in the vocal tract

Let us begin by considering a simple case of a lossless tube:

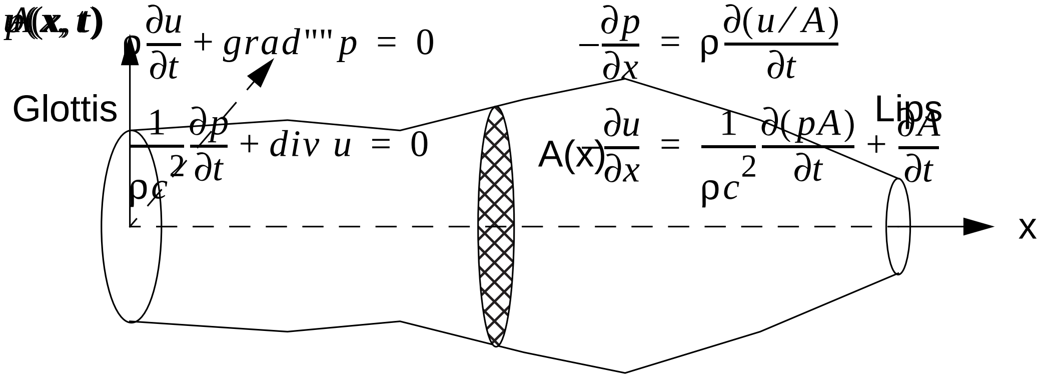
$A = A(x, t)$, $p = p(x, t)$

$$\rho \frac{\partial u}{\partial t} + grad\, p = 0$$

$$\frac{1}{\rho c^2} \frac{\partial p}{\partial t} + div\, u = 0$$

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

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



For frequencies that are long compared to the dimensions of the vocal tract (less than about 4000 Hz, which implies a wavelength of 8.5 cm), sound waves satisfy the following pair of equations:

$$\rho\frac{\partial(u/A)}{\partial t} + grad''''\,p = 0 \qquad\qquad -\frac{\partial p}{\partial x} = \rho\frac{\partial(u/A)}{\partial t}$$

or

$$\frac{1}{\rho c^2}\frac{\partial p}{\partial t} + \frac{\partial A}{\partial t} + div\ u = 0 \qquad\qquad -\frac{\partial u}{\partial x} = \frac{1}{\rho c^2}\frac{\partial(pA)}{\partial t} + \frac{\partial A}{\partial t}$$

where

$p = p(x, t)$	is the variation of the sound pressure in the tube
$u = u(x, t)$	is the variation in the volume velocity
ρ	is the density of air in the tube (1.2 mg/cc)
c	is the velocity of sound (35000 cm/s)
$A = A(x, t)$	is the area function (about 17.5 cm long)

Uniform Lossless Tube

If $A(x, t) = A$, then the above equations reduce to:

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

The solution is a traveling wave:

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

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

which is analogous to a transmission line:

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

What are the salient features of the lossless transmission line model?

where

Acoustic Quantity	*Analogous Electric Quantity*
p - pressure	v - voltage
u - volume velocity	i - current
ρ/A - acoustic inductance	L - inductance
A/(ρc^2) - acoustic capacitance	C - capacitance

The sinusoisdal steady state solutions are:

$$p(x,t) = jZ_o \frac{\sin[\Omega(l-x)/c]}{\cos[\Omega l/c]} U_G(\Omega) e^{j\Omega t}$$

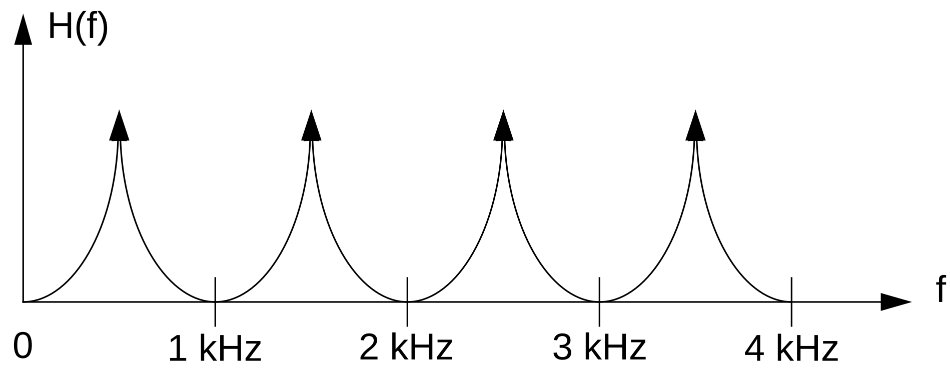
$$u(x,t) = \frac{\cos[\Omega(l-x)/c]}{\cos[\Omega l/c]} U_G(\Omega) e^{j\Omega t}$$

where $Z_0 = \frac{\rho c}{A}$ is the characteristic impedance.

The transfer function is given by:

$$\frac{U(l,\Omega)}{U(0,\Omega)} = \frac{1}{\cos(\Omega l/c)}$$

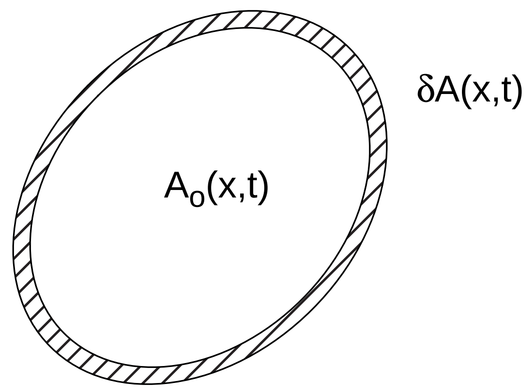
This function has poles located at every $\frac{(2n+1)\pi c}{2l}$. Note that these correspond to the frequencies at which the tube becomes a quarter wavelength: $\left(\frac{\Omega l}{c} = \frac{\pi}{2}\right) \Rightarrow \left(\Omega = \frac{c}{4l}\right)$.



Is this model realistic?

Effects of Losses

What do we predict the effects of yielding walls to be?



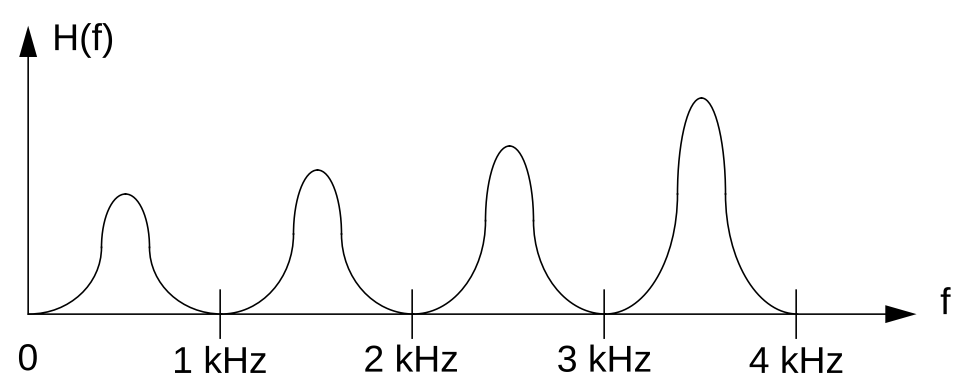
Use perturbation analysis:

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

We can develop a model that relates δA(x,t) to pressure:

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

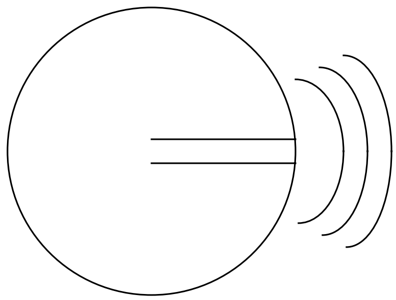
and solve for the new transfer function. But we can easily predict the effect of this:



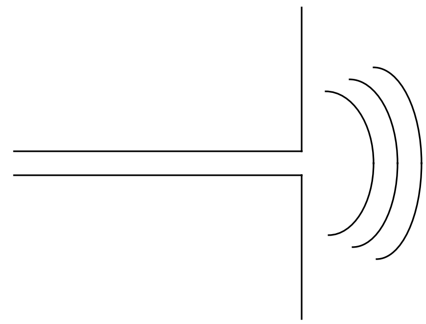
What would you expect to be the effect of friction and thermal losses?

Lip Radiation

How is the sound pressure wave within the vocal tract coupled into the air?



Radiation from a spherical baffle



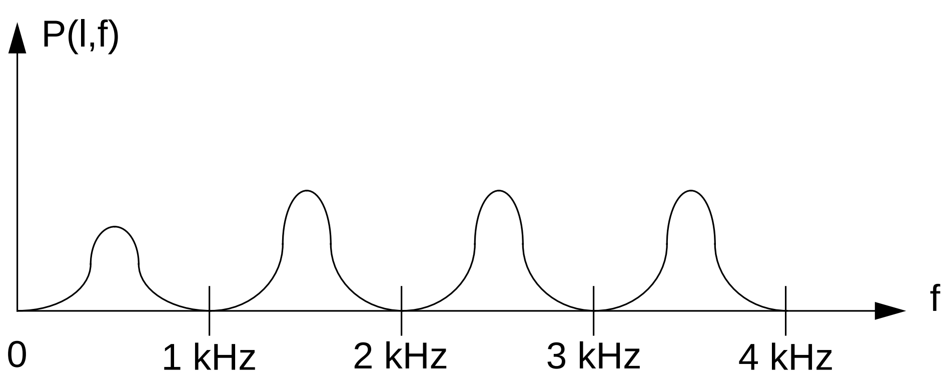
Radiation from an infinite plane baffle

Net effect is to place a complex load on the system:

$$Z_L(\Omega) = \frac{j\Omega L_r R_r}{R_r + j\Omega L_r} \quad \text{and} \quad P(l, \Omega) = Z_L(\Omega) U(l, \Omega)$$

where $R_r = 128/9\pi^2$ and $L_r = 8a/3\pi c$, and a is the radius of the opening.

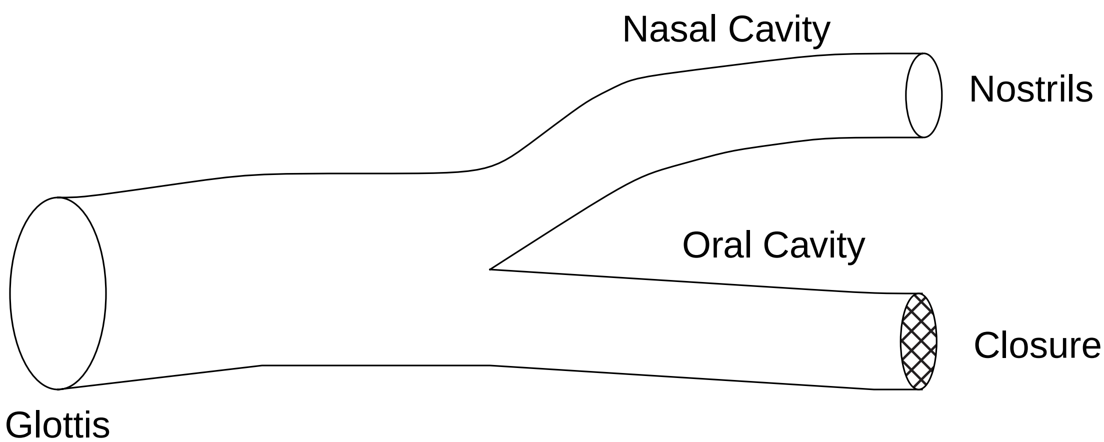
This impedance acts as a short circuit at low frequencies, and an imaginary impedance at high frequencies. The next effect on the volume velocity is to act as a highpass filter and to attenuate low frequencies. Lip radiation introduces a zero in the spectrum at DC and broadens the bandwidths at higher frequencies.



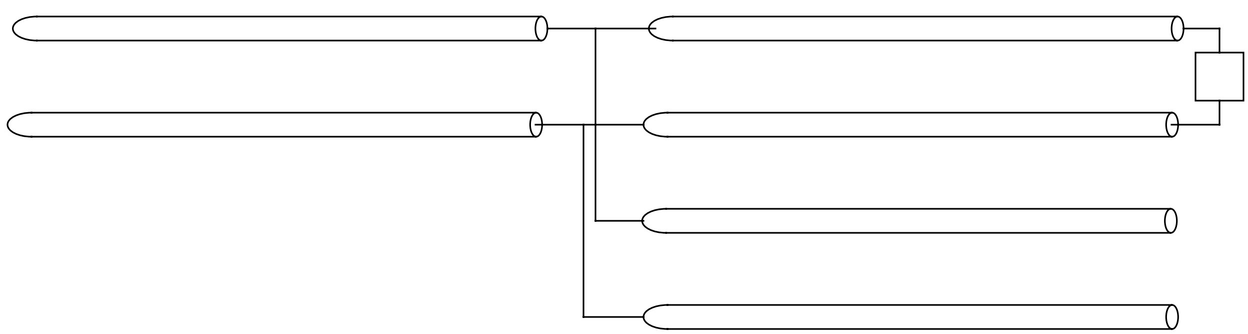
Nasal Coupling

How is the sound pressure wave within the vocal tract coupled into the air?

We also must worry about the nasal cavity, especially for labial sounds for which the mouth is closed during sound production.



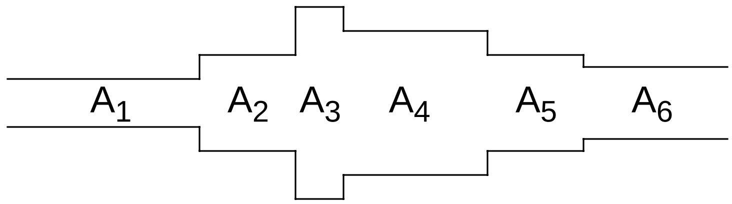
This is the equivalent of placing a transmission line in parallel with the vocal tract (oral cavity). What will the effect be?



The net effect is to produce a zero in the spectrum at about 1 kHz. As a result, nasal sounds (such as “m” and “n” in American English) have very little high frequency energy.

Piecewise Linear Approximations For The Vocal Tract

Consider the following approximation to the vocal tract area function:



Recall,

$$p_k(x, t) = \frac{\rho c}{A_k}[u_k^+(t - x/c) + u_k^-(t + x/c)]$$

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

For the kth section, if we apply the boundary conditions:

$$p_k(l_k, t) = p_{k+1}(0, t)$$

$$u_k(l_k, t) = u_{k+1}(0, t)$$

We can combine these two equations to show:

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

where $\tau_k = l_k / c$.

We can define a reflection coefficient for the kth junction:

$$r_k = \frac{u_{k+1}^+(t)}{u_{k+1}^-(t)} = \frac{A_{k+1} - A_k}{A_{k+1} + A_k}$$

It is easy to show that the reflection coefficients are bounded: $-1 \le r_k \le 1$.

The velocity can be expressed in terms of the reflection coefficients:

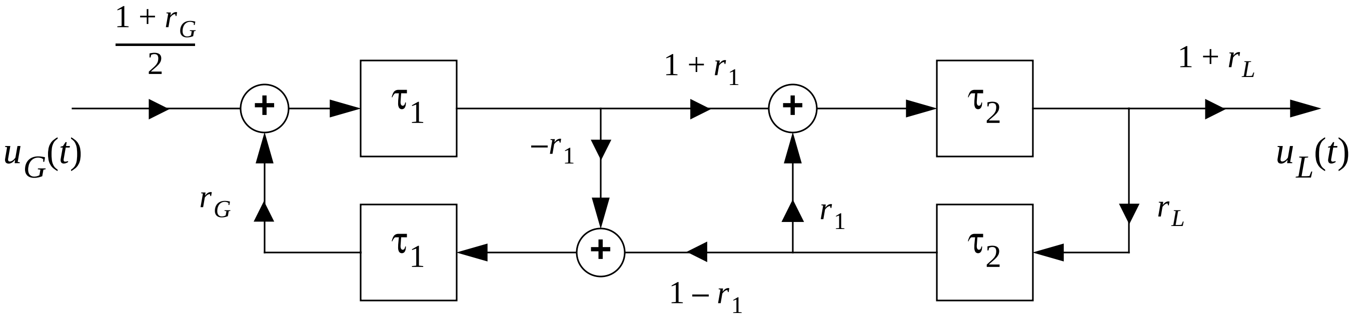
$$u_{k+1}^+(t) = (1 + r_k)u_k^+(t - \tau_k) + r_k u_{k+1}^-(t)$$

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

Ultimately, we will relate $\{r_k\}$ to a discrete model of the velocity profile.

Resonant Frequencies

Consider a two tube approximation to the vocal tract:

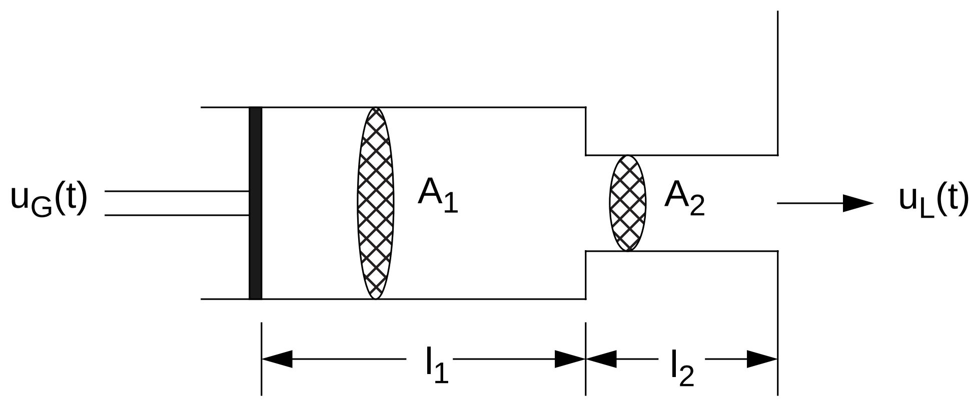


The frequency response of this system is:

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

What does this tell us about the frequency response?

If we consider the case $r_G = r_L = 1$:

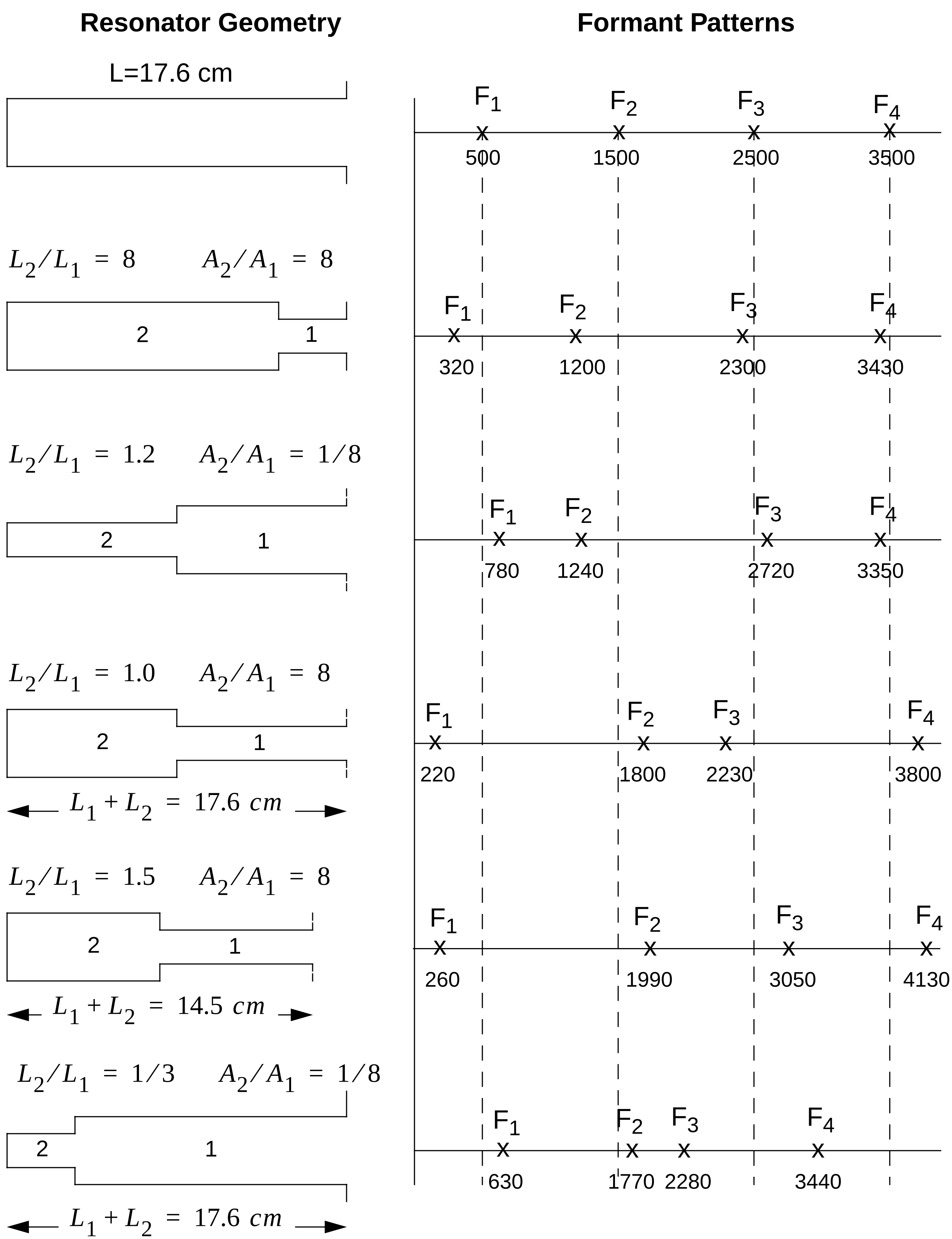


For this system, the poles are located at values that satisfy the equation:

$$\frac{A_1}{A_2}\tan(\Omega\tau_2) = \cot(\Omega\tau_1)$$

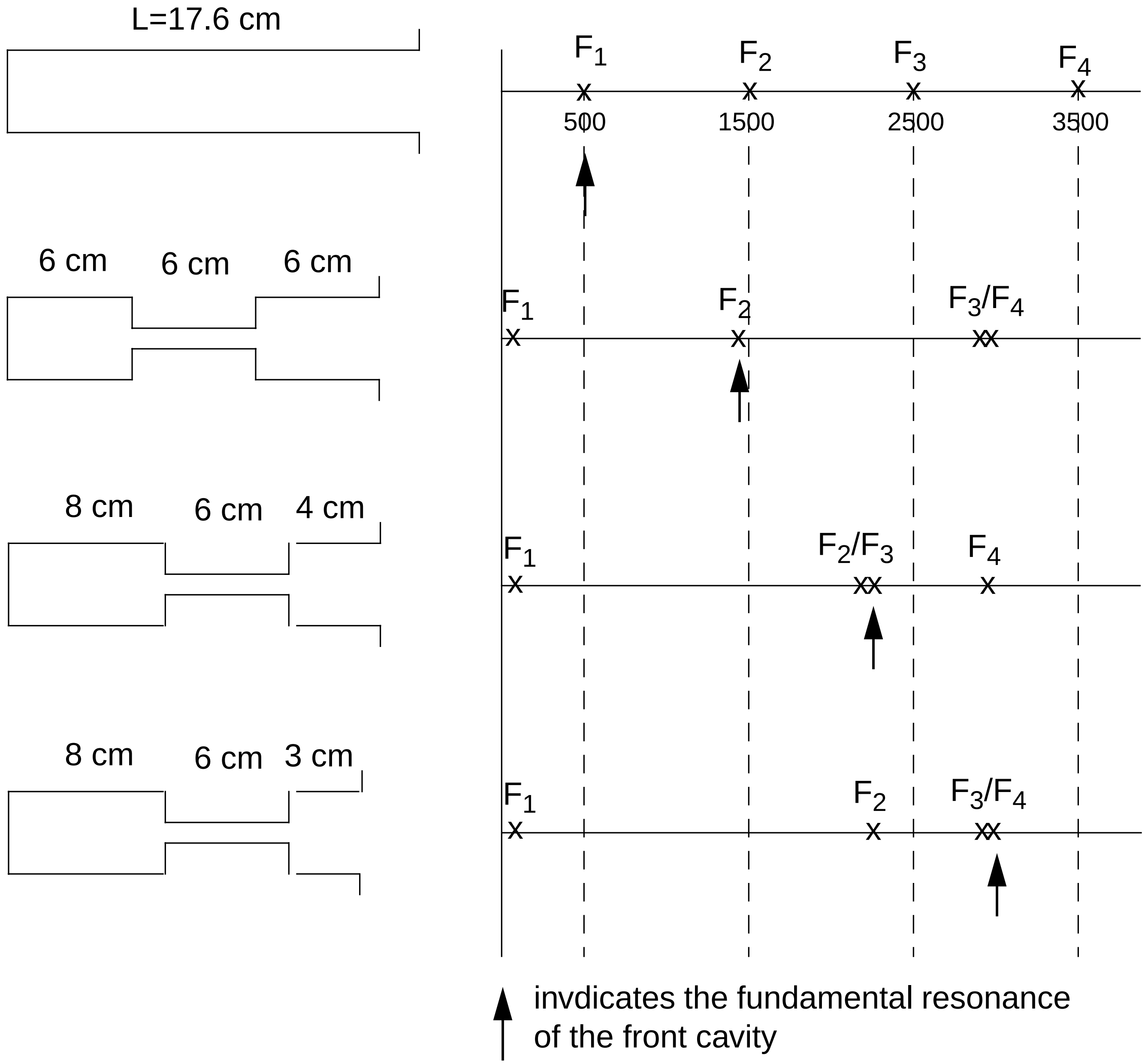
How does this compare to a single lossless tube?

Poles must be found through numerical analysis - nonlinear equation.


Resonator Geometry
Formant Patterns
L=17.6 cm
F1 F2 F3 F4
500 1500 2500 3500
L2/L1 = 8 A2/A1 = 8
2 1
F1 F2 F3 F4
320 1200 2300 3430
L2/L1 = 1.2 A2/A1 = 1/8
2 1
F1 F2 F3 F4
780 1240 2720 3350
L2/L1 = 1.0 A2/A1 = 8
2 1
L1 + L2 = 17.6 cm
F1 F2 F3 F4
220 1800 2230 3800
L2/L1 = 1.5 A2/A1 = 8
2 1
L1 + L2 = 14.5 cm
F1 F2 F3 F4
260 1990 3050 4130
L2/L1 = 1/3 A2/A1 = 1/8
2 1
L1 + L2 = 17.6 cm
F1 F2 F3 F4
630 1770 2280 3440

Three-Tube Models

Resonator Geometry **Formant Patterns**



Transfer Function of the Lossless Tube Model

Recall, $V(\Omega) = \frac{U_L(\Omega)}{U_G(\Omega)}$. In the discrete domain, we can write: $V(z) = \frac{U_L(z)}{U_G(z)}$.

Following our derivation of the wave equation, we can express the transfer function for a lossless tube as follows:

$$\boldsymbol{U}_k = \boldsymbol{Q}_k \boldsymbol{U}_{k-1}$$

where

$$\boldsymbol{U}_k = \begin{bmatrix} U_k^+(z) \\ U_k^-(z) \end{bmatrix} \quad \text{and} \quad \boldsymbol{Q}_k = \begin{bmatrix} \frac{z^{1/2}}{1+r_k} & \frac{-r_k z^{1/2}}{1+r_k} \\ \frac{-r_k z^{1/2}}{1+r_k} & \frac{z^{-1/2}}{1+r_k} \end{bmatrix}$$

The combined transfer function is a product of these matrices.
The net result is a transfer function that can be expressed as:

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

where

$$D(z) = \begin{bmatrix} 1 & -r_G \end{bmatrix} \begin{bmatrix} 1 & -r_1 \\ -r_1 z^{-1} & z^{-1} \end{bmatrix} \cdots \begin{bmatrix} 1 & -r_N \\ -r_N z^{-N} & z^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We can write $D(z)$ in a simpler form:

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

Why is this important?