EECS225d Audio Signal Processing in Humans and Machines **Spring 1995**

Lecture 4: January 25, 1995

Lecturer: Prof. Ben Gold *Scribe: Jeff Gilbert*

Speech Production Models

4.1 Physical Modeling

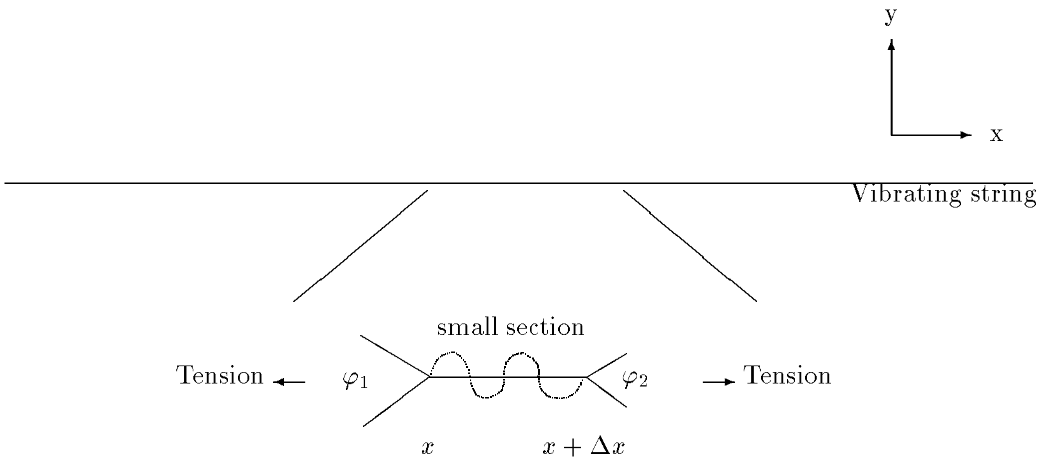
Speech and music can be modeled similarly by physical modeling of the speaker or instrument. For example, the sound of a trumpet is produced by excitation at the mouthpiece followed by a filtering by the shape and configuration of the trumpet body and valves. Similarly, the sound of speech is produced by excitation in the vocal chords which is then filtered by the vocal tract, mouth, and nasal cavity. Both speech and many instruments can be described via similar wave equations as we will see.

Different strings and stringed instruments have different modes of operation which lead to different equations for *Plucked strings* (for example, guitar), *Bowed strings* (for example, cello), *Struck strings* (for example, piano). There are also flexible bar instruments such as the xylophone which have their own set of equations. Also membrane instruments such as the drum have another set of equations.

The **Acoustic tube model** works for instruments such as the trumpet, clarinet, saxophone, as well as the vocal chords.

4.2 Analysis of waves in a string

Consider the problem of a string moving only in the y direction – a *transverse wave*, as seen below:



To simplify, we make the following assumptions:

- No stiffness in string
- Tension and mass are distributed and constant
- Δy displacement is small
- No gravity or friction

From the laws of physics, the force (F) due to the tension (T) in the figure is given by:

$$F = T(\tan\varphi_1 - \tan\varphi_2) \quad \text{with:} \quad \tan\varphi_1 = \left.\frac{\partial x}{\partial y}\right|_x \quad \text{and} \quad \tan\varphi_2 = \left.\frac{\partial x}{\partial y}\right|_{x+\Delta x} = \frac{\partial x}{\partial y} + \frac{\partial^2 y}{\partial x^2}\partial x$$

$$F = T\left(\frac{\partial x}{\partial y} + \frac{\partial^2 y}{\partial x^2}\partial x - \frac{\partial x}{\partial y}\right) = T\frac{\partial^2 y}{\partial x^2}\partial x$$

Thus the difference net force, caused by tension is $T\frac{\partial^2 y}{\partial x^2}\partial x$. Now we can use Newton's law relating force, mass, and acceleration: F=m$\cdot$a. The mass is the mass per unit length, or mass density, ϵ, times the length ∂x so m $= \epsilon\partial x$. Acceleration is defined as $\frac{\partial^2 y}{\partial t^2}$ so the net result is:

$$T\left(\frac{\partial^2 y}{\partial x^2}\right)\partial x = \epsilon\partial x\frac{\partial^2 y}{\partial t^2}$$

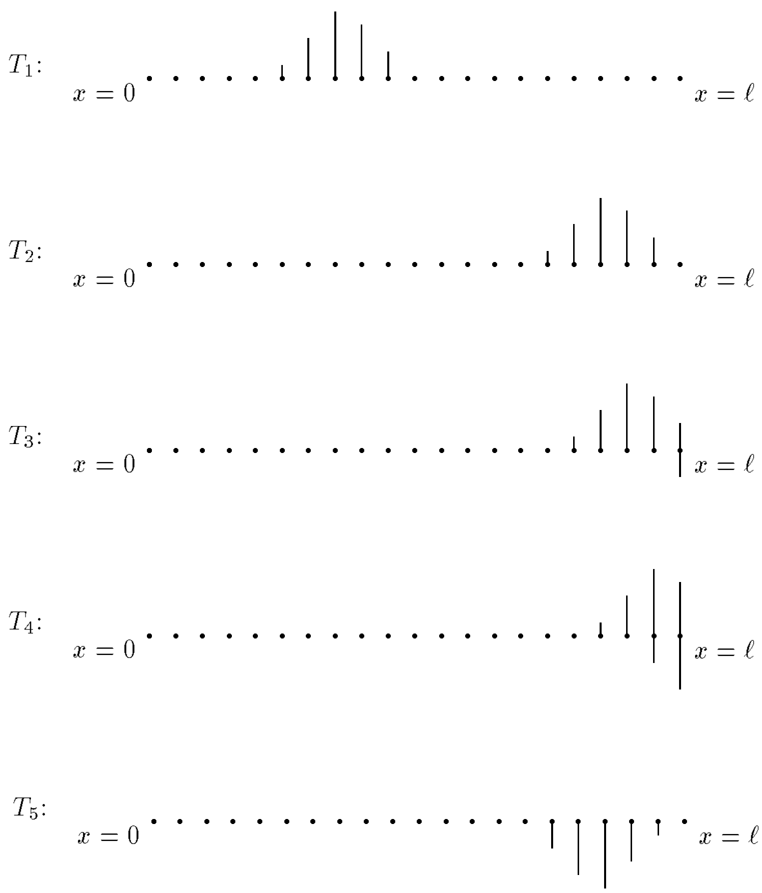
which can be rewritten as the **WAVE EQUATION**:

$$\boxed{\frac{\partial^2 y(x,y)}{\partial x^2} = \frac{1}{c^2}\cdot\frac{\partial^2 y}{\partial t^2}}$$

C can be interpreted as a velocity and $c^2 \equiv \frac{T}{\epsilon}$. This linear differential equation can then be solved by guessing that it is the product of a function of x and a function of t. I.e. $y(x,t) = f_1(x)\cdot f_2(t)$. Since it is a wave, we guess that it travels with time and thus $y(x,t) = f(x - ct)$ where $y(x,t)$ is the displacement of the string at point x at time t and f is any doubly differentiable function. F just determines the wave shape. ($y(x,t) = f(x + ct)$ will also be a solution.)

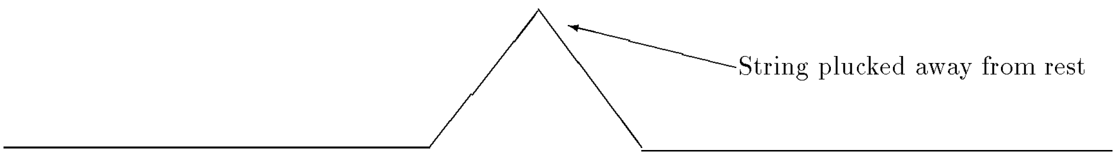
4.2.1 Boundary Conditions

But what about the *boundary conditions*? What happens at the ends of the string? To simplify, we will just consider discrete samples in space – discrete x. (See figure below). At time t_1, we see the non-zero part of the wave in its initial configuration. It has not yet reached the $x = \ell$ boundary at which $y(\ell) = 0$. At time t_2, the wave has traveled to the right a bit but has still not yet hit the boundary. Finally, at time t_3, the non-zero part of the wave has come to the $x = \ell$ part. The boundary condition forces this to be 0 so a *backward* or *reflected* traveling negative wave is created to make a net effect of $y(\ell) = 0$. At time t_4, the forward wave has traveled further right and the backward wave has more left. Since the reflected wave is of opposite sign, the two combine and cause some cancellation. Eventually, at time t_5 (in case you have not guessed, the times t_1, t_2, t_3, t_4, and t_5 are not evenly spaced), the forward wave is gone and the reflected wave remains and travels left. The wave would be reflected back into the positive y displacement at $x = 0$ and thus would continue indefinitely if it were a lossless system.

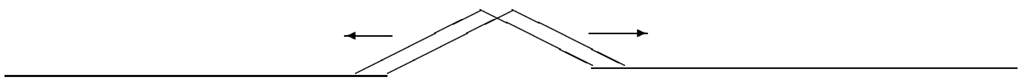


4.2.2 Another Example

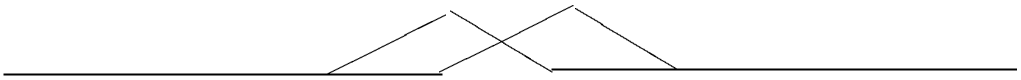
Consider an infinite string (no boundary conditions) plucked somewhere in the middle:



Since we cannot assume that it will go either way, we must assume that it goes both ways. Pretend that it is composed of two plucks - each of half the displacement and traveling opposite directions. The following occurs:



becomes:



which looks like:

until it becomes:

4.3 Analysis of waves in an acoustic tube

Now imagine that instead of a string, we have a tube filled with air. The air will have pressure and velocity at each time and place which are represented by $p(x,t)$ and $u(x,t)$ respectively. For the acoustic tube, we also have the constant ρ which is the density of the medium (air) in units of mass per unit volume. Furthermore, κ is a constant related to the elasticity of the medium. Using these constants, the speed of sound, c, can be determined with $c^2 = \frac{1}{\rho\kappa}$. (The speed of sound in air is 344 meters per second.) The pressure and volume of a wave in an acoustic tube can then be related by:

$$\boxed{\begin{gathered} \frac{\partial^2 p(x,y)}{\partial x^2} = \frac{1}{c^2} \cdot \frac{\partial^2 p}{\partial t^2} \\ \text{and} \\ \frac{\partial^2 u(x,y)}{\partial x^2} = \frac{1}{c^2} \cdot \frac{\partial^2 u}{\partial t^2} \end{gathered}}$$

4.3.1 Differences between string and acoustic tube models

- In the string model, it is a *transverse* wave – the wave moves in a direction perpendicular to the fluctuation.

- In the acoustic tube model, the wave is *longitudinal.* The wave travels in the same direction as the fluctuation.

4.3.2 Relationship between pressure and velocity

Next we investigate the relationship between the pressure and velocity of a wave traveling in an acoustic tube. Assume that the wave consists of one part moving to the right and one part moving to the left. E.g.:

$$\underbrace{u(x,t)}_{\text{total wave}} = \underbrace{u^+(t-\frac{x}{c})}_{\text{wave traveling right}} - \underbrace{u^-(t+\frac{x}{c})}_{\text{wave traveling left}}$$

Note that the u^+ and u^- waves are subtracted as their velocities are in opposite directions. The solution for pressure is similarly:

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

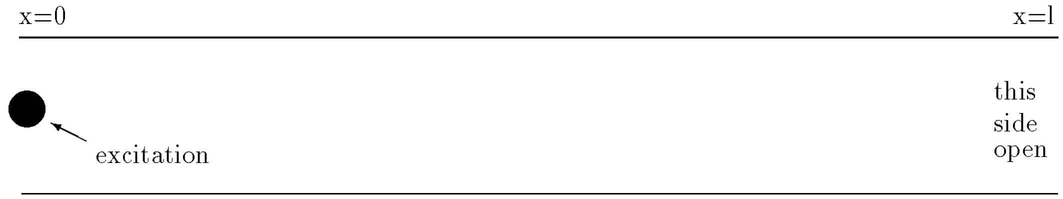
Here K is a constant which can be shown to be equal to ρc which is akin to the impedance of the system. Also note that the pressure goes as the sum of the u^+ and u^- waves, since the velocities are in opposite directions, if they have the same sign it means that they are more dissimilar and will thus lead to a higher pressure. Now if we replace K with Z_0 we get:

$$p(x,t) = Z_0\left[u^+(t-\frac{x}{c}) + u^-(t+\frac{x}{c})\right]$$

4.4 Finding the Modes of Vibration

<u>Fourier's Theorem</u> Any doubly differentiable function is a solution to the wave equation and can be decomposed into the sum of harmonics.

Assume a tube of length ℓ is excited at the left side by a sinusoidal function of the form $Ue^{j\Omega t}$. The boundary condition would be $u(0,t) = Ue^{j\Omega t}$. U is the magnitude of the excitation function. The tube is open at the right side so there is no pressure there, and thus the boundary condition is $p(\ell,t) = 0$. See below:



This solves to:

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

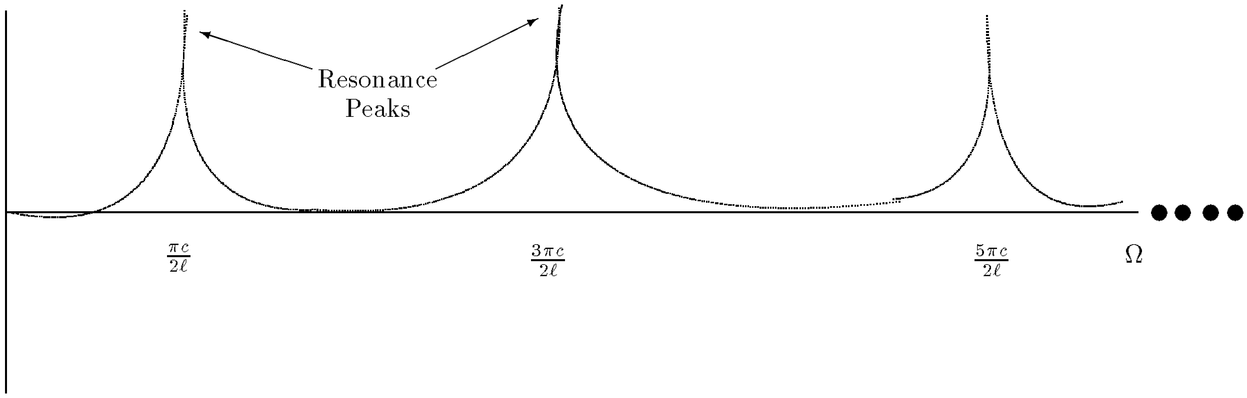
which yields potentially complex coefficients:

$$p(x,t) = Z_0\left[k^+e^{j\Omega(t-\frac{x}{c})} + k^-e^{j\Omega(t+\frac{x}{c})}\right]$$

Now if we solve to find k^+ and k^- we get:

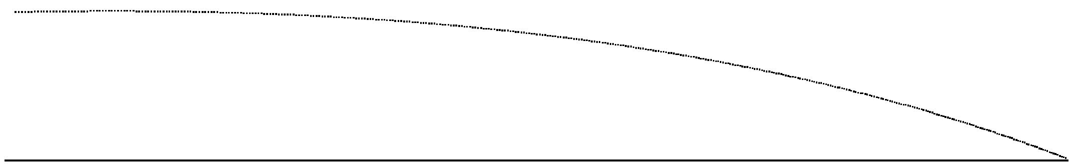
$$u(x,t) = Ue^{j\Omega t} \cdot \frac{\cos \frac{\Omega}{c}(\ell - x)}{2 \cos \frac{\Omega \ell}{c}} \qquad \text{where } \Omega = \frac{\pi c}{2\ell}, \frac{3\pi c}{2\ell}, \frac{5\pi c}{2\ell}, \ldots$$

Hence Ω corresponds to the modes of the oscillation which would appear as resonance peaks in the frequency response: (In a lossless system, these would be infinite.)

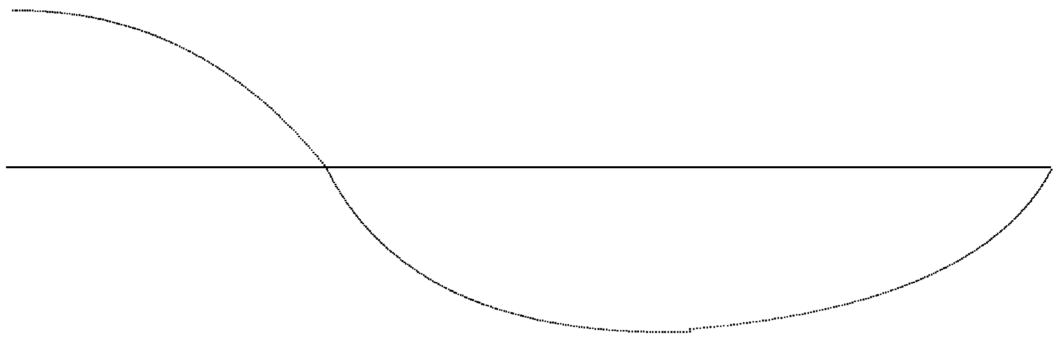


4.4.1 Physical String Analogy

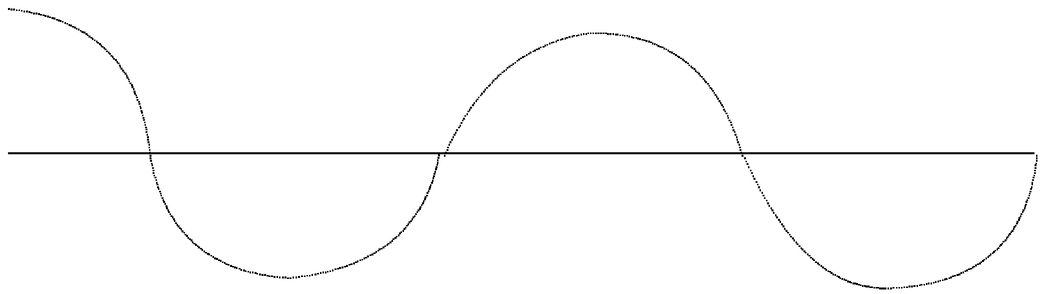
In a string, the modes would appear as waves with $\frac{1}{4}$, $\frac{3}{4}$, $\frac{5}{4}$, ... periods over the length of the string. I.e. the first mode for $\Omega = \frac{\pi c}{2\ell}$ would look like:



The second mode for $\Omega = \frac{3\pi c}{2\ell}$ would look like:



and the third mode for $\Omega = \frac{5\pi c}{2\ell}$ would look like:



and so on...