Lecture 11: February 10, 1995

Lecturer: Professor Nelson Morgan *Scribe: Kristine Ma*

11.1 Overview

The topic of this lecture is on room acoustics. For simplicity, we will assume that wave motions are originated from a point source. In this lecture, we will discuss the effect of room boundaries on a sound wave, the phenomenon of reverberation (i.e. the bouncing around of a sound wave and its attenuation), the effect of reverberation on speech intelligibility, and the early echo response - another effect on perception.

11.2 Sound Waves

The speed of sound is

$$c = 331.4 + 0.6\Theta \ m/s$$

where Θ is in degrees celsius. At room temperature, the speed of sound is 1133 ft/sec, which is approximately 1 msec/ft. If the wall is located at 10 feet away, it will take 20 msec for the sound wave to bounce back. A 20 msec delay is too short a duration for human to sense an echo. If the wall is located at 50 feet away, it will take 100 msec for the sound wave to bounce back, a long enough gap to induce an echo sensation.

In a real environment, the speed of sound differs in different parts of the room. This is due to factors such as temperature variations, movements of entities around the room, etc. Thus the equation for the speed of sound as given above is a gross approximation; however, for a sound wave at mid frequency range, this linear model is often sufficiently accurate.

11.2.1 1-D Wave Equation

The 1-D wave equation relating spatial (in x-direction), temporal (in t), and pressure (in p) is

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

It has a general solution of the form

$$p(x,t) = F(ct - x) + G(ct + x),$$

where F(ct-x) represents a right-going wave, and G(ct+x) represents a left-going wave. For sinusoidal functions, λ = c/f (i.e. wave length = speed of sound divided by frequency).

To give some physical intuition, if we assume that

$$c \approx 1000 ft/sec,$$

then

a 20 Hz wave has $\approx$ 50 feet long wave length
a 1 KHz wave has $\approx$ 1 foot long wave length
and, a 20 KHz wave has $\approx$ 0.5 inch long wave length

Thus our head is a pretty big barrier to high frequency sound waves. This is sometimes called head shadowing.

11.2.2 Spherical Wave Equation

The wave equation in spherical coordinate is

$$\frac{\partial^2 p}{\partial r^2} + \frac{2}{r}\frac{\partial p}{\partial r} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}.$$

One solution is a complex exponential,

$$p(r,t) = P_0 \frac{e^{j(wt-kr)}}{r},$$

where p is the pressure, r is the distance between the source and the receiver, P_0 is the amplitude of the sinusoidal sum, t is time, $k = \frac{2\pi}{\lambda}$, and $\omega = kc$. Note that the pressure is inversely proportional to r.

11.3 Intensity

Intensity is defined as the amount of sound energy flowing across a unit area surface in a second:

$$I = \bar{pv} = \frac{\bar{p^2}}{\rho_0 c},$$

where ρ_0 is the medium density, c is the speed of sound as before, and $\rho_0 c$ is the characteristic impedance. For a sinusoid,

$$I = \frac{P_0^2}{2\rho_0 c}.$$

So $I \propto p^2$, and for spherical wave $p \propto \frac{1}{r}$ implies $I \propto \frac{1}{r^2}$.

11.4 dB Sound Levels

Conventionally, the difference between sound energy levels is measured in decibels:

$$L = 20log_{10}|\frac{p_1}{p_2}| = 10log_{10}|\frac{I_1}{I_2}|.$$

The reference values are usually chosen so that it corresponds to typical threshold of hearing at 1 kHz, namely

$$p_2 = 2\text{x}10^{-5}N/m^2 \qquad and \qquad I_2 = 10^{-12}W/m^2.$$

The dB levels then become Sound Pressure Level (SPL) and Intensity Level (IL), respectively. They are about the same for travelling plane or spherical waves.

11.5 Typical Power Sources

Since we are dealing with speech, we will assume that the point source propagates hemispherically rather than spherically. The followings are typical power values for various speech source. The number in parenthesis indicates the Sound Pressure Level of the sound wave flowing across a one meter square area at 16 inches away from the source.

- Whispered speech: 1 nW (30 dB SPL)
- Average for speech: 10 uW (70 dB SPL)
- Loud speech: 200 uW (83 dB SPL)
- Shouting: 1mW (90 dB SPL)

If we ignore boundaries, SPL would be 6 dB lower for doubled distance. Note that shouting and whispered speech differ by 60 dB, which is a factor of a million ratio in power!

This indicates that sound pressure level is not directly related to human perceivable loudness. However, its cube root could be used as an approximation to loudness, i.e. 10 dB SPL difference doubles loudness. However, loudness is frequency dependent. Figure 13.10 displays the equal loudness level contours expressed in decibels relative to $10^{-12} watt/m^2$. Notice that the apparent loudness increases as the intensity level increases, but they are not the same. Perceivable loudness is weaker for lower frequencies, and the curve is flatter for louder sound. The "loudness" control found on some stereos as opposed to the "volume", does some filtering to adjust for human sensitivity characteristics. Figure 6.2 plots three such filters response. Weighting curve A is typically used for low loudness level, where a large adjustment on energy level need to be made on the low frequencies.

11.6 Room Acoustic Factors

We will consider three room acoustic factors: room modes, losses at boundaries, and early reflection patterns.

11.6.1 Room Modes

Modelling a room as a box involves the same mathematics as for strings and horns. At characteristic frequencies, the acoustic can be modelled as standing waves. Recall that the natural resonances of a resonator is inversely proportional to its dimensions, thus room resonances are mostly significant to low frequency signals. At high frequencies, the mode characteristic essentially becomes a continuum since the number of modes below frequency f is proportional to f^3.

11.6.2 Boundary Losses

Boundary loss refers to the lost of energy when sound wave strikes a boundary. The lost energy (fraction of energy that doesn't reflects back) is often turned into heat or simply transmitted through the boundary. Absorption coefficients are use for surfaces characterization, and they are frequency dependent (see Table below). Absorption in air has an intensity factor of e^{-mr}, where m is 0.0013 per meter at 1 kHz, and 0.021 at 4 kHz. These two numbers are at 50% relative humidity.

Material	Frequency, cycles/sec		
	125	500	2000
Acoustic paneling	0.16	0.50	0.80
Acoustic plaster	0.30	0.50	0.55
Brick wall, unpained	0.02	0.03	0.05
Draperies, light	0.04	0.11	0.30
Draperies, heavy	0.10	0.50	0.82
Felt	0.13	0.56	0.65
Floor, concrete	0.01	0.02	0.02
Floor, wood	0.06	0.06	0.06
Floor, carpeted	0.11	0.37	0.27
Glass	0.04	0.05	0.05
Marble or glazed tile	0.01	0.01	0.02
Plaster	0.04	0.05	0.05
Rock wool	0.35	0.63	0.83
Wood paneling, pine	0.10	0.10	0.08

Table 11.1: Listed values are only representative in that actual values also depend on mounting and thickness of the material.

11.6.3 Reverberation Time

After 100-200 msec, reflections typically become dense. The 1st order description of reverberation is energy decay. This is typically measured by the time to 60 dB decay (RT60). Ignoring air absorption,

$$RT60 = \frac{0.163V}{S\bar{\alpha}} \quad in\ meters,$$

$$and\quad RT60 = \frac{0.049V}{S\bar{\alpha}} \quad in\ feet,$$

where S, V, and $\bar{\alpha}$ are the surface area, volume, and average absorption of the room, respectively.

Including air effects,

$$RT60 = \frac{0.049V}{S\bar{\alpha} + 4mV} \quad in\ feet,$$

where m is the air absorption term given in section 11.6.2. This air term typically dominates at very high frequencies, and it is irrelevant for low frequencies.

11.6.4 Energy Growth

For steady state noise starting at time $t = 0$, intensity is

$$I = \frac{W}{S\bar{\alpha}}(1 - e^{-\frac{cS\bar{\alpha}}{4V}t})$$

which corresponds to an exponential growth with time constant $\frac{4V}{cS\bar{\alpha}}$, and a steady state intensity of

$$I = \frac{W}{S\bar{\alpha}},$$

where the noise is generated with power W. As time approaches infinity, the growth reaches steady state.

11.6.5 Effect on Intelligibility

- Energy is larger than without reverberation - walls help to propagate the sound though they may introduce side effects.
- Decay stretches each sound out in time.
- Masks following sounds with decaying energy. Figure 14.2 plots intensity level in a living room as produced by successively spoken syllables *pah, pah, pah, pah*. Notice that the decaying *ah* has more energy than the onset of the next *p* sound. Thus large reverberation can hurt intelligibility.
- Colorations change the spectrum.

11.6.6 Example: 247 Cory

- Dimensions: 34 x 24 x 16 (feet)
- Volume: 13056 cubic feet
- Surface area: 3488 square feet
- Mid-frequency RT60: 2.0 second (in empty room)
- We infer $\bar{\alpha}$ of about 0.1
- Air absorption less than 10% for 1 kHz, about 50% for 4 kHz

For a 10 μW source, the steady state intensity is

$$I = \frac{W}{S\bar{\alpha}} = 0.028\mu W/ft^2 = 0.30\mu W/m^2.$$

This steady state intensity is about 55 dB SPL. At 200 msec after cutoff, the transient intensity level is about 49 dB. Will this interfere? Yes, it does. And Professor Morgan demonstrated it by suddenly lowering his voice, letting the background noise cover up what he was saying, given the long reverberation time.

11.6.7 Early Reflections

Prior to the arrival of a significant energy density, reflections are sparse. These initial sparse delay pattern colors short-term spectrum of the loudest and the earliest sound. The Initial Time Delay Gap (ITDG), from sound onset to the first echo, is a critical parameter. From a survey study, the best concert halls rated by conductors are those having a 15 ms ITDG. It was said that this 15 ms ITDG yields a sense of intimacy for the conductor. (Note that the 15 ms ITDG is an average number for the hall. Many seats will have longer or shorter delay time.)

Figure 13 plots an anechoic (free from echoes and reverberations) direct impulse response. Figure 15 is the direct impulse response of Cory Room 247. Comparing these two graphs, Cory 247 have a 2 msec ITDG. One can often judge the size of a room from its early echo response.

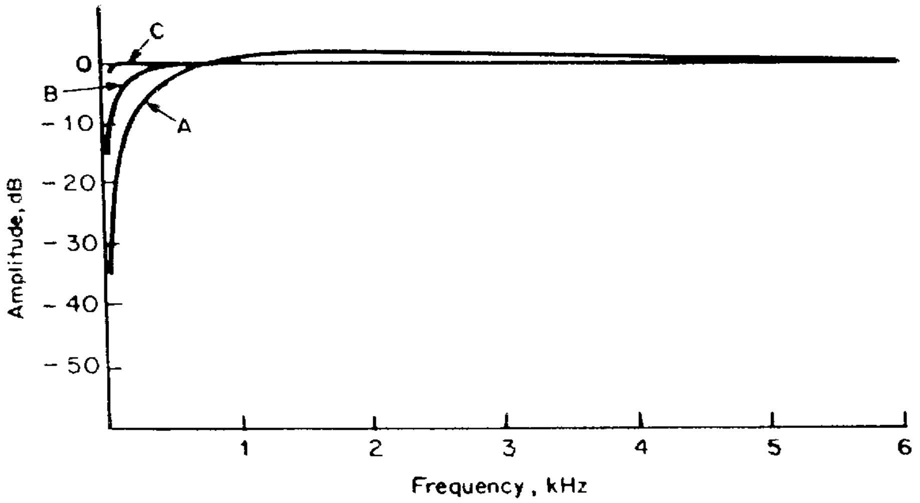


FIG. 6.2. A, B, C weighting curves.

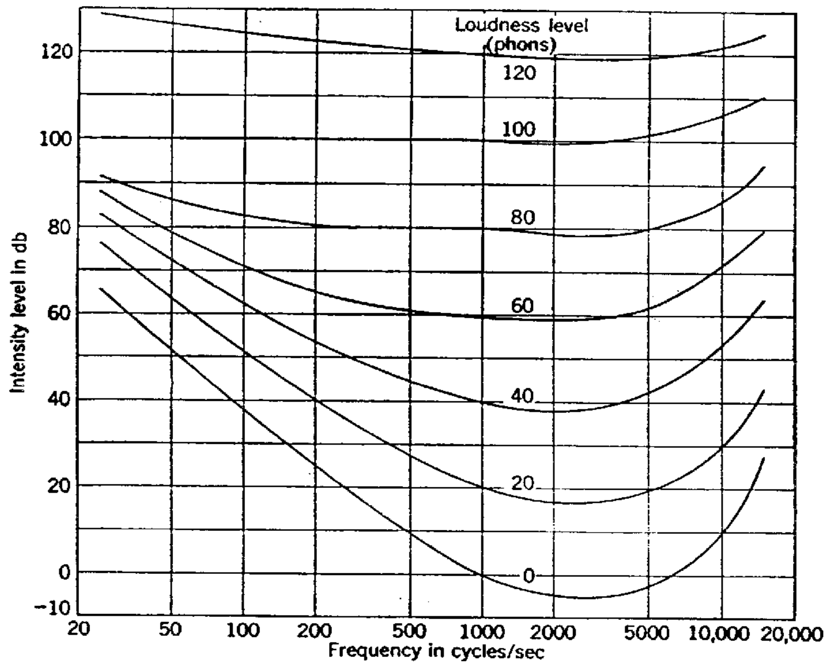
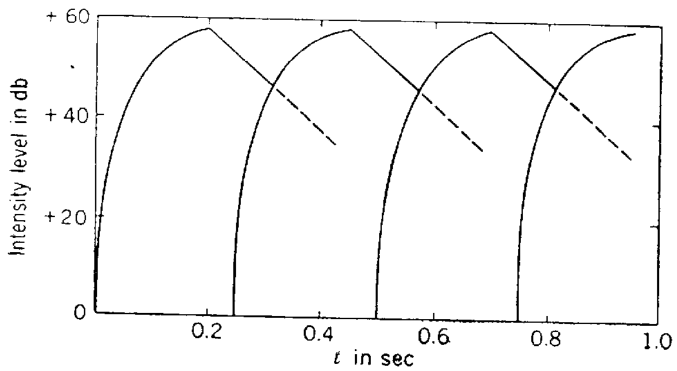


Fig. 13.10. Equal loudness level contours expressed in decibels relative to 10^{-12} watt/m^2.



14.2. Intensity level in a live room as produced by successively spoken syllables.

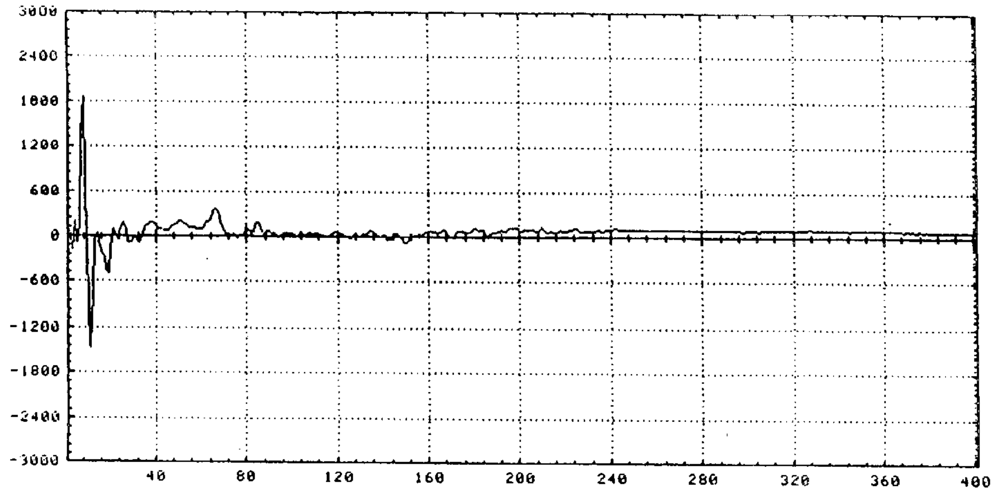


Figure 13: Anechoic direct impulse response, Rogers speaker

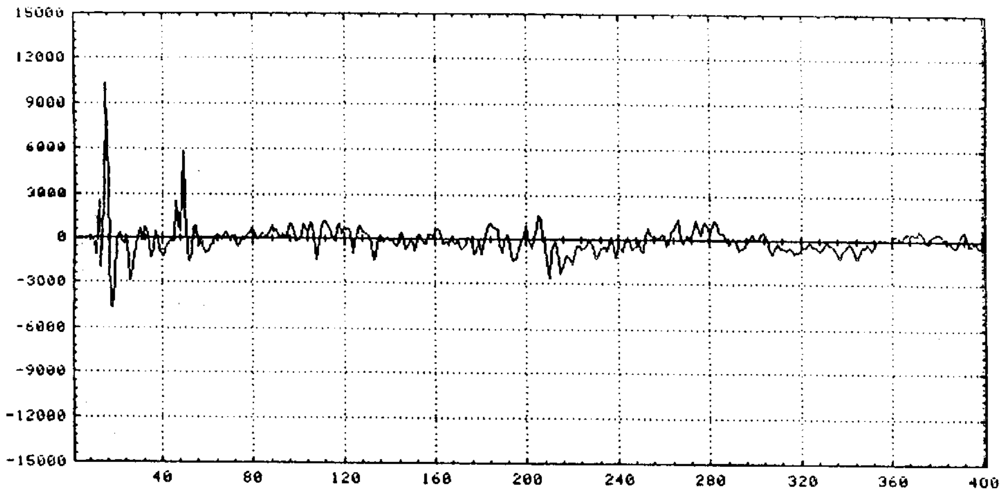


Figure 15: Room 247 direct impulse response, Rogers speaker