SCI220 – Foundations of Musical Acoustics
Cogswell Polytechnical College
Fall 2008

Week 7 – Class Notes

Room Acoustics (FMA)

Chapter 11: Room Acoustics I - Excitation of the Modes and the Transmission of Impulses

A room is a three-dimensional region containing air. The air contained in a room has both elasticity and mass. These two qualities allow for different oscillation modes that have their own vibrational shape and characteristic frequency.

11.1 Sound Pressure: A Way of Describing the Characteristic Oscillatory Modes of Room Air

Microphones function by having their diaphragm pushed back and forth by the air pressure associated with a particular sound and converting this motion into an electrical signal that is a direct measure of the pressure exerted by the air molecules. In a similar fashion, out hearing system operates on signals passed by the motion of the eardrum. In such manner, we can conveniently use sound pressure fluctuations as a way of describing the air oscillations in a room. We can further display these fluctuations as displacement diagrams much in the same way as we have described strings and two-dimensional areas.

11.2 Excitation of Room Modes by a Simple Source

By popping a balloon or setting off a firecracker we can impulsively excite the air modes in a room. For sinusoidal excitation of the room air we need a way to inject and extract air in the room at the desired driving frequency. In this case a loudspeaker proves to be ideal.

The strength of the excitative influence on the air moving within a cavity is labeled as the *source strength*. A *simple source* is one whose aperture is very tiny compared to the displacement between nodal regions of the characteristic room oscillations. The changes in room excitation brought about by widening the source aperture are exactly like that of a string excited by a widened hammer.

We can observe the following assertions:

1. a given oscillatory mode in a room is maximally excited by a simple source if the source is located at one of the points of maximum pressure fluctuation belonging to that mode. There will be no excitation if the source is located at the position of a pressure node.
2. Sinusoidal excitation of a mode by means of a simple source produces a maximum oscillatory pressure amplitude when the excitation has a driving frequency that matches the natural frequency of the mode.
3. Lightly damped room modes, when started and left to ring, will run for many oscillations before the pressure amplitude has fallen appreciably. Such modes respond strongly to the excitation frequency only when excitation frequency closely matches the characteristic frequency of the mode.
4. Each particular room mode responds sinusoidally to some extent regardless of the driving frequency. The resulting oscillation will have the characteristic shape belonging to the mode (after the initial transient) and will oscillate at the driving frequency.
5. All of the transient behavior of room air responding either to sinusoidal driving by a simple source or to impulsive excitation by explosions shows pressure variations which follow the same rules that apply to mechanical oscillation of masses, strings, bars, plates, and membranes.

11.3 Detection of Room Modes by a Microphone or by the Ear: Interchangeability of Source and Detector

It is trivial to state that the amplitude of the microphone's electrical signal will fall to zero when it is located at a pressure node; conversely, the output microphone signal will be at a maximum at a room location where is a pressure maximum corresponding to a mode. This point will be called the pressure antinode.

Our ears behave in similar ways to a microphone and we could expect our hearing to behave in the exact same way as described for the microphone. However, because we are not exciting just one mode of a room but many, we cannot find points of silence in a room or points or marked loudness.

The modes of oscillation characteristic to a system are most favorably excited as precisely the same points as those where the oscillation is most strongly observable. There appears to be a considerable similarity between the efficacy of excitation and the ease of detection of an oscillation. Meaning that the points of excitation by a simple source and of detection by a small microphone are precisely interchangeable. If a simple source, A is located at one point in a room and a small microphone, B is located at another, then the pressure signal at B due to the action of a flow source at A is exactly the same as the pressure signal to be observed at A if the source is moved to B.

11.4 Measured Steady-State Response Curves for a Room

In figure 11.2 (p.178), we can see a graph showing the average number of room modes lying close enough to a given sinusoidal excitation frequency to be excited to more than half their maximum amplitude. In this case, for any excitation frequency above 200Hz several dozen of the room modes are strongly excited by the source.

The total number of active modes at any frequency is proportional to the volume of the room. An increased damping increases $W_{1/2}$ proportionately and thus makes a corresponding change in the number of modes that can be excited.

A response curve displays the randomly varying combination of the strongly excited modes as “heard” by the microphone along with the more smoothly varying contributions from the “tails” of the resonance curves of tens of thousands of other room modes which, despite being weakly excited, make up by sheer numerousness an appreciable part of the measured response. An average curve of the response can be calculated by simply integrating all of the contributions of all of the weakly responding high frequency modes.

The mean spacing of the peaks in the response curve depends on the halving time $T_{1/2}$ for the decay of oscillations. A long decay time is associated with a steady-state pressure response curve that has many peaks over any interval of frequency, shorter decay times are associated with a smaller number of peaks. A uniform response curve over a frequency range shows that the decay time is nearly the same for all modes over the measured frequency range.

11.5 The Influence of Furniture and Moving Objects on Room Modes

The presence of furniture and people in a room will rearrange the frequencies of the various modes and also change the oscillatory pressure distributions which characterize them. Although such alterations are minute in magnitude, the total addition of the thousands of strongly and weakly excited room modes will provide a significant net effect.

1. For any particular frequency of excitation and source location there are microphone positions in a room at which the detected sound pressure is particularly insensitive to the effects arising when objects are moved around in a room. There is generally a strong transmission of sound to the microphone at

such positions, and things are little changed if the frequency is altered slightly.

2. For any given excitation frequency and source location, there are a few microphone positions in the room at which sound pressure is extremely low. These regions are very small. When objects are moved around in the room there are enormous fluctuations in the microphone signal. The position in the room of such points of minimum sound pressure and wild signal fluctuation are considerably displaced is a small change is made in the excitation frequency.

11.6 Room Responses: Some Apparent Problems

It may seem that some of the previous statements will lead us to think that the properties of a room may significantly affect the possibility of a well-defined tone color. Depending on where the listener is located (or by symmetry on where the source is located) the strengths of the various partial components of the tone apparently can have any relationship whatever to one another at the listener's ears.

We are tempted to experiment with an array of microphones set up in a room to detect an “unspoiled” signal. However, this turns out to be false since the electrical signal measured is an addition of sinusoids (from each microphone) that has its individual amplitude and delayed arrival time. We can not just an add up the sinusoids as done in a graphical manner.

11.7 Transient Response of Rooms to Sinusoidal Excitation

The build-up and the decay of sound in a room are dominated by the strongly excited modes, which somehow manage to give a total response that looks like that of a single spring-mass system taken by itself.

We can state the following:

1. The transient behavior observed at any point in a room after the excitation is turned off is very similar to that observed immediately after the source is turned off.
2. The duration of the onset and decay transients is as long, or longer than, what we might expect on the basis of the known halving time for individual room modes.
3. When the excitation frequency is set to a value that makes the microphone signal insensitive to the position of objects moving around the room, the onset and turn-off transients have a form similar to the build-up and decay of a single mode acting by itself.
4. When the driving frequency is set to produce maximum sensitivity of the microphone signal to rearrangements of furniture, the onset and turn-off transients are in the form of irregular bursts of sound.

5. At the driving frequency other than of statements 3 and 4, the transient behavior is intermediate between the two extreme forms.
6. Moving the source and the microphone to new positions in the room causes a total rearrangement of the frequencies at which the various forms of transient and steady-state response take place.

11.8 Response to Impulsive Excitation I: Signal Delays and Reverberation

When a microphone is moved farther away from the source it will detect a signal that has been affected by the averaged room response and will take a longer time to "detect" the signal. This measured signal will be significantly different from the original signal.

However, the speed in which the direct signal travels will be the same as the reflected signals regardless of the excitation type.

If many observations of the decay of sound in many points in a room, it is possible to construct a room average decay curve for the sound amplitude. This curve will show a definite halving time behavior.

The reverberation time is defined as the time required for the averaged sound pressure in a room to die down one-thousandth of its initial amplitude. The reverberation time T_{rev} is 9.97 times the halving time $T_{1/2}$, and follows that the bandwidth $W_{1/2}$ over which a given mode is strongly excited can be calculated by:

$$W_{1/2} = (3.8/T_{rev})\text{Hz}$$

11.9 Response to Impulsive Excitation II: Reflections and Scattering

For a delayed signal, the time delay between the first and second appearances of the disturbance is simply the time required for sound to make one round trip from the microphone to the wall and back. The pressure amplitude of the second pulse is reduced relatively to the first because, 1) it has traveled an additional distance in making the round trip , and 2) because of the dissipative effects produced during the reflection process.

The shape of a reflected pulse, from a perfectly flat and reflective wall, matches that of the direct sound. The eventual complication in the shape of the pressure disturbance arise from the adding up of successively delayed overlapping replicas of the original impulse.

For irregular walls, or rooms with objects in it, the reflected sound will be scattered and will produce a signal which shape is not similar to the direct signal. Small objects can have a considerable effect on sound scattering.

Chapter 12: Room Acoustics II – The Listener and the Room

An *anechoic chamber* is a room fitted with arrays of rock-wool wedges arranged to prevent any echoes from bouncing off the walls, floor, and ceiling.

12.1 Hearing Sustained Sounds in a Room

Any instrument played at a given loudness will act as a well-defined source. The source strength of each partial are definite, and the ratio between the amplitude of the first harmonic partial and the subsequent partials are also well defined and reproducible. Meaning that the excitation acting on the room has a stable and definite nature.

Any point in the room where the ear might find itself has its own relationship between source strength and measured sound pressure, this relationship being different for every frequency component that may be emitted by the source. The vibration recipe that relates the amplitudes of the partials of a given instrument's tone at the listener's ears will differ from point to point in the room.

Our nervous system is able to operate in many ways simultaneously. One thing it can operate in is make a preliminary assessment of the strength of the partials of a musical tone by forming an averaged amplitude measurement based upon the different signals received from both ears.

If the sound field in a room is explored by means of a slowly moving microphone which uses a fixed frequency, we can see that in the trace chart of different excitation frequencies high frequency excitation is associated with closely spaced random wiggles in the sound pressure trace. For low frequencies, on the other hand, the microphone must move much farther to go from one part of the fluctuation to the next.

Half wavelength connotes the width of a hump in a room mode, and if we could excite on of these modes by itself, the travels of our microphone would give us a smooth varying trace on the chart, with each hump following the next in proper order. In an actual room, however, the superposition of contributions from the thousands of modes leads to great irregularity. A look in the chart would allow us to predict only a few steps ahead. The relationships between two points a half wavelength or more apart is totally random: there is no chance to predict the sound pressure at one of these points on the basis of what we find at the other point. For points closer that half a wavelength is it possible to make a rough prediction. An approximate measurement of the human head to find the distance between our ears may be used along with the half wavelength number to show that only at frequencies above 1000Hz our ears are spaced enough to get independent views of the room sound for averaging purposes.

Furthermore, our nervous system not only able to make running averages of the signals at the two ears for determining the various partial strengths, but also to pile up information from both ears over a short period of time and average all of it, so as to exploit fluctuations arising from moving objects in a room. Also, the human body is of such a size that ordinary small motions of a listener or player are sufficient to provide appreciable help to our aural averaging mechanism (for steady sounds only with frequency components above 500 Hz).

12.2 The Role of Early Reflections: The Precedence Effect

Low frequency partials are sent out with equal strength at all directions. Each of our ears is supplied with the direct sound within a few milliseconds, and is also supplied with the first six echoes off the wall, floor, and ceiling of the room. The pressure disturbances reflected from these surfaces will not be altered.

Higher frequency partials from any instrument do not excite the room modes in the manner of a simple source. The sound will be sent out in a more focused manner in the direction where the instrument is pointed at (say the bell). If the instrument is pointed at us, the direct sound from the higher partials will be heard stronger than the reflected sounds. Under these conditions we can get an aggregate impression of the sound that is somewhat different from what is received by the rest of the listener's in a room.

The precedence or Haas effect is described by the following:

1. the ear will combine a set of reduplicated sound sequences and hear them as though they were a single entity, provided that a) they all arrive within less than 35ms of the arrival of the direct sound and b) the sound pressure recipes of all members of the set are sufficiently alike.
2. The singly perceived entity is heard as though all of the later arrivals are piled upon the first one without any delay. Meaning that the perceived time of arrival is the set is that belonging to the direct sound and that its loudness is augmented.
3. The apparent position of the source of the perceived sound will coincides with the position of the source.
4. The functioning of the precedence effect persists even when the later arrivals in the set of sounds have pressure amplitudes that are larger than that of the first signal, provided that these are no more than 3 times the first sound.

12.3 Localization by the Ears of Sound Sources in a Room

When we listen to music in a concert hall or have a conversation in a noisy room, it is possible for us to focus our attention on particular sources of sound with high accuracy. This can be done much more easily in a room than outdoors because we can get many clues regarding the position of the source and its actions by the way the sounds are reflected and scattered by the room.

The simplest clues to the position of a musical source to be expected by a monaural listener are the time relationship between the arrival of the direct sound and the first dozen echoes. The spreading of sound around the listener's head due to scattering provides additional directional clues. The amplitudes and shapes of these pressure signals depend on the direction from which they arrive to the head.

The listener's head scatters the sound in such a way that it alters the shape of the impulse, thus the pressure distributions at various points around the head is different for high frequency signal components than for low frequency components.

The experience of listening though headphones as opposed to a “live” listening context can be for some an uncomfortable and frustrating experience. This is due to the reduced ability to hear subtle details that arise from the sound projected and its interaction with the room.

When a sound comes to us, we make a quick and rough determination of the position of the source and of the kind of sound being generated, based on our determination of the first-arriving train of signals. We then make use of the precedence effect in taking the later arriving part of the sound during a 35ms interval to reconfirm and elaborate a picture of what is happening. While one part of our process melts signals together over this interval of time, the other parts are doing fine frequency and time analysis that extract information that is being put together.

When the delay is 60ms or more, the conflict between what the the subject hears and what he is trying to say makes it impossible to talk properly. Meaning that the echoes perceived at this time interval will sound clearly distinguishable and will become a distraction and hindrance to understanding.

Appendix

Room Mode Calculation

Axial Modes - Involve two parallel surfaces - opposite parallel walls, or the floor and ceiling. These are the strongest modes.

Tangential Modes - Involve two sets of parallel surfaces - all four walls, or two walls the ceiling and the floor. These are about half as strong (energy) as the axial modes (−3 dB).

Tangential Modes - Involve two sets of parallel surfaces - all four walls, or two walls the ceiling and the floor. These are about half as strong (energy) as the axial modes (−3 dB).

To calculate the frequencies of the axial, oblique and tangential modes, we use the following formula:

$$f=(c/2)\sqrt{((nx/L)^2+(ny/W)^2+(nz/H)^2)}$$

f = Frequency of the mode in Hz
c = Speed of sound 343 m/s at 20°C (68°F)
nx = Order of the mode of the room length
ny = Order of the mode of the room width
nz = Order of the mode of the room height
L, W, H = Length, width, and height of the room in meters

General Criteria for Room Acoustics

1. *Clarity* – each note should arrive cleanly, crisply, and unobscured. This is specially important for speech in which word intelligibility is important.
2. *Uniformity* – listener's in all parts of the hall should hear nearly the same sound and there should be no dead spots.
3. *Envelopment* – the listener should not feel separated from the source but rather bathed in sound from all sides; yet the sound source position should be identifiable by both vision and hearing.
4. *Freedom from echo* – no reflections from the walls should be perceived as separate echoes. All reflections should blend together smoothly.
5. *Reverberation* – the continuation of the sound after the source stops should have an appropriate loudness relative to the loudness of the original sound and should

have a pleasing rate.

6. *Performer satisfaction* – the stage must be free from distracting echoes and have enough enclosure to provide a sense of good communication between the performers in a group.
7. *Freedom from noise* – outside traffic noise and internal noise (from AC system) should be kept to a minimum.

Reverberation Calculation

Remember: the reverberation time T_{rev} is the time in which a decaying sound level drops 60dB below its original level.

$$T_{rev} = (0.16V)/ S_e$$

where V is the room volume in m^3 and S_e is the effective absorption area in m^2. Furthermore, S_e can be calculated in the following manner:

$$S_e = \sum(\alpha n Sn) = \alpha 1\, S1 + \alpha 2\, S2 + \alpha 3\, S3 + ...$$

where α is the absorption coefficient and S is the surface area.

Reverberant Sound Levels

Reverberation sound intensity (or level) is calculated by

$$I_{rev} = (36PT_{rev})/V$$

where P is the source power and V is room volume.

Spatial Perception

Sound localization can be done by the following mechanisms:

1. Interaural intensity differences: directionality of high frequencies vs. omnidirectional nature of low frequencies.
2. Sound arrival time differences from one ear to the other.
3. Phase differences (works best for frequencies below 1500Hz).
4. Outer ear shape enhances directionality of high frequencies and produces a difference in strength when source changes from front to back.