Acoustics of Small Rooms, Home Listening Rooms, Recording Studios

The acoustical properties of small rooms – used *e.g.* for listening to music, watching movies, *etc*. in a person's home, or *e.g.* sound recording studios – differs considerably from that of large rooms – auditoriums, concert halls, cathedrals, lecture halls, *etc*. primarily in the reverberation times (typically T_{60} < ½ second for a small room) and also room resonances. The "mix" of direct sound *vs*. early reflected sound vs. reverberant sound is different for small *vs*. large rooms. In a large room, first-arrival times of the early reflected sound are typically on the order of ~ 50-80 ms after the direct sound, whereas for small rooms, the first-arrival times of the early reflected sound are typically on the order of ~ few *ms* after the direct sound. Additionally, and especially so in home environments, the sound absorption properties of the room often are significantly higher than in large rooms, due to the presence of carpeting on the floor, window curtains on walls, *etc*. Thus, the acoustic "intimacy" of the small room often makes it difficult to emulate the acoustics associated with that of a larger space, *e.g.* when listening to recorded music.

The Sabine formula $T_{60} = 0.161V/A$ holds for small rooms, and shows that for fixed small volume *V*, the reverberation time can be increased by reducing the absorption *A* of the room. However, if you've ever been in an empty room in a house, *e.g.* with no carpeting or drapes/curtains present, because of the short first-arrival times associated with a small room, the reverberant properties of a small empty room are starkly different than that of an auditorium. The short *vs*. long decay time associated with sound in small *vs*. large rooms provides important auditory information/clues to the listener about the size and nature of the room.

For rectangular rooms, the eigen-frequencies associated with the axial, tangential and oblique-mode room resonances $f_{lmn} = \frac{1}{2}v\sqrt{(l/L_x)^2 + (m/L_y)^2 + (n/L_z)^2}$ with $l, m, n = 0, 1, 2, 3, 4, 5, ...$ will also be commensurately higher than those associated with an auditorium, also contributing to the perceived acoustic differences between small *vs*. large rooms. The higher-frequency room resonances accompanying small *vs*. large rooms thus "color" the sound of recorded music being listened to in a small room differently than *e.g.* in an auditorium-type live-sound environment.

Most listeners in a small room will likely be situated such that they are ~ 2 *m* or more away from loudspeakers located in the small room. At low frequencies, the directivity factor *Q* of loudspeakers is reduced {due to diffraction effects} and the room absorption, *A* is typically low in small rooms at low frequencies {*e.g.* carpeting absorbs sound relatively poorly at low frequencies}, thus a listener in a small room is often in the reverberant field of the room at low frequencies, *i.e.* typically $Q/4\pi r^2 < 4/A$ at low frequencies. At higher frequencies, the directivity factor *Q* of the loudspeakers increase as well as the sound absorption *A* of the room such that for a typical listening distance of *r* ~ 3-4 *m*, $Q/4\pi r^2 > 4/A$ at higher frequencies, further contributing to the listener's perception that small rooms are "dead"-sounding, relative to large auditorium/concert halls, *etc*.

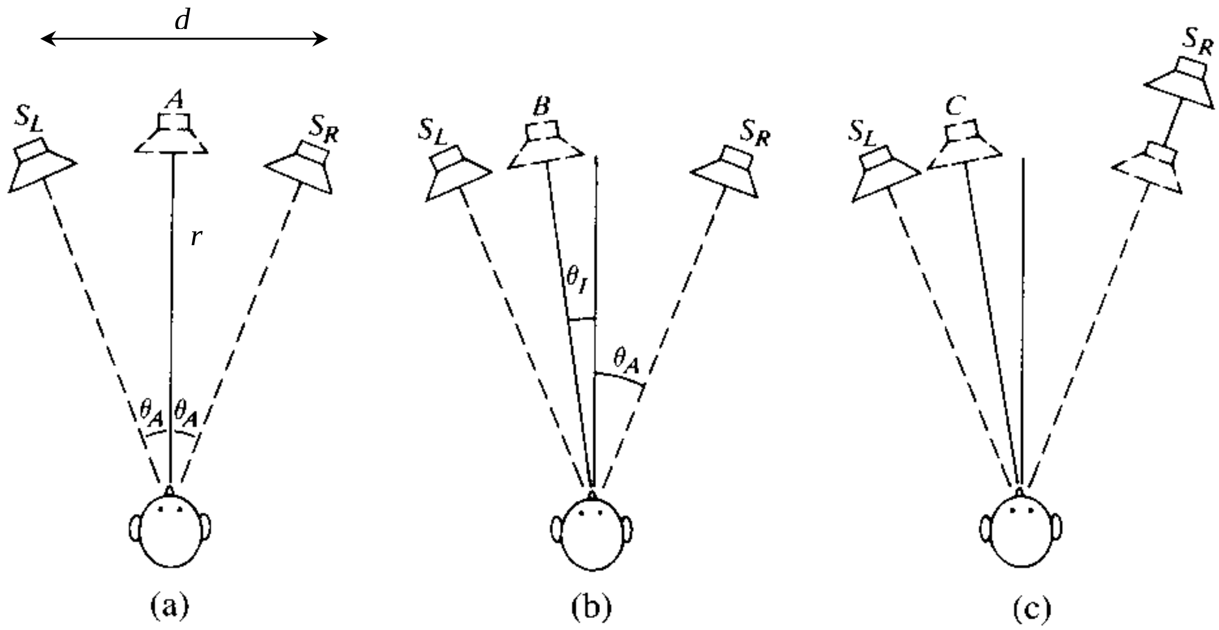
In a small listening room such as in a house, an audiophile likely enjoys listening to music recorded in stereo (*i.e.* L & R-channel sound), or perhaps a enjoys watching a movie, or recordings of live music *e.g.* on a DVD with the 5.1 surround-sound – *i.e.* requiring a multiple channel/multiple speaker home theater sound system.

As discussed in previous P406 lecture notes on human hearing, at low frequencies ($100 < f < 1500$ *Hz*) our main clue to the direction/location of a sound source is the inter-aural time difference – *i.e.* the difference in arrival times/phase information at our two ears, whereas at higher frequencies, the inter-aural intensity difference (*IID*) dominates our ability to localize high-frequency sounds. Below $f \sim 100$ *Hz*, we have increasing difficulty in localizing sounds {a consequence of which *e.g.* is that only a single sub-woofer is needed in the 5.1 surround sound scheme for low frequencies}.

Before launching into a discussion of high-fidelity stereo and/or 5.1 surround sound systems, we first discuss some aspects of how humans, with their binaural hearing and neural sound-processing networks perceive sounds from two or more sound sources…

<u>Human Perception of Sound From Two Loudspeakers, Fed by a Monophonic Signal:</u>

A listener located at a distance *r* on the median plane equidistant from two identical loudspeakers separated a transverse distance *d* apart from each other, and fed by a common {monophonic} signal, perceives a sound "image" located on the median plane, at location *A*, as shown in diagram (a) of the figure below:



If instead the signal strengths of the two speakers are <u>not</u> equal – *e.g* the left speaker's signal is louder than that from the right speaker, the sound "image" in the mind of the listener will shift towards the louder (left) speaker, *e.g.* to location *B* as shown above in diagram (b). The angle θ_I of the sound "image" shift with respect to the median plane can be calculated from the equation:

$$\sin\theta_I = \left(\frac{p_L(\vec{r}) - p_R(\vec{r})}{p_L(\vec{r}) + p_R(\vec{r})}\right)\sin\theta_A = \left(\frac{p_L(\vec{r}) - p_R(\vec{r})}{p_L(\vec{r}) + p_R(\vec{r})}\right)\frac{d/2}{\sqrt{r^2 + (d/2)^2}}$$

where $p_L(\vec{r})\left(p_R(\vec{r})\right)$ are the over-pressure amplitudes associated with the sounds coming from the left (right) loudspeakers, respectively, evaluated at the listener's position $\vec{r}$.

Similarly, if instead, the phase of a sine-wave signal output from one of the loudspeakers is shifted relative to the other, the sound "image" in the mind of the listener will shift toward the speaker that is ahead/leading in phase (modulo 2π), as we demonstrated in the P406 POM lectures a while back, for the phenomenon of consonance/dissonance. See also *e.g.* Matt Gilson's Fall Semester, 2000 P406 Final Project Report, posted on the P406 website at:

http://courses.physics.illinois.edu/phys406/406pom_student_projects_fall00.html

If one of the two loudspeakers is instead *e.g.* moved to a greater radial distance away from the listener, the sound "image" also moves toward the nearer sound source, as shown above in diagram (c). If the RHS source is farther way by more than ~ 1/3 *m*, corresponding to an arrival time difference, $\Delta t > 1\,ms$, the sound image then coincides with the LHS sound source (S_L). However, if the sound from the RHS sound source (S_R) is made louder than that from the LHS sound source (S_L), to compensate for it being farther away, then the sound "image" moves back towards the median plane.

Thus, it is possible to trade of pressure amplitude/*SPL* for time delay/phase information, within certain limits.

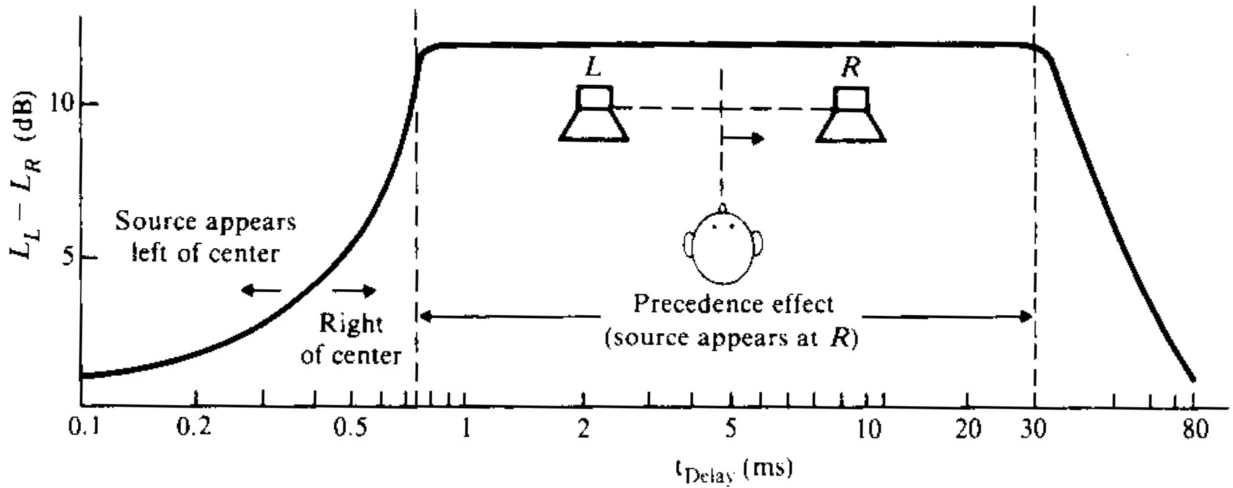
The extent to which trade-off of pressure amplitude *vs*. time delay/phase information works, and the so-called ***sound trading ratio***, R_T defined as the difference in arrival time divided by the equivalent difference in *SPL*, at this time have not been completely/fully-established, however experimental results obtained thus far seem to indicate that the sound trading ratio R_T is clearly frequency-dependent, since our ability to localize very low frequency sounds ($f < 100$ *Hz*) is increasingly poor, whereas the $100 < f < 1500$ *Hz* region it is mainly due to arrival time differences (or relative phase information) of the sound at our two ears in, whereas the intensity / loudness / sound pressure level difference dominates at high frequencies.

Note that this trade-off is also not perfectly complete/equivalent, in that, although at low- and mid-frequencies, a large fraction of the sound "image" shift due a change in distance from a sound source can be compensated by a change in loudness/*SPL*, the sound "image" {position *C* in diagram (c) above} apparently cannot be completely/perfectly restored to the median plane.

Disagreement currently exists between trading ratio experiment results. At low frequencies, *e.g.* $f \sim 200$ *Hz*, trading ratios ranging from $R_T\left(f \sim 200Hz\right) \sim 60 - 150\,\mu s/dB$ have been reported, whereas at $f \sim 500$ *Hz*, trading ratios ranging from $R_T\left(f \sim 500Hz\right) \sim 10 - 200\,\mu s/dB$ have been reported…

Experiments to investigate the nature of sound localization of human hearing can easily be carried out using *e.g.* a home stereo sound system and providing a common signal, *e.g.* output from a signal function generator to the sound system, adjusting the stereo balance control to try to compensate *e.g.* for changing the listener's position from his/her nominal mid-plane location, or to compensate *e.g.* for moving one speaker away from its nominal symmetry position. These experiments can also be carried out as a function of frequency, loudness level, …

The figure below shows the time/*SPL* difference trading and also the approximate range of time and SPL differences over which the *precedence* {*aka Haas*} *effect* applies.



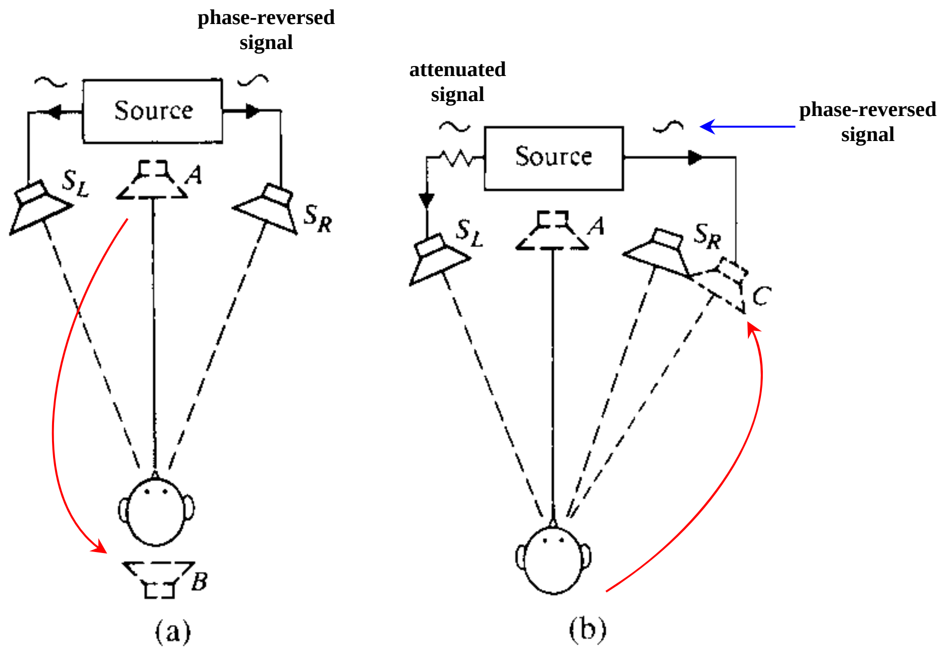
The horizontal axis is the time delay *e.g.* of a pulse to the LHS loudspeaker is delayed relative to the RHS loudspeaker. The vertical axis is the *L–R* difference in sound pressure level (*SPL*) due to the SPL of the LHS loudspeaker exceeding that from the RHS loudspeaker.

The ascending curve on the left-hand side of the graph indicates the approximate combinations of *L–R* time delay *vs*. *L–R SPL* difference that will center the source image at the median plane of the loudspeakers. Note that when $\Delta SPL(L\text{-}R) \equiv L_p^L - L_p^R \gtrsim 15\,dB$, it is impossible to compensate completely with <u>*any*</u> time delay, $t_{Delay}(L\text{-}R)$. Conversely, note that when the time delay $t_{Delay}(L\text{-}R) \gtrsim 1\,ms$, the <u>*precedence*</u> <u>*effect*</u> defeats time/intensity trading.

{A fun experiment that demonstrates the nature of the precedence effect associated with human hearing is for two people to go into a small, but "live" (*i.e.* highly-reverberant) room, close the door and have one person, located somewhere in the room close their eyes, while the other person walks slowly around the room, occasionally clapping his/her hands (once), to launch a sharp, short sound impulse into the room. The listener will have no problem localizing the sounds, due to his/her brain's ability to discern/process the inter-aural time difference sound information. If *e.g.* a single, continuous, sine-wave single-frequency type sound source is instead used, the listener will have a great deal of difficulty localizing such a sound source.}

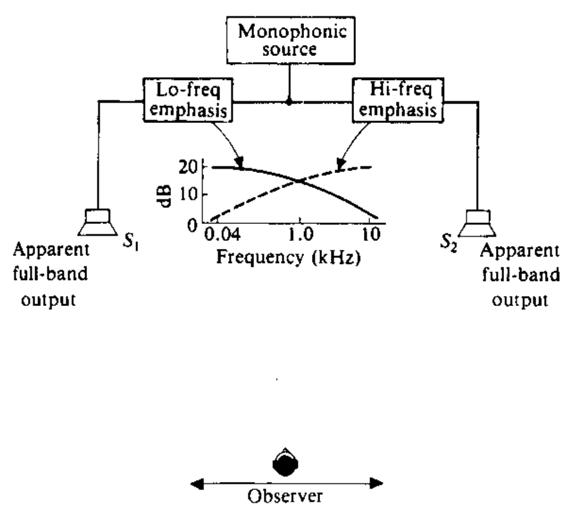
Note that a listener seated only a distance of ~ 1 *ft* (~ 0.3 *m*) closer to the RHS loudspeaker would experience such a $t_{Delay}(L\text{-}R) \sim 1\,ms$ time delay – which has the unfortunate consequence that the stereophonic effect works best only within a limited region of space known as the "sweet spot", due to the constructive interference of the sound waves output from the two loudspeakers with each other.

If the <u>phase</u> <u>polarity</u> of one of the loudspeakers is <u>reversed</u> by 180^o, if the listener is located on the median plane (as before) the sound "image" will move from position *A* at the nominal median plane to a position *B*, which can be inside, or in back of the listener's head, as shown in diagram (a) of the figure below:



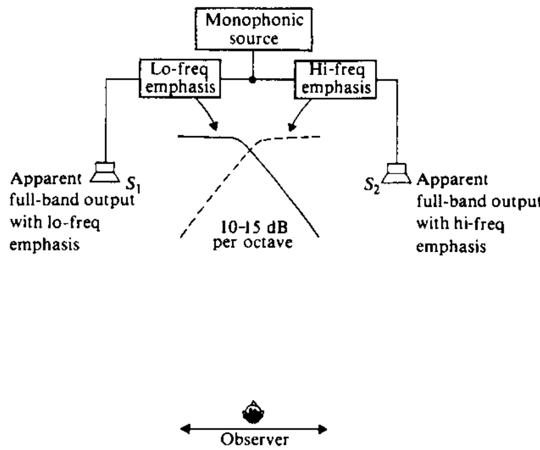
If <u>additionally</u>, *e.g.* the sound from the <u>LHS</u> loudspeaker signal is <u>reduced/attenuated</u> sufficiently, then the location of the sound "image" may shift from position *B* to position *C*, beyond/outboard of the RHS loudspeaker, as shown in diagram (b) of the above figure.

Instead of providing a common/identical single-frequency sine-wave type signal to the two loudspeakers (although with different signal strengths and/or phases), if a <u>spectrum</u> of common frequencies, but with different spectral <u>emphasis</u> is input to the *L* vs. *R* channels, then the sound "image" will appear to be <u>spatially</u> <u>broadened</u>/<u>wider</u>. For example, if the RHS channel is given a <u>slowly</u>-varying high frequency emphasis, while the LHS channel is given a <u>slowly</u>-varying low-frequency emphasis, as shown in diagram (a) of the figure below, the sounds from <u>both</u> speakers will appear to the listener to have a <u>flat</u> frequency spectrum, but the sound "image" will appear to be <u>spatially</u> <u>broadened</u>/<u>wider</u> than that associated with inputting a flat-frequency spectra into both speakers! Furthermore, the listener can shift away from the median position without losing this auditory effect.



(a)

If however, the crossover between low-frequency/high-frequency emphasis between the LHS *vs*. RHS loudspeakers is abrupt (*i.e*. a roll-off of > 10-15 *dB*/octave), as shown in diagram (b) in the figure below, there will be a noticeable difference in the timbre of the two sound sources – *i.e*. while both speakers will have the same ***apparent*** output/loudness level, the LHS speaker will have noticeable low-frequency emphasis, whereas the RHS speaker will have noticeable high-frequency emphasis.



(b)

Other differences between two sound sources, other than the spectrum shape are also found to broaden the sound "image" perceived by a human listener. One effective way to achieve sound "image" broadening is to add reverberation to one of the two sound sources (*e.g*. the left channel), but not to the other source (!)

Thus, three important properties of sounds heard from multiple speakers, strongly influenced by differences in loudness level/SPL, arrival time difference/relative phase, spectral differences and asymmetry in reverberant sound are:

(1) The degree of fusion of the two separate sounds into a single sound "image"
(2) The broadening of the fused sound "image"
(3) The spatial displacement of the fused sound "image"

<u>High-Fidelity Sound</u>:

An ideal high-fidelity sound system, independent of the nature/scheme/type of such a sound system, should have the following six sonic attributes:

> (1) The frequency range (*i.e.*) bandwidth of the sound system should be able to faithfully/accurately reproduce <u>all</u> of the original frequency components in the original recorded sound; the sound spectrum of the reproduced sound should be identical to that of the original recorded sound.
>
> (2) The reproduced sound should be free (insofar as possible) of distortion, inter-modulation distortion and/or transient distortion, as well as any/all types of noise.
>
> (3) The reproduced sound should have loudness and dynamic range equivalent to that of the original recorded sound.
>
> (4) The reproduced sound should not unduly introduce any significant frequency-dependent <u>phase</u> <u>shifts</u> that are not present in the original recorded sound.
>
> (5) The spatial sound pattern/sound "image" of the original sound should be faithfully reproduced.
>
> (6) The frequency-dependent, spatial and temporal reverberation characteristics of the original sound should be faithfully preserved in the reproduced sound.

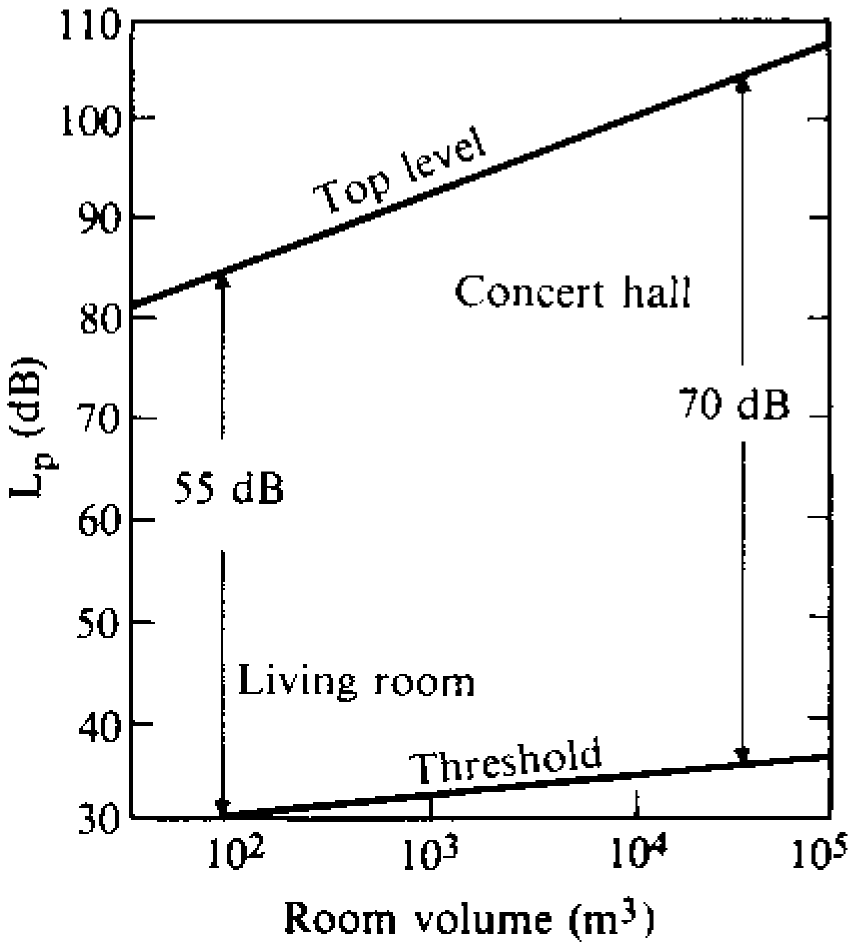
No real sound system exists that perfectly/fully simultaneously satisfies all six of these sonic attributes – the devil is (always) in the details of everything – *i.e.* all pieces of equipment in the signal path, from the medium that was used for recording the original sound and its accompanying transducer, if any (*e.g.* vinyl LP's, magnetic tape, …), the preamplifier, frequency equalization (*i.e.* tone controls and or graphic equalizer), the power amplifier(s), the loudspeakers (their crossover networks, if any), the details of the design of the speaker enclosures, the placement of the loudspeaker enclosures in the room, the details of the room acoustics and finally, the location of the listener.

Many modern high fidelity sound systems have enough power to reproduce the peak sound levels *e.g.* heard in an actual concert hall, *i.e.* around ~ 100 *dB* and more. However, a home sound system that outputs sound pressure levels of ~ 85 *dB* will in fact sound quite loud in the smaller listening room of a house, as compared to a voluminous concert hall.

Modern high-fidelity sound systems quote very low distortion, inter-modulation and transient distortion figures, as well as impressive signal-to-noise figures, compared those of hi-fi sound equipment manufactured only a couple-few decades ago, when they are operated within their rated output.

The dynamic range in a listening room is limited by the so-called <u>tolerable</u> <u>top</u> <u>level</u> and by the threshold that can be heard above background noise levels, which in a home listening environment may be around ~ 25–30 *dB* *vs*. ~ 30–35 *dB* in a concert hall.

The figure shown below gives a general indication of the loudness and dynamic range at which music may be heard in rooms of various sizes.



It can be seen that a listening room in a home typically has ~ 55 *dB* of listenable/tolerable dynamic range *vs*. ~ 70 *dB* of dynamic range for a concert hall. If the sound pressure level exceeds the top level curve at any point, the (average) listener's response is "it's too loud"… The threshold curve is associated with the minimum adequate signal-to-noise levels associated with average/typical listening rooms of varying room volume *V*.

Not all tone controls/graphic equalizers have acceptable phase-shift attributes at their band-edges, and transient response {"you gitz what you payz for"}. Similarly, 2-way/3-way/4-way loudspeaker sound enclosures with passive cross-over networks may also have unacceptable phase-shifts and transient response at the cross-over frequency points.

Attempts to improve the ambience or spatial-temporal characteristics of reproduced sound in small listening rooms have led to the development of a variety of room expanders, stereophonic spreaders and shifters, *etc*. These are often ignored by hi-fi sound enthusiasts/audiophiles…

<u>Single, Stereo and Multi-Channel Sound Reproducing Systems</u>:

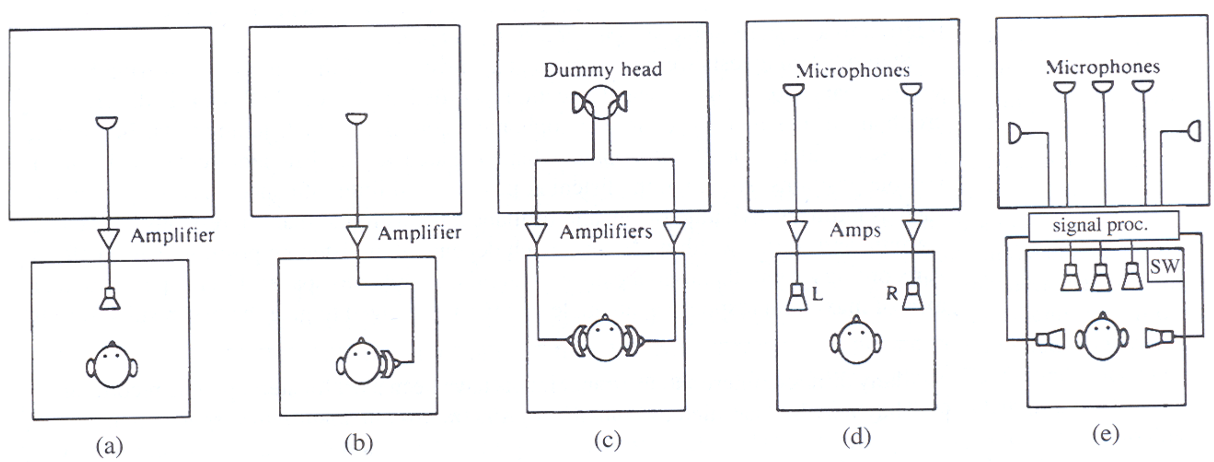
A ***<u>monophonic</u>*** sound system consists of recording sound/music with a single microphone, playing it back with a single amplifier and a single loudspeaker in a listening room, as shown in diagram (a) of the figure below. Ambience is provided by the acoustic characteristics of the listening room. The performance of a high-quality monophonic sound system in a listening room with excellent room acoustics should <u>not</u> be underestimated, however, these days not many good monophonic sound systems exist any longer, much less monophonic recordings….

A ***<u>monaural</u>*** sound system differs from a monophonic sound system in that sound is fed to only one ear (via an earphone) of a listener, as shown in diagram (b) of the figure below. This type of sound system is used primarily *e.g.* in telecommunications and *e.g.* psycho-acoustics experiments. It is most definitely not a hi-fi sound system.

A ***<u>binaural</u>*** sound system simulates human hearing and uses two identical microphones installed at the ear-locations of a dummy human head, the signals from which are independently amplified and heard by the listener via stereo headphones, as shown in diagram (c) of the figure below. One disadvantage of earlier binaural sound systems is that the dummy head cannot be rotated, whereas a human head can do so, thus binaural recordings tend to sound as if the sound source is <u>inside</u> the listener's head, rather than coming from outside. Modern/state-of-the-art virtual-reality type binaural sound systems can compensate for head movement using head-tracking devices. Listening to stereophonic-recorded music with stereo headphones tends to produce a greatly-exaggerated stereo effect that is interesting, but not realistic – the sound source "image" usually appears to be inside, or above the head, which is not true binaural sound reproduction, because most likely the microphones used in the stereophonic recording were not positioned at/in the ears of a dummy head.

A ***<u>stereophonic</u>*** sound system uses independently-amplified signals recorded from two identical microphones fed to two *L*/*R* loudspeakers in the listening room, as shown in diagram (d) in the figure below. We will discuss stereophonic sound systems more below.

A ***<u>surround-sound</u>*** system uses independently amplified signals recorded from multiple microphones fed to multiple speakers in the listening room, as shown in diagram (e) in the figure below. We will discuss surround-sound systems more below.



<u>Stereophonic Sound Systems</u>:

Today, the stereophonic (*aka* "2.0") sound reproduction system is the most popular, and perhaps also the most successful spatial-temporal hi-fi sound reproduction system. There are many ways that have been developed over the years for the <u>*recording*</u> of stereophonic sound. Early experiments with the recording of stereophonic sound took place in the 1930's, carried out by Harvey Fletcher and colleagues at Bell Labs in the U.S. and by Alan Blumlein and colleagues at EMI (Electric and Musical Industries, Ltd) in England. The BBC broadcast the first stereophonic recording in December, 1925. Walt Disney Studios ***Fantasia*** (1940) was the first commercial motion picture to have stereophonic sound. We list/describe various microphone arrangements that have been developed and used for stereophonic sound recording:

A.) The ***X-Y system*** uses a coincident pair of cardioid-pattern pressure microphones with their symmetry axes at an opening angle of 135^o.

B.) The ***Stereophonic system***, or ***Blumlein pair*** uses a coincident pair of bi-directional, figure-of-eight pattern {*i.e.* ***differential*** pressure} microphones with their symmetry axes at an opening angle of 90^o.

C.) The ***MS* (*midside*) *system*** uses a coincident pair consisting of one forward-pointing cardioid pressure microphone and a sideways pointing bi-directional/figure-of-eight/***differential*** pressure microphone. The ***sum*** and ***difference*** of the signals output from these two types of microphones are recorded as the right and left stereo channels.

D.) The ***ORTF system*** uses a nearly-coincident pair of cardioid pressure microphones spaced ~ 16.5 *cm* apart (the same distance as the typical/average human ear separation distance) with their symmetry axes at an opening angle of 110^o. ORTF is the French broadcasting system. In a stereophonic recording/listening test held several years ago, the *ORTF* system was judged to give the best results overall.

E.) The ***NOS system*** uses a nearly-coincident pair of cardioid pressure microphones spaced ~ 30 *cm* apart with their symmetry axes at an opening angle of 90^o. NOS is the Dutch broadcasting system.

F.) The ***A-B***, or ***Spaced-Pair system*** uses a pair of microphones spaced several feet apart. The microphones can have any response pattern, however omni-directional pressure microphones seem to be the most popular. Note that if the microphone spacing is too great, it tends to give an exaggerated stereo effect, which increases with increasing mic separation.{The ***Painkillers*** record their live performances using this method, courtesy of *A-Roosta Records* →}

G.) The ***OSS* (*Optimal Signal Stereo*)**, or ***Baffled-Pair system*** uses two omni-directional pressure microphones separated by ~ 36 *cm* with a disk-shaped baffle in between them. The disk-shaped baffle creates a sound shadow, shielding each microphone from the other, and is often a hard disk ~ 35 cm in diameter with ~ 1 cm sound absorbent foam on each side of the disk, also known as the Jecklin disk, named after Jürg Jecklin {one-time chief sound engineer of Swiss radio}. The two omni-directional microphones are spaced ~ 36 *cm* apart (~ 2× the human ear separation distance) with the symmetry axes of the microphones parallel to each other, and parallel to the plane of the disk.

In another ***OSS*** stereophonic recording scheme, the two omni-directional microphones are separated only by ~ 10 *cm*/few inches from each other, but are angled away from each other, each at an angle of ~ 20^o from the plane of the disk, as shown in the figure below:

The specific arrangement of microphones used for stereophonic recording of sound/music is always a matter of taste – "*beauty is in the ear of the beholder*". For some people, the coincident microphone techniques sound dry and/or analytical – *i.e*. "too correct". However, a sensation of spaciousness can be created *e.g*. by introducing some signal delay between the two microphones, or by using spaced microphone pairs.

In recording studios, the sound from each instrument is recorded with its own microphone (sometimes more than one microphone is used, *e.g.* for drums, Hammond B3 organ/Leslie rotating speaker sounds, …), the recorded signals are then mixed together and then mastered to create a stereophonic signal.

One important criterion for realism in reproduced sound is that the spatial-temporal aspects of the original sound should be reproduced by the sound system. Obviously, stereophonic sound systems accomplish this to a much greater degree than monophonic systems, however the listener must be seated in a the "sweet spot" of the stereophonic sound field in order to take full advantage of the stereophonic effect, which is usually located in the median plane between the *L* and *R* speakers, for a rectangular listening room, and such that the angular separation of the speakers, viewed from the listener's position is ~ 40–90 degrees, which frequently presents difficulties *e.g.* in arranging a home living room for good stereophonic sound.

In a rectangular room, the optimal location for speakers is usually in the corners of the end wall of the room, because corner placement enhances the low-frequency sound (due to pressure anti-nodes in the corners of the room for the various room modes). The figure below shows the "sweet spot" favorable listening area associated with three different loudspeaker arrangements for a rectangular room with $L \times W$ dimension ratio of 3:2. The optimal arrangement is that shown in diagram (a), which has the overall largest "sweet spot", and with the speakers in the corners of the end wall of the room. Note however, that if the angular separation of the speakers is *too* narrow, the sound "image" will appear to be monophonic rather than stereophonic.

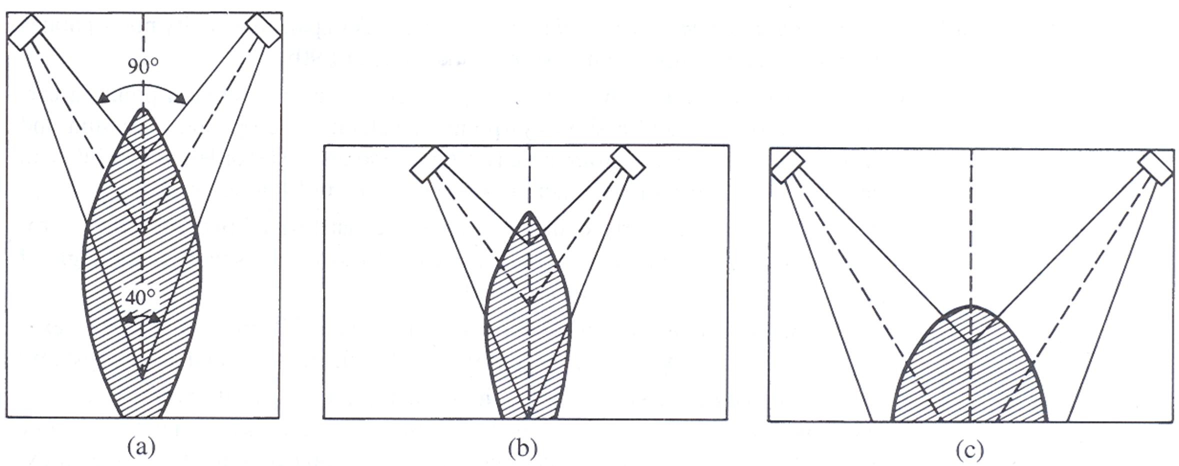


FIGURE 25.9 Favorable stereo listening areas for three different loudspeaker arrangements in a rectangular room with dimension in the ratio 3 : 2. (After Rossing 1981.)

An "improved" version of the stereo "2.0" sound system is the stereo "2.1" sound system, which simply augments the two main *L*/*R* loudspeakers with a subwoofer, as shown in the figure below:

Human binaural hearing does not do well in the spatial localization of low frequency sounds output from L/R speakers ($f < 100\ Hz$), hence the 2.1 sound system simply routes the *L*/*R* low frequencies to the single subwoofer, freeing up this task for the *L*/*R* speakers – their design can then be optimized for reproduction of all higher frequencies….

The Sound Field In Small Listening Rooms:

We can walk blindfolded into any room and very quickly, *qualitatively* assess that room's sound characteristics, because the neural sound processing center(s) in our brains have been programmed by years of auditory experiences to do so. When first-reflected sounds follow closely on the heels of the direct sound, an auditory impression of smallness is created, whereas if they arrive commensurately later, as in the case of an auditorium or concert hall, a feeling of spaciousness is created. Various attempts have been made at creating some of the acoustic features of a concert hall in a home listening room, *e.g*. using stereophonic expanders, electronic reverberation, *etc*. Placing additional speakers in the room also can help to create a feeling of spaciousness to some degree, since then direct sound arrives from several directions, rather than just two.

As mentioned/discussed earlier, in a small listening room, at low frequencies, nearly all listeners are in the reverberant sound field, whereas at high frequencies, the effect would depend on where the listener was seated – closer (or not) to the direct sound(s) emanating from the *L*/*R* stereo speakers, somewhere along the median plane between the speakers in the above figure.

This problem is compounded by the fact that in small listening rooms, due primarily to carpeting on the floor, the sound absorption *A* of the room is greater at high frequencies. While the spatial location of the sound "image" is determined by the precedence effect in association only with the direct sound(s) from the *L*/*R* stereo speakers, the tonal balance perceived by the listener appears to be derived from the total sound heard by the listener. This implies that the sound heard in the room is best described by the *power* radiated by each speaker at a given frequency, $dP(f)/df$ (*Watts*/*Hz*), rather than the {on-axis} sound pressure level, $SPL = L_p$ (*dB*).

In experiments measuring and comparing the acoustic spectral characteristics of high-quality hi-fi stereophonic sound reproduction systems in small listening rooms to actual concert halls, the mid-band (250-2500 *Hz*) characteristics were found to be comparable, whereas below 250 *Hz*, the low-frequency spectral response of the stereo hi-fi system in small listening rooms was significantly below that of the concert halls, attributed to inadequate stiffness of the walls, windows, etc. of the small listening room, which causes them to absorb low-frequency sound by vibrating sympathetically. At high frequencies, the average sound levels in small listening rooms was higher than in concert halls. Thus, these experiments suggest that in order to emulate the tonal balance heard in a concert hall, the *EQ* of a hi-fi stereophonic sound system in a small listening room should be boosted in the lower frequencies, and cut somewhat at the higher frequencies.

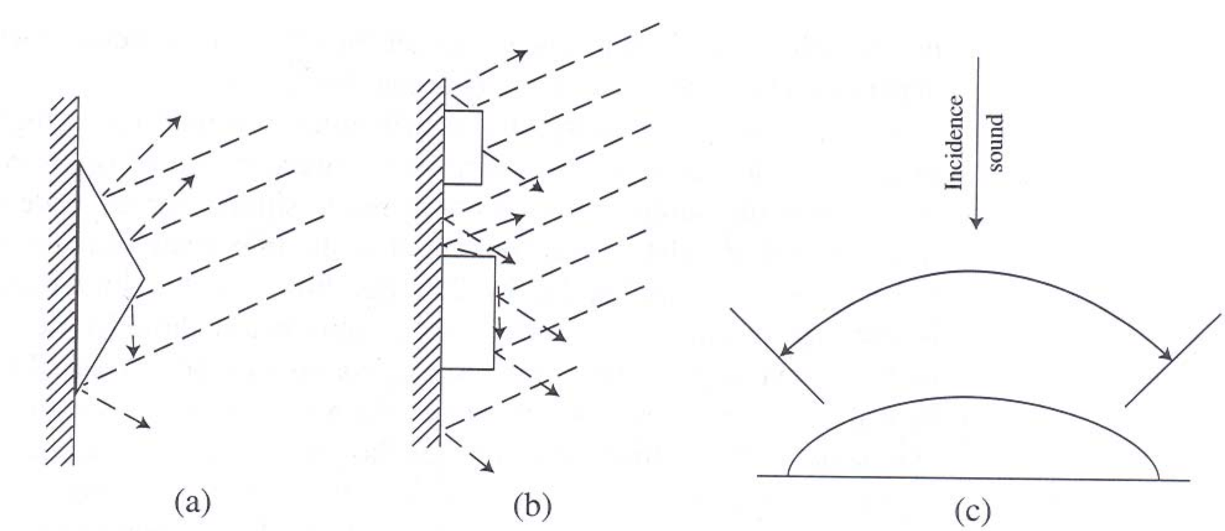
As also mentioned earlier, placing the speakers in the corners of endwall of the small listening room will preferentially help boost the low frequencies, since the corners of the room are pressure anti-nodes of the various room modes, by as much as ~ 9 *dB* in some situations. The corner walls + floor (or ceiling) of the room form a sort of frequency-dependent pyramidal horn that increases the efficiency of sound radiation at low frequencies. The frequency-dependence can be minimized by placing the loudspeaker such that the distance from the woofer cone to nearby walls – *i.e*. reflecting surfaces differ by at least a factor of 2× {*n.b*. this is also important for placement *e.g*. of bass guitar amp speaker cabinets, for gigs in smaller venues…}

The Use of Sound Diffusers In Small Listening Rooms:

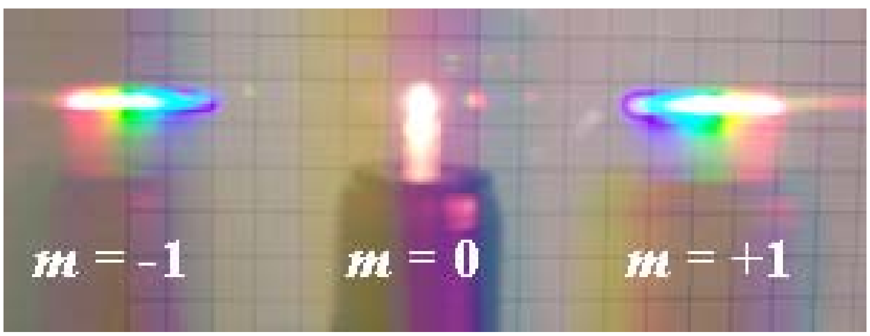
The total sound(s) that we hear in any given room at a given instant in time are a combination of direct sound(s) from sound sources in the room + indirect, reverberant sounds from multiple reflections in the room associated with direct sounds output from the sound source(s) that were produced at earlier times. For a small room, the time delays (direct *vs*. reverberant sound(s)) are characteristically shorter than for large rooms – concert halls, auditoriums, *etc*.

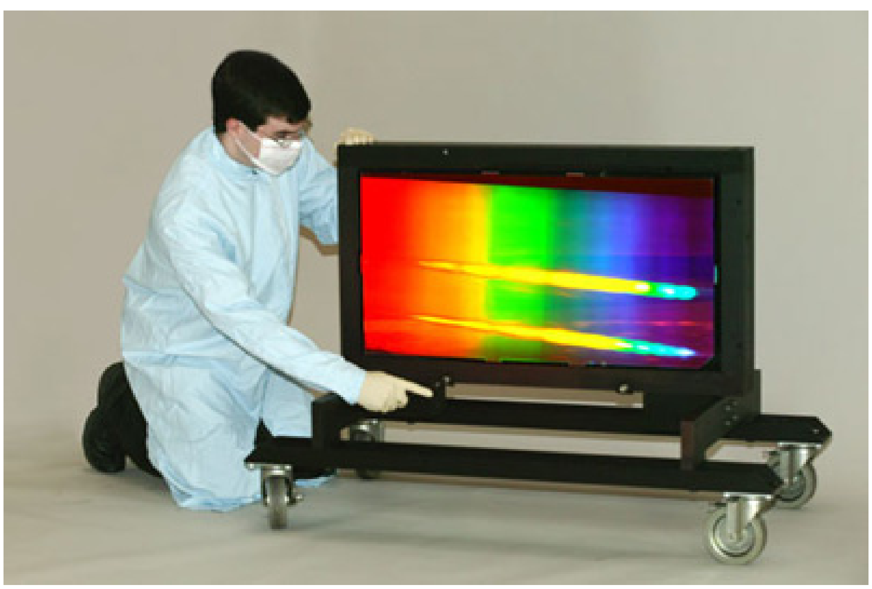
The reverberant sound also does not have the same frequency spectrum as that associated with the direct sound, for two reasons – frequency-dependent absorption of the sound by various internal surfaces in the room and also the excitation of room modes. Additionally, in small rooms oftentimes the sound at a given frequency f is absorbed before a uniform energy density $w(f)$ of reverberant sound is obtained throughout the room. Thus, the dynamical evolution of the reverberant sound field in a small room in evolving from the initial direct sound to a steady-state can be quite different than for large rooms. Furthermore, in a small listening room, *e.g*. a living room in a house, almost always the room is filled with other items – sofas, coffee tables, lamps, chairs, etc. all of which reflect & absorb the sound in a myriad of ways, from these additional objects located at different places in the room, resulting in even more complexity associated with the reverberant sound field in a small listening room.

The judicious use of sound diffusers in a small room helps/aids in creating a more uniform reverberant sound field in a small listening room, hopefully approximating that associated with a larger room. Whereas flat walls and concave surfaces tend to direct the sound, convex and/or rough surfaces will instead scatter the sound in several, possibly many directions, thereby helping to even out/make more uniform the reverberant sound field. Geometrical shapes attached to room surfaces (*i.e*. walls, floor and/or ceiling) help to scatter and diffuse the sound. Triangular, rectangular and/or semi-cylindrical protrusions on these room surfaces help to scatter the sound in many directions, thereby helping create a diffuse/more uniform reverberant sound field, as shown in the figure below:

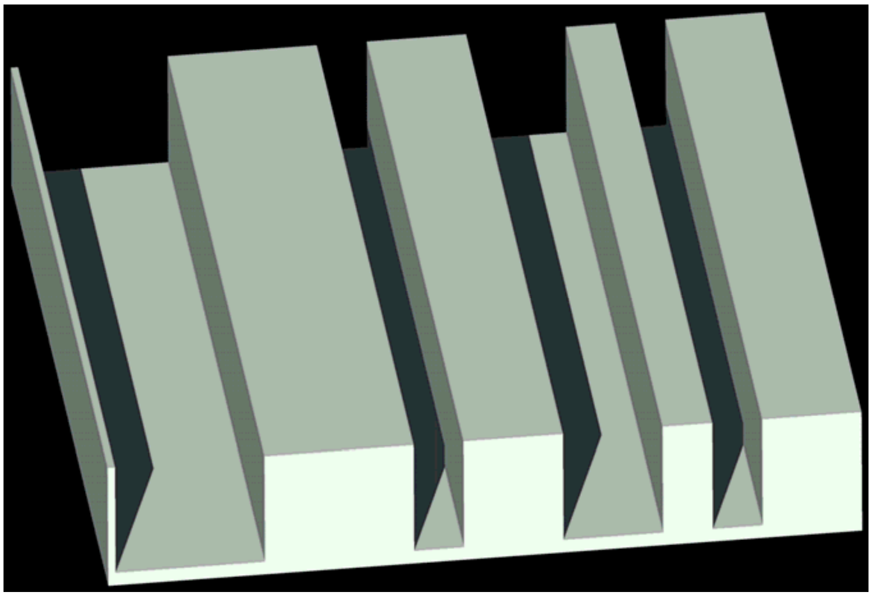


In P406 Lecture Notes 3 (*p*. 8-11), we discussed the phenomenon of sound diffraction – through apertures and around obstacles. Sound waves *e.g.* passing through a narrow opening spread out due to diffraction of the sound – the wavelength of the sound, size of the aperture / opening and the geometrical shape of opening dictate how the sound spreads out, and how much it spreads out in passing through the aperture. With two or more openings, both diffraction and interference phenomena occur, the latter can be constructive/destructive, or anywhere in between, depending on the relative phase of the waves at a given observation/listening point in 3-D space. A diffraction grating with many narrow, parallel slits illuminated either by light or sound is one example of such phenomena. Diffraction gratings also work for ***reflection*** of light (and/or sound!) off of the surface of a reflection diffraction grating of width L, consisting of N_s finely spaced parallel grooves, each separated by a very small distance $d = L/N_s \sim \lambda$, the wavelength of light. For light at normal incidence on the reflection diffraction grating, the diffracted light has maxima at an angle(s) $\pm\theta_m$ from the normal to the surface of the diffraction grating, given by the formula $d\sin\theta_m = m\lambda$ where the integer $m = 0, \pm1, \pm2, \pm3, \ldots$ {the value of $|m|$ is known as the ***order*** of the diffraction}. The following two pictures respectively show the image of a MagLite flashlight's lightbulb viewed through a transmission diffraction grating, and the image of a white light source viewed from a large reflection diffraction grating. Note that if the angle of incidence of the light with respect to the normal is θ_i, it is straightforward to show that the above formula becomes: $d\left(\sin\theta_m + \sin\theta_i\right) = m\lambda$.





In 1975, Manfred R. Schroeder, an acoustician, proposed the use of an acoustic diffraction grating (*aka* phase-grating) as an effective sound diffuser, *e.g*. for use in small listening rooms. The theory of ***phase-grating*** sound diffusers is based on ***number theory***, Schroeder used a mathematical scheme known as ***maximum length sequences***, which are a stream of fixed-length digital 1's and 0's with some interesting statistical properties. He built a prototype sound diffuser using a piece of sheet metal bent into the necessary geometrical pattern of digital 1's and 0's to confirm his theory of acoustical scattering from such an object; it looked similar to that shown in the figure below.

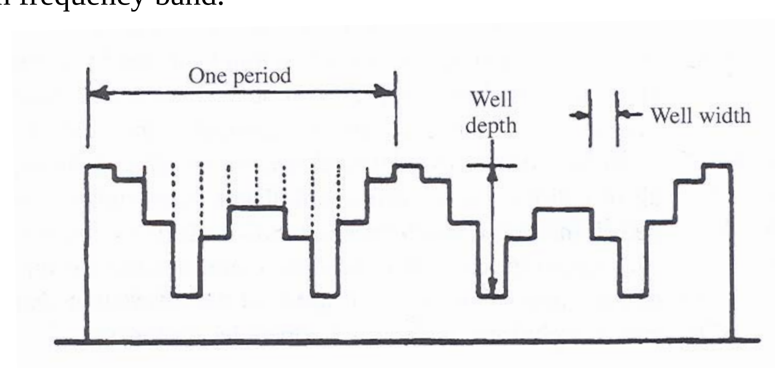


A 1-Dimensional Maximum Length Sequence (MLS) Phase-Grating Sound Diffuser

Note that such a MLS/PGD sound diffuser scatters the sound only in one direction – *e.g*. if the grooves are vertical, the scattering of sound is in the horizontal direction, *e.g*. similar to what a diffraction grating does in scattering visible light.

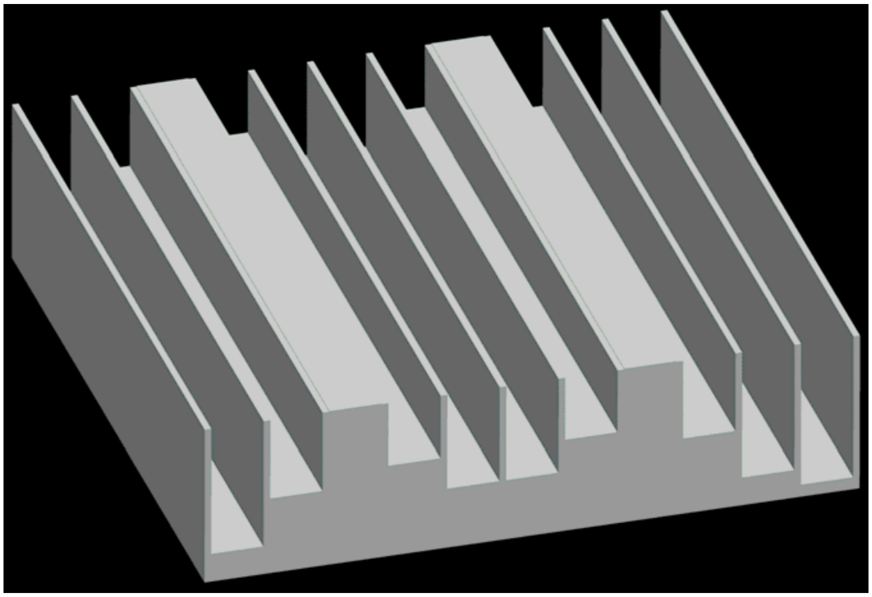
Since Schroeder's initial work much additional theoretical and experimental work has gone into the development of many new types of sound diffusers – so-called ***quadratic-residue*** diffusers (QRD's) and ***primitive-root*** diffusers (PRD's), building on Schroeder's initial theory of ***phase-grating*** sound diffusers.

The figure below shows a cross sectional view of a 1-D ***quadratic residue*** type phase-grating sound diffuser, consisting of a structure that has a repeating sequence of wells that scatter sound within a certain frequency band.



The maximum depth of the wells determines the effective low frequency limit of the diffusers. The well depth should be 1½ times the wavelength at the lowest frequency. The highest frequency scattered is determined by the well width, which is half a wavelength at the highest frequency. The actual sequence of wells used is determined by number theory.

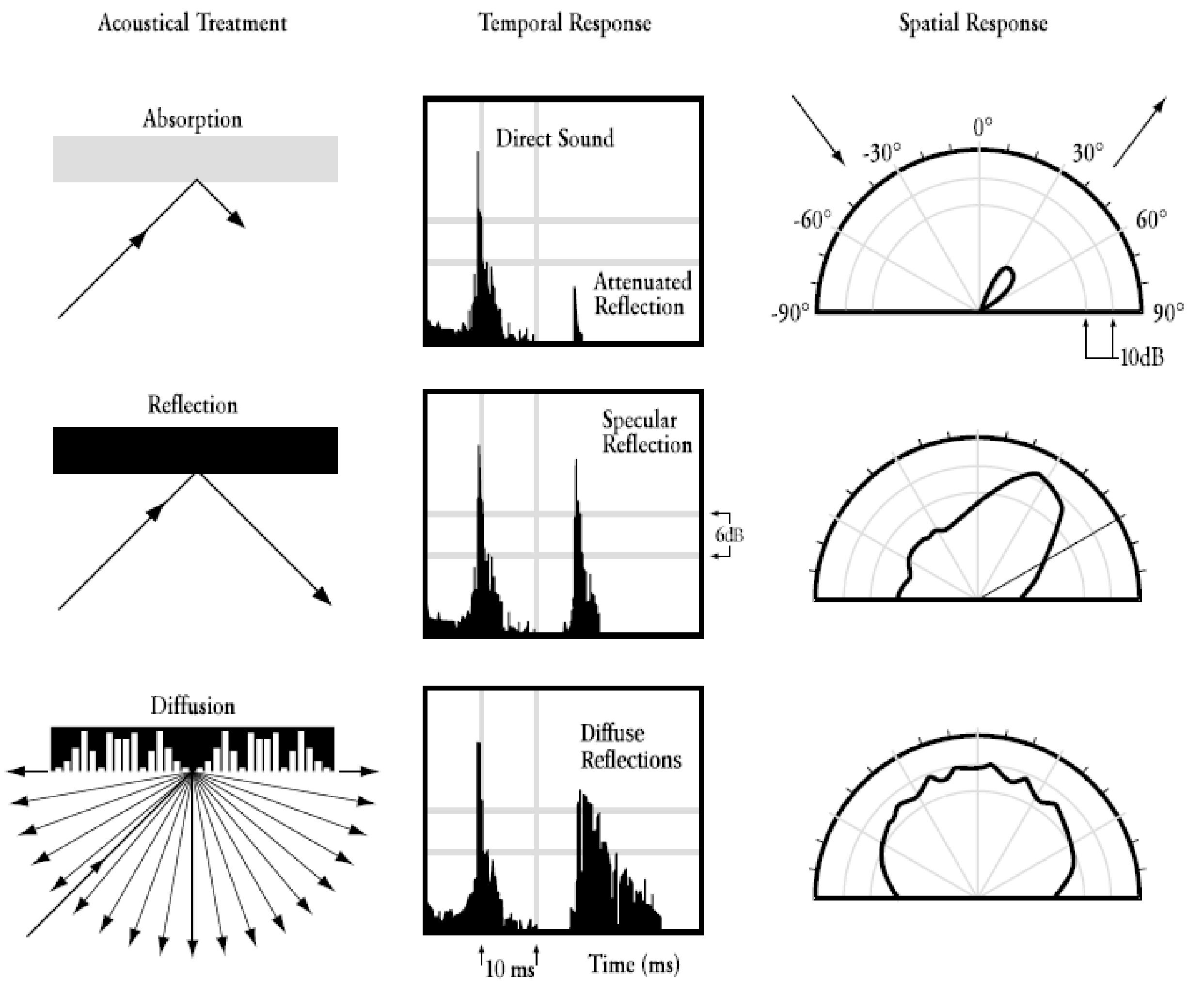
A 3-D view of a 1-D Quadratic Residue sound diffuser is shown in the figure below:



A 1-Dimensional Quadratic Residue Phase-Grating Sound Diffuser (QRD)

Schroeder-type QRD sound diffusers have been installed *e.g.* in Carnegie Hall in NYC to improve the acoustics there by eliminating echoes from the back wall of this concert hall, as shown in the figure below:

The following figure shows the efficacy of the use of a ***phase-grating*** sound diffuser – their temporal and spatial/angular response on scattering sound in all directions, compared to conventional/flat-surface sound absorbers and/or simple reflection from a planar surface:

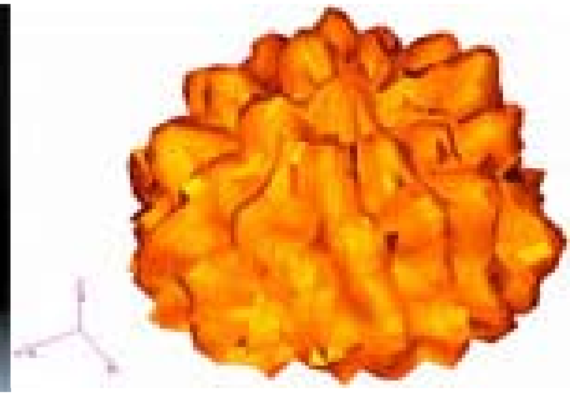


The interested reader is encouraged to Google much additional information on state-of-the-art sound diffuser technology that exists out on the *WWW*, *e.g.* see/visit the RPG Diffusor, Inc. website http://www.rpginc.com/, which has many technical papers on sound diffuser technology and other interesting information posted there. A number of DIY phase-grating sound diffuser websites also exist, along with phase-grating sound diffuser calculators.

Examples of other types of sound diffusers that have recently been developed:

2-D Quadratic Residue Diffuser (QRD)

2-D Primitive Root Diffuser (PRD)



3-D Balloon Plot for the Primitive Root Diffuser

An example of the use of QRD-type sound diffusers in a home listening room application:

Note the Flying Vee Uke!

Sound Absorption In Small Listening Rooms:

As we have seen, the sound absorption $A = Sa$ (m^2) in a room depends primarily on the frequency-dependent absorptive properties associated with the materials used for the six surfaces of room – *i.e.* the walls, floor and ceiling. The Sabine formula tells us that the reverberation time T_{60} is proportional to the volume to surface area ratio, V/S. For small listening rooms, the volume to surface area ratio V/S is usually small (compared *e.g.* to a concert hall or auditorium) and hence the reverberation time T_{60} for a typical small room is usually quite short.

A home listening room usually also has furniture, whose upholstery adds significantly to the absorption A of the room, and often has a carpeted floor, which likewise contributes to the overall A of the room. When listening to recorded or broadcast music in a small home listening room, such music often has the accompanying reverberation signature of the concert hall or recording studio in which it was recorded. Thus, in order to fully appreciate the sonic ambience of the original recording, the listening room insofar as possible should be almost free of reverberation in order not to unduly "color", or otherwise distort the sound of the original recording of the music.

Porous materials such as drapery/curtains, carpets, glass fiber and acoustical tile absorb sound energy very well at high frequencies, whereas materials commonly used in home construction such as wood, glass, gypsum board (drywall) and plaster on lath absorb sound energy very well at low frequencies. Thus, in afore-hand/custom home building, an architect-acoustician can consciously/deliberately design a quality home listening room by judicious choice of the design of the room and of the materials used in the construction of the room.

If room resonances, especially at low frequency, are problematic, another type of sound absorber that capitalizes on the {time-reversed!} principle of operation of a Helmholtz resonator can be used to provide sound absorption over a selected frequency band.

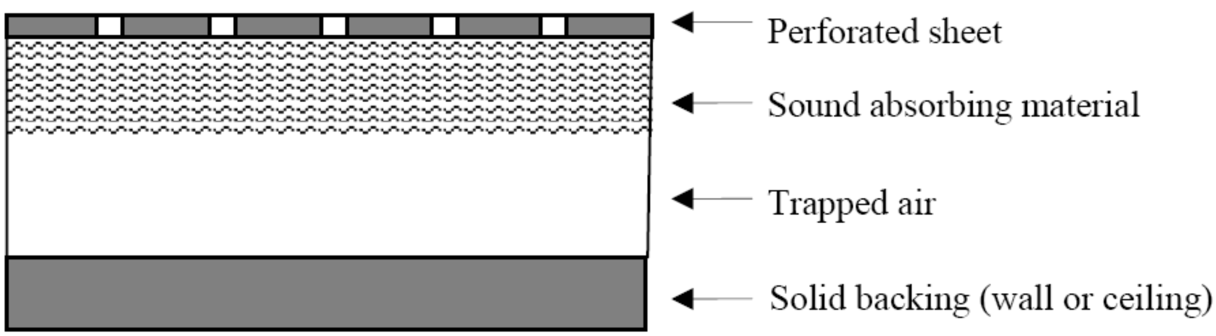
Recall that a Helmholtz resonator has a {fundamental} resonance frequency of $f_r = (v/2\pi)\sqrt{A_h/Vh'}$ where v = 343 m/s = speed of sound, $A_h = \pi r^2$ = cross sectional area (m^2) of the hole in the neck of the resonator, V = volume (m^3) of the resonator, $h' = h + \delta_{end}$ where h = length (m) of the neck of the resonator and $\delta_{end} \sim 1.7r$ is the so-called end correction. The Q-factor associated with the Helmholtz resonator is $Q_r = 2\pi\sqrt{V\left(h'/A\right)^3} = f_r/\Gamma_r$ where $\Gamma_r = f_r/Q_f = f_{hi} - f_{low} = FWHM$ of the resonance.

Sound energy from the room at/near the resonant frequency f_r of the Helmholtz resonator enters through the neck of the Helmholtz resonator and is trapped/stored inside it. By energy conservation, the energy stored in the Helmholtz resonator initially came from the room, thus there must be correspondingly less energy at this resonant frequency left in the room!

If the inside of Helmholtz resonator cavity is then made absorptive (*n.b.* increasing the resonant width Γ_r /decreasing the Q_r -factor), the stored sound energy inside the resonator is (ultimately) dissipated as heat energy (albeit very small amounts thereof).

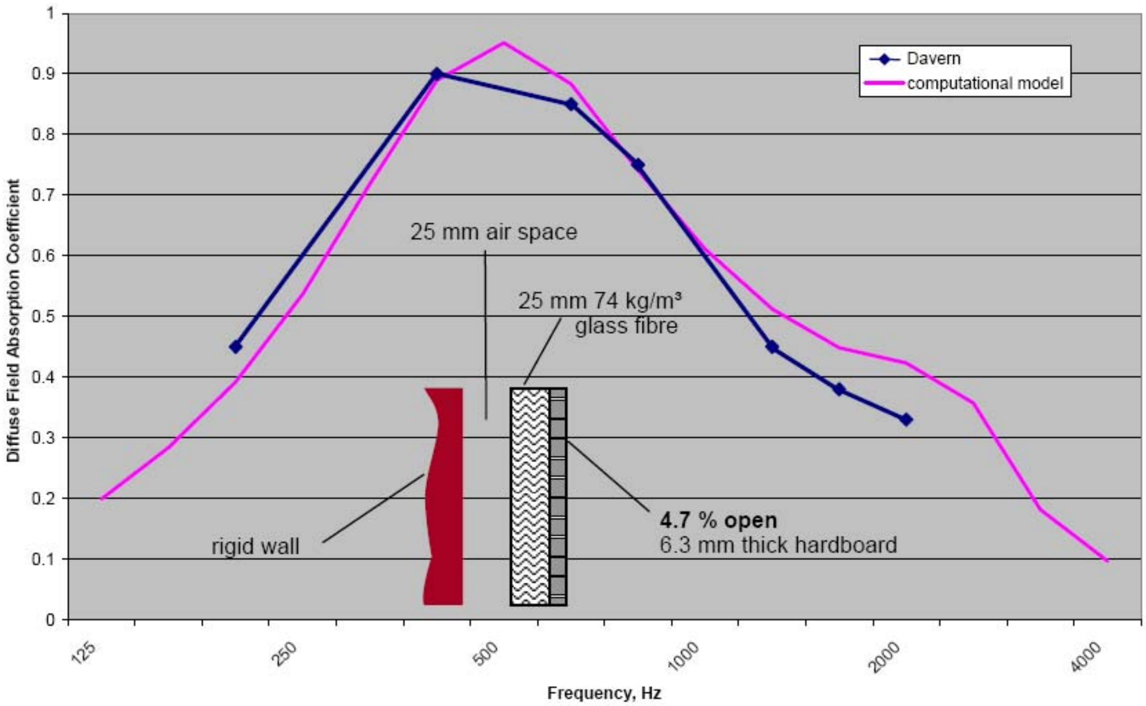
Thus, sound energy from problematic, low-frequency room mode resonances (*n.b.* oftentimes these are the lower-frequency axial modes {100, 010, 001, *etc*.}) in a small listening room can be ameliorated via the use of Helmholtz-type resonator/absorbers, strategically placed within the room – where pressure anti-nodes for these room modes exist, *e.g.* along/at the walls, for the axial modes, and/or in the corners of the room for the tangential and/or oblique modes. Note that some centuries-old churches in Scandinavia used clay pots embedded in the walls to act as Helmholtz resonators to control their low-frequency room resonances!

Rather than placing actual Helmholtz-type resonators around a home listening room, a more practical way to achieve the same type of low-frequency absorption is to use so-called perforated (or micro-perforated) panel absorbers, which operate on the same principle as a Helmholtz resonator. These so-called distributed Helmholtz resonator devices are made by covering *e.g.* a rectangular box-type cavity with a perforated panel, and using one (or two) layers of absorbing materials (*e.g.* fiberglass insulation) inside the cavity, as shown in the figure below:

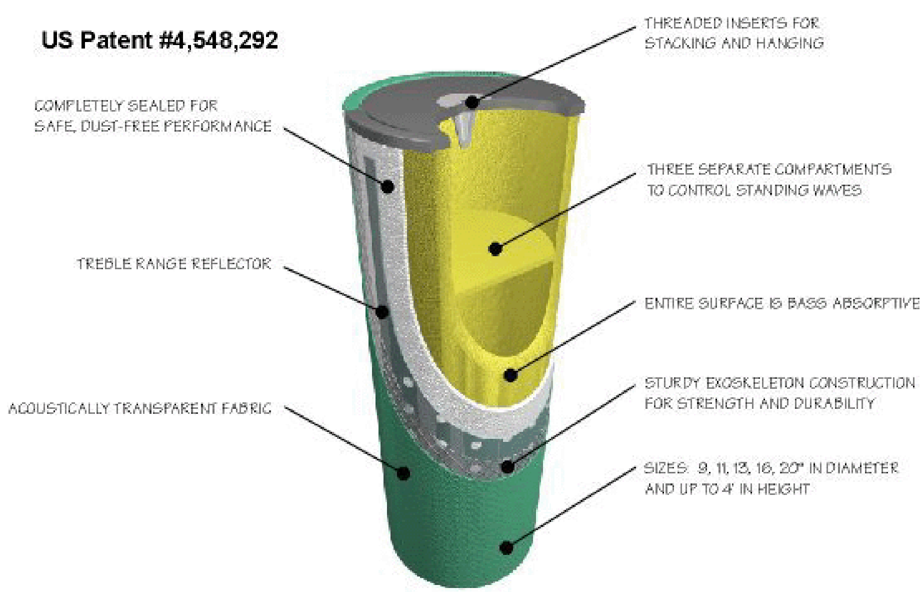


The resonance frequency of the fundamental associated with the perforated panel absorber is $f_r = (v/2\pi)\sqrt{N_h A_h/(Vh')}$ where v = 343 *m/s* is the speed of sound, N_h = # of holes on the panel, $A_h = \frac{1}{4}\pi d_h^2$ is the cross-sectional area of each of the holes in the perforated sheet, h' is the effective hole length (= thickness of perforated sheet, h + 0.85×hole diameter, d_h), $V = L \cdot W \cdot D$ is the internal volume (m^3) of the perforated panel absorber, L and W (m) are its <u>transverse</u> dimensions, $D = D_{air} + D_{abs}$ is the internal <u>depth</u> of the perforated panel absorber. Defining the <u>hole</u> <u>fraction</u> of the perforated sheet as $F_h \equiv N_h A_h / A_{sheet} = N_h A_h / L \cdot W$, the expression for the perforated panel absorber's fundamental resonance frequency is also $f_r = v/2\pi \sqrt{F_h/(Dh')}$.

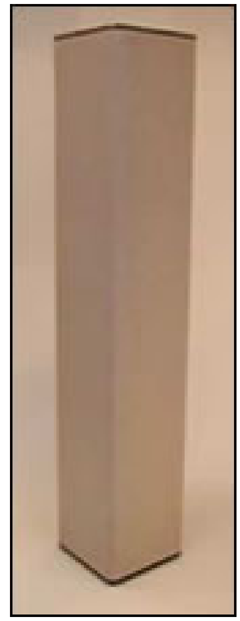
A plot of the measured *vs*. calculated sound absorption coefficient $a(f)$ vs. f is shown below for a typical multi-layer perforated panel absorber whose fundamental resonance frequency was chosen to be f_r ~ 500 *Hz*.



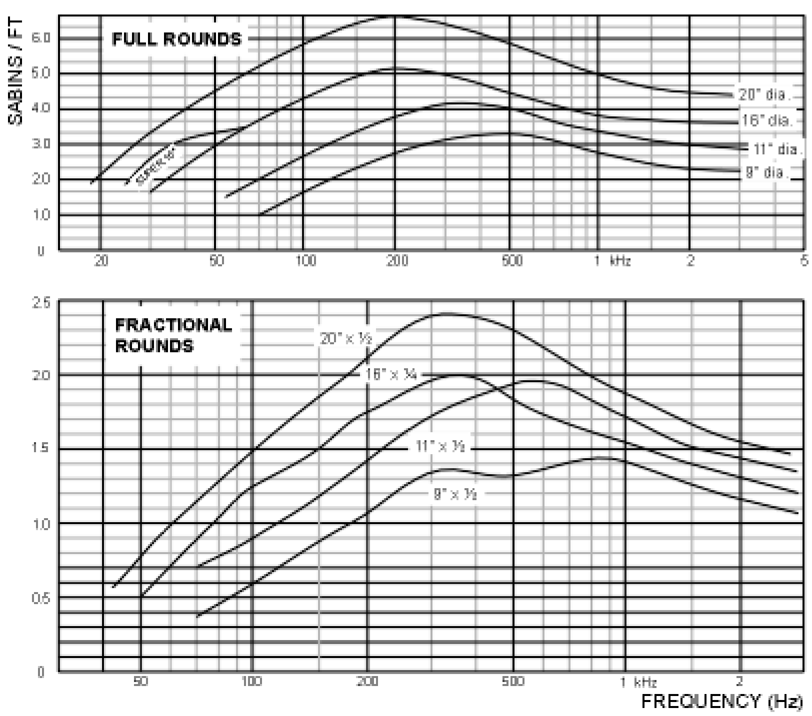
Another type of low-frequency sound absorber is known as a bass trap, which utilizes the "lossy" open-closed organ pipe cavity-type resonance as its principle of operation. The bass trap has alternating layers of absorbent, porous materials (*e.g.* fiberglass insulation) and air to absorb frequencies which have ¼-wavelengths equal to the depth of the bass trap, *i.e.* $D_{bt} = \lambda/4$.
For a depth $D_{bt} = 1\ m$, a bass trap absorbs frequencies $f_{bt} = v/\lambda = v/4D_{bt} = 343/4 \simeq 86\ Hz$.
The construction of a typical bass trap is shown in the figure below:



Bass traps can be built in a variety of shapes – *e.g.* cylinders, hemi-cylinders, towers, … as can be seen in the figures shown below:



The absorption $A(f)$ (Sabins/ft) *vs.* f for commercially-available full-, half- and quarter-round bass traps of varying diameter is shown in the graphs below:



Note that the physics of bass traps also has applications *e.g.* in automobile muffler design!

To summarize the above discussion of various types of sound absorbers, and to further clarify their generalized principle of operation: ***any*** type of acoustic resonant cavity (or structure) can be modified for use as an acoustic resonant absorber – this statement has far reaching consequences.

You may have learned *e.g.* in an *E&M* physics course that "a good (poor) emitter of radiation is also a good (poor) absorber of radiation", perhaps in the context *e.g.* of black body/thermal radiation. Precisely the same statement is applicable for acoustic radiation!

Why is this true? It is due to the fact, that at the microscopic level, sound waves/sound vibrations of any/all kinds manifestly (also) involves the electromagnetic (*EM*) interaction of atoms and molecules with each other, just as black body/thermal *EM* radiation does.

A fundamental symmetry property of the *EM* interaction, at the microscopic level {*i.e.* the exchange of virtual photons between electrically charged particles – here for acoustics, between atoms and molecules, even if overall they are electrically neutral – they are composite particles made up of point-like negative-charge electrons and positive-charge nuclei} processes involving the *EM* interaction manifestly obey time-reversal invariance – *i.e.* the picture (or movie) of an *EM* process running backwards in time is *indistinguishable* from that for the same process running forwards in time. Hence, here in an acoustical physics setting, it can be seen that an efficient radiator of sound energy will also be/can be made to be an efficient absorber of sound energy, because of/due to the manifest time-reversal invariant nature of the *EM* interaction at the microscopic scale. This may seem to be trivial statement, but it in fact is by no means the case, since we know of another fundamental force of nature – the weak interaction (*e.g.* responsible for radioactivity/beta-decay of nuclei) which manifestly violates time-reversal invariance in certain situations – *e.g.* the weak decays of neutral *K* and *B* mesons!

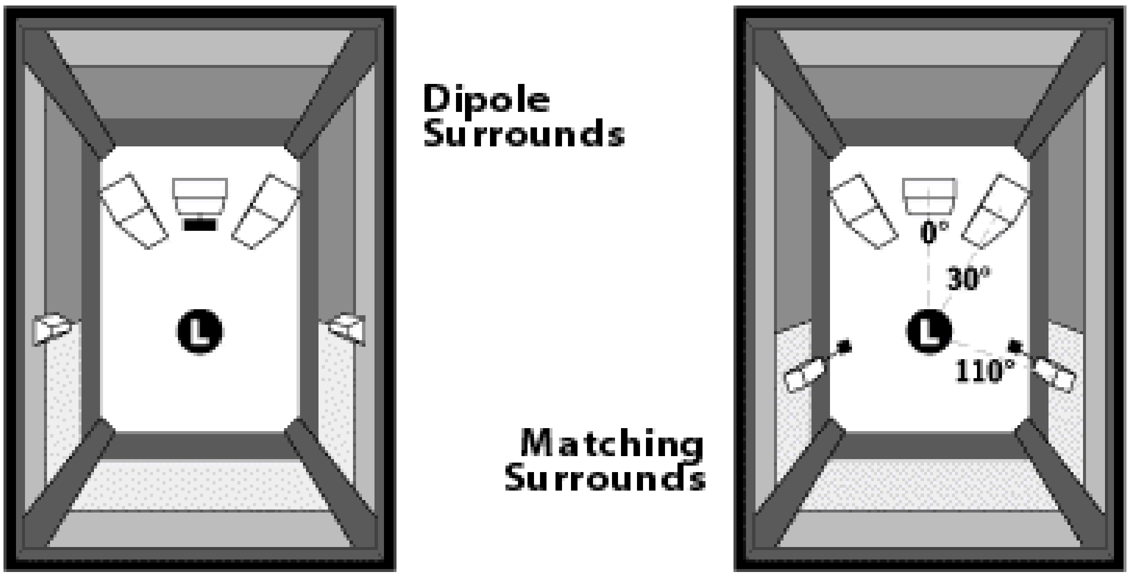
Home Theater & Surround-Sound Systems:

For today's home theater, their design is such that typically the room used for home theater entertainment is systematically somewhat larger than that of the average hi-fi home listening room, however such rooms are still small in comparison to concert halls, auditoriums, etc. Acoustically, the goal of a home theater is to replicate that of a commercial movie theater, which often uses the 5.1 surround-sound system – hence home theaters will have this also.

The 5.1 surround-sound system uses 5 separate loudspeakers – left, right, center, left surround and right surround, and a subwoofer (the .1 of 5.1). The center speaker is optional in some 5.1 S-S recordings, but is important in motion pictures, *e.g.* for speech dialog between characters/actors.

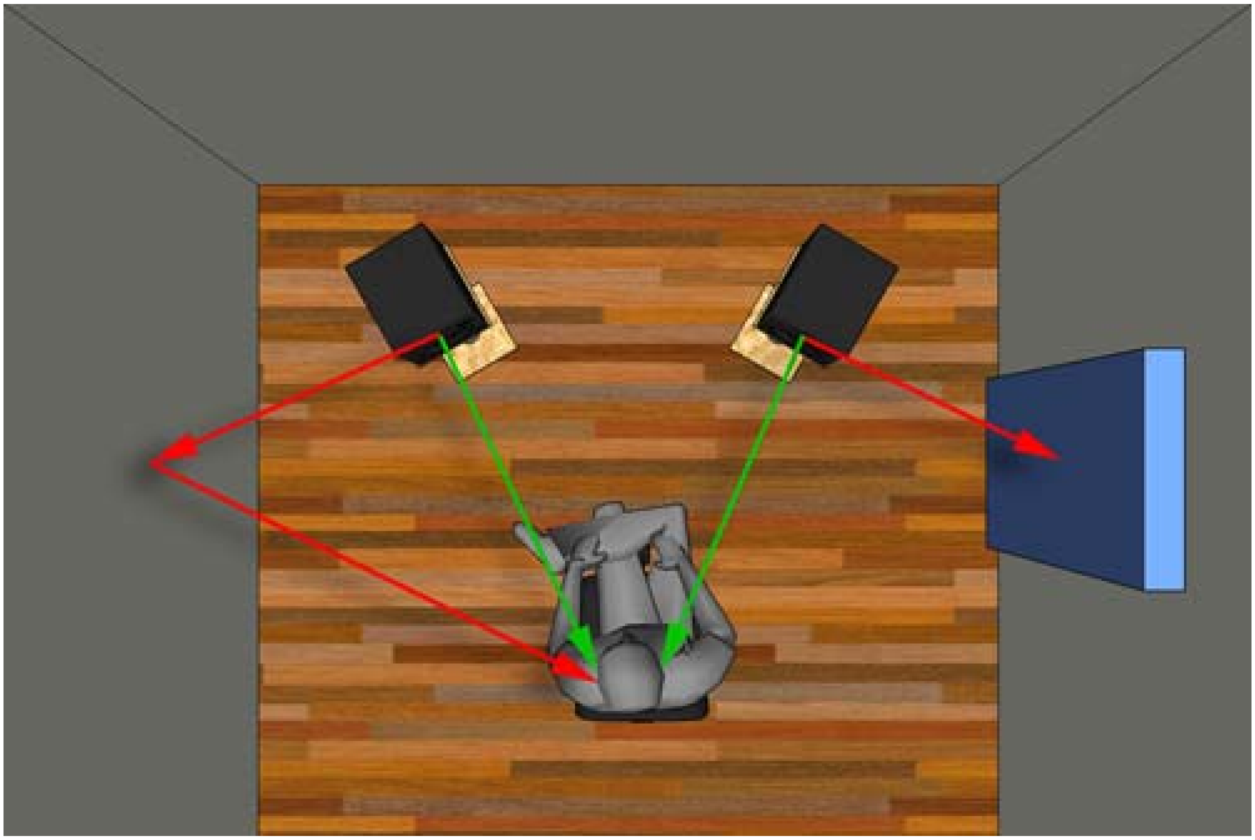
For hi-fi stereophonic home listening rooms, we discussed the importance of the listener being in the "sweet spot" of the sound "image", located on the median plane between the *L* & *R* speakers (*p.* 12 of these lecture notes / Fig 25.9 *p.*578 of SoS textbook). In home theaters, this is impractical (as it is in commercial movie theaters) because there often are many people wanting to watch a movie, and they all can't fit into the "sweet spot" together/at the same time. This is precisely why the center speaker in 5.1 S-S is used primarily for speech dialog – it is centrally localized.

The *L* & *R* surrounds can either be (*a*) placed as so-called dipole surrounds (USA THX-format), where they are located on the side walls directly to the left and right of the listener, at ±90° to the median plane of the system (*i.e.* 180° apart from each other), with all 5 speakers equidistant from the listener, as shown in the *LHS* figure below, or (*b*) the *L* & *R* surrounds can be placed as a so-called matching surrounds (European ITU-format), where they are again located on the side walls, but at the somewhat larger angle of ±110° to the median plane of the system (*i.e.* 140° apart from each other), again with all 5 speakers equidistant to the listener, as shown in the *RHS* figure below:

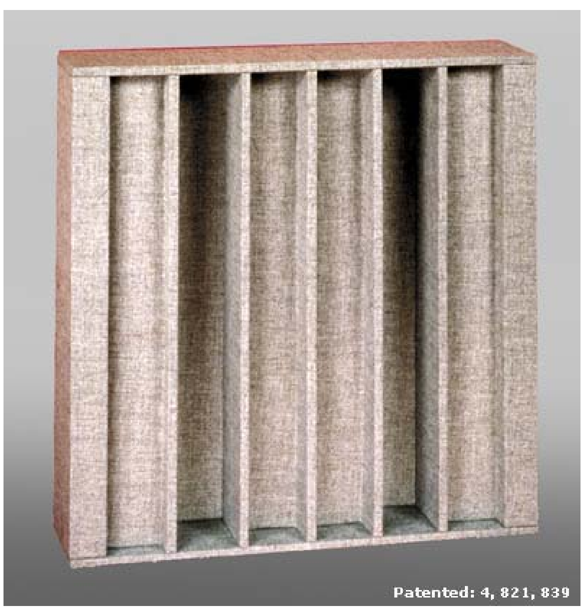


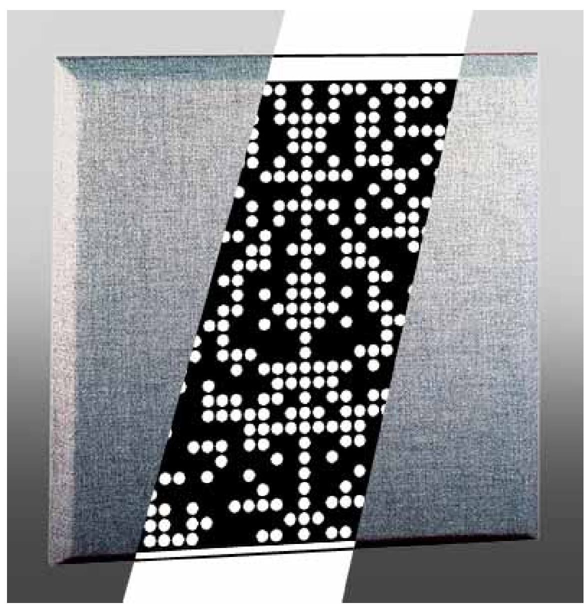
The *L* & *R* front speakers in both home theater configurations are located at ±30° to the median plane of the system (*i.e.* 60° apart from each other), with the center speaker located on the median plane of the system, as it is in commercial movie theaters (but located behind the screen).

While early sound reflections in a concert hall or large auditorium enhance the overall sound, giving rise to feelings of intimacy and ambience in the ears of concert-goers, for small listening rooms and/or home theaters, early sound reflections interact adversely with the direct sound from the 5.1 S-S system, resulting in comb-filtering – *i.e.* a "hilly" rather than flat frequency response – one which has peaks and dips in the sound spectrum due to partial constructive/destructive interference at certain frequencies. The spatial "image" effect(s) achieved in 5.1 surround-sound systems are achieved primarily via signal processing rather than via the room acoustics of the home theater, and so early sound reflections can detract/distract from the intended original audio signals emanating from the 5.1 S-S system, corrupting the original sound stage. Hence, *e.g.* the use of phase-grating sound diffusers on the walls of the home theater can be very helpful in dispersing the sound energy associated with the early reflections, thereby significantly alleviating these problems, as shown in the figure below:



Abfusors™ and diffusorbers/diffsorbers™ are respectively sound absorbent phase-grating panels and sound diffusing perforated panel absorbers that have recently been developed and are now commercially available for such uses, as shown in the figures below:

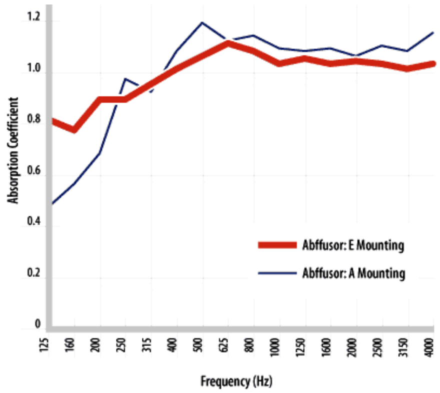


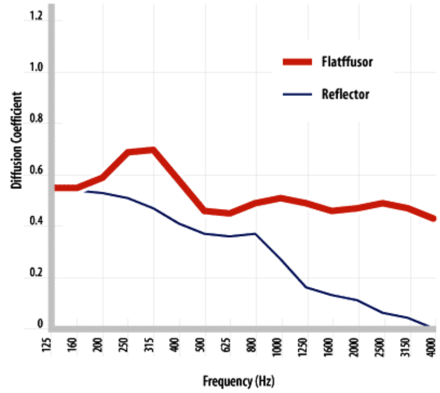


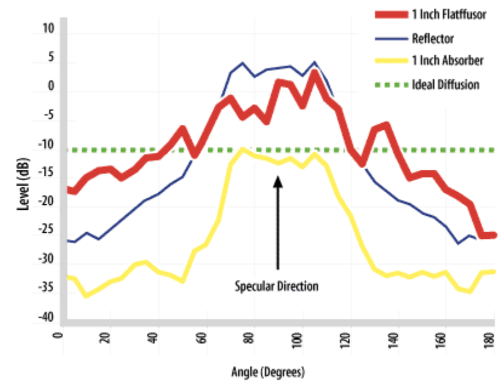
RPG Abfusor™

RPG "Flatfusor" ™ Diffusorber

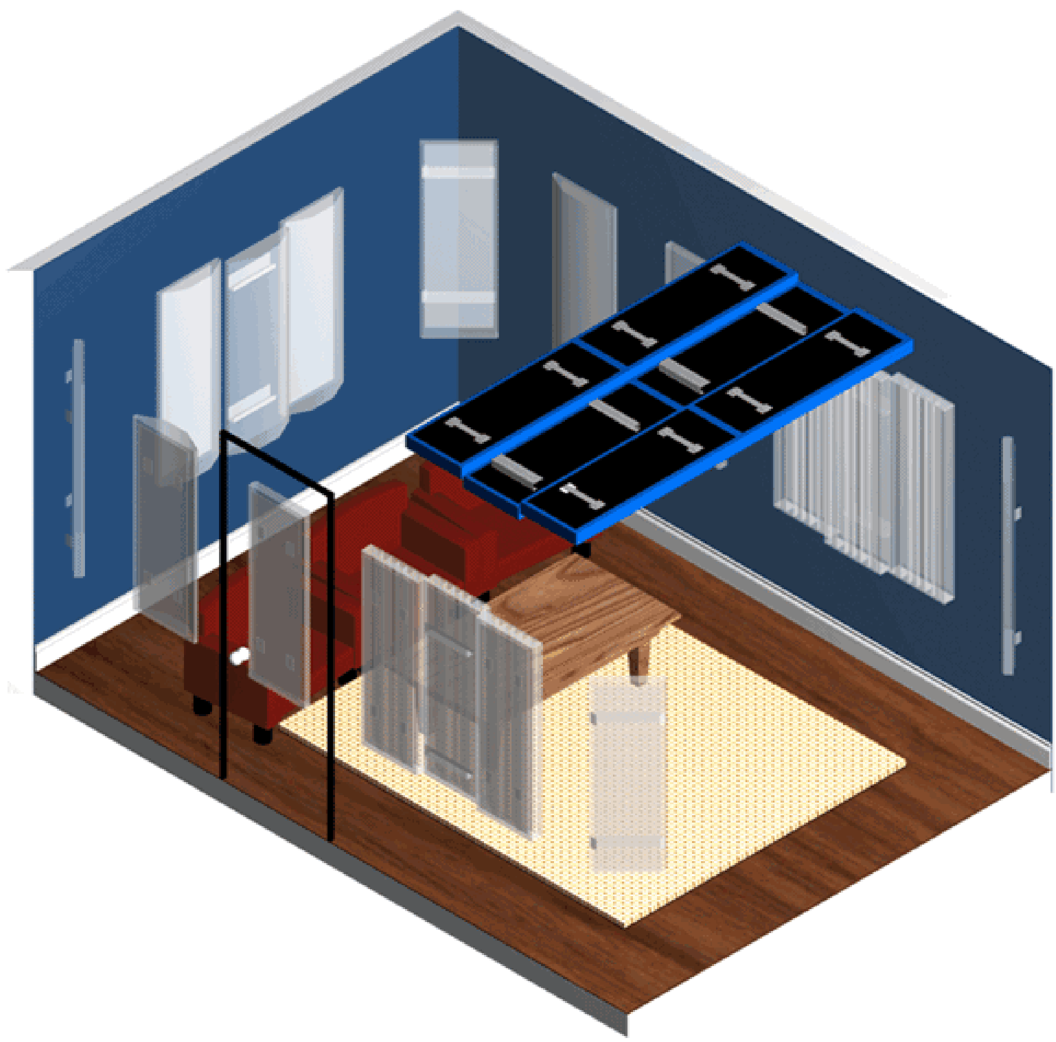
The absorption coefficient of the RPG abfusor™ and diffusion coefficient and polar response of RPG's zero-depth "Flatfusor" ™ diffusorber are shown in the figures below:



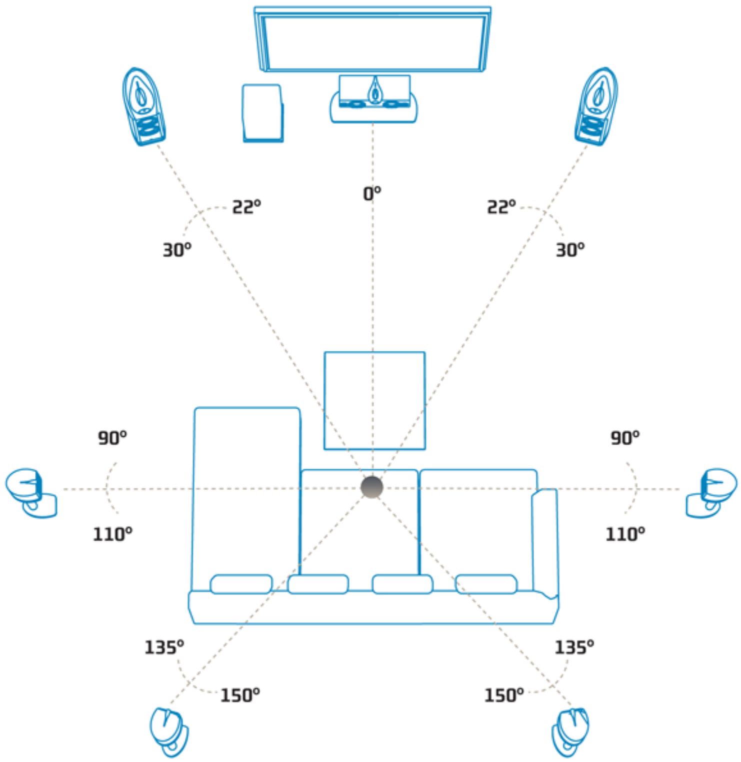




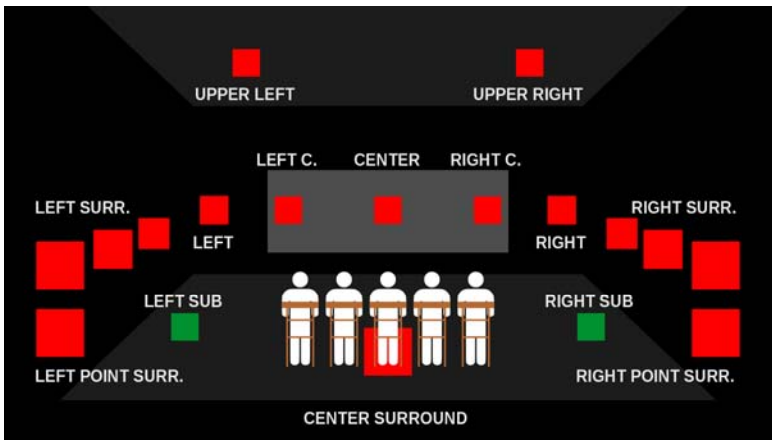
A complete, "top-down" design for excellent home theater acoustics, utilizing a variety of state-of-the-art, strategically-placed sound diffusing/sound absorbing panels might look something like that shown in the figure below:



More recently, 6.1 and 7.1 surround-sound systems have been developed, the latter of which uses 7 separate loudspeakers – front left/right, center, left/right surrounds, left/right rear speakers and a subwoofer, as shown in the figure below:



A 10.2 channel surround system ("twice as good as 5.1") has also been developed (primarily for use in commercial theaters), which has 14 channels total, as shown in the figure below:



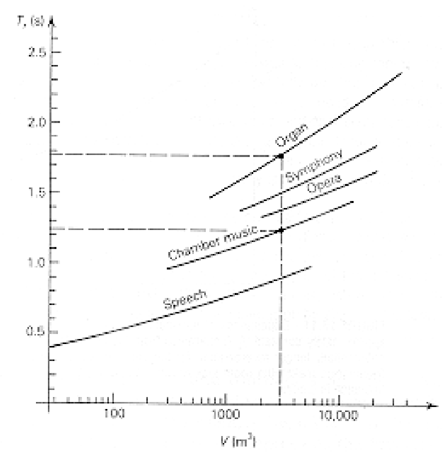
A 22.2 channel surround system has also been developed, for ultra-high definition television, which uses 24 speakers, arranged in 3 layers – a middle layer of 10 speakers, and upper layer of 9 speakers, a lower layer of 3 speakers and 2 subwoofers.

<u>Sound Recording Studios</u>:

Sound recording studios vary widely in size, scope/need and in design. Small, home recording type studios or ones used *e.g*. for recording soloists or ensembles may be as small as $V \sim 100\ m^3$ ($1000\ ft^3$), whereas a small chamber-music studio, with volume $V \sim 1000\ m^3$ ($35{,}000\ ft^3$) can accommodate a small orchestra, choir, or instrumental ensembles. A large music studio, such as the one shown in the figure below, would have a volume $V \sim 2000\ m^3$ ($70{,}000\ ft^3$) or even more.

The recording studio must be large enough so that the musicians feel comfortable/at ease playing their music, however sound reflection path length(s) to the microphones must be kept as short as possible.

Reverberation time is carefully controlled/tuned in recording studios. From the Sabine formula $T_{60} = 0.161V/A$, it (obviously) depends linearly on the room volume V, but also depends on the type/style/genre of music being recorded, as can be seen from the figure on the right.



Reverberation times in recording studios are usually shorter than those found in concert halls. In chamber music recording studios, reverberation times are typically ~ 0.9 to 1.2 *s*, whereas in larger recording studios, the reverberation times are typically ~ 1.2 to 2.4 *s*. Movable panels with variable absorption can be used in recording studios to alter the reverberation time.

One example of such panels is a rotatable panel, flat on one side with absorptive material, the other side being convex in shape and treated *e.g.* with hardboard, to make it reflective. The reverberation time of the room is increased (decreased) by rotating the reflective (absorptive) side out – *i.e.* towards the inside of the room.

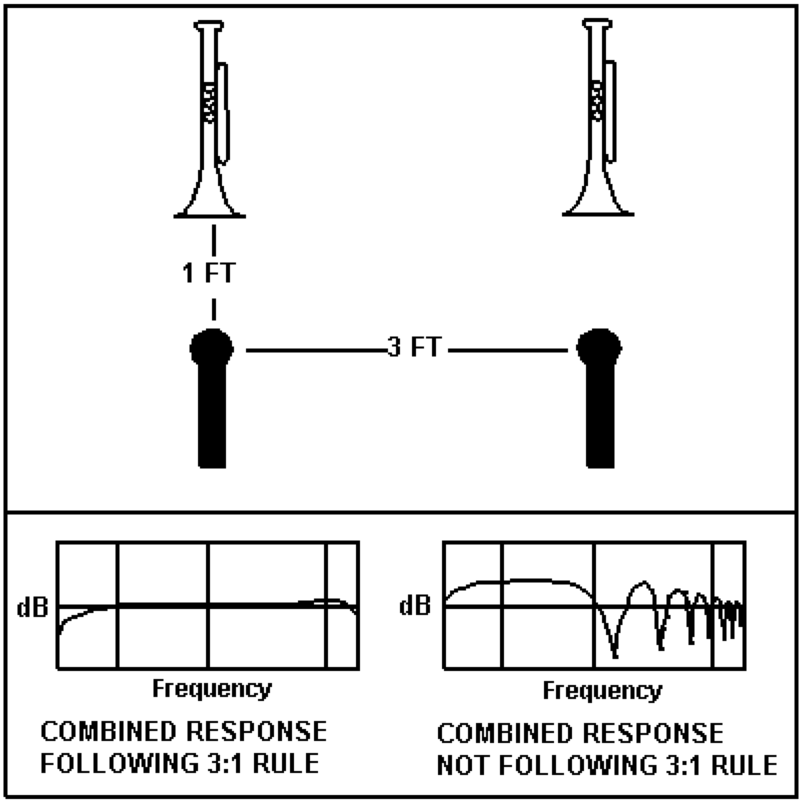
Control of the axial/tangential/oblique room mode resonances and scattering reflections is critical in a recording studio, far more so than for home listening rooms/home theaters, via use of strategically placed sound diffusers (note the phase-grating diffuser and other sound absorbing panels on the walls of the recording studio in the above photo). In order to suppress flutter echoes (rapid-fire echoes associated primarily with axial modes between parallel opposing hard walls, with periodicity $T = L/v$, originating from transient/impulse-type/short-duration sounds – *e.g.* hits on a snare drum) and to distribute room resonance frequencies more evenly, recording studios are often consciously built with irregular shaped *vs*. all-parallel walls (note the irregular ceiling and sound diffusers in the above photo).

As mentioned previously, sound diffuser panels as well as sound absorber panels, traps, *etc*. are used in combination to control reflections and resonances in the recording studio. The optimal placement of sound diffusers and absorbers will depend on the geometrical details of the shape of the recording studio, but again, generally speaking, the optimal placement for bass traps will be along walls/in corners of the room – at the pressure anti-nodes of the low-frequency standing waves of the room.

Another important parameter is the Initial Time Delay (ITD) – which is the time difference between the direct sound and the first reflected sound reaching the recording microphone. The ITD helps determine the intimacy of the music, and which is controlled by the relative placement of (*a*) the musician performer, (*b*) the microphone and (*c*) the surface on which the first sound reflection occurs.

Two additional important factors in recording studio design (and operation) are noise isolation and ambient noise level(s). Ambient noise in the recording studio must be kept as low as humanly possible. It makes no sense to locate a recording studio *e.g.* near a busy train station or heavy industries, so a site environmental noise survey should be done afore hand. Specifications for noise isolation are written that drive the construction details of the studio for walls, doors, windows, and, since rooms require adequate ventilation, and thus HVAC noise levels pose significant design considerations for recording studios.

Sound isolation within the recording studio is often called for. An overall, ensemble-type sound is usually desired for orchestras and choirs, requiring recording microphones to be placed a distance from the musical group so that sounds blend together naturally before reaching the microphones. For soloists and smaller ensembles, the close-miking technique is often used to record the individual performer's sounds, thus requiring mixing in post-recording production. In such situations, it is (highly) undesirable for the sound of one performer to be picked up by the microphone of another. The so-called *rule-of-three* is often used to ensure that the distance of one musician to any other microphone is at least 3× the distance to his/her own microphone, in order to suppress unwanted so-called "comb-filtering" frequency-dependent constructive/ destructive interference effects, as shown in the figure below:



For rock music, with its intrinsically higher sound levels, extra isolation is required, especially for recording the drummer playing his/her drum set – a special, isolated room called a drum cage is often used for this. An isolated booth is often also used for vocalists in rock bands.

Control Rooms In Recording Studios:

Two activities take place in the control room of a sound recording studio – sound recording engineer(s) record in real time the (live) music being played in the sound studio room, and then mix (and sometimes master) the recorded music in post-recording sessions, using the studio's mixing console and other associated sound recording electronics. Depending on the type/style/genre of music, the music from individual musicians may be recorded separately / individually from each other (*i.e.* at different times), or as a group/ensemble/whole orchestra.

The acoustical requirements of a control room differ significantly from that of the sound recording studio itself. Usually (but not always) the size of the control room is smaller than that of the sound recording studio. The sound recording engineer needs to be able to hear the sounds being recorded (or already-recorded sounds) played back via a pair of so-called reference monitoring loudspeakers (for a stereophonic recording) – which are very high fidelity, flat-response stereo loudspeakers, which ideally do not color or otherwise distort/change the recorded sound(s) in any manner whatsoever.

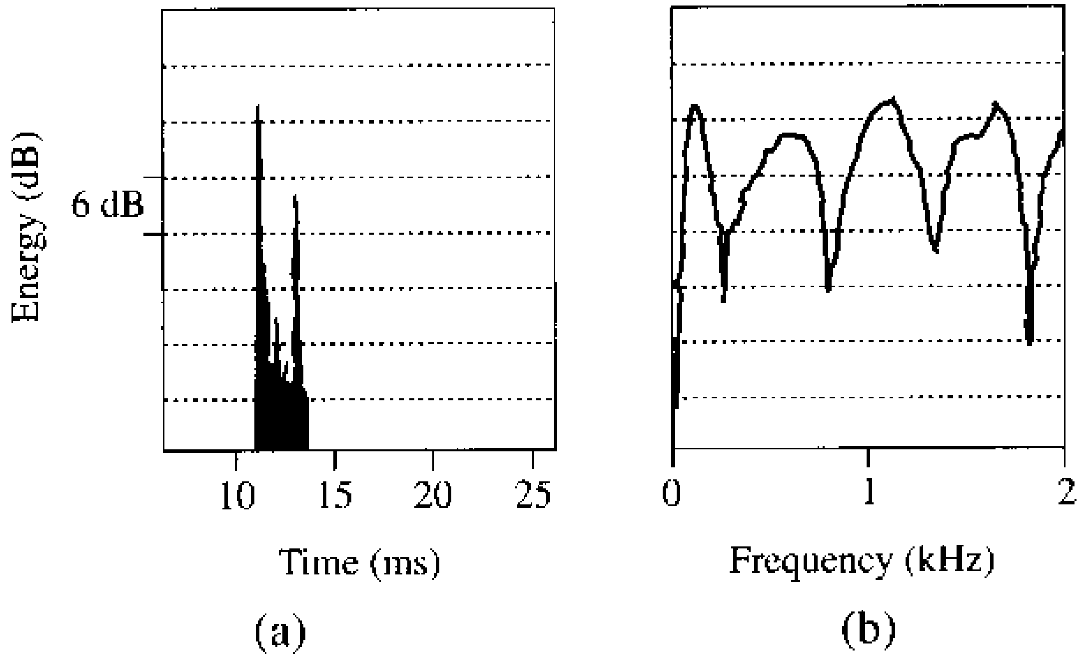
A key issue in any recording studio is transferability: the ability of a mix to be transferred to any/all other listening environments outside the original recording studio. In order for a mix to faithfully transfer to a wide range of acoustical environments, the original mix must be created in a room with minimal acoustic distortion. The four sources of acoustic distortion are (*a*) modal emphasis, (*b*) speaker boundary interference, (*c*) comb filtering, and (*d*) sparse reflection density. Professional recording engineers will attest to the importance of mixing in an acoustically well-designed room. Hence RPG Inc's slogan: "If you can't take the room out of your mix, you can't take your mix out of the room".

The acoustics of the control room need to be such that the sound recording engineers can ideally hear the direct sound from the reference monitors with no interference/coloration of the direct sound by reflected/reverberant sound in the control room. Very often, the mixing console is located at the front of the control room, directly in front of/near a ~ large window viewing the musicians in the recording studio. The thickness and construction of this window is important because if it is not thick/absorptive enough, it can transmit the live sound from the recording studio into the control room, thereby obscuring/interfering with the sound coming from the engineer's reference monitors.

The positioning of the stereo pair of reference monitors relative to the listening position of the sound recording engineer is very important, just as it is in a home listening room, for accurate L-R stereo-image positioning of the recorded stereo signal(s).

The early reflections of the direct sound from the reference monitors off of the walls, floor and ceiling of the control room can cause problems/interfere with the direct sound of the reference monitor in several ways. Sound reflections from the nearby surfaces at the front of the room could back to the sound recording engineers position could be as short as ~ 1-5 *msec*, and can adversely color/affect the sound recording engineer's perception of the direct sound coming from the reference monitors. For this reason, very often the front portion of the control room has much sound absorption *A* associated with it.

Early reflections from surfaces at/near the front of the control room can also interfere constructively/destructively with the direct sound coming from the reference monitors – this interference is known as comb filtering – arising due to phase differences of direct *vs*. early reflected sound's path lengths, manifesting itself as constructive interference peaks and destructive interference dips distributed across the audio frequency spectrum, as shown in the figure below. Again, absorbing the early sound reflecting off of the surfaces near the front portion of the control room helps to suppress frequency-dependent comb-filtering type interference effects.

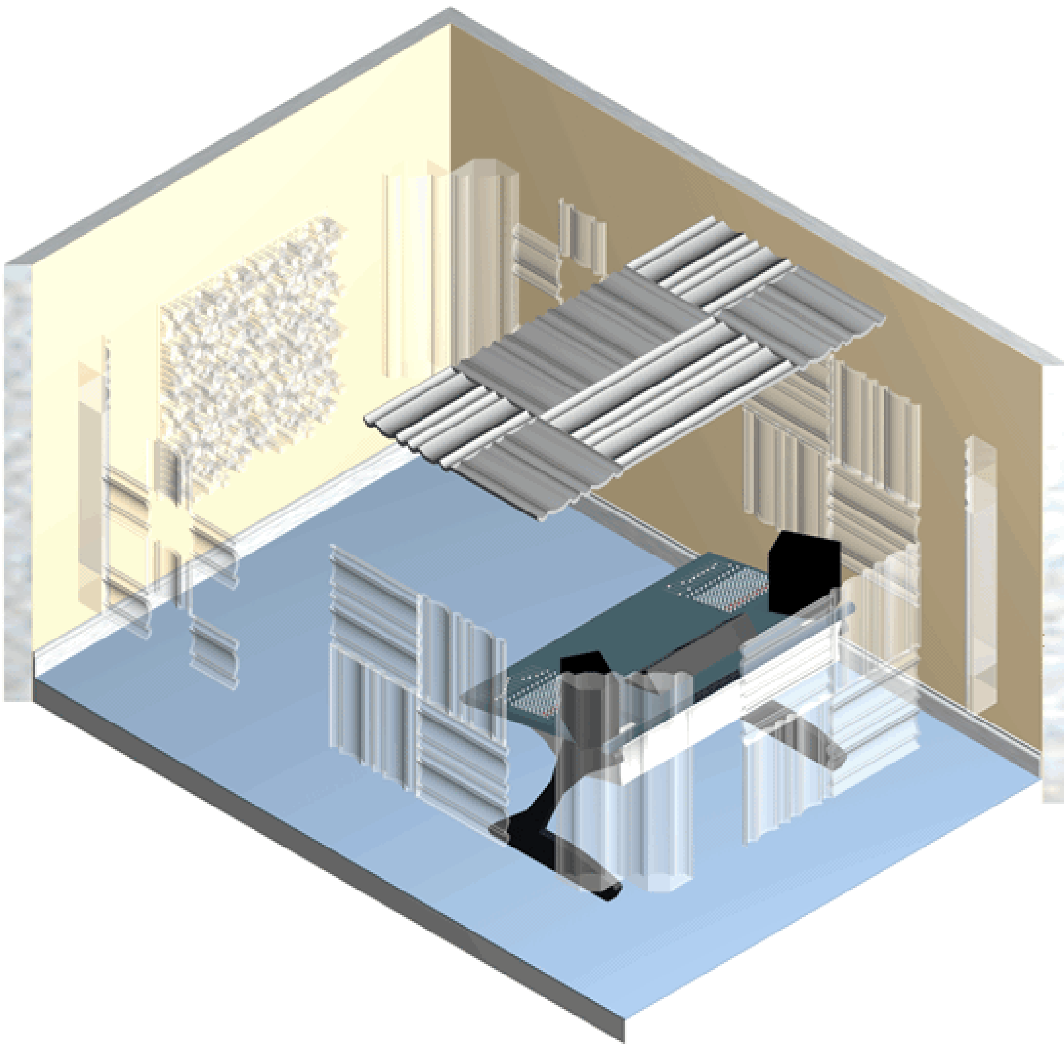


An intimately related interference effect is known as <u>speaker</u> <u>boundary</u> <u>interference</u> - the coherent interference between the direct sound emanating from a reference monitor loudspeaker and the reflections from the room it is in, in particular, usually from the corner immediately next to the loudspeaker. This type of audio distortion, like comb-filtering, can also occur over the entire frequency spectrum, but is usually more problematic at lower frequencies.

The control room's boundaries – walls, floor, ceiling if highly acoustically reflecting, mirror the sound coming from the loudspeaker, forming virtual sound sources behind these room surfaces. These first-reflection virtual sound sources then interfere constructively/destructively to varying degrees with the direct sound from the loudspeaker, depending on the amplitude and phase relationship between the direct sound *vs*. reflected sound(s) at the listening position. If *e.g.* a loudspeaker is located ~ 1 *m* from each surface in the corner in the control room, then there will be a total of 11 virtual images of the loudspeaker formed in the room! Subsequent reflections of the sound will produce even more virtual images, located behind the first 11. If the walls of the sound room were perfectly reflecting, in the steady-state, there would be an infinite number of images of the loudspeaker formed in each corner of the room, each fading off into the distance, just as in the case of light, for a real room of mirrors. Moving the loudspeaker farther away from the nearest adjacent corner will lower the frequency of the first destructive interference notch, and if far enough away, hopefully it will be below the lower cutoff frequency of the loudspeaker. However, *e.g.* for a cutoff frequency of 20 *Hz* this distance is ~ 5 *m*!

The nature of the reverberant sound field associated with a recording studio's control room is also critical. It is best that the {frequency-dependent} reverberation time of the control room be relatively short – shorter than that of the reverberation time of the recording studio itself, so that the natural reverberation effects of the recording studio can be clearly heard by the recording engineer.

The behavior/nature of the reverberant sound field's acoustic room modes associated with the control room is also critical – again usually for the lowest frequency modes. Afore-hand design *e.g.* of the geometrical shape of the control room can help to reduce such problems. Frequently, the rear portion of a control room has many diffusing-type surfaces (as discussed above) to spread out/disperse the sound waves reflecting off of the wall surfaces at the rear of the control room. Frequency-specific sound absorbing resonant cavities (such as those discussed above) can be placed *e.g.* in the corners of the control room to specifically absorb/damp problematic room modes of the control room. A well-designed "top-down" control room in a recording studio might look something like that shown in the figure below:



Legal Disclaimer and Copyright Notice:

Legal Disclaimer:

The author specifically disclaims legal responsibility for any loss of profit, or any consequential, incidental, and/or other damages resulting from the mis-use of information contained in this document. The author has made every effort possible to ensure that the information contained in this document is factually and technically accurate and correct.

Copyright Notice:

The contents of this document are protected under both United States of America and International Copyright Laws. No portion of this document may be reproduced in any manner for commercial use without prior *written* permission from the author of this document. The author grants permission for the use of information contained in this document for *private*, *non-commercial* purposes only.