Recording and Reproduction over Two Loudspeakers as Heard Live

Part 1: Hearing, Loudspeakers and Rooms

Siegfried Linkwitz[1], Don Barringer[2]

[1] Linkwitz Lab, Corte Madera, CA 94925, USA
sl@linkwitzlab.com

[2] Linkwitz Lab, Arlington, VA 22201, USA
pagebarringer@verizon.net

ABSTRACT

Innate hearing processes define the realism that can be obtained from reproduced sound. An unspecified system with two loudspeakers in a room places considerable limitations upon the degree of auditory realism that can be obtained. It has been observed that loudspeakers and room must be hidden from the auditory scene that is evoked in the listener's brain. Requirements upon the polar response and the output volume capability of the loudspeaker will be discussed. Problems and solutions in designing a 3-way, open baffle loudspeaker with piston drivers will be presented. Loudspeakers and listener must be symmetrically placed in the room to minimize the effects of reflections upon the auditory illusion.

1. INTRODUCTION

Two-channel stereo sound reproduction in domestic environments has a long history. Today the vast majority of loudspeakers are basically boxes with two or more drive units. They cover a frequency range from at least 50 Hz to 15 kHz on-axis, though generally at lower than live volume levels. Bass frequencies are usually augmented by a vent resonance. Common loudspeaker architectures and design procedures, aided by computer models and software, have led to a generic loudspeaker

sound. It is readily recognized as different from unamplified sound.

Loudspeakers are listened to in rooms of various shapes and sizes. They are often placed against walls and in corners for practical or aesthetic reasons, though this will deteriorate the sound quality. The only requirement for 2-channel playback is a triangular setup of loudspeakers and listener to obtain the stereo effect. Dedicated audiophiles experiment with this basic arrangement to maximize the size of the "sweet spot". Below about 150 Hz the resonant modes of the room tend to become dominant and annoying. Various electronic or mechanical remedies are applied to tame the loudspeaker and room interaction. With the typical loudspeaker and room combination we have come to a plateau in sound reproduction accuracy. We have seldom realized what is possible with 2-channel stereo, because loudspeaker and room are not optimally designed for each other.

Two-channel stereo sound reproduction is different from Ambisonics or Wave-field Synthesis. The goal in those systems is to accurately recreate the sound field that was sampled by the microphones. In Ambisonics the sound field is reconstructed for one point in space. The reproduction errors are small in a certain spatial volume around this point. A listener with his head inside the volume will perceive the originally recorded acoustic scene. In Wave-field Synthesis the aim is to accurately reproduce the original sound field on a closed surface in space. Again, the listeners are inside the volume defined by the surface and presented with the original acoustic scene. In both cases many more than two loudspeakers are required.

What we hear is basically what our ear-brain-memory perceptual apparatus can reconstruct from the air pressure signals at the two ears. In the case of 2-channel stereo we try to create a believable illusion of nonexistent sources. The creation of auditory phantom sources depends upon signals from the two loudspeakers and their reflections in the room. The interaction between loudspeakers and room must be controlled in such a way that the ear/brain receives necessary and sufficient cues for producing the auditory illusion. Distracting cues must be minimized.

Ambiphonics is a system that cancels the potentially distracting cross-talk signals from left and right loudspeakers at each ear. It works effectively for a precise listener placement and over a small volume

around its head. Room reflections, though, present misleading cues at other locations since they contain the cancellation signals.

In the following we will briefly recall known hearing processes that relate to a listener with two loudspeakers in a room. The necessary loudspeaker characteristics will be explained and how to obtain them for low, mid and high frequency reproduction. Loudspeakers and listener must be placed in a specified relation to the room boundaries. The room should have certain reflective properties.

Under those conditions sound reproduction can be accurate in translating an electrical signal stream into an auditory experience without distractions from loudspeakers and room, but it may not sound live. For that to occur the recording must have captured appropriate signals and cues. A suitable recording process will be discussed in detail in the second part of this paper [1].

2.　　HEARING

Various factors determine an auditory experience. They are described by terms like ITD, ILD. HRTF, head rotation, precedence effect, auditory stream analysis, etc. [2, 3, 4, 5]. The auditory experience is subject to attention, expectation, suggestion, memory, cognition, preference, emotion, etc. Together they shape a subjective auditory experience of the acoustic event [6, 7].

The performance of Beethoven's 9th symphony in a concert hall is an acoustic event. The sound pressure at any point in the hall is the sum of direct sounds from many instruments and of a multitude of slowly decaying sound reflections from various other directions. As I sit in my favorite seat and close my eyes I have an auditory experience from hearing the two streams of pressure variations that impinge upon left and right eardrums. Figure 1.
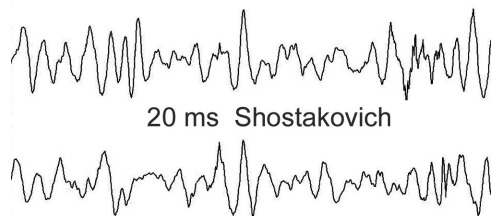
Figure 1: Left and right ear air pressure signals as primary inputs for an auditory experience. A significant amount of information about direction, distance, size and meaning of a sound source is gathered within 50 ms.

Hearing is an amazingly efficient perceptual process that evolved for survival and communication. Direction, distance, size and meaning of multiple sound sources are rapidly detected, tracked or ignored in the two streams of sound. The auditory scene that I perceive from my seat in the hall is subject to those processes and to my memory of similar scenes, and to my emotions, and to the visual scenery. Together they assure the reality of the perceived event [8, 9, 10, 11, 12].

A recording of my ear signals, when played back over two loudspeakers in my living room, will not automatically recreate the auditory experience from the concert hall. There are now only two physical sources of sound in my room. They produce a set of reflections from the surfaces in the room. The reflections decay faster than in the concert hall. The sound field of the concert hall is not replicated at any point in my living room. The two streams of pressure variation that impinge upon left and right ear are different from those that were recorded. Yet these streams carry sufficient cues from which my ear-brain apparatus is able to construct an auditory illusion that has many similarities to hearing Beethoven's 9th symphony from my favorite seat in the hall. The auditory event, though, is perceived as being an illusion.

Our hearing has evolved to detect and track real sources in different environments and not to be misled by illusions. Therefore, to strengthen the illusion of hearing an orchestra in a concert hall, the two loudspeakers must not draw attention upon themselves. Likewise the listening room, which reflects and modifies the sounds from the loudspeakers must not draw attention upon

itself. If the indirect and attenuated sounds from the room are sufficiently delayed and carry no new information as to the location of the phantom sources, then the ear/brain perceptual apparatus can safely withdraw attention from them. Reflections are not ignored, though, and are important for externalizing the phantom sources created by the direct sound streams from the two loudspeakers in the living room.

For example, identical pink noise signals fed to left and right loudspeakers and listened to at equal distances from both loudspeakers will create a phantom noise source midway between the two loudspeakers. A fuzzy ball of noise is perceived in the plane of the loudspeakers. If the room were reflection-free, i.e. anechoic, then the noise would have been perceived inside the head. Here the hearing apparatus has been presented with a confusing situation. Evolutionary speaking, we are unfamiliar with being equal-distant from dual mono sources. We localize them as a single source inside the head. To resolve the ambiguity requires additional information. Room reflections provide the cues that the source is at some distance and it is therefore externalized. Highly directional loudspeakers in a very absorptive room sometimes produce a center phantom image that is in front of the loudspeakers.

If, in the example, the loudspeakers and listener had formed an equilateral triangle with a 2.4 m base and the listener moves his head about 80 mm to the right, then the timbre of the pink noise becomes darker. In this position the left ear is about equidistant from both loudspeakers, but the right ear is closer to the right loudspeaker and further away from the left one. The path length from the left loudspeaker to the right ear has increased, to the right loudspeaker it has decreased. The difference is about $\lambda/2$ at 2 kHz, which leads to a notch in the pink noise spectrum and to signal cancellation at the right ear [13]. The phantom source has shifted towards the right loudspeaker. It rapidly collapses into the right speaker upon further head shift to the right. The ear/brain perceptual apparatus has now detected the location of a physical source, which appears to have caused the phantom noise source in the room. The left loudspeaker output becomes sufficiently delayed to merely add to the volume and spaciousness of the right loudspeaker's sound.

It is important to note that turning the head towards the left or right loudspeaker, when seated at the apex of the equilateral triangle, does not move the phantom source

from the midway position, nor does it change its timbre. The ear signals change but the ear/brain perceptual apparatus compensates for it [14].

The perception changes when the pink noise sources that drive left and right loudspeakers are independent from each other. A diffuse cloud of sound is formed that fills the room. The loudspeakers are difficult to localize even with large head movements. If the loudspeakers have a frequency independent polar response, then the timbre of the noise will not change when listened to from different locations in the room.

Pink noise is a revealing test signal for detecting differences, but it is not known what it actually should sound like because of its unfamiliar nature. The sounds of human voices, acoustic instruments, animals, machines, traffic, bells, thunder, etc. in their normal environments are very familiar to us. When recorded and played back over two loudspeakers in a room the sound streams at the listener's ears usually carry the necessary cues for the ear/brain to create a recognizable auditory scene. Unlike with pink noise a centered female voice will not sound different with sideways head movement, but the phantom person will shift position.

Recordings must capture both, the sound source and the response of the environment or venue to be recognizable and believable. The brain appears more than willing to fill in missing pieces to bring realism to the auditory experience. That process is made easier if distracting cues from loudspeakers and room are minimized. Under those conditions the size of the phantom image can even be taller than the loudspeakers and go much deeper than the distance to the wall behind the loudspeakers. Whether recorded outdoors, in a concert hall or a small club, the spaciousness of the venue becomes clearly audible. It also becomes audible and is perceived as unnatural when the recording is a collage of multiple sub-spaces around instruments and voices. It can be particularly annoying when the acoustic perspective slides during a performance to bring a particular instrument closer and then the scene fades back to the previous perspective. Furthermore, the width of the phantom image can be wider than the distance between the loudspeakers and without a hard edge, but this depends strongly upon the recording technique used. Phantom sources appear essentially behind the plane formed by the two loudspeakers. This is not surround sound, but the strong spatial rendering in front of the listener allows for a convincing auditory experience when combined with near realistic volume levels, and heard with eyes closed.

3. LOUDSPEAKERS

Loudspeakers will draw attention upon themselves when their frequency response and output volume capabilities are inadequate to reproduce the recorded acoustic scene. The on-axis frequency response is most important. It must be flat and without hidden resonance. Usually overlooked in their importance are the horizontal and vertical polar frequency responses. They define how the loudspeaker illuminates the room with sound, which then contributes to what we hear [15]. The horizontal frequency response should be frequency independent and variations in the vertical response should be kept to a minimum. This requires acoustically small and closely spaced source elements. Room reflections play an important role in phantom source placement and out-of-head localization. Frequency independent room illumination requires that a loudspeaker is either omnidirectional or uniformly directional at all frequencies. Typical box loudspeakers are omnidirectional at low frequencies, become completely forward firing at high frequencies and often do not change monotonically in directionality in-between. This behavior is undesirable because room illumination is not uniform for all frequencies.

Omnidirectional loudspeakers can be built fairly easily for uniform response up to a few kHz. Since they radiate in all directions they also generate a maximum number of room reflections. Thus the auditory experience becomes more dependent upon the room's reflective properties than it would be for a directional loudspeaker. Omnidirectional loudspeakers should be listened to from close distance.

A dipole presents the simplest concept for building a loudspeaker that is uniformly directional over the complete range of frequencies, from bass to tweeter. A cardioid loudspeaker poses some added difficulties. We will therefore concentrate here on the dipole loudspeaker and illustrate practical approaches to obtain the desired low, mid and high frequency dipolar behavior.

The output volume capability of a loudspeaker is important because our hearing mechanism is non-linear. The distortions it creates are part of normal hearing. It will only sound realistic, when the sound pressure at the

ear drum reaches near realistic levels. The well-known equal loudness contours (Fletcher-Munson) describe the perceptual change with sound pressure level and frequency. In particular at low frequencies sufficient pressure levels must be generated to contribute properly to timbre balance. But, if the loudspeaker generates too much signal distortion, then it will merely sound loud and unpleasant but not realistic. Loudspeakers must be designed to handle large signals adequately.

3.1. Dipole Model

An ideal dipole consists of two point sources of equal strength and opposite polarity [16]. The sources are separated by a distance D. Figure 2. The magnitude of the frequency response at distances where the rays from the two sources are essentially parallel is given by:

$$H = 2 \sin(\pi\, D/\lambda\, \cos\alpha) \tag{1}$$

where α = off-axis angle, D = polar distance,
λ = wavelength = v / F, with v = 340 m/s and F in Hz.

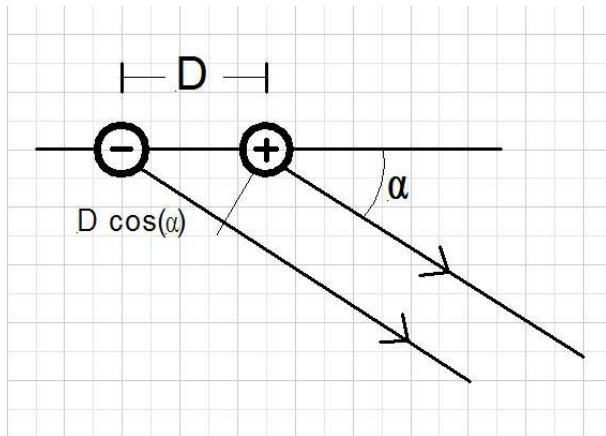Figure 2: Ideal dipole model. The two sources are separated by a distance D. The summed output is zero for $\alpha = 90^0$.
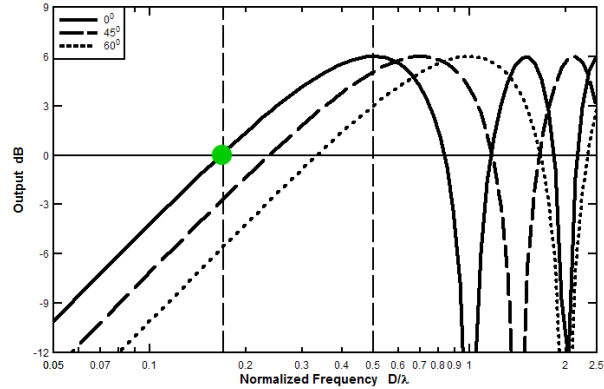

Figure 3: Frequency response of the ideal dipole at 0, 45 and 60 degrees off-axis. The output equals that of a single source at F_{equal} where $D/\lambda = 0.17$.

Figure 3 shows a plot of (1) on log-log scale. The frequency axis is in D/λ normalized units. The response rises from low frequencies at a rate of 6 dB/octave. At $D/\lambda = 1/6 = 0.17$ the magnitude of the on-axis response equals that of a single source. The corresponding frequency is:

$$F_{equal} = v/(6\,D) \tag{2}$$

F_{equal} is of practical value for estimating the required volume displacement for reproducing the lowest frequencies.

At $D/\lambda = 0.5$ and $\alpha = 0^0$ where the two sources are a half-wavelength apart, their outputs add maximally for a 6 dB peak in the frequency response. At 45^0 and 60^0 off-axis angles the response peak occurs at higher frequencies. The response is 3 dB and 6 dB down at low frequencies for 45^0 and 60^0 angles.

At $D/\lambda = 1.0$ and $\alpha = 0^0$ where the two sources are one wavelength apart, their outputs subtract. The pattern of addition and subtraction repeats as frequency increases further. The peaks and nulls occur at progressively higher frequencies as the off-axis angle α increases.

Figure 4 shows the polar response for different normalized frequencies D/λ. The initial figure-of-eight pattern widens with increasing D/λ and the on-axis magnitude increases to +6 dB. Above this frequency the

on-axis response drops and reaches a null at D = λ. For all frequencies there is a null at α = 90^0 because the two sources are equidistant from the observer. The widening in polar pattern is useful as it counteract the increasing directivity with frequency of piston electro-dynamic drivers. Thus drivers with larger diameters can be used than would be normally chosen for conventional box loudspeakers.
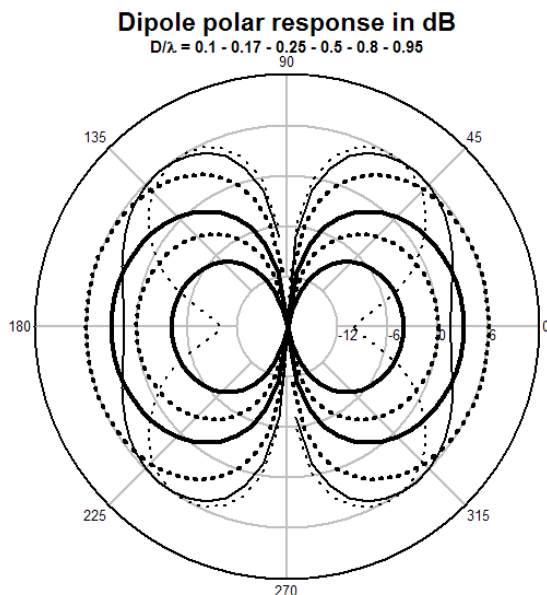
Dipole polar response in dB
D/λ = 0.1 - 0.17 - 0.25 - 0.5 - 0.8 - 0.95

Figure 4: Theoretical polar response increasing in amplitude and widening with increasing normalized frequency D/λ.

Practical dipole implementations follow the model very closely at low frequencies where the physical distance between the two sources is small compared to the radiated wavelength.

$$H = 2\pi F\, D/v\, \cos\alpha \quad \text{for } D << \lambda/\pi \qquad (3)$$

At higher frequencies the widening in the polar response may be noticeable but the nulls are not. In practical structures the complete cancellation of on-axis sound is prevented by increasing radiator size and the variation in path length D. There remains though the null at 90^0 for symmetrical (+) and (-) source arrangements.

The on-axis sound pressure level SPL_{dipole} that is generated by a dipole source under free-space conditions and referred back to 1 m distance from the source can be estimated from:

$$SPL_{dipole} = SPL_{mono} - 20 \log(F_{equal} / F) \qquad (4)$$

for $F < F_{equal}$ (4) where

$$SPL_{mono} = -37.6 + 20 \log(S_d) + 20 \log(X_{max}) + 40 \log(F) \qquad (5)$$

with SPL in dB, piston area S_d in cm^2, peak excursion X_{max} in mm, frequency F in Hz

3.2. Dipole Woofer Design

While it is easy at lower frequencies and longer wavelengths to obtain perfect dipole behavior it also becomes increasingly difficult to generate high sound pressure levels. The path length D should be as long as possible but it is limited by how high in frequency the woofer will be used. The other limitation is the maximum linear volume displacement that can be obtained from the chosen drivers. Rather than stacking two sealed box woofers at a distance D behind each other, as the ideal case might suggest, it is more efficient to use an open baffle. The H-frame configuration provides a simple solution and has proven itself [17]. Figure 5.

Two drivers are used for increased output volume and to reduce distortion. The two pistons move in tandem in the same direction relative to the H-frame. Push-pull mounting of the drivers reduces even order distortion by more than 10 dB.

Output volume can be estimated. Assuming that the effective source separation is D = 305 mm (12 inch) then from (2) above: F_{equal} = 189 Hz.

Assuming furthermore a total piston area S_d = 1000 cm^2 and X_{max} = 10 mm, then from (4) and (5) above and at a frequency of 50 Hz:

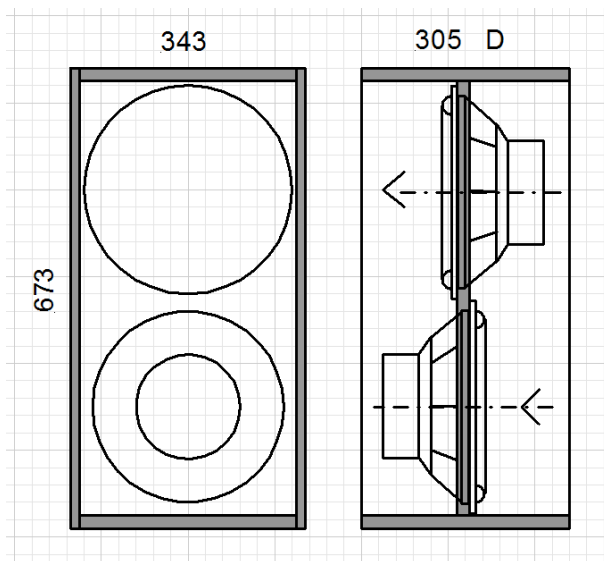$$SPL_{mono} = 110 \text{ dB and } SPL_{dipole} = 98 \text{ dB}$$

Figure 5: H-frame dipole woofer construction. Drivers are configured for even order distortion reduction.

In this example the open baffle output at 50 Hz is 12 dB lower than that from a closed baffle driver configuration with the same volume displacement.

The useable frequency range of the H-frame woofer is limited at the high end by a resonance. It is created by a mismatch between the impedance of the acoustic waveguide and its opening. It occurs when $D/2 = \lambda/4$. Figure 6. The resonant peak can be equalized with a notch filter. Varying the contour of the H-frame opening between upper and lower drive units can also flatten the peak.
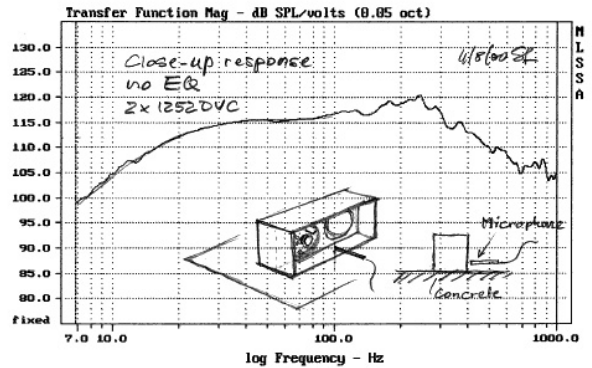

Figure 6: H-frame dipole woofer frequency response measurement in the opening plane of the baffle.

3.2.1. Woofer Response Equalization

Dipole woofer response equalization by boosting lower frequencies is easily accomplished with an active circuit. It is a first order RC lowpass filter with gain. In the example above, the dipole woofer drive signal had to be boosted at 6 dB/octave rate to obtain the 98 dB SPL and X_{max} of 10 mm. The amount of boost was $(F_{equal} / F) = 3.8$ times or 11.6 dB at 50 Hz. Typically it does not require much power from the amplifier at the lowest frequencies to generate the maximum displacement that a driver is capable of. Operation will be near the driver's mechanical resonance where the impedance rises sharply. It does require large voltage swings from the amplifier. The maximum output from a dipole woofer is usually displacement limited at the low end and amplifier limited at the high end of its frequency range. At the low end it is preferable that the amplifier output voltage clips before the drivers hit their mechanical stops to prevent damage to the drivers. Since the woofer is usually positioned on the floor one can add 6 dB to the SPL estimates for half-space radiation. Still, below 50 Hz, it may become necessary in some situations to cross over to a sealed box subwoofer to keep the dipole woofer size reasonable.

Instead of using electronic equalization it is also possible to obtain some degree of correction towards a flat frequency response by using low frequency drivers with a weak motor and high Q_0. Figure 7. Below F_0 the frequency response rolls off initially at greater than 18 dB/octave. This approach uses stored energy to boost low frequencies and is similar in that respect to vented

enclosures. Neither design is desirable for building an accurate loudspeaker because it stretches the time response and distorts the input signal. The increase in response above F_0 can be corrected in the crossover to the midrange unit.
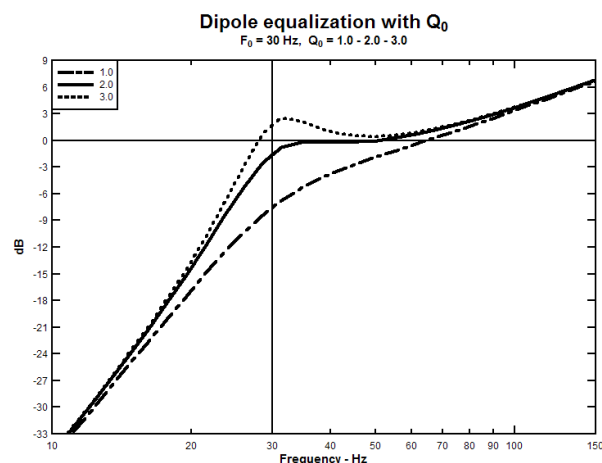

Figure 7: Passive equalization of a dipole woofer by using a driver with $Q_0 = 2.0$.

Using drivers with $Q_0 = 0.7$ and equalizing them with a first order lowpass of corner frequency F_0 yields a response that rolls of at 18 dB/octave below F_0. A slower roll-off rate is desirable to reduce the group-delay accumulation and to increase the realism of the lowest bass frequencies. If Q_0 is less than 0.5 and equalization is carried out to F_0, then the roll-off below F_0 starts at 12 dB/octave rate. Depending on Q_0 the turn into an 18 dB/octave slope can occur as low as 5Hz.

Other forms of dipole woofer construction are certainly feasible. The W-frame provides force cancellation and thus vibration cancellation in a very compact form factor, albeit in a more complex housing [18].

3.3.   Dipole Midrange Design

A rectangular baffle with a single driver mounted on the axis of symmetry works well for a dipole midrange. Figure 8. The baffle may be folded for structural rigidity, but the wings must be kept short in order not to create cavities that store acoustic energy.
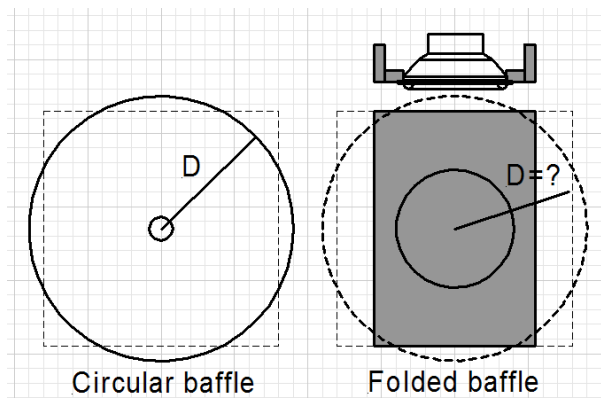

Figure 8: The path length D is closely defined for a dipole point source in a circular baffle. It can only be estimated for a folded baffle with a real driver.

The mid-frequency portion of a dipole loudspeaker usually has three ranges with different behavior due to driver and baffle size. First, at low frequencies, the behavior follows the ideal dipole. Here the driver's volume displacement and the baffle shape, or path length D, determine how low a crossover frequency to the woofer can be chosen.

Next up in frequency is the transition range to the $D/\lambda > 0.5$ region. Usually the slope becomes steeper than 6 dB/octave before the frequency response turns towards being flat. The increase in slope is thought to be due to the asymmetry of front and rear radiation and to the driver becoming more directional. The rear radiation response contains a lowpass component due to obstruction from basket and magnet. Figure 9.

Third, the nulls at and above $D/\lambda = 0.5$ are not observed in practice. Even for a circular disk baffle with the driver mounted in its center, not enough of the rear radiation diffracts around the baffle edge to perfectly cancel the front on-axis radiation. Drivers with adequate lower frequency SPL capability are simply too large to radiate uniformly into front and rear hemispheres over a +/- 180^0 angle. This can be seen from the general polar diagram for a driver with sealed back. Figure 10.
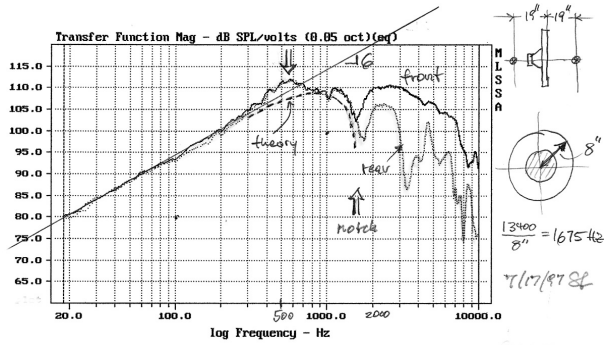
Figure 9: Example of front and rear frequency responses for a driver on a circular disk baffle.

Remnants of the dipole nulls though can be observed depending on the baffle contour. The circular baffle is a worst case. Figure 9
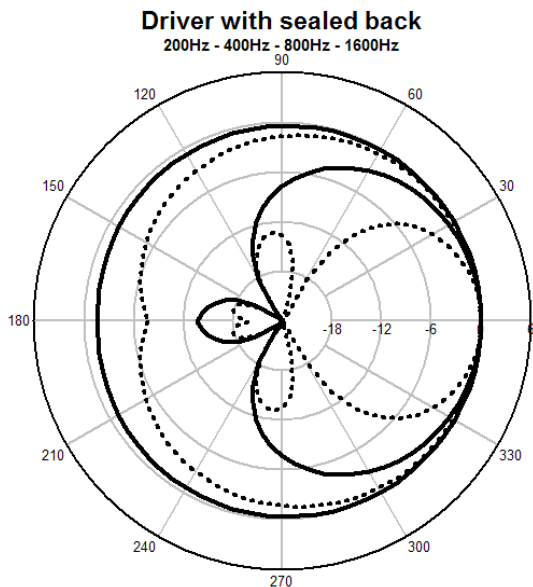

Figure 10: Polar response example for a driver with a sealed back. The response follows a sin(x)/x function because of the acoustically increasing piston size with frequency.

Equalization of the dipole roll-off and transition to flat response requires more than a shelving lowpass filter response A. Figure11. A bridged-T network response B can provide the increased slope rate and a notch. More flexibility in adjustment is obtained from a combination of shelf and notch filter response C [19].
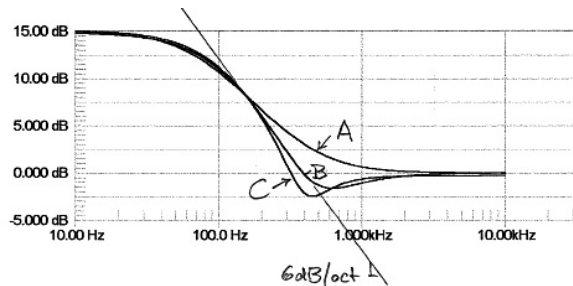

Figure 11: Midrange dipole equalization with active electronics to provide gain and match driver sensitivities.

3.4. Dipole Tweeter Design

To obtain dipole radiation in the kHz frequency range can pose problems because there are few suitable high frequency driver architectures. Planar or ribbon drivers that are open in the back tend to have either insufficient volume displacement and high distortion or they beam at higher frequencies due to their acoustically large size. The transition from midrange to tweeter must occur at a sufficiently low frequency where the midrange driver's polar response on the open baffle is still well behaved and blends with the tweeter's response. Two back-to-back dome tweeters on a baffle of the same width as that of the midrange can meet the polar response and output demands [20]. Figure 12. The baffle is typically too wide for the rear tweeter to affect the on-axis response of the front tweeter. The rear tweeter influences the frontal hemisphere radiation only at large off-axis angles. As frequency increases front and rear tweeters become so directional that they do not interact. At all frequencies this generates a dipole like radiation pattern. It has a cancellation notch at about 90^0 off-axis that widens with increasing frequency.

A combination of the described woofer, midrange and tweeter dipole elements can give an acoustically small dipole source that has the necessary output volume capability for realistic sound reproduction.
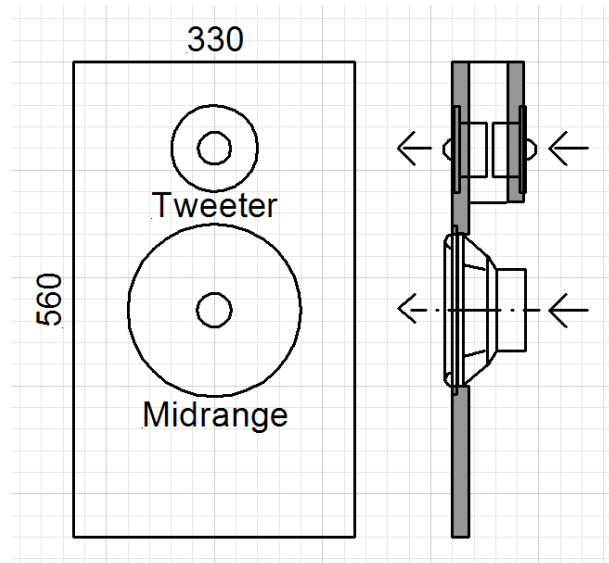

Figure 12: Dipole tweeter configuration to blend with the polar frequency response of the open-baffle midrange driver.

low frequencies. The loudspeaker is omnidirectional. With increasing frequency it becomes progressively more directional. Figure 14.
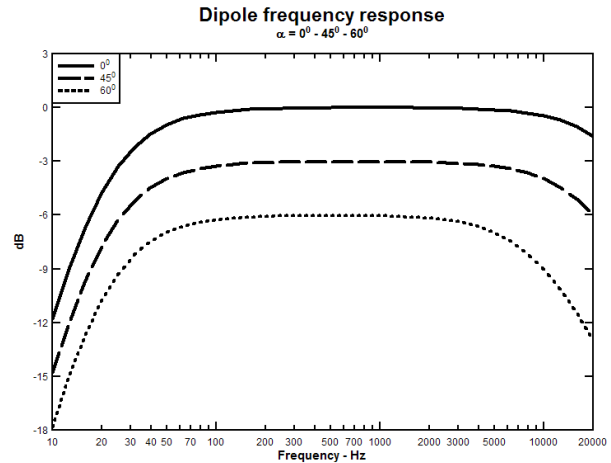

Figure 13: Free-field dipole frequency response at 0, 45, and 60 degrees off-axis. The response shows constant directivity at low frequencies, which increases slowly towards high frequencies.

3.5. Compared to other Architectures

Traditional planar dipole loudspeakers suffer from insufficient volume displacement and large dimensions. The result is inadequate bass output, uneven frequency response and multi-directional high frequency radiation when radiating planar surfaces become acoustically too large. The irregular polar response makes optimum room placement difficult. These loudspeakers, though, show the benefits of dipole radiation and clearly differentiate themselves from conventional box loudspeakers and this despite their lack of dynamic range.

The general difference between open baffle and box loudspeakers can be seen in the off-axis frequency response. The open-baffle or dipole response decreases uniformly from its on-axis response with increasing off-axis angles α. Figure 13. The box loudspeaker frequency response, though, is independent of angle at
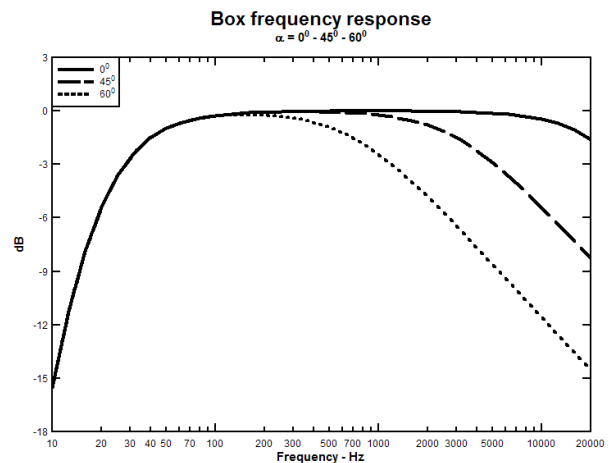

Figure 14: Free-field box loudspeaker response at 0, 45 and 60 degrees off-axis. The response is omnidirectional at low frequencies and becomes increasingly directional at high frequencies.

The difference becomes even more visible when integrating the radiated sound pressure over a spherical surface with the loudspeaker in the center. When both types have the same on-axis frequency response, then the total power radiated by the box loudspeaker is 4.8 dB higher at low frequencies. Towards high frequencies the power response rolls off much sooner and becomes ultimately 3 dB lower than that of the dipole due to a lack of tweeter rear radiation. Figure 15.
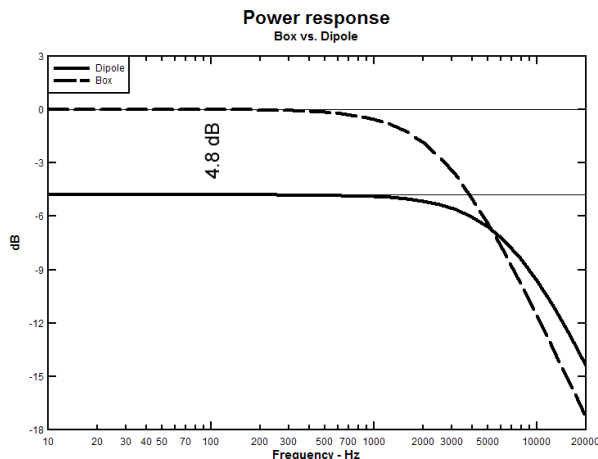

Figure 15: Free-field power response of dipole and box loudspeaker at the same on-axis sound pressure level. The box has 4.8 dB higher power radiation at low frequencies and rolls off sooner towards high frequencies for the same size drivers as the dipole.

4.   ROOMS

In domestic size rooms a loudspeaker's polar response determines the spectral energy distribution. Reflections and room modes can have a large influence upon tonal balance and phantom image creation. In large venues or outdoors the direct sound from the loudspeaker is most important. There the polar response is shaped to cover an audience with maximum sound volume and efficiency. Acoustic interference and phantom image formation must be avoided.

When listening to two loudspeakers in a room the ears receive the sum of the direct sound signals from the loudspeakers and from the multitude of reflections and room modes that were stimulated by the loudspeakers. Evolutionary forces have shaped the hearing apparatus such that it continuously analyzes the sound streams at the ears for signals of interest and of requiring attention. It can separate these out to some degree from the streams of sounds that hold no interest and are non-threatening. They become the background and context in which the interesting stuff happens.

Loudspeakers with a frequency independent polar response will generate reflections that have essentially the same spectral content as the direct sound. The reflections are attenuated and delayed copies of the direct sound. They contain information about the reflective, diffusive and absorptive characteristics of the room. This is not new information to the listener who carries already an acoustic impression of the room. It is formed very quickly from various sounds and noises and pre-consciously upon entering the room or from having spent time in it.

The perceptual separation of direct and reflected sounds begins to occur when the reflections are delayed by more than 5 ms as described by the precedence effect. Furthermore, when the reflections are attenuated copies of the direct sound and carry known information about the listening room, the hearing processes can safely suppress them from the auditory experience. For this effect to occur the loudspeakers must be placed properly in the room. They should be located at least 1 m away from the nearest wall. A uniform horizontal spread of the auditory scene is assured by placing the loudspeaker and listener triangle symmetrical with respect to the acoustical room boundaries. Figure 16. The room itself should have acoustic properties that are not overly absorbent to maintain a uniform spectral balance. A reverberation time between 400 ms and 600 ms is appropriate for dipole loudspeakers. No special room treatment is required, just the normal "stuff of life". Strong discrete reflections that affect phantom image formation should be treated with diffusive elements rather than absorptive ones to maintain the spectral balance of the reverberant sound.

It has been observed that dipole woofers reproduce bass more realistically than box loudspeakers for reasons that are not fully understood. They also appear to generate fewer problems with exaggerated bass due to room modes [21].
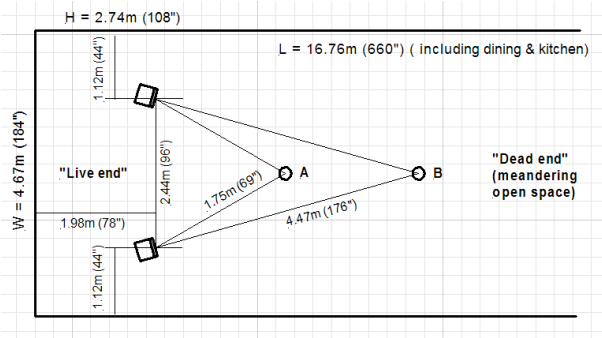
Figure 16: Symmetrical setup of loudspeakers and listener relative to the room boundaries. A minimum distance of 1 m from the tweeter to large reflecting surfaces should be maintained.

5.   RECORDING

The described dipole loudspeaker-listener triangle and room setup reveals much about the recording technique used in the production of a CD. With masking effects of the listening room removed one can clearly hear the spatial context in which the recording took place. This can be an extremely enjoyable auditory experience as it adds realism to the perceived phantom sources. It can also be distracting if the recording is a collage of multiple spaces that do not form a spatial continuum. Disappointingly this is the case for a majority of recordings. The close-up acoustic perspective that was used for the recording becomes obvious. In almost all cases it is not a perspective that the listener would ever have experienced, but an artificial creation that apparently is meant to serve specific purposes.

A question has been raised and answered. Is it even possible to record in such a way that it sounds live when reproduced over two loudspeakers in a domestic room? Live in the sense that a person would have heard something very similar at the recording venue. To develop the appropriate technique a recording engineer would be well served to use the described dipole loudspeaker and playback setup for monitoring and mix-down. It could be argued that by doing so the recording engineer now produces recordings that sound realistic only under the specified reproduction conditions [12]. In practice, however, this would not be a limitation. Under these conditions the engineer's hearing apparatus receives accurate cues for recreating a

known auditory scene. Those cues have been imbedded in the recording. Thus when played back over different loudspeakers and in different rooms, the cues are still available to be perceived. They may be masked, though, to varying degree by the interaction of these loudspeakers with the room via a non-uniform radiation pattern or distorted in other ways. In such cases the sound reproduced by the loudspeakers would fall short of creating the full range of phantom images and auditory illusions that are possible. The same would happen with any recording. The potential for sounding live can be imbedded in the recording. How much of it is recovered has always been and will always be dependent upon the consumer's commitment to playback accuracy. It is the recording engineer's responsibility to prevent his own monitor system from becoming the limiting factor.

A novel recording process will be described in a companion paper [1]. Instrumental to that process was the described monitor loudspeaker. It served as a reliable tool for evaluating the microphone selection, their setup and the mix-down of tracks. Figure 17.

6.   CONCLUSIONS

The auditory experience that is created by listening to two loudspeakers in a room depends primarily upon the sound pressure stream at each ear. These streams are not an accurate replica of the sound that was picked up by the recording microphones. Cross-talk between the ear signals and reverberation in the room degrade the pressure streams. The ear/brain hearing apparatus can recover information to a surprising degree and create believable phantom sources. This process is made easier when the room's response to the loudspeakers is a delayed and attenuated copy of the first arriving sound at the ears.

Loudspeakers can be built that have the necessary frequency-independent polar response. In their simplest form these are either omnidirectional or dipole loudspeakers. The omnidirectional source is less desirable because of the larger number of possible sound reflections in the room.

Uniformity of the polar response requires sources that are acoustically small. Dipolar woofer, midrange and tweeter configurations based on piston electro-dynamic drivers can generate realistic sound pressures at adequately low non-linear distortion levels.
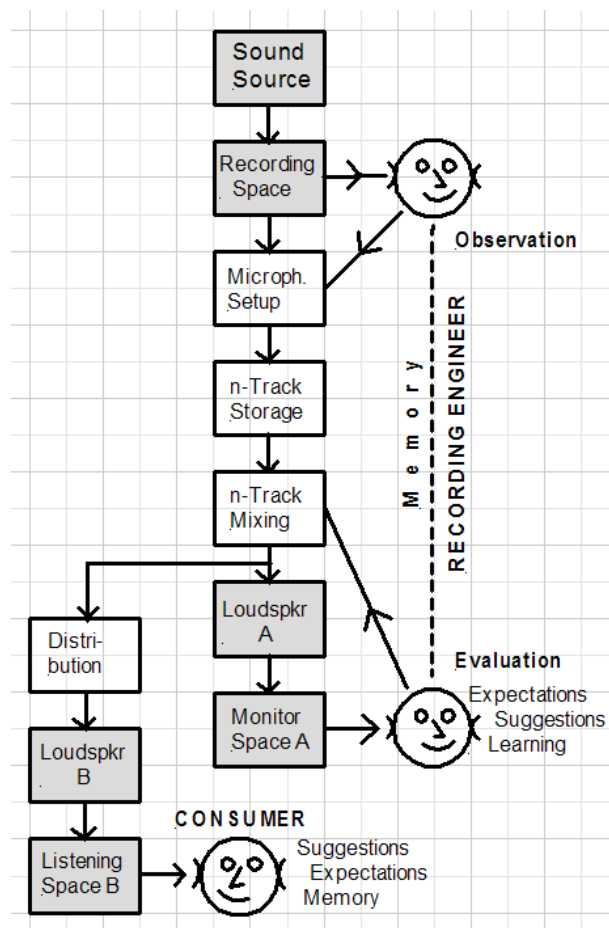
Figure 17: Processing steps in recording and reproduction. Accurate monitor loudspeakers A are essential for breaking the "circle of confusion" [13].

A dipole source requires large amounts of equalization to compensate its acoustic short-circuit and inherently non-flat frequency response. Such equalization is best realized with active electronic circuits. The loudspeaker then becomes an active system with a power amplifier for each driver and with a low-level electronic crossover/equalizer. This approach minimizes individual amplifier power requirements and allows combining drive units of different sensitivities.

The loudspeakers must be set up symmetrically in the room with the tweeter at least 1 m distant from adjacent walls. This is a hindrance in many situations. It is not a new requirement, though, and the vast majority of loudspeakers in use today would benefit from it as well.

Acoustically small dipole loudspeakers reveal more of a recording and they should stimulate a revisit of recording practices. Sound recording and reproduction as heard live requires a system solution. Current practices are sub-optimal but could be easily improved.

The described dipole loudspeakers would minimize the need for room treatment and room equalization. It would become apparent that the room is rarely the problem, but the traditional loudspeaker is. Focus would shift from loudspeakers and associated equipment to the acoustic qualities of recordings. Today there are few sources for simply buying uniformly directional loudspeakers. With some motivation and effort they can be built from the ample information and resources provided on the Internet. It has happened in the last five years to universal acclaim.

7.   ACKNOWLEDGEMENTS

8.   REFERENCES

[1] Don Barringer, Siegfried Linkwitz, "Recording and Reproduction over two Loudspeakers as Heard Live – Part 2: Recording Concepts and Practices", presented at the AES 126th Convention, Munich Germany, 2009 May 7-10 (Preprint xxxx)

[2] Jens Blauert, "Spatial Hearing – The Psychophysics of Human Sound Localization", The MIT Press, 1997

[3] Albert S. Bregman, "Auditory Scene Analysis – The Perceptual Organization of Sound", The MIT Press, 1999

[4] Jon Boley, "Auditory Component Analysis", presented at the 121st Convention, San Francisco, October 5-8, 2006 (Preprint 6925)

[5] Brian C.J. Moore, "An Introduction to the Psychology of Hearing", Academic Press, 1997

[6] Daniel J. Levitin, "This Is Your Brain on Music – The Science of a Human Obsession", Dutton, 2000

[7] Oliver Sacks, "Musicophilia – Tales of Music and the Brain", Vintage Books, 2008

[8] William A. Yost, "The Cocktail Party Problem: Forty Years Later", Chapter 17 in "Binaural and Spatial Hearing in Real and Virtual Environments" R. Gilkey and T. Anderson (Eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, 1997

[9] William M. Hartmann, "Listening in a Room and the Precedence Effect", Chapter 10 in "Binaural and Spatial Hearing in Real and Virtual Environments", R. Gilkey and T. Anderson (Eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, 1997

[10] Siegfried Linkwitz, "Room Reflections Misunderstood?", presented at the AES 123rd Convention, New York, USA, 2007 October 5-8, (Preprint 7162)

[11] Barry Blesser, Linda-Ruth Salter, "Spaces Speak, Are You Listening? – Experiencing Aural Architecture", MIT Press, 2007

[12] Peter Damaske, "Acoustics and Hearing", Springer, 2008

[13] Floyd E. Toole, "Sound Reproduction – Loudspeakers and Rooms", Focal Press, 2008

[14] Guenther Theile, "On the Naturalness of Two-Channel Stereo Sound", JAES, Vol.39, No. 10, 1991 October

[15] Mark F. Davis, "Loudspeaker Systems with Optimized Wide-Listening-Area Imaging", JAES, Vol. 35, No. 11, 1987 November

[16] Leo L. Beranek, "Acoustics", McGraw-Hill, 1954

[17] H-Frame Woofer, http://www.linkwitzlab.com/proto.htm#PW1

[18] W-Frame Woofer, http://www.linkwitzlab.com/woofer.htm

[19] Midrange 6 dB/oct slope compensation, http://www.linkwitzlab.com/filters.htm#8

[20] Dipole Tweeter arrangement, http://www.linkwitzlab.com/orion++.htm

[21] Siegfried Linkwitz, "Investigation of Sound Quality Differences between Monopolar and Dipolar Woofers in Small Rooms", presented at the AES 105th Convention, San Francisco, USA, 1998 September 26-29, (Preprint 4786)