Audio Engineering Society

Convention Paper

Presented at the 133rd Convention
2012 October 26–29 San Francisco, CA, USA

A Model for Rendering Stereo Signals in the ITD-Range of Hearing

Siegfried Linkwitz

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

ABSTRACT

Live sounds at a concert have spatial relationships to each other and to their environment. The specific microphone technique used for recording the sounds, the placement and directional properties of the playback loudspeakers, and the room's response determine the signals at the listener's ears and thus the rendering of the concert recording. For the frequency range, in which Inter-aural Time Differences dominate directional hearing, a free-field transmission line model will be used to predict the placement of phantom sources between two loudspeakers. Level panning and time panning of monaural sources are investigated. Effectiveness and limitations of different microphone pairs are shown. Recording techniques can be improved by recognizing fundamental requirements for spatial rendering. Observations from a novel 4-loudspeaker configuration are presented. The setup provides enhanced spatial rendering of 2-channel sound.

1. INTRODUCTION

We hear naturally occurring sounds in a spatial context. Whether outdoors or indoors we derive information about direction, distance and size of a sound source from its interaction with the environment and our body. This allows us to differentiate between multiple simultaneous sounds, to assign meaning and to direct attention, most of it pre-consciously. The brain processes eardrum pressure variations within milliseconds to assure a proper response and to stay out of danger.

When we listen to music in the safety and comfort of a concert hall we appreciate the artistry of the musicians and the conductor in rendering a composition and we are drawn in by the response of the hall. The hall reflects the orchestra's sound waves in ever-varying time, frequency and amplitude reverberation patterns depending upon which musical instruments or groups predominate at the moment and where we are seated.

A recording should capture much of the spatial and temporal interplay between orchestra and hall, between groups of musicians and soloists as might have been heard live. This requires an understanding of how the microphone signals will be rendered in playback. For instance, an equilateral triangle loudspeaker and listener stereo setup in a room is a rendering configuration, which will produce significantly different ear signals from a given recording than a pair of headphones.

1.1. Binaural recording and rendering

The greatest fidelity in recording and rendering is obtained by working directly with the air pressure variations at the eardrums, by recording them and reproducing an exact copy, Figure 1, [1]. Playback renders HRTF cues accurately, except that the ear signals do not change when the listener turns his head. The aural scene is experienced as being outside of the head, but only when the listener's HRTF matches closely to that of the recording and associated playback equalization. Otherwise and without head turning cues, the aural scene hovers close to the ears or is experienced inside the head.

Conversely, reproduction over loudspeakers places the origin of sound at large distance from the ears. Room reflections add to identify or detract from the source. Ear signals change with movement of the head and the aural scene is externalized.

Figure 1 A recording and playback technique to generate identical sound signals at the eardrums. The same microphone tube next to the eardrum is used for recording and then to equalize the headphones.

1.2. Spatial hearing

Spatial hearing and in particular directional hearing can be described by arrival time and frequency response changes at the ears for different angles of sound incidence and by the brain's interpretation of inter-aural changes with head movement, namely by HRTF, ITD and ILD changes [2-6]. ILD dominates above 3 kHz where head diffraction and pinna shape affect the ear signals. ITD dominates at frequencies below 800 Hz where the head size is less than one-half wavelength. At long wavelengths the incident sound waves will differ primarily in time of arrival at the ears. Amplitude differences due to diffraction will be insignificant.

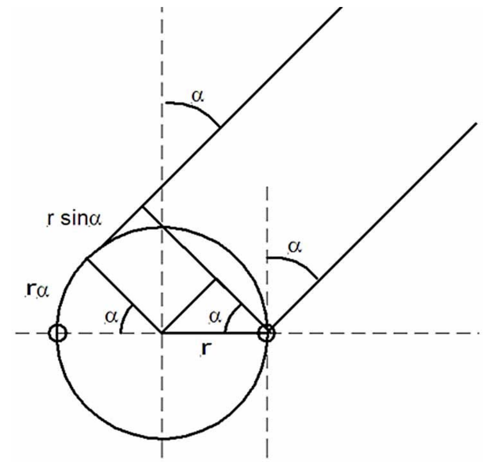


Figure 2 Derivation of inter-aural time delay for a solid sphere with sound incident at angle α.

With these simplifications numerical ITD values can be derived from a solid sphere model of the human head, Figure 2. With the speed of sound c and the angle of sound incidence α the inter-aural time delay becomes:

$$ITD = (r/c)(\alpha + \sin\alpha) \quad (1)$$

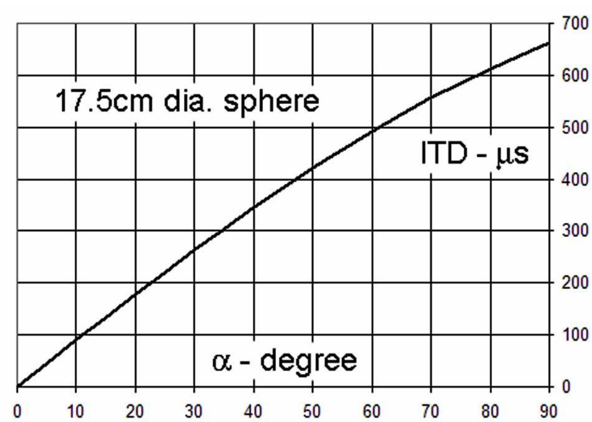


Figure 3 Calculated low frequency ITD for a solid sphere as function of sound incidence angle α.

The delay is 263 µs for a sphere of 17.5 cm diameter and 30^0 sound incidence, Figure 3. The delay of 263 µs will be used in the rendering model of Figure 6 for an analysis of phantom source placement between two loudspeakers, when the direct sound arrives from $+/-30^0$ at the listener's ears. The principles involved with level panning and time panning of a mono signal and with the spatial rendering of stereo signals will be explained.

Rendering of high frequency signals, in the ILD range of hearing, follows more intricate principles. Those signals pull towards left or right loudspeakers, unless the loudspeakers have wide dispersion. ITD in combination with head turning confirmation tends to dominate phantom source positioning on program material.

2. RENDERING IN A REVERBERANT SPACE

How a recording is rendered, the aural scene created in the listener's mind, depends upon the characteristics of the loudspeakers used, their placement relative to large surfaces in the room, the room's acoustic properties and the listener-and-loudspeaker configuration.

2.1. Live vs. amplified voice

When recorded anechoic voice is played back over a single loudspeaker we usually recognize that it is not live. This might be due to volume level and acoustic quality. Volume can be easily adjusted to a realistic level. Misleading aural cues from frequency response and radiation pattern of the loudspeaker can be minimized. Also loudspeaker cabinet resonances and driver non-linear distortion must be kept low. The off-axis frequency response is the most difficult parameter to control. It determines where and how the room responds to the loudspeaker. If the response differs from that of a real human voice, then we know it was a loudspeaker. We readily recognize the nature of human voice in different surroundings.

2.2. Withdrawing attention from room and loudspeakers

Radiation from a loudspeaker inevitably causes reflections, reverberation and room modes, Figure 4. If room surfaces are highly absorptive, then the indirect sounds will decay quickly. Mixing studios are built acoustically dead for analytical work. They are not considered to be ideal places for music enjoyment. Home listening environments are usually more live, which means that reflections, modes and reverberation may interfere with the rendering of the aural scene, if inappropriately excited by the loudspeakers.

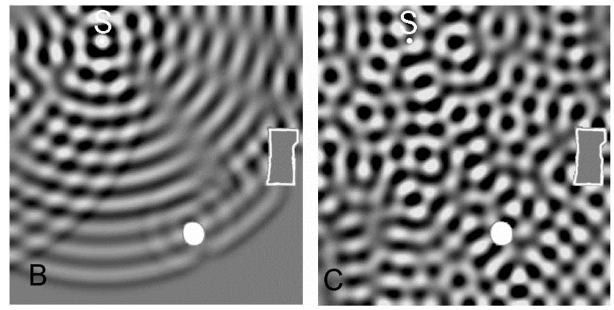


Figure 4 A single source S and two objects in a rectangular room. Initial wall reflections and diffraction by the objects (B), which over time lead to the reverberant sound field (C)

We have found that loudspeakers should be placed at least 1 m from large reflecting surfaces. Reflections are then delayed by at least 6 ms, which allocates them perceptually to sound streams that are separate from the direct loudspeaker sound. Furthermore, if the loudspeaker has a frequency independent polar pattern, such as omni-directional (monopole), bi-directional (dipole) or cardioid, then the delayed room reflections are essentially copies of the direct sound or modified in a familiar way. Under these conditions the brain can withdraw attention from the room, which adds redundant and not misleading information to the direct sound, and focus on the information contained in the direct sound [7].

A single loudspeaker is difficult to hide aurally, but when two loudspeakers produce phantom sources, then attention can be drawn away from them. Again, their polar response is critical, assuming that non-linear effects, energy storage and cabinet diffraction have been minimized.

2.3. Optimum listening distance

The direct sound level from a loudspeaker decreases inversely with distance, Figure 5. The reverberant sound level in the room builds up and reaches equilibrium between supplied and dissipated acoustic energy. Direct and reverberant sound levels become equal at the Reverberation Distance [8]. An increase of loudspeaker directivity or a decrease in room reverberation time lead to a larger reverberation distance and thus to a larger ratio of direct to reverberant sound level.

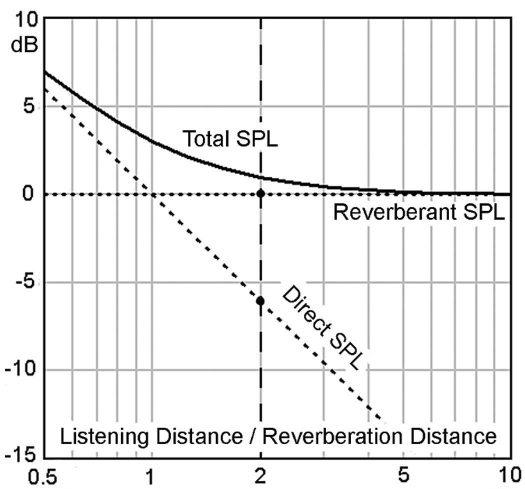


Figure 5 Sound pressure level distribution in a room when normalized to the reverberation distance

We have found that the maximum listening distance should not exceed twice the reverberation distance for optimum spatial imaging. This ensures that the room reverberated sound level masks the direct sound by less than 6 dB and can be ignored perceptually in favor of the direct sound.

A typical living room with 80 m^3 internal volume and 400 ms reverberation time, in the midrange above the 135 Hz Schroeder Frequency, has a reverberation distance of 0.8 m. This assumes a monopole or omni-directional loudspeaker as source A dipole loudspeaker increases this to 1.4 m giving an optimum listening distance of less than 2.8 m for greater than 6 dB direct-to-reverberant sound level.

2.4. The living room is not a concert hall

Reverberation times are much shorter in a living room than a concert hall and first reflections arrive much sooner. Reflected and reverberated sound in the living room should not be heard as such but room modes can make it difficult to render low frequency sounds naturally. The aural scene should only be derived from cues in the direct sound of the loudspeakers, including cues about the spatial distribution of the acoustic scene and its environment. Musical instruments differ in their directivity. They excite the hall’s response to varying degrees. Reverberation strength and timbre change with orchestration and that should also be heard in the living room [9-11]. When listening to a recording one should be able to answer the question: Where am I?

3. RENDERING MONO SIGNALS

Monaural signals are typically pan-pot mixed by ear to a position between left and right loudspeakers. Left and right loudspeaker signals are changed in relative amplitude and/or relative delay to accomplish this.

3.1. The rendering model

The following electrical circuit model is used to analyze the “ear signals” of a solid sphere with loudspeakers at +/-30^0 in front of it. The model is only valid for the ITD range of hearing, because diffraction and blockage effects are not included. The transmission to and summation of the loudspeaker signals at the ears is modeled for free-field conditions, Figure 6. The assumed ear spacing is 17.5 cm, causing 263 µs relative delay, Figure 3. The frequency response at the ear points will be investigated for level and group delay variations. The model was spot-checked for accuracy of some unexpected results by graphical vector addition of the ear signals.

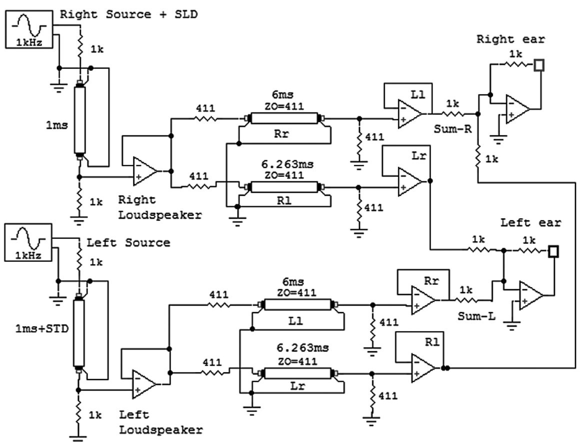


Figure 6 Electrical model for transmission of loudspeaker signals to the ears

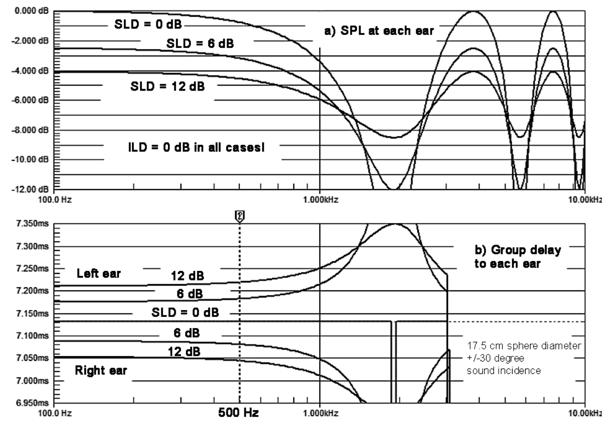


Figure 7 Ear signal frequency response (a) and group delay (b) for monaural sources at +/-30^0 and ear spacing of 17.5 cm. Left and right loudspeaker signals differ by 0 dB, 6 dB and 12 dB in level.

3.2. Level panning

The frequency response of the signal sum at each ear shows some surprises. For a monaural signal, where left and right loudspeaker outputs differ only in level and not in phase, left and right ear signals become identical in level, ILD = 0, Figure 7a. Crosstalk has canceled the level differences between the loudspeakers but the ear signal levels decrease with increasing level difference between the loudspeakers. The level difference between the two loudspeakers is converted into an arrival time difference, ITD, at the ears, Figure 7b. At low frequencies the time difference between the ears is proportional to the amplitude ratio between the loudspeakers, to their sound level difference in dB. For the following discussion we use the delay variation with source level difference at 500 Hz to illustrate basic rendering properties, Figure 8.

A real source would produce the ITD of Figure 3 for different angles of incidence α. We can therefore now relate the time difference in Figure 7b to the angle γ of a phantom source between the loudspeakers as a function of source level differences, Figure 8.

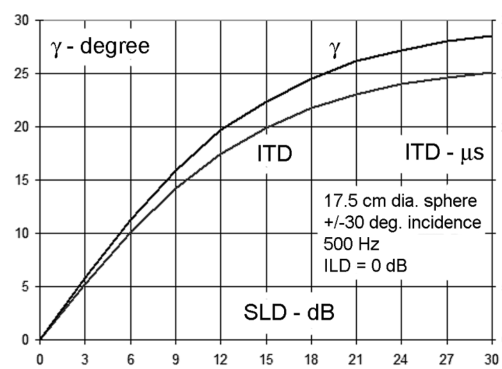


Figure 8 ITD and phantom source angle γ at 500 Hz vs. source level difference in dB between two loudspeakers at +/-30^0. This is the level panning law as derived from the low frequency sphere model of a head.

Left and right ear signal levels are identical, as seen in Figure 7a, but they decrease with increasing SLD, or increasing amount of panning towards one loudspeaker, Figure 9. A center phantom source is about 5 dB stronger.

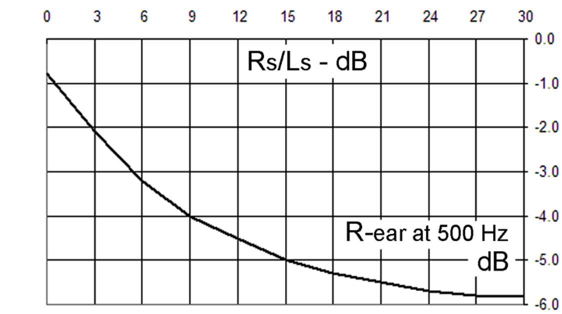
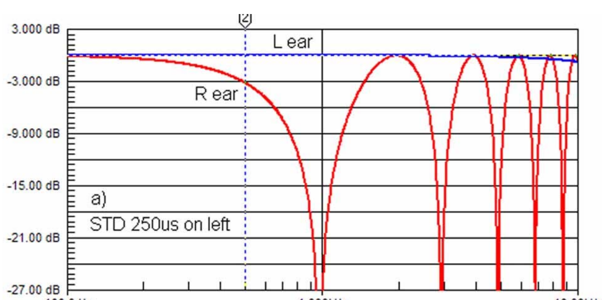


Figure 9 Phantom source level at 500 Hz as a function of right and left loudspeaker level differences.

At frequencies above 1 kHz the ear signals show periodic patterns of destructive interference, Figure 7. The model can only indicate the trend of the true behavior in the higher frequency range, because it does not account for sound blockage and other diffraction effects.

3.3. Time panning

Here the levels of left and right loudspeaker signals are identical. If one of the two channels is delayed with respect to the other, then the frequency response of the ear signals indicates a very different behavior than in Figure 7. Left and right loudspeaker signals interfere at the ears depending upon their phase differences and differently at each ear, Figure 10. Regardless of the delay between the source signals the inter-aural delay is zero. Timing differences between the loudspeaker signals do not produce corresponding arrival time differences at the ears, but amplitude differences due to a comb filter response!



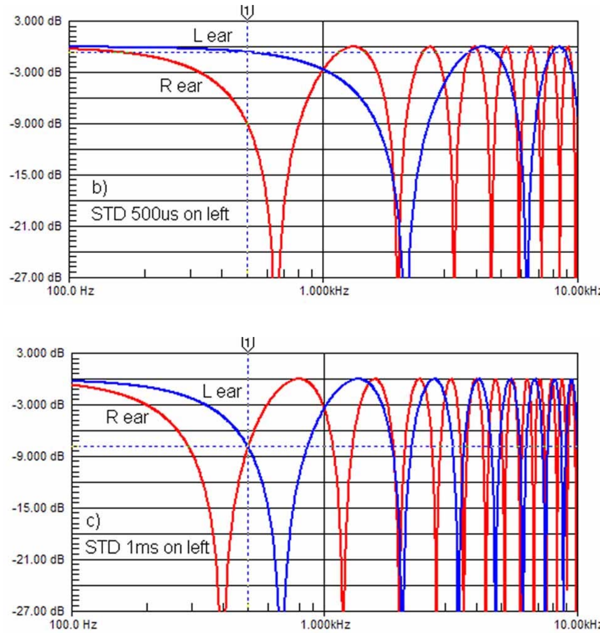


Figure 10 Frequency response of ear signals when the left source is delayed by 250 us (a), 500 us (b) and 1 ms (c) for SLD = 0 dB

Thus timing differences between the sources cause level differences between the ears, which are highly frequency dependent! At 500 Hz, for instance, L_{ear} changes slowly with increasing delay of the left source. R_{ear} heads towards cancellation at 750 µs. L_{ear} is always larger than R_{ear}. Even in the absence of ITD a phantom source is likely perceived as being left of center, towards the delayed source, Figure 11.

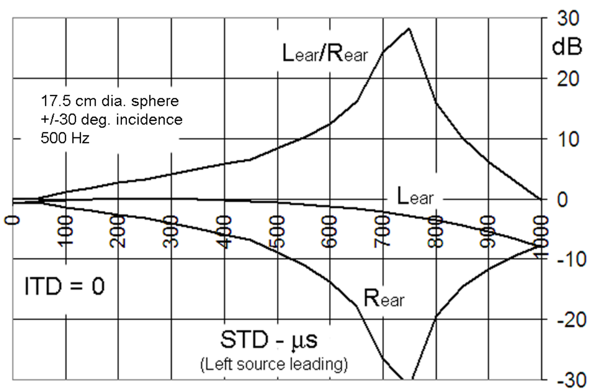


Figure 11 Left and right ear signal levels and their ratio in dB at 500 Hz vs. time difference between

sources

In general it is safe to say that time panning cannot generate well-defined phantom sources [12]. At best it will produce a phase instability that is mistaken as reverberation and spaciousness. Music and voice signals occupy a time dependent bandwidths. Time panning will render them diffusely in the ITD range of hearing. In the ILD range the head geometry changes the interference patterns of the loudspeaker signals at the ears. Due to pinna and precedence effects some directionality can be expected from time panned sources in the ILD range.

4. RENDERING STEREO SIGNALS

Stereo signals are created by either a down-mix from multiple sound tracks or are the output signals from a microphone pair. In a down-mix it is difficult to preserve a coherent spatial and sonic relationship between sound sources and their acoustic environment. Artificial reverberation is often applied to the mix to mimic a unifying acoustic space. A single, coincident microphone pair will have a coherent view of an acoustic event. How this view is rendered over two loudspeakers depends upon the polar response behavior of the individual microphone capsules and the placement of the pair relative to the sound sources and their reverberant environment. The placement of phantom sources between and at the loudspeakers has been described by the Stereo Recording Angle using an empirical time and level panning relationship [13, 14]. Here we investigate the placement of phantom sources based upon ITD for a solid sphere model.

4.1. XY microphone pairs

If the two microphones of a pair are directional, are coincident and point in different directions, then their outputs differ in magnitude only. Left and right microphone outputs are in phase with each other, provided that there is no path length difference between them and their frequency responses are identical. But the microphone outputs can have opposite polarity for certain angles of sound incidence, Figure 12.

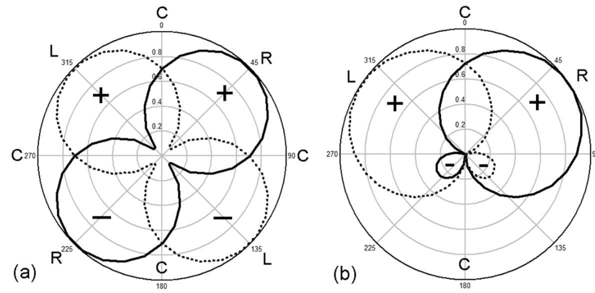


Figure 12 Polar responses of two bidirectional microphones at 90^0 (a) and two supercardioid microphones at 110^0 subtended angle (b).

For example, the R microphone in Figure 12a outputs +0.7 for 90^0 sound incidence, while the L microphone outputs –0.7. A similar situation occurs in Figure 12b for incidence angles greater than 90^0.

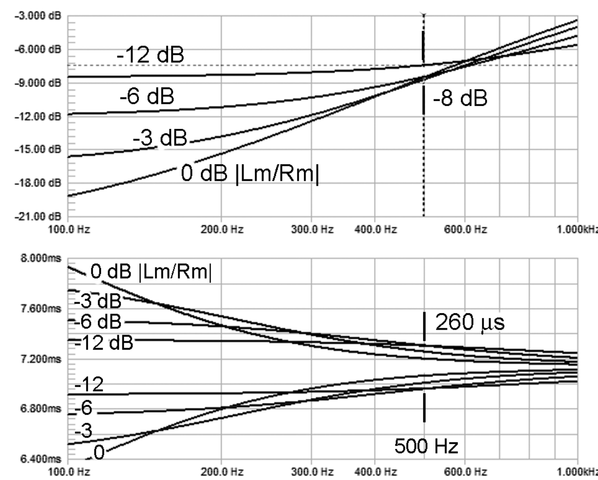


Figure 13 Ear signals due to opposite polarity microphone signals of different magnitude ratios. Magnitude at left and right ears, top, and arrival times, bottom.

For microphone signals of opposite polarity and different level the rendering model provides the ear signal responses in Figure 13. In general, opposite polarity loudspeaker signals produce large timing differences at the ears, as if coming from left or right loudspeaker only. Each ear receives identical signal levels, but they change with frequency and left to right level difference. In the following we restrict the analysis to 500 Hz in order to simplify the discussion of different microphone pairs. We assume that the magnitude of the ear signals is fixed at –8 dB and the ITD is 260 µs for angles α where the microphone signals are of opposite polarity.

When the two microphone output signals are in-phase, then they produce phantom sources in locations between and at the corresponding loudspeakers, depending upon the ratio of left and right microphone outputs as derived from SLD values in Figure 8. Left and right microphone output signals and their ratios as a function of sound incidence angle α are shown in Figure 14 for a pair of bidirectional microphones at 90^0 subtended angle. Here the polar response of Figure 12a is plotted on a linear scale for the angle α.

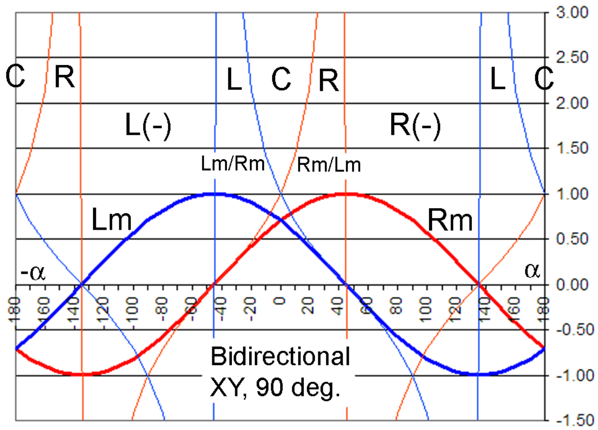


Figure 14 Bidirectional microphones at 90^0 subtended angle. Left and right output amplitudes Lm and Rm. The Lm/Rm and Rm/Lm ratio curves define where an incident sound at angle α appears as phantom source with angle γ in-between L and R loudspeakers, in the center C, or completely in L and R loudspeaker locations according to Figure 8.

In general each microphone will pick up signals from any angle in front, below, above and behind it. All those signals are rendered as phantom sources in-between, behind and at the loudspeakers, rarely in front of the loudspeakers or outside of them. In the following analysis only the horizontal plane will be considered. Rendering is symmetrical to the 0^0 centerline between the loudspeakers. Four microphone polar patterns will be investigated, Figure 15.

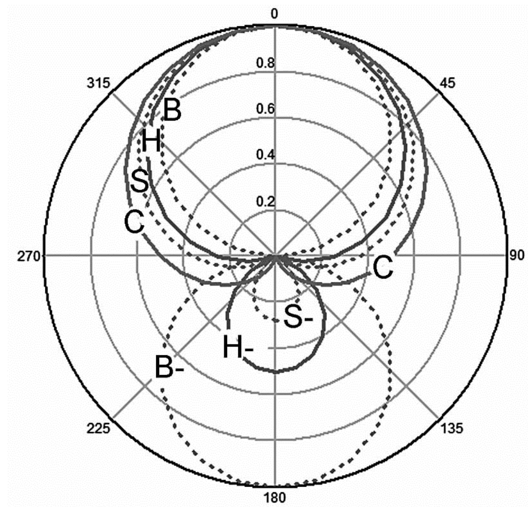


Figure 15 Polar patterns of Bidirectional, Hypercardioid, Supercardioid and Cardioid microphones

The microphone output signal ratio for the XY pair at angle α in Figure 14 leads to the phantom angle γ from Figure 8. The phantom source level is determined from Figure 9 and Figure 14 relative to the summed microphone output at the center. The information is processed in a spreadsheet (not shown) and displayed graphically in Figure 16, which shows the phantom source angle and the phantom source level as function of sound incidence angle α.

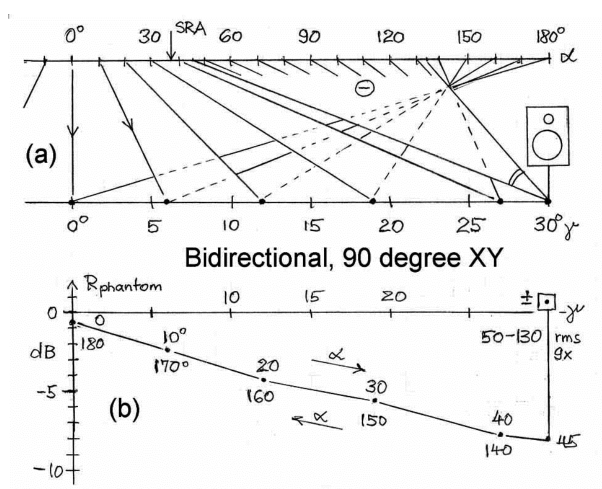


Figure 16 Bidirectional microphone pair at 90^0 subtended angle. Translation of sound incidence angle α

to phantom source angle γ (a) and phantom source level (b) from center point to right loudspeaker.

The phantom sources between the two loudspeakers represent the -45^0 to +45^0 and the –135^0 to +135^0 ranges of signals incident to the microphones. Microphone rear signals from 135^0 to 180^0 are mapped between left speaker and center, those between 180^0 and -135^0 between center and right speaker. Signals between +/-(45^0-135^0) incidence angle are rendered as opposite polarity signals of large level difference by left and right loudspeakers, Figure 16a.

The level of the phantom sources in Figure 16b is derived from Figure 9 for level panning and multiplied by the microphone polar response as in Figure 14, using the sum of absolute values for Lm and Rm as a function of α. Opposite polarity signals of large level difference between 50^0 and 130^0 are then added in rms fashion, Figure 16b.

The 90^0-bidirectional pair maps essentially a +/-40^0 source signal range between the loudspeakers. Opposite polarity signals from 45^0 to 135^0 add up to strong out-of-phase mono signals at +/-30^0 phantom angle. The phantom source level decreases about 7 dB between center and either loudspeaker. The 6 dB width of the phantom level response is +/-35^0.

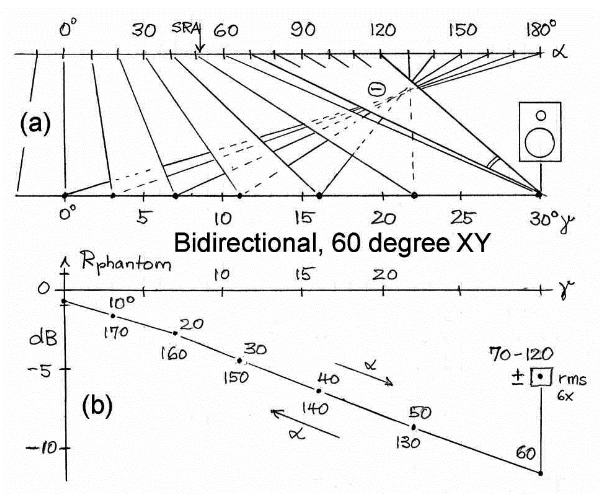


Figure 17 Bidirectional microphone pair at 60^0 subtended angle. Translation of sound incidence angle α to phantom source angle γ (a) and phantom source level (b) from center point to right loudspeaker.

With a narrower subtended angle of 60^0 the source pickup range widens to 120^0, Figure 17. The phantom level though changes by 11 dB between center and either loudspeaker. The 6 dB width becomes +/-42^0. The opposite polarity signals between the loudspeakers decrease compared to the 90^0 subtended angle bidirectional microphone pair.

Bidirectional microphone pairs provide a relatively narrow rendering focus and no front-to-back discrimination. They also render a strong and spatially diffuse mono components from each loudspeaker direction.

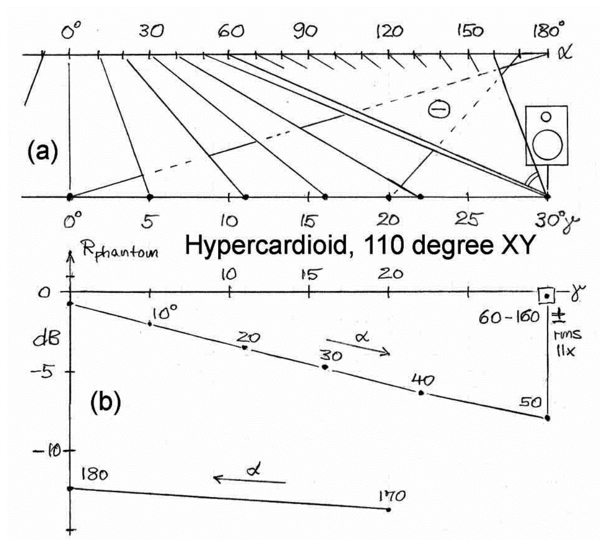


Figure 17 Hypercardioid microphone pair at 110^0 subtended angle. Translation of sound incidence angle α to phantom source angle γ (a) and phantom source level (b) from center point to right loudspeaker.

The hypercardioid XY pair at 110^0 shows a slightly wider recording angle than the bidirectional pair at 90^0 and nearly identical phantom level distribution, Figure 17. Phantom level width within 6 dB is +/-42^0. Rear pickup covers a narrower angle and is attenuated by over 10 dB due to a reduced rear lobe of each microphone, Figure 15.

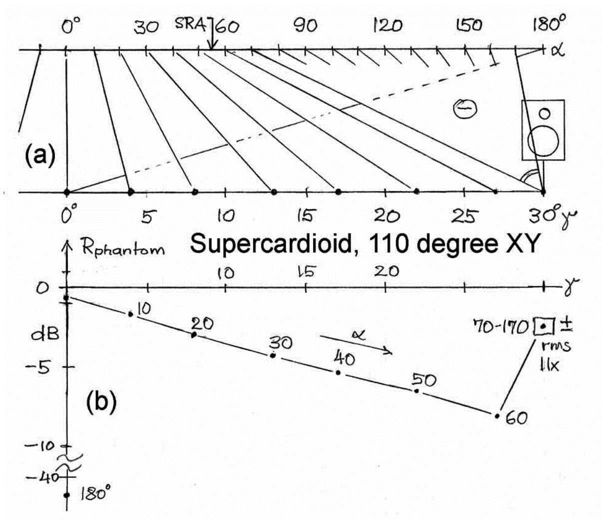


Figure 18 Supercardioid microphone pair at 110^0 subtended angle. Translation of sound incidence angle α to phantom source angle γ (a) and phantom source level (b) from center point to right loudspeaker.

The supercardioid microphone has a rear lobe that is even further suppressed than that of the hypercardioid. Consequently the phantom from 180^0 in the rear is attenuated over 40 dB. Opposite polarity signals, though are picked up over a wide range of +/-(70^0 to 170^0). The 6 dB phantom level width is 50^0.

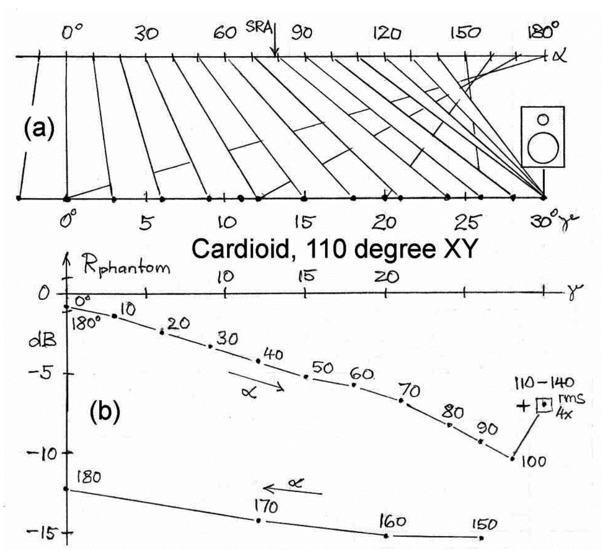


Figure 19 Cardioid microphone pair at 110^0 subtended angle. Translation of sound incidence angle α to phantom source angle γ (a) and phantom source level (b) from center point to right loudspeaker.

Ideal cardioid microphones do not have a rear lobe. A coincident cardioid microphone pair at 110^0 subtended angle maps signals from -110^0 to $+110^0$ into the +/-30^0 loudspeaker arc. The 6 dB phantom level width is +/-70^0. Signals between 110^0 and 140^0 are essentially summed into the right loudspeaker as monaural signals, Figure 19. Actually, the mono component does not change level significantly with changes in subtended angle. Signals from 140^0 to 180^0 are rendered in reverse direction, backward to the center.

The supercardioid pair would appear to be a most useful configuration for recording. It combines a reduced pickup angle with a minimum of rendering artifacts except for opposite polarity signals between +/-(70^0 to 170^0).

4.2. Rendering with AB microphone pairs

Microphone pairs, which are not coincident, will pick up a single source with phase and time differences in addition to amplitude differences between their outputs. According to the ITD model results in Figure 10, Figure 11 and Figure 13 this will lead to interference of the loudspeaker signals at the ears with inconsistent cues for phantom source direction and size. This will diffuse the aural scene and generate an impression of spaciousness that may be pleasing, but it will not yield naturally sounding spatial relationships. Widely spaced microphone pairs will essentially produce left and right mono loudspeaker signals.

4.3. Towards object oriented recording

The ITD model indicates that spatial rendering is best obtained from level panned mono signals and from the outputs of coincident pairs of directional microphones. Signals from spaced microphones or time panned mono signals lead to unrealistic spaciousness. It should be possible to use a single coincident XY microphone pair as the basis for rendering the spatial relationships between individual sound sources and their interaction with the reverberant sound field of the recording venue, Figure 20. The pair must be placed at some distance from the acoustic sources to minimize level differences between near and far instruments and to capture the width and height of the acoustic scene.

If necessary for clarity or audibility, individual sources or groups of instruments (a through k) are recorded as mono signals and then panned to the proper location in the phantom scene established by the XY pair. Electrical signal delay may be needed, if the required amplitude approaches the level of the XY pair output. When more than one microphone (f, g, h) is used to record a larger group of sound sources, then signal overlap and timing differences between the microphone outputs will lead to a loss of clarity when panned to left, center, right respectively. The microphones need to be close to the source to minimize leakage. The panned level of the microphones must be kept low not to override the precedence of the XY outputs. Two microphones (a, b) for a single source will add diffuseness, but only to the high frequency spectrum when placed close together. Widely spaced microphones (A, B) render a diffuse sound depending upon their distance from sources. They are useful for low frequency pickup when further away from the orchestra.

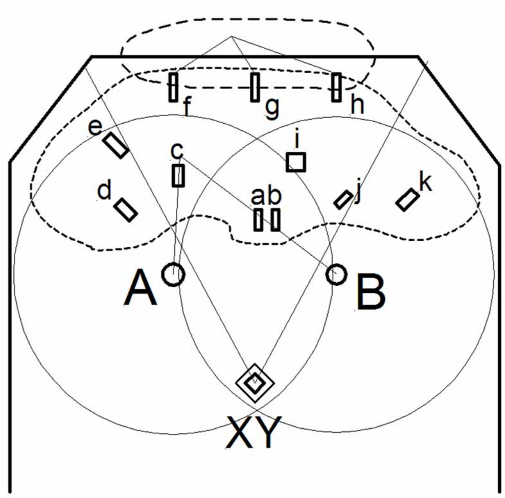


Figure 20 Microphone setups for preserving natural spatial relationships between sound sources and recording venue

5. STEREO WITH FOUR LOUDSPEAKERS

Headphones provide analytical clarity and sonic detail. Level panned recordings, though, are not rendered properly because they do not generate ITD cues [15]. The perceived aural scene is perceived as inside the head and distances to sources are foreshortened. Loudspeakers place phantom sources at or behind the line between the loudspeakers and not inside or near the head. Clarity and detail, though, can be lost with loudspeaker reproduction due to the room reverberated sound field and room modes.

As an experiment, a recording was played back simultaneously over loudspeakers and headphones. The sense of natural spaciousness was greatly enhanced, when spreading the headphones some distance away from the ears and letting the loudspeaker sound level predominate. A more practical approach led to the construction of two small omni-directional loudspeakers, which were placed close to the listening chair [16].

Figure 21 Stereo rendering with two acoustically small dipoles at standard distance and two omni-directional loudspeakers covering the ITD frequency range in close proximity to the listener

The main loudspeakers are positioned at $+/-30^0$ and the support loudspeakers at about $+/-65^0$ from the center line, Figure 21. Sound from the support loudspeakers rolls off below 200 Hz and above 1.5 kHz.

Addition of support loudspeakers to the free-field model in Figure 6 indicates an increase in phantom angle γ at a given source level difference SLD, Figure 22. The angle widens even further when the support loudspeaker output precedes the main loudspeakers. Eventually the model shows to combing effects similar to those of Figure 10.

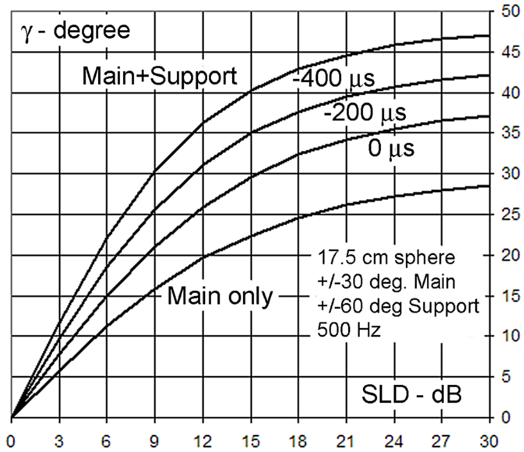


Figure 22 ITD for two monaural sources at +/-30^0 and with the addition of two synchronous sources at +/-55^0

Listening to a variety of recordings indicated that the support speakers must be operated below a threshold volume level where they are not noticed by themselves, but still contribute spatial enhancement. Their measured output level, including room reverberation, is about 6 dB below that of the main loudspeakers, making the effect subtle but realistic. Surprisingly, electrical delay of the support loudspeaker output can be removed without audible consequences. The optimum listening distance for the effect lies within two reverberation distances from the main loudspeakers. Spaciousness remains at larger listening distances, but phantom source locations become less defined.

The setup appears to render XY microphone and level panned stereo recordings optimally. It shows the spatial flaws in spaced microphone recordings. It could be a tool for recording engineers to judge the overall quality of a mix [17].

It must be pointed out that the four loudspeakers are set up with greater than 1m distance from large reflecting surfaces. The main dipole loudspeakers and the support monopoles exhibit essentially frequency independent radiation patterns and constant directivities. Thus their reverberation distances are nearly constant over a wide frequency range. Delayed room reflections therefore have similar spectral content as the direct loudspeaker signals, which allows the brain to withdraw attention from the room and all four loudspeakers. Rendering the stereo signal with four loudspeakers widens the phantom scene, even approaching +/-90^0 for some recordings. Image depth and height are increased. Phantom images become spatially better defined. Increasing the volume level decreases the distance to the phantom scene and magnifies it towards a natural size, which greatly enhances realism and enjoyment.

6. CONCLUSIONS

- Stereo recording and rendering must be considered as a unit, if communication of natural spatial information is important.
- A sphere model of the human head can provide qualitative insight into the rendering of sound over two loudspeakers in the ITD frequency range of hearing.
- Coincident microphone and level panned single microphone techniques yield spatially defined phantom sources.
- Spaced microphones and time panned single microphone techniques yield spatially diffuse phantom sources.
- An experimental setup with four loudspeakers acts like a spatial acoustic magnifier. It could be a useful tool for evaluating a recording mix.
- The quality of a recording can only be judged after it has been rendered. Recordings, which preserve natural spatial relationships, lead to greater enjoyment.

7. REFERENCES

[1] David Griesinger, Frequency Response Adaptation in Binaural Hearing, 126th AES Convention, Munich 2009, Preprint 7768

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

[3] Zachary A. Constan, William M. Hartmann, On the detection of dispersion in the head-related transfer function, JASA, Vol. 114, No. 2, August 2003

[4] Eric Benjamin, An experimental Verification of Localization in Two-Channel Stereo, 121st AES Convention, San Francisco 2006, Preprint 6968

[5] Jonas Braasch, A binaural model to predict position and extension of spatial images created with standard sound recording techniques, 119th AES Convention, New York 2005, Preprint 6610

[6] Francis Rumsey, Spatial Audio, Focal Press, 2005

[7] Siegfried Linkwitz, Room Reflections Misunderstood?, 123rd AES Convention, New York 2007, Preprint 7162

[8] Heinrich Kuttruf, Room Acoustics, John Wiley & Sons, 1973

[9] Juergen Meyer, Acoustics and the performance of Music, Springer, 5th Edition 2009

[10] Barry Blesser, Linda-Ruth Salter, Spaces speak, are you listening? Experiencing aural architecture, The MIT Press, 2007

[11] Barry Blesser, An Interdisciplinary Synthesis of Reverberation Viewpoints, JAES, Vol. 49, No. 10, 2001 October

[12] Stanley Lipshitz, Stereo Microphone Techniques: Are the Purists Wrong?, JAES, Vol. 4, No. 9, 1986 September

[13] Michael Williams, Microphone Arrays for Stereo and Multichannel Sound Recording, Editrice Il Rostro, 2004

[14] Helmut Wittek, "Image Assistant", JAVA applet for determining the "Stereo Recording Angle", www.hauptmikrofon.de

[15] Benjamin B. Bauer, Stereophonic Earphones and Binaural Loudspeakers, JAES, Vol. 9, Issue 2, pp. 148-151; 1961 April

[16] Siegfried Linkwitz, WATSON Stereo Enhancement Loudspeakers, www.linkwitzlab.com/Watson/watson.htm

[17] Jens Blauert, Ute Jekosch, A Layer Model of Sound Quality, JAES, Vol;. 60, No. 1-2, 2012 January/February