



---

# Audio Engineering Society Convention Paper

Presented at the 111th Convention  
2001 September 21–24 New York, NY, USA

*This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42<sup>nd</sup> Street, New York, New York 10165-2520, USA; also see [www.aes.org](http://www.aes.org). All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.*

---

## AMBIOPHONICS

### Achieving Physiological Realism in Music Recording and Reproduction

By Ralph Glasgal  
Ambiophonics Institute  
4 Piermont Road  
Rockleigh, New Jersey 07647  
[www.ambiophonics.org](http://www.ambiophonics.org)  
glasgal@ambiophonics.org

#### ABSTRACT

Ambiophonics, is the logical successor, to stereophonics, 5.1, 6.0, 7.1, 10.2, or Ambisonics in the periphonic recording and reproduction of frontally staged music or drama. We show how, optimally, only two recording media channels, driving a multi-speaker surround Ambiophonic system, can consistently generate a "you are there" sound field that the domestic concert hall listener can sense has normal binaural physiological verisimilitude. Ambiophonics can deliver palpable realism even from standard two media channel recordings such as the existing library of LPs, CDs, DVDs, or SACDs or via super-wide-stage recordings made using an Ambiophone.

#### INTRODUCTION

The goal of Ambiophonics is to deliver to one or two home listeners a realistic replica of the concert hall experience. Ambiophonics, a public domain technology supported by the non-commercial Ambiophonics Institute, combines an exploitation of under-appreciated psychoacoustic principles with the basic rules of good musical performance-space design to create believable concert-hall soundfields in typical home listening rooms. Ambiophonics moves the listener into the same space as the performers by accommodating to the uniqueness of individual pinna characteristics, minimizing ambient interaural

correlation at the listening area, abandoning the traditional stereo loudspeaker triangle (including the center speaker in 5.1), generating early reflections and reverberant fields from stored real-hall impulse responses, eliminating comb filtering due to front-loudspeaker crosstalk, and using room correction/treatment technology to insure that the listening room does not impair the illusion.

Ambiophonics is not meant to be used for movies or video and makes no serious provision to record or reproduce direct sound sources that are not frontal. It also assumes that the acoustic venue remains static for the duration of the performance and the recording

session. That is, an orchestra recording the first movement of Bruckner's 5th Symphony in Carnegie Hall does not move to Avery Fisher Hall to play the second movement.

Ambiophonics is a binaural technology based recording and reproduction system for music that abjures phantom imaging, combines real hall ambience convolution, pinna angle preservation, and recording microphone compensation (optional) [1] to provide exceptionally realistic reproduction of 2-channel music sources such as LPs, CDs, as well as 5.1 Dolby Digital or DTS material. This is possible because 2-channel recordings themselves are not inherently Blumlein stereophonic. That is, they do not contain the crosstalk rays or angle distortions that will be produced by the loudspeakers of the usual stereo triangle and they do not know that their sound sources will not be localized binaurally but rather by using the phantom image hearing mechanism of the antiquated stereo triangle.

After a review of the pertinent acoustic and psychoacoustic principles involved in the normal process of listening to music in halls and theaters we show how these well documented psychoacoustic and acoustic principles can be used to derive the key elements of the Ambiophonic paradigm. These include the Ambipole, a comb-filtering-free front-speaker pair, real hall impulse convolution, (the process by which even a studio recording can seem to be sounding in one of the worlds great concert halls) Surrstats, (a surround speaker design that facilitates the natural reproduction of hall ambience) room correction to ensure that listening room anomalies do not distort the Ambiophonic sound field, and the Ambiphone, a main microphone arrangement conceived to make new two-channel surround music recordings that are optimum for Ambiophonic playback.

One could make a good case that there is nothing novel in Ambiophonics technology. I have been able to find every one of its basic precepts supported elsewhere in the literature and some of the references go back decades. The remarkable thing is that this basic knowledge of acoustics and psychoacoustics was so thoroughly ignored as the sound recording and reproduction industry moved from monophonic to stereophonic to surround (but still phantom imaging) music systems.

## REALISM IN SOUND REPRODUCTION

In this paper, realism in the reproduction of classical and pop live music and virtual reality compositions is

understood to mean the generation of a sound field realistic enough to satisfy any normal ear-brain system that it is in the same space as the performers, that this is a space that could physically exist and that the sound sources in this space are as easy to locate and that the ambience as full bodied as at a live concert. Note that physiological realism does not necessarily equate to accuracy. For instance an opera recording made in La Scala but reproduced as if it were in Covent Garden can seem to be fully realistic even if inaccurate. Appropriateness is also not equivalent to realism. Thus recording Mahler's second symphony in the Concertgebouw and then reproducing it in The Wigmore may be ghostly sounding but nevertheless excruciatingly realistic.

Likewise, it is important that localization be as effortless and as natural as in everyday hearing rather than be exaggerated by spot mic's and pan processing. Note that if a stage is 120 degrees in the hall but reproduces as if it is 140 or 100 degrees it can still be realistic since such changes in stage width are equivalent to changing one's seat in an auditorium. We will see however, that Ambiophonics can be exceedingly precise in maintaining such stage perspectives and can even allow one to choose the perspective they are most comfortable with. [2] Fig. 3

The binaural recording/reproduction ideal states that if you deliver to the entrance of the ear canals precisely what those ear canals would have sensed had they been in a good seat at the live event then realism is guaranteed. I believe that two-media channel Ambiophonics does meet this exacting binaural standard for the case where there are no direct sound sources in the rear half of the horizontal plane and no direct sound sources far above or below it. (You can of course, always use multichannel formats such as DVD-A to carry applause or antiphonal direct offstage sound.) For hall ambience, Ambiophonics is as binaurally periphonic as you wish it to be in your living room and is limited only by your budget and how much space you have for surround speakers. We will now consider the acoustic and psychoacoustic elements that need to be understood in order to construct Ambiophonic systems. We will consider the frontal stage sound and the surrounding ambient sound field separately.

## FRONTAL PSYCHOACOUSTIC ELEMENTS

It is well known that human beings localize by using both their pinnae and their head shadow. [3] In this paper we will define the term HRTF to include head/torso shadowing only. Pinna functionality will

generally refer to the ability of humans to localize using just one ear. HRTFs are responsible for all interaural differences in phase/time and level and thus the localization that can be inferred from such interaural time and level differences. By contrast, each pinna is independently sensitive to direction and it provides a higher frequency localization system, dependent on interaural differences only to the extent that one pinna can (maybe) confirm the localization sensed by the other one.

The pinnae are, in essence, direction finders. They map the direction of incident sound into a very complex pattern of peaks and nulls in the frequency domain from about 1000Hz up. [8,9] As you can see from figures 1 and 2, the patterns are very complex and vary from individual to individual. Snow and Moir [4,5] have shown that the ability to localize complex transient rich sound such as music is greater in the region above 1000 Hz than below it. Indeed many people can localize to less than 1 degree in this frequency range even though the only interaural clue for a sound dead ahead is that there are no interaural clues.

However, Ambiophonics does not presume or even need to fully understand the physiology of the hearing mechanism. We make the assumption that if we deliver to the ears a physically correct binaural sound field that the ears will respond accordingly and sense realism. Thus it is not necessary to accept the precept that the pinna are superior to the HRTF. It is only important that nothing be done to interfere with either HRTF or pinna functionality in the recording/reproduction process. Thus we can state the first law of Ambiophonics. In any recording/reproduction chain there shall be only one set of pinnae and those pinnae must be the home listener's. Remember that the pinnae are really direction finders. This means that to the greatest extent possible all recorded sound must impinge on the home listener from the correct direction. This includes both direct stage and hall ambience sound.

The second law of Ambiophonics is that there shall be at least one but only one head shadow (HRTF) in a recording/reproduction chain. It is desirable but not essential that this HRTF be the home listener's. Head shadowing is quite variable depending on the position of the head, up or down, right or left. Also sound passes over the head, under the chin, around the back and across the nose. Therefore it is difficult for the brain to distinguish between a sound shadow due to its head shape, its attitude or due to a combination of the two. Thus, the difference between individual head shapes is not nearly as significant as the

differences between pinna. In front stage recording and reproduction, Ambiophonics does make some slight use of the fact that putting your pinna on someone else's head does not impair the sense of realism to any great extent.

Another factor to consider in this regard is that, in a concert hall, head shadowing of the direct stage sound for most of the front stage is not appreciable except for the stage extremities. But this is not true for hall ambience. We will see that the indirect hall ambience sound can and should be delivered using both the listener's own HRTF and pinnae. In the case of hall ambience, sound does come from virtually every direction and so little HRTF or pinna compromise should be tolerated. Fig. 4

### **THE PSYCHOACOUSTICS OF THE STEREO TRIANGLE**

Ambiophonic theory deals with three main parameters of staged music sound recording and reproduction. They are the capture and delivery of a binaurally correct front stage image, the recreation of real early reflections and reverberant tails, and control of the listening room environment. We first consider the traditional but psychoacoustically incorrect method and then the binaurally correct capture of the direct stage sound of an orchestra or an opera being performed in one of the worlds great concert halls or opera houses.

It was well known, even to Alan Dower Blumlein, that the stereo loudspeaker triangle had and still has serious psychoacoustic defects. In this paper stereo refers only to the use of the equilateral loudspeaker triangle not the Greek sense of three dimensions. First the Blumlein stereo effect depends on a "sonical" illusion usually referred to as phantom imaging. (A distinction is made here between virtual images and phantom images. The term "phantom" is applied only to images not localized by the usual binaural hearing mechanism i.e. images that are generated from sound vectors that could not exist in nature.) This is analogous to an optical illusion. Like optical illusions, phantom images are intriguing and useful but they are fluky and not realistic. Note that mere localization is no guarantor of realism. For example, a recording of a solo piano played back over a single speaker is easy to localize but seldom sounds real. If the position of a lion on the savanna had to be detected by our African ancestors by cutting the lion in half and having it roar from both 30 degrees to the left and right I doubt that ancestor worship would be possible. If physiological realism

is the goal, we want to avoid any reliance on the phantom image (panning) mechanism.

The stereophonic triangle has other defects as well. The stereo triangle crosstalk (where the left signal reaches the right ear slightly delayed and HRTF filtered and vice versa) generates comb filtering artifacts from about 800 Hz up whenever there is any correlation between the signals at the speakers. [14] It is not just the crosstalk per se but the comb filtering patterns it engenders that are so deleterious. These combing effects mimic the mapping function of the pinna in the sense that the comb filtering peaks and dips are not unlike the peaks and valleys caused by the contours of the pinna. [Figs. 1&2]

The physiology of how the comb filtering interferes with the normal pinna functionality and the extent of the deterioration is, to my knowledge, not well documented. I believe these combing effects cause listeners to sense that the stereo field is not real, and can also cause listening unease or fatigue. It also makes stereo systems very sensitive to speaker placement, phase or delay differences, head attitude, room reflections, etc. because even small changes in delay or geometry radically alter the pattern of the comb filtering and therefore the perception of the sound stage. Similarly, untreated room reflections in the sub-millisecond region can cause combing and thus distort pinna directional cues. In any case, whether I have it right or wrong, Ambiophonics takes no chances and makes no compromises. It strives to eliminate all sources of comb filtering in playback by limiting room reflections and front loudspeaker crosstalk.

You can easily hear the effects of comb filtering in any stereo triangle by playing white noise and walking in front of the speakers from one side to the other with one ear facing the speakers. If you also cover the ear facing away from the speakers to block room reflections, the effect is especially audible. As you pass through the area where your open ear is equidistant from each speaker you will hear several seemingly minor changes in level and tonal coloration but these slight audible effects are not insignificant in relation to our sense of realism.

The worst effect of stereo crosstalk is certainly the comb filtering. But the size of our head limits comb filtering to frequencies over 800 Hz or so. [14] The question arises as to whether stereo crosstalk below this frequency affects central stage frontal realism seriously enough to require its elimination. At very low bass frequencies say 90 Hz and down bass cannot be localized. In a concert hall, either a central

or a sideways-located low bass instrument will illuminate both ears equally. Likewise, two stereo speakers with identical bass signals will deliver the same signal to both ears at sixty degrees or any degrees. (In the case of coincident microphone recordings, mid-bass crosstalk actually enhances separation by converting the head-width-induced phase shifts between the ears of tens-of-degrees into audible level differences between the ears. But this is the only exception to the rule that stereo crosstalk is always undesirable.)

In the range of say 200 Hz to 800 Hz we can assume negligible pinna influence. But at the higher end of this frequency range the HRTF is still responsible for both amplitude and timing shadows that decline steadily in detectability with increasing wavelength. Looking again at the stereo triangle, and a central sound source, a signal from the right stereo speaker will reach the left ear about half a millisecond later than at the right ear and be attenuated by the head shadow. This delayed signal is added or subtracted depending on phase to the slightly earlier and slightly louder, but identical, signal coming from the left speaker to the left ear. Thus, compared to the live concert experience, there is a boost (or a dip) of a bit less than 3dB in parts of this frequency range. At higher frequencies this effect degenerates into comb filtering and at very low frequencies it approaches +3dB asymptotically.

Whether this midrange stereo triangle crosstalk is seriously detrimental to realism is still an open question. But since these extra sound rays are absent in the concert hall let us assume that whatever means are used below on to replace the stereo triangle will allow little crosstalk above 150 Hz. Trying to eliminate crosstalk at very low frequencies where there is no such interaural function is deleterious, (because of excessive signal levels), costly, and hardly necessary in virtual sound applications.

The stereo triangle also limits the stage width to the angle between the speakers. This is obvious, since if only the left speaker is on during a sound originating from the extreme stage left, then one hears just this speaker and localizes to it as in everyday binaural hearing. Thus if the orchestral stage is, say 120-degrees, as seen from the fifth row center, it will be reproduced as if it were only 60 degrees wide in most stereophonic systems.

There is also the question of pinna incident angle error. Since all stereo sound comes from plus or minus 30 degrees, the pinnae transmit this pattern to the brain (if the comb filtering doesn't interfere). One

pinna says the sound is over there at 30 degrees to the left and the other pinna says the sound is 30 degrees to the right. The interaural HRTF cues are similarly unnatural. But most people resolve this situation by sensing a source directly in front of them. Again these anomalies cause us to sense that the sound is not real and explains why virtual reality panning algorithms are more impressive using earphones or ear-speakers as in 3D IMAX.

The stereo triangle has a limited sweet area in the direction along the line toward or away from the plane of the speakers. The angle to the speakers will change as one moves along this line and this will change the comb filtering, the pinna angle error, the midrange HRTF, and at the extremities, even cause a loss of phantom imaging, leading to either mono or a hole in the middle. However, the deterioration of the stereo effect as one moves sideways away from the sweet spot seems mild to many listeners. I would maintain that this is because, since stereo (and LRC of 5.1) is quite unrealistic to start with, moving away from the ideal listening point doesn't seem to be much of a sacrifice to a lot of home listeners.

In stereo, and also to a large extent in 5.1, rear and side (including ceiling) ambience comes willy-nilly from the front speakers. As we will see below, only natural frontal stage ambience should come from the front speakers.

5.1 surround sound is essentially two stereo triangles side by side each rotated offside by 15 degrees. In live music, 5.1 depends on phantom imaging for localization on each side of the center speaker and has the same problems with comb filtering and pinna angle of incidence error that two-speaker stereo has. It also has one other interesting psychoacoustic flaw. Phantom imaging really depends on frontal speaker symmetry in the horizontal plane. If you listen to a stereo system sitting 15 degrees sideways, your phantom-imaging ability declines in accuracy. Again, the movie-reasonable LRC arrangement is fine for panned mono dialog and mono surround sound effects or a contrived (3 mono) string trio but it cannot deliver front stage musical realism any better than, or I would say, as well as the traditional stereo triangle. G. Theile [13] offers some heroic means to improve LRC phantom imaging in LRC systems, as does D. Griesinger [12].

## AMBIPOLES

Let us assume that we have two super directional, ideal, point source loudspeakers and place them head spaced directly in front of a listener. Then the sound

from the left loudspeaker only impinges on the left ear and the sound from the right loudspeaker goes only to the right ear at least over the about 150 Hz and up frequency range of interaural and pinna interest. In this configuration, little sound goes past the nose to reach the wrong ear and, for central stage sources, perhaps up to 15 degrees to each side the incident angles to both pinnae are reasonably correct. If the signals to the speakers have been recorded, either as described below or are reasonably unprocessed, then each ear hears what the recording microphone (or the mix) heard and phantom imaging of the central stage is not needed. There is also no possibility of comb filtering since we postulate that the speakers are sharply focused on the ears in the comb-filtering region. Even if there were some crosstalk, the path length difference is so small with the speakers so close together, that the combing onset moves up in frequency and is much less audible.

If the speaker's beams are truly collimated, then in theory, the sweet area extends from very near the speakers backward to infinity. If the beam is say twice the diameter of the pinna, one can rotate one's head, as in the concert hall, and the stage remains frontal and fixed. If we also assume that the beams are even larger at just under a head size in diameter then one can also move sideways a head's width without compromising stage realism.

My own Ambipole does exhibit exactly these characteristics (with the surround hall ambience speakers on, see below). One is free to move about more than in a seat in a concert hall. There is also room for six or even more listeners along the centerline. Off the centerline, it sounds as like any stereo or LRC system, except that one can get up and walk around and still be in a concert hall or cathedral with the stage up front. An interesting phenomena is that as one moves back the perspective seems to shift proportionally until at the rear of the room one seems to have arrived at the balcony. In other words Ambiophonics reproduces depth cues and behaves as though there is a critical hall radius.

I call such speaker arrangements Ambipoles. A simple Ambipole can be formed at virtually no cost by taking an eight inch thick seat cushion, holding it on edge on your lap, and listening to two speakers say about six feet away, head-spaced directly in front of you. [11,14] With virtually any stereo recording of live or virtual music you will hear a greatly enlarged and natural stage often with enhanced depth. Ideally Ambipoles are best constructed using speakers that are time coherent, don't spray a lot of sound around the room that could be reflected back to

the listening area, and are line sources so that the angle to the ear of all the drivers is consistent. Flat or slightly concave Electrostatic panels (Ambiostats) are good in this application, as are ribbons and narrow speaker designs. However, in practice, the barrier effect is very obvious with just about any type of speaker.

Since barriers are of finite size they do not eliminate crosstalk at frequencies much below 400Hz. Thus the fact that a wide front stage can be produced by even a small mechanical Ambipole would confirm that low end crosstalk is not profoundly detrimental and that higher frequency fast transient cues for the localization of music are more significant than lower frequency slower ones. [4,5]

Very small speakers such as satellite types with a separate subwoofer make excellent Ambiophonic studio monitors. The small speakers do behave like point sources and a small hinged barrier can be placed to make it easy to accurately monitor a recording session without crosstalk, phantom imaging or pinna distortion. [Fig 5 & 6]

Of course mechanical barriers are not to everyone's taste. If you measure the impulse response of a great sounding barrier and then write software to duplicate what it does down to a somewhat lower frequency, you can replace the barrier with software running on an Ambiovolver. [11,22]

Ambipole software can also include head shadow adjustments to compensate for the differing microphone arrangements used in recordings such as OCT, INS, ORTF, M/S, Decca Tree, etc. [11] Even if one does not know how the recording was made (panned or more likely spot mic'd to death) one can pick a correction curve that seems best. In practice, however, I have found that, to my ears at least, the Ambipole works exceptionally well for 90% of the classical music recordings I have without such tweaking. I would go further and state that any so-called stereo recording (or 5.1) sounds better played Ambiophonically no matter how it was recorded.

Software Ambipoles work by generating delayed, opposite polarity crosstalk cancelling signals in an infinite series. [10] Since the cancellation in this case is not based on collimation it is possible to use an omnidirectional or a hemidirectional speaker and enlarge the sweet area substantially in the horizontal direction. I have tried this with Ikonoklast speakers that have a cone shaped tweeter and been very successful. As one moves off center the amplitude at each ear from each speaker remains the same but

eventually timing errors reduce the effectiveness of the crosstalk cancellation. Ikonoklasts with larger diameter cones operating down to below 1000 Hz would likely be hard to beat in this application. As described above, an interesting aspect of my own virtual hall is that it seems to scale like a physical concert hall. As you move to the Ikonoklast Ambipole you feel closer to the stage. By the time you have moved back ten feet or so you seem to be in the last row.

## AMBIOPHONES

Before we end our consideration of the realistic reproduction of the front stage, we must consider whether there is a recording method that is ideal for Ambiophonic reproduction via the Ambipole and the ambience surround speakers described below. Indeed there is. [2] Fig.3 In an ideal recording and reproduction chain, we have indicated that the rule is that there should be only one set of pinna (your own) and one HRTF. If one takes a head shaped sphere and places two omni directional microphones on the surface where the ears would be one has introduced a reasonable replica of a head shadow into the front stage recording chain. Fig. 7 Remember that this is necessary because in the reproduction of the front stage left and right channels the Ambipole speakers are directly in front and so there is no head shadow for stage sounds coming from the sides. A head size sphere or head shaped microphone without pinna can perform this function. Indeed such a microphone already exists and is commercially available as the Schoeps KFM-6. [15]

If you simply place this microphone fifth row center you will record and then reproduce the same perspective that a listener at that seat would have had. Fig. 3 Normally, one cannot do this because most microphones would pick up more ambience than direct sound and this ambience coming from the front speakers would make the musicians seem to be performing in a sewer. Thus the Ambiphone is shielded to the rear, the extreme sides, and overhead by sound absorbent material.

The Ambiphone picks up a normal binaural sound signal from the front of the stage including frontal early reflections and reverberation tails. While not the most important source of reverberation, proscenium ambience is desirable, easily captured by a baffled Ambiphone, Fig. 7, and such frontal ambience has been shown to enhance the value of hall reverb coming from all the other directions. [17] In Ambipole reproduction, both direct sound and frontal hall ambience come from the proper direction

(the front) and the head sphere provides level and time delay cues for realistic localization and spatial sense at the far sides.

However, there is one fault in this method where sound sources at the extreme sides from 70 to 90 degrees are present. Although the head shadow is made correct, the pinna are being irradiated by the Ambipole dead ahead and as the angle gets wider the pinna eventually win out and the stage never gets to 180-degrees. One can only say that in real concert halls and opera houses only rarely are there direct sound sources at such extreme angles. The effect of this slight stage narrowing is akin to moving back a few rows in the hall. Finally one should note that the stage width produced by an Ambipole/Ambiophone combination is easily two and one half times wider than stereo or 5.1 can produce. As described next, the addition of the surround hall ambience speakers does appear to provide enough pinna stimulation to largely restore the missing extremities.

### CONCERT HALL AMBIENCE BASICS

In order to recreate concert hall ambience in a home living room, it is advisable to understand the essential nature of concert halls to a reasonable degree. Again, since not everything is known about concert hall psychoacoustics, Ambiophonics makes as few compromises or assumptions as possible to avoid any inadvertent diminution in realism. But let us briefly review the state of the art in hall acoustics.

Yoichi Ando, [21] established that halls sound better to most concert goers if the IACC, Interaural Cross Correlation factor is as small as possible. That is, the ambient field should be as different between the ears as possible. He also indicated that some very early reflections are favored if they come from the front 55 degrees although this may have more to do with apparent source width on the stage than with hall spaciousness. He and Griesinger [12] also showed why different halls or hall parameters favor different types of music. That is, if a hall is ideal for Mozart it is unlikely to be so for Mahler. Thus in an Ambiophonic system we have an option to improve on the live concert hall experience if we provide a means to choose the best hall for the recording we are playing at the moment.

Note that reverberation that comes from directly over head or directly from the rear or from anywhere in the medial plane is identical at the ears with an IACC of 1. One would therefore expect that such ambience is not as interesting as reflections from other directions and this is indeed the case. [19] Our

relative insensitivity to over head reflections could be considered intuitively obvious to coin a phrase. Cavemen, living mostly outdoors, despite their descriptive name, would not be expected to evolve a keen sense of spaciousness based on signals from an above that has no above. Thus when setting up a home system, if one has limited resources, it is better not to become fixated on height to the exclusion of horizontal plane diversity. Of course, raising side speakers is fine.

No two researchers or architects agree on what makes the best hall. For instance Ando's reflections from 55 degrees might better come from 90 degrees depending on whether you like hall early reflection ambience to affect the apparent source width as well as the sense that you are sitting in a large interesting space. There is no question that most listeners favor reverberation that has a strong lateral component but directional variety is regarded as the real spice of concert life. [17] We will see that Ambiophonics makes no assumptions as to ideal hall design. Quite the contrary. This review is designed to show the futility of any such compromises or preconceived notions.

For example, researchers talk about listener envelopment or LEV. If LEV means a realistic feeling that you are in a concert hall I have no argument with the term. However, like phantom imaging, artificially generating a sense of LEV does not necessarily equate to physiological verisimilitude. There are many ways to increase subjective LEV. [12] Not all of them sound binaurally correct.

In particular it is ludicrous to think that two or four surround channels as in 5.1, or 7.1, even though providing a modicum of LEV, can emulate a concert hall or produce even marginally acceptable hall realism. Damaske and Ando [6] say five is the bare minimum but bare minimums are not the Ambiophonic way. Furoya et al [7] have shown that concert hall listeners are pleased when ambience comes from lateral, rear, overhead and frontal directions in that order of importance and that such ambience should be as uncorrelated or diverse as possible. They also state that diverse late arriving ambience from such non-lateral directions is measurably significant in our perception of and preference for spaciousness in a concert hall. Thus if there were two (for low IACC) ambient fields coming from each of these directions, the minimum for realism would be eight.

Hanyu and Kimura [17] go even further and seem to have shown rather neatly that realistic LEV is the

result of mutual interactions between uncorrelated reflections coming from the sides, the rear, the front and the top. However, it is clear from the work of Evjen, Bradley and Norcross [16], Marshall and Barron, [19], and Bradley and Soulodre [24] that the lateral reflections coming from the side of the listener in the horizontal plane are the most influential. Again, the fact that early lateral reflections will not produce much of a sense of spaciousness is intuitively obvious since short delays imply walls nearby and thus not much space. Also, if early reflections are frontal, it is a given that such reflections cannot contain much information about the surrounding container. It is clear that just early or late lateral reflections alone are not sufficient to create a truly realistic concert hall ambience. This is why using just two surround speakers as in 5.1 is inadequate for classical music but reasonable for movies.

Again there is no unanimity on these points. If there were, all new halls would sound alike. Morimoto et al [18] find that both early reflections and late reflections enhance the feeling of being in a hall and can do so without seeming to increase the size of instruments on the stage. They also found that halls sound better and better as more and more of the late reverb comes from the rear rather than from the front. However, the key finding here is that rear reverb alone cannot produce a concert hall effect. They also conclude that spatial distribution of reflections is important in hall design and therefore this is important in Ambiophonic hall design.

## THE AMBIOVOLVER

If one looks at the literature as a whole, one must conclude that any preconceived notion as to the number, placement, level, and reflection parameters used to fabricate a hall in the home listening room will prove incorrect. The proper method is not to assume but to measure the impulse responses of the world's great halls and then allow the domestic concert hall denizen to select that hall that complements the music they are about to listen to and that suits the room and the number of speakers they can afford and have room to site.

The impulse responses of halls can now be easily measured before, during or after a recording session [1] and included with the CD or stored on CD-ROM or placed on a web site. The process of applying the hall impulse response to the left and right stage signals to recover the early reflections and late reverberation tails and feed these signals to the appropriate, directionally correct loudspeakers is

called convolution. We call the computing device that does this in real time an Ambiovolver. [20, 22] In actuality, an Ambiovolver is an all purpose device that can convolve the left and right channels to calculate the signals for the front Ambipole, to calculate any speaker or room corrections needed, and to compute all the hall reflections for as many surround speakers as you have placed and have the processing power to support.

From an information theory standpoint it makes little sense to record ambience. A concert hall is really just an analog computer doing thousands of mathematical operations on sound vectors for every note. But these operations are invariable for a given hall and for the duration of not only one performance but for all sound launched from its stage. Why record its response over and over again using microphones. The procedure for measuring the hall response from several positions on the stage at several of the best seats in a hall is well documented. This is in contrast to the technical confusion that reigns when the question of where to put hall ambience microphones and how to mix them comes up. [13]

Most surround sound recording engineers use extra microphones to record hall ambience or use artificial reverb devices and mixers to create two or four or more surround speaker channels. Basically, either of these methods is doomed to failure. There is no known practical recording microphone technique that can capture hall ambience without losing its directional components or becoming contaminated with direct sound.

The other great advantage of using Ambiophonics and an Ambiovolver when reproducing music is that only two media channels are needed for storing the music. Ambiophonic editing and mastering is even easier than is the case for stereo.

## ROOM TREATMENT

The Ambiophonics Institute also strongly recommends room treatment and room/speaker correction. DSP based room correctors are now widely available and can correct most speaker responses and eliminate the worst of the bass room modes. At the high frequencies, absorbent room treatment is useful to avoid erroneous early reflections of direct sound. However, the presence of four (or hopefully more) convolved surround ambience speakers mostly swamps the  $R_t$  of the room. [1] The small room essentially rereverberates the hall tails adding a few tenths of a second to the

convolved reverberation time and that is difficult to detect. However, a similar case cannot be made for spurious very early room reflections and so room and recording studio treatment remains highly desirable.

## CONCLUSION

Ambiophonics has been shown [1,2,11] to be remarkably compatible with and able to reproduce existing two channel classical music LPs, CDs, SACDs, and DVDs with realism, despite the variety of microphone, equalization, mixing, and panning techniques employed. Ongoing comparative listening tests performed by the University of Parma [1] so far confirm a preference for Ambiophonics over stereophonics, 5.1 surround and a synthesized virtual reality version of B-format Ambisonics, all derived from the same two channel live master recording. All comparisons were made using frontal music sound samples.

The Ambiovolver parcels out the early reflections and the reverberant tails to the appropriate speakers in the domestic concert hall. Fig. 8 The process is scalable and the numbers, locations, and frequency responses, of the surround speakers are not critical. The process of convolution, in contrast to using hall microphones during a recording session, nicely insures that no direct sound can get into the surround speakers. Recording engineers do not have to worry about the ratio of direct to reverberant sound in the hall or main microphones. The main microphones can be placed without regard to the hall's critical radius. Convolution of a stored impulse response makes it unnecessary to use microphones to record hall sound and also eliminates the need for more than two media (SACD, DVD) surround tracks.

In the case of Ambiophonics, as opposed to 5.1 or stereo mixing, it is normally up to listeners to decide how great a hall they need to recreate to be satisfied. They are free to select both the hall and the number of ambience speakers. Note that even a poor hall can seem as real as a good hall. However, there is nothing to prevent recording engineers from providing the impulse response of the hall they recommend or stating the file number of the hall in the eventual Internet library they wish the listener to use. It is also possible or inevitable that the media and player will control the convolver in this regard. The process would then be transparent to the unskilled user.

Another major advantage of hall ambience convolution over trying to record, store, and deliver multichannel surround ambience is that both the

locations and the number of the surround speakers in the home system are flexible, scalable, and not critical.

There is the minor question of reproducing surround applause when concerts in front of audiences are recorded. As in loudspeaker stereo, the Ambiophonic method will cause any rearward direct sound picked up by a microphone to come from the front and after convolution to also correctly come from the surround speakers. If this is considered to be a serious defect, then it is possible for record producers to code discs in the future to mute the front channels during applause intervals or whenever a convolved rearward direct sound effect is desired. Eventually disc codes would also be able to control the Ambiovolver without having to use any additional media channels to steer rear sound effects to specific surround speakers without including hall reverberation.

Even if the source material is not acoustic because it has been synthesized using virtual reality methods and panning algorithms, it is still better to reproduce such music without stereo psychoacoustic distortions such as crosstalk and pinna angle error. Such electronic music can also be convolved to set it in a pleasing, lively ambience.

In the Ambiophonic Domestic concert hall, one is free to move ones head without prejudice just as in a concert hall. One can even get up, and walk about within the circle of surround speakers (hopefully horizontal line sources or panels) but still feel that one is in a hall with a stage up front.

Once you have heard what an Ambipole, combined with room correction and an Ambiovolver can do for ordinary LPs, CDs, SACD, or DVDs it is hard to be satisfied with 2.0, 5.1, 6.0, 7.1, or 10.2 sound reproduction. There is a potential problem when playing music recorded in surround. One must mix down or extract two equivalent left and right front channels that have a minimum of rear hall sound contained in them. To the extent that this may prove difficult, virgin two-channel media will likely outperform present 5.1 or similar surround media.

A theoretical defect of hall convolution is that the impulse response is not measured for every position on the stage just usually the middle left, the middle right, and the sometimes the center. In theory every instrument position on the stage generates a different impulse response at the perfect seat. However, even if you measured 70 or 80 different stage impulse responses you could not convolve them with the instruments at those locations because at most you

only have a left side signal, a right side signal, and, in 5.1, a center signal. I can only maintain, without proof, that in the absence of direct sound, the differences in the impulse responses from instruments just a few meters apart are not detectable by the human hearing mechanism, at least to the extent that they nullify the sense of realism.

It is tempting to think that with a six channel or more digital delivery system like DVD-A you could record six or more front stage signals to get extreme width and convolve for richer ambience. The temptation should be resisted because there is no way to construct a multi-microphone and mixing system that wouldn't violate the laws of Ambiophonics and binaural technology governing the rules of one pinna, one head, no comb filtering, and strict pinna angle direction preservation particularly in the middle. It is possible to use directional microphones with such extra channels for mono direct frontal height if corresponding elevated speakers are added but this has little application in classical music recording.

I believe once recording engineers become familiar with the use of Ambiovolver that they will find it much more realistic, trouble free, and cost effective to derive the surround signals for standard 5.1 or 6.0 recordings from the impulse response of the hall than to site microphones and hope to capture hall sound during the live session. Once the classical music public becomes used to programming their own halls at home, live concert and opera recording engineers will have much less to fret about.

The development of digital signal processors and algorithms able to process digital audio in real time, without audible harmonic distortion or noise, has made it feasible and practical for music lovers to enjoy and recording engineers to deliver greater realism in music recording. Recordings made with the Ambiphone or the Schoeps KFM-6 have been shown to provide binaural realism and a normal perspective when coupled with an Ambipole and an Ambiovolver. Such Ambiophonic recordings require no spot microphone support, panning algorithms, artificial reflections, HRTF manipulation and not only need just two media channels but work better if there are just two.

Since even existing LP and CD discs reproduce well Ambiophonically, the future of Ambiophonics rests as much or more with the reproduction side than the recording side. It is hoped that the audio industry will begin to offer dedicated processors (Ambiovolver) for the home market that include effective crosstalk cancellers tailored to work with

Ambipole speaker types, room correction, and equipped to access the hall impulse response libraries needed to create great sounding domestic concert halls.

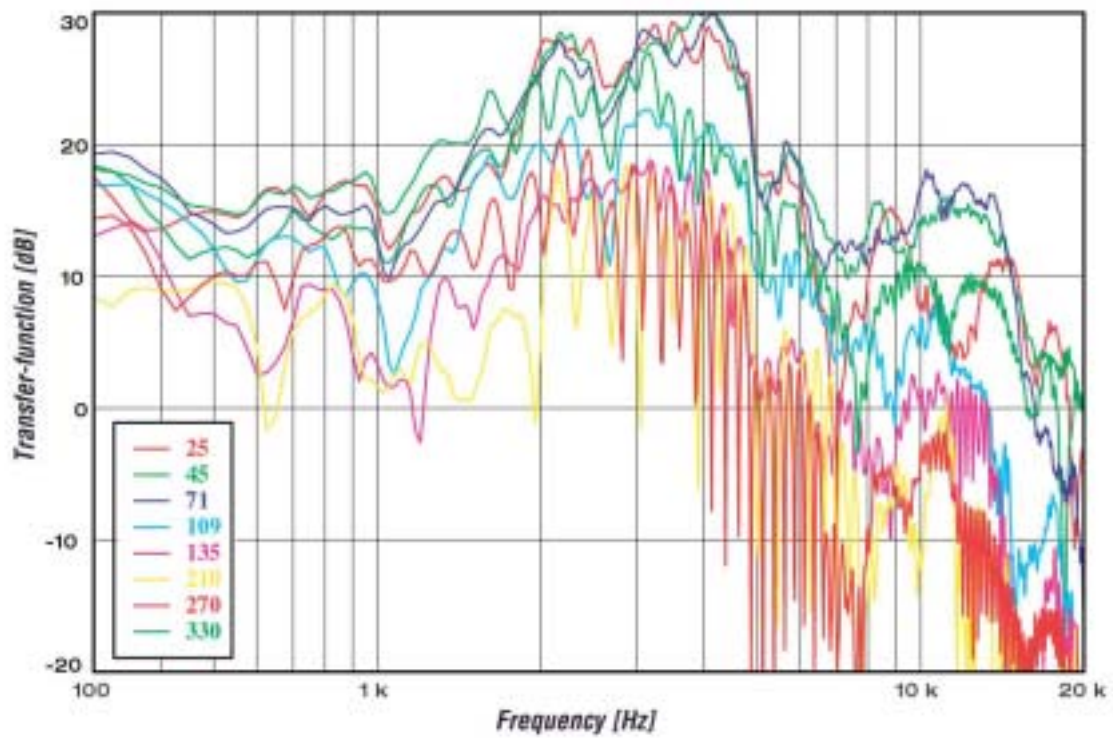
## ACKNOWLEDGEMENTS

If Ambiophonics remained just a theory it would be useless. Major contributions to advancing this technology and making Ambiophonics a living reality have come from Angelo Farina, Ole Kirkeby, Anders Torger, Robin Miller, Enrico Armelloni, and Jose Javier Lopez. They represent Italy, Denmark, Sweden, USA, and Spain and the Universities of Parma, Southampton, and Valencia. Their support has been unstinting.

## REFERENCES

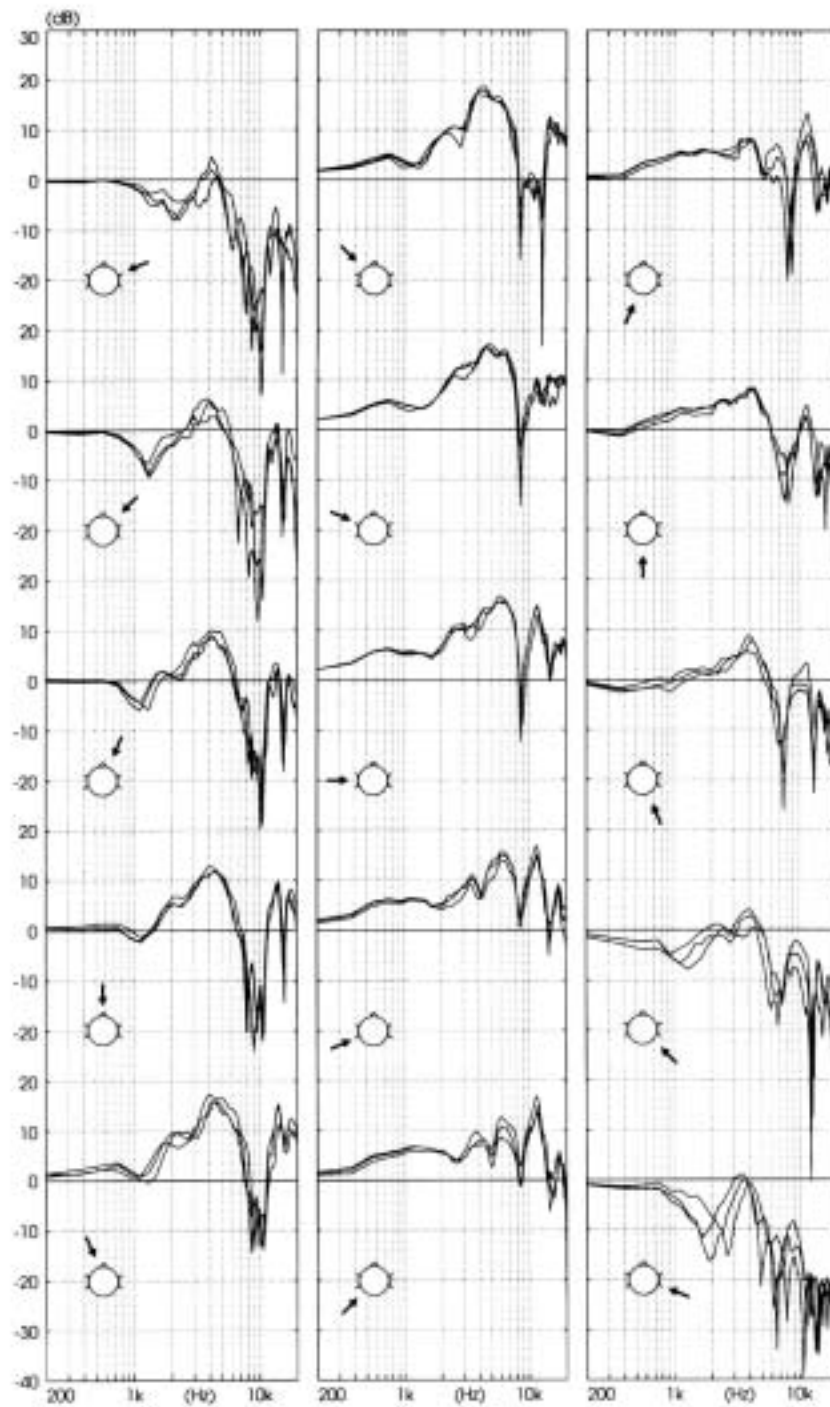
- [1] Angelo Farina, Enrico Armelloni, Ralph Glasgal. Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music. Proceedings of the AES 19th International Conference, 2001
- [2] Ralph Glasgal. The Ambiphone. Derivation of a Recording Methodology Optimized for Ambiophonic Reproduction. Proceedings AES 19th Conference on Surround Sound Techniques, 2001  
[www.ambiophonics.org](http://www.ambiophonics.org)
- [3] Jens Blauert. Spatial Hearing, 1997 Edition, MIT Press. [www.mitpress.mit.edu](http://www.mitpress.mit.edu)
- [4] James Moir. Stereophonic Reproduction. Stereophonic Techniques, An AES Anthology 1986.
- [5] William B. Snow. Basic Principles of Stereophonic Sound. Stereophonic Techniques, An AES Anthology 1986.
- [6] Damaske & Ando. Interaural Crosscorrelations for Multichannel Loudspeaker Reproduction. Acoustica, Volume 27.
- [7] Furuya, Fujimoto, Choi Young Ji, Higa. Arrival Direction of Late Sound and Listener Envelopment. Applied Acoustics.  
[www.elsevier.com/locate/apacoust](http://www.elsevier.com/locate/apacoust)
- [8] Henrik Møller, et al. Head-Related Transfer Functions of Human Subjects, J. Audio Eng. Soc., May 1995.
- [9] Ronald M. Aarts. Phantom Sources Applied to Stereo-Base Widening. J. Audio Eng. Society, March 2000.
- [10] Kirkeby, Nelson, Hamada. The Stereo Dipole. J. Audio Engineering Society, May 1998.
- [11] Ralph Glasgal. Ambiophonics, 2nd Edition.  
[www.ambiophonics.org](http://www.ambiophonics.org)
- [12] David Griesinger. The Theory and Practice of Perceptual Modeling and other papers  
<http://world.std.com/~griesngr/threedpm.pdf>

- [13] G. Theile. Multichannel Natural Music Recording Based on Psychoacoustic Principles. AES preprint 5156, Sept. 2000
- [14] T. M. Bock, D. B. Keele. The Effects of Interaural Crosstalk on Stereo Reproduction. 1986 AES Preprints 2420A & B.
- [15] G.Theile. On the Naturalness of Two-Channel Stereo Sound. J. Audio Eng. Society, Oct. 1991.
- [16] P. Evjen, J.S. Bradley, S.G. Norcross. The Effect of Late Reflections From Above and Behind on Listener Envelopment 1999.  
[www.elsevier.com/locate/apacoust](http://www.elsevier.com/locate/apacoust)
- [17] Toshiki Hanyu, Sho Kimura. A New Objective Measure for Evaluation of Listener Envelopment focusing on the Spatial Balance of Reflections 1999.  
[www.elsevier.com/locate/apacoust](http://www.elsevier.com/locate/apacoust)
- [18] Masayuki Morimoto, Kazuhiro Iida, Kimihiro Sakagami. The Role of Reflections from Behind the Listener in Spatial Impression 1999.  
[www.elsevier.com/locate/apacoust](http://www.elsevier.com/locate/apacoust)
- [19] A.H. Marshall, M. Barron. Spatial Responsiveness in Concert Halls and the Origins of Spatial Impression, 1999.  
[www.elsevier.com/locate/apacoust](http://www.elsevier.com/locate/apacoust)
- [20] Anders Torger, Angelo Farina. Real-Time Partitioned Convolution for Ambiophonics Surround Sound. Mohonk, IEEE Conference Proceedings, Oct. 2001.
- [21] Y. Ando. Concert Hall Acoustics. Springer Verlag, 1984.
- [22] Jose Xavier Lopez. PC-Based, Real-Time, Multichannel Convolver for Ambiophonic Reproduction. Proceedings, AES 19th Conference on Surround Sound Techniques, 2001
- [23] Bradley John S.,Soulodre,Gilbert A. The Influence of Late Arriving Energy on Spatial Impression. J. Acoustic Soc. Am. Vol 97. No.4, April 1995



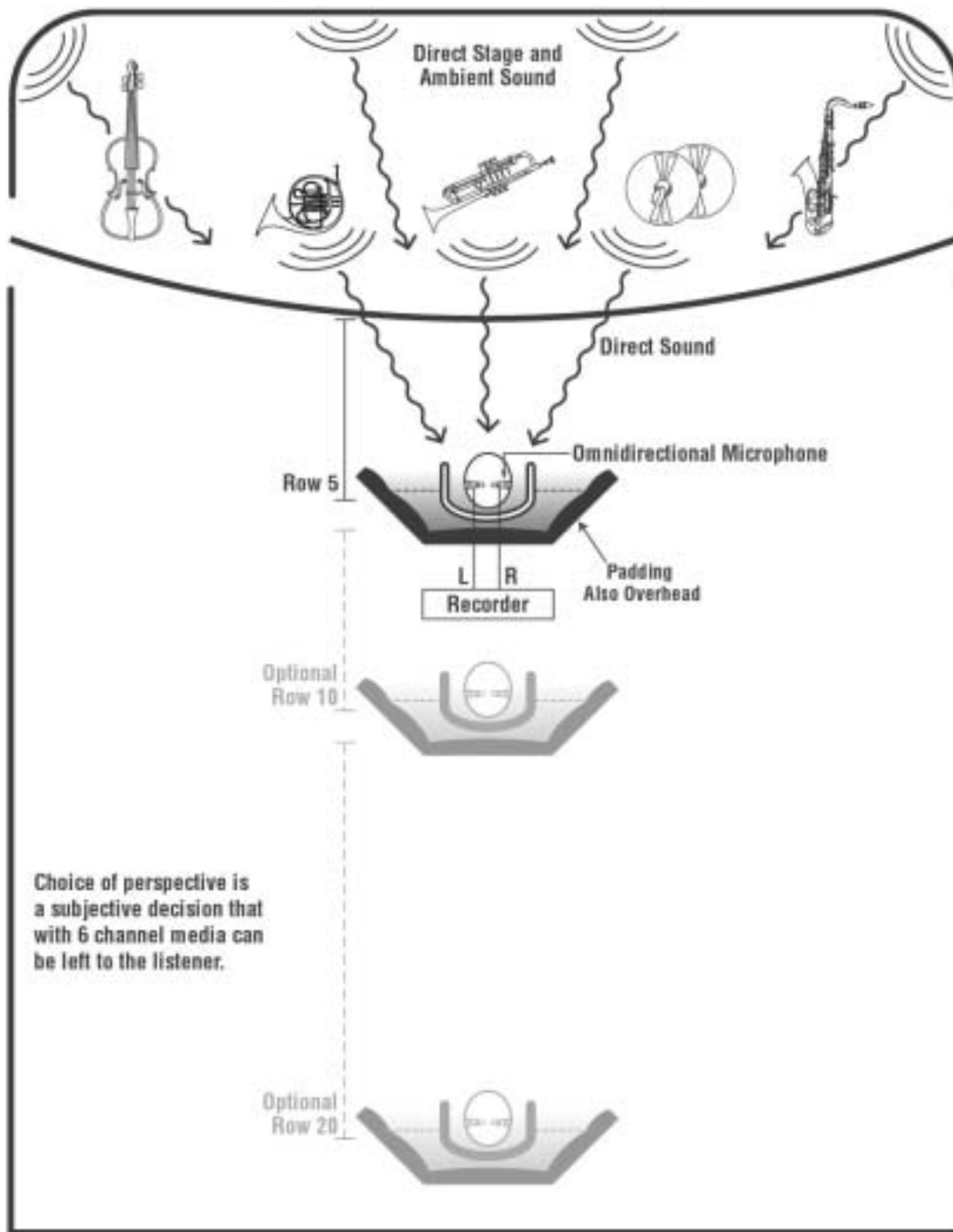
Pinna and Head Shadow Response Around the Horizontal Plane for a Single Individual  
*(courtesy of Ronald M. Aarts)*

**Figure 1** Observe the Complexity of the Pinna Direction Finding Function



Three repetition measurements of HRTFs for left ear for the horizontal plane.  
 (after Møller)

Figure 2. Head-Shadow and Pinna Functions Combined



**Figure 3: Recording with an Ambiophone**

Since Ambiophonics only requires two media channels, several two-channel versions of the same performance can be included on multichannel media allowing a home listener to choose the stage perspective they prefer

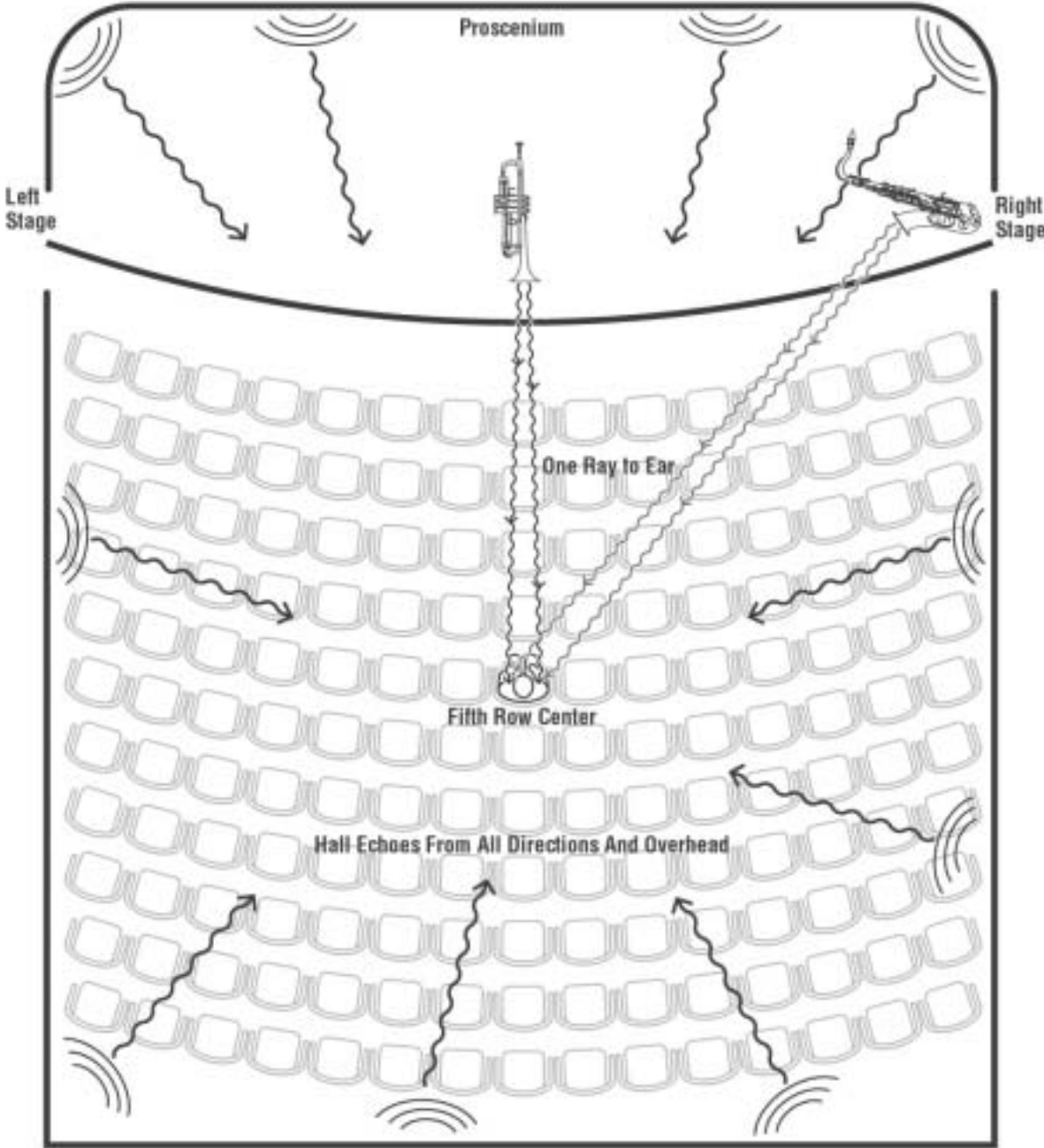


Figure 4 Normal Concert Hall Listening

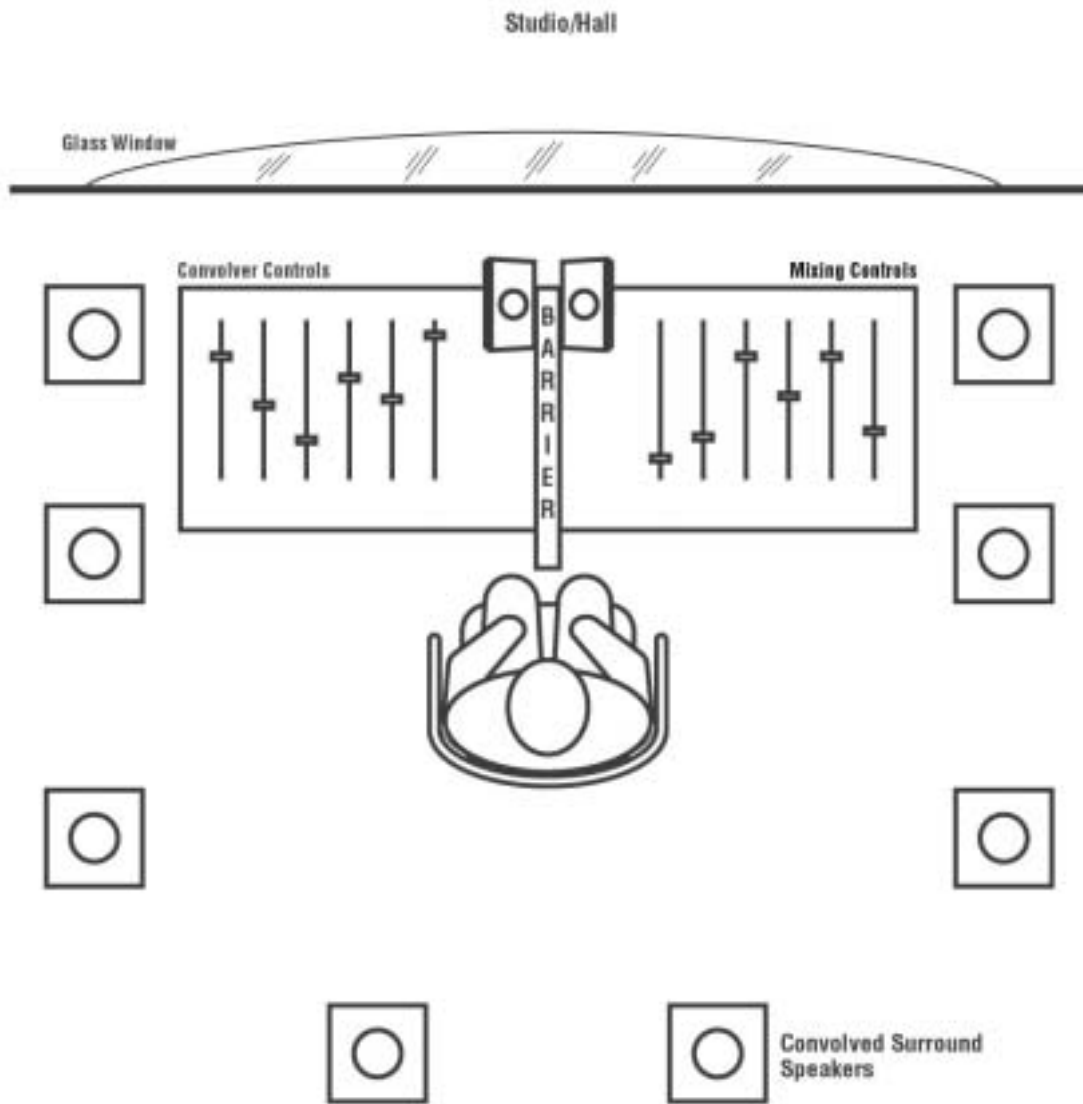
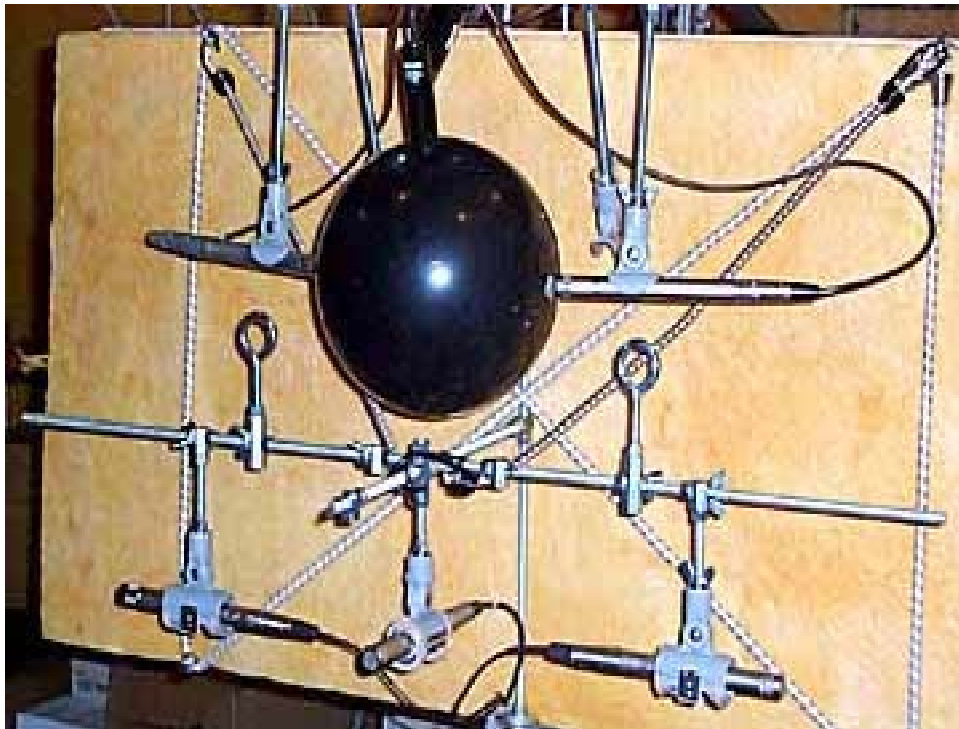


Figure 5 An Ambiophonic Recording Studio Monitoring Station



**Figure 6 Robin Miller at Ambiopole Monitoring Station**



**Figure 7 Ambiophone Mounted Against Sound Absorbing Panel Over OCT Recording Microphones**

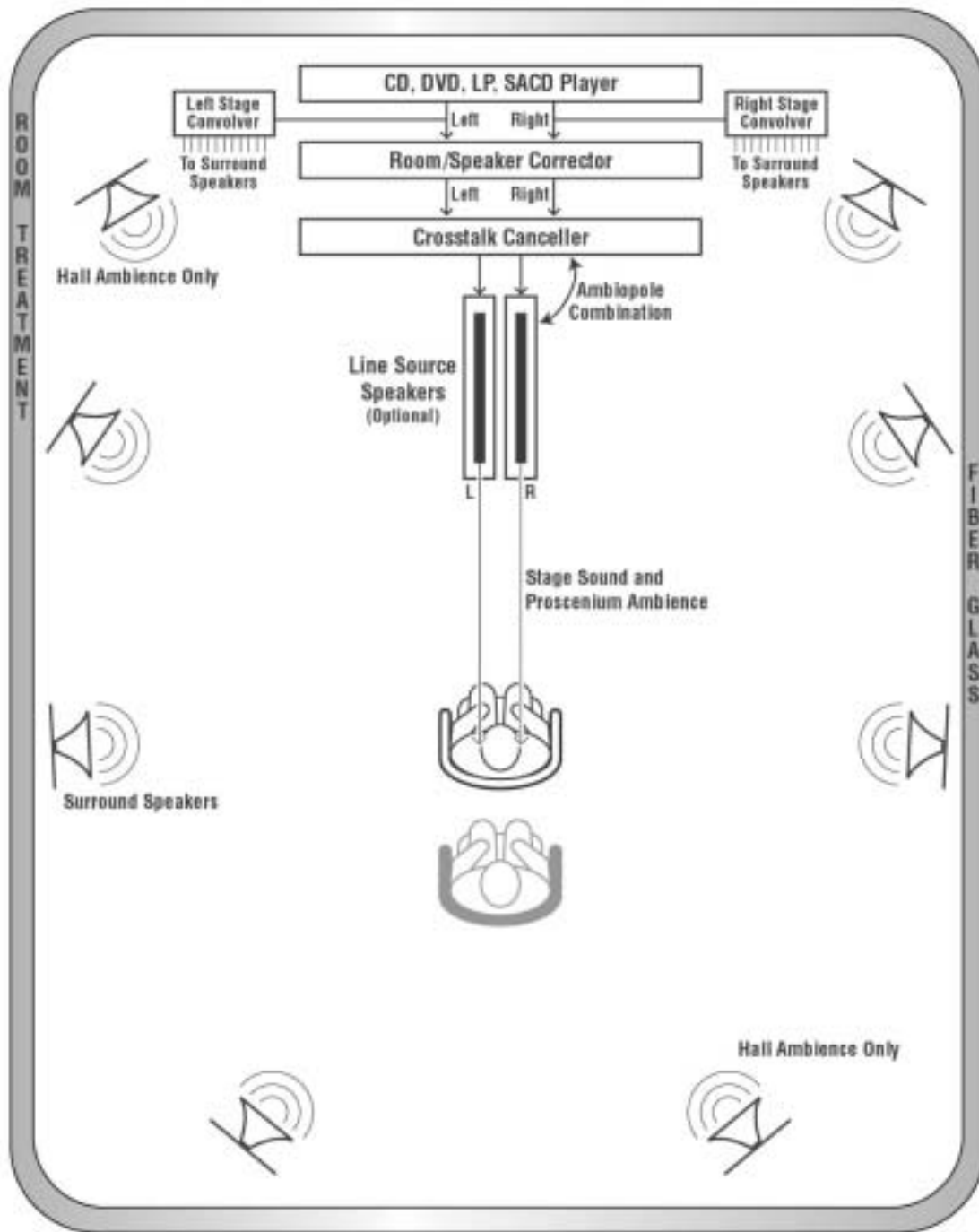


Figure 8 An Ambiophonic Domestic Concert Hall