AES 24th International Conference on
Multichannel Audio
June 26-28, 2003, Banff, Canada

SURROUND AMBIOPHONIC RECORDING AND REPRODUCTION

RALPH GLASGAL

Ambiophonics Institute
4 Piermont Road
Rockleigh, New Jersey, 07647
glasgal@ambiophonics.org
www.ambiophonics.org

ABSTRACT

Ambiophonics, Panorambiophonics, and Periambiophonics are related surround sound paradigms that reliably deliver up to full 360-degree spherical localization for both direct and ambient sound via two, four, or six DVD/SACD/MLP/DTS/Dolby/ADAT coding/media channels. They reproduce old or new, standard, 2, 4, 6, or ITU 5.1-channel music discs with unprecedented spatial realism and binaural-like localization accuracy via direct sound radiating front/rear/overhead stage-producing Ambiopoles and virtually any desired number of ambience surround speakers. Alternatively, superior acoustic recordings can be made using the described Ambiophones (or using algorithms if fabricated) to capture images of startling depth and presence for music in the round, 3D movie sound tracks, virtual reality, or electronic music soundscapes. Six-channel Periambiophonics adds elevated direct sound to the fully spherical hall ambience vectors already provided by basic Ambiophonics which drives essentially any number of hall ambience speakers regardless of their positions. All the versions of surround Ambiophonics easily deliver a "you-are-there", psychoacoustically correct, home listening experience, via home theater media, albeit best limited to one or two listeners. .

INTRODUCTION

Ambiophonics is a comprehensive sound recording/reproduction methodology, that like or unlike Stereophonics, Ambisonics, THX 5.1 or Wavefield Synthesis, prescribes hardware/software that scrupulously insures that the well known tenets of human binaural hearing [Appendix 1] are rigorously catered to so as to achieve psychoacoustic and physiological verisimilitude for one or two home listeners/viewers who seek and value "you-are-there" realism.

Ambiophonics combines crosstalk-free speaker pairs (Ambiopoles), surround speaker ambience derived from measured hall impulse responses via a convolver (Ambiovolver) and room/speaker correction/treatment to generate a binaurally correct sound field similar to wavefield synthesis. Ambiophonics creates a concert hall stage and hall from just two media channels as found on CDs, MP3s or LPs feeding a single Ambiopole. Panorambiophonics requires four media channels as provided by multichannel DVDs or SACDs each pair feeding its own Ambiopole. Periambiophonics uses six media channels as in DVD-A, DTS-EX, etc. feeding three Ambiopoles. In each type of system additional hall ambience surround speakers may also be driven via a single Ambiovolver and this is strongly recommended where music is concerned.

A single Ambiopole in front easily produces a stage of about 160-degrees in width. A single Ambiopole to the rear of the listener produces a similar rear stage width. A remarkable property of the Ambiopole software we have developed is that when both front and rear Ambiopoles are working together, they blend and the front and rear stages widen to the full 180-degrees. Thus, 360-degrees of horizontal localization becomes easily attainable for recordings

made with Ambiophones or synthesized. A third or even more Ambiopoles can be elevated over the front and/or rear Ambiopoles to add full width stages high in the air and again there is vertical fill between the stages although the extent of this phenomenon has yet to be fully investigated.

The most basic Ambiophonic theory [See Ref. list] is meant to allow previously recorded two channel media such as CDs, MP3s, and LPs to be reproduced without the well known limitations of the traditional 60-degree stereo triangle [Appendix 1], to deliver an uncompromised full width direct sound stage from two center-front speakers (an Ambiopole) and to provide real diffuse but still directional hall ambience to almost any number or location of surround speakers including elevated speakers.

It became obvious in the early development of Ambiophonics that existing stereo microphone techniques could be revised to produce better two channel recordings. Thus, on the recording side, the Ambiophone, a novel, baffled microphone arrangement, takes advantage, when recording, of the knowledge that the playback will be Ambiophonic. (not via the stereo equilateral triangle although Ambiophone recordings are actually backward compatible and sound quite normal in standard stereo) The Ambiophone also assumes that both the amplitude and the directional attributes of the early reflections and reverberant tails of the hall will be properly directed to the appropriate frontal Ambiopole and surround speakers. (Indeed, this is possible even in the case of non-Ambiophone recordings if the recorded or added reverb, unfortunately mixed into the direct frontal sound, is not too intrusive.)

After a brief review of the basics, this paper is devoted to advanced versions of Ambiophonics which take into account the 5.1, 6.0, 7.1, Dolby/THX and DTS coding/media/speaker arrangements. Special matrix coding developed by Robin Miller of Filmakers Studios [ref 17] allows two or four media channel Ambiophonic/Panorambiophonic recordings to be played back via the 5.1 speaker arrangement with clearly superior, if not ideal (compared to Ambiophonic) verisimilitude over such ITU systems. In reverse, standard 5.1 discs may also be played Panorambiophonically (described in detail below) in a manner analogous to the Ambiophonic playback of ostensibly stereo CDs or LPs, but, in this case, using front and rear Ambiopoles and surround ambience speakers driven by a hall impulse response Ambiovolver. Most 5.1 movie and music DVDs or music SACDs reproduce exceptionally well this way especially when compared with the ITU 5.1 standard speaker arrangement. [Appendix 1]
Panorambiophonics, described below, uses four channel coding/media such as Dolby, DTS, SACD, or DVD-A to deliver an easily localizable 360-degree direct sound stage as in movies, or, for concerts, a very wide front stage that, if in a hall, automatically includes horizontal 360 degree hall ambience. A four channel recording mic, the Panorambiophone, has been designed to make such recordings. Only four speakers (two Ambiopoles) are used in Panorambiophonic reproduction to reproduce all horizontal plane direct sound and horizontal hall ambience with full circle normal binaural localization physiology. Where the direct sound recording has been made in a dry or small studio, it is possible to enhance the reproduction of these front and rear direct sound fields by adding ambience surround speakers driven by an hall Ambiovolver as in standard Ambiophonics.

Periambiophonics adds a third elevated Ambiopole to Panorambiophonics to provide for a full direct sound stage in all dimensions including some height. The elevated Ambiopole can be used for direct sound reproduction or ambience. In the latter case this allows a concert-hall direct sound performances to be recreated in a home with just three speaker pairs and no surrounds.. Using three direct sound Ambiopoles allows movies, virtual reality, games and soundscapes to sound more like the live experience. Furthermore, Periambiophonics can combine six-channel Periambiophone recording, and the front, rear, and elevated Ambiopoles, with an Ambiovolver to add virtually any desired number of surround speakers so as to deliver physiological verisimilitude of a concert hall experience that also includes rear or overhead direct sound sources to a home listener via standard DVD/SACD media. Clearly, both Panorambiophonics and Periambiophonics are well suited to capture, create and reproduce 3D electronic music or virtual reality projects.

This paper reviews the theory, techniques, and features, of the hardware and software required to make these various kinds of Pan/Peri/Ambiophonic recordings and to reproduce these as well as stereo CDs and the various multichannel surround media.
.

REVIEW OF BASIC AMBIOPHONICS (Fig. 1)

.The simplest form of Ambiophonics is meant for the playback of ordinary stereo CDs, LPs, SACDs, MP3s, cassettes, stereo TV, etc. In stereo, the front stage is created between the speakers, in Ambiophonics the stage is created from the speakers

outward and so can be much wider. The Ambiopole speaker pair form an angle to the listener of from twenty to thirty degrees. One can sit anywhere along the line between the speakers and can stand or recline, turns one head, lean etc. Some omnidirectional speakers produce larger sweet areas but in general the best Ambiopoles can only accommodate one or two listeners, perhaps best one behind the other for golden eared results.

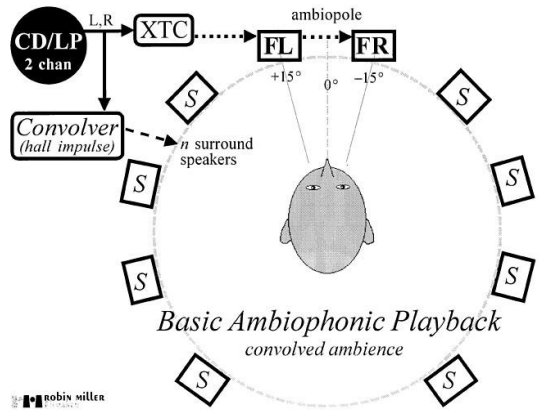

Figure 1

The scattered surround speakers are fed hall ambience signals calculated for both the left and right channels by a computer which we call an Ambiovolver. The Ambiovolver has stored within it the impulse responses of some of the great halls, churchs, and auditoriums of the world and more such hall signatures are being accumulated all the time. One simply selects the hall best suited to the recording or the actual hall where the recording was made. The Ambiovolver is told the location of all the surround speakers in the room and it then generates the appropriate reflections and feeds them to a surround speaker that can then mimic a concert hall wall. In this way the levels, frequency responses, and the directionalities of the reverberant field are maintained. I have driven up to 24 surround speakers this way and, while clearly overkill, the results are gratifying. This is in contrast to normal 5.1 practice where recorded hall ambience whether from front, rear, overhead or the side is lumped together and launched from just two surround speakers. The attached references describe Ambiovolver design, hall impulse response measuring procedures and hall acoustic properties.

It is always desirable to keep the listening room early reflection characteristics under control. Absorptive panels are quite effective. However, since the direct sound speakers are so close together and aimed forward, they are easier to position than for stereo or 5.1. Bad room acoustics are actually less of an issue in Ambiophonics than in stereo. Except that when room reflections interfere in stereo most listeners can hardly notice or care enough. In Ambiophonics the average listener will notice bogus early reflections and feel deprived. It is possible to adjust the Ambiovolver to compensate for listening room late reverb tails.

2 CHNL AMBIOPHONIC RECORDING (fig. 3)

While for many people, with large CD collections, basic Ambiophonics will sound as good as they wish, others will find enjoyment in the improvement that can be achieved by making recordings specifically meant to be played back over Ambiophonic systems. The Ambiophone recording microphone assembly was designed to make this feasible. (fig. 2)

The derivation of the Ambiophone has been described in detail in the earlier referenced papers. Basically it is a head shaped ball with two omnidirectional microphones mounted flush where the ear canals would be. The microphone is baffled. That is it faces forward and is shielded from sound originating from overhead, the rear or the extreme sides.

Fig. 2 Ambiophone Prototype Before Placement of Omni Capsules Within the head

The microphone is placed first to fifth row center depending on taste. The perspective one hears during reproduction is the same as if one were at the mic position during the recording session. The usual considerations of hall radius or ratios of direct to reverberant sound do not apply here since the mic is baffled. Since all hall ambience will be generated

from this or other great hall impulse responses, it is not necessary to record hall reverb during the recording session. The Ambiophone must also collect horizontal frontal or proscenium ambience since this indirect sound should emerge from the Ambiopole with the direct stage sound. The head shape of the Ambiopole provides the Interaural Level Difference for sounds from the stage sides. Otherwise, the Ambiopole, being centered in front of the home listener, would not provide this. The Ambiophone captures both correct ILD and ITD (see Appendix) compared to coincident microphone techniques, spaced omnis, spot mic mixing, etc. The Schoeps KFM-6 turns out to be a good match for an Ambiophone, if baffled during use.

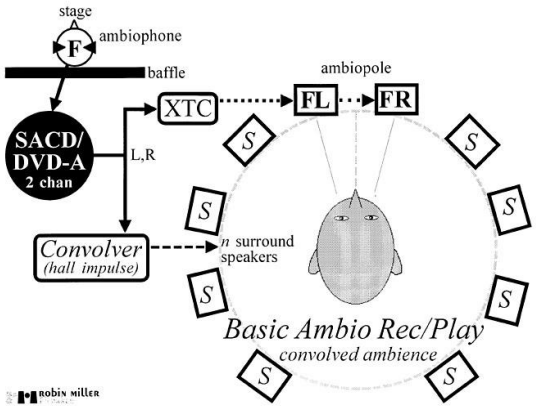

Fig. 3. Two Channel Ambiophonic System

Where classical music, reproduced in the home environment is concerned, two channel Ambiophonic recording and reproduction should satisfy even the most golden-eared audiophiles. Ambiophonics appears to entirely swamp the digital sampling differences between 2 channel media such as CD, SACD and DVD-A. It would be an interesting double blind project to see if the different media can actually be distinguished when Ambiophonic conditions prevail during both recording and playback..

AMBIO PLAYBACK OF 5.1 DISCS (fig. 4)

Home theater surround movies or music recordings can be played back Ambiophonically rather than stereophonically in a manner analogous to the playback of CDs and LPs, discussed above. (The psychoacoustic disadvantages of the LCR reproduction scheme are reviewed in the appendix.) The left and right frontal 5.1 channels can be fed directly to a crosstalk canceller and thence to an Ambiopole. An Ambiopole can also be used simultaneously as a center speaker pair simultaneously and so it is an easy matter to add the unaltered center information to the other two Ambiophonic signals. One can also imagine a home theater processor equipped to drive two center speakers with a switch to go between stereo LCR and the CC Ambiopole. Certainly, it is easier to set up the front part of a home theater system using just two center speakers 20 or 30 degrees apart so as not to stand in front of the TV screen than setting up three speakers that must be equidistant and spaced symmetrically. Also, for home TV, viewers like to be centered so the major supposed advantage of the LCR arrangement seems of limited value.

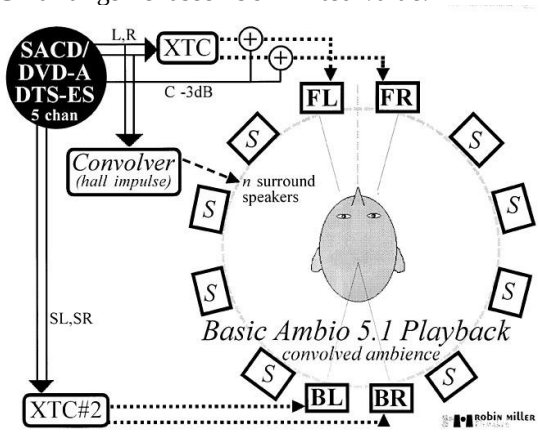

Fig. 4. Ambiophonic Playback of 5.1 Material

The two rear surround channels go to a second crosstalk canceller and a rear Ambiopole. For many movies this arrangement produces a rear stage with excellent localization over some 160 degrees in the rear. Of course, this applies only to movies that were recorded in stereo in the rear, not just dual mono or fabricated sound effects, or ambience. Music DVDs often include real hall ambience captured during the performance and again, if not mono, can provide an ambient field spread across the rear. While not ideal, since ceiling rear and frontal ambience comes from this horizontal rear arc, this effect is better than the standard ITU plus and minus110-degree arrangement whose properties are discussed in the Appendix.

Better yet, for much music, where there are no instruments or vocals in the rear, the rear surround channels can be ignored and the more natural ambience generated by the Ambiovolver can be used instead free from the constraints of the ITU surround speaker position mandate. It is also possible to use both the Ambiovolver and the rear Ambiopole simultaneously. For those who want applause coming from the rear, this arrangement works well for both live music and movies.

PANORAMBIOPHONICS (Fig. 5)

If a four channel medium such as SACD, DVD, or DTS-CD is available then it is possible to record direct sound sources over 360-degrees in the horizontal plane. A special microphone assembly which we have called a Panorambiophone (or Panambiophone and Panambio for short) is used to capture signals appropriate for reproduction via one front Ambiopole and one rear Ambiopole. Front and rear Ambiopoles do merge seamlessly. We have already demonstrated to hundreds of visitors that the combination of the Panambiophone and the two Ambiopoles does indeed allow normal binaural localization over the full circle including the 90-degree positions at the extreme sides. This localization is possible for most individuals even if their heads are fixed. But it is also true that slight head movements focus the azimuth for some individuals at some problem angles mostly to the rear.

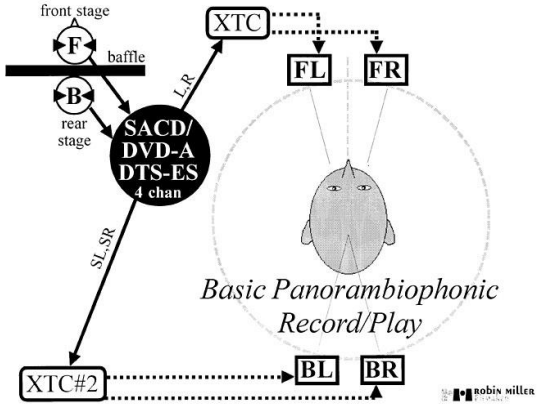

Fig. 5 Panorambiophonics Delivers Full Circle Localization via 4 Channels and Speakers

The Panambiophone consists of two Ambiophones placed one behind the other but still both facing in the forward direction. The two head shaped balls must be placed nose to baffle to nape since if there is too much separation the differential delay of slightly off 90-degree direct sound sources will cause comb filtering when reproduced. A baffle between the two Ambiophones insures that the front stage is picked up mainly by the front ball and the rear stage is mainly heard by the rear ball. For concert music both balls should be protected from overhead reflections. The Panambiophone, like the Ambiophone is placed at the best seat in the house or at the center of the sonic action.

In reproduction, two crosstalk cancellers feed two Ambiopoles and listeners should sit, stand, or recline on the line between them for best effect. Off line seating still yields interesting front/back localization but the exact angles are unpredictable.

During symphonic recording, the rear Ambiophone picks up the rear half horizontal hall ambience while the front Ambiophone automatically captures the front half direct and ambient sound. Thus, one can have a reasonable you-are-there 360 degree ambience sound experience with just four speakers if the recording venue was not just a studio or an auralization computer. This methodology should be compared with the random difficulties encountered using other ambience pickup microphone arrangements such as IRT, Fukada, Williams, Decca, etc.

ADVANCED PANAMBIO (Fig. 6)

If the Panambiophone is used in the concert hall during a live performance, the ambient field cannot be captured with precision because the overhead ambience is either mixed in with the horizontal components or excluded if the microphone is shielded from the ceiling. Likewise in circular direct sound recordings made in studios or on location, it is difficult to include a realistic ambient field as part of the four channel medium. Thus, there are good reasons to use an impulse response, an Ambiovolver, and surround speakers to enhance the performance of basic Panambio.

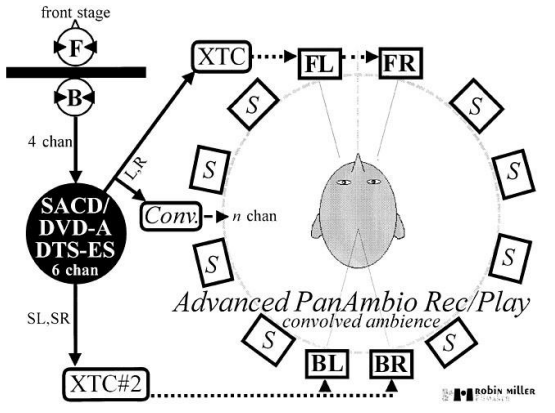

Fig. 6 Panorambiophonics Including Both Circular Direct Sound and Hall Ambience

An example of this is Robin Miller's Panorambiophonic recording of an outdoor parade. The bands pass by in front and the crowd shouts and claps come from behind and from the sides. But when you add impulse response derived reflections from the surrounding office buildings to the mix the scene becomes that much more vivid. Another

example is Robin Miller's studio recording of country music with a boisterous audience present. The Ambiopoles take care of all the direct, front and rear, tunes and shouts but the Ambiovolver transports the whole scene to Nashville.

PERIAMBIOPHONICS (Fig. 7)

Periambiophonics adds height to the Ambiophonic mix. This requires another pair of media channels but ADAT, DVD-A and DTS-ES are examples of commercially available systems capable of delivering sufficient data to create close to full upper hemisphere soundscapes. Another baffled Ambiophone head is used to capture an elevated front stage and this signal pair needs to be fed to a crosstalk canceller and an elevated Ambiopole. If one has four pairs of media channels available, via ADAT for instance, then the two front stages and the two rear stages can produce virtually anything desired. To date only frontal and rear elevated stage merging has been tested and Periambiophonics is still a work in progress. The real issue is whether there is a viable commercial home market for such a direct sound technology.

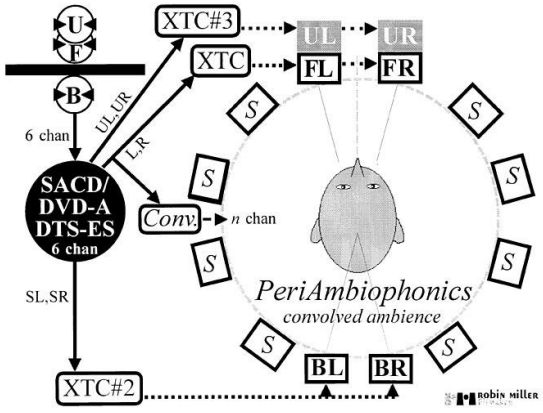

Fig. 7 Periambiophonics Provides An Elevated Stage

Figure 7 also shows that convolved ambience can provide periphonic envelopment. As discussed above all the various Ambiophonic methods can employ the Ambiovolver to produce signals for surround speakers at any azimuth or elevation if the impulse response used has been taken in three dimensions.

AMBISONICS + AMBIOPHONICS (Fig. 8)

In the absence of a three ball Ambiophone or impulse response expertise, it is possible to make live recordings using a single Ambiophone facing forward to catch the stage and an Ambisonic WXYZ coincident microphone just behind it to only record hall ambience in all dimensions. The Ambisonic array must be baffled to prevent it from picking up frontal direct sound. Six media channels are also required to store this version of periphonic sound.

Instead of an Ambiovolver an Ambisonic decoder is required to deliver the ambience to the surround speakers. Normally, Ambisonic surround speakers must be symmetrically arranged about the listening spot. However, when only ambient sound is being handled Ambisonically, the requirements are less stringent. Also the more speakers used the better the results.

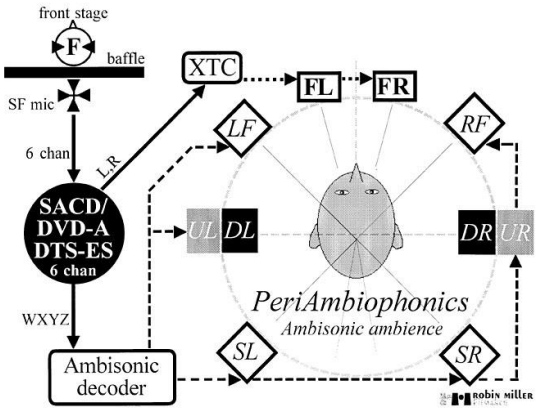

Fig. 8 Ambisonic Techniques Can Be Used to Provide Surround Ambience With Height

Since Ambisonic technology and some hardware and software has been available since the 70s, this route may be more attractive to researchers than the Ambiovolver approach. The advantage of the Ambiovolver however, is that hall ambience need not be recorded during the performance. There is also the complexity of the various Ambisonic four-capsule microphones and control units. Ambisonic techniques are also often used to capture hall impulse responses

CONCLUSION

It took 25 years for stereophonics to seriously begin to replace monophonics. It is likely that a similar period will be required for Ambiophonics to replace or at least supplement stereophonics and its twin brother 5.1. But the development of digital signal processors and algorithms able to process digital audio in real time, without audible distortion or noise, has now made it feasible and practical for music/movie lovers to enjoy and recording engineers to deliver greater physiological verisimilitude in music and video recording. Recordings already made

with the various varieties of Ambiophones can be demonstrated to all who are interested or doubting. Ambiophonics provides binaural realism and a normal stage perspective when reproduced via one or more Ambiopoles. Ambiovolver driven surround speakers easily provide surround ambience without requiring media bandwidth or recording session bother. Ambiophonic recordings should need no spot microphone support, panning algorithms, artificial reflections, or HRTF manipulation and consume just two media channels for classical music or four or six if rear and/or overhead sound stages are desired. Best of all, not only is Ambiophonic reproduction of existing CDs and LPs superior to stereo triangle reproduction but Ambiophonic surround reproduction of 5.1 DVDs and SACD is also psychoacoustically superior and easier to implement than the ITU 5.1 speaker arrangement. However, Ambiophonics is for domesticity and is not suitable for large group listening applications.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] Ralph Glasgal. Ambiophonics, 2nd Edition. www.ambiophonics.org
[2] William B. Snow. Basic Principles of Stereophonic Sound. Stereophonic Techniques, An AES Anthology 1986.
[3] James Moir. Stereophonic Reproduction. Stereophonic Techniques, An AES Anthology 1986.
[4] Jens Blauert. Spatial Hearing, 1997 Edition, MIT Press. www.mitpress.mit.edu.
[5] 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
[6] Ralph Glasgal. The Ambiophone. Derivation of a Recording Methodology Optimized for Ambiophonic Reproduction. Proceedings AES 19th Conference on Surround Sound Techniques, 2001 www.ambiophonics.org
[7] Ralph Glasgal. Ambiophonics, Achieving Physiological Realism in Music Recording and Reproduction. Preprint 5426 AES 111th Convention Dec. 2001
[8] Jens Blauert. Spatial Hearing, 1997 Edition, MIT Press. www.mitpress.mit.edu.
[9] Kirkeby, Nelson, Hamada. The Stereo Dipole. J. Audio Engineering Society, May 1998.
[10] T. M. Bock, D. B. Keele. The Effects of Interaural Crosstalk on Stereo Reproduction. 1986 AES Preprints 2420A & B.
[11] G.Theile. On the Naturalness of Two-Channel Stereo Sound. J. Audio Eng. Society, Oct. 1991.
[12] Masayuki Morimoto, Kazuhiro Iida, Kimihiro Sakagami. The Role of Reflections from Behind the Listener in Spatial Impression 1999. www.elsevier.com/locate/apacoust
[13] A.H. Marshall, M. Barron. Spatial Responsiveness in Concert Halls and the Origins of Spatial Impression, 1999. www.elsevier.com/locate/apacoust
[14] Anders Torger, Angelo Farina. Real-Time Partioned Convolution for Ambiophonics Surround Sound. Mohonk, IEEE Conference Proceedings, Oct. 2001.
[15] Jose Xavier Lopez. PC-Based, Real-Time, Multichannel Convolver for Ambiophonic Reproduction. Proceedings, AES 19th Conference on Surround Sound Techniques, 2001
[16] Anders Torger. Ambiovolver Instruction Manual Link. www.ambiophonics.org
[17]R Miller, "Transforming Ambiophonic + Ambisonic 3D Surround Sound to & from ITU 5.1/6.1," Preprint, AES 114th Convention, Amsterdam, March 2003.
[18] R Miller, "Compatible PanAmbiophonic 4.1 and PerAmbiophonic 6.1 Surround Sound for Advanced Television Beyond ITU 5.1," SMPTE 144th Technical Conf. & Exhibition, Pasadena CA. Oct. 02
[19] R Miller, "Contrasting ITU 5.1 and Panor-ambiophonic 4.1 Surround Sound Recording Using OCT and Sphere Microphones," AES 112th Convention, Munich, May 2002, preprint #5577.
[20] Robin Miller, "PerAmbiolating 360°: Panorambiophonic 4.0" multichannel evaluation DTS-CD, "PerAmbiolating 360°: Companion ITU 5.0," multichannel evaluation DTS-CD, with print insert, "Surround" pages, www.filmaker.com.
[21] R. Miller, "Ambiophonic Demonstration: AES 19th, Bavaria 2001" 2-chan. CD with print insert, and PS online Ambio demonstration, "Surround" pages, www.filmaker.com.

APPENDIX

I proclaim these psychoacoustic truths to be self evident.

Human sound localization is possible using three and only three sonic clues (not counting bone conduction)

1. Time, including phase and transient edge, differences between the ears. This ITD includes the precedence effect.

2. Sound level differences between the ears. (ILD)

3. Single and twin eared pinna direction finding effects.

Each of these mechanisms is only effective in a specific frequency range but they overlap and the predominance of one over the other also depends on genetics, the nature of the signal, i.e. sinewave, pink noise, music, or venue, etc.

For a full range complex sound such as music, experienced live, all three mechanisms are always in play and normally agree. By definition such an experience is said to be realistic or, better phrased for the creative and artistic recording fraternity, said to yield guaranteed physiological verisimilitude. If the three mechanisms are not consistent then we often make errors in localization such as in most earphone listening where the interference with the pinna and head shadow usually result in internalization even if the ITD, including some ILD crosstalk, is perfect.

Before we get to Stereophony, let me discuss the relative strengths of the three mechanisms listed above. Snow and Moir in their classic papers showed that localization of complex signals in the pinna range above 1000 Hz was superior by a few degrees, to localization that relied solely on complex lower frequencies. That is, their subjects could localize bands of high frequencies to within one half a degree but only to one or two degrees at lower frequencies. The accuracy of localization, in general, declines with frequency until at 90 Hz or so, as Bose has demonstrated, it goes to zilch. Remember this when we get to discuss crosstalk.

It is important for understanding the workings of Stereophony that you are convinced that all three mechanisms are significant and I would suggest, with Keele, Snow and Moir, that the Pinnae are first among equals. You should satisfy yourself on some of this by running water in a sink to get a nice complex high frequency source. Close your eyes to avoid bias, block one ear to reduce ILD and ITD, and see if you can localize the water sound with just the one open ear. Point to the sound, open your eyes, and like most people you will be pointing correctly within a degree or so. With both ears you should be right on despite having a signal too high in frequency to have much ITD or ILD. But with two pinnae agreeing and the zero ILD clue, the localization is easily accurate.

Again, if a system like stereo or 5.1 cannot deliver, the ITD, ILD and Pinna cues intact without large errors it cannot ever deliver full localization versimilitude for signals like music. If the cues are inconsistent, localization may occur but it is fragile, it may vary with the note or instrument played, and such localization is usually accompanied by a sense that the music is canned, lacks depth, presence, etc. Mere localization is no guarantee of fidelity.

Let us now look at the stereo triangle in reproduction and the microphones used to make such recordings and see what happens to the three localization cues. Basically Stereophonics is an audible illusion, like an optical illusion. In an optical illusion the artist uses two dimensional artistic tricks to stimulate the brain into seeing a third dimension, something not really there. The Blumlein stereo illusion is similar in that most brains perceive a line of sound between two isolated dots of sound. Like optical illusions, where one is always aware that they are not real, one would never confuse the Stereophonic illusion with a live binaural experience. For starters, the placement of images on the line is nonlinear as a function of ITD and ILD, and the length of the line is limited to the angle between the speakers. (I know, everyone, including Blumlein, has heard sounds beyond the speakers on occasion but diatribe space is limited.)

I want to get to the ILD/ITD phantom imaging issue involved in this topic. But let us first get the pinna issue tucked away. No matter where you locate a speaker, high frequencies above 1000 Hz can be detected by the pinna and the location of the speaker will be pinpointed unless other competing cues override or confuse this mechanism. In the case of the stereo triangle the pinna and the ILD/ITD agree near the location of the speakers. Thus in 5.1 LCR triple mono sounds fine especially for movie dialog. In stereo, for central sounds, the pinna angle impingement error is overridden by the brain because the ITD and the ILD are consistent with a centered sound illusion since they are equal at each ear. The brain also ignores the bogus head shadow since its coloration and attenuation is symmetrical for central sources and not large enough to destroy the stereo sonic illusion. Likewise, the comb-filtering due to crosstalk, in the pinna frequency region, interferes with the pinna direction finding facility thus forcing the brain to rely on the two remaining lower frequency cues. All these discrepancies are

consciously or subconsciously detected by golden ears who spend time and treasure striving to eliminate them and make stereo perfect. Similarly, the urge to perfect 5.1 is now manifest.

Consider just the three front speakers in 5.1. Unless we are talking about three channel mono, we really have two stereo systems side by side. Remember, stereo is a rather fragile illusion. If you listen to your standard equilateral stereo system with your head facing one speaker and the other speaker moved in 30-degrees, you won't be thrilled. The ILD is affected since the head shadows are not the same with one speaker causing virtually no head shadow and the other a 30 degree one. Similarly the pinna functions are quite dissimilar. (In the LCR arrangement the comb-filtering artifacts now are at their worst in two locations at plus and minus 15-degrees instead of just around 0-degrees as in stereo) Thus for equal amplitudes (such as L&C) where a signal is centered at 15 degrees, as in our little experiment, the already freakish stereo illusion is badly strained. Finally, the ITD is still okay and partly accounts for the fact that despite the center speaker there is still a sweet spot in almost all home 5.1 systems. Various and quite ingenious 5.1 recording systems try to compensate for some of these errors but the results are highly subjective and even controversial. It is also probably lucky that in 5.1 recording, it is difficult to avoid an ITD since a coincident main microphones is seldom used in this environment.

TECHNICAL DIGRESSION FOR RECORDING ENGINEERS.

Before getting to side imaging, there are some points on the relationship between microphones and reproductive crosstalk that should be elucidated. Whether crosstalk is beneficial or not depends on what frequency range you are talking about and thus what localization method you are relying on. At high frequencies, in the pinna range, stereo speaker crosstalk is obviously not a benefit. There is no way that this unpredictable pattern of peaks and valleys can enhance localization in a stereo or LCR system This is true whether spaced or coincident mic's are used.

Stereo crosstalk can cause a phase shift at frequencies below where comb filtering predominates. That is, two sinewave signals with slightly different delays but with comparable amplitudes will combine to form a new sinewave with some different amplitude and phase angle. I maintain that the phase part of this change is inaudible from 90 Hz on down, nonexistent for the central 10-degrees and virtually non-existent for images from the far right or left, and thus of doubtful audibility in between or in LCR systems. Stereo crosstalk cannot create an ITD for transients captured in coincident mic recordings but it can shift the phase of midbass and low bass. But there is no evidence that the small phase shifts of this type are audible or affect localization. If spaced mic's are used, then there is an ITD and crosstalk has little deleterious effect but likewise no benefit.

The ILD is a slightly different story. In the low bass, say below 90 Hz, the phase difference between the direct sound and the crosstalk sound is too small (heads are too small) to cause any significant change in phase and thus change amplitude at an ear when the two signals are added together. So regardless of the microphone used, low bass crosstalk is not the issue. Again, I maintain that the very low bass energy at both ears remains almost the same even if the left and right signals are different in amplitude as in coincident mic'ing. As Blumlein observed, as the frequency goes up the path length difference is equivalent to larger phase angles and so, if there is a difference in amplitude, but not phase, between the speakers, the signals will go up at one ear and down at the other as the signals are combined on each side of the head. Clearly if the phase shift gets to 90-degrees on up this same crosstalk mechanism becomes detrimental. This boost in mid bass separation is only applicable to phantom stereo images around 15 degrees. In the center there is no crosstalk amplitude asymmetry to take advantage of and at 30 degrees where the speakers are, hopefully, the stereo separation ensures that the crosstalk has little to add to or subtract from.

If spaced microphones are used, the ILD at low frequencies may be minimum especially for omnis. But let us assume that above 90 Hz there is a substantial ILD as well as an ITD. In this case the LF effect of the crosstalk phase change is sort of unpredictable. Again in the 15 degree region there could be enhancement of bass separation but the ITD induced phase shift could counter this. In summary, crosstalk is really only desirable in the case of coincident mic stereo recordings, as Blumlein wrote, and only if restricted to frequencies below 300 Hz or so as I claim.

SURROUND SOUND LOCALIZATION

Let us consider surround sound localization. Obviously, if a mono signal is placed at 110 degrees it can be localized using pinna, ILD, and ITD even when facing forward. Between the two rear surround

speakers you have effectively a stereo pair spanning 140 degrees. In such a situation, if there is a lot of high frequency energy, the pinna will localize to the speakers and it will be difficult for some individuals to hear sound directly behind or in the central rear region. (The new rear surround channel can fix this, but the LCR anomalies as above will then apply.) However, if there is a real ITD and a real ILD between the rear speakers it is theoretically possible to hear a wide stage to the rear as in the frontal stereo illusion. However the crosstalk, and thus the comb-filtering, is extreme at this angle and it starts at a lower frequency thus interfering with the ILD at 800 Hz or lower. If there is an ITD this can help but then the speakers must be properly placed or delay adjusted. Obviously, if 140-degree spacing was a good way to make a stereo stage, front or rear, it would have been done this way long before now.

Finally, let us see what happens when we try to image from the front side speaker to a speaker at 110 degrees on the same side while facing forward. In the case of the pinnae, the pinna facing the speakers can localize to each speaker discretely if the signals are different. If they are correlated or identical, the brain will use some other cue to localize. There may be some gifted individuals who can localize high frequency phantoms between the speakers using one pinna but I can't do it. The higher frequencies also go around the head to produce a head shadow and this at least allows the brain to decide the source is at the loud side.

If there is a time difference, then the two signals from each speaker reach the exposed ear canal and add together to produce garbage and a head shadowed version of this time garbage also reaches the far ear. Basically, regardless of the recorded TD, the ITD the brain perceives is always the ITD based on one's ear spacing. However this is sufficient to localize to the louder side but makes localization between the speakers wishful thinking.

If there is a level difference, then the two signals from each speaker reach the exposed ear canal and add together to produce garbage and a head shadowed version of this level difference garbage also reaches the far ear. Basically regardless of the recorded LD, the ILD the brain perceives is always the ILD based on one's head shadow. However, as above, this is sufficient to localize to the louder side but makes localization between the speakers wishful thinking.

That the above scenario is more or less correct is attested to by the fact that the industry keeps adding more speakers to correct these defects. We have the rear speaker, height speakers, and the 7.1 and 10.2 proposals, etc.