tmt27

expertise in audio media

27. Tonmeistertagung
VDT International Convention

22.11.– 25.11.2012 Köln | Cologne

Tagungsbericht
Conference Proceedings

ISBN 978-3-9812830-3-7

Properties of captured 3D audio / Auro 3D room signals

(Eigenschaften von mikrofonierten Raumsignalen bei 3D Audio / Auro 3D)

Translation from German script

Lasse Nipkow
Silent Work LCC / lasse.nipkow@silentwork.com

Abstract

In the case of aurophonic playback systems such as Auro 3D, the property of the room sound is significantly more important than in the case of 2-channel stereo, particularly in the case of classical music. With stereo, direct and room sound can only be reproduced from the two available loudspeakers. As a result, the room sound is partially covered by the direct sound and can therefore not be perceived as well. With Auro 3D 8.0, all loudspeakers are usually used for room sound signals, among other things. Due to the involvement of the four rear loudspeakers in particular, the room sound can be perceived much more differentiated than with stereo, because little or no direct sound is reproduced from the rear.

Although the room sound, which is reproduced from the front loudspeakers in Auro 3D recordings, is less perceptible in combination with direct sound, it still has an important function: it plays a decisive role in the auditory connection between the front and rear loudspeaker layers. If the room sound reproduced by the front speakers is reduced by 10dB or more per speaker compared to the rear speakers, the auditory connection between front and rear falls apart. Therefore, the level balance of the reproduced room sound signals plays an important role when mixing Auro 3D recordings.

1. Introduction

Spatiality is created in 2CH stereo by capturing room signals with a correlation around zero. The spatiality appears evenly distributed on the connecting line between the two loudspeakers, see **Fig. 1a**.

If the playback system is expanded to Auro 3D 8.0 and spatial signals with a correlation around zero are reproduced from all loudspeakers, the spatiality appears as a volume between all loudspeakers, see **Fig. 1b** [1]. A listener can move freely in this volume without the perceived spatiality changing noticeably; the sweet spot is eliminated on this spatial sound layer [2]. The perceived space essentially corresponds to that of the captured recording space. This can go so far that a listener who is in the next room with the door open has the auditory impression that the playback room is the recording room.

ISBN 978-3-9812830-3-7

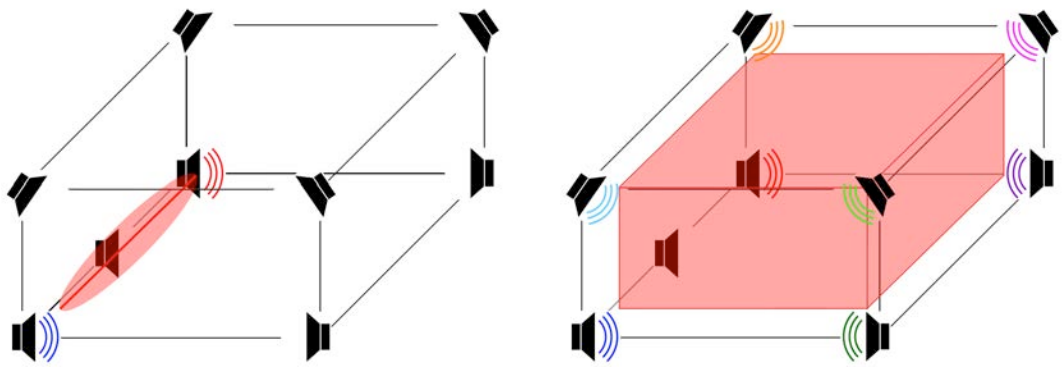


Fig. 1a* and *1b: Perception of stereo and Auro 3D 8.0 / 9.0 spatial representation

2. Different auditory perceptions of microphoned spatial signals with the Auro 3D 9.0 playback system

With an Auro 3D 9.0 speaker configuration, not only captured room signals are reproduced, but a combination of direct sound and room sound. As mentioned at the beginning, when playing classical music, there is a preference to reproduce direct sound from the front. Because of this fact, rules can be established for the properties of captured room signals in order to get the most impressive or natural-sounding result possible.

The loudspeaker signals for the Auro 3D 9.0 setup are defined in **Fig. 2**.

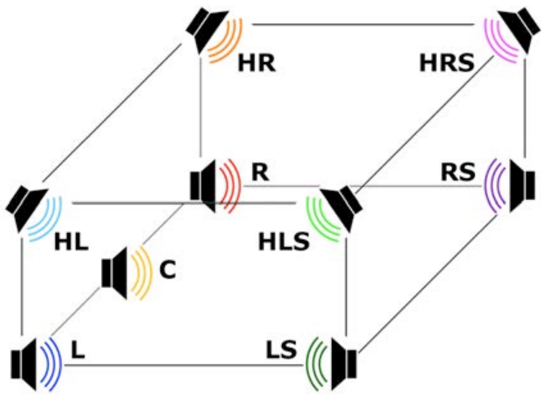


Fig. 2: Allocation of signals to the loudspeakers in Auro 3D 9.0

2.1. Increased sensitivity to direct sound

The author's listening tests have shown that unintentional direct sound coming from behind and above (HLS/HRS, see **Fig. 2**) is perceived as particularly annoying. This can be percussive parts of instruments, echo effects due to uncorrected propagation times of room signals affected by direct sound, or general background noise such as crackling. It is therefore advisable to reproduce signals from that direction that contain as less as possible direct sound. Conversely, it hardly matters to use captured room signals for the front channels (L/C/R), which include direct sound components; these are usually largely covered by the direct sound from the front.

ISBN 978-3-9812830-3-7

2.2. Influence of early reflections in the front upper speakers (HL/HR)

Tonmeisters who have dealt with Auro 3D for the first time and have gained experience with it have described a significant improvement in the timbre of the instruments as one of the main advantages over Surround Sound and 2-channel stereo [2] [3]. The recordings were made using large-AB microphone systems, which roughly correspond to the proportions of the loudspeaker set-up. Thus, the HL/HR channels contain a mixture of room and direct sound from a distance of several meters from the sound body, see **Fig. 3**. According to the author's previous experience, however, this phenomenon does not occur if the loudspeaker signals HL/HR only contain room sound with a low proportion of direct sound or early reflections. Phantom sources between HL/L and HR/R with coherent signals lead not to the phenomenon described either, since identical signals only lead to sound discoloration [4] [5]. It is therefore reasonable to assume that this is due to the influence of early reflections or direct sound, which is interpreted as early reflections.

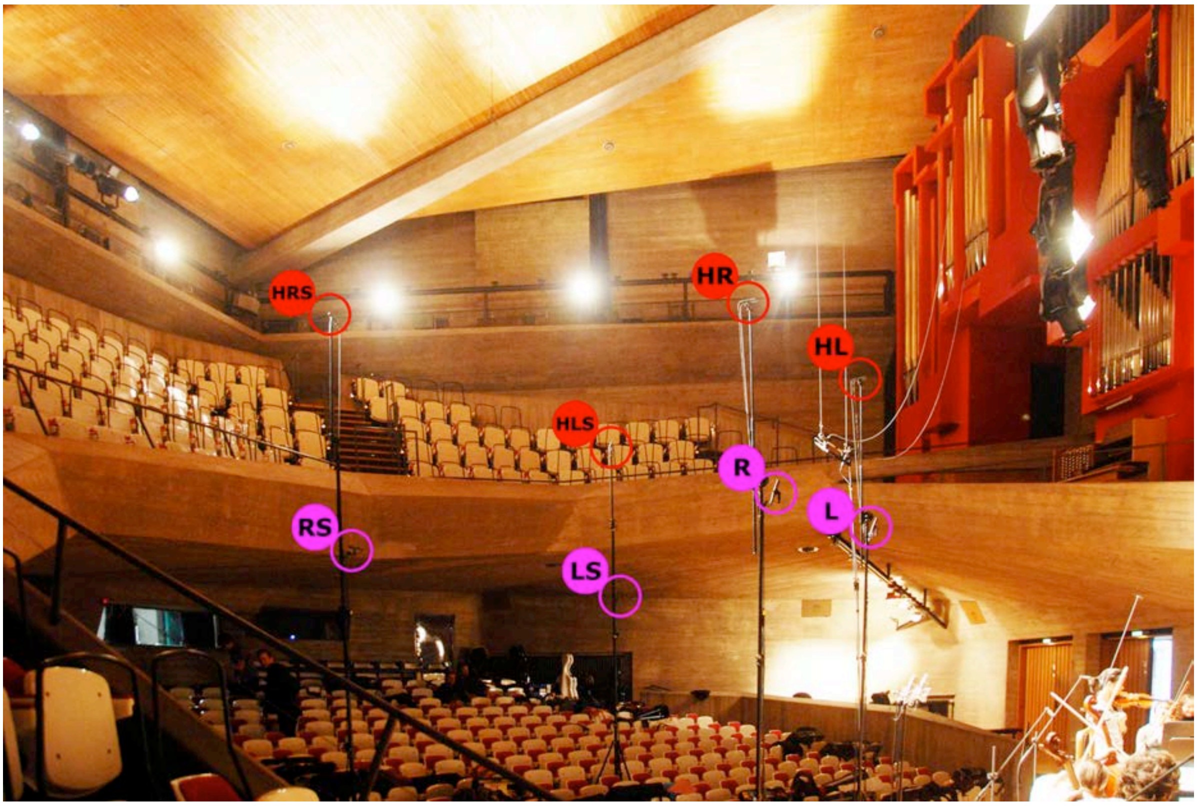


Fig. 3: Auro 3D large AB main microphone system at the HMT Hannover (May 12, 2010) [2]

3. Capturing of room sound

The main goal when capturing room sound is to obtain the sound of the recording room. This is not to be confused with the image: The image of sound sources means that phantom sound sources arise between two loudspeakers; accordingly, the correlation of those direct sound components must be close to 1. Room sound, on the other hand, leads to spatiality and envelopment. The correlation of the room sound signals has to be around zero [6]. Reproduced room sound components with a correlation equal to 1 sound very unnatural, especially if all loudspeakers of an Auro 3D 8.0 / 9.0 setup reproduce identical signals [1].

ISBN 978-3-9812830-3-7

If the above-mentioned condition of a correlation around 0 is met by means of suitable capturing, all room signals (possibly after adjustment with an equalizer) sound very similar, even if the microphone systems are constructed very differently and capture the sound of the recording room with very different methods and directional characteristics. This similarity of sound is important for the human hearing to perceive the signals with low correlation from different directions as a auditory line (2-channel stereo), horizontal plane (5.0 Surround Sound) or volume (Auro 3D 8.0 / 9.0), see **Fig. 1a** and **1b** [1]. This phenomenon occurs with atmospheric sounds such as distant city noise or applause as well as with synthesizer pad sounds or broadband noise.

The following sections describe characteristics of room signals that meet the requirements for the four loudspeaker pairs L/R, LS/RS, HL/HR and HLS/HRS of the Auro 3D setup. In addition, the microphone systems that can be used to obtain those signals are explained. The author used all systems for recordings and their signals were meticulously analyzed.

3.1. Room sound with maximum envelopment (applicable for all channels)

In order to achieve the best possible envelopment in a sound recording, room signals are required that have a correlation around zero down to low frequencies (<200 Hz) [7]; this requires strong channel separation across the entire spectrum to ensure the independence of the individual microphone signals [5]. One method of meeting this requirement is to use a large AB system, where the microphones are placed as far apart as possible, such as on the side walls of a large room, see **Fig. 4**.

Fig. 4: Microphones in the side aisles with large AB systems in the Hofkirche Luzern, CH

If several stereo microphone systems are used in order to obtain sufficient spatial signals for playback, the individual large spacing AB systems must also be at a large distance from one another. In **Fig. 4**, the microphones on the side walls are about 7m apart. If music is played with percussive instruments such as a snare drum, the propagation times need to be aligned to avoid annoying echoes, see **Fig. 5a** and **5b**.

It is particularly important to ensure that as little as possible direct sound is captured for the HLS/HRS loudspeaker signals. Therefore, the microphone system that is furthest away from the sound body can also be suitable.

ISBN 978-3-9812830-3-7

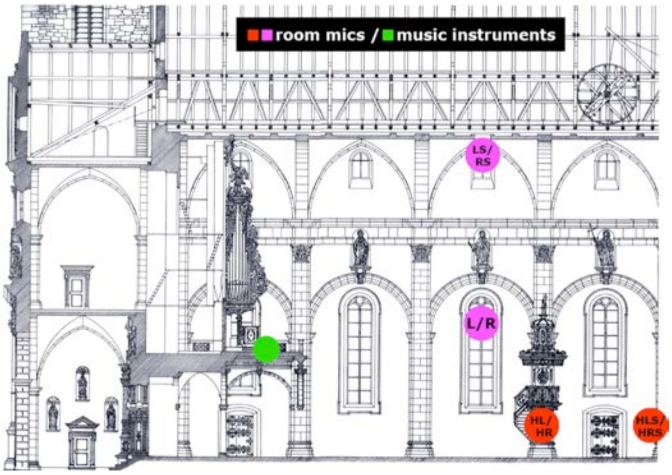


Fig. 5a: Position of the room microphones in the Hofkirche Luzern, CH (July 3, 2009);
Fig 5b: Time correction during mixing

3.2. Early lateral reflections with low transit time (applicable to all channels)

Another method for capturing room signals with low correlation is a system with two supercardioids, which are directed towards the side walls at a small distance from each other and are on the main axis to the sound source. Alternatively, a Sennheiser MKH 800 TWIN double-capsule microphone can also be used and two supercardioids, aligned towards the front and rear, can be constructed during the mix, see **Fig. 6a** and **6b**. Due to the super-cardioid directional characteristic, frequencies <50Hz drop sharply in level; this reduces correlated components of low frequencies. The diffuse field correlation (DFC) is approximately zero at an axis angle of 180°, see **Fig. 7** [8].

Fig. 6a and 6b: Experimental arrangement of Schoeps CCM-41V and
Sennheiser MKH 800 TWIN supercardioids in the Jesus Christ Church in Berlin-Dahlem

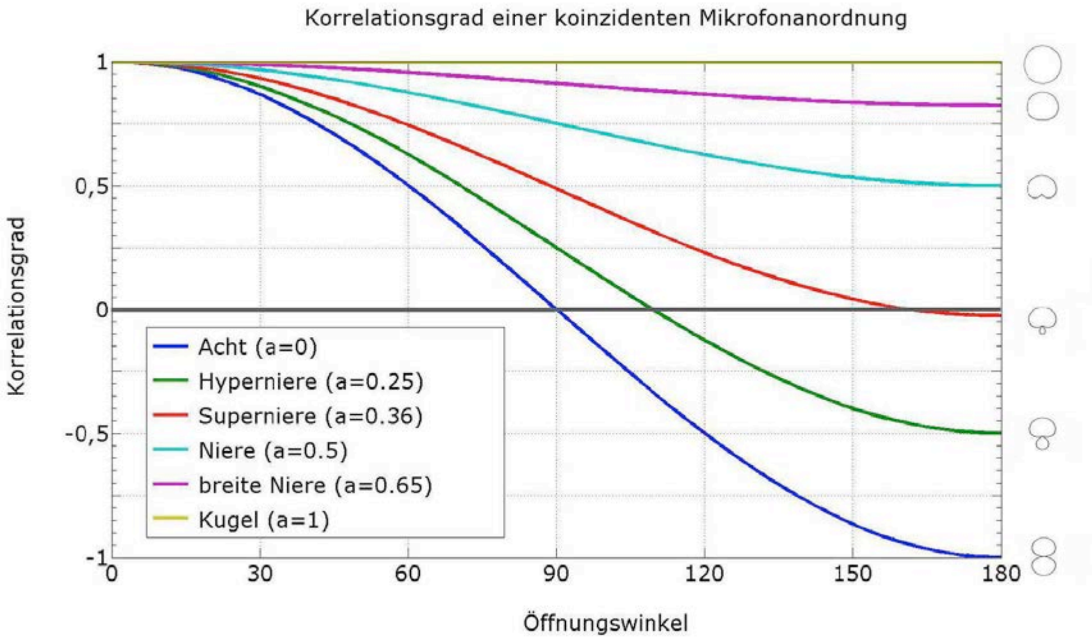


Fig. 7: Relationship between degree of correlation and directivity and angle between two coincident microphones [8]

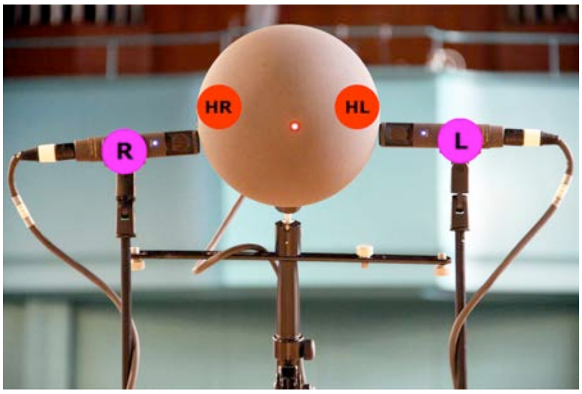
3.3. Room sound with a low proportion of direct sound (especially for HLS/HRS)

As mentioned several times above, the loudspeaker signals HLS/HRS should have as few direct sound components as possible. One way of specifically avoiding direct sound is to use the null angle of microphones with pressure gradient characteristics [9].

The following test arrangement was set up for this purpose: Schoeps KFM-6 sphere microphone as a separator and 2x Sennheiser MKH 800 TWIN for obtaining the direct and room signals, see **Fig. 8a** and **8b**. The system works as follows:

- The direct sound is captured by supercardioid polar patterns, which are directed towards the sound body - these are constructed during the mix [10]. Due to the high directivity of the supercardioids (focus ratio = 5.72 dB), early lateral reflections and reverberation from behind are strongly attenuated. The arrangement on the acoustical separator results in a strong channel separation, especially at high frequencies (spectral differences L ↔ R) [11]. Due to its properties, the subsystem described is particularly suitable for sound bodies that are captured without spot microphones. As expected, those signals of the supercardioid microphones sound much less reverberant than the signals of the KMF-6.
- The room sound is captured by a cardioid directional characteristic. The characteristics constructed during mixing are directed away from the sound body. The extinction angle of this stereophonic arrangement points exactly to the body of the sound; accordingly, there is an enormously high attenuation of the direct sound. As with direct sound, there is also strong channel separation for room signals at high frequencies. The author's listening tests have shown that the present system has significantly less direct sound than the system with two supercardioids directed to the side walls. The propagation times between the L and R channels, on the other hand, are similarly small.

ISBN 978-3-9812830-3-7



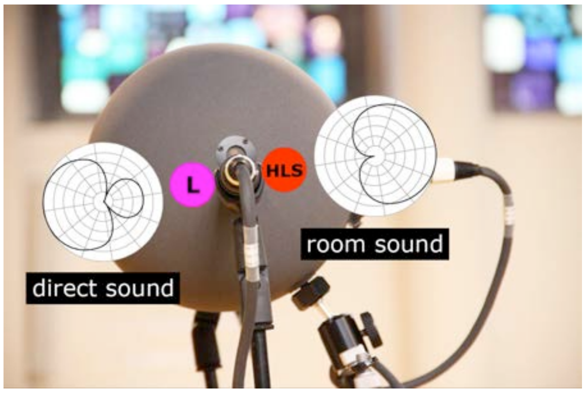


Fig. 8a and 8b: Experimental arrangement of the main microphone system for direct and room sound w. Schoeps KFM-6 and 2x Sennheiser MKH 800 TWIN in Jesus Christ Church

The distance between the violinist Vladyslava Luchenko and the microphone system is considerable at 4.5m. This was chosen at the musician's request for an optimal violin sound, see **Fig. 9**. The unique acoustics of the Jesus Christ Church allow such large distances for recordings without any muddiness being audible. In addition, this results in the following geometric fact: If a sound body moves or turns to the left and right in front of a stereo microphone system, this can have a fatal effect on the image. This is much more pronounced when microphones are placed close to the musicians than when the musicians and the microphone are farther apart, and it leads to unnatural phenomena with string instruments that have a complex directional characteristic [12].

Fig. 9: Arrangement of violinist Vladyslava Luchenko and main microphone system in the Jesus Christ Church Berlin-Dahlem (September 2, 2012)

ISBN 978-3-9812830-3-7

3.4. Room sound with a large proportion of direct sound (HL/HR)

As described in chapter 2.2, early reflections in the loudspeaker signals HL/HR can significantly improve the timbre of the captured instruments. The author's experiments with organ recordings in the Hofkiche Luzern have shown that a similar effect occurs when low-level direct sound from spot microphones is mixed into the HL/HR channels [13]. However, that direct sound does not exist in the L/C/R channels; they are registers positioned high up in the organ, i.e. Récit and Oberwerk, see **Fig. 10**. Some registers of the main organ are located in chambers that are acoustically separated from each other; in general, the registers are spatially distributed over several floors. On the one hand, this means that the leakage between the spot microphones is very small, and on the other hand, all the organ pipes in the chambers can be heard about equally well on the spot microphones, because the sound is evenly distributed in those small rooms with thin wooden walls due to reflections and so with the organ playing ensures high sound pressure. Very good results can be achieved with dynamic microphones such as the model Shure SM7B or Electrovoice RE-20 in the present configuration.

Fig. 10: Récit and Oberwerk microphones in the main organ of the Hofkirche Luzern, CH (ping pong stereophony for speaker signals HL/HR)

The following conclusion can be drawn on the basis of the present results: For the phenomenon described in Chapter 2.2, sound is required that differs sufficiently from the direct sound of the L/C/R channels so that uncontrolled elevation effects do not occur. Spot microphone signals in the HL/HR channels are certainly possible, but result in those signals being localized from above; in many cases this is not desirable for reasons of sound aesthetics.

The directional bands shows, that almost only high frequencies (around 8 kHz) are responsible for perception from above [6], see **Fig. 11**. From numerous experiments with sound recordings in different rooms, the author has found that boosted high frequencies at early reflections leads to increased brilliance of the sound; with direct sound, on the other hand, this often leads to a harsh, sharp sound [14]. Therefore, if early reflections are reproduces from above, according to the author's hypothesis, this increases the brilliance of the sound. This also leads to the conclusion that the upper channels HL/HR should tend to be provided with signals that actually have larger portions in the upper range of the spectrum.

ISBN 978-3-9812830-3-7

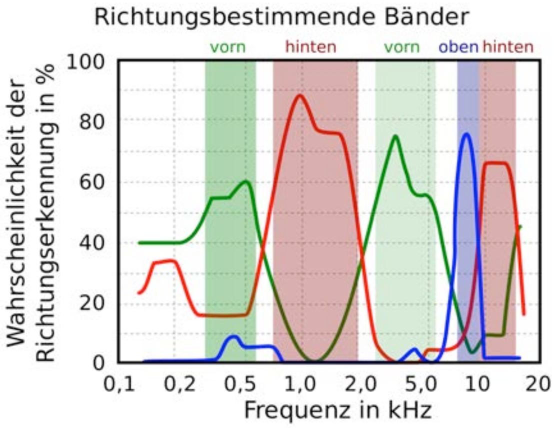


Fig. 11: Directional bands according to Blauert [15]

When recording classical music, the high frequency components contribute to the fine structure of the sound. In the case of orchestral music, instruments prevail in the room that have a large proportion of high frequencies, such as radiating trumpets or percussion instruments (including cymbals). With these instruments, the reverberation distance in the direction of their axis is very large compared to others [12].

As mentioned several times in Chapter 3, there is a requirement for the loudspeaker signals HL/HR that the correlation of the room signals must be around zero. Therefore, it makes sense to use a microphone system that works with large spacing AB. It makes sense to position a microphone system very close to the edge of the stage to capture high frequencies of the direct sound - a main microphone would also be placed at this distance, but on the 0° axis and not on the side walls.

In his report on Auro 3D, Gregor Zielisky mentions that there is a risk of hearing the instruments from above, especially the woodwind instruments, with the height microphones due to the radiation behavior of certain instruments [2]. This can be avoided by placing the HL/HR microphones at the same height as the L/R and LS/RS microphones. An example of such a recording is shown in **Figs. 12 and 13**. It is a live recording from September 27, 2012 with the Tonhalle Orchestra Zurich and chief conductor David Zinman; the Auro 3D recording was created in cooperation with the Swiss radio DRS 2 and Silencium Musikproduktion GmbH.

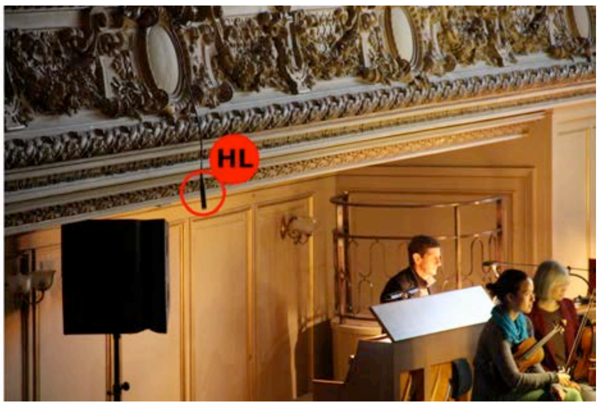


Fig. 12: Microphones on the balcony for HL/HR signals in the Zurich Tonhalle (September 27, 2012)

ISBN 978-3-9812830-3-7

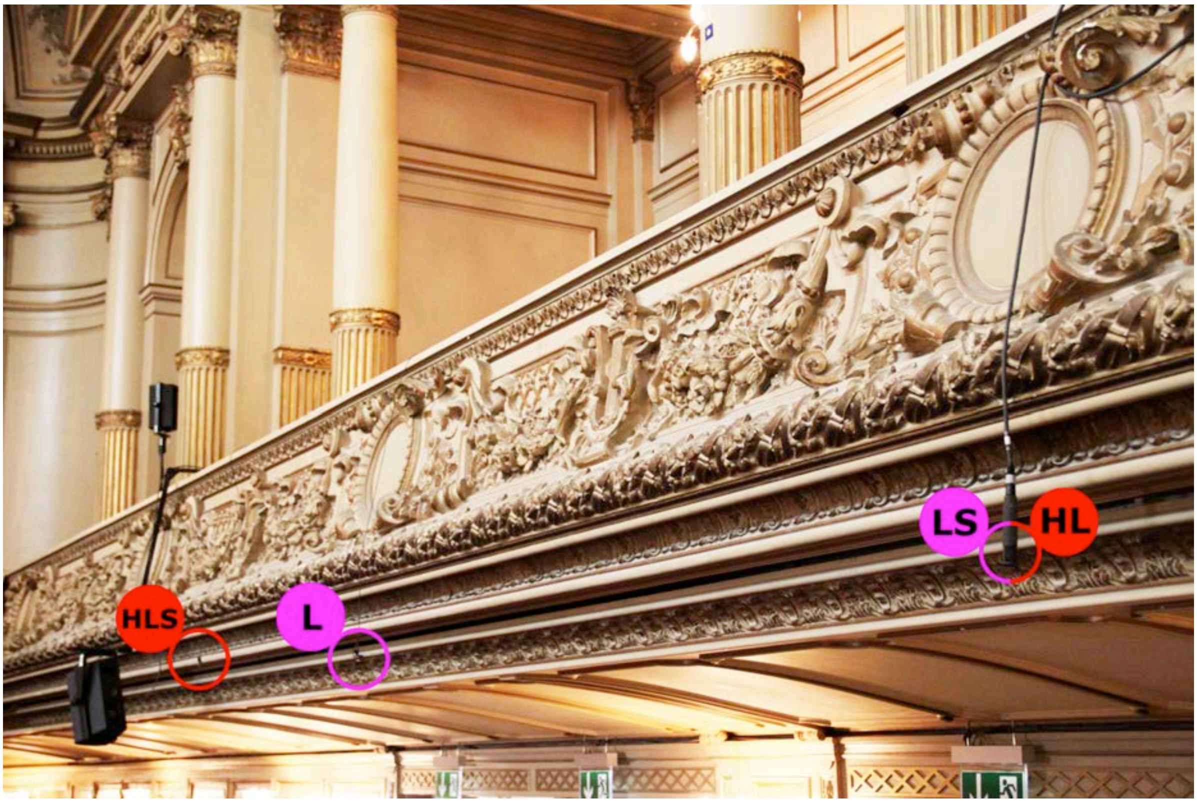


Fig. 13: Position of the room microphones on the balcony in the Zurich Tonhalle (September 27, 2012). The Sennheiser MKH 800 TWIN capsules are used here as kidneys.

As explained at the beginning of Chapter 3, room signals from the same room generally sound very similar. This can be checked as follows: To do this, the room signal pairs assigned to the L/R, LS/RS and HLS/HRS loudspeakers are swapped around - the HL/HR signal pair cannot be relocated because those signals contain a large amount of direct sound. If the recording is edited so that the channel assignment changes while the recording is in progress, no changes should be discernible, see **Fig. 14**. This shows that the recording of room signals from different height levels is possible, but not mandatory.

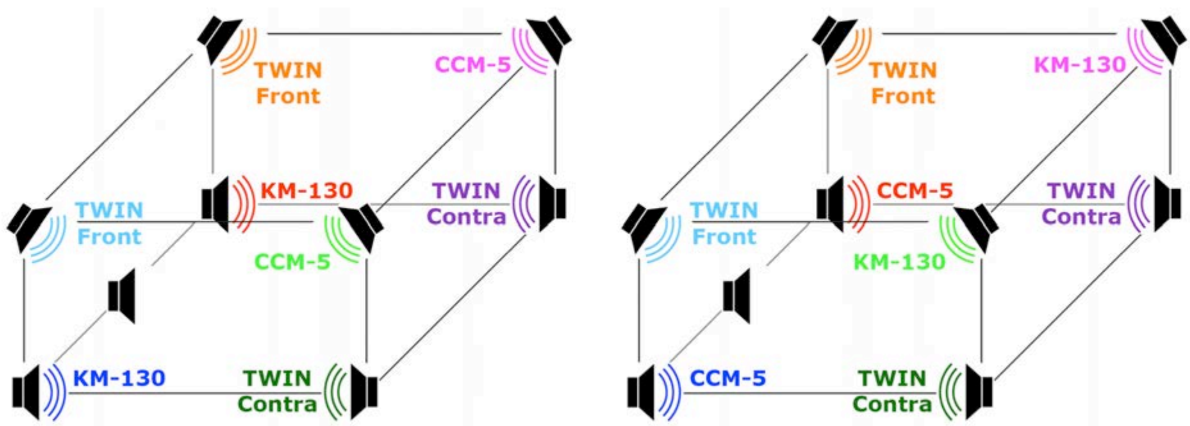


Fig. 14: Switching the signal allocation of the room signals to loudspeakers: should usually not be noticeable.

ISBN 978-3-9812830-3-7

3.5. Direct sound (L/C/R)

Even if direct sound does not belong to room sound, a note should be given at this point. With Auro 3D recordings, the L/C/R channels should have as few early reflections as possible if the proportion of room signals is relatively high. Thus, the room sound of the recording is easily perceptible. If the proportion of early reflections is too high, this leads to an intransparent sound of the instruments.

The strength of the recording concept described lies in the contrast between the clarity of the direct sound and the envelopment by room sound as well as the natural-soundig timbre of the instruments due to the early reflections from the HL/HR channels, see **Fig. 15**.

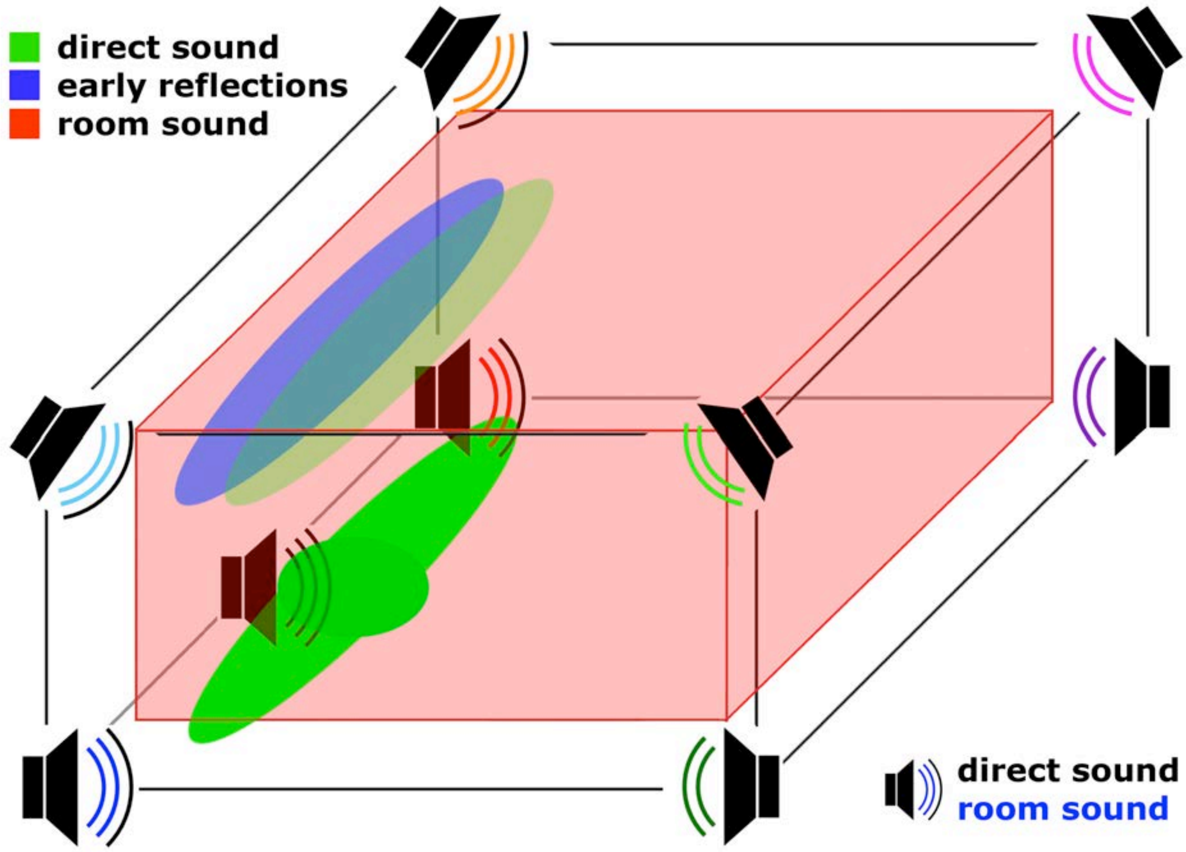


Fig. 15: Sound concept for loudspeaker signals of an Auro 3D Classic recording

4. Energy balance in room signals

It is a challenge to create an auditory connection between the front and rear speaker planes for both Surround Sound 5.0 and Auro 3D 9.0. As explained in the introduction, room signals with a correlation around zero result in a balanced room representation. A balanced energy ratio of the room sound holds the two speaker planes together: It is easy to see that the auditory connection decreases when the balance of the room sound between the front and the back is changed significantly, see **Fig. 16**. In the extreme case – no more room sound is reproduced from the front – the perceived room is reduced to a plane between the rear four speakers LS/RS and HLS/HRS.

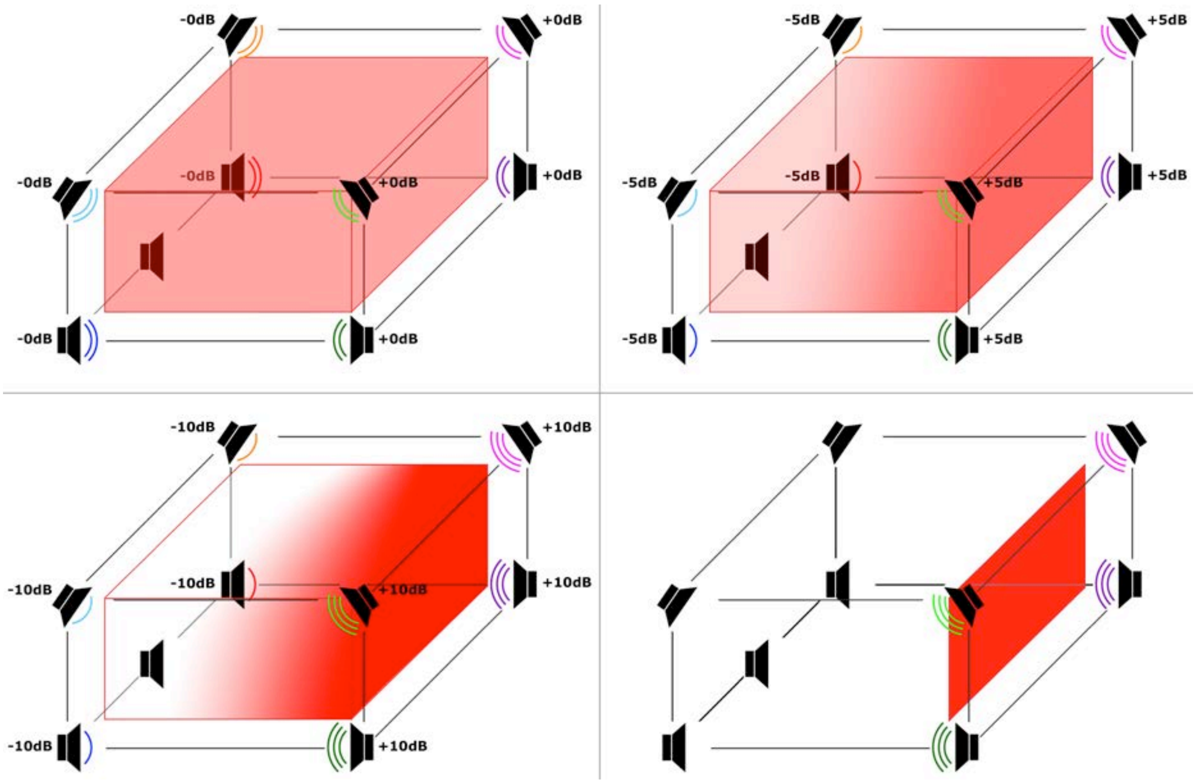


Fig. 16: Perceived spatial impression with increasing imbalance of room signals with the Auro 3D 8.0 / 9.0 loudspeaker setup

In the case of a sound recording, direct sound can usually be heard in addition to the room sound, see sound concept chapter 3.5. During playback, direct sound and surround sound are superimposed. Direct sound cannot influence the perception of the room sound because direct sound contains no room information. Accordingly, direct sound reproduced from the front cannot lead to an auditory connection with the room sound reproduced exclusively from behind, see **Fig. 17**.

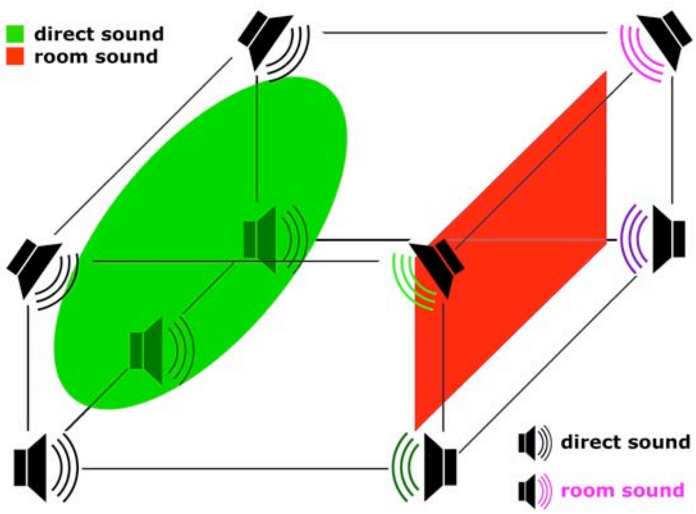


Fig. 17: Perceived sound impression with complete separation between direct sound at the front and room sound at the rear (Auro 3D 9.0 loudspeaker setup)

ISBN 978-3-9812830-3-7

It is a challenge to evenly balance the room signal level in the front channels L/R and HL/HR, since the room sound is partially covered by the direct sound, see **Fig. 18**. This can only be done precisely if direct and room sound can be recorded almost completely separately from each other or if the room signal levels are already balanced on the recording tracks by choosing the same microphone types and identical settings on the microphone amplifiers.

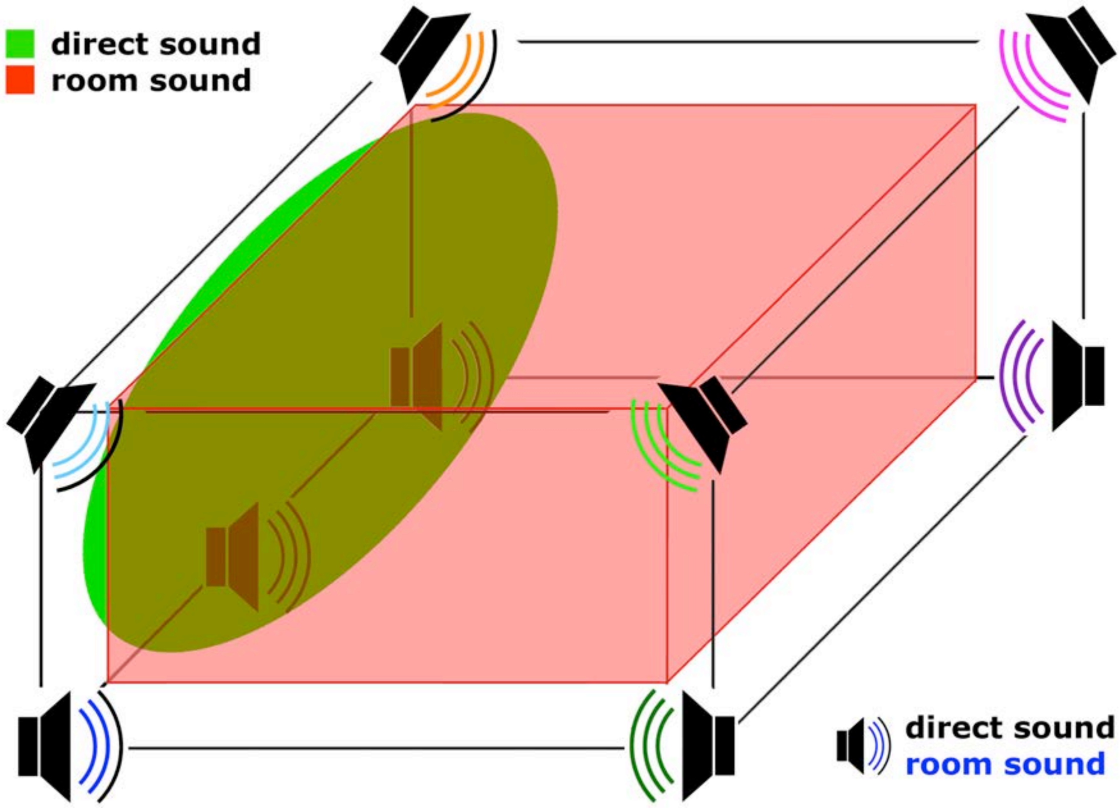


Fig. 18: Perceived sound impression with correct balancing of room signals together with direct sound (Auro 3D 9.0 loudspeaker setup)

5. Conclusions

This article shows that the psychoacoustic relationships between recording and playback of 3D Audio / Auro 3D content are very complex. On the one hand, it is not trivial to make a recording satisfactorily, because general conditions such as the quality of the room acoustics, outside noise, the possibility of appropriate miking (including the time factor) or simply inadequate knowledge and insufficient experience during recording or mixing have an even greater impact than with 2-channel stereo or Surround Sound 5.0. On the other hand, with Auro 3D, sound experiences and thus musical delights can be made possible that literally outshine 2-channel stereo recordings – after all, there are two more dimensions available for tonal creativity!

ISBN 978-3-9812830-3-7

6. Acknowledgments

The author would like to thank all the contributors who made it possible to realize extensive and further recordings for this article. Special thanks go to Christof Faller, who was always available for suggestions, discussions and technical implementation, and to Helmut V. Fuchs, who gave the author significant and kind support in realizing the recording in the Jesus Christ Church in Berlin-Dahlem.

7. List of References

[1] Nipkow, L. (2010): *‚Angewandte Psychoakustik bei 3D Surround Sound Aufnahmen'*, 26. Tonmeistertagung, Leipzig, Proceedings.

[2] Zielinsky, G. (2011): *‚More reality mit Auro 3D'*, VDT-Magazin, Heft 2, 24-27

[3] Albinska-Frank, M. (2011): *‚Auro 3D ist ein Erlebnis'*, VDT-Magazin, 2011, Heft 2, 27-30

[4] tho Pesch, P. (*2010)*: *‚Die Lokalisation von Phantomschallquellen im oberen Halbraum: Untersuchungen zur Erweiterung der Binauralen Raumsynthese'*, Verlag Dr. Müller.

[5] Theile, G.; Wittek, H. (2011): *‚Die dritte Dimension für Lautsprecher-Stereofonie'*, VDT-Magazin, Heft 2, 31-37.

[6] Blauert, J. (1974): *‚Räumliches Hören'*, S. Hirzel Verlag, Stuttgart.

[7] Griesinger, D. (1998): *‚General overview of spatial impression, envelopment, localization, and externalization'*, 15th International AES Conference, Copenhagen, Proceedings 136-149.

[8] Wittek, H. (2006): *‚M/S Techniques for Stereo and Surround'*, 24. Tonmeistertagung, Proceedings.

[9] Jecklin, J. (2002): *‚Sound, Image und Space. Mehrkanalige Aufnahmetechnik'*, avguide.ch, magazine.

[10] Sennheiser: product information to MKH 800 TWIN, *http://de-de.sennheiser.com/mkh-800-twin-ni*

[11] Schoeps: product information to KFM-6, *http://www.schoeps.de/de/products/kfm6*

[12] Meyer, J. (1999): *‚Akustik und musikalische Aufführungspraxis'*, Verlag Erwin Bochinsky, 4th revised edition.

[13] Sieber, W. (2009); Swiss Brass Consort: *‚Rhapsody in Blue'*, CD production, Phonoplay International.

[14] Genuit, K. (2010): *‚Sound-Engineering im Automobilbereich: Methoden zur Messung und Auswertung'*, Springer-Verlag Berlin Heidelberg.

[15] Skyhead, *http://de.wikipedia.org/wiki/Blauertsche_Bänder*

ISBN 978-3-9812830-3-7

tmt27
expertise in audio media

22.11.– 25.11.2012 Köln | Cologne

Tagungsbericht der | Conference Proceedings
27. Tonmeistertagung

Impressum

Released by Verband Deutscher Tonmeister e. V.

Editors: Wolfgang Hoeg, Ernst Rothe, Günther Theile

Layout: Andrea Krahmer

© 2013 Verband Deutscher Tonmeister e. V.

All rights reserved.

ISBN 978-3-9812830-3-7