Chapter 2
Auditory Events of Multi-loudspeaker Playback

It is evident that until one knows what information needs to be presented at the listener’s ears, no rational system design can proceed.

Michael A. Gerzon 1976 and AES Vienna [1] 1992.

Abstract This chapter describes the perceptual properties of auditory events, the sound images that we localize in terms of direction and width, when distributing a signal with different amplitudes to one or a couple of loudspeakers. These amplitude differences are what methods for amplitude panning implement, and they are also what mapping of any coincident-microphone recording implies when reproduced over the directions of a loudspeaker layout. Therefore several listening experiments on localization are described and analyzed that are essential to understand and model the psychoacoustical properties of amplitude panning on multiple loudspeakers of a 3D audio system. For delay-based recordings or diffuse sounds, there is some relation, however, it is found to be less stable for the desired applications. Moreover, amplitude panning is not only about consistent directional localization. Loudness, spectrum, temporal structure, or the perceived width should be panning-invariant. The chapter also shows experiments and models required to understand and provide those panning-invariant aspects, especially for moving sounds. It concludes with openly-available response data of most of the presented listening experiments.

Starting from classic listening experiments on stereo panning by Leakey [2], Wendt [3], and pairwise horizontal panning by Theile [4], this chapter explores the relevant perceptual properties for 3D amplitude panning and their models. Important experimental studies considered here are for instance those by Simon [5], Kimura [6], F. Wendt [7], Lee [8], Helm [9], and Frank [10, 11]. By the experimental results, it
is possible to firmly establish Gerzon’s [1] $E$, $\mathbf{r}_E$ and $\|\mathbf{r}_E\|$ estimators for perceived loudness, direction, and width that apply to most stationary sounds in typical studio and performance environments.

### 2.1 Loudness

At a measurement point in the free field, the same signal fed to equalized loudspeakers of exactly the same acoustic distance would superimpose constructively (+6 dB).

In a room with early reflections and a less strict equality of the incoming pair of sounds (typical, slight inaccuracy in loudspeaker/listener position, different mounting situations, different directions in the directivities of ears and loudspeakers), the superposition can be regarded as stochastically constructive (+3 dB) in particular at frequencies that aren’t very low.

For the above reasoning, typical amplitude panning rules try to keep the weights distributing the signal to the loudspeakers normalized by root of squares instead of normalizing to the linear sum, in order to obtain constant loudness ([12], VBAP):

$$g_l \leftarrow \frac{g_l}{\sqrt{\sum_{l=1}^{L} g_l^2}}.$$  \hspace{1cm} (2.1)

**Loudness Model.** If all loudspeakers are equalized, located at the same distance to the listener, and fed by the same signal with different amplitude gains $g_l$, a constructive interference could be expected so that the amplitude becomes [1]

$$P = \sum_{l=1}^{L} g_l.$$  \hspace{1cm} (2.2)

However, the interference stops to be strictly constructive as soon as the room is not entirely anechoic, the sitting position is not exactly centered, or even for anechoic and centered conditions at high frequencies, when the superposition at the ears cannot be assumed to be purely constructive anymore. Then it is better to assume a less well-defined, stochastic superposition in which a squared amplitude is determined by the sum of the squared weights [1]:

$$E = \sum_{l=1}^{L} g_l^2.$$  \hspace{1cm} (2.3)

Therefore, the most common amplitude panning rules use root-squares normalization to obtain a loudness impression that is as constant as possible.

The measure $E$ seems to be most useful when designing and evaluating amplitude-panning or coincident microphone techniques. It is not surprising that the ITU-R
2.1 Loudness

BS.1770-4\textsuperscript{1} uses the $Leq(RLB)$ measure as a loudness model: it is essentially the RMS level after high-pass filtering, cf. [13], which is closely related to the $E$ measure detected from loudspeaker signals.

An interesting refinement was proposed by Laitinen et al. [14], which uses a measure $\sqrt{\sum_{i=1}^{L} g_i^p}$ in which the exponent $p$ is close to 1 at low frequencies under anechoic conditions and close to 2 at high frequencies/under reverberant conditions.

2.2 Direction

In the early years of stereophony, researchers investigated the differences in delay times and amplitudes required to control the perceived direction. Below, only experiments are considered that did not use fixation of the listener’s head.

2.2.1 Time Differences on Frontal, Horizontal Loudspeaker Pair

The dissertation of K. Wendt in 1963 [3] shows notably accurate listening experiments done on ±30° two-channel stereophony using time delays, in which listeners indicated from where they heard the sounds for each of the tested time differences. H. Lee revisited the properties in 2013 [8], but with musical sound material and an experiment, in which the listener adjusted the time differences until the perceived direction matched the one of a corresponding fixed reference loudspeaker, Fig. 2.1.

The time differences are seldom applicable to reliable angular auditory event placement: auditory images are strongly frequency-dependent (not shown here) and therefore unstable for narrow-band sounds. Leakey and Cherry showed 1957 [2] that time-delay stereophony loses its effect under the presence of background noise.

2.2.2 Level Differences on Frontal, Horizontal Loudspeaker Pair

K. Wendt’s [3] and H. Lee’s [8] experiments deliver insights in sound source positioning with ±30° two-channel stereophony, however this time with level differences.

As opposed to Fig. 2.1, in which auditory image panning with time differences were characterized by statistical spreads of up to 15°, level-difference-based panning is clearly smaller in the spread of perceived directions than 10°, Fig. 2.2.

Signal dependency. Wendt [3] described the signal dependency of panning curves on various transient and band-limited sounds, and Lee [8] for musical sounds. A new

\textsuperscript{1}https://www.itu.int/rec/R-REC-BS.1770-4-201510-I/en Algorithms to measure audio programme loudness and true-peak audio level (10/2015).
Fig. 2.1 K. Wendt’s experiment [3] used an angular marks helping to specify the localized direction (left). Right shows results for time differences between impulse signals fed to loudspeakers, no head fixation (diagram shows means and standard deviation; the standard deviation was interpolated for the figure). In gray: Results of the time-difference adjustment experiment of Lee [8] using musical material (25, 50, 75% quartiles, symmetrized diagram).

Fig. 2.2 Wendt’s [3] results to crack (impulsive) signals with level differences and without head fixation (the figure shows means and standard deviation; standard deviation was interpolated to plot this figure). In gray: Results of Lee’s [8] level-difference adjustment experiment with musical sounds (25, 50, 75% quartiles, symmetrized diagram).

A comprehensive investigation on frequency dependency was carried out by Helm and Kurz [9]. With level differences \{0, 3, 6, 9, 12\} dB and third-octave filtered pulsed pink noise at \{125, 250, 500, 1k, 2k, 4k\} Hz, they showed that the perceived angle pointed at by the listeners using a motion-tracked pointer was similar between the broad-band case and third-octave bands below 2 kHz. In bands above 2 kHz, smaller level differences cause a larger lateralization, see interpolated curves in Fig. 2.3.
2.2.3 Level Differences on Horizontally Surrounding Pairs

Successive pairwise panning on neighboring loudspeaker pairs is typically used to pan auditory events freely along the loudspeakers of a horizontally surrounding loudspeaker ring. The classical research done specifically targeted at such applications was contributed by Theile and Plenge 1977 [4]. They used a mobile reference loudspeaker with some reference sound that could be moved to match the perceived direction of a loudspeaker pair playing pink noise with level differences at different orientations with respect to the listener’s head. There is also the experiment of Pulkki [15] using a level-adjustment task, in which levels were adjusted as to match the auditory event to one of a reference loudspeaker at three different reference directions and for different head orientations. A comprehensive experiment was done by Simon et al. [5], who used a graphical user interface displaying the floor plan of a 45°-spaced loudspeaker ring to have the listeners specify the perceived direction. Martin et al. in 1999 [16] used a graphical user interface showing the floor plan of a 5.1 ring in their experiment, and last but not least, Matthias Frank used a direct pointing method to enter the perceived direction [10] in one of his experiments.

As the experiments did not seem to yield consistent results, a comprehensive level-difference adjustment experiment with 24 loudspeakers arranged as a horizontal ring was done in [17] and partially repeated later in [11], see results in Fig. 2.4. In the repeated experiment [11] it became clear that in the anechoic room, a large amount of the differently pronounced localization biases can be avoided by encouraging the listeners to do front-back and left-right head motion by a few of centimeters, whenever there is doubt.
2.2.4 Level Differences on Frontal, Horizontal to Vertical Pairs

Quite extensively, T. Kimura investigates the localization of auditory events between frontal, vertical ±13.5° loudspeaker pairs in 2012 [6, 18]. The work of F. Wendt in 2013 [7, 19] also investigates a slant and vertical loudspeaker pair, Fig. 2.5. Kimura uses pulsed white noise, Wendt uses pulsed pink noise.

Obviously, the horizontal spread is always smaller than the vertical spread and the spread does not align with the direction of the loudspeaker pair. The largest vertical spread appears for the vertical loudspeaker pair.

2.2.5 Vector Models for Horizontal Loudspeaker Pairs

A weighted sum of the loudspeakers’ direction vectors \( \theta_1, \theta_2 \) could be conceived as simple linear model of the perceived direction, using a linear blending parameter \( 0 \leq q \leq 1 \)

\[
    r = (1 - q) \theta_1 + q \theta_2. \quad (2.4)
\]

The parameter \( q \) adjusts where the resulting vector \( r \) is located on the connecting line between \( \theta_1 \) and \( \theta_2 \). On frontal loudspeaker pairs, localization curves typically run through the middle direction \( q = \frac{1}{2} \) for level differences of 0 dB. If only one
loospeakers is active, the result is either of the loudspeaker directions, thus the parameter is \( q = 0 \) or \( q = 1 \).

**Classical definitions.** As the simplest choice for \( q \), one could insert \( q = \frac{g_2}{g_1 + g_2} \) or \( q = \frac{g_1^2}{g_1^2 + g_2^2} \) to get the vector definitions as weighted average using either the linear or squared gains according to [1]:

\[
\mathbf{r}_V = g_1 \theta_1 + g_2 \theta_2, \quad \mathbf{r}_E = \frac{g_1^2 \theta_1 + g_2^2 \theta_2}{g_1^2 + g_2^2}. \tag{2.5}
\]

For both models, equal gains \( g_1 = g_2 \) yield \( q = \frac{1}{2} \), and also the endpoints with \( g_2 = 0 \) or \( g_1 = 0 \) correspond to \( q = 0 \) or \( q = 1 \), respectively. However, the slope of the \( \mathbf{r}_E \) vector is steeper than the one of the \( \mathbf{r}_V \). For instance, if \( g_2 = 2 g_1 \), the vector \( \mathbf{r}_V \) lies on \( q = 2/3 \) of the line between \( \theta_1 \) and \( \theta_2 \), while \( \mathbf{r}_E \) lies at \( q = 4/5 \) of the connecting line.

The \( \mathbf{r}_V \) vector for the \( \pm \alpha \) loudspeaker pair at the directions \( \theta_{1,2} = (\cos \alpha, \pm \sin \alpha) \) corresponds to the tangent law [20], whose formal origin lies in a model of summing localization based on a simple model of the ear signals, cf. Appendix A.7. The equivalence of this law to the vector model follows from the tangent \( \tan \varphi \) as ratio of the \( y \) divided by \( x \) component of the \( \mathbf{r}_V \) vector, \( \tan \varphi = \frac{g_1 \sin(\alpha) + g_2 \sin(-\alpha)}{g_1 \cos(\alpha) + g_2 \cos(\alpha)} = \frac{g_1 - g_2}{g_1 + g_2} \tan \alpha \).
Adjusted slope. Differently steep curves were fitted by an adjustable-slope model [17]

\[ r_\gamma = \frac{|g_1|^\gamma \theta_1 + |g_2|^\gamma \theta_2}{|g_1|^\gamma + |g_2|^\gamma}, \quad (2.6) \]

which uses \( \gamma = 1 \) for \( r_V \) and \( \gamma = 2 \) for \( r_E \). Figure 2.6 compares the prediction by \( r_V \), \( r_E \), and \( r_\gamma \) to frequency-dependently perceived directions in frontal horizontal pairs, to perceived directions in a lateral stereo pair, and to perceived directions in a frontal pair that is either horizontal or vertical, using various studies mentioned above.

Practical choice \( r_E \). While a specific exponent \( \gamma \) closely fitting the experimental data may vary, a constant value is preferable. Figure 2.6 indicates that in most cases focusing on \( r_E \) is reasonable and sufficiently precise, see also [11].
2.2.6 Level Differences on Frontal Loudspeaker Triangles

V. Pulkki [21] and F. Wendt [7, 19] investigated localization properties for frontal loudspeaker triplets with level differences, see Fig. 2.7. Both used pulsed pink noise in their experiments. While V. Pulkki used an indirect adjustment task to evaluate VBAP control angles to obtain auditory events directionally matching the respective reference loudspeakers, F. Wendt uses a direct pointing method. Wendt’s experiments indicate that loudspeaker triplets with three different azimuthal positions yield a smaller spread in the indicated direction than such with vertical loudspeaker pairs (not the case in Pulkki’s experiments).

2.2.7 Level Differences on Frontal Loudspeaker Rectangles

F. Wendt [7, 19] moreover presents experiments about frontal loudspeaker rectangles, again using a pointer method and pulsed pink noise, Fig. 2.8. Again it seems that arrangements avoiding vertical loudspeaker pairs exhibit a smaller statistical spread in the responses.
Fig. 2.8 Wendt’s experiments about frontal loudspeaker rectangles showing two-dimensional 95% confidence (solid) and standard deviation ellipses (dotted). The experimental setup of this and above-mentioned experiments is shown. Left: each of the corner loudspeakers is raised once by +6 dB in level, right: both left/right loudspeaker levels are raised once by \{+3, +6\} dB, and both top/bottom pairs are once raised by +6 dB.

### 2.2.8 Vector Model for More than 2 Loudspeakers

For more than two active loudspeakers and in 3D, a vector model based on the exponent $\gamma = 2$ yields the $\mathbf{r}_E$ vector [1]

$$
\mathbf{r}_E = \frac{\sum_{l=1}^{L} g_l^2 \theta_l}{\sum_{l=1}^{L} g_l^2}.
$$

### 2.2.9 Vector Model for Off-Center Listening Positions

At off-center listening positions, the distances to the loudspeakers are not equal anymore, resulting in additional attenuation and delay for each loudspeaker depending on the position. For stationary sounds, this effect can be incorporated into the energy vector by additional weights $w_{r,i}$ and $w_{\tau,i}$.
The weight $w_{r,l}$ models the attenuation of a point-source-like propagation $\frac{1}{r}$. The reference distance is the distance to the closest loudspeaker at the evaluated listening position, thus the weight of each loudspeaker results in

$$w_{r,l} = \frac{1}{r_l}.$$  

(2.9)

The incorporation of delays into the energy vector requires a transformation that yields the weights $w_{\tau,l}$ for each loudspeaker. It is reasonable that these weights attenuate the lagging signals in order to reduce their influence on the predicted direction. An attenuation of $\frac{1}{4} \text{dB} \cdot \frac{\text{ms}}{}$ is known from the echo threshold in [22], similarly [23], and has successfully been applied for the prediction of localization in rooms [24]. The weight of each loudspeaker is calculated as $\tau_l = \frac{c}{r_l}$ in seconds at the listening position under test

$$w_{\tau,l} = 10^{-\frac{1000}{4\cdot20}\tau_l}.$$  

(2.10)

Further weights can be applied in order to model the precedence effect in more detail, as proposed by Stitt [25, 26]. Listening test results in [27] compared the differently complex extensions of the energy vector and revealed that the simple weighting with $w_{r,i}$ and $w_{\tau,i}$ is sufficient for a rough prediction of the perceived direction in typical playback scenarios.

The left side of Fig. 2.9 shows the predicted directions by the energy vector for various listening positions when playing back the same signal on a standard stereo loudspeaker pair with a radius of 2.5 m. The absolute localization error can be calculated from the difference of the predicted direction and the desired panning direction. The right side of Fig. 2.9 depicts areas with localization errors within 4 ranges: $0^\circ$ ... $10^\circ$ (white, perfect localization), $10^\circ$ ... $30^\circ$ (light gray, plausible localization), $30^\circ$ ... $90^\circ$ (gray, rough localization), and $>90^\circ$ (dark gray, poor localization).

Concerning a single playback scenario, i.e. a single panning direction on a loudspeaker setup, the perceptual sweet area for plausible playback can be estimated by the area with localization errors below $30^\circ$. For the prediction of a more general sweet area, the absolute localization errors can be computed for all possible panning directions in a fine grid of $1^\circ$ and averaged at each listening position as shown in Fig. 2.10.
Fig. 2.9 Predictions of perceived directions by the energy vector for different listening positions in a standard stereo setup with two loudspeakers playing the same signal. Gray-scale areas on the right indicate listening areas with predicted absolute localization errors within different angular ranges.

Fig. 2.10 Predictions of mean absolute localization errors by the energy vector in a standard stereo setup for panning directions between $-30^\circ$ and $30^\circ$.

2.3 Width

M. Frank [10] investigated the auditory source width for frontal loudspeaker pairs with 0 dB level difference and various aperture angles, as well as the influence of an additional center loudspeaker on the auditory source width. The response was given by reading numbers off a left-right symmetric scale written on the loudspeaker arrangement (Fig. 2.11).

Figure 2.11 (right) shows the statistical analysis of the responses. Obviously the additional center loudspeaker decreases the auditory source width.

Auditory source width is difficult to compare for different directions and also single loudspeakers yield auditory source widths that vary with direction. Still, a relatively constant auditory source width is desirable for moving auditory events. For static auditory events, the narrowest-possible extent can be desirable.
2.3 Width

The angle $2 \arccos \| r_E \|$ describes the aperture of a cap cut off the unit sphere perpendicular to the $r_E$ vector, at its tip, from the origin, see Fig. 2.12. As the $r_E$ vector length is between 0 (unclear direction) and 1 (only one loudspeaker active), this angle stays between $180^\circ$ and $0^\circ$.

M. Frank’s experiments about the auditory source width [10, 28] showed that stereo pairs of larger half angles $\alpha$ were also heard as wider. The length of the $r_E$ vector gets shorter with the half angle $\alpha$. In a symmetrical loudspeaker pair $\theta_{12}^T = (\cos \alpha, \pm \sin \alpha)$ with $g_1 = g_2 = 1$, the $y$ coordinate of the $r_E$ vector cancels and its length is

$$||r_E|| = r_{E,x} = \cos \alpha.$$  

The corresponding spherical cap is same size as the loudspeaker pair $2 \arccos ||r_E|| = 2\alpha$. However, only $\frac{5}{8}$ of the size was indicated by the listeners of the experiments, which yields the following estimator of the perceived width:

$$ASW = \frac{5}{8} \cdot \frac{180^\circ}{\pi} \cdot 2 \arccos ||r_E||.$$  \hspace{1cm} \text{(2.11)}

Fig. 2.12  Cap size associated with $r_E$ length model for L+R (left plot) and L+R+C (right plot)
For an additional center loudspeaker \( g_3 = 1 \), \( \theta^T = (1, 0) \), the estimator yields

\[ \| r_E \| = r_{E,x} = \frac{1}{3} + \frac{2}{3} \cos \alpha, \]

an increase matching the experiments as \( \arccos \| r_E \| < \alpha \), see Figs. 2.13 and 2.12.

### 2.4 Coloration

Despite research primarily focuses on the spatial fidelity of multi-loudspeaker playback, the overall quality of surround sound playback was found to be largely determined by timbral fidelity (70%) [29]. Loudspeakers in a studio or performance space are often characterized by different colorations that are caused by different reflection patterns (most often the wall behind the loudspeaker). When changing the active loudspeakers, or their number, these differences become audible. On the one hand, static coloration, e.g. the frequency responses of the loudspeakers, can typically be equalized. On the other hand, changes in coloration during the movement of a source cannot be equalized easily and yield annoying comb filters.

Although coloration is often assessed verbally [30], we employ a simple technical predictor based on the composite loudness level (CLL) by Ono [31, 32]. The CLL spectrum predicts the perceived coloration and is calculated from the sum of the loudnesses of both ears in each third-octave band. Studies about loudspeaker and headphone equalization show that differences in third-octave band levels of less than 1dB are inaudible by most listeners [33, 34]. This criterion can also be applied for the perception of coloration, i.e., differences between CLL spectra of less than 1dB are assumed to be inaudible.

Pairwise panning between loudspeakers results in a single active loudspeaker for source directions that coincide with the direction of a loudspeaker and two equally loud loudspeakers for source directions exactly between two neighboring loudspeakers, cf. Fig. 2.14. In the second case, the different propagation paths from the two loudspeakers to the ears create a comb filter. This comb filter is not present for sources
played from a single loudspeaker. Thus, moving a source between the two directions yields noticeable coloration. This is in contrast to static sources, for which Theile’s experiments [35] indicated that they are perceived without coloration.

The actual shape of the afore-mentioned comb filter depends on the angular distance between the loudspeakers. The first notch and its depth decreases with the distance. This implies that coloration increases for playback with higher loudspeaker densities.

A similar comb filter is created when using a triplet of loudspeakers with the same loudspeaker density as the pair, e.g. L, C, R compared to C, R. In order to avoid a strong increase in source width or annoying phasing effects, the outmost loudspeakers L and R are strongly reduced in their level, typically around -12dB compared to loudspeaker C. In doing so, the similarity of the comb filters yields barely any coloration when moving a source between the two directions, cf. Fig. 2.15.

Judging from what is shown above, it appears beneficial to activate always a few loudspeaker to stabilize the coloration, as opposed to using just one loudspeaker and moving the playback to another one. Keeping the number of simultaneously active loudspeakers more or less constant does not only prevent coloration of source movements, it also yields a more constant source width. Because of this relation between coloration and source width, the fluctuation of \( ||r_E|| \) is also a simple predictor of panning-dependent coloration.

In general, the strongest coloration is perceived under anechoic listening conditions. In reverberant rooms, the additional comb filters introduced by reflections help to conceal the comb filters due to multi-loudspeaker playback.
2.5 Open Listening Experiment Data

Experimental data from azimuthal localization in frontal and lateral loudspeaker pairs Figs. 2.3 and 2.4, azimuthal/elevational localization in horizontal, skew, and vertical frontal pairs Fig. 2.5, triangles Fig. 2.7, and quadrilaterals Fig. 2.8 are available online at https://opendata.iem.at in the listening experiment data project, as well as the data to the width experiment in Fig. 2.11.

The opendata.iem.at listening experiment data project contains evaluation routines to analyze the 95%-confidence intervals symmetrically based on means, standard deviations and the inverse Student’s t-distribution CIMEAN.m, or more robustly based on median and inter-quartile ranges CI2.m and Student’s t-distribution, or for two-dimensional data analysis robust_multivariate_confidence_region.m. The MATLAB script plot_gathered_data.m reads the formatted listening experiment data and its exemplary code generates figures like the above.

In order to support others providing own listening experiment data, the MATLAB functions write_experimental_data.m read_experimental_data.m are provided on the website.

References

1. M. Gerzon, General metatheory of auditory localization, in prepr. 3306, Convention Audio Engineering Society (1992)
2. D.M. Leakey E.C. Cherry, Influence of noise upon the equivalence of intensity differences and small time delays in two-loudspeaker systems. J. Acoust. Soc. Am. 29, 284–286 (1957)
3. K. Wendt, Das Richtungshören bei Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitsterophonie, Ph.D. Thesis, RWTH-Aachen (1963)
4. G. Theile, G. Plenge, Localization of lateral phantom sources. J. Audio Eng. Soc. 25(4), 96–200 (1977)
5. L. Simon, R. Mason, F. Rumsey, Localisation curves for a regularly-spaced octagon loudspeaker array, in prepr. 7015, Convention Audio Engineering Society (2009)
6. T. Kimura, H. Ando, 3D audio system using multiple vertical panning for large-screen multiview 3D video display. ITE Trans. Media Technol. Appl. 2(1), 33–45 (2014)
7. F. Wendt, M. Frank, F. Zotter, Panning with height on 2, 3, and 4 loudspeakers, in Proceedings of the ICSA (Erlangen, 2013)
8. H. Lee, F. Rumsey, Level and time panning of phantom images for musical sources. J. Audio Eng. Soc. 61(12), 978–988 (2013)
9. J.M. Helm, E. Kurz, M. Frank, Frequency-dependent amplitude-panning curves, in 29th Tonmeistertagung (Köln, 2016)
10. M. Frank, Phantom sources using multiple loudspeakers in the horizontal plane, Ph.D. Thesis, Kunstuni Graz (2013)
11. M. Frank F. Zotter, Extension of the generalized tangent law for multiple loudspeakers, in Fortschrritte der Akustik - DAGA (Kiel, 2017)
12. V. Pulkkki, Virtual sound source positioning using vector base amplitude panning. J. Audio Eng. Soc. 45(6), 456–466 (1997)
13. G. Soulodre, Evaluation of objective measures of loudness, in prepr. 6161, 116th AES Convention (Berlin, 2004)
14. M.-V. Laitinen, J. Vilkamo, K. Jussila, A. Politis, V. Pulkkki, Gain normalization in amplitude panning as a function of frequency and room reverberance, in Proceedings 55th AES Conference (Helsinki, 2014)
15. V. Pulkkki, Compensating displacement of amplitude-panned virtual sources, in 22nd Conference Audio Engineering Society (2003)
16. G. Martin, W. Woszczyk, J. Corey, R. Quesnel, Sound source localization in a five-channel surround sound reproduction system, in prepr. 4994, Convention Audio Engineering Society (1999)
17. F. Zotter, M. Frank, Generalized tangent law for horizontal pairwise amplitude panning, in Proceedings of the 3rd ICSA, Graz (2015)
18. T. Kimura, H. Ando, Listening test for three-dimensional audio system based on multiple vertical panning, in Acoustics-12, Hong Kong (2012)
19. F. Wendt, Untersuchung von phantomschallquellen vertikaler lautsprecheranordnungen, M. Thesis, Kunstuni Graz (2013)
20. H. Clark, G. Dutton, P. Vanerlyn, The ‘stereosonic’ recording and reproduction system. Repr. J. Audio Eng. Soc. Proc. Inst. Electr. Eng. 1957 6(2), 102–117 (1958)
21. V. Pulkkki, Localization of amplitude-panned virtual sources ii: two- and three-dimensional panning. J. Audio Eng. Soc. 49(9), 753–767 (2001)
22. B. Rakerd, W.M. Hartmann, J. Hsu, Echo suppression in the horizontal and median sagittal planes. J. Acoust. Soc. Am. 107(2) (2000)
23. P.W. Robinson, A. Walthier, C. Faller, J. Braasch, Echo thresholds for reflections from acoustically diffusive architectural surfaces. J. Acoust. Soc. Am. 134(4) (2013)
24. F. Zotter, M. Frank, M. Kronlachner, J. Choi, Efficient phantom source widening and diffuseness in ambisonics, in EAA Symposium on Auralization and Ambisonics (Berlin, 2014)
25. P. Stitt, S. Bertet, M. van Walstijn, Extended energy vector prediction of ambisonically reproduced image direction at off-center listening positions. J. Audio Eng. Soc. 64(5), 299–310 (2016)
26. P. Stitt, S. Bertet, M. Van Walstijn, Off-center listening with third-order ambisonics: dependence of perceived source direction on signal type. J. Audio Eng. Soc. 65(3), 188–197 (2017)
27. E. Kurz, M. Frank, Prediction of the listening area based on the energy vector, in Proceedings of the 4th International Conference on Spatial Audio (ICSA) (Graz, 2017)
28. M. Frank, Source width of frontal phantom sources: perception, measurement, and modeling. Arch. Acoust. 38(3), 311–319 (2013)
29. F. Rumsey, S. Zieliński, R. Kassier, S. Bech, On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality. J. Acoust. Soc. Am. 118(2), 968–976 (2005), http://link.aip.org/link/?JAS/118/968/1
30. S. Choisel, F. Wickelmaier, Evaluation of multichannel reproduced sound: scaling auditory attributes underlying listener preference. J. Acoust. Soc. Am. 121(1), 388–400 (2007)
31. K. Ono, V. Pulkki, M. Karjalainen, Binaural modeling of multiple sound source perception: methodology and coloration experiments. Audio Eng. Soc. Conv. 111, 11 (2001), http://www.aes.org/e-lib/browse.cfm?elib=9884
32. K. Ono, V. Pulkki, M. Karjalainen, Binaural modeling of multiple sound source perception: Coloration of wideband sound. Audio Eng. Soc. Conv. 112 4 (2002), http://www.aes.org/e-lib/browse.cfm?elib=11331
33. A. Ramsteiner, G. Spikofski, Ermittlung von Wahrnehmbarkeitsschwellen für Klangfarbenunterschiede unter Verwendung eines diffusfeldentzerrten Kopfhörers (Fortschritte der Akustik, DAGA, 1987)
34. M. Karjalainen, E. Piirilä, A. Järvinen, J. Huopaniemi, Comparison of loudspeaker equalization methods based on DSP techniques. J. Audio Eng. Soc. 47(1/2), 14–31 (1999), http://www.aes.org/e-lib/browse.cfm?elib=12117
35. G. Theile, Über die Lokalisation im überlagerten Schallfeld, Ph.D. dissertation, Technische Universität (Berlin, 1980)

Open Access  This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.