

# Chapter 7

## Compact Spherical Loudspeaker Arrays



*Le haut-parleur anonymise la source réelle.*

Pierre Boulez [1], ICA, 1983.

*...adjustable radiation naturalize[s] alien sounds by embedding them in the natural spatiality of the room.*

OSIL project [2], InSonic, 2015.

**Abstract** This chapter introduces auditory objects that can be created by adjustable-directivity sources in rooms. After showing basic positioning properties in distance and direction, we describe physical first- and higher-order spherical loudspeaker arrays and their control, such as the loudspeaker cubes or the icosahedral loudspeaker (IKO). Not only static auditory objects, but such traversing space by their time-varying beam forming are considered here. Signal dependency and different practical setups are discussed and briefly analyzed. This young Ambisonic technology brings new means of expression to sound reinforcement, electroacoustic or computer music.

While surrounding Ambisonic loudspeaker arrays play sound from outside the listening area into the audience, compact spherical loudspeaker arrays play sound into the room from a single position. Directivity adjustable in orientation and shape can be used to steer sound beams in order to excite wall reflections in the given, acoustic environment. The directional shapes and orientations of such beams are all controlled by—guess what—Ambisonic signals. Despite the huge practical difference, both applications do not only share the spherical harmonics that lend their shapes to Ambisonic signals: The control of radiating sound beams employs nearly the same model- or measurement-based radial steering filters as those of compact higher-order Ambisonic microphones.

The works of Warusfel [3], Kassakian [4], Avizienis [5], Zotter [6, 7], Pomberger [8], Pollow [9], Mattioli Pasqual [10] established the electroacoustic background technology required to describe compact spherical loudspeaker arrays built with electrodynamic transducers. The early works on auditory objects were written by Schmeder [11], Sharma, Frank, and Zotter [2, 12, 13]. And some contemporary results were found in the project “Orchestrating the Space by Icosahedral Loudspeaker” (OSIL) between 2015 and 2018 [14–19].

## 7.1 Auditory Events of Ambisonically Controlled Directivity

### 7.1.1 Perceived Distance

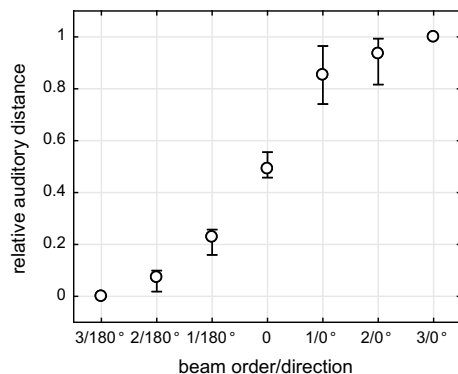
Laitinen showed in [20] that increasing the directivity of a listener-facing loudspeaker array from omnidirectional to second order was able to create auditory events that were perceptually closer than the physical distance to the loudspeaker array. The experimental results can be explained by the increase of the direct-to-reverberant energy ratio, as the sound beam of the directional source does not as much excite room reflections.

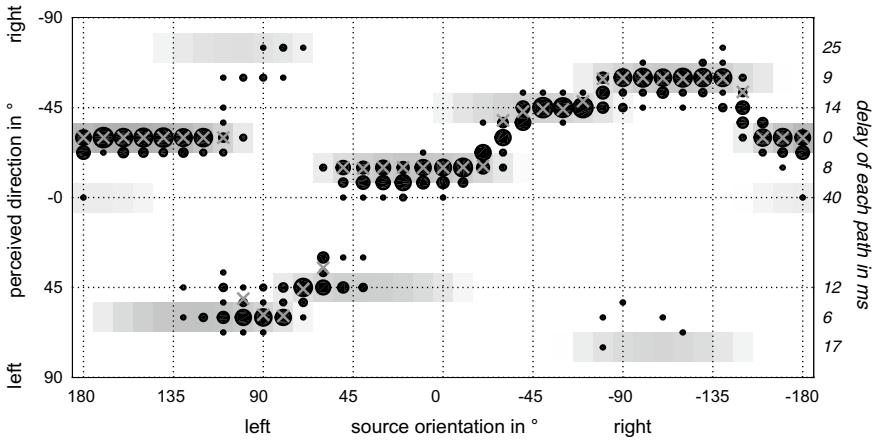
Wendt extended Laitinen's work by experiments employing a simulation of a third-order directional source in a virtual room (third-order image source model) played back by a loudspeaker ring in an anechoic room [16]. He could show that the perceived distance between the listener and the higher-order directional source could not only be controlled by the order of the directivity pattern but also by the orientation of the source (towards the listener, away from the listener). Beams projecting sounds away from the listener were perceived behind the source, cf. Fig. 7.1. Again, the perceptual results could be modeled by simple measures known from room acoustics.

### 7.1.2 Perceived Direction

Using a similar room simulation, the study in [21] asked participants to indicate the perceived direction of an auditory event created by a third-order directional source. The results showed that for different source orientations, listeners perceived auditory objects at directions that often did not coincide with the sound source, but with the delayed reflection paths, cf. Fig. 7.2. Perceived directions focused on the direct sound and the three first reflections after 6, 8, and 9 ms. For some orientations, still even

**Fig. 7.1** Perceived distance (medians and 95% confidence intervals) depends on order and direction of a static source emitting sound with  $\max-r_E$  directivity ( $180^\circ$  towards listener,  $0^\circ$  away from listener) in a simulated environment





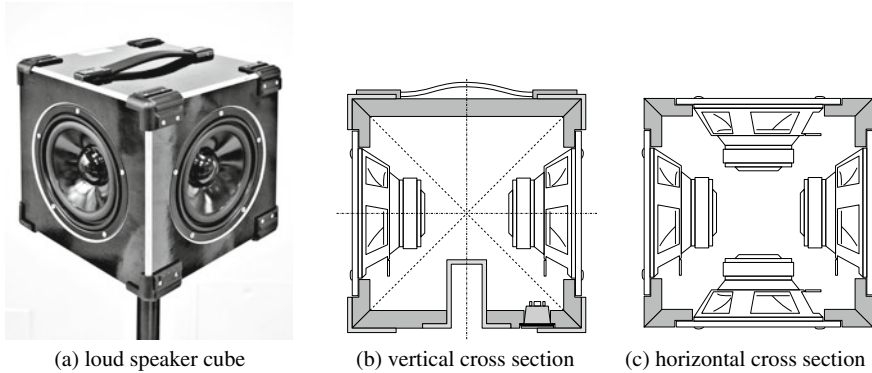
**Fig. 7.2** IKO perceived directions (black circles, radii indicate relative amount of answers) and modeling (gray crosses), 3rd-order max- $r_E$  beam, 2nd-order image source model. Gray shading in the background indicates level of each path

the second-order reflections at 12 and 14 ms were dominating localization. However, the influence of later reflections is reduced by the precedence effect. The perceived directions can be modeled by the extended energy vector originally developed for off-center listening positions in surrounding loudspeakers arrangements, as also shown in [17]. Experiments in [22] showed that panning between a reflection and the direct sound creates auditory objects in between. When applying the appropriate delay and gain to the direct sound to compensate for the longer path of the reflection, the localization curves are similar to those of standard stereo using a pair of loudspeakers.

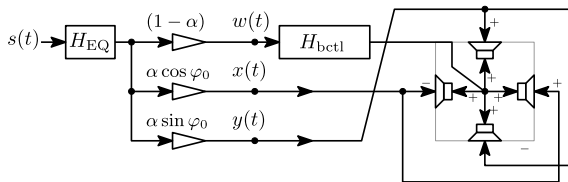
## 7.2 First-Order Compact Loudspeaker Arrays and Cubes

The simplest way of creating a loudspeaker array with adjustable directivity in a practical sense is a cube with loudspeakers on its plane surfaces, as suggested by Misdariis [23]. Restricting the directivity control to two dimensions reduces the number of loudspeaker drivers to four and facilitates to equip the array with a carrying handle on top and a flange adapter at the bottom, cf. [24] and Fig. 7.3.

**Directivity control.** First-order Ambisonics utilizes monopole and dipole modes, which directly translate to the corresponding far-field radiation patterns. These modes can easily be created due to the cubic shape by either playing of all four drivers in phase or the opposing drivers out of phase, cf. Fig. 7.4. Nevertheless, the frequency responses of such monopole and dipole modes need to be equalized to enable their phase- and magnitude-aligned superposition in the far field. Filters and measurement data of cube loudspeakers built at IEM [24] are freely available on <http://phaidra.kug.ac.at/o:67631>.



**Fig. 7.3** Design of a loudspeaker cube: prototype, and vertical and horizontal cross section plots

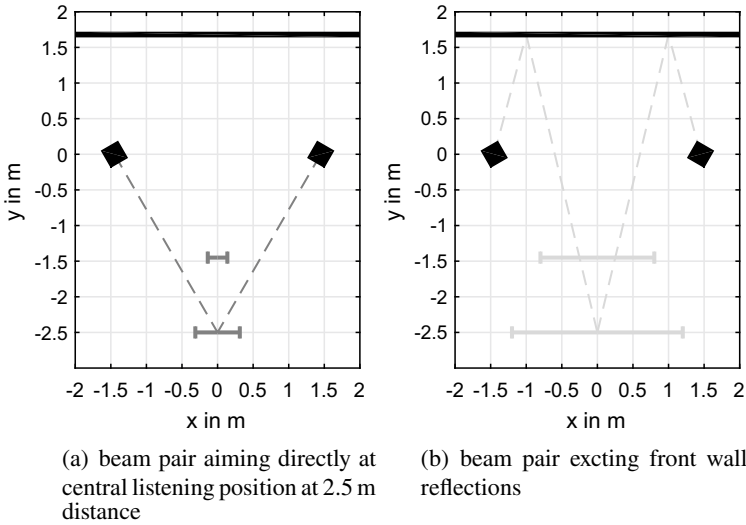


**Fig. 7.4** System controlling the monopoles and dipole modes of the loudspeaker cubes, to accomplish first-order beamforming with the shape parameter  $\alpha$  and beam direction  $\varphi_0$

To overcome the compressive effort of interior volume changes at low frequencies, the filter  $H_{bctl}$  in Fig. 7.4 equalizes the smaller velocity of the loudspeaker cones when driven omnidirectionally to the velocity when driven in dipoles as a first step, and as a second step, it attenuates the monopole pattern slightly to account for its more efficient radiation at low frequencies. The filter  $H_{EQ}$  is a general equalizer required to obtain a flat frequency response,  $0 \leq \alpha \leq 1$  is a first-order omni to dipole beam-shape parameter, and  $\varphi_0$  is the beam direction. The filter  $H_{bctl}$  can be specified as a 5th-order IIR filter purely based on geometric and electroacoustic parameters [19].

**Direct and indirect sound with two cubes.** The study in [19] examined the width of the listening area for the creation of a central auditory object between a pair of loudspeaker cubes cf. Fig. 7.5. Steering the two beams directly at the listener yielded a narrow listening area that increased with the distance to the loudspeakers, similar as known from typical stereo applications, cf. Fig. 2.9. A much wider listening area is achieved by steering the beams to the front wall to excite reflections. To this end,  $\max\text{-}r_E$  (super-cardioid) beams were chosen and oriented in a way to ideally suppress direct sound from the loudspeaker cubes at the listening position. The proposed setup of two loudspeaker cubes can be used to play back stable L, C, R channels of a surround production without the need of an actual center loudspeaker.

**Surround with depth:** Together with the distance control described by Laitinen [20], the stable in-between auditory image has been used in [19] to establish a *surround-*

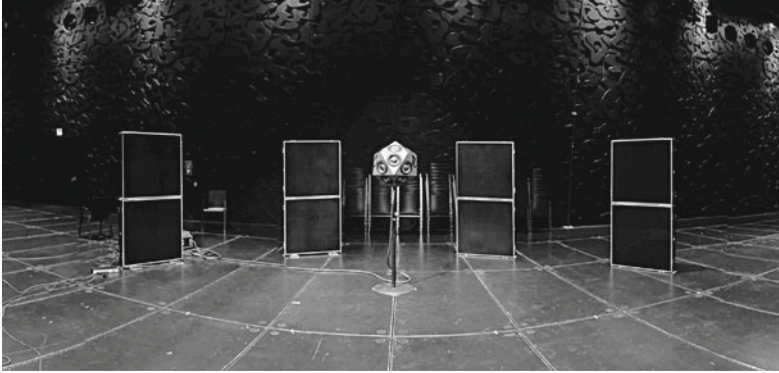


**Fig. 7.5** Width of the listening area for a central auditory object at two distances from a pair of loudspeaker cubes with different orientation of max- $r_E$ /super-cardioid beams

*with-depth* system consisting of a quadraphonic setup of four loudspeaker cubes. As first layer, it uses the direct sounds from the 4 loudspeakers from  $\pm 45^\circ$  and  $\pm 135^\circ$  together with the 4 in-between images at  $0^\circ$ ,  $\pm 90^\circ$ , and  $180^\circ$  to obtain 8 directions for third-order Ambisonic surround panning. As a second layer for depth, *surround with depth* uses 4 cardioid beams pointing into the 4 room corners to provide the impression of distant sounds. Blending between those two layer is used to control the distance impression of surround sounds.

### 7.3 Higher-Order Compact Spherical Loudspeaker Arrays and IKO

With transducers mounted on spheres or polyhedra, higher-order radiators can be built. Typically, those are Platonic solids such as dodecahedra or icosahedra, as they can easily be manufactured from equal-sided polygons cf. Fig. 7.6. Often, the loudspeakers are also mounted onto a common interior volume. Hereby, the higher-order modes can be controlled at reduced impedance of the inner stiffness, however, this also causes acoustic coupling of the transducer motions. Typically, multiple-input-multiple-output (MIMO) crosstalk cancellers are employed to suppress the coupling and to control the velocity of the transducer cones. If this is accomplished, the acoustic radiation can be modeled and equalized by the spherical cap model, cf. [6, 15, 25, 26].



**Fig. 7.6** Powerful icosahedral loudspeaker array (IKO by IEM and Sonible) and reflecting baffles in Ligeti concert hall, in preparation of an electroacoustic music concert

**Cap model.** Higher-order loudspeaker arrays on a compact spherical housing are modeled by the spherical cap model. It assumes for the exterior air that the radial surface velocity is a boundary condition consisting of separated spherical cap shapes of the size  $\alpha$  centered around the directions  $\{\theta_l\}$ , each unity in value. These idealized transducer shapes driven by the transducer velocities  $v_l$  compose the surface velocity

$$v(\theta) = \sum_{l=0}^L u(\theta_l^T \theta - \cos \frac{\alpha}{2}) v_l. \quad (7.1)$$

Here,  $u(\zeta)$  denotes the unit step function that is unity for  $\zeta \geq 0$  and zero otherwise. The surface velocity distribution can be decomposed into spherical harmonics as

$$v(\theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\theta) \sum_{l=0}^L w_{nm}^{(l)} v_l. \quad (7.2)$$

The coefficients  $w_{nm}^{(l)}$  of the  $l$ th cap are defined by spherical convolution Eq. (A.56) of a Dirac delta  $\delta(\theta_l^T \theta - 1)$  pointing to the cap center with a zenithal cap  $u(\cos \vartheta - \cos \frac{\alpha}{2})$ :

$$w_{nm}^{(l)} = w_n Y_n^m(\theta_l), \quad (7.3)$$

where  $Y_n^m(\theta_l)$  are the coefficients expressing the Dirac delta, extended to a cap by weighting with  $w_n$ . The term  $w_n = 2\pi \int_{\cos \frac{\alpha}{2}}^1 P_n(\zeta) d\zeta$  is derived in Eq. (A.60)

$$w_n = 2\pi \begin{cases} -\frac{P_{n+1}(\cos \frac{\alpha}{2}) - \cos \frac{\alpha}{2} P_n(\cos \frac{\alpha}{2})}{n}, & \text{for } n > 0, \\ 1 - \cos \frac{\alpha}{2}, & \text{for } n = 0. \end{cases} \quad (7.4)$$

**Decoder.** Without radiation control yet, any low-order target spherical harmonic  $n \leq N$  can be synthesized as velocity pattern  $\phi_{nm}$  by superimposing the spherical cap coefficients  $w_{nm}^{(l)}$  with suitable transducer velocities  $v_l$ , i.e.  $\phi_{nm} = \sum_l w_n Y_n^m(\theta_l) v_l$ . We write a matrix/vector notation with the matrix  $\mathbf{Y} = [\mathbf{y}(\theta_1), \dots, \mathbf{y}(\theta_L)]$  containing the spherical harmonics  $\mathbf{y}(\theta) = [Y_n^m(\theta)]_{nm}$  sampled at the transducer positions  $\{\theta_l\}$  to represent Dirac deltas pointing there, and  $\mathbf{w} = [w_n]_{nm}$  to represent the cap shape,

$$\boldsymbol{\phi} = \text{diag}\{\mathbf{w}\} \mathbf{Y} \mathbf{v}. \quad (7.5)$$

As long as the order  $N$  up to which coefficients are controlled is low enough  $L \geq (N+1)^2$  and transducers are well-distributed, perfect control is feasible. The corresponding velocities are found by solving a least-squares problem, see Appendix A.4, Eq. (A.63), yielding the right inverse of the  $N$ th-order cap-coefficient matrix,

$$\mathbf{v} = \mathbf{Y}_N^T (\mathbf{Y}_N \mathbf{Y}_N^T)^{-1} \text{diag}\{\mathbf{w}_N\}^{-1} \boldsymbol{\phi}_N = \mathbf{D} \text{diag}\{\mathbf{w}_N\}^{-1} \boldsymbol{\phi}_N. \quad (7.6)$$

The right inverse  $\mathbf{D} = \mathbf{Y}_N^T (\mathbf{Y}_N \mathbf{Y}_N^T)^{-1}$  is a mode-matching decoder, cf. Eq. (4.40).

**Exterior problem.** The radiated sound pressure is described by the exterior problem denoted by the coefficients  $c_{nm}$  in Eq. (6.13) and the spherical Hankel functions  $h_n^{(2)}(kr)$ . To relate it to a time-derived surface velocity at the array radius  $r = a$ , we derive the exterior solution with regard to radius  $\frac{\partial p}{\partial r} = k \frac{\partial p}{\partial kr} = -ikc \rho v$ , cf. Eq. (6.2),

$$v(\boldsymbol{\theta}) = \frac{i}{\rho c} \sum_{n=0}^{\infty} \sum_{m=-n}^n h_n^{(2)}(ka) Y_n^m(\boldsymbol{\theta}) c_{nm}. \quad (7.7)$$

Comparing Eq. (7.2) to Eq. (7.7) yields  $c_{nm} = \rho c [i h_n^{(2)}(ka)]^{-1} \sum_{l=0}^L w_n Y_n^m(\theta_l) v_l$ , the coefficients to calculate the radiated pressure. Far away, we replace the spherical Hankel function that approaches  $h_n^{(2)}(kr) \rightarrow i^{n+1} k^{-1} e^{-ikr}$  by the term  $i^{n+1} k^{-1}$  in Eq. (6.13) so that the radiated far-field sound pressure  $p \propto \sum i^{n+1} k^{-1} Y_n^m c_{nm}$  becomes

$$p(\boldsymbol{\theta}) \propto \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\boldsymbol{\theta}) \frac{i^n w_n}{k h_n^{(2)}(ka)} \sum_{l=0}^L Y_n^m(\theta_l) v_l. \quad (7.8)$$

### 7.3.1 Directivity Control

The spherical harmonics coefficients of the far-field sound pressure pattern in Eq. (7.8) are controlled by the cap velocities  $v_l$

$$\psi_{nm} = \frac{i^n w_n}{k h_n^{(2)}(ka)} \sum_{l=0}^L Y_n^m(\theta_l) v_l, \quad (7.9)$$

and we desire to form the directional sound beam they represent according to a max- $r_E$  pattern  $a_n Y_n^m(\theta_0)$  yielding radiation focused towards  $\theta_0$

$$\psi_{nm} = a_n Y_n^m(\theta_0). \quad (7.10)$$

To find suitable cap velocities  $v_l$ , we equate the model Eqs. (7.9) and (7.10). In matrix/vector notation never used the equation is

$$\text{diag}\{[i^n w_n k^{-1} / h_n^{(2)}(ka)]_{nm}\} \mathbf{Y} \mathbf{v} = \text{diag}\{[a_n]_{nm}\} \mathbf{y}(\theta_0). \quad (7.11)$$

The diagonal matrix on the left is easy to invert, and for patterns up to the order  $n \leq N$ , the mode-matching decoder  $\mathbf{D}$  of Eq. (7.6) already gives us a way to define velocities inverting the matrix  $\mathbf{Y}_N$  from the right. The preliminary solution becomes

$$\Rightarrow \mathbf{v} = \mathbf{D} \text{diag}\{[i^{-n} w_n^{-1} k h_n^{(2)}(ka) a_n]_{nm}\} \mathbf{y}_N(\theta_0). \quad (7.12)$$

***On-axis equalized, sidelobe-suppressing directivity control limiting the excursion.***

The inverse cap shape coefficient  $w_n^{-1}$  and the max- $r_E$  weight  $a_n$  can be regarded as a part of the radiation control filters  $i^{-n} k h_n^{(2)}(ka)$ . The expression  $i^{-n-1} (ka)^2 h_n^{(2)}(ka)$  of compact spherical microphone arrays (Sect. 6.6) qualitatively differs by a factor  $k$ . Practical implementation of radiation control filters and their regularization is therefore quite similar to radial filters of spherical microphone arrays. There are three main differences, as explained in [15]:

- With loudspeaker arrays, it is rather the excursion that is limited, which primarily entails a different strategy of adjusting the filter bank cut-on frequencies, which due to size are at lower frequencies where group-delay distortions are less disturbing, and linear-phase implementations would cause avoidably long delays.
- Moreover, instead of cut-on filter slopes of  $(n + 1)$ th order required for noise removal in signals obtained from spherical microphone arrays, limited excursion requires cut-on slopes of at least  $(n + 3)$ th order, i.e. 4th order to cut on the 1st-order Ambisonic signals. Thereof, one additional order is caused by the qualitative difference of  $k^{-1}$  in radial filters, and another order by the conversion of velocity to excursion by a factor  $(i\omega)^{-1}$ .
- Finally, instead of diffuse-field equalization that is useful for surround sound playback of spherical microphone array signals, it is more useful to equalize spherical sound beams on-axis (free field).

On-axis equalization yields a different scaling of the sub-band max- $r_E$  weights

$$a_{n,b} = \begin{cases} P_n \left( \cos \frac{137.9^\circ}{b+1.51} \right) \frac{\sum_{n=0}^N (2n+1) P_n \left( \cos \frac{137.9^\circ}{N+1.51} \right)}{\sum_{n=0}^b (2n+1) P_n \left( \cos \frac{137.9^\circ}{b+1.51} \right)}, & \text{for } n \leq b, \\ 0, & \text{otherwise.} \end{cases} \quad (7.13)$$



Typically, cut-on frequencies for compact spherical loudspeaker arrays are low, and linear-phase filterbanks would require long pre-delays. It is useful to employ Linkwitz-Riley filters for the crossovers, to get a low-latency implementation. To emphasize the similarity to Eq. (6.22), we write Linkwitz-Riley filters [27] as combination of an all-pass  $A^m$  with twice the phase response of an  $m$ th-order Butterworth low-pass combined either with the magnitude-squared low-pass response  $[1 + (\omega/\omega_c)^{2m}]^{-1}$  or high-pass response  $(\omega/\omega_c)^{2m}[1 + (\omega/\omega_c)^{2m}]^{-1}$  per crossover. Such a minimum-phase crossover is of even order, so that the minimum-order cut-on slope must be rounded up to the next even order  $2\lceil \frac{b+3}{2} \rceil$ . Plain high/low crossovers would be in-phase unless combined with further crossovers to form narrower bands. However, an in-phase filterbank is obtained after inserting the product of all all-passes in every band, cf. [28]. Although non-minimum-phase, this is still low-latency. For the band  $b$  containing Ambisonic orders  $0 \leq n \leq b$ , the modified filterbank is

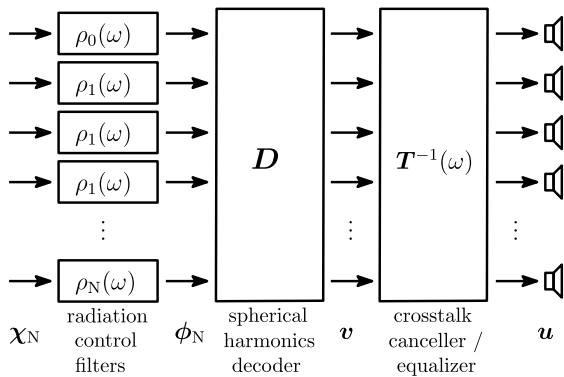
$$H_b(\omega) = \frac{\left(\frac{\omega}{\omega_b}\right)^{2\lceil \frac{b+3}{2} \rceil}}{1 + \left(\frac{\omega}{\omega_b}\right)^{2\lceil \frac{b+3}{2} \rceil}} \frac{1}{1 + \left(\frac{\omega}{\omega_{b+1}}\right)^{2\lceil \frac{b+4}{2} \rceil}} \prod_{b'=0}^N A_{\omega_{b'}}^{\lceil \frac{b'+3}{2} \rceil}(\omega). \tag{7.14}$$

The sum  $\sum_b H_b(\omega)$  is considered to be sufficiently flat, so that the radial filters for compact spherical loudspeaker arrays using Eqs. (7.8), (7.13), (7.14) become

$$\rho_n(\omega) = \left[ \sum_{b=n}^N a_{n,b} H_b(\omega) \right] i^{-n} w_n^{-1} h_n^{(2)}(ka) e^{ika}. \tag{7.15}$$

Figure 7.7 shows the block diagram to control compact spherical loudspeaker arrays by Ambisonic input signals, including the radiation control filters, the decoder, and a voltage-equalizing crosstalk canceller feeding the loudspeakers.

**Fig. 7.7** Signal processing for higher-order compact spherical loudspeaker array control: the Ambisonic directivity signals  $\chi_N(t)$  run through radiation control filters  $\rho_n(\omega)$ , a decoder yielding desired velocities  $\mathbf{v}(t)$ , and a crosstalk canceller/equalizer provides suitable output voltages  $\mathbf{u}(t)$



### 7.3.2 Control System and Verification Based on Measurements

**Velocity equalization/crosstalk cancellation.** In the frequency domain, laser vibrometer measurements, cf. Fig. 7.8a, characterize the physical multiple-input-multiple-output (MIMO) system of transducer input voltages  $u_l(\omega)$  to transducer velocities  $v_l(\omega)$

$$\mathbf{v}(\omega) = \mathbf{T}(\omega) \mathbf{u}(\omega), \quad (7.16)$$

including the effect of acoustic coupling through the common enclosure. Corresponding open measurement data sets<sup>1</sup> can be found online, as described in [18]. Theoretically, the frequency-domain inverse of the matrix  $\mathbf{T}(\omega)$  can be used to equalize and control the transducer velocities with acoustic crosstalk cancelled, as indicated in Fig. 7.7,

$$\mathbf{u}(\omega) = \mathbf{T}^{-1}(\omega) \mathbf{v}(\omega). \quad (7.17)$$

In practice, this is only useful up to the frequency at which the loudspeaker cone vibration breaks up into modes, so typically below 1 kHz.

**Control system:** The entire control system with Ambisonic signals  $\chi_N(\omega)$  as inputs uses Eqs. (7.6), (7.15), (7.17)

$$\mathbf{u}(\omega) = \mathbf{T}^{-1}(\omega) \mathbf{D} \text{diag}\{\rho(\omega)\} \chi_N(\omega). \quad (7.18)$$

**Directivity measurement.** It is useful to characterize the directivity obtained by measurements to verify the results; high-resolution  $648 \times 20$  measurements  $\mathbf{G}(\omega)$  of the IKO are found online<sup>1</sup>. The sound pressure can be decomposed with the known directional sampling by left-inversion of a spherical harmonics matrix  $\mathbf{Y}_{17}^T$ , see Appendix A.4, Eq. (A.65, which can be up to 17th order on a  $10^\circ \times 10^\circ$  grid in azimuth and zenith:

$$\mathbf{p}(\omega) = \mathbf{G}(\omega) \mathbf{u}(\omega), \quad \Rightarrow \boldsymbol{\psi}_{17}(\omega) = (\mathbf{Y}_{17}^T)^\dagger \mathbf{p}(\omega). \quad (7.19)$$

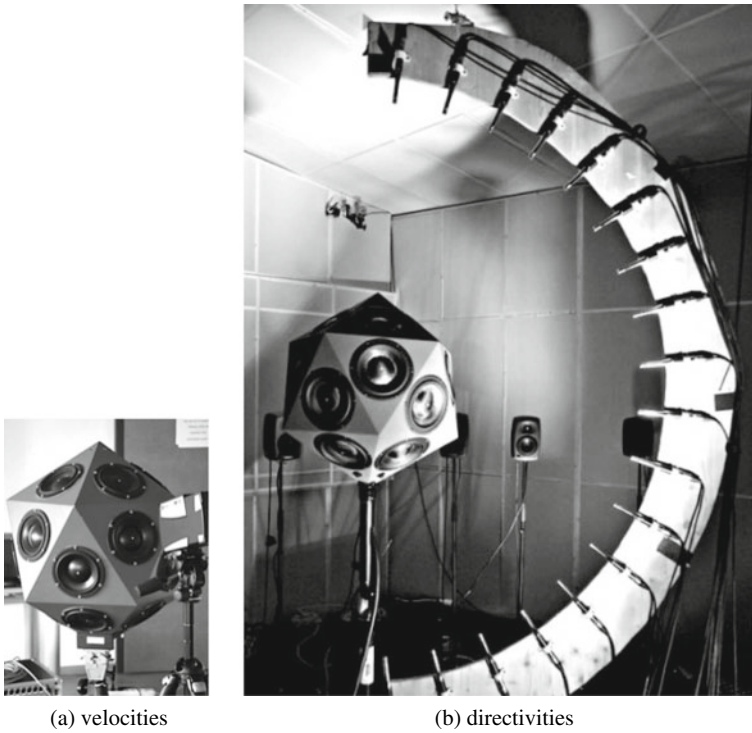
With the highly resolved spherical harmonics coefficients, polar diagrams or balloon diagrams can be evaluated at any direction

$$p(\boldsymbol{\theta}, \omega) = \mathbf{y}_{17}(\boldsymbol{\theta})^T \boldsymbol{\psi}_{17}(\omega), \quad (7.20)$$

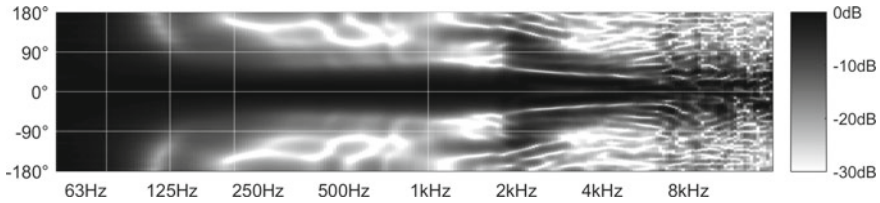
given any control system delivering suitable voltages  $\mathbf{u}$  for beamforming, as e.g. obtained by Eq. (7.18).

---

<sup>1</sup><http://phaidra.kug.ac.at/o:67609>.



**Fig. 7.8** Measurements on the IKO as a MIMO system in terms of transducer output velocities (left) and radiation patterns (right) depending on the transducer input voltages



**Fig. 7.9** Horizontal cross section of the IKO's directivity/dB over frequency/Hz and azimuth/degrees when beamforming to 0° azimuth on the horizon, with radiation control filters above, with filterbank frequencies (38, 75, 125, 210) Hz

To inspect the frequency-dependent directivity, a horizontal cross section is shown in Fig 7.9. The beamforming gets effective above 100 Hz and a beam width of  $\pm 30^\circ$  is held until 2 kHz. The filterbank starts the 0th order above 38 Hz, and with 75, 125, 210 Hz, 1st, 2nd, and 3rd order are successively added including on-axis equalized max- $r_E$  weightings. Above 2 kHz both spatial aliasing and modal breakup of the transducer cones affect directivity. However, these beamforming-direction-dependent distortions are often negligible in typical rooms.

## 7.4 Auditory Objects of the IKO

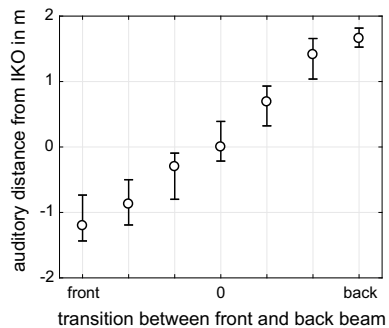
### 7.4.1 Static Auditory Objects

The study in [16] showed that distance control by changing the directivity and its orientation can also be achieved with the IKO in a real room, cf. Fig. 7.10. The experiments used stationary pink noise and could create auditory objects nearly 2 m behind the IKO, which corresponds to the distance between the IKO and the front wall of the playback room.

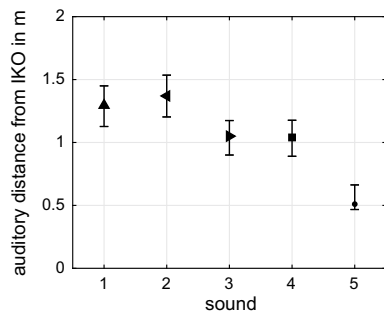
The maximum distance of auditory objects created by the IKO is strongly signal-dependent. Experiments in [14] showed that the auditory distance of pink noise bursts decreased for shorter fade-in times, while the fade-out time had no influence, cf. Fig. 7.11. A transient click sound was perceived even closer to the IKO. This can be explained by the precedence effect, that favors the earlier direct sound over the reflected sound from the walls. While this effect is strong for transient sounds, it is inhibited for stationary sounds with long fade-in times.

However, the precedence effect can even be reduced for transient click sounds by simultaneous playback of a masker sound that reduces the influence of the direct sound [29]. In comparison to no masker, playing a pink noise masker doubles the auditory distance, cf. Fig. 7.12. Using the room noise as a masker by playing the

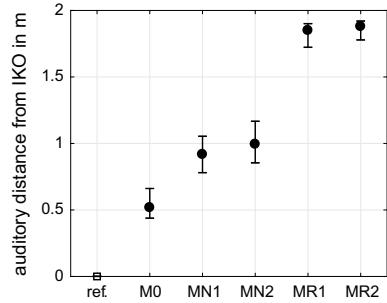
**Fig. 7.10** Distance control using the IKO in the IEM CUBE



**Fig. 7.11** Signal-dependent distance of auditory objects created by the IKO: markers indicate short or long fade-in and fade-out of pink noise bursts, · indicates transient click sound



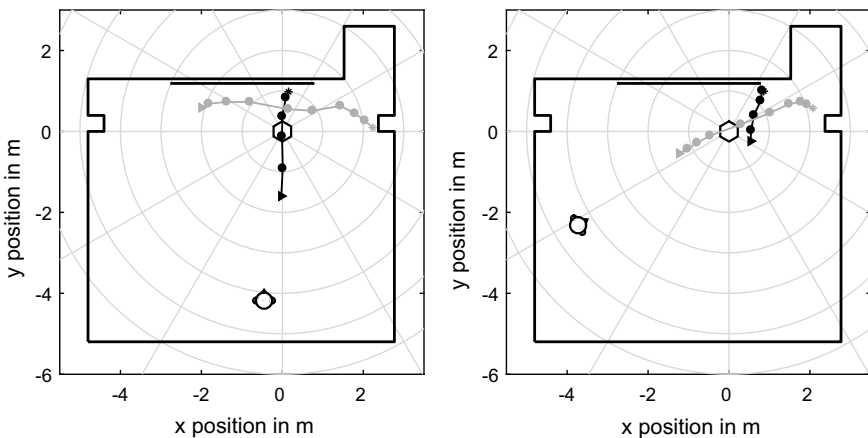
**Fig. 7.12** Distance of auditory objects created by the IKO playing a transient click sound in dependence of maskers: no masker (M0), noise at different levels (M1, M2) and playback at very low level (room masker MR1 and MR2)



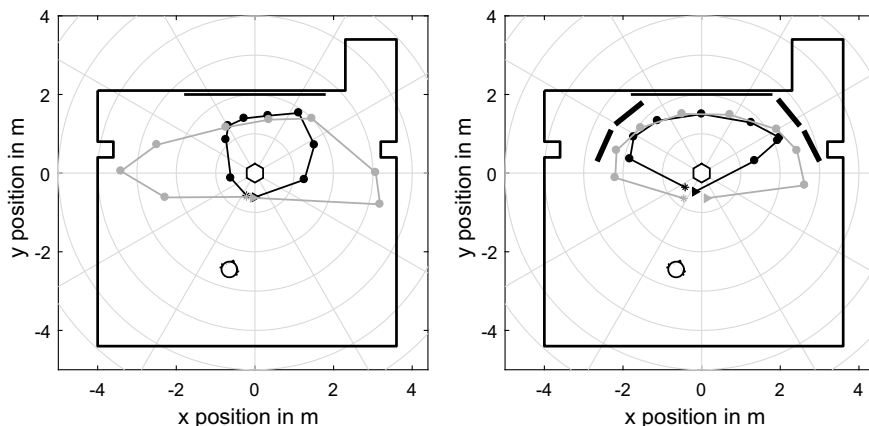
target sound very softly further increases the distance and yield a perception that is detached from the IKO.

### 7.4.2 Moving Auditory Objects

The studies in [14, 15] extended the previous listening experiments towards simple time-varying beam directions, such as from the left to the right, front/back or circles. To report the perceived locations of the moving auditory objects, listeners used a touch screen that showed a floor plan of the room, including the listening position and the position of the IKO. They had to indicate the location of the auditory object’s trajectory every 500 ms. The perceived trajectories depend on the listening position, but they can always be recognized, cf. Fig. 7.13. The empirical knowledge was



**Fig. 7.13** Average perceived locations for each 500 ms step during front/back-movement (dark gray) and left/right-movement (light gray) at two listening positions, triangle indicates start and asterisk end of the trajectory



**Fig. 7.14** Average perceived locations for each 500 ms step during circular movement of transient sound (dark gray) and stationary noise (light gray) without and with additional reflectors, triangle indicates start and asterisk end of the trajectory

applied in the artistic study in [14] about body-space relations, composing sounds that are spatialized with different static directions and simple movements.

For concerts, the artistic practice evolved to set the IKO up together with reflector baffles, cf. Fig. 7.14. A recent study in [30] investigated their effect on the perception of moving transient and stationary sounds. The baffles obviously reduce the signal-dependency by contributing more additional reflection paths, contrasting the direct sound.

## 7.5 Practical Free-Software Examples

### 7.5.1 IEM Room Encoder and Directivity Shaper

The IEM `Room Encoder` VST plug-in, cf. Fig. 4.36, can not only be used to simulate the room reflections of an omnidirectional sound source based on the image-source method, but it also supports directional sound sources. As format, it employs Ambisonics with ACN ordering and adjustable normalization up to seventh order. Thus, it enables to utilize data from directivity measurements or even directional recordings done with a surrounding spherical microphone, e.g. to put real instrument recordings into the virtual room.

As an alternative, the IEM `Directivity Shaper`, cf. Fig. 7.15 provides simple means to generate a frequency-dependent directivity pattern from scratch and to apply it on a mono input signal. This is useful to generate the typical rotary speaker effect of a Leslie cabinet.

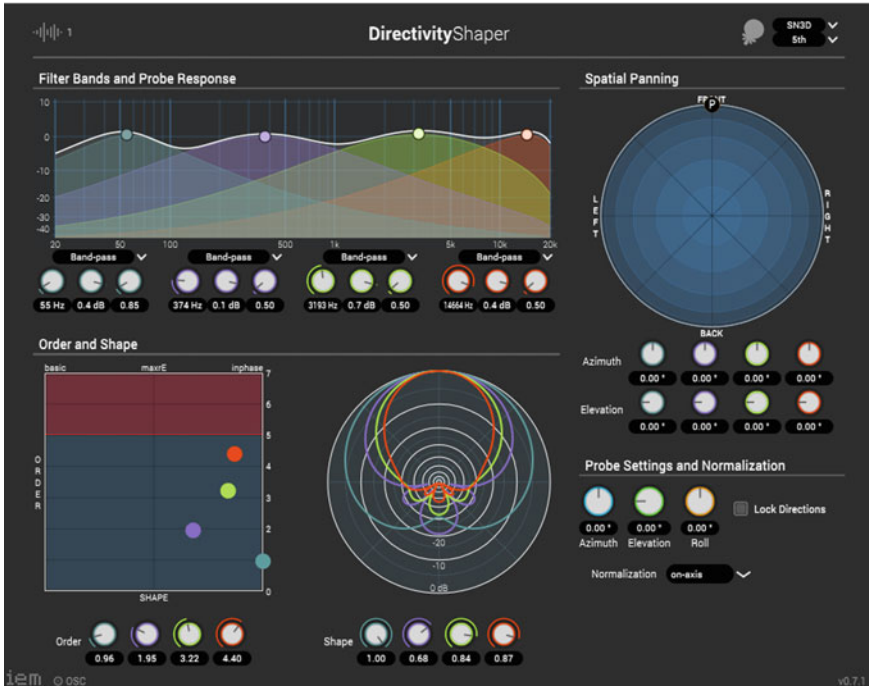


Fig. 7.15 IEM Directivity Shaper plug-in

### 7.5.2 IEM Cubes 5.1 Player and Surround with Depth

As shown in Fig. 7.5, a pair of loudspeaker cubes can create a stable auditory event in between them to replace an actual center loudspeaker. In order to play back an entire 5.1 production, the IEM cubes 5.1 Player plug-in extends this approach by two additional beams to the side walls for the surround channels, cf. Fig. 7.16. The plug-in provides a control of the shape, direction, and level of all beams, as well as a delay compensation for the reflection paths.

Surround sound with depth can be realized with a quadraphonic setup of four loudspeaker cubes and a combination of the cubes Surround Decoder and multiple Distance Encoder plug-ins, cf. Fig. 7.16. For each source, the Distance Encoder controls position and distance, i.e. the blending between the two layers. The output of the plug-in is a 10-channel audio stream including 7 channels for third-order (inner layer) and 3 for first-order 2D Ambisonics (outer depth layer). The cubes Surround Decoder plug-in decodes the 10-channel audio stream and distributes the signals to the 16 drivers of four loudspeaker cubes. For each loudspeaker cube, the directions to excite direct and reflected sound of the inner layer and the diffuse sound of the depth layer can be adjusted in order to adapt to the playback environment. Additionally, the directivity patterns for direct, reflected, and diffuse sound beams



Fig. 7.16 IEM cubes 5.1 Player, cubes Surround Decoder and Distance Encoder plug-ins

can be controlled, as well as a delay to compensate for the longer propagation paths of the reflected sound. The plug-ins are available under <https://git.iem.at/audioplugins/CubeSpeakerPlugins>.



### 7.5.3 IKO

Spatialization using the IKO can use a similar infrastructure of plug-ins as surrounding loudspeaker arrays. Ambisonic encoder plug-ins, such as the `ambix_encoder` or the IEM `StereoEncoder` or `MultiEncoder`, create the third-order Ambisonic signals that are subsequently fed to a decoder. Decoding to the IKO requires the processing steps as shown in Fig. 7.7: radiation control filters in the spherical harmonic domain, decoding from spherical harmonics to transducer signals, as well as crosstalk cancellation and equalization of the transducers. This processing can be summarized in a  $16$  (spherical harmonics up to third order)  $\times$   $20$  (transducers) filter matrix. Convolution can be done efficiently using the `mcfx_convolver` plug-in. Filter presets for the IKO can be found under <http://phaidra.kug.ac.at/o:79235>.

## References

1. P. Boulez, L'acoustique et la musique contemporaine, in *Proceedings of the 11th International Congress on Acoustics*, vol. 8 (Paris, 1983), <https://www.icacommission.org/Proceedings/ICA1983Paris/ICA11%20Proceedings%20Vol8.pdf>
2. M. Frank, G.K. Sharma, F. Zotter, What we already know about spatialization with compact spherical arrays as variable-directivity loudspeakers, in *Proceedings of the inSONIC2015* (Karlsruhe, 2015)
3. O. Warusfel, P. Derogis, R. Causse, Radiation synthesis with digitally controlled loudspeakers, in *103rd AES Convention* (Paris, 1997)
4. P. Kassakian, D. Wessel, Characterization of spherical loudspeaker arrays, in *Convention paper #1430* (San Francisco, 2004)
5. R. Avizienis, A. Freed, P. Kassakian, D. Wessel, A compact 120 independent element spherical loudspeaker array with programmable radiation patterns, in *120th AES Convention* (France, Paris, 2006)
6. F. Zotter, R. Höldrich, Modeling radiation synthesis with spherical loudspeaker arrays, in *Proceedings of the ICA* (Madrid, 2007), <http://iaem.at/Members/zotter>
7. F. Zotter, Analysis and synthesis of sound-radiation with spherical arrays, PhD. Thesis, University of Music and Performing Arts, Graz (2009)
8. H. Pomberger, Angular and radial directivity control for spherical loudspeaker arrays, M. Thesis, Institut für Elektronische Musik und Akustik, Kunstuni Graz, Technical University Graz, Graz, A (2008)
9. M. Pollow, G.K. Behler, Variable directivity for platonic sound sources based on spherical harmonics optimization. *Acta Acust. United Acust.* **95**(6), 1082–1092 (2009)
10. A.M. Pasqual, P. Herzog, J.R. Arruda, Theoretical and experimental analysis of the behavior of a compact spherical loudspeaker array for directivity control. *J. Acoust. Soc. Am.* **128**(6), 3478–3488 (2010)
11. A. Schmeder, An exploration of design parameters for human-interactive systems with compact spherical loudspeaker arrays, in *1st Ambisonics Symposium* (Graz, 2009)
12. G.K. Sharma, F. Zotter, M. Frank, Orchestrating wall reflections in space by icosahedral loudspeaker: findings from first artistic research exploration, in *Proceedings of the ICMC/SMC* (Athens, 2014), pp. 830–835
13. F. Zotter, M. Frank, A. Fuchs, D. Rudrich, Preliminary study on the perception of orientation-changing directional sound sources in rooms, in *Proceedings of the Forum Acusticum* (Kraków, 2014)

14. F. Wendt, G.K. Sharma, M. Frank, F. Zotter, R. Höldrich, Perception of spatial sound phenomena created by the icosahedral loudspeaker. *Comput. Music J.* **41**(1) (2017)
15. F. Zotter, M. Zaunschirm, M. Frank, M. Kronlachner, A beamformer to play with wall reflections: the icosahedral loudspeaker. *Comput. Music J.* **41**(3) (2017)
16. F. Wendt, F. Zotter, M. Frank, R. Höldrich, Auditory distance control using a variable-directivity loudspeaker. *MDPI Appl. Sci.* **7**(7) (2017)
17. F. Wendt, M. Frank, On the localization of auditory objects created by directional sound sources in a virtual room, in *VDT Tonmeistertagung* (Cologne, 2018)
18. F. Schultz, M. Zaunschirm, F. Zotter, Directivity and electro-acoustic measurements of the IKO, in *Audio Engineering Society Convention 144* (2018), <http://www.aes.org/e-lib/browse.cfm?elib=19557>
19. T. Deppisch, N. Meyer-Kahlen, F. Zotter, M. Frank, Surround with depth on first-order beam-controlling loudspeakers, in *Audio Engineering Society Convention 144* (2018), <http://www.aes.org/e-lib/browse.cfm?elib=19494>
20. M.-V. Laitinen, A. Politis, I. Huhtakallio, V. Pulkki, Controlling the perceived distance of an auditory object by manipulation of loudspeaker directivity. *J. Acoust. Soc. Am.* **137**(6) EL462–EL468 (2015)
21. F. Zotter, M. Frank, Investigation of auditory objects caused by directional sound sources in rooms. *Acta Phys. Pol. A* **128**(1-A) (2015), <http://przyrbwn.icm.edu.pl/APP/PDF/128/a128z1ap01.pdf>
22. F. Zagala, J. Linke, F. Zotter, M. Frank, Amplitude panning between beamforming-controlled direct and reflected sound, in *Audio Engineering Society Convention 142* (Berlin, 2017), <http://www.aes.org/e-lib/browse.cfm?elib=18679>
23. N. Misdariis, F. Nicolas, O. Warusfel, R. Caussee, Radiation control on multi-loudspeaker device : La timée,” in *International Computer Music Conference* (La Habana, 2001)
24. N. Meyer-Kahlen, F. Zotter, K. Pollack, Design and measurement of a first-order, in *Audio Engineering Society Convention 144* (2018), <http://www.aes.org/e-lib/browse.cfm?elib=19559>
25. F. Zotter, H. Pomberger, A. Schmeder, Efficient directivity pattern control for spherical loudspeaker arrays, in *ACOUSTICS08* (2008)
26. F. Zotter, A. Schmeder, M. Noisternig, Crosstalk cancellation for spherical loudspeaker arrays, in *Fortschritte der Akustik - DAGA* (Dresden, 2008)
27. S.P. Lipshitz, J. Vanderkoy, In-phase crossover network design. *J. Audio Eng. Soc.* **34**(11), 889–894 (1986)
28. S. Lösler, F. Zotter, Comprehensive radial filter design for practical higher-order ambisonic recording, in *Fortschritte der Akustik – DAGA Nürnberg* (2015)
29. J. Linke, F. Wendt, F. Zotter, M. Frank, How masking affects auditory objects of beamformed sounds, in *Fortschritte der Akustik, DAGA* (Munich, 2018)
30. J. Linke, F. Wendt, F. Zotter, M. Frank, How the perception of moving sound beams is influenced by masking and reflector setup, in *VDT Tonmeistertagung* (Cologne, 2018)

**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.

