Principles and Considerations to Controllable Focused Sound Source Reproduction

Markus Guldenschuh, Alois Sontacchi, Franz Zotter, and Robert Höldrich

***Abstract*—Within this paper different common approaches are discussed which have the potential to establish a controllable sound field within a restricted area based on loudspeaker setups. Therefore the usage of headphones which is demanding over long time periods can be avoided. In the case of air traffic control at controller working positions this invention might improve the working process and therefore positively influence the safety conditions. It will be shown how the excitation of the diffuse sound field and the radiation towards neighbouring workspaces can be minimised. In order to enable an appropriate freedom of movement the usage of beam-steering methods and further kinds to produce audio sculptures are discussed. Simulation results outline the abilities and limitations for the case of nearfield beamforming.**

***Index Terms*—acoustic beamforming, sound field, focus**

I. Introduction

COMMUNICATION between aircraft pilots and controllers play one of the major roles within the air traffic management. Therefore reliability and security of the communication line are of greatest importance. Beside limited and degraded communication channels both communication participants have to handle with their tasks in rather loud environments. Noise levels around 75 dB A-weighted are common. These circumstances further emphasise bad working conditions. Caused by the traffic growth, the workload of the controller increases rapidly, too. That is why attention has to be paid to the controllers working conditions due to their indirect impact on the safety in the ATC.

An essential part of the work in an air traffic control tower is the radio communication to airplane cockpits. As the usage of headphones over long time periods is demanding, loudspeakers are used which cause crosstalk between neighbouring working positions, so that it becomes even harder to understand the strongly band limited radio traffic and increase the interfering ambient noise level. An approach to this problem is the usage of loudspeakers with a higher directivity, in order to minimise the excitation of the diffuse sound field and to reduce the radiation towards neighbouring workspaces. In order to provide the controllers at their working positions (CWP) an appropriate freedom and naturalness a sophisticated sound reproduction system is requested. This novel audio system should claim for an isolated sound reproduction at the different working positions, whereby it is able to track and follow the position of a defined controller at the CWP. Furthermore controller and co-controller should have the possibility to work within there own isolated communication zone but also should have access to a shared area where information exchange with the partner is possible without losing the contact to the own zone. The vision of two crossing so called"sound bubbles"is depicted in fig. 1. As above already stated, loudspeakers with high directivity might be able to fulfil the demand not to excite the diffuse sound field and to reduce the crosstalk to other working positions but it suffers form its inflexibility. A first improvement can be realized by the usage of a concave shaped acoustical mirror in addition to a proper located loudspeaker with high directivity. The loudspeaker will radiate the sound towards the mirror whereby the reflections will constructively interfere at a specific location in space (focus point) which is defined by the shape of the mirror and the geometrical arrangement of the system. This phenomenon is definitely nothing new and was at the latest used in the ancient Greece. More flexibility concerning the tracking of the controller position can be established by adding a number of static arranged loudspeakers in the vicinity of the mirror. By simple panning laws, feeding these loudspeakers with different distinct gains, the location of the focus point can be arbitrarily chosen within a specific area.

Fig. 1. Schematic depiction of so called "Sound bubbles"- Acoustically isolated/transparent working positions.

In the scientific audio community several audio rendering approaches have been developed due to the existing practicability of digital signal processing to establish synthesised sound fields with defined characteristics. Commonly approaches can be assigned to functional, physical equivalents or in between as hybrid realisations. In case of approaches based on physical equivalents they can further be differentiated between attempting to produce global or local reproduction areas. Furthermore in case of local reproduction –in the vicinity of the human ears –the reproduction via headphones (binaural) or even loudspeaker (transaural) can be distinguished.

Considering the global approaches, two somehow related

Manuscript received September 15, 2008. This work was supported in part by the Eurocontrol under Research Grant Scheme - Graz, (08-120918-C).

All authors are with the Institute of Electronic Music and Acoustics, University of Music and dramatic Arts Graz, 8010 AUSTRIA, (corresponding author to provide phone: M. Guldenschuh: ++43/316-389-5219; fax: ++43/316-389-3117; e-mail: guldenschuh@iem.at).

major strategies can be named. On the one hand Ambisonics which was first invented independently by Gerzon [1] and Cooper [2] was primarily developed to provide audio rendering even in the third dimension. It convinces with a compact notation and properties concerning the handling of a decomposed sound field based on a plane wave model. In [3] the theoretical framework up to arbitrary order (also called higher order Ambisonics, HOA) and the major properties of the thereby realisable sound fields are discussed. Nevertheless, HOA in its primary convention is not able to render sound sources at arbitrary distant locations. Therefore in [4] the first attempt has been suggested to overcome this drawback. An alternative approach based on a reformulation concerning the primary assumptions (spherical wave model) first invented in [5] and further revisited and theoretical examined in [6]. Based on further investigations in [7], [8] and [9] concerning arbitrary loudspeaker arrangements some limitations and restrictions stated in the primary invention can be negotiated. Therefore some more potential might be found within this sound field description approach.

On the other hand Wave Field Synthesis which was initially invented in the late 80's by Berkhout [10] and further developed at the TU Delft. The basic idea is related to the Huygens' Principle which states, that an arbitrary wave front may be considered as a secondary source distribution. Regarding the propagating wave from the given wave front we cannot differentiate if it was either emitted by the original sound source (the primary source) or by a secondary source distribution along this wave front. As a consequence, the secondary source distribution may be substituted for the primary source, in order to reproduce the primary sound field. Based on this physical background we can state that WFS aims at reproducing sound waves by (distributed) loudspeaker arrays. In [11] an excellent review concerning derivation, system description and properties can be found.

Considering approaches providing local reproduction via loudspeaker the invention of the stereo-dipole [12] has to be named. Binaural recorded or pre-processed audio material is emitted by two loudspeakers. In order to transmit the signal of the left loudspeaker only to the left ear and the signal of the right loudspeaker to the right ear the crosstalk of the left speaker to the right ear and the right speaker to the left ear has to be removed. The process of crosstalk cancellation is resolved by an inversion of the transfer path matrix which is determined through the geometrical arrangement of loudspeakers and listening position. Therefore modelling or even better measurements of the required transfer paths are required. Utilising only two loudspeakers will lead to an ill-conditioned transfer path matrix, in general. Improvements can be achieved by frequency dependent signal distributions over various loudspeakers (cf. [13]). Beside that, in [14], the possibility of enabling a higher listener mobility based on adaptive systems have been investigated.

Summarising the advantages of the above described global approaches it can be stated, that caused by the analytical description alterations of the sound source position can be easily adapted. Regarding the disadvantages of these approaches practical implementations are tied with huge technical equipment cost and high computation load. In the case of local sound field reproduction approaches the hardware requirements can be reduced dramatically. Caused by the non-existing analytical description the cost raises when arbitrary tracked listing positions should be taken into account.

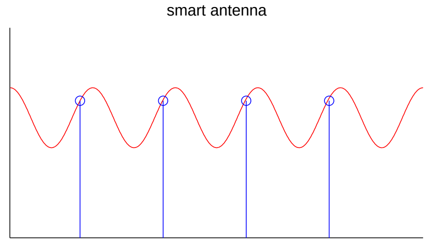

Fig. 2. An array of antennas samples a wave in the distance of a wavelength.

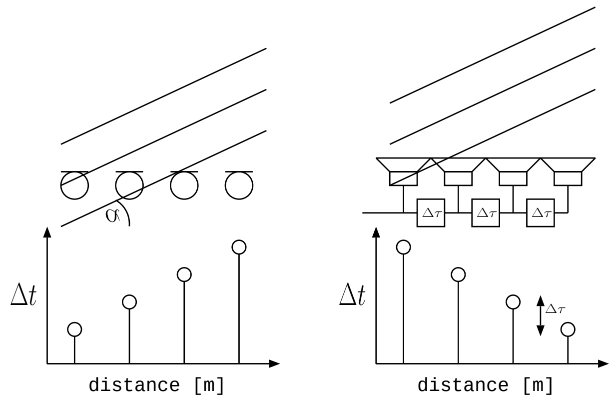

Fig. 3. Recording and reproduction of an incident wavefront.

This article explores realisation capabilities of practical implementations by examining simulations of computer models. Within the next section beamforming methods are explored and section III will discuss the integration of crosstalk cancellation. The performance of beamforming will be examined in IV and section V will provide an outlook to the time reversal mirror (TRM) before the methods are resumed in the Conclusion.

II. Beamforming

A. Far field beamforming

Beamforming was initially used in telecommunication techniques for smart antennas. If several antennas are in a line in the direction of the incident wave with an interelementary distance of a multiple of the wavelength, the arriving wave is in phase on every antenna. (See fig. 2.) Thus the signal can be amplified by constructive addition. The same principle is applied in acoustics, for either reception with a microphone array or transmission with a loudspeaker array. Fig. 3 illustrates acoustical far field beamforming which emphasises plane waves in certain direction.

The microphone array in fig. 3 records the plane wave, for which the time delay of arrival corresponds to the angle of incidence α. For a unique assignment of α, the spatial

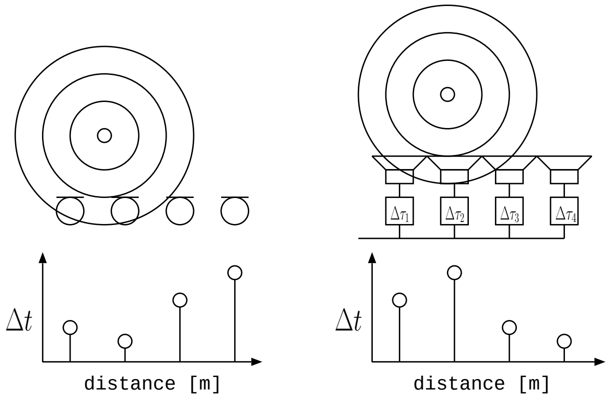

Fig. 4. Recording and reproduction of a point source. Compared to far field beamforming the delay times are not constant anymore.

sampling theorem has to be fulfilled

$$\Delta x \leq \frac{\lambda}{2\sin(\alpha_{max})}, \tag{1}$$

wherein Δx is the distance between the elements (microphones or loudspeakers) and λ the wavelength. In an analogue way, a plane wave can be produced with a delay line between the loudspeakers. However, in the case of a sound field concentration in a desktop environment, the generation of a focused wave is desirable. Therefore we consider near field beamforming in the next section.

B. Near field beamforming

Near field beamforming aims to focus the sound field somewhere in the vicinity of the loudspeaker array. It uses geometrical properties (i.e. the distance from loudspeaker to focus point) to calculate the delay time of the loudspeaker signals, such that the waves coincide in phase in the focus point. This means that the loudspeakers can be driven with a rather low signal because, in the focus point, all the signals will be constructively added and therefore amplified acoustically. As a consequence, the listener who is sitting in the focus point hears an appropriate loud signal, while the sound pressure in the rest of the area is low. Hence it fulfills the requirements we stated in the introduction. Fig. 4 shows the main differences to far field beamforming. The delay lines have to be independent for each loudspeaker channel and there is no unique incident angle anymore. As a consequence, eq. (1) does not apply. The spatial resolution now depends on the distance from the target point to the array, too. To guaranty aliasing prevention eq. (1) can be generalised to the worst case

$$\Delta x \leq \frac{\lambda}{2}. \tag{2}$$

The bandwidth in ATC reaches from 300 to 2500 Hz. As a consequence, the distance between the array elements must not be larger than 7 cm to prevent spatial aliasing. Fig. 5 shows a directivity diagram of a linear array. The actual beam that is directed to the focus point is denoted as main lobe, whereas spatial aliasing appears in the form of strong side lobes. For a better understanding of the relation between the array properties, directivity patterns and spatial aliasing, it is worth to throw a glance at acoustics and signal theory.

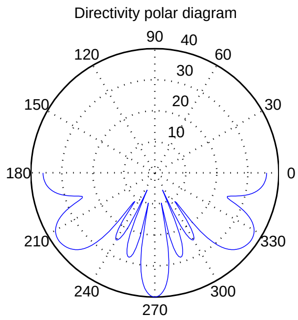

Fig. 5. Directivity polar plot with the main lobe in the middle, and 4 strong side lobes at 330 and 210 degree and at the x-axis, respectively.

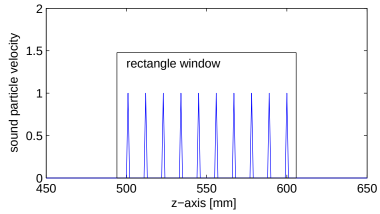

Fig. 6. The sound particle velocity distribution of a line array on the z axis. There is only sound particle velocity at the discrete positions of the loudspeakers. The array has finite length. Hence the pulse train is limited through a rectangle window.

The radiation of a linear array can be described by a cylindrical wave expansion in the far field. We express an infinite equispaced linear array as a particle velocity distribution $\nu(z)$ along a cylinder in the form of a pulse train [15]. The Fourier transform of $\nu(z)$ yields the k space spectrum $\Psi(k_z)$ [15].

$$\Psi(k_z) = \int \nu(z) e^{-ik_z z} dz, \tag{3}$$

with k_z defined in eq. (4). The array itself, of course, is not infinite. Therefore the pulse train has to be multiplied with a window. A pulse train is not modified through a Fourier transform, but a multiplication with a rectangle window corresponds to a convolution with a sinc function in k space. The windowed pulse train is shown in fig. 6 and the result of the convolution is depicted in fig. 7.

The wavenumber in z direction is given as [15]

$$k_z = \frac{\omega}{c} \cos(\alpha), \tag{4}$$

hence $\Psi(k_z)$ is is a measure of directivity. The scaling of the k_z-axis depends on the angular frequency ω. The higher the frequency, the higher is the angular resolution on the k_z-axis. For evaluating the directivity, we are only interested in the angles form $-\pi$ to π or from -90 degree to 90 degree, respectively. A finer resolution therefore leads to a larger evaluation window. The main and side lobes of the sinc function correspond to the main and side lobes in the directivity patterns. If we increase the distance between the loudspeakers, the pulse train in the k space gets narrower, which means that the sinc functions move closer together.

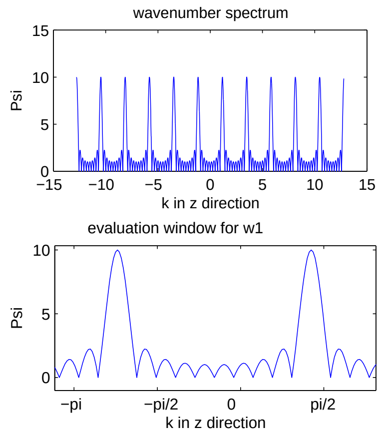

Fig. 7. The wavenumber spectrum of a finite array. The k axis has to be scaled for a certain angular frequency ω_1. In this scaling the wavenumber spectrum can be evaluated from $-\pi/2$ to $\pi/2$.

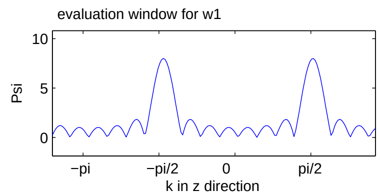

Fig. 8. Increasing the loudspeaker distances makes the sinc functions move together. Aliasing occurs if two main lobes appear in one evaluation window ($-\pi$ to π); which is the case in this figure.

Aliasing occurs if more than one main lobe appears in the evaluation window. (See fig. 8.)

The delay times between the array elements correspond to a multiplication of the pulse train with a linear phase term e^{izk_z}. With this multiplication, the directivity pattern is modulated towards the target direction.

The knowledge of windowing properties helps to avoid strong side lobes. Instead of a rectangle window, the array can be truncated by a smoother window that piecewise reduces the amplification of the loudspeakers. Fig. 9 shows the result of a triangle window in the k space. The triangle window suppresses the side lobes, but widens up the main lobe. Thus

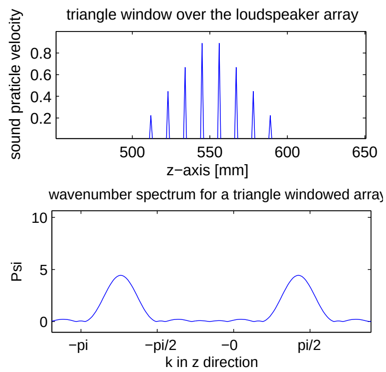

Fig. 9. A triangle window over the array and its result in the k space. The sidelobes are completely compressed, but therefore the main lobes are widened up.

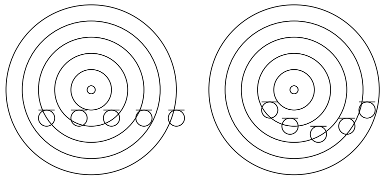

Fig. 10. Comparison of a linear and a circular array in terms of spatial sampling of a spherical wave.

the decision for a certain window is always a tradeoff between a sharp focus and reduced side lobes.

If we recall eq. (1) we notice that a higher maximal frequency is allowed if the plane wave fronts are parallel to the line array (i.e. if $\alpha = 0$). The analogue case for a spherical wave would be a semicircular array with the centre at the point source location. Fig. 10 shows that a circular array samples a spherical wave indeed much better than a linear one. As a consequence we consider a curved loudspeaker array to be more adequate to produce a focused sound field. Fig. 11 compares the sound field of a linear array with a curved one. Both simulations were done for 4500 Hz with an array-length of one meter and 15 array elements. The figure shows two main advantages of the curved array over the linear one. First, the pressure of the sound field behind the focus point (i.e. between the array and the focus) is much lower than in the case of the linear array. Secondly, the two side lobes moved out of the region of interest.

However, moving the focus outside of the aperture yields poor results: A wide main lobe and several side lobe pointing towards the centre. These effect are shown in fig. 12. Please note, that in the following all simulation results and discussions are made for a circular shaped array.

The strength of the side lobe can be decreased, by attenuating the loudspeaker signal at the side lobe side. In other words, the window over the array is shifted such that it follows the focus. We call this technique *sliding aperture*. Fig. 13 shows the results of a triangle window. As discussed above, the smooth Bartlett window suppresses side lobes but widens the main lobe.

The problem of the widened main lobe can only be solved by increasing the aperture of the array, i.e. increasing its radius. This of course leads to a higher number of required loudspeakers if the sampling theorem should be fulfilled. At this point, also elliptical arrays should be considered as tradeoff between a linear and a circular array.

The principle of near field beamforming is that the waves superpose constructively in the focus point. Therefore it is evident that the sound pressure in the focus point increases with the number of loudspeakers. Hence the sharpness of the focus depends on the number of loudspeakers. The width of the focus however also depends on the wavelength. The focus of a low frequency signal will be much broader, than the one of a high frequency signal. To enable a constant focus width throughout the whole bandwidth, the array has to be windowed in dependence of the frequency. Whereas a smaller number of loudspeakers will be used for higher frequencies.

One demand for an ATC communication setup is that

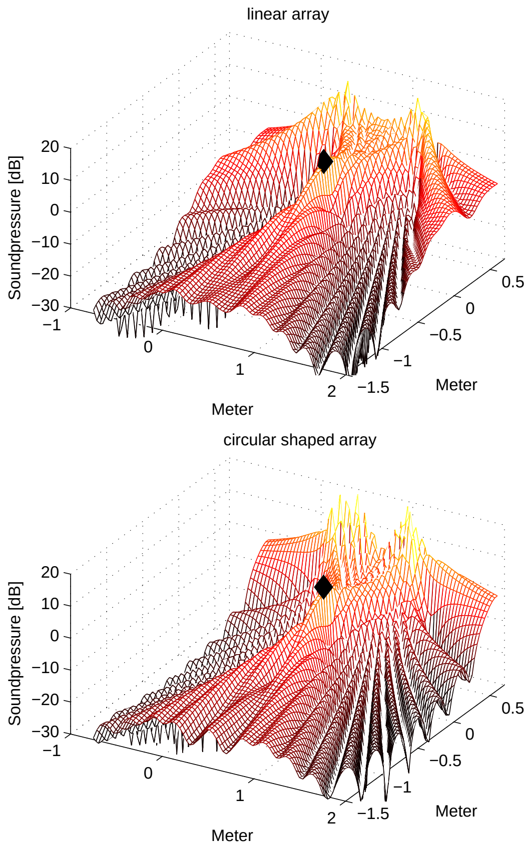

Fig. 11. Soundfield reproduction by a linear and a circular array with the focus on the black dot. There are strong side lobes in the linear case that do not appear for the circular shaped array.

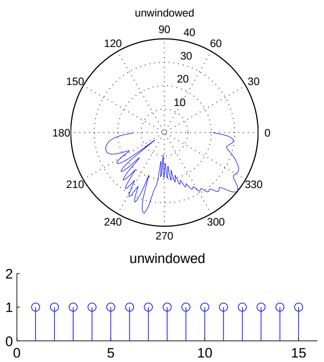

Fig. 12. Polar plot for a circular shaped array with the focus point outside of the aperture. Two consequences can be seen: 1. The side lobes moved into the centre. 2. The main lobe got very broad.

the controllers are free to move. The great advantage of the beamforming approach is its analytical expression that only bases on geometrical properties. Therefore it is very easy to enable an adaptive system. Common head tracking technologies enable the determination of the position and the orientation of the users. With this data a new focus position can be updated very fast. E.g. for a tracking system via webcam the only mobility restrictions are given by the length of the array and the focal length of the camera's lense. A scheme of these restrictions is given in fig. 14.

Until now, we only considered one beam that should be

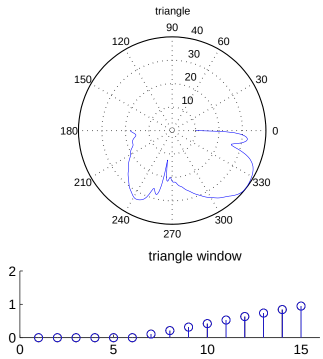

Fig. 13. The effect of a sliding aperture. The side lobes are reduced.

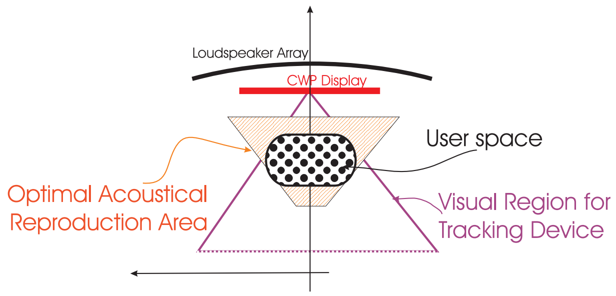

Fig. 14. For an array length of 1.4 m and a wide angle camera, a user space of approximately 1.4 x 1.5 m can be drawn.

directed to the listener such that the focus point coincides with the listeners head. This would lead to the emersion of a mono source in the head of the listener, which of course will not cause a very natural impression. Producing two focus points in the vicinity of of the listener's ears will enable the synthesis of a spatialised sound. As it can be read in the following section any source direction can be encoded onto a mono source if two beams are produced.

III. Transaural stereo

Spatialised hearing works above all over interaural time differences (ITDs) and interaural intensity differences (IIDs). But also diffraction and shadowing caused by the head and reflections from the pinna play an important part; especially when it comes to front back or elevational distinctions. In general the acoustical free-field path from a sound source to the ears is described by the head related transfer functions (HRTFs). In the first line, they are dependent on the incident angles of the sound source. If near field HRTFs are concerned, also the radius has to be considered.

Convolving a mono source with a pair of HRTFs[1] for a certain direction produces a binaural signal. In other words the HRTFs encode the mono source with directional information. Besides it is possible to keep the virtual sound source at fixed position, even if the user is turning his head. In [16] it is well described how Ambisonics encoding can be used to compensate head rotations without the need of large HRTF data bases.

[1]HRTFs for the right and the left ear, respectively.

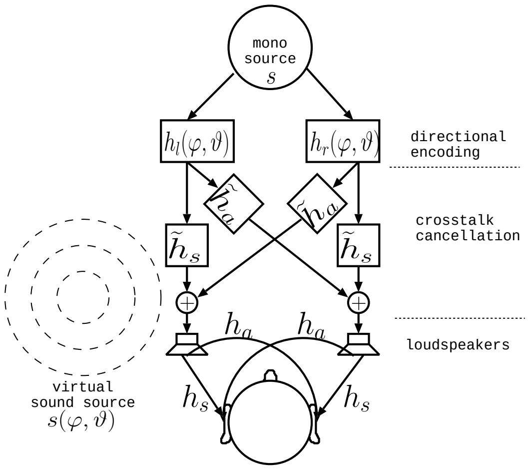

Fig. 15. Block diagram of encoding a mono source to a virtual spatialised sound source with crosstalk cancellation. $\tilde{h}_a = -h_a/(h_s^2 - h_a^2)$ and $\tilde{h}_s = h_s/(h_s^2 - h_a^2)$.

Binaural signals are stereo signal, whereas the left channel must be transmitted to the left ear only and the right channel to the right ear. The ideal transmitter for such a signal is of course a headphone as it prevents crosstalk between the channels. But as we stated in the introduction, the usage of headphones is hardly acceptable for a whole working day. Thus we propose the usage of crosstalk filters to still be able to generate spatialised audio.

Crosstalk cancellation bases on the knowledge of the HRTFs from a loudspeaker to left and right ear of the listeners. Fig.15 shows crosstalk cancellation for a simple case. The depicted lattice filter compensates for the crosstalk pathes h_a as well as for the direct pathes h_s. Its simplicity lies in the symmetry of the layout and in the limitation to two loudspeakers. For the beamforming approach however, several loudspeakers have to be considered. In [17] one of the authors proposes an adaptive crosstalk cancellation for the usage with loudspeaker arrays.

IV. Simulations of Near Field Beamforming with Multiple Foci

In section II-B we declared that the distance between the loudspeakers should not exceed 7 cm to prevent spatial aliasing up to 2500 Hz. For an array-length of 2 meter this would lead to 28 array elements. Simulations in (fig. 16) show excellent results with only 23 loudspeakers, if the windowing concepts introduced in sec. II-B are applied. This is possible because the incident angle is far from being 0 degrees. Thus it is not necessary to calculate the worst case, like it is proposed in eq. (2).

From the spatial aliasing point of view it is therefore no problem to provide the listeners with two beams. Fig. 17 shows four beams, whereas the array was truncated with a rectangle window. It can be seen, that their main lobes are very narrow and the side lobes are still 10 dB under the main lobes.

The above described simulation were done for a frequency of 1000 Hz, which (in usual room conditions) corresponds to a wavelength of 34 cm. Please not, that the physical limits of the focus width equals $\frac{\lambda}{2}$. The lower cut off frequency for ATC

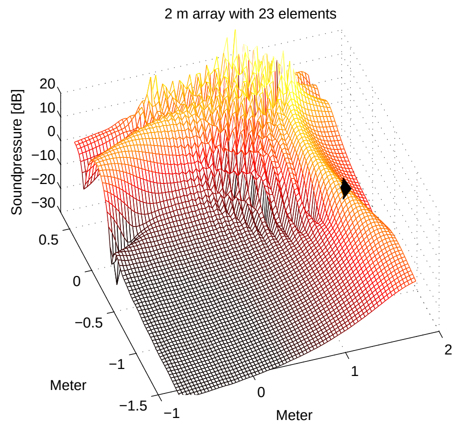

Fig. 16. A 23 elements array for a 2500 Hz beam. The only side lobe appears in the back where it does not influence the region of interest.

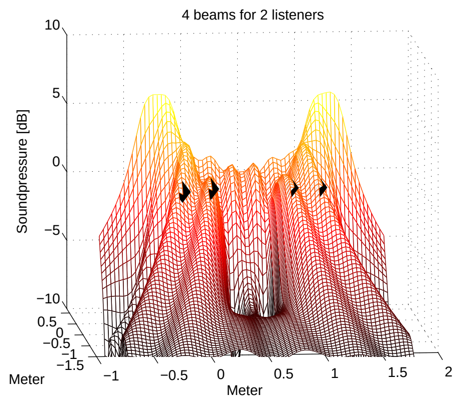

Fig. 17. 4 beams of 1000 Hz to the ears of 2 listeners. The ear positions are denoted as black dots. It can be seen, that the 4 beams are very well distinguishable. Also, there are clearly two separate sound bubbles.

communication bandwidth is 300 Hz. This frequency has a wavelength of more than 1 meter. Fig. 18 shows that the two beams of these frequency are not distinguishable any more. Regarding one beam alone in fig. 19 however shows that it does not reach the second user as long as both users are in a distance of more than half a meter from each other. The sound pressure level (SPL) decay at the position of the second users

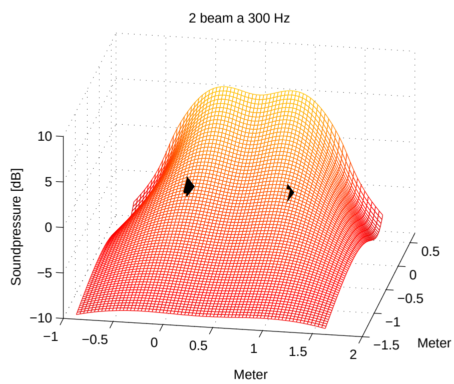

Fig. 18. 2 beams of 300 Hz have a focus of more than half a meter and occupy the whole sound field.

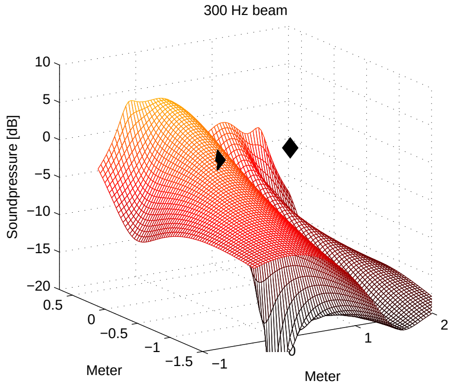

Fig. 19. Here, one 300 Hz beam is shown. The beam does not reach the second user (denoted as black dot) as he/she is in a distance of more than a half meter.

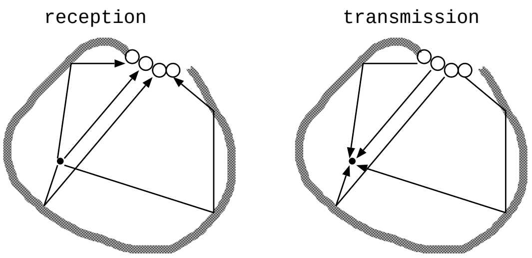

Fig. 20. Reverberation cavity closed by microphones (receive mode) or loudspeakers (transmit mode).Adapted from [18]

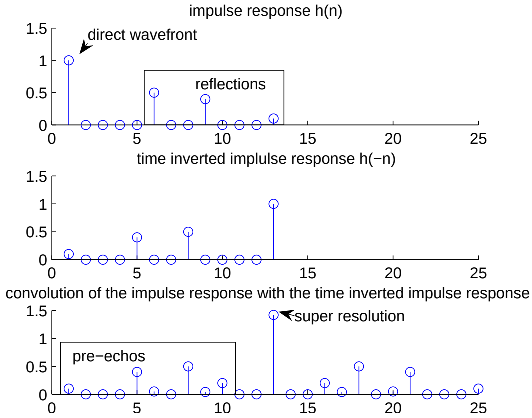

Fig. 21. Convolution of a schematic impulse response with its time reversal.

is -10 dB. Hence the low frequency beams will not disturb the neighbouring sound bubble, but it will not be possible to generate two beams for two user in the low frequency range. Thus the concept of transaural stereo, introduced in the previous section, can only be applied for high frequencies.

Our simulation program also allows to calculate the directivity and hence the distance factor γ of the beamformer. In general γ increases with the frequency. However if spatial aliasing occurs, this tendency is compensated. For a simulation of a 16 element array, γ varies between 3 and 5. Knowing γ allows to calculate the sound pressure of the diffuse field which is excited by the beam.

$$L_r = L_d - 10\log P_0 - 103\,\mathrm{dB} - 10\log\gamma + 20\log r - 10\log A \qquad (5)$$

For a sound pressure level $L_d = 60\,\mathrm{dB}$ at the focus point in a distance of $r = 1\,\mathrm{m}$, $\gamma = 4$, and with a surrounding reflecting surface of $A = 100\,\mathrm{m}^2$, the sound pressure level of the diffuse field is only 51 dB. Further 4 dB can be gained easily, if absorbing material is put behind the user (considering that the listener himself will also absorb sound energy).

V. Time reversal mirror

This section will introduce an alternative approaches to sound source focussing. The principle of the time reversal mirror can be explained over a reverberating cavity that, in a first step, is closed by an array of microphones like it is shown in fig. 20. At some point in this cavity a pulsive sound is emitted. If the signals recorded by the microphones are played back time inverted[2], the sound propagates back and focuses on the initial emitter point. Mathematically expressed, we have L recorded impulse responses $h_l(t)$ and their time inverted $h_l(-t)$.

In acoustics, the properties of the propagation path from source to sink are the same as from sink to source. Thus, if the microphones are exchanged for loudspeakers, the way back to the focus point is as well described by $h_l(t)$. Evaluating the played back time inverted signals $h_l(-t)$ in the focus point, results in a convolution with the original impulse responses:

$$s_l(t) = h_l(-t) * h_l(t) \qquad (6)$$

In order to focus any signal $x(t)$ in the given point, it has to be convolved with $h_l(-t)$ of the corresponding loudspeaker. Thus the loudspeaker signals $y_l(t)$ are calculated as

$$y_l(t) = x(t) * h_l(-t), \qquad (7)$$

and the signals in the focus point follow to

$$s_l(t) = x(t) * h_l(-t) * h_l(t). \qquad (8)$$

The frequency pendant of equation (8) is

$$S_l(\omega) = X(\omega) H_l^*(\omega) H_l(\omega), \qquad (9)$$

where $H_l^*(\omega)$ is the conjugate transposed of $H_l(\omega)$ and their product is called the time reversal operator. It can be seen, that the input signal $X(\omega)$ is filtered with the squared norm of the frequency response of the path from microphone to loudspeaker, including room characteristics as well as the characteristics of the electroacoustical transducers. To avoid this filtering, the time reversal operator should be equalised with $\frac{1}{|H_l(\omega)|^2}$. However, this equalisation is dangerous, if $|H_l(\omega)|^2$ contains values close to zero.

Impulse responses in a reverberating cavity as well as in ATC workplaces will not only contain the direct wavefront but also some reflexions over walls and desktops. A schematic impulse response with a direct wavefront and reflection parts is drawn in figure 21. These reflexions will appear as pre-echos, when played back time inverted. On the other hand,

[2]Whereby the microphones are replaced by loudspeakers with identical spatial properties. (I.e. the spatial sensitivity of the microphones corresponds with the spatial radiation pattern of the loudspeakers.)

after having run through the reflection pathes again,[3] they are also constructively added with the directional part. This results in a main impulse that is even stronger than the original first wavefront has been. This effect is called super resolution. Summing up, this means that reflections help to focus the sound into a specific point, but they also cause pre-echos that are disturbing especially for plosive sounds [18]. The occurrence of pre-echos and their audibility depends on the loudspeaker arrangement, the room size and its acoustical properties, further on the reproduced focus point and on the position of the listener in this situation. The impact of ATC desktops on the audibility of pre-echos will be a topic for further investigations.

The quality of the TRM depends on the reflection properties of the surrounding and, similarly to the beamforming approach, on the number of loudspeakers. It also profits from the bandwidth of the signal, as a higher bandwidth increases the decorrelation outside of the focus point.

Time reversal mirroring can also produce multiple focus points, if the impulse response from each focus point to each loudspeaker is known. An adaptive system that is able to follow a moving user needs to know all the impulse responses of the area in which the user can move. According to eq. (2) the maximal distance between the impulse response measurement points for a bandwidth up to 2500 Hz is 7 cm. In an area of one square meter approximately 210 impulse responses would be needed to determine.

VI. Conclusion

Simulation results have shown that near field beamforming is very well suited to produce a focused sound field. Only one consideration has to be taken, which is to prevent strong side lobes. This can be worked out by choosing a useful array shape and appropriate windowing techniques. It is furthermore no problem to generate several distinct focus points, keeping in mind the physical limitations given by the half wavelength of the played back signals.

For a further enhancement of the reproduced sound, we proposed the usage of the transaural stereo approach. The mono sources can be binaural encoded such that they contain directional information. This technique demands two separate beams for each ear of the listeners and can thus only be used for high frequencies, where the beams are narrow enough.

Finally an alternative sound focussing techniques was discussed. TRM makes use of reverberation to naturally increase the soundpressure in the focus point, whereas echos will disturb the sound field in the beamforming approach. The big disadvantage of this method is that transfer functions from every possible user position to the loudspeakers have to be known. Finding an analytical expression for such a transfer function matrix will also be a field of future investigations.

References

[1] M. Gerzon, “Periphony: With-height sound reproduction,” *JAES*, vol. 21, no. 1, January/February 1973, (first presented at the 2nd AES Convention of the Central Europe Section, Munich, Germany, March, 1972).

[2] G. Cooper, “Tetrahedral ambiophony,” *Studio Sound*, June 1970.

[3] M. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, November, 2005.

[4] A. Sontacchi and R. Höldrich, “Further investigations on 3d sound fields using distance coding,” *DAFx01 Proceedings, Limerick, Ireland*, 2001.

[5] J. Daniel, “Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format,” *In 23rd Int. AES Conf., Copenhagen, Denmark*, May 23-25, 2003.

[6] J. Ahrens and S. Spors, “Focusing of virtual sources in higher order ambisonics,” *124th AES Convention, Amsterdam, The Netherlands*, May 17-20, 2008.

[7] A. Laborie, R. Bruno, and S. Montoya, “Reproducing multi channel sound on any speaker layout,” *118th AES Convention, Barcelona, Spain*, May 28-31, 2005.

[8] J. Hannemann and K. Donohue, “Virtual sound source rendering using a multipole-expansion and method-of-moments approach,” *J. Audio Eng. Soc.*, vol. 56, no. 6, June, 2008.

[9] A. Sontacchi, F. Zotter, and R. Höldrich, “3d sound field rendering under non-idealized loudspeaker arrangements,” *Acoustics 08, Paris, France*, June 29-July 4, 2008. [Online]. Available: pdf-Fileavailableathttp://iem.at/projekte/publications/paper/hoared

[10] A. Berkhout, “A holographic approach to acoustic control,” *J. Audio Eng. Soc.*, pp. 977–995, 1988.

[11] S. Spors, R. Rabenstein, and J. Ahrens, “The theory of wave field synthesis revisited,” *Audio Eng. Soc. Conv. Paper*, vol. 124th Conv. Amsterdam, May 2008.

[12] O. Kirkeby, P. Nelson, and H. Hamada, “The ‘stereo dipole’–binaural sound reproduction using two closely spaced loudspeakers,” *102nd AES Convention, Preprint 4463(I6)*, 1997.

[13] T. Takeuchi and P. Nelson, “Optimal source distribution system for virtual acoustic imaging,” *AES 110th Convention, Amsterdam, The Netherlands*, May 12-15, 2001.

[14] W. G. Gardner, “Head tracked 3-d audio using loudspeakers,” *1997 IEEE ASSP Workshop*, October 1997.

[15] E. Williams, “Fourier acoustics,” *London, Academic Press*, 1998.

[16] M. Noisternig, A. Sontacchi, T. Musil, and R. Höldrich, “A 3d ambisonic absed binaural sound reproduction system,” *AES 24th Int. Conf. on Multichannel Audio*, 2003.

[17] A. Sontacchi and T. Musil, “Demonstrator for a controllable focused sound source reproduction,” *submitted to 7th Eurocontrol Innovative Research Workshop and Exhibition*, Dez. 2008.

[18] S. Yon, M. Tanter, and M. Fink, “Sond focussing in rooms i: The time-reversal approach,” *J. Acoustical Society of America*, vol. 113, no. 3, pp. 1533–1543, March 2003.

[3] Which is being convolved with the original impulse response.