Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

The use of parametric arrays for transaural applications

Raphaël Greff

A-Volute, 1 rue du transit 59650 Villeneuve d'Ascq, France

PACS: 43.66.Pn

ABSTRACT

The cancellation of transaural acoustic crosstalk is an essential and critical feature of all virtual auditory displays based on HRTF (Head-Related Transfer Functions) when loudspeakers are used for listening. Generally, a satisfying crosstalk cancellation is achieved within an extremely small sweet spot. Reducing the angle between the loudspeakers with respect to the listener increases the controlled area in a high frequency range, especially on the front-back axis, but makes harder the cancellation and equalization at lower frequencies, the amount of energy required to achieve the cancellation in this frequency range being prohibitively high. The directivity of employed loudspeakers has a direct impact on the transaural acoustic crosstalk. The narrower the directivity is, the lower the crosstalk level should be. An alternative to the use of a signal processing step to cancel the crosstalk would be to use highly directional loudspeakers to physically reduce the levels of indirect path responses, *i.e.* from each loudspeaker to the corresponding contralateral ear, compared to those of direct ones, *i.e.* from each loudspeaker to the corresponding ipsilateral ear. Devices known as parametric arrays employ the nonlinearity of the air to create audible sound from inaudible ultrasound, resulting in an extremely directive, beamlike wide-band acoustical source. This paper investigates the potential use of a pair of parametric arrays for HRTF-based transaural applications.

INTRODUCTION

The following paper presents a theoretic investigation of the use of parametric arrays for transaural audio applications. The basic idea behind this study is to see if the highly directive audible sound beam generated by such a device may illuminate one ear only, or at least may induce by nature a sufficiently low-level transaural acoustic crosstalk, to overcome the need of an additional crosstalk cancellation processing .

We will first review the basic principles and practical implementations of transaural audio and the parametric array; then, we will focus on the evaluation of the transaural crosstalk reduction that we may expect from the use of parametric array loudspeakers instead of conventional loudspeakers.

TRANSAURAL AUDIO

Basic principle

Transaural audio is a method used to deliver binaural signals to the ears of a listener using stereo loudspeakers. The basic principle of this technique consists in filtering the binaural signal in such a way that the subsequent stereo presentation, *i.e.* the transaural signal played back over the pair of loudspeakers, produces the original binaural signal at the listener's ears. The technique was first put into practice by Schroeder and Atal [1] and later generalized by Cooper and Bauck [2, 3]. The transaural processing can be decomposed into a binaural processing step and a crosstalk cancellation step, but it is often reduced to this latter processing.

In general, a stereo (loudspeaker-based) listening situation may be described by the schematic representation of Figure 1, where x_L and x_R are the signals feeding the loudspeakers, and y_L and y_R are the signals at the ears of the listener.

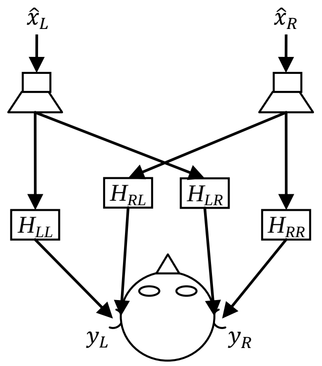

Figure 1. Transfer functions from loudspeakers to ears in a stereo listening situation.

The whole system can be described by the following matrix equation:

$$y = H\hat{x} \tag{1}$$

with:

$$y = \begin{bmatrix} y_L \\ y_R \end{bmatrix}, \mathrm{H} = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix}, \hat{x} = \begin{bmatrix} \hat{x}_L \\ \hat{x}_R \end{bmatrix} \tag{2}$$

H_{LL}, H_{RL}, H_{LR} and H_{RR} are the HRTF representative of the listening situation, and describe the four acoustical paths, from each loudspeaker to each ear. If x is the binaural signal that we wish to deliver to the listener's ears, the system transfer matrix must be inverted to obtain the correct transaural signal:

$$\hat{x} = \mathrm{H}^{-1}x \tag{3}$$

with:

$$\mathrm{H}^{-1} = \frac{1}{H_{LL}H_{RR} - H_{LR}H_{RL}} \begin{bmatrix} H_{RR} & -H_{RL} \\ -H_{LR} & H_{LL} \end{bmatrix} \quad (4)$$

This expression of the system inversion leads to the general form of the transaural crosstalk cancellation processing (see Figure 2).

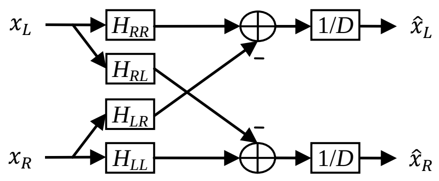

Figure 2. General form of the transaural crosstalk canceller, where $D = H_{LL}\, H_{RR} - H_{LR}\, H_{RL}$.

Practical implementation

Conventional listening situation may be assumed as exhibiting of left-right symmetry (symmetrical arrangement of the loudspeakers and symmetrical morphology of the listener). Under this assumption, the transaural filters can be specified in terms of ipsilateral ($H_i = H_{LL} = H_{RR}$) and contralateral ($H_c = H_{LR} = H_{RL}$) transfer functions:

$$\mathrm{H}^{-1} = \frac{1}{{H_i}^2 - {H_c}^2} \begin{bmatrix} H_i & -H_c \\ -H_c & H_i \end{bmatrix} \quad (5)$$

In [2], Cooper and Bauck proposed an efficient "shuffler" implementation of the crosstalk canceller assuming symmetrical listening conditions (see Figure 3). This implementation involves computing the sum and difference of the left and right binaural signals, filtering them, and forming the sum and difference of the resulting filtered signals. These sum and difference operation are accomplished by the unitary shuffler matrix U. The system inversion may be expressed in shuffler form as following:

$$\mathrm{H}^{-1} = \mathrm{U} \begin{bmatrix} \frac{1}{H_i + H_c} & 0 \\ 0 & \frac{1}{H_i - H_c} \end{bmatrix} \mathrm{U}^{-1} \quad (6)$$

with:

$$\mathrm{U} = \mathrm{U}^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (7)$$

One of the most important observations they made is that the derived sum ($H_i + H_c$) and difference (H_i - H_c) filters have almost the same excess phase characteristics. Since this phase affects both sum and difference signals equally, the crosstalk cancellation processing can be reduced to a simple minimum-phase equalization of the sum and difference signals.

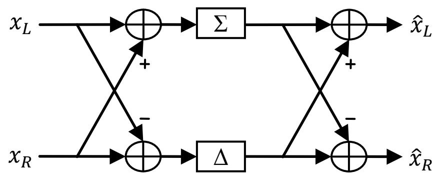

Figure 3. Shuffler form of the transaural crosstalk canceller, where $\Sigma = H_i + H_c$ and $\Delta = H_i$ - H_c.

To overcome the problem of morphological dependencies of the HRTF, and to avoid the need of individualized transaural filters, the angle between the loudspeakers may be reduced. When the angle decreases, the ratio between the frequency response of the sum and difference transaural filters becomes less dependent on listener's morphology. The use of an angle spacing of less than 10° leads to the commonly named "stereo dipole" configuration. The problem of such a loudspeaker arrangement is that the equalization of the difference signal requires a prohibitive amount of energy in a relative broad bass frequency range.

PARAMETRIC ARRAY LOUDSPEAKERS

Basic principle

A parametric array loudspeaker generates a high intensity ultrasonic beam to obtain audible sound waves, having the property of being highly directive, like the ultrasonic beam is. Such a device exploits an effect known as self-demodulation to create extremely directive low frequency sounds which would otherwise require an enormous array of audible frequency transducers.

Self-demodulation occurs when nonlinearities of a compressible medium cause high frequency wave components to interact. This interaction produces a secondary wave whose frequency components correspond to the sums and the differences of the individual frequency components of the primary wave.

This phenomenon was first analyzed by Westervelt [4], and is well known as "nonlinear interaction of sound waves" or "scattering of sound by sound". Based on Lighthill's arbitrary fluid motion equation:

$$\frac{\partial^2 \rho}{\partial t^2} - c_0^2 \nabla^2 \rho = \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j} \quad (8)$$

where ρ is the fluid density, T_{ij} the stress tensor and c_0 is the small signal sound velocity – he derived an inhomogeneous wave equation which is satisfied by the sound pressure of the secondary wave produced by the nonlinear interaction:

$$\nabla^2 p_s - \frac{1}{c_0^2} \frac{\partial^2 p_s}{\partial t^2} = -\rho_0 \frac{\partial q}{\partial t} \quad (9)$$

with:

$$q = \frac{\beta}{\rho_0^2 c_0^4} \frac{\partial p_i^2}{\partial t} \quad (10)$$

where p_s is the secondary wave pressure, p_i the primary wave pressure, β the nonlinear fluid parameter and ρ_0 the small signal fluid density. The solution of this equation may be expressed by the superposition integral of the Green's function and the virtual second source – right side term of equation (9):

$$p_s = \frac{\rho_0}{4\pi} \iiint_v \frac{1}{|r - r'|} \frac{\partial q}{\partial t}\left(r, t - \frac{|r - r'|}{c_0}\right) dr' \quad (11)$$

where r is the observation point position vector, r' the source position vector and v the nonlinear interaction space.

Practical implementation

The self-demodulation induced by the nonlinear interaction of the ultrasonic waves produces new frequency components at the combination of sums and differences of their individual frequency components, in a process akin to AM demodulation.

According to the work of Yoneyama *et al*. [7], which referred to studies of Westervelt [4], Berktay [5], and Muir and Villette [6] on the parametric array for underwater sonar applications, parametric array loudspeakers generally use the AM-modulation scheme for generating the driving signal. Berktay stated that a collimated and plane primary wave consisting of an AM-modulated wave of pressure p_i given by:

$$p_i = p_0 E\left(t - \frac{r}{c_0}\right) e^{-\alpha r} \sin\left(\omega_0 \left(t - \frac{r}{c_0}\right)\right) \qquad (12)$$

where p_0 is the amplitude of the primary ultrasonic beam pressure, E the modulation envelope composed by audible frequency components, ω_0 the ultrasonic carrier angular frequency, α the medium attenuation factor – will demodulate, creating an audible virtual source in the primary ultrasonic beam given by:

$$q = \frac{\beta p_0^2}{2\rho_0^2 c_0^4} e^{-2\alpha} \frac{\partial E^2}{\partial t}\left(t - \frac{r}{c_0}\right) \qquad (13)$$

Assuming the primary ultrasonic beam cross section area is S, the audible secondary sound wave p_s can be expressed as following:

$$p_s = \frac{\beta p_0^2 S}{16\pi \rho_0 c_0^4 \alpha r} \frac{\partial^2 E^2}{\partial t^2}\left(t - \frac{r}{c_0}\right) \qquad (14)$$

DSBWC AM modulation scheme

Berktay's analysis leads directly to the conventional DSBWC (Double Side Band With Carrier) modulation technique presented by Yoneyama. Using the driving signal s_{in} given by:

$$s_{in}^{DSB} = \left(1 + mg(t)\right) \sin(\omega_0 t) \qquad (15)$$

where g is the considered audio signal to be diffused by the parametric array loudspeaker, and m the modulation depth – the resulting audible signal s_{out} is given by:

$$s_{out}^{DSB} = Kh(t) * \left(mg(t) + \frac{1}{2} m^2 g^2(t)\right) \qquad (16)$$

with K a frequency-independent gain (related to the properties of the parametric array, the properties of the air and the on-axis distance of the observation point from the array), and h a 12db/octave slope filter.

The major problem of DSBWC is the relative high distortion rate: The desired demodulated signal and the undesired quadratic nonlinearity are proportional to m and m^2 respectively. Hence, to produce adequate sound pressure, m should be large enough, but this in turns results in higher distortion. Distortions are of two kinds: Inter-side-band distortions and intra-side-band distortions.

In [8], Pompei presented a simple theoretic method to reduce distortion effects. He proposed to apply Kite's pre-processing algorithm to the modulation envelope in association with the conventional AM modulation (17). Kite's algorithm consists in applying a double time integral to compensate for the double time derivative and a square root to compensate for the quadratic nonlinearity (according to Berktay's model).

$$s_{in}^{SR/DSB} = \sqrt{\left(1 + m \iint g(t)\right)} \sin(\omega_0 t) \qquad (17)$$

Attractive in theory, this solution is still less than ideal in practice. For being totally removed, all the harmonics generated by the square root operation need to be reproduced. However, this requirement can only be realized if a very wide band parametric array loudspeaker is used which is generally not feasible. In practice, the effective impact of Pompei's approach is a strong attenuation of the lower side band, thus a reduction of the inter-side-band distortion, but not of the intra-side-band distortion.

SSBWC AM modulation scheme

A more efficient way to avoid inter-side-band distortions is to adopt a SSBWC (Single Side Band With Carrier) modulation scheme. The main characteristic of SSB is the presence of a unique side band. Having just one side band, the inter-side-band distortion is null. Moreover, Tan *et al*. [9] have shown that some intra-side-band distortion components are removed. SSBWC modulation processing may be expressed as following:

$$s_{in}^{SSB} = \left(1 + mg(t)\right) \sin(\omega_0 t) + mg^*(t) \cos(\omega_0 t) \qquad (18)$$

where g^* is the Hilbert transform of g.

TRANSAURAL APPLICATION OF PARAMETRIC ARRAY LOUDSPEAKERS

Loudspeaker configuration

One of the only ways to make use of a pair of parametric array loudspeakers for transaural applications is by positioning them orthogonal to the median plane of the listener, each device pointing directly to the corresponding ipsilateral ear. This configuration should maximize the benefits of their sharper directivity patterns, and hence reduce naturally the transaural crosstalk. Unfortunately, this disposition does not allow using arrays of radius greater than the one of the listener's head, keeping in mind that the wider a loudspeaker is, the narrower its directivity should be. Such an arrangement of the loudspeakers is comparable to the one of a stereo dipole, as explained previously.

Modelling of the directivity

In our study, we investigated the use of a device of radius 100mm, assuming listener's head was of same size. One of the more convenient ways to model the directivity of the audible byproducts of a parametric array is to make use of Muir and Villette numerical solution [6], which describe the pressure field of sum and difference waves generated by the interaction of two monochromatic carriers in the far-field of an infinitely baffled circular piston source. Figure 4 exposes the geometry of the problem, and presents the terms used in the next equations.

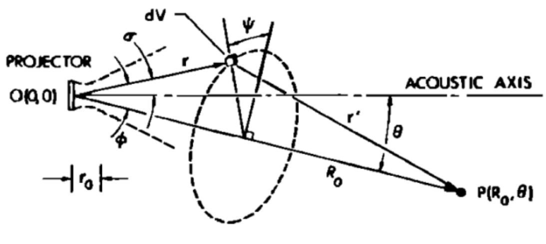

Figure 4. Geometry of the problem as exposed by Muir and Villette – image from [6]

The primary wave pressure p_i emitted by the piston may be expressed by:

$$p_i(r,\sigma) = \sum_{n=1}^{2} D(k_n a \sin\sigma) \ldots$$
$$\frac{r_0}{r} p_n e^{-\alpha_n r} \cos(\omega_n t - k_n r) \tag{19}$$

with:

$$D(k_n a \sin\sigma) = \frac{2J_1(k_n a \sin\sigma)}{k_n a \sin\sigma} \tag{20}$$

The solution for the secondary wave pressure p_s, valid for either the sum or the difference wave, is presented by Muir and Villette in the following equation form:

$$p_s(R_0,\theta) = \frac{2p_1 p_2 r_0^2 \omega_s^2 \beta}{\pi \rho_0 c_0^4 a^2} \int_{r_0}^{R_0} \int_0^{\varphi_e+\theta} \int_0^{\pi} \ldots$$
$$D\big(k_1 a \sin\sigma\,(\psi,\varphi,\theta)\big) D\big(k_2 a \sin\sigma(\psi,\varphi,\theta)\big) \ldots \tag{21}$$
$$\frac{e^{-((\alpha_1+\alpha_2)-ik_s)r-(ik_s-\alpha_s)r'(r,\varphi))}}{r'(r,\varphi)} \sin\varphi \, d\psi d\varphi dr$$

The distance r_0 corresponds to the limit between near-field and far-field, and may be expressed by $r_0 = 3a^2/4\lambda$, where λ is the wavelength of one of the two carriers. Muir and Villette validated this numerical solution by experimentations.

In our study, we fixed $f_1 = \omega_1/c$ at 40kHz and made f_2 vary from 40,5kHz to 56kHz. For 40kHz, we found r_0 = 875mm, and chose R_0 = 1m, in the vicinity of the near-field/far-field boundary (the closer from the array the listener is, the greater the angle at which the contralateral ear is illuminated, the more the transaural crosstalk should be reduced). For saving computation time, we decided to evaluate the difference wave pressure field just at octave band center frequencies and at angles varying from 0° to 30° by 5° steps. As Muir and Villette observed that the secondary wave directivity was much smoother compared to the one of the carriers (no null), we assumed that interpolated data should be sufficient to provide a relevant tendency. 2D spline-based interpolation was used to refine the results at any desired frequency or angle.

Figure 5 presents the directivity pattern of our theoretic parametric array loudspeaker, and Figure 6 compares it to the directivity pattern of the carrier wave. We may see that the secondary wave directivity is not always narrower than the primary wave. In our case, the parametric array directivity for frequencies less than 2kHz is wider than for the carrier frequency at 40kHz.

Figure 7 presents a comparison between the directivity patterns at the same frequencies of the secondary nonlinear audible byproducts and the primary linear audible waves, that the same baffled piston were able to emit (radius of 100mm, distance of 1m). The narrower directivity of the secondary nonlinear wave is clearly visible, and the difference with the primary linear wave increases with frequency. From these results, we may expect a relevant reduction of the transaural crosstalk by using the nonlinear effect of the air instead of the direct linear wave emission.

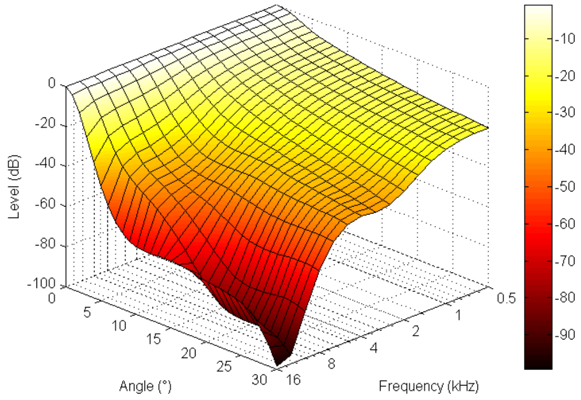

Figure 5. Directivity pattern of the audible byproducts emitted by a baffled piston of radius 100m at a distance of 1m, for a carrier frequency of 40kHz.

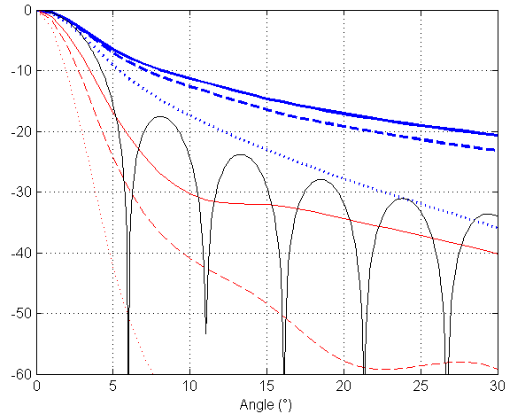

Figure 6. Comparison of the directivity patterns of the audible byproducts and the carrier wave – colored: difference wave; black: carrier at 40kHzk – blue: 0,5kHz; dashed blue: 1kHz; dotted blue: 2kHz; red: 4kHz; dashed red: 8kHz; dotted red: 16kHz.

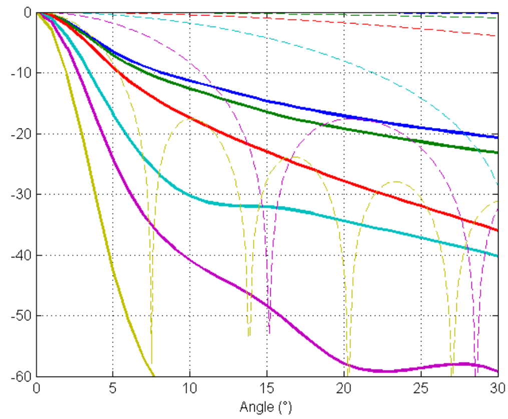

Figure 7. Comparison of the directivity patterns of primary linear wave and secondary nonlinear wave at the same frequencies – colored: nonlinear wave; dashed colored: linear – blue: 0,5kHz; green: 1kHz; red: 2kHz; cyan: 4kHz; magenta: 8kHz; yellow: 16kHz.

Transaural crosstalk reduction

To evaluate the impact of the directivity on the transaural crosstalk, we assumed that the transaural crosstalk may be structurally composed of the crosstalk caused by a monopole source in presence of listener's head, and the crosstalk caused by the directivity of the considered source located at the same position as the monopole source, but the head being acoustically transparent. This assumption was not verified, but it is based on the fact that the diffraction effects caused by the head have a lesser impact on the HRTF for sources located in front of the listener. This assumption also permitted us to evaluate rapidly the reduction of the transaural crosstalk that we may expect from the use of such a parametric array.

At 1m, the contralateral ear of a 100mm radius spherical head is illuminated by the parametric array at an angle of about 10° (for ears located at ±90° on the sphere). Figure 8 shows the expected transaural crosstalk reduction relative to the use of a baffled piston as a parametric array loudspeaker instead of a conventional loudspeaker.

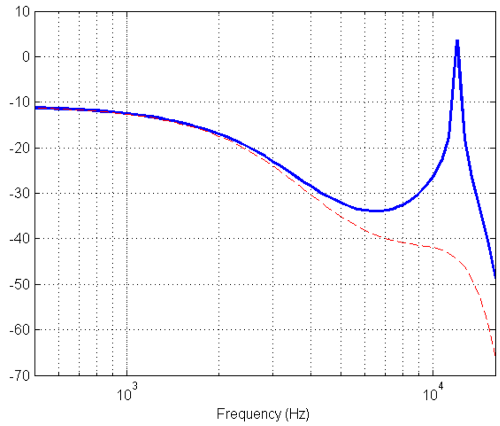

Figure 8. Transaural crosstalk reduction induced by the parametric array – blue: compared to a baffled piston source of same size used in a linear fashion; dashed red: compared to a monopole source.

The sharp peak exhibited by the blue line is due to a null in the directivity pattern of the primary linear wave response of the baffled piston. Conventional loudspeakers are generally made of more than one baffled piston, their directivity tending to the one of a monopole source in the ±30° range. Thus, the reduction of the transaural crosstalk induced by our theoretic parametric array may better be described by the dashed red line. We see that the reduction is of more than 20dB for frequencies above 3kHz.

CONCLUSION

A transaural crosstalk of at least -20dB is required for enabling transaural audio system to render spatial effects. Our simulations have shown that the use of parametric array loudspeakers in a stereo dipole configuration may enable this low level for frequencies above 3kHz. Hence, for this range of frequencies, it may be possible to diffuse directly binaural 3D audio via a pair of parametric arrays without making use of a transaural crosstalk cancellation step. For frequencies lower than 3kHz, this processing step is still required, but it may involve less energy for achieving a desired crosstalk level than conventional loudspeakers do.

REFERENCES

1 M.R. Schroeder and B.S. Atal, "Computer simulation of sound transmission in rooms" *IEEE Conv. Record* **7**, 150–155 (1963)
2 D.H. Cooper and J.L. Bauck, "Prospects for transaural recording" *J. Audio Eng. Soc*. **37**, 3–19 (1989)
3 J.L. Bauck and D.H. Cooper, "Generalized transaural stereo and applications" *J. Audio Eng. Soc*. **44(9)**, 683–705 (1996)
4 P.J. Westervelt, "Parametric acoustic array" *J. Acoust. Soc. Am.* **35(4)**, 535–537 (1963)
5 H.O. Berktay, "Possible exploitation of nonlinear acoustics in underwater transmitting applications" *J. Sound Vib*. **2**, 435–461 (1965)
6 T.G. Muir and J.G. Villette, "Parametric acoustic transmitting arrays" *J. Acoust. Soc. Am.* **52**, 1481–1486(1972)
7 M. Yoneyama *et al*., "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design" *J. Acoust. Soc. Am.* **73(5)**, 1532–1536(1983)
8 F.J. Pompei, "The use of airborne ultrasonics for generating audible sound beams" *J. Audio Eng. Soc*. **47(9)**, 726–731 (1999)
9 E.L. Tan. *et al*., "Distortion analysis and reduction for the parametric array" *Proceedings of the 124th Audio Eng. Soc. Conv.* (2008)