Phantom Sources Applied to Stereo-Base Widening*

RONALD M. AARTS, AES Fellow

Philips Research Laboratories, 5656 AA Eindhoven, The Netherlands

A stereo-base widening system producing a good effect, achievable with a simple analog circuit as well as with a digital one, is presented. The system is derived from a combination of the traditional approach using head-related transfer functions (HRTFs), and by using a simple model where ideal loudspeakers and an acoustically transparent subject’s head are assumed. Both the HRTF approach and the simple model are discussed, and some special cases for the latter are considered. Merging these models with some experimentally determined adaptations led to a practical widening system. The system produces a pleasant and natural sound, particularly for voices and normal mixed stereo recordings. The widening circuit is tuned in such a way that for common stereo recordings it gives the same tonal balance as conventional stereo reproduction. For sound effects as in action movies, the system produces extreme effects, which was a deliberate design goal of the system.

0 INTRODUCTION

It is hardly possible to image sound reproduction today without stereophonic techniques, and it is to the credit of both the technology and human binaural hearing capabilities that a single pair of loudspeakers can evoke auditory perspectives so convincingly. Since the first stereophonic demonstration at the Paris Opera [1] in 1881, a whole range of improvements to the system have been suggested, of which Blumlein’s patent [2] is an early example. Other early examples of interest are de Boer’s papers [3]–[5], of which [5] is concerned with the problem of the distance between the loudspeakers. As early as 1946, this distance had to be minimized “in order to obtain a compact arrangement.” This served as a mechanical solution, but numerous electronic solutions have also been proposed (see, for example, [6]–[9]).

Although there is a vast amount of literature on general stereophony (early publications include [3], [11]–[15] whereas [16]–[19] are more recent), it is beyond the scope of this section to discuss stereophony in general. Many of the “classic” papers are included in [10].

It is the purpose of this engineering report to discuss a method enabling a virtual widening of the loudspeaker separation—to make a compact setup as in portable audio devices—by using virtual loudspeakers. The idea of creating virtual loudspeaker sources was introduced over 30 years ago. Schroeder and Atal [20], [21] outlined a method for generating a phantom source. These ideas were elaborated by Damaske [22]–[24] and are known as TRADIS (true reproduction of all directional information by stereophony). A vast amount of literature followed, in the form of papers (see, for example, [25], [26]), but mostly through patents. According to a patent reviewer [27]: “If the energy and intelligence spent on jazzing up stereophonic reproduction in automobiles during the past 10 years had instead been devoted to the space program, we would now have a colony on Mars.”

The aim of the following sections is to derive a practical method enabling a virtual widening of the loudspeaker separation. Various approaches for the widening problem will be discussed—a digital filter (using many multiplications) with the head-related transfer function (HRTF) approach in Section 2, and a second one by using a simple model where ideal loudspeakers are assumed.

1 GENERAL APPROACH

In order to create phantom sources $LS_{L_{ph}}$ and $LS_{R_{ph}}$, the left and right source signals, denoted by $V_L$ and $V_R$, were processed as shown in Fig. 1, where for reason of simplicity, only $LS_{R_{ph}}$ is drawn. The notations for a left phantom source follow analogously. The rationale of this processing is that the same pressure is obtained in both ears when the real loudspeakers $LS_L$ and $LS_R$ are playing as in the situation when the phantom loudspeakers $LS_{L_{ph}}$ and $LS_{R_{ph}}$ are playing. If the pressure at the eardrums generated by real loudspeakers is the same as that which would be generated by the phantom sources, and there are not other (such as visual) cues, then the subject will perceive no difference, and a virtual enlargement of the aperture $\alpha$ to $\beta$ is realized.

* Manuscript received 1998 April 7; revised 1999 May 5 and 1999 December 14.
The transfer function $H^p_{LR}$ is from the left loudspeaker to the right ear, with an angle $\alpha$ between the loudspeaker and the median plane. The other functions used are as indicated in Fig. 1, with the first (uppercase) index corresponding to the source and the second (lowercase) index to the ear. The pressures in the left and right ears caused by the real loudspeakers LS$_L$ and LS$_R$ are

\[ p_{LR} = V_{LR}H^p_{LR} + V_{LR}H^p_{L1} \]  
\[ p_{RL} = V_{RL}H^p_{RL} + V_{RL}H^p_{R1} \]  

where $V_{LR}$ and $V_{RL}$ are the filtered source signals fed to loudspeakers LS$_L$ and LS$_R$, respectively. The pressure at the ears caused by the phantom source LS$_{RPh}$ is

\[ p_{LRPh} = V_{LR}H^p_{R1} \]  
\[ p_{RLPh} = V_{RL}H^p_{R2} \]  

The combination of Eqs. (1)-(4) and the loudspeaker voltage relations given by

\[ V_{LR} = H_{L1}V_L + H_{L2}V_R \]  
\[ V_{RL} = H_{R1}V_R + H_{R2}V_L \]  

yield the filter functions $H_{R1}$, $H_{L1}$, $H_{R2}$, and $H_{L2}$,

\[ H_{R1} = \frac{H^p_{RL}H^p_{L1} - H^p_{R1}H^p_{L1}}{\Delta} \]  
\[ H_{L1} = \frac{H^p_{RL}H^p_{R1} - H^p_{R1}H^p_{R2}}{\Delta} \]  

where $\Delta = H^p_{RL}H^p_{R1} - H^p_{R1}H^p_{L1}$.

It appears that all functions $H$ have a common denominator $\Delta$. This term might give rise to complications in the implementation of $H$ due to a singularity at certain frequencies. However, the term $\Delta$ can be left out without changing the position of the phantom source, but it will then lead to timbre differences. In some cases, as will be discussed in Section 3.1, the numerators of Eqs. (7)-(10) contain their denominator as a factor, which gives very simple filter functions.

If these transfer functions [given by Eqs. (7)-(10)], are implemented, then the subject will perceive a virtual enlargement of the aperture $\alpha$ to $\beta$.

If there is symmetry in the median plane, $H_{L2} = H_{R1}$, and $H_{L1} = H_{R2}$. Then let $H_{R1} = H_{L1} = H_1$, $H_{R2} = H_{L2} = H_2$, and Eqs. (7)-(10) reduce to

\[ H_1 = \frac{H^p_{RL}H^p_{R1} - H^p_{R1}H^p_{L1}}{\Delta_1} \]  
\[ H_2 = \frac{H^p_{RL}H^p_{R2} - H^p_{R2}H^p_{R1}}{\Delta_1} \]  
\[ \Delta_1 = (H^p_{R1})^2 - (H^p_{R2})^2 \]  

Depending on the model used for the transfer functions $H$, various solutions can be obtained, as will be discussed.

---

Fig. 1. Setup for generating phantom sources. Using filters $H_{L1}-H_{R1}$ and "real" loudspeakers LS$_L$ and LS$_R$, a virtual enlargement of the aperture $\alpha$ to $\beta$ can be achieved. For reason of simplicity, only one phantom source (LS$_{RPh}$) is drawn, the notations for a left phantom source follow analogously. $p_l$, $p_r$—sound pressure at ears, $D_s$—head diameter.
2 MEASURED TRANSFER FUNCTIONS

The transfer functions used in Eqs. (1)–(4) were measured with the aid of the setup shown in Fig. 2 in an anechoic chamber. The results are plotted in Fig. 3. Invoking Eqs. (11)–(13), functions $H_1$ and $H_2$ were computed; the results are presented in Figs. 4 and 5, respectively. As Figs. 4 and 5 show, the required filter characteristics are rather complex. Due to possible changes in the position of the subject’s head during sound reproduction with respect to the head position during the measurement of the HRTFs, severe errors may be introduced in functions $H_1$ and $H_2$.

An adaptive method would be more suitable for computing $H_1$ and $H_2$. The adaptation process is described as follows. Initially three loudspeakers are used, one unfiltered at the position of the phantom source $L_{SP}$, the other two filtered. Microphones positioned near the subject’s ears are used to control the filter coefficients until the pressure measured is at a minimum level. The coefficients are then frozen and the third loudspeaker is removed. When the two loudspeakers are driven through the reversed sign of the function of the filter, only the illusion of a third source results. The method thus described can hence be regarded as an active noise-cancellation solution. We will not elaborate on this noise-cancellation principle; see [28], [29] for more details. A simple model will be discussed in the following section.

3 SIMPLE TRANSFER FUNCTIONS

In order to create phantom sources $L_{SP}$ and $L_{SP}$, the left and right signals $V_L$ and $V_R$ were processed as indicated in Fig. 1. When ideal loudspeakers and an acoustically transparent subject’s head (with a diameter $D_h = 2r$) are assumed, then the transfer function $H_{LR}$, from the left loudspeaker to the right ear, with an angle $\alpha$ between the loudspeaker and the median plane, is modeled as

$$H_{LR} = \frac{\exp(-jkD_\alpha)}{D_\alpha} \quad (14)$$

where $k$ is the wave number and $D_\alpha$ is the distance between the left loudspeaker and the right ear. If symmetry around the median plane is assumed, all loudspeakers are arranged in an arc with a radius $R$, $R \gg r$, and the common phase term exp$(-jkR)$ may be omitted. Therefore the pressure in the left and right ears generated by the real loudspeakers $L_{L}$ and $L_{R}$ is

$$p_{LR} = \frac{V_L \exp(-jk\sin \alpha)}{R} + \frac{V_R \exp(jk\sin \alpha)}{R} \quad (15)$$

![Fig. 2. Setup for measuring HRTFs. For purposes of verification and measurement a real source is present at the place of the phantom source $L_{SP}$ (center). The left source $L_{L}$, (not visible) is present at the left of the monitor.](image)

![Fig. 3. Measured HRTFs, including loudspeaker response (which was flat within ±2 dB in the range of the plot) from right loudspeaker to right ear. Legend indicates angle of loudspeaker with respect to median plane.](image)
AARTS ENGINEERING REPORTS

\[ P_{R,L} = \frac{V_{R,R} \exp(jkr \sin \alpha)}{R} + \frac{V_{L,R} \exp(-jkr \sin \alpha)}{R} \]  \ \ \ (16)

where \( V_{L,R} \) and \( V_{R_R} \) are the voltages applied to the two loudspeakers, supplied by the filters \( H_1 \) and \( H_2 \), where, as in the assumed symmetrical setup, \( H_1 = H_{L1} = H_{R1} \) and \( H_2 = H_{L2} = H_{R2} \).

The pressure generated by the phantom source \( LS_{RPh} \) is

\[ \rho_{LPh} = V_{R,R} \exp(-jkr \sin \beta) \]  \ \ \ (17)

\[ \rho_{RPh} = V_{R,R} \exp(jkr \sin \beta) \]  \ \ \ (18)

If Eq. (15) is equated to Eq. (17), Eq. (16) is equated to Eq. (18), and Eqs. (5) and (6) are implemented, then the following holds for filters \( H_1 \) and \( H_2 \):

\[ H_1 = \frac{\sin[kr(\sin \alpha + \sin \beta)]}{\sin[2kr(\sin \alpha)]} \]  \ \ \ (19)

\[ H_2 = \frac{\sin[kr(\sin \alpha - \sin \beta)]}{\sin[2kr(\sin \alpha)]} \]  \ \ \ (20)

It appears that the functions \( H \) have a common denominator \( \sin[2kr(\sin \alpha)] \). This term might give rise to complications in implementing \( H \) due to a singularity at certain frequencies. The physical explanation for this singularity is that for certain frequencies in relation to the geometry, the path length difference between a loudspeaker and both ears becomes equal to a wavelength of the sinusoidal input signal.

![Fig. 4. Transfer functions of filter \( H_1 \) using HRTFs and Eq. (11). Legend indicates angle \( \beta \) of phantom loudspeaker with respect to median plane; "real" sources were at \( \alpha = \pm 10^\circ \).](image)

![Fig. 5. Transfer functions of filter \( H_2 \) using HRTFs and Eq. (12). Legend indicates angle \( \beta \) of phantom loudspeaker with respect to median plane; "real" sources were at \( \alpha = \pm 10^\circ \).](image)
3.1 Special Cases

If \( A = k r \sin \alpha \) and \( h = \sin \beta / \sin \alpha \), then

\[
H_1 = \frac{\sin((1 + h)A)}{\sin(2A)}
\]

(21)

\[
H_2 = \frac{\sin((1 - h)A)}{\sin(2A)}
\]

(22)

For odd integer values of \( h \) the numerators of Eqs. (21) and (22) contain their denominator as a factor, as can be shown.

Eqs. (21) and (22) are related to Chebyshev polynomials. Setting \( \cos \theta = x \), the expressions

\[
T_n(x) = \cos(n\theta)
\]

\[
U_n = \frac{1}{n + 1} \left( T_{n+1}(x) - \frac{\sin((n + 1)\theta)}{\sin\theta} \right)
\]

\( n = 0, 1, 2, \ldots \)  (23)

are the Chebyshev polynomials in \( x \) of degree \( n \). Thus if \( \theta = 2A \) and \( (n + 1)2 = 1 \pm h \) it appears that for odd integer values of \( h \) the transfer functions \( H_{1,2} \) can be written as a cosine polynomial, which leads to the following identities:

\[
\frac{\sin(4n\theta)}{\sin(2\theta)} = 2 \sum_{i=1}^{n} \cos((2i - 1)2\theta)\]

(24)

\[
\frac{\sin(2\theta(2n + 1))}{\sin(2\theta)} = 1 + 2 \sum_{i=1}^{n} \cos(4i\theta)\]

(25)

As an example two values for \( h \) will be used. For \( h = 3 \) we have

\[
H_{1h=3} = 2 \cos(2A)
\]

(26)

\[
H_{2h=3} = -1
\]

(27)

Here we observe the usual, well-known effect of the introduction of a 180° phase-shifted "crosstalk" between the channels (see, for example, [30]). The "main path" \( H_{1h=3} \) then becomes weaker than the "cross path" \( H_{2h=3} \), which can produce annoying effects, such as an unstable stereo image due to the listener's head movements. These equations are real (the imaginary part is equal to zero). However, after multiplication by a common phase term, they can be easily implemented with finite impulse response (FIR) filters. The same result [Eqs. (26), (27)] would be obtained for a small-angle and small-A-value approximation of Eqs. (19) and (20).

For \( h = 5 \) we find

\[
H_{1h=5} = 1 + 2 \cos(4A)
\]

(28)

\[
H_{2h=5} = -2 \cos(2A)
\]

(29)

Then the cross term \( H_2 \) becomes frequency dependent. Using the relations between \( U_n(x) \) and \( T_n(x) \), see [31, §22.5], the recursive relation

\[
H_{1h} = 2 \cos(2A)H_{1h-2} - H_{1h-4}
\]

(30)

can be derived, with \( H_{1h=1} = 0 \), \( H_{1h=1} = 1 \), and \( |h| = 1, 3, 5, \ldots \), and similarly for \( H_2 \), or by using the relation

\[
H_{2h} = -H_{1h-2}
\]

(31)

Thus for odd integer values of \( h \) Eqs. (21) and (22) can be realized with the aid of simple FIR filters, as will be discussed in the following section.

4 IMPLEMENTATIONS OF STEREO WIDENING

Two electronic circuits realizing the widening system have been designed, an analog component version and a digital one. Both circuits are based on the simple model of Section 3. According to this model the required filters can be implemented by FIR filters using Eqs. (21) and (22), as shown in Fig. 6. The value of integer constants \( m_0, \ldots, m_3 \) (determining the delay between the taps) and the coefficients \( c_0, \ldots, c_5 \), are chosen so that the filter response gives a best fit to Eqs. (21) and (22). The transfer functions of these functions are plotted in Fig. 7 for \( D_s = 2r = 17.5 \) cm and \( \alpha = 10^\circ \). However, the model of Section 3 does not take into account the influence of the pinna, the ear canal, the head and torso, and so on. Therefore the filter shown in Fig. 6 (for \( h = 3 \)) was tuned "by hand," so that the timbre was as close as possible to normal stereo reproduction for various types of music, while the widening remained. The model is inadequate to describe the behavior of true HRTFs at higher frequencies than, say, 3 kHz, since an acoustically transparent head was assumed. Therefore we see large deviations at high frequencies between the model "tuned" response.

Of course some of the relevant parameters can be made accessible to users, who can then adapt them to their own preference. The so obtained transfer function was transferred to a circuit using analog components, as discussed in the next section.

4.1 Simple Analog Filter

The tuned filter response as just described was fitted experimentally to a circuit using analog components. The bass response was slightly increased, which was previously not possible in the (for intermediate purpose only, digital) FIR filter (Fig. 6) due to its short length. The amplitude response of the widening circuit is shown in Fig. 8. The label \( H_s \) refers to the output of one of the channels, while both channels are driven by the same signal. The line marked \( S \) is the level at which the circuit has the same loudness as ordinary stereo. The phase response is shown in Fig. 9. It appears that the phase of the sum signal is rather flat, which explains the lack of coloration, especially for mono components in the
Fig. 6. FIR filter setup for $H_1$ and $H_2$ given by Eqs. (21) and (22) using simple model where ideal loudspeakers and an acoustically transparent subject’s head are assumed. $c_0$-$c_5$—coefficients (tap weights) of FIR filter; $m_0$-$m_3$ integer constants; $f_s$—sampling frequency; $T = 1/f_s$—sampling time; $mT = \tau$—delay between taps $c$.

Fig. 7. FIR filter transfer functions $H_1$ and $H_2$ for $h = 1, 3, 5$ in Eq. (21), yielding $H_{1,h=1} = -H_{2,h=3}$, $H_{1,h=3} = -H_{2,h=5}$, $H_{1,h=5} = -H_{2,h=7}$. Ideal loudspeakers and an acoustically transparent subject’s head are assumed. $D_0 = 2r = 17.5$ cm and $\alpha = 10^\circ$. 
(stereo) signal. This level was determined empirically by a panel of five experienced listeners, using popular and classical music. Finally the analog filter was fitted to a digital IIR filter, which could be done with a deviation within one half of a decibel, consisting of order 5 and 4 for the $H_1$ and $H_2$ filters, respectively.

5 CONCLUSIONS

A stereo-base widening system producing a good effect has been presented. The "averaging" between the traditional HRTF approach and a simple model appeared to be a practical method to derive such a system. The system is applicable to TV/monitor sets (multimedia), portable audio, and midi sets. It produces a pleasant and natural sound, particularly for voices. For normal stereo recordings it produces the same tonal balance as is obtained without the circuit.

6 REFERENCES


![Graph](image1)

**Fig. 9.** Phase response of widening circuit.

![Graph](image2)

**Fig. 9.** Phase response of widening circuit.


THE AUTHOR
Ronald Aarts was born in Amsterdam, The Netherlands, in 1956. He received the B.Sc. degree in electrical engineering in 1977 and the Ph.D. degree in 1994 from Delft University of Technology.

In 1977 he joined Philips Research Laboratories, Eindhoven, The Netherlands, in the Optics group. There he was engaged in research into servos and signal processing for use in both Video Long Play players and Compact Disc players. In 1984 he joined the Acoustics group and worked in the development of CAD tools and signal processing for loudspeaker systems. In 1994 he became a member of the DSP group and became engaged in the improvement of sound reproduction by exploiting DSP and psychoacoustical phenomena.

He has published a number of technical papers and reports and is the holder of several patents in the aforementioned fields. He was a member of the organizing committee and chairman for various conventions. He is a senior member of the IEEE, a fellow of the AES, the NAG (Dutch Acoustical Society), and the Acoustical Society of America. He is a past chairman of the Dutch Section of the AES.