The "Stereo Dipole"—A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers*

OLE KIRKEBY, AES Member, AND PHILIP A. NELSON, AES Member

Institute of Sound and Vibration Research, University of Southampton, Highfield, SO17, IBJ, UK

AND

HAREO HAMADA, AES Member

Department of Electrical and Communications Engineering, Tokyo Denki University, Tokyo 101, Japan

A system comprising only two closely spaced loudspeakers can create very convincing virtual images around a single listener. It is demonstrated that such a system is very robust with respect to head movement, and that the processing does not introduce any excessive artifacts. In practice, the loudspeakers ought to be "pair matched" in order to ensure accurate imaging.

0 INTRODUCTION

Under certain circumstances it is possible to give a listener the impression that there is a sound source, referred to as a "virtual source," at a given position in space where no real sound source exists. One way to achieve this is to ensure that the sound pressures that are reproduced at the listener's ears are the same as the sound pressures that would have been produced there by a real source at the same position as the virtual source [1]–[4]. With N loudspeakers it is in principle possible to control the sound field at N target points [5], and so with only two loudspeakers placed symmetrically in front of a single listener it is possible to reproduce any pair of desired signals at the two ears of the listener. In practice, however, it is advantageous that the desired signals be reproduced accurately not only at the listener's ears, but also in the vicinity of those two points since this will allow moderate head movement without letting the illusion of the virtual image break down.

In order to determine the signals that must be input to the loudspeakers, a suitable processing scheme must be applied to the source material. If the source material is a pair of binaural signals as, for example, the result of a dummy-head recording [6], [7], then it must be passed through a crosstalk cancellation network [8]. If the source material is monophonic, such as a speech signal, it must be passed through a pair of filters whose characteristics depend on the intended position of the virtual source [9].

The purpose of a crosstalk cancellation system is to produce a specified desired signal very accurately at one ear of the listener, while nothing is heard at the other ear [10]. Once this can be achieved, any pair of signals can be produced at the ears of a listener. The first method for the cancellation of crosstalk is probably that patented by Atal et al. [8] in 1966. Their method used analog techniques, and it was based on a free-field model that did not account for the presence of a listener in the sound field. Since then more sophisticated methods, some based on digital signal processing techniques, have been developed [11]–[14].

It seems that with a few notable exceptions [15], [16] most workers have concentrated on virtual source imaging systems using widely spaced loudspeakers, typically spanning an angle of 60° as seen by the listener. The aim of this paper is to show that it is an advantage to position the two loudspeakers close together so that they span only 10° as seen by the listener. We refer to such a loudspeaker arrangement as a "stereo dipole," even though strictly speaking it would be more accurate to use the term "stereo monopole–dipole" [17]. A simple free-field model, identical to the one used by Atal et al., is used to demonstrate the physics underlying the problem.

1 PROBLEM DESCRIPTION

1.1 Geometry

The geometry of the problem is shown in Fig. 1. Two loudspeakers (sources), separated by a distance \( \Delta s \), are positioned on the \( x_1 \) axis symmetrically about the \( x_3 \) axis. We imagine that a listener is positioned \( r_0 \) meters away from the loudspeakers directly in front of them. The ears of the listener are represented by two microphones, separated by the distance \( \Delta M \), and they are also positioned symmetrically about the \( x_3 \) axis. Note that when we refer to the "left" or "right" loudspeaker or ear or microphone, we consider this to be relative to the listener's point of view as indicated in Fig. 1. The loudspeakers span an angle \( \theta \) as seen from the position of the listener. Only two of the four distances from the loudspeakers to the microphones are different: \( r_1 \) is the shortest (direct path), \( r_2 \) is the furthest (crosstalk path). The inputs to the right and left loudspeakers are denoted by \( V_1 \) and \( V_2 \), respectively, the outputs from the right and left microphones are \( W_1 \) and \( W_2 \), respectively. It will later prove convenient to introduce the two variables

\[
g_c = \frac{r_1}{r_2} \tag{1}
\]

which is a gain that is always smaller than 1, and

\[
\tau_c = \frac{r_2 - r_1}{c_0} \tag{2}
\]

which is a positive delay corresponding to the time it takes the sound to travel the path length difference \( r_2 - r_1 \).

1.2 Electroacoustic Transfer Functions

1.2.1 Definitions

When the system is operating at a single frequency, we can use complex notation to describe the inputs to the loudspeakers and the outputs from the microphones. Thus, we assume that \( V_1, V_2, W_1, \) and \( W_2 \) are complex scalars. Unless otherwise stated, we use uppercase italic letters to denote frequency variables, lowercase italic letters to denote time-domain variables, lowercase bold italic to denote vectors, and uppercase bold italic to denote matrices. The loudspeaker inputs and the microphone outputs are related through the two transfer functions

\[
C_1 = \begin{bmatrix} W_1 \\ V_1 \end{bmatrix}_{V_1 = 0} = \frac{W_2}{V_2}_{V_1 = 0} \tag{3a}
\]

and

\[
C_2 = \begin{bmatrix} W_2 \\ V_1 \end{bmatrix}_{V_1 = 0} = \frac{W_1}{V_2}_{V_1 = 0} \tag{3b}
\]

Using these two transfer functions, the output from the microphones as a function of the inputs to the loudspeakers is conveniently expressed as a matrix–vector multiplication,

\[
w = Cv \tag{4}
\]

where

\[
w = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \\ C_2 & C_1 \end{bmatrix}, \quad v = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{5}
\]

1.2.2 Free-Field Conditions

The sound field \( P_{mo} \) created by a monopole in a free field is given by

\[
P_{mo} = j \omega \rho_0 q \frac{\exp(-jkr)}{4\pi} \tag{6}
\]

where \( \omega \) is the angular frequency, \( \rho_0 \) is the density of the medium, \( q \) is the source strength, \( k \) is the wave number \( \omega/c_0 \), \( c_0 \) being the speed of sound, and \( r \) is the distance from the source to the field point. If \( V \) is defined as

\[
V = \frac{j \omega \rho_0 q}{4\pi} \tag{7}
\]

then the transfer function \( C \) is given by

\[
C = \frac{P_{mo}}{V} = \frac{\exp(-jkr)}{r} \tag{8}
\]

Note that \( V \) has the dimension of source "acceleration." This quantity is the time derivative of the source strength, which is the variable usually used in textbooks on acoustics.

2 CROSSTALK CANCELLATION AND VIRTUAL SOURCE IMAGING

The aim of the system shown in Fig. 1 is to reproduce a pair of desired signals \( D_1 \) and \( D_2 \) at the microphones. Consequently we require \( W_1 \) to be equal to \( D_1 \), and \( W_2 \) to be equal to \( D_2 \). The pair of desired signals can be specified with two fundamentally different objectives in mind: crosstalk cancellation or virtual source imaging. In both cases, two linear filters \( H_1 \) and \( H_2 \) operate on a single input \( D \), and so

\[
v = Dh \tag{9}
\]

where

\[
h = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \tag{10}
\]

This is illustrated in Fig. 2. Perfect crosstalk cancellation [Fig. 2(a)] requires that a signal be reproduced perfectly at one ear of the listener while nothing is heard at the other ear. If we want to produce a desired signal \( D_2 \) at the listener's left ear, then \( D_1 \) must be zero. Virtual source imaging [Fig. 2(b)], on the other hand, requires
that the signals reproduced at the ears of the listener be identical (up to a common delay and a common scaling factor) to the signals that would have been produced there by a real source.

2.1 Crosstalk Cancellation under Free-Field Conditions

2.1.1 Analysis

It is advantageous to define $D_2$ to be the product of $D$ and the phase factor $\exp(\textstyle{jkr}_1)$ since this guarantees that the time responses corresponding to the frequency response functions $V_1$ and $V_2$ are causal. (In the time domain this causes the desired signal to be delayed, but it does not affect its “shape.”) By solving the linear equation system

$$C_r = \begin{bmatrix} 0 \\ D \end{bmatrix} \exp(-jkr_1)$$  \hspace{1cm} (11)

for $\nu$, we find

$$\nu = Dr_1 \left[ \frac{1}{1 - g_c^2 \exp(-j2\omega \tau_c)} \right] \exp(-j2\omega \tau_c).$$  \hspace{1cm} (12)

In order to find the time response of $\nu$, we rewrite the term $1/(1 - g_c^2 \exp(-j2\omega \tau_c))$, using the power series expansion [18, p. 153]

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \cdots, \quad |z| < 1.$$  \hspace{1cm} (13)

The result is

$$\nu = Dr_1 \left[ \frac{-g_c \exp(-j\omega \tau_c)}{1} \right] \sum_{n=0}^{\infty} g_c^{2n} \exp(-j2\omega \tau_c).$$  \hspace{1cm} (14)

After an inverse Fourier transform [19, ch. 2], we can now write $\nu$ as a function of time,

$$\left[ \begin{array}{c} \nu_1(t) \\ \nu_2(t) \end{array} \right] = r_1 \left[ \begin{array}{c} -g_c \delta(t - \tau_c) \\ \delta(t) \end{array} \right] \sum_{n=0}^{\infty} g_c^{2n} \delta(t - 2n\tau_c).$$  \hspace{1cm} (15)

Fig. 1. Geometry of loudspeaker—microphone arrangement. $\theta$—angle spanned by loudspeakers as seen from center of listener's head; $r_0$—distance from this point to center between loudspeakers.

Fig. 2. Definitions of necessary transfer functions, signals, and filters. (a) Crosstalk cancellation. (b) Virtual source imaging.
where * denotes convolution and δ is the dirac delta function [19, sec. 1.1.4 and 1.1.5]. The summation represents a decaying train of delta functions. The first delta function occurs at time \( t = 0 \), and adjacent delta functions are \( 2\tau \) apart. Consequently, as recognized by Atal et al. [8], \( v(t) \) is intrinsically recursive, but even so it is guaranteed to be both causal and stable as long as \( d(t) \) is causal and stable. The solution is readily interpreted physically in the case where \( d(t) \) is a pulse of very short duration (more specifically, much shorter than \( \tau \)). First, the left loudspeaker sends out a pulse which is heard at the listener’s left ear. At time \( \tau \), after reaching the left ear, this pulse reaches the listener’s right ear, where it is not intended to be heard, and consequently between adjacent pulses also increases. This evidently is heard at the listener’s right ear. It is clear from Fig. 3 that as \( \theta \) increases, the overlap between adjacent pulses also increases. This evidently makes \( v_1(t) \) and \( v_2(t) \) smoother, and it is intuitively obvious that if \( f_0 \) is very large, the ringing frequency is suppressed almost completely, and both \( v_1(t) \) and \( v_2(t) \) will be simple decaying exponentials (decaying in the sense that they both return to zero for large \( t \)). However, it is also intuitively obvious that by increasing \( f_0 \), the

For the three loudspeaker spans of 60, 20, and 10° this approximation gives the three \( f_0 \) values of 1.8, 5.4, and 10.8 kHz (rule of thumb: \( f_0 \approx 100 \text{kHz} \) divided by loudspeaker span in degrees), which are in good agreement with the exact values. According to Eq. (18), \( f_0 \) tends to infinity as \( \theta \) tends to zero. It can be shown that the limiting case is equivalent to a monopole and a point dipole, both positioned at the origin of the coordinate system [17]. A crosstalk cancellation system based on this source combination is in a sense optimal since the reproduced sound field does not contain any “ringing” (this will be illustrated later).

It is clear from Fig. 3 that as \( f_0 \) increases, the overlap between adjacent pulses also increases. This evidently makes \( v_1(t) \) and \( v_2(t) \) smoother, and it is intuitively obvious that if \( f_0 \) is very large, the ringing frequency is suppressed almost completely, and both \( v_1(t) \) and \( v_2(t) \) will be simple decaying exponentials (decaying in the sense that they both return to zero for large \( t \)). However, it is also intuitively obvious that by increasing \( f_0 \), the

2.1.2 Source Inputs

Fig. 3 shows the input to the two sources for three different loudspeaker spans of 60, 20, and 10°. The distance to the listener is 0.5 m (which is probably a reasonable estimate of the average distance that the user of a multimedia computer will sit away from the screen). The microphone separation, which approximates the head diameter, is 180 mm. The desired signal is a Hanning pulse (one period of a cosine, sometimes also referred to as raised cosine) specified by

\[
d(t) = \begin{cases} 
(1 - \cos \omega_0 t)/2, & 0 \leq t \leq 2\pi/\omega_0 \\
0, & \text{all other } t 
\end{cases}
\]  

(16)

where \( \omega_0 \) is chosen to be 2\( \pi \) times 3.2 kHz (the spectrum of this pulse has its first zero at 6.4 kHz). For the three loudspeaker spans of 60, 20, and 10°, the corresponding ringing frequencies \( f_0 \) are 1.9, 5.5, and 11 kHz, respectively. If the listener does not sit too close to the sources, \( \tau \) is well approximated by assuming that the direct path and the crosstalk path are parallel lines, in which case

\[
\tau \approx \frac{\Delta M}{c_0} \sin \left( \frac{\theta}{2} \right).
\]  

(17)

If in addition we assume that the loudspeaker span is small, then \( \sin(\theta/2) \) can be simplified to \( \theta/2 \), and so \( f_0 \) is well approximated by

\[
f_0 \approx \frac{c_0}{\Delta M} \frac{1}{\theta}.
\]  

(18)

Fig. 3. Time response of two source input signals required to achieve perfect crosstalk cancellation at listener’s right ear for three loudspeaker spans \( \theta \). \( v_1 \) — heavy lines; \( v_2 \) — light lines. (a) \( \theta = 60^\circ \); (b) \( \theta = 20^\circ \); (c) \( \theta = 10^\circ \). Note how overlap increases as \( \theta \) decreases.
low-frequency contents of both $v_1(t)$ and $v_2(t)$ are increased. Consequently, in order to achieve perfect crosstalk cancellation with a pair of closely spaced loudspeakers, a very large low-frequency output is necessary. This undesirable property is caused by the underlying physics of the problem, and there is no obvious way to avoid it.

2.1.3 Reproduced Sound Field

Fig. 4 shows the sound field reproduced by four different source configurations: the three loudspeaker spans of 60, 20, and 10°, and the sound field generated by a point monopole and a point dipole. The sound fields plotted in Fig. 4(a)-(c) are those generated by the source inputs plotted in Fig. 3. Each of the four plots contains nine "snapshots," or frames, of the sound field. The frames are listed sequentially in a "reading sequence" from top left to bottom right. Top left is the earliest time ($t = 0.2/c_0$), bottom right is the latest time ($t = 1.0/c_0$). The time increment between each frame is $0.1/c_0$, which is equivalent to the time it takes the sound to travel 100 mm. The normalization of the desired signals ensures that the left loudspeaker starts emitting sound at exactly $t = 0$; the right loudspeaker starts emitting sound $\tau_c$ seconds later. Each frame is calculated at 101 x 101 points over an area of 1 m x 1 m ($-0.5 \text{ m} < x_1 < 0.5 \text{ m}, 0 < x_2 < 1$). The positions of the loudspeakers and the microphones are indicated by circles. Values greater than 1 are plotted as white, values smaller than $-1$ are plotted as blacks, values between $-1$ and 1 are shaded appropriately.

Fig. 4(a) illustrates the crosstalk cancellation principle when $\theta$ is 60°. It is easy to identify a sequence of positive pulses from the left loudspeaker and a sequence of negative pulses from the right loudspeaker. Each of the two pulse trains is emitted with the ringing frequency of 1.9 kHz. Only the first pulse emitted from the left

Fig. 4. Sound field reproduced by four different source configurations adjusted to achieve perfect crosstalk cancellation at listener's right ear. (a) $\theta = 60^\circ$. (b) $\theta = 20^\circ$. (c) $\theta = 10^\circ$. (d) Monopole–dipole combination.
The loudspeaker is actually "seen" at the left microphone; consecutive pulses are canceled out at both microphones. However, many copies of the original Hanning pulse are seen at other locations in the sound field, even very close to the two microphones, and so this setup is very sensitive to head movement.

When the loudspeaker span is reduced to 20° [Fig. 4(b)], the reproduced sound field becomes simpler. The desired Hanning pulse is now beamed toward the left microphone, and a similar "line of crosstalk cancellation" extends through the position of the right microphone. The ringing frequency is now present as a ripple behind the main wavefront.

When the loudspeaker span is reduced even further to 10° [Fig. 4(c)], the effect of the ringing frequency is almost completely eliminated, and so the only sound pressure disturbance seen at most locations in the sound field is a single attenuated and delayed copy of the original Hanning pulse. This indicates that reducing the loudspeaker span improves the system's robustness with respect to head movement. Note, however, that very close to the two monopoles, the large low-frequency output starts to show up as a near-field effect.

Fig. 4(d) shows the sound field reproduced by a monopole–dipole source combination. This source combination avoids ringing completely, and so the reproduced sound field is very "clean." Just as for the two monopoles spanning 10°, it also contains a near-field component as expected. Note the similarity of the plots in Fig. 4(c) and (d). This means that little is gained by moving the sources closer together for this particular choice of desired signal.

2.2 Virtual Source Imaging Under Free-Field Conditions

2.2.1 Analysis

In principle it is a trivial task to create a virtual source once it is known how to implement a crosstalk cancellation system. The idea is to solve the crosstalk cancellation problem for each ear, and then to add the two solutions together. It turns out that in practice it is far easier for the loudspeakers to create the signals due to a virtual source than to achieve perfect crosstalk cancellation at one point.

The virtual source imaging problem is illustrated in Fig. 2(b). We imagine that a monopole source is positioned somewhere in the listening space. The transfer functions from this source to the listener's ears are of the same type as $C_1$ and $C_2$ [see Eqs. 3(a) and 3(b)], and they are denoted by $A_1$ and $A_2$. Consequently,

$$A_1 = \frac{W_1}{D} \tag{19a}$$

and

$$A_2 = \frac{W_2}{D}. \tag{19b}$$

As in the crosstalk cancellation case, it is convenient to normalize the desired signals $D_1$ and $D_2$ in order to ensure causality of the source inputs. In addition, we want the loudest of the two desired signals to be equal to $D$.

It must not be scaled no matter what the distance is from the virtual source to the center of the listener's head. Thus if we assume that the virtual source is positioned to the left of the listener, the desired signals are defined as

$$\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} A_1/A_2 \\ 1 \end{bmatrix} \exp(-jkr). \tag{20}$$

It is convenient to introduce a positive gain factor $g_1$ and a positive time delay $\tau_1$ such that

$$\frac{A_1}{A_2} = g_1 \exp(-j\omega\tau_1). \tag{21}$$

The definitions of $g_1$ and $\tau_1$ are equivalent to those of $g_s$ [see Eq. (1)] and $\tau_s$ [see Eq. (2)]. Using $g_1$ and $\tau_1$, and assuming that the sound from the virtual source propagates under free-field conditions, we can write the time-domain responses of the desired signals as

$$\begin{bmatrix} d_1(t) \\ d_2(t) \end{bmatrix} = \begin{bmatrix} g_1d(t - \tau_1) - \tau_1) \\ d(t - \tau_1) \end{bmatrix}. \tag{22}$$

This corresponds to two positive pulses: the first one is picked up at the left ear with an amplitude of 1, the second occurs at the right ear $\tau_1$ seconds later with an amplitude of $g_1$, which is slightly smaller than 1. As in the crosstalk cancellation case, the source inputs can be calculated by solving $C_2 = d$ for $v$, and the time-domain responses can then be determined by taking the inverse Fourier transform. The result is that each source input is now the convolution of $d(t)$ with the sum of two decaying trains of delta functions, one positive and one negative.

This happens because the sources have to reproduce two positive pulses rather than just one. Consider the expressions for $v_1(t)$ and $v_2(t)$ given by Eq. (15). The two source inputs are very similar, but they have opposite signs, $v_1(t)$ is negative and $v_2(t)$ is positive. Together, they produce a pulse of amplitude 1 at the listener's left ear. The inputs necessary to reproduce a pulse of magnitude $g_1$ at the listener's right ear will also be very similar, but the roles of $v_1(t)$ and $v_2(t)$ are reversed. Consequently, $v_2(t)$ is negative and $v_1(t)$ is positive. This means that the resulting expression for $v_1(t)$ will contain both a positive and a negative part, and so will $v_2(t)$.

What happens, then, is that the positive part of $v_1(t)$ combined with the negative part of $v_2(t)$ produces the pulse at the listener's left ear, whereas the negative part of $v_1(t)$ combined with the positive part of $v_2(t)$ produces the pulse at the listener's right ear. This is illustrated in Fig. 5.

2.2.2 Source Inputs

Fig. 5 shows the source inputs equivalent to those plotted in Fig. 3 (three different loudspeaker spans: 60, 20, and 10°), but for a virtual source imaging system rather than a crosstalk cancellation system. The virtual source is positioned at $(0.5 \, \text{m}, \, 0 \, \text{m})$, which means that it is at an angle of 45° to the left relative to straight front.
as seen by the listener. When $\theta$ is 60° [Fig. 5(a)], both
the positive and the negative pulse trains can be seen
clearly in $v_1(t)$ and $v_2(t)$. However, as $\theta$ is reduced to
20° [Fig. 5(b)], the positive and negative pulse trains
start to cancel out. This is even more evident when $\theta$ is
10° [Fig. 5(c)]. In this case the two source inputs look
roughly like square pulses of relatively short duration
(this duration is mainly determined by $\tau_s$). The advantage
of the canceling that occurs when the positive and
negative parts of the pulse trains combine is that it
greatly reduces the low-frequency content of the source
inputs, and this is why virtual source imaging systems
in practice are much easier to implement than crosstalk
cancellation systems. Even so, low frequencies are still
boosted for virtual images at positions well outside the
angle spanned by the two loudspeakers.

### 2.2.3 Reproduced Sound Field

Fig. 6 shows another four sets of nine “snapshots” of
the reproduced sound field. It is equivalent to Fig. 4,
but for a virtual source at (0.5 m, 0 m) (indicated in the
bottom right-hand corner of each frame) rather than for a
crosstalk cancellation system. As in Fig. 4, the plots
demonstrate that the reproduced sound field becomes
simpler as the loudspeaker span is reduced. In the limit
[Fig. 6(d)] there is no ringing and only the two pulses
corresponding to the desired signals are seen in the
sound field.

In conclusion, the reproduced sound field will be simi-
lar to that produced by a monopole–dipole combination
as long as the highest frequency component in the de-
sired signal is significantly smaller than the ringing fre-
cquency $f_0$. The ringing frequency can be increased even
further by reducing the loudspeaker span $\theta$, but this will
also increase the low-frequency contents of $V_1$ and $V_2$,
and this is undesirable in practice.

### 3 PRACTICAL IMPLEMENTATION

When the free-field transfer functions are replaced by
more realistic head-related transfer functions (HRTFs),
it becomes necessary to consider the problem of in-
vverting an ill-conditioned system that contains
minimum-phase components. We have developed a
number of digital filter design methods [11], [20]–[23]
that are appropriate for this purpose. All of these meth-
ods determine a matrix of digital finite impulse response
(FIR) filters that are optimal in a quantifiable sense.

At low frequencies, the crosstalk cancellation problem
is ill-conditioned. Consequently, each of the filters in
the crosstalk cancellation network is likely to boost low
frequencies by 30 dB or more. When the outputs from
these filters are added together, most of the low-frequency
energy ought to cancel out and form a set of loudspeaker
input signals with relatively well-behaved frequency re-
 sponses (dynamic range 15 dB or less). This can only
happen if the digital signal processing does not introduce
any rounding or truncation errors, and if no uncorrelated
low-frequency noise is introduced.

It is very important that the two loudspeakers have al-
most identical frequency responses (not just their ampi-
tude responses, but also their phase responses must be
the same). As a rule of thumb, “pair matching” to within
±0.5 dB in amplitude and ±5° in phase is more than
sufficient to ensure accurate and symmetric imaging.

### 4 CONCLUSIONS

When only two loudspeakers placed symmetrically in
front of a single listener are used to control the sound
field around the head of the listener, the area over which
the sound field can be controlled is larger when the two
loudspeakers are close together than when they are far
apart. In addition, moving the loudspeakers closer to-
gether helps to suppress the “ringing frequency” and
its harmonics.

As the loudspeaker span is reduced, an increasingly
large boost of low frequencies is required in order to
create a virtual image at a position well outside the angle
spanned by the loudspeakers. In practice, a loudspeaker
span of $10^\circ$ is a good choice; it ensures that the ringing frequency is above 10 kHz without boosting low frequencies excessively.

5 REFERENCES


---

**THE AUTHORS**

Ole Kirkeby received an M.Sc. in electric engineering from the Technical University of Denmark in 1989, and a Ph.D. from the Institute of Sound and Vibration Research, University of Southampton, UK, in 1995. From 1989 to 1991 he was a technical writer at Bruel & Kjaer's headquarters outside Copenhagen. After the completion of his Ph.D., he spent eight months working with Professor Hareo Hamada at the Acoustics Laboratory, Tokyo Denki University, Japan. Dr. Kirkeby's research has concentrated on developing digital signal processing methods for multichannel sound reproduction.

Philip Nelson graduated from the University of Southampton in 1974 with first class honours in mechanical engineering. He went on to study the aerodynamic production of sound for the degree of Ph.D. at the Institute of Sound and Vibration Research at the University of Southampton. From 1978 he worked for four years in industry on a wide range of practical problems in noise and vibration control. He was appointed lecturer in ISVR in 1982, and was promoted to senior lecturer in 1988 and professor in 1994. Dr. Nelson's main research interests are in the fields of acoustics, vibrations, and signal processing and he is author or coauthor of over 150 technical publications, including 2 books and over 50 papers in refereed journals. Many of his publications deal with the active control of sound and vibration.

Hareo Hamada was born in Kagoshima, Japan, in 1954. He received the M.Eng. degree in 1980 and the Dr.Eng. degree in 1983, both from Tokyo Denki University. Since 1985, he has been an associate professor at Tokyo Denki University. His research activities are in the application of digital signal processing to the field of acoustics. He received the Satomedal from the Acoustical Society of Japan in 1988 for his paper on "Active Control of Sound." His current field of interest relates to binaural signal processing and adaptive signal processing as applied to the active control of sound and vibration.