

Optimal Source Distribution for Virtual Acoustic Imaging

T. Takeuchi and P. A. Nelson

ISVR Technical Report No 288

February 2000



SCIENTIFIC PUBLICATIONS BY THE ISVR

Technical Reports are published to promote timely dissemination of research results by ISVR personnel. This medium permits more detailed presentation than is usually acceptable for scientific journals. Responsibility for both the content and any opinions expressed rests entirely with the author(s).

Technical Memoranda are produced to enable the early or preliminary release of information by ISVR personnel where such release is deemed to be appropriate. Information contained in these memoranda may be incomplete, or form part of a continuing programme; this should be borne in mind when using or quoting from these documents.

Contract Reports are produced to record the results of scientific work carried out for sponsors, under contract. The ISVR treats these reports as confidential to sponsors and does not make them available for general circulation. Individual sponsors may, however, authorize subsequent release of the material.

COPYRIGHT NOTICE

(c) ISVR University of Southampton All rights reserved.

ISVR authorises you to view and download the Materials at this Web site ("Site") only for your personal, non-commercial use. This authorization is not a transfer of title in the Materials and copies of the Materials and is subject to the following restrictions: 1) you must retain, on all copies of the Materials downloaded, all copyright and other proprietary notices contained in the Materials; 2) you may not modify the Materials in any way or reproduce or publicly display, perform, or distribute or otherwise use them for any public or commercial purpose; and 3) you must not transfer the Materials to any other person unless you give them notice of, and they agree to accept, the obligations arising under these terms and conditions of use. You agree to abide by all additional restrictions displayed on the Site as it may be updated from time to time. This Site, including all Materials, is protected by worldwide copyright laws and treaty provisions. You agree to comply with all copyright laws worldwide in your use of this Site and to prevent any unauthorised copying of the Materials.

UNIVERSITY OF SOUTHAMPTON
INSTITUTE OF SOUND AND VIBRATION RESEARCH
FLUID DYNAMICS AND ACOUSTICS GROUP

Optimal Source Distribution for Virtual Acoustic Imaging

by

T Takeuchi and P A Nelson

ISVR Technical Report No. 288

February 2000

Authorized for issue by
Professor P A Nelson
Group Chairman

© Institute of Sound and Vibration Research

Acknowledgements

We would like to thank Dr Keith Holland for helpful discussions. This research was supported by Yamaha Corporation, Alpine Electronics and Hitachi Ltd. T. Takeuchi is supported by Kajima Corporation.

Contents

ACKNOWLEDGEMENTS	ii
ABSTRACT	iv
1 INTRODUCTION.....	1
1.1 INTRODUCTION	1
1.2 PRINCIPLES OF MULTI-CHANNEL SOUND CONTROL WITH SYSTEM INVERSION	2
2 FUNDAMENTAL PROBLEMS OF EXISTING SYSTEMS	4
2.1 INVERSE FILTER MATRIX	4
2.2 LOSS OF DYNAMIC RANGE	7
2.3 ROBUSTNESS TO ERROR IN THE PLANT AND THE INVERSE FILTERS	8
2.4 ROBUSTNESS TO REFLECTIONS	10
2.5 REGULARISATION	10
3 A SYSTEM TO OVERCOME THE PROBLEMS.....	12
3.1 PRINCIPLE OF THE PROPOSED SYSTEM	12
3.1.1 <i>Principle of the Optimal Source Distribution ("OSD" system)</i>	12
3.1.2 <i>Aspects of the proposed system</i>	14
3.2 PRACTICAL DISCRETE SYSTEM.....	15
3.2.1 <i>Frequency range and span for discretised transducer pairs</i>	16
3.2.2 <i>Consequence of the discretisation of variable source span</i>	17
3.2.3 <i>Considerations for the low frequency region</i>	18
3.3 EXAMPLES OF DISCRETE (MULTI-WAY) "OSD" SYSTEM.....	19
3.3.1 <i>"3-way" systems and more</i>	19
3.3.2 <i>"2-way" systems</i>	21
3.3.3 <i>"1-way" systems</i>	22
3.4 COMMENTS ON MULTI-REGION SYSTEMS.....	23
3.5 CONSIDERATIONS FOR CROSS-OVER FILTERS AND INVERSE FILTERS	24
3.5.1 <i>"2 by 2" plant matrix</i>	24
3.5.2 <i>Multiple "2 by 2" plant matrix</i>	24
3.5.3 <i>"2 by (2 × multiple)" plant matrix</i>	25
3.5.4 <i>Type of filters</i>	25
3.6 COMMENTS ON MULTI-CHANNEL SYSTEMS.....	26
4. SUMMARY	28
REFERENCES	28
APPENDIX.....	31

Abstract

When binaural sound signals are presented with loudspeakers, the system inversion involved gives rise to number of problems such as, for example, loss of dynamic range and a lack of robustness to small errors in control performance. The amplification required by the system inversion results in loss of dynamic range. The inverse filters obtained are likely to contain large errors around certain frequencies where the inversion is ill-conditioned. Regularisation is often used to design practical filters but this also results in poor control performance around the frequencies at which ill-conditioning occurs.

These problems for such systems are investigated in relation to the angle spanned by the sources and this has resulted in the proposal of a new system, the Optimal Source Distribution (or “OSD” system), which overcomes these problems by means of variable transducer span. A practical solution to realise a variable transducer span by discretisation is also described. The discretisation expands the frequency region to be used but the adverse effect is small since the plant transfer function matrix is well-conditioned over a wide frequency range around the ideal frequency. The “OSD” system makes use of multiple pairs of transducers for each frequency range with corresponding transducer spans in order to eventually cover the whole audible frequency range. Several examples of the multi-way “OSD” system are demonstrated which in practice produce a very robust system over the whole frequency range. The relationship to the “Stereo Dipole” system is also described.

1 INTRODUCTION

1.1 Introduction

Binaural technology [1]-[3] is often used to present a virtual acoustic environment to a listener. The principle of this technology is to control the sound field at the listener's ears so that the reproduced sound field coincides with what would be produced when he is in the desired real sound field. One way of achieving this is to use a pair of loudspeakers (electro-acoustic transducers) at different positions in a listening space with the help of signal processing to ensure that appropriate binaural signals are obtained at the listener's ears. [4]-[8]

One of the objectives of this report is to investigate a number of problems which arise from the multi-channel system inversion involved in such as binaural synthesis over loudspeakers. A basic analysis with a free field transfer function model illustrates the fundamental difficulties which such systems can have. The amplification required by the system inversion results in loss of dynamic range. The inverse filters obtained are likely to contain large errors around ill-conditioned frequencies. Regularisation is often used to design practical filters but this also results in poor control performance around ill-conditioned frequencies. Further analysis with a more realistic plant matrix, where the sound signals are controlled at a listener's ears in the presence of the listener's body (pinnae, head...), demonstrates that this is still the case.

A new concept of sound control system, referred to as the Optimal Source Distribution (or "OSD" system), which overcomes these fundamental problems is proposed in this report. The practical system realised is very robust over a wide frequency range (e.g. the whole audible

frequency range) and this is described with many examples as well as the underlying theoretical principle.

1.2 Principles of multi-channel sound control with system inversion

System inversion is often used for multi-channel sound control. The principle of such systems is described below with 2-channel binaural reproduction over loudspeakers as an example for convenience in later analysis and is illustrated in Fig. 1. Independent control of two signals (such as binaural sound signals) at two points (such as the ears of a listener) can be achieved with two electro-acoustic transducers (such as loudspeakers), by filtering the input signals to the transducers with the inverse of the transfer function matrix of the plant. The signals and transfer functions involved are defined as follows. Two monopole transducers produce source strengths defined by the elements of the complex vector $\mathbf{v} = [v_1(j\omega) \ v_2(j\omega)]^T$. The resulting acoustic pressure signals are given by the elements of the vector $\mathbf{w} = [w_1(j\omega) \ w_2(j\omega)]^T$. This is given by

$$\mathbf{w} = \mathbf{C}\mathbf{v} \tag{1}$$

where \mathbf{C} is a matrix of transfer functions between sources and receivers. The two signals to be synthesised at the receivers are defined by the elements of the complex vector $\mathbf{d} = [d_1(j\omega) \ d_2(j\omega)]^T$. In the case of audio applications, these signals are usually the signals that would produce a desired virtual auditory sensation when fed to the ears (Fig. 1). They can be obtained, for example, by recording sound source signals \mathbf{u} with a recording head or filtering signals \mathbf{u} by matrix of

synthesised binaural filters \mathbf{A} . Therefore, a filter matrix \mathbf{H} which contains inverse filters is introduced so that $\mathbf{v} = \mathbf{H}\mathbf{d}$ where

$$\mathbf{H} = \begin{bmatrix} H_{11}(j\omega) & H_{12}(j\omega) \\ H_{21}(j\omega) & H_{22}(j\omega) \end{bmatrix} \quad (2)$$

and thus

$$\mathbf{w} = \mathbf{C}\mathbf{H}\mathbf{d} \quad (3)$$

For convenience in later analysis, we also define the control performance matrix \mathbf{R} given by

$$\mathbf{R} = \mathbf{C}\mathbf{H} \quad (4)$$

The filter matrix \mathbf{H} can be designed so that the vector \mathbf{w} is a good approximation to the vector \mathbf{d} with a certain delay. [9][10]

2 Fundamental problems of existing systems

The system inversion involved gives rise to a number of problems such as, for example, loss of dynamic range and sensitivity to errors. A simple case involving the control of two monopole receivers with two monopole transducers (sources) under free field conditions is first considered here. The fundamental problems with regard to system inversion can be illustrated in this simple case where the effect of path length difference dominates the problem. A matrix of Head Related Transfer Functions (HRTFs) is also analysed as an example of a more realistic plant. In such a case, the acoustic response of the human body (pinnae, head, torso and so on) also comes to affect the problem. A symmetric case with the inter-source axis parallel to the inter-receiver axis is considered for an examination of the basic properties of the system. The geometry is illustrated in Fig. 2.

2.1 Inverse filter matrix

In the free field case, the plant transfer function matrix can be modelled as

$$\mathbf{C} = \frac{\rho_0}{4\pi} \begin{bmatrix} e^{-jkl_1}/l_1 & e^{-jkl_2}/l_2 \\ e^{-jkl_2}/l_2 & e^{-jkl_1}/l_1 \end{bmatrix} \quad (5)$$

where an $e^{j\omega t}$ time dependence is assumed with $k = \omega/c_0$, and where ρ_0 and c_0 are the density and sound speed. When the ratio of and the difference between the path lengths connecting one source and two receivers are defined as $g = l_1/l_2$ and $\Delta l = l_2 - l_1$,

$$\mathbf{C} = \frac{\rho_0 e^{-jk l_1}}{4\pi l_1} \begin{bmatrix} 1 & g e^{-jk \Delta l} \\ g e^{-jk \Delta l} & 1 \end{bmatrix} \quad (6)$$

Now consider the case

$$\mathbf{d} = \frac{\rho_0 e^{-jk l_1}}{4\pi l_1} \begin{bmatrix} D_1(j\omega) \\ D_2(j\omega) \end{bmatrix} \quad (7)$$

i.e., the desired signals are the acoustic pressure signals which would have been produced by the closer sound source and whose values are either $D_1(j\omega)$ or $D_2(j\omega)$ without disturbance due to the other source (cross-talk). This enables a description of the effect of system inversion as well as ensuring a causal solution. The elements of \mathbf{H} can be obtained from the exact inverse of \mathbf{C} and can be written as

$$\mathbf{H} = \mathbf{C}^{-1} = \frac{1}{1 - g^2 e^{-2jk \Delta l}} \begin{bmatrix} 1 & -g e^{-jk \Delta l} \\ -g e^{-jk \Delta l} & 1 \end{bmatrix} \quad (8)$$

When $l \gg \Delta r$, we have the approximation $\Delta l \approx \Delta r \sin \theta$ where 2θ is the source span (hence $0 < \theta \leq (\pi/2)$) and under these conditions,

$$\mathbf{H} = \frac{1}{1 - g^2 e^{-2jk\Delta r \sin \theta}} \begin{bmatrix} 1 & -ge^{-jk\Delta r \sin \theta} \\ -ge^{-jk\Delta r \sin \theta} & 1 \end{bmatrix} \quad (9)$$

The magnitude of the elements of \mathbf{H} ($|H_{mn}(j\omega)|$) show the necessary amplification of the desired signals produced by each inverse filter in \mathbf{H} . The maximum amplification of the source strengths can be found from the 2-norm of \mathbf{H} ($\|\mathbf{H}\|$) which is the largest of the singular values of \mathbf{H} , where these singular values are denoted by σ_o and σ_i (see the Appendix). Thus

$$\|\mathbf{H}\| = \max(\sigma_o, \sigma_i) \quad (10)$$

$$\text{where } \sigma_o = \frac{1}{\sqrt{(1 - ge^{-jk\Delta r \sin \theta})(1 - ge^{jk\Delta r \sin \theta})}} \text{ and } \sigma_i = \frac{1}{\sqrt{(1 + ge^{-jk\Delta r \sin \theta})(1 + ge^{jk\Delta r \sin \theta})}}$$

σ_o and σ_i correspond to orthogonal components of the inverse filters. σ_o corresponds to the amplification factor of the out-of-phase component of the desired signals and σ_i corresponds to the amplification factor of the in-phase component of the desired signals. Plots of σ_o , σ_i and $\|\mathbf{H}\|$ with respect to $k\Delta r \sin \theta$ are illustrated in Fig. 3. The examples throughout this paper uses typical value of the distance between the adult human ears for Δr . As seen in Eq. (10) and Fig. 3, $\|\mathbf{H}\|$ changes periodically and has peaks where k and θ satisfy the following relationship with even values of the integer number n .

$$k\Delta r \sin \theta = \frac{n\pi}{2} \quad (11)$$

The singular value σ_o has peaks at $n = 0, 4, 8, \dots$ where the system has difficulty in reproducing the out-of-phase component of the desired signals and σ_i has peaks at $n = 2, 6, 10, \dots$ where the system has difficulty in reproducing the in-phase component.

2.2 Loss of dynamic range

In practice, since the maximum source output is given by $\|\mathbf{H}\|_{\max}$, this must be within the range of the system in order to avoid clipping of the signals. The required amplification results directly in the loss of dynamic range illustrated in Fig. 4. The level of the output source signal (\mathbf{v}) and the resulting level of the acoustic pressure (\mathbf{w}) are plotted both with and without system inversion assuming that the maximum output level and dynamic range of the system are the same. The given dynamic range is distributed into the system inversion and the remaining dynamic range which is to be used by the binaural auditory space synthesis, and also most importantly, by the sound source signal itself. The frequency of the peaks do not affect the amount of dynamic range loss but the magnitude of the peaks do. The dynamic range loss is defined by the difference between the signal level at the receiver with one monopole source and the signal level reproduced by two sources having the same maximum source strength when the system is inverted. Since $\|\mathbf{H}\|$ here is normalised by the case without system inversion by Eq. (7), the dynamic range loss Γ is given by

$$\Gamma = \|\mathbf{H}\|_{\max} = \frac{1}{1-g} \quad (12)$$

The dynamic range loss given by Eq. (12) as a function of source span is shown in Fig. 5. Since $g \approx 1 - \Delta r \sin \theta / l$, Γ can be approximated as

$$\Gamma \approx \frac{l}{\Delta r \sin \theta} \quad (13)$$

as a function of θ . Fig. 5 and Eq. (13) show that the larger the source span, the less is the dynamic range loss.

2.3 Robustness to error in the plant and the inverse filters

Eq. (1) implies that the system inversion (which determines \mathbf{v} and leads to the design of the filter matrix \mathbf{H}) is very sensitive to small errors in the assumed plant \mathbf{C} (which is often measured and thus small errors are inevitable) where the condition number of \mathbf{C} , $\kappa(\mathbf{C})$, is large. In addition, since

$$\mathbf{v} = \mathbf{C}^{-1} \mathbf{w} \quad (14)$$

and $\kappa(\mathbf{C}^{-1}) = \kappa(\mathbf{C})$, the reproduced signals \mathbf{w} are less robust to small changes in the inverse of the plant matrix \mathbf{C}^{-1} , hence \mathbf{H} , where $\kappa(\mathbf{C})$ is large.

The condition number of \mathbf{C} is given by

$$\kappa(\mathbf{C}) = \|\mathbf{C}\| \|\mathbf{C}^{-1}\| = \|\mathbf{C}\| \|\mathbf{H}\| = \|\mathbf{H}^{-1}\| \|\mathbf{H}\| = \max \left(\sqrt{\frac{(1 + ge^{-jk\Delta r \sin \theta})(1 + ge^{jk\Delta r \sin \theta})}{(1 - ge^{-jk\Delta r \sin \theta})(1 - ge^{jk\Delta r \sin \theta})}}, \sqrt{\frac{(1 - ge^{-jk\Delta r \sin \theta})(1 - ge^{jk\Delta r \sin \theta})}{(1 + ge^{-jk\Delta r \sin \theta})(1 + ge^{jk\Delta r \sin \theta})}} \right) \quad (15)$$

and is shown in Fig. 6. As seen in Eq. (15) and Fig. 6, $\kappa(\mathbf{C})$ has peaks where Eq. (11) is satisfied with an even value of the integer number n . The frequencies which give peaks of $\kappa(\mathbf{C})$ are consistent with those which give the peaks of $\|\mathbf{H}\|$.

Around the frequencies where $\kappa(\mathbf{C})$ is large, the system is very sensitive to small errors in \mathbf{C} and \mathbf{H} . The calculated inverse filter matrix \mathbf{H} is likely to contain large errors due to small errors in \mathbf{C} and results in large errors in the reproduced signal \mathbf{w} at the receiver. Even if \mathbf{C} does not contain any errors, the reproduction of the signals at the receiver is too sensitive to the small errors in the inverse filter matrix \mathbf{H} to be useful. On the contrary, $\kappa(\mathbf{C})$ is small around the frequencies where n is an odd integer number in Eq. (11). Around these frequencies, a practical and close to ideal inverse filter matrix \mathbf{H} is easily obtained. For the same value of n , the robust frequency range becomes lower as the source span becomes larger. With a logarithmic frequency scale, which is related to the perceptual attributes of the human auditory system, the frequency range of robust inversion is more or less constant for different source spans for the same value of n , even though it looks wider for smaller source spans on a linear frequency scale.

2.4 Robustness to reflections

The amplification by the inverse filter also results in severe reflection [11]. Fig. 7 shows an example of far field sound pressure level produced by the control sources with reference to the receiver directions. The horizontal axis is the inter-source axis and the receivers (ears) are at the directions of the vertical axis. At frequencies where Eq. (11) is satisfied with an even value of the integer number n , as in this example, the sound pressure levels at directions other than receiver directions can be significantly larger (typically +30dB ~ 40dB) than those at the receiver directions (0dB and -∞dB). The amount of this excessive radiation is the same as the amount of dynamic range loss as in Eq. (12) and Fig. 5. When the environment is not anechoic, as is normally the case, this obviously results in severe reflections and the control performance of the system is deteriorates severely.

2.5 Regularisation

It is possible to reduce the excess amplification and hence the dynamic range loss by means of regularisation, where the pseudo inverse filter matrix \mathbf{H} is given by

$$\mathbf{H} = [\mathbf{C}^H \mathbf{C} + \beta \mathbf{I}]^{-1} \mathbf{C}^H \quad (16)$$

where β is a regularisation parameter. The regularisation parameter penalises large values of \mathbf{H} and hence limits the dynamic range loss of the system. Since $\|\mathbf{H}\|$ is normalised by the case without system inversion by Eq. (7), the regularisation parameter limits the dynamic range loss to less than about

$$\Gamma \approx -10 \log_{10} \beta - 6 \text{ (dB)}$$

(17)

However, the regularisation parameter intentionally, hence inevitably, introduces a small error in the inversion process. This gives rise to a problem for filter design for frequencies where $\kappa(\mathbf{C})$ is large. An example of this is illustrated in Fig. 8. The dynamic range loss is reduced by regularisation from about 27dB (without regularisation) as in Fig. 8a to 14dB as shown in Fig. 8b ($\beta = 10^{-2}$). However, it can be clearly seen that the control performance of the system deteriorates around the frequencies where n is an even integer number in Eq. (11). The contribution of the correct desired signals (R_{11} and R_{22}) is reduced only slightly but the contribution of the wrong desired signals (R_{12} and R_{21} , the cross-talk component) is increased significantly. In other words, the system has little control (cross-talk cancellation) around these frequencies. This problem is significant at lower frequencies ($n < 1$ in Eq. (11)) in the sense that the region without cross-talk suppression is large, and at higher frequencies ($n > 1$ in Eq. (11)), in the sense that there are many frequencies at which the plant is ill-conditioned. With an equivalent dynamic range loss, making the source span larger leads to a better control performance at lower frequencies but a poorer performance at higher frequencies (Fig. 9a). On the contrary, making the source span smaller leads to better control performance at higher frequencies but poorer performance at lower frequencies (Fig. 9b).

3 A system to overcome the problems.

As discussed above, there is a trade-off between dynamic range, robustness and control performance. However, a system which aims to overcome these fundamental problems is proposed in what follows.

3.1 Principle of the proposed system

3.1.1 Principle of the Optimal Source Distribution (“OSD” system)

Eq. (11) can be rewritten in terms of the source span 2θ as

$$2\theta = 2 \arcsin\left(\frac{n\pi}{2k\Delta r}\right) \quad (18)$$

As seen from the analysis above, systems with the source span where n is an odd integer number in Eq. (18) give the best control performance as well as robustness. This implies the optimal source span must vary as a function of frequency.

We now consider a pair of conceptual monopole transducers whose span varies continuously as a function of frequency in order to satisfy the requirement for n to be an odd integer number in Eq. (18). This is illustrated in Fig. 10. The source span becomes smaller as frequency becomes higher. With this assumption, Eq. (9) becomes

$$\mathbf{H} = \frac{1}{1+g^2} \begin{bmatrix} 1 & -jg \\ -jg & 1 \end{bmatrix} \quad (19)$$

Note that $\|\mathbf{H}\| = 1/\sqrt{2}$ and $\kappa(\mathbf{C}) = 1$ for all frequencies. Therefore, there is no dynamic range loss compared to the case without system inversion. In fact, there is a dynamic range gain of 3dB since the two orthogonal components of the desired signals are $\pi/2$ out of phase. The error in calculating the inverse filter is small and the system has very good control over the reproduced signals. The system is also very robust to the changes in plant matrix. The far field sound pressure levels produced by the transducer pair at the all directions are always smaller than those at the receiver directions. An example when $n = 1$ are shown in Fig. 11. Therefore, it is also robust to reflections in a reverberant environment. In practice, the directivity of transducers helps to reduce the effect of reflection further.

Also note that when $l \gg \Delta r$, $g \approx 1$ so

$$\mathbf{H} \approx \frac{1}{2} \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix} \quad (20)$$

This implies that independent control of the two signals is nearly achieved just by addition of the desired signals with a $\pi/2$ relative phase shift between them.

This principle requires a pair of monopole type transducers whose position varies continuously as frequency varies. This might, for example, be realised by exciting a triangular

shaped plate whose width varies along its length. The requirement of such a transducer is that a certain frequency of vibration is excited most at a particular position having a certain width such that sound of that frequency is radiated mostly from that position (Fig. 12).

3.1.2 Aspects of the proposed system

From Eq. (18), the range of source span is given by the frequency range of interest as can be seen from Fig. 10. A smaller value of n gives a smaller source span for the same frequency. Therefore, the smallest source span θ_h for the same high frequency limit is given by $n = 1$ and this is about 4° to give control of the sound field at two positions separated by the distance between two ears (about 0.13m for KEMAR dummy head) up to a frequency of 20kHz.

Eq. (11) can also be rewritten in terms of frequency as

$$f = \frac{nc_0}{4\Delta r \sin \theta} \quad (21)$$

The smallest value of n gives the lowest frequency limit for a given source span. Since $\sin \theta \leq 1$,

$$f \geq \frac{nc_0}{4\Delta r} \quad (22)$$

i.e., the physically maximum source span of $2\theta = 180^\circ$ gives the low frequency limit, f_l , associated with this principle. A smaller value of n gives a lower low frequency limit so the system given by n

$= 1$ is normally the most useful among those with an odd integer number n . The low frequency limit given by $n = 1$ of a system designed to control the sound field at two positions separated by the distance between two ears is about $f_l = 300 \sim 400$ Hz.

3.2 Practical discrete system

In practice, a pair of conceptual monopole transducers whose span varies continuously as a function of frequency is currently not available. However, it is possible to realise a practical system based on this principle by discretising the transducer span as illustrated in Fig. 13. With a given span, the frequency region where the amplification is relatively small and plant matrix \mathbf{C} is well conditioned is relatively wide around the optimal frequency. Therefore, by allowing n to have some width, say $\pm \nu$ ($0 < \nu < 1$), which results in a small amount of dynamic range loss and slightly reduced robustness, a certain transducer span can nevertheless be allocated to cover a certain range of frequencies where control performance and robustness of the system is still reasonably good. Consequently, it is possible to discretise the continuously varying transducer span into a finite number of transducer spans. Such a practical system can also be interpreted as making use of better-conditioned frequencies only and excluding ill-conditioned frequencies by limiting the frequency range to be used for a certain transducer span. By making use of different transducer spans for different frequency ranges, it is possible to construct a practical system which can cover a wide frequency range (most of the audible frequency range in fact) with a few sets of pairs of transducers with different transducer spans.

This principle is extremely useful and practical because a single transducer which can cover the whole audible frequency range is not practically available either. Therefore, this principle also

gives the ideal background for multi-way systems for binaural reproduction over loudspeakers which maximise the frequency range to be covered. It should be noted that this is still a simple “2 channel” control system where only two independent control signals are necessary to control any form of virtual auditory space. This in principle can synthesise an infinite number of virtual source locations with different source signals with any type of acoustic response of the space. The difference from the conventional 2 channel system is that the two control signals are divided into multiple frequency bands and fed into the different pairs of driver units with different spans.

3.2.1 Frequency range and span for discretised transducer pairs.

The condition number $\kappa(C)$ of the plant matrix plotted as a function of frequency and source span is shown in Fig. 14 for the audible frequency range (20Hz ~ 20kHz). It is important to design the system to ensure a condition number that is as small as possible over a frequency range that is as wide as possible. Therefore, the transducer spans for each pair of transducers in each frequency range can be decided to ensure that the smallest possible values of ν are used over the all frequency range of interest above f_l (See 3.2.2)

Fig. 15 shows the condition number of the more realistic plant matrix with HRTFs. The HRTFs were measured with the KEMAR dummy head at MIT Media Lab [12] and the loudspeaker response was deconvolved later. A similar trend can clearly be seen as in the free field case. However, additional “ill-conditioned frequencies” can be observed around 9kHz and 13kHz where the HRTFs have minima. It is possible that the signal to noise ratio of the data around these frequencies is poor. It should also be noted that where the incidence angle θ is small, the peak frequencies obtained with the HRTF plant matrix are similar to that of the free field plant

with the receiver distance $\Delta r \approx 0.13$. This corresponds to the shortest distance between the entrances of the ear canals of the KEMAR dummy head. However, where the incidence angle θ is large, the peak frequencies obtained with the HRTF plant matrix are similar to that of the free field plant with the receiver distance $\Delta r \approx 0.25$. This is a much larger distance than the shortest distance between the entrances of the ear canals of the KEMAR dummy head and is probably a result of diffraction around the head.

Fig. 16 shows the dynamic range loss as a function of frequency and source span. It is also possible to discretise, i.e., decide the transducer spans and frequency ranges to be covered by each pair of driver units (i.e. range of n), in terms of a tolerable dynamic range loss. The dynamic range loss of the entire system is now given by the maximum value among the values given by each discretised transducer span.

3.2.2 Consequence of the discretisation of variable source span

It should be noted that the low frequency limit f_l given by odd integer numbers n in Eq. (22) is extended towards a lower frequency by discretisation because now the region for frequency and transducer span where n is not an integer number is also used. For example, a practical system discretised from the ideal system with $n = 1$ can now make use of the region $1 - \nu < n < 1 + \nu$ so that the low frequency limit is given by $n = 1 - \nu$.

As can be seen from Fig. 10, in the higher frequency range where the source span is very small, the frequency range to be covered is very sensitive to small differences in transducer span. On the contrary, it is very insensitive to the source span at lower frequencies. Consequently, the

range of practical span for the low frequency units is very large, which can practically be anywhere from 60° to 180° with only a very slight increase of f_l .

A system with a smaller n gives a wider region with the same performance on a logarithmic scale as can be seen in Fig. 14 ~ Fig. 16.

3.2.3 Considerations for the low frequency region

At the frequencies below f_l ($n < 1-\nu$) where $\kappa(\mathbf{C})$ is larger than other frequencies, robustness of the system and the requirement for dynamic range loss are more severe than at other frequencies. When f_l is reasonably low, where interaural difference is not crucial for binaural reproduction, one can avoid system inversion and simply add a single sub-woofer unit for this frequency region to avoid the extra dynamic range loss required by this region.

It is also possible to cover this sub-low frequency region with the lowest frequency pair of units by making use of regularisation to limit the amplification, and hence without too much dynamic range loss, without sacrificing robustness for other frequencies. The robustness to errors and cross-talk performance with regularisation in the frequency range below f_l is not as good as the other frequencies as a result of the ill-conditioning of the plant matrix \mathbf{C} . However, there can still be reasonable cross-talk suppression available.

The cross-talk cancellation performance in this region is very sensitive to the allocated dynamic range loss. If less dynamic range loss is allowed, a larger regularisation parameter is needed to suppress the amplitude of the inverse filter, and this results in cross-talk. Therefore, it is possible to design the system by selecting the required low frequency cross-talk cancellation

performance. As an example, Fig. 17 illustrates the cross-talk cancellation performance as a function of frequency and source span when 20dB dynamic range loss is allocated for system inversion. When more dynamic range loss is allowed, the greater is the cross-talk cancellation performance obtained for the whole frequency/span region.

When large dynamic range can not be allocated to system inversion, a large value of regularisation parameter is necessary. Even if reasonable cross-talk suppression is not available, the low frequency pair can still work as a sub-woofer. In this case, although the control performance deteriorates severely, $\|\mathbf{R}\|$ and hence the norm of the reproduced signal is the same as that without regularisation. This means that although the system has difficulty in reproducing the out-of-phase component of the desired signal, it still can produce the in-phase component as well as before. This is beneficial in binaural reproduction since the difference between the two desired signals are normally not so large and sometimes negligible in the very low frequency range.

3.3 Examples of discrete (multi-way) “OSD” system

3.3.1 “3-way” systems and more

Examples of 3-way systems with $0 < n < 2$ are illustrated in Fig. 18. These examples aim to ensure a condition number that is as small as possible over a frequency range that is as wide as possible. Therefore, the transducer spans (2θ) for the high frequency units and the low frequency units were chosen at two extreme positions which gives $\nu = 0.7$. A pair of high frequency units spanning 6.2° is chosen to cover the frequency range up to 20kHz ($n = 1.7$) while a pair of low frequency units spanning 180° is chosen to cover as low a frequency as possible. The span for the

mid frequency units is 32° . The dynamic range loss of about 7dB can be achieved with 3 pairs of units (Fig. 19). This arrangement gives $f_l \approx 110\text{Hz}$ ($n = 0.3$ with low frequency pair) and a sub-woofer may be added to deal with the range below this frequency. The cross-over frequencies are given by $n = 0.3$ and $n = 1.7$ ($\nu = 0.7$) for each pair of units and at around 600Hz and 4kHz. The far field sound pressure level produced by the transducer pairs becomes maximum at the cross-over frequencies and is shown in Fig. 20. Note that those of around the middle of the frequency range for each transducer pair are as shown in Fig. 11 .

By limiting the amplification of the low frequency pair for frequencies below f_l to 7 dB with regularisation, the low frequency units can also cover frequencies down to about 100Hz with reasonable cross-talk cancellation of more than 20dB and cover below 100Hz with reduced interaural difference.(Fig. 21).

When more dynamic range loss is allowed, it is possible to use smaller regularisation parameters hence low frequency cross-talk performance improves (Fig. 22). By allowing dynamic range loss of 13dB, the low frequency units spanning 180° can cover frequencies down to 20 Hz with more than 20dB cross-talk suppression.

Alternatively, it is possible to use a smaller ν , i.e., transducer spans to improve the robustness of the system in the higher frequency range at the expense of the low frequency cross-talk performance, there being plenty to spare in the previous example (Fig. 22). An example of this strategy is described in the following section for “2-way” systems.

As the variable transducer span is discretised more finely, e.g., by using 4-way or 5-way systems and so on, the smaller the width of $n (\pm \nu)$ becomes. Hence, the system becomes more robust at frequencies above f_l . However, the performance gain becomes smaller and smaller as the number of driver units is increased. Obviously, the finer the discretisation, the closer the design is to the principle of the continuously variable transducer span. However, the number of driver pairs increases and hence the trade-off between performance gain and cost becomes more significant.

3.3.2 “2-way” systems

An example of a 2-way system with $0 < n < 2$ is illustrated in Fig. 23 and Fig. 24. This example is again designed to ensure small condition numbers over a wide frequency range so the transducer spans were chosen at 6.9° and 120° which gives $\nu \approx 0.9$. A dynamic range loss of about 18dB can be achieved with only 2 pairs of units without regularisation. A pair of mid-high frequency units spanning 6.9° is used to cover the frequency range up to 20kHz ($n = 1.9$) while a pair of mid-low frequency units spanning 120° gives a value of f_l of about 20Hz ($n = 0.1$ with low frequency pair). The cross-over frequency is given by $n = 0.1$ and $n = 1.9$ for each pair of units and is at around 900Hz. The maximum far field sound pressure level produced by the transducer pairs at the cross-over frequencies are shown in Fig. 25. Note again the sound pressures produced around the middle of the frequency range for each transducer pair are as shown in Fig. 11 .

As discretisation becomes coarser, the more frequency regions become severely ill-conditioned. It is possible to reduce transducer spans to improve robustness at higher frequencies at the expense of the low frequency cross-talk performance. Fig. 26 shows another example of a 2-way system which is obtained by omitting the pair of woofer units from the 3-way system ($\nu \approx$

0.7) described in the previous section. The dynamic range in this example is maintained to be the same as that in the previous example of the 2-way system (as in Fig. 24) by means of regularisation. The span for the high frequency units is 6.2° . The span for mid-low frequency units is 32° which also covers the frequency range below $f_l \approx 600\text{Hz}$ with a cross-talk cancellation performance of more than 20dB. The mid-low frequency pair can also cover the range below 200Hz where the cross-talk cancellation performance becomes less than 20dB. The cross-over frequency is now at around 4kHz. The conditioning above $f_l \approx 600\text{Hz}$ is as good as the 3-way system and it can be seen that the condition number becomes very small compared to the previous example illustrated in Fig. 24.

3.3.3 “1-way” systems

The coarsest discretisation is given by an example of a 1-way virtual acoustic imaging system with $0 < n < 2$ as illustrated in Fig. 27 and Fig. 28. The transducer span is 7.2° . The benefit available is very limited for a 1-way system with this principle. Since the frequency range to be covered with a single pair of transducers is the whole audible frequency range (20Hz ~ 20kHz), the width of n is nearly ± 1 ($\nu = 0.998$). The dynamic range loss is more than 40dB and very large condition numbers are notable in the wide range of low frequencies and at the high frequency end. When regularisation is used to limit the dynamic range loss to 18dB, the cross-talk cancellation performance below 1kHz is less than 20dB (Fig. 29).

This is not practical anyway since a practical single transducer which can be used over this frequency range is not available. It is possible to come to a compromise design to reduce the

width of $n (\pm \nu)$ by sacrificing the high and low frequency ranges which a practical full-range unit can not cover.

3.4 Comments on multi-region systems

It is also possible to compromise further to utilise two or more regions of n . Then there is no distinction from conventional systems. However, it is still possible to optimise their performance by utilising a similar discussion to that presented above but extending it into multiple regions of n . This approach is beneficial when one attempts to cover a wider frequency range with a smaller number of transducer pairs. The “Stereo Dipole” system [13] which has a pair of transducers spanning 10° is one such system. The simplest example with a single pair of transducers utilising the regions of $0 < n < 2$ and $2 < n < 4$ is illustrated in Fig. 30 and Fig. 31. The frequency range of 20Hz ~ 20kHz is covered with a single pair of transducers spanning 14° . The required amplification is about 40dB so the example illustrated is regularised to 18dB dynamic range loss. It can be seen that the cross-talk cancellation performance in the low frequency range is improved from the 1-way system in Fig. 29. This example shows more than 20dB cross-talk cancellation performance down to about 400Hz (which was 1kHz in Fig. 29). However, there is an additional unusable region around 10kHz ($1+\nu < n < 3-\nu$) where the system has little control and is not robust.

It is also possible to match this unusable region to the frequencies where HRTFs have minima ($\|C\|$ is small) since inversion of minima requires further amplification in \mathbf{H} and dynamic range loss. In addition, the position of minima in the higher frequency range can vary considerably

between individuals. [14] Therefore, it may not be practical to provide inversion at these frequencies where the HRTFs used for filter design have minima.

3.5 Considerations for cross-over filters and inverse filters

Cross-over filters (low pass, high pass or band pass filters) are used to distribute signals of the appropriate frequency range to the appropriate pair of driver units of the multi-way “OSD” system. Since an ideal filter which gives a rectangular window in the frequency domain can not be realised practically, there are frequency regions around the cross-over frequency where multiple pairs of driver units are contributing significantly to the synthesis of the reproduced signals w . Therefore, it is important to ensure this “cross-over region” is also within the region of this principle.

3.5.1 “2 by 2” plant matrix

If the plant matrix C is obtained when including a cross-over network as illustrated in Fig. 32, it consists of a single 2 by 2 matrix of electro-acoustic transfer functions between two outputs of the filter matrix H and two receivers which contain the responses of the cross-over networks and the interaction between different pairs of driver units around the cross-over frequency. The plant matrix C for inverse filter design can also contain the transducer responses and the acoustic response of human body and the surrounding environment. The obtained 2 by 2 inverse filter matrix H designed from this plant matrix C automatically compensates for all those responses contained in order to synthesise the correct desired signals at the listener’s ears.

3.5.2 Multiple “2 by 2” plant matrix

Alternatively, one can design inverse filter matrices $\mathbf{H}_1, \mathbf{H}_2, \dots$ for plants $\mathbf{C}_1, \mathbf{C}_2, \dots$ of each pair of driver units (Fig. 33). The cross-over filters for each pair of driver units ensure that the signals contain the corresponding frequency range of the signals for the particular pair of units. In this case, around cross-over frequencies, a virtual acoustic environment is synthesised with two different inverse filter matrices. Since both reproduced signals at the ears synthesised with both pairs of driver units are correct, the correct desired signals are reproduced at the ears as a simple sum of those two (identical but different in level) desired signals, provided that the cross-over filters behave well. Since the system inversion is now independent of the cross-over filters, the cross-over filters can also be applied to signals prior to the input to the inverse filters which can be after (Fig. 33b) or even before the binaural synthesis.

3.5.3 “2 by $(2 \times \text{multiple})$ ” plant matrix

It is also possible to obtain the plant matrix \mathbf{C} as a 2 by $(2 \times m)$ matrix where m is a number of driver pairs (Fig. 34). The system is underdetermined and a $(2 \times m)$ by 2 matrix of the pseudo inverse filter matrix \mathbf{H} is given by

$$\mathbf{H} = \mathbf{C}^H [\mathbf{C}\mathbf{C}^H + \beta \mathbf{I}]^{-1} \quad (23)$$

where β is a regularisation parameter. This solution ensures that the “least effort” (smallest output) of the transducers is used in providing the desired signals at the listener’s ears. The net result is similar to the case with a single 2 by 2 plant matrix inversion described in section 3.5.1.

3.5.4 Type of filters

In any case, the cross-over filters can be passive, active or digital filters. Obviously, when the cross-over filters are applied prior to the inverse filters, they can also be applied prior to the binaural synthesis filters \mathbf{A} in Fig. 1. If they are digital filters, they can also be included in the same filters which implement the system inversion in the exactly the same way as the filters for binaural synthesis. As Eq. (20) suggests, the inverse filter matrix \mathbf{H} can also be realised as analogue (active or passive) filters when the “OSD” principle is approximated reasonably well by means of fine discretisation or an ideal variable transducer such as that depicted in Fig. 12.

3.6 Comments on multi-channel systems

When the cross-over filters are not used, then the problem becomes a conventional multi-channel system, contrary to the “OSD” system which is a multi-way system. In this case where m is a number of driver pairs, the plant matrix is again a 2 by $(2 \times m)$ matrix of electro-acoustic transfer functions between $(2 \times m)$ outputs of the filter matrix \mathbf{H} and 2 receivers where $(2 \times m)$ is the number of channels. The pseudo inverse filter matrix \mathbf{H} is given by Eq. (23). The obtained inverse filter matrix \mathbf{H} is a $(2 \times m)$ by 2 matrix which distributes signals automatically to different drivers so that least effort is required. As an example, the magnitude of the elements of \mathbf{H} ($|\mathbf{H}_{mn}(j\omega)|$) which has 6 channels of transducers at the same position as the drivers used for the examples of the 3-way “OSD” systems with $\nu = 0.7$ are plotted in Fig. 35. The property of multi-

channel inversion is beneficial in that frequencies at which there are problems such as ill-conditioning and minima of HRTFs are automatically avoided. On the other hand, with the absence of the cross-over filters, multi-channel systems do not have some of the merit of the “OSD” system.

One of the important advantages is that of the “OSD” system being a multi-way system. The inversion of multi-channel systems ensures that most of the lower frequency signals are distributed to the pair of units with larger span since the condition numbers of the pair are always smaller than the loudspeaker pairs with smaller span at low frequencies. However, some of the higher frequency signals are also distributed to the pairs of units with larger span since there are a number of frequencies for which the larger span gives a smaller condition number due to its periodic nature. This requires the pairs with larger span to produce a very wide frequency range of signals, which is not practical.

Another merit of the “OSD” system, which being a 2-channel system, is also lost in a multi-channel system. Only two independent output signals, hence only two channels of amplifiers, are required for a passive cross-over “OSD” system whereas the same number of channels of amplifiers as number of driver units are always required for a multi-channel system.

4. Summary

A new 2-channel sound control system is proposed which overcomes the fundamental problems with system inversion by utilising a variable transducer span.

This system can be realised in practice by discretising the theoretical continuously variable transducer span which results in multi-way sound control system.

Even though basic principles and properties have been explained with a 2-channel system as an example, the same principle can be applied to multi-channel case as multi-way multi-channel systems.

When the variable transducer span is well approximated, it may be possible to achieve a virtual source synthesis with a simple gain change and phase shift.

References

- [1] J. Blauert, Spatial Hearing; The Psychophysics of Human Sound Localization (MIT Press, Cambridge, MA, 1997).
- [2] H. Møller, "Fundamentals of Binaural Technology," *Appl. Acoust.* **36**, 171-218 (1992).
- [3] D. R. Begault, 3-D Sound for Virtual Reality and Multimedia (AP Professional, Cambridge, MA, 1994).
- [4] M. R. Schroeder, B. S. Atal, "Computer Simulation of Sound Transmission in Rooms," *IEEE Intercon. Rec.* Pt7, 150-155 (1963).

- [5] P. Damaske, "Head-related Two-channel Stereophony with Reproduction," J. Acoust. Soc. Am. **50**, 1109-1115 (1971).
- [6] H. Hamada, N. Ikeshoji, Y. Ogura And T. Miura, "Relation between Physical Characteristics of Orthostereophonic System and Horizontal Plane Localisation," Journal of the Acoustical Society of Japan, (E) **6**, 143-154, (1985).
- [7] J. L. Bauck and D. H. Cooper, "Generalized Transaural Stereo and Applications," J. Acoust. Soc. Am. **44** (9), 683-705 (1996).
- [8] P. A. Nelson, O. Kirkeby, T. Takeuchi, and H. Hamada, "Sound fields for the production of virtual acoustic images," J. Sound. Vib. **204** (2), 386-396 (1997).
- [9] P.A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Inverse Filter Design and Equalisation Zones in Multi-Channel Sound Reproduction," IEEE Trans. Speech Audio Process. **3**(3), 185-192 (1995).
- [10] O. Kirkeby, P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Local Sound Field Reproduction Using Digital Signal Processing," J. Acoust. Soc. Am. **100**, 1584-1593 (1996).
- [11] T. Takeuchi, P.A. Nelson, O. Kirkeby and H. Hamada, "The Effects of Reflections on the Performance of Virtual Acoustic Imaging Systems", pp. 955-966, in Proceedings of the Active 97, The international symposium on active control of sound and vibration, Budapest, Hungary, August 21-23, 1997, OPAKFI
- [12] B. Gardner, and K. Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone," MIT Media Lab Perceptual Computing - Technical Report No. 280 (1994).
- [13] O. Kirkeby, P. A. Nelson, and H. Hamada, "Stereo Dipole," UK Patent Application, 9603236.2, 1996.

- [14] T. Takeuchi, P.A. Nelson, O. Kirkeby and H. Hamada, "Influence of Individual Head Related Transfer Function on the Performance of Virtual Acoustic Imaging Systems", 104th AES Convention Preprint 4700 (P4-3).
- [15] S. Barnett, *Matrices - Methods and Applications* (Oxford University Press, Oxford, 1990).

Appendix

When the desired signals are defined as Eq. (7), this effectively normalises the plant transfer function matrix \mathbf{C} with respect to the acoustic pressure signals which would have been produced by the closer of two sound sources to the receiver point. Then this normalised plant transfer function matrix \mathbf{C} can be written as

$$\mathbf{C} = \begin{bmatrix} 1 & g e^{-jk\Delta l} \\ g e^{-jk\Delta l} & 1 \end{bmatrix} \quad (\text{A } 1)$$

It is possible to express \mathbf{C} with unitary matrices \mathbf{U} and \mathbf{V} such that

$$\mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \quad (\text{A } 2)$$

where $\mathbf{\Sigma}$ is the diagonal matrix whose elements are singular values of \mathbf{C} , σ_1 and σ_2 [15]. The singular values of \mathbf{C} can be found from the square roots of eigenvalues of $\mathbf{C}^H \mathbf{C}$.

$$\mathbf{C}^H \mathbf{C} = \begin{bmatrix} 1 & g e^{jk\Delta l} \\ g e^{jk\Delta l} & 1 \end{bmatrix} \begin{bmatrix} 1 & g e^{-jk\Delta l} \\ g e^{-jk\Delta l} & 1 \end{bmatrix} = \begin{bmatrix} 1+g^2 & g(e^{jk\Delta l} + e^{-jk\Delta l}) \\ g(e^{jk\Delta l} + e^{-jk\Delta l}) & 1+g^2 \end{bmatrix} \quad (\text{A } 3)$$

The eigenvalues of $\mathbf{C}^H \mathbf{C}$ are given by

$$\lambda_{1,2} = (1 + ge^{jk\Delta l})(1 + ge^{-jk\Delta l}), (1 - ge^{jk\Delta l})(1 - ge^{-jk\Delta l}) \quad (\text{A } 4)$$

Therefore, the singular values of \mathbf{C} are given by

$$\sigma_{1,2} = \sqrt{\lambda_{1,2}} = \sqrt{(1 + ge^{jk\Delta l})(1 + ge^{-jk\Delta l})}, \sqrt{(1 - ge^{jk\Delta l})(1 - ge^{-jk\Delta l})} \quad (\text{A } 5)$$

It is possible to find an orthonormal set of eigenvectors \mathbf{x}_i such that

$$\mathbf{C}^H \mathbf{C} \mathbf{x}_i = \sigma_i^2 \mathbf{x}_i \quad (\text{A } 6)$$

Therefore,

$$\mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad (\text{A } 7)$$

The vectors comprising \mathbf{U} are given by

$$\mathbf{y}_i = \frac{1}{\sigma_i} \mathbf{C} \mathbf{x}_i \quad (\text{A } 8)$$

Hence

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{\frac{1+ge^{-jk\Delta l}}{1+ge^{jk\Delta l}}} & \sqrt{\frac{1-ge^{-jk\Delta l}}{1-ge^{jk\Delta l}}} \\ \sqrt{\frac{1+ge^{-jk\Delta l}}{1+ge^{jk\Delta l}}} & -\sqrt{\frac{1-ge^{-jk\Delta l}}{1-ge^{jk\Delta l}}} \end{bmatrix}$$

(A 9)

The singular value decomposition of \mathbf{C} may therefore be written as

$$\begin{aligned} \mathbf{C} &= \mathbf{U}\Sigma\mathbf{V}^H \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{\frac{1+ge^{-jk\Delta l}}{1+ge^{jk\Delta l}}} & \sqrt{\frac{1-ge^{-jk\Delta l}}{1-ge^{jk\Delta l}}} \\ \sqrt{\frac{1+ge^{-jk\Delta l}}{1+ge^{jk\Delta l}}} & -\sqrt{\frac{1-ge^{-jk\Delta l}}{1-ge^{jk\Delta l}}} \end{bmatrix} \begin{bmatrix} \sqrt{(1+ge^{jk\Delta l})(1+ge^{-jk\Delta l})} & 0 \\ 0 & \sqrt{(1-ge^{jk\Delta l})(1-ge^{-jk\Delta l})} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

(A 10)

Note that

$$\mathbf{C}^{-1} = [\mathbf{U}\Sigma\mathbf{V}^H]^{-1} = \mathbf{V}\Sigma^{-1}\mathbf{U}^H$$

(A 11)

since $\mathbf{V}^H\mathbf{V} = \mathbf{I}$, $[\mathbf{V}^H]^{-1} = \mathbf{V}$, $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ and $\mathbf{U}^{-1} = \mathbf{U}^H$. Therefore,

$$\begin{aligned}
\mathbf{H} &= \mathbf{C}^{-1} \\
&= \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H \\
&= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{(1+ge^{jk\Delta l})(1+ge^{-jk\Delta l})}} & 0 \\ 0 & \frac{1}{\sqrt{(1-ge^{jk\Delta l})(1-ge^{-jk\Delta l})}} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{\frac{1+ge^{jk\Delta l}}{1+ge^{-jk\Delta l}}} & \sqrt{\frac{1+ge^{jk\Delta l}}{1+ge^{-jk\Delta l}}} \\ \sqrt{\frac{1-ge^{jk\Delta l}}{1-ge^{-jk\Delta l}}} & -\sqrt{\frac{1-ge^{jk\Delta l}}{1-ge^{-jk\Delta l}}} \end{bmatrix}
\end{aligned}
\tag{A 12}$$

Hence

$$\begin{aligned}
\sigma_i &= \frac{1}{\sqrt{(1+ge^{-jk\Delta r \sin \theta})(1+ge^{jk\Delta r \sin \theta})}} \\
\sigma_o &= \frac{1}{\sqrt{(1-ge^{-jk\Delta r \sin \theta})(1-ge^{jk\Delta r \sin \theta})}}
\end{aligned}
\tag{A 13}$$

are the singular values of the inverse filter matrix \mathbf{H} .

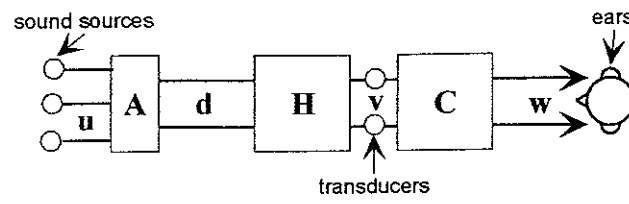


Fig. 1 Block diagram for multi-channel sound control
with system inversion

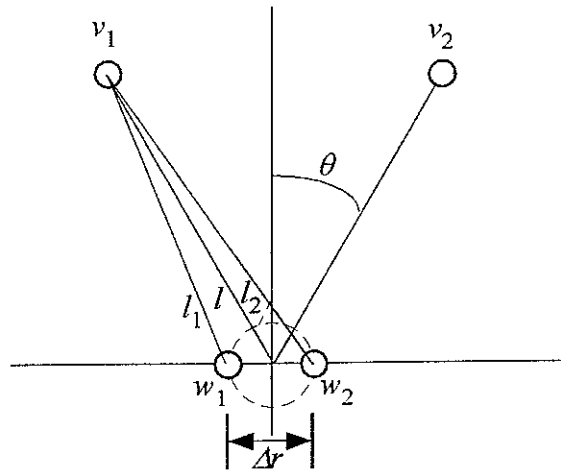
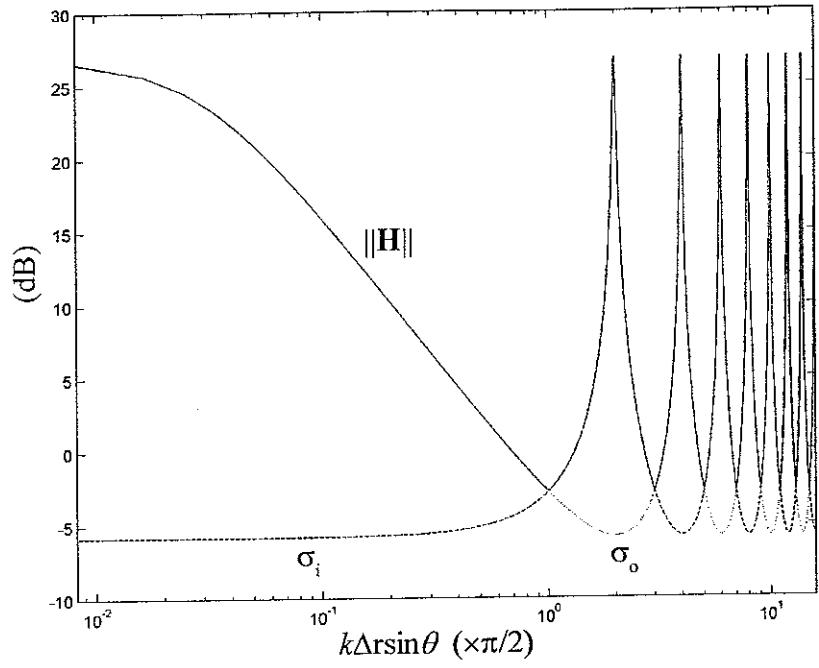
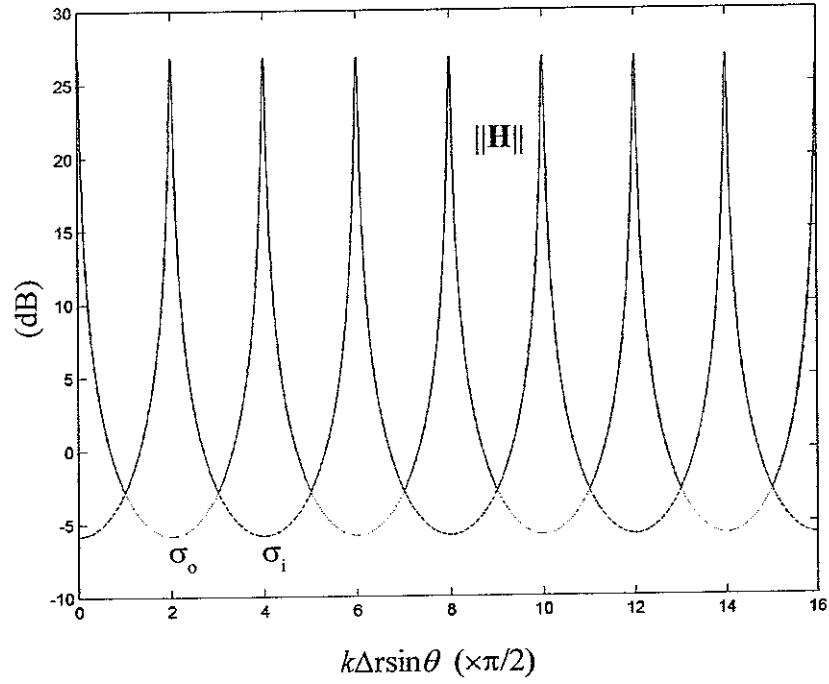


Fig. 2 The geometry of a 2-source 2-receiver system under investigation.



a)



b)

Fig. 3 Norm and singular values of the inverse filter matrix \mathbf{H} as a function of $k\Delta r \sin \theta$. a) Logarithmic scale.

b) Linear scale.

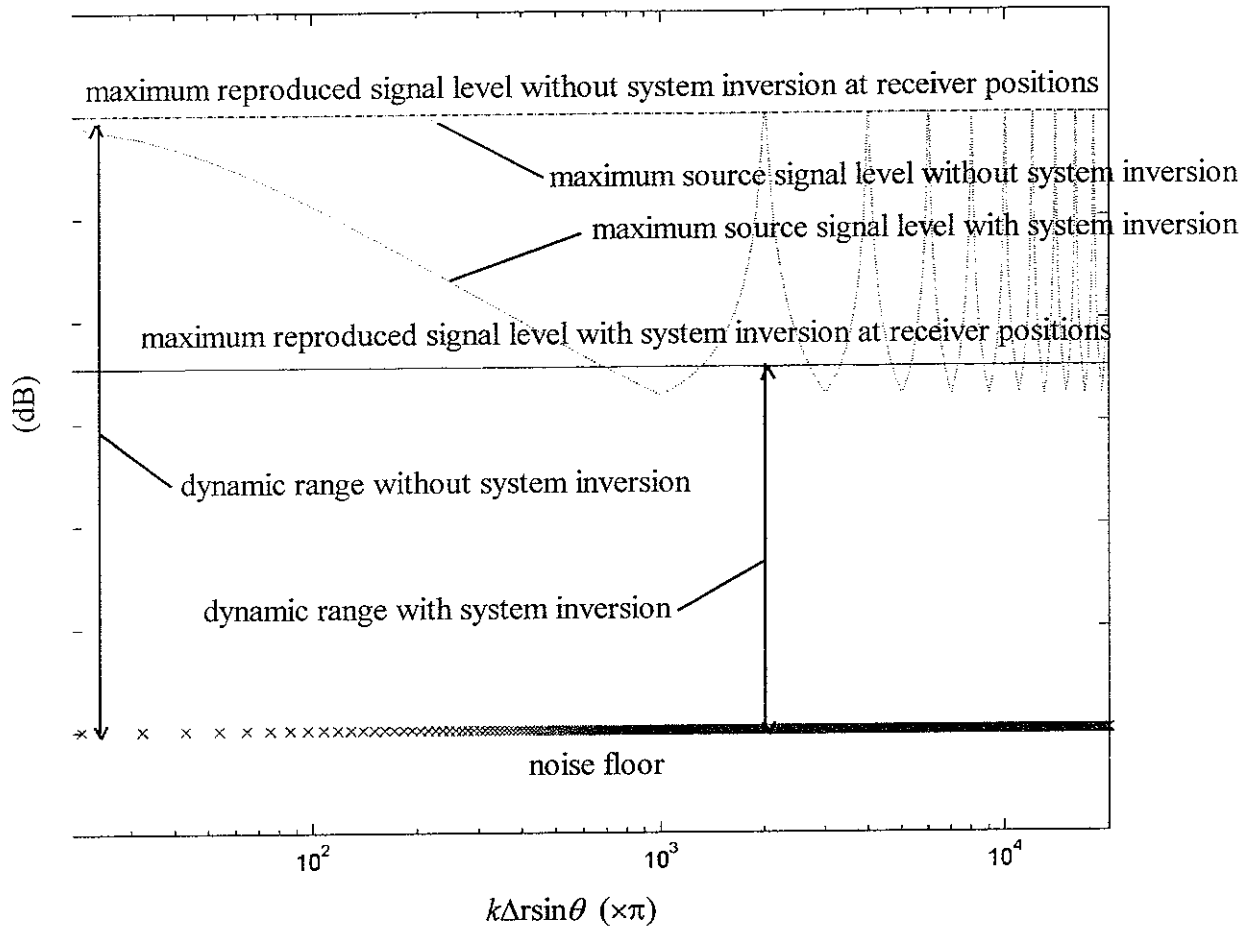


Fig. 4 Dynamic range loss due to system inversion.

Source signal levels are normalised by the maximum source signal level without system inversion as 0 dB.

Reproduced signal levels are normalised by the reproduced signal level without system inversion as 0 dB.

Noise floor is plotted at -60dB for an aid to illustrate dynamic range.

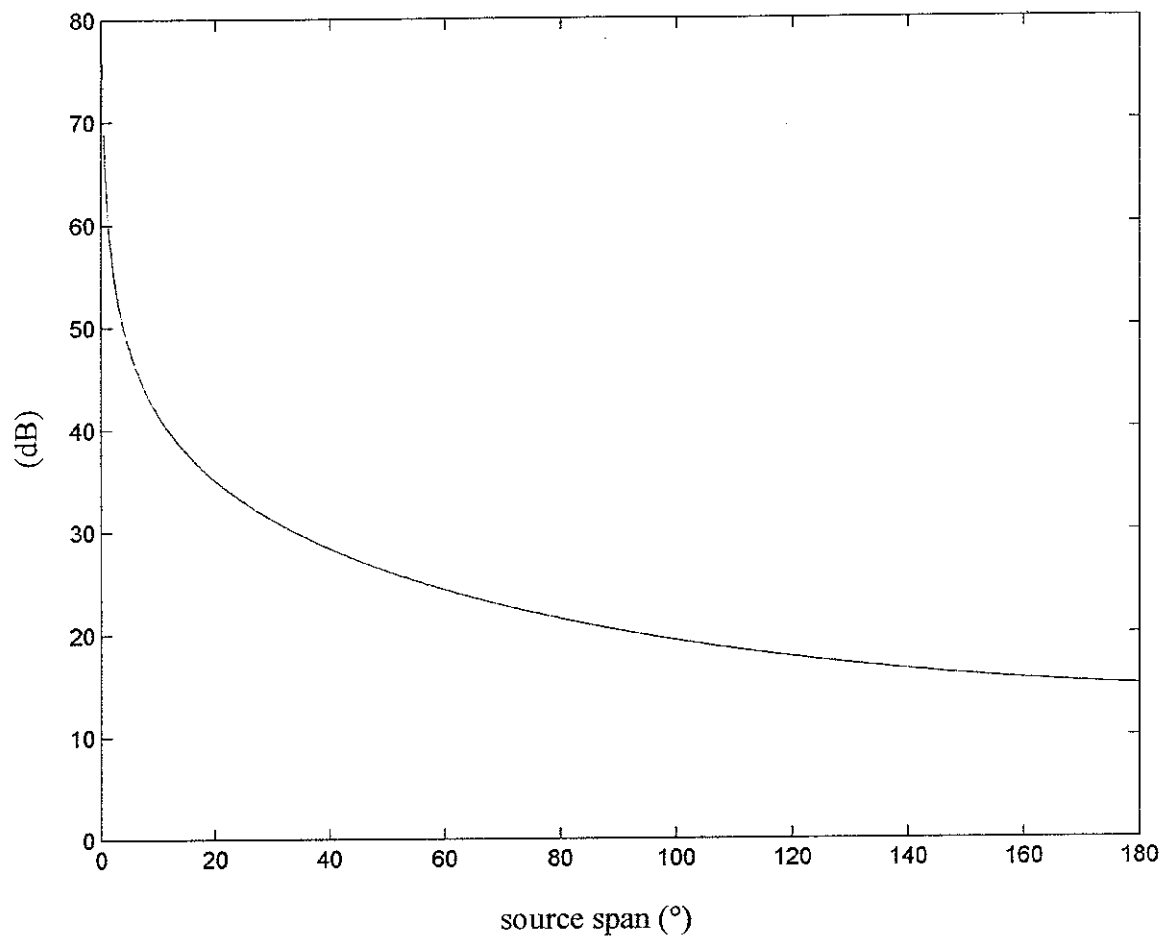
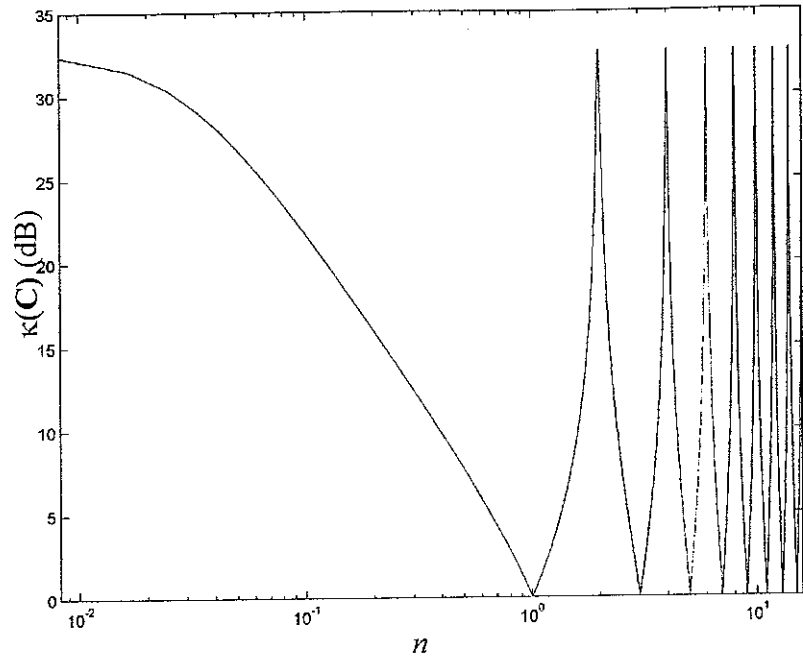
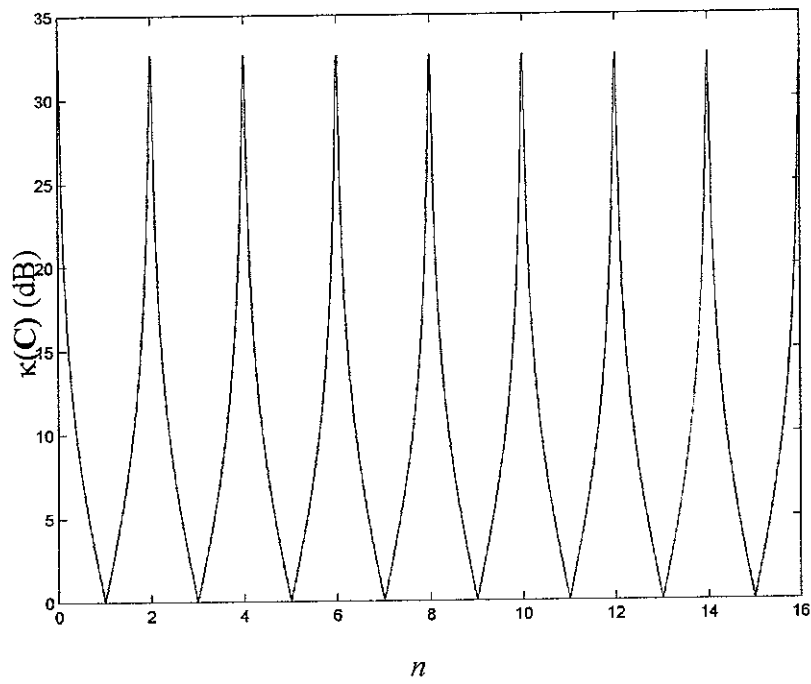


Fig. 5 Dynamic range loss as a function of source span.



a)



b)

Fig. 6 Condition number $\kappa(C)$ as a function of n . a) Logarithmic scale. b) Linear scale.

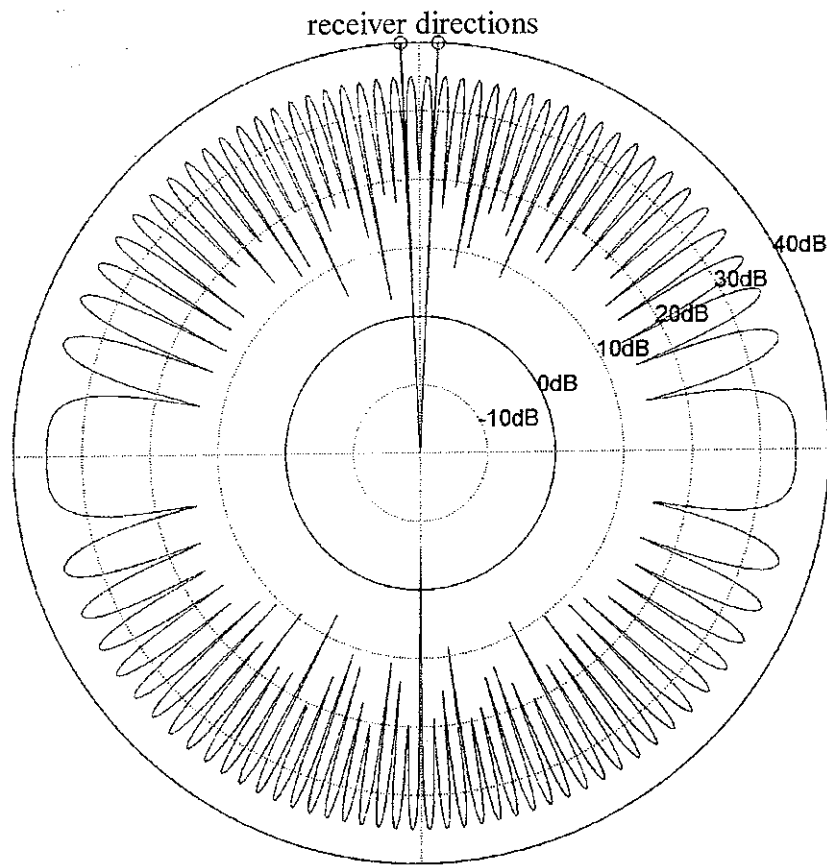
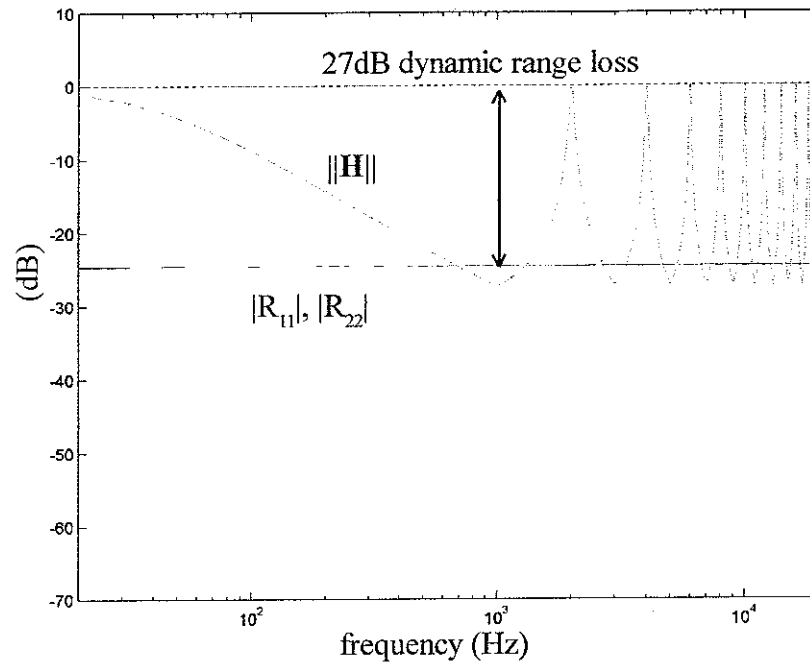
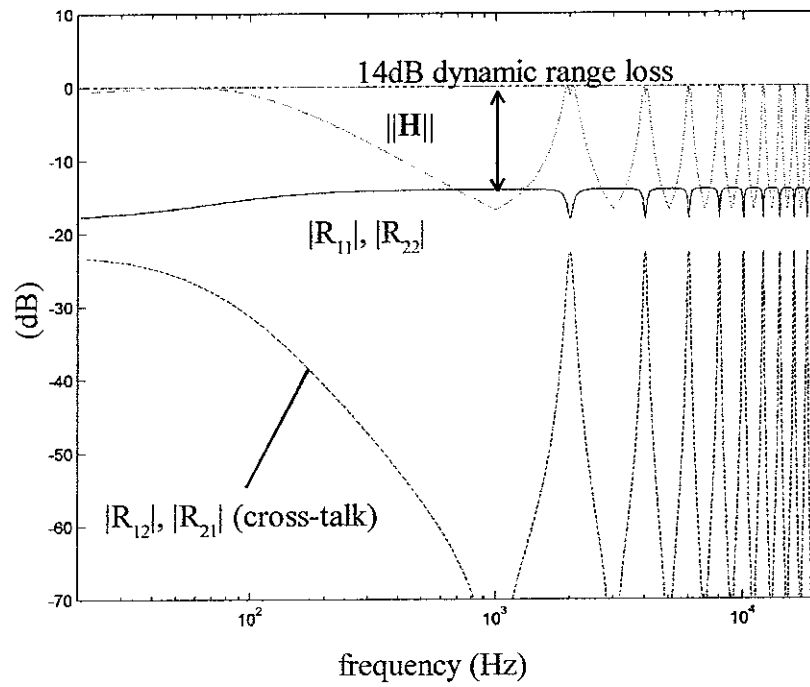


Fig. 7 Far field sound pressure level produced by the control sources with reference to the receiver directions (0dB and $-\infty$ dB), $n = 4$.



a)



b)

Fig. 8 Dynamic range improvement and loss of control performance with regularisation. a) Without regularisation. b) With regularisation

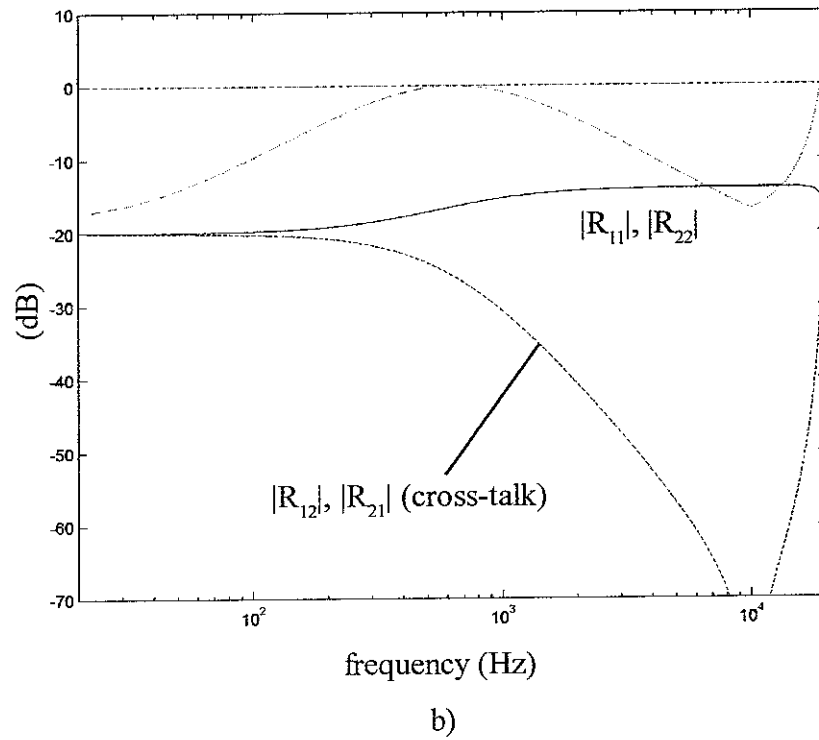
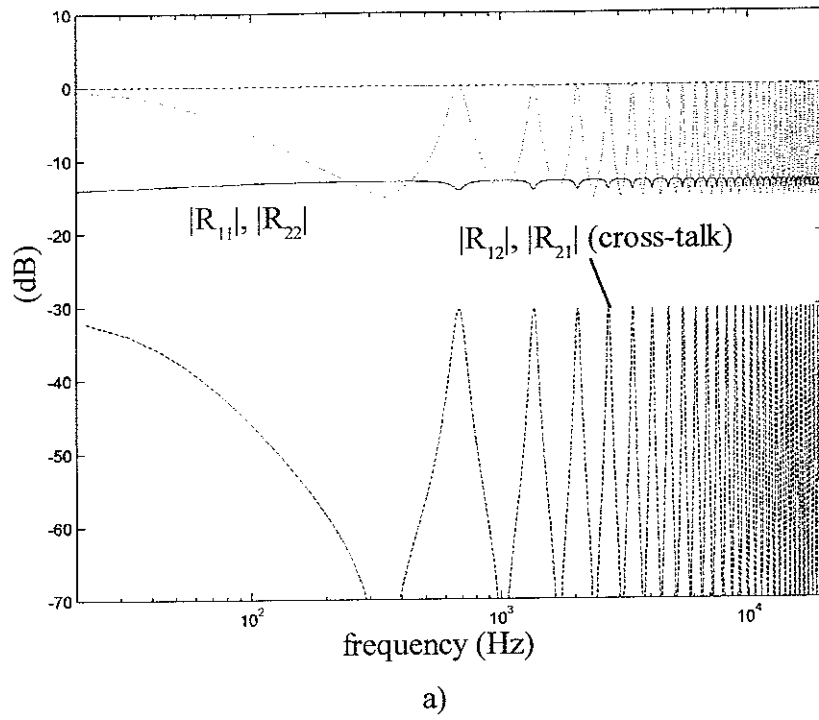


Fig. 9 Effect of changing source span. a) Larger source span. b) Smaller source span.

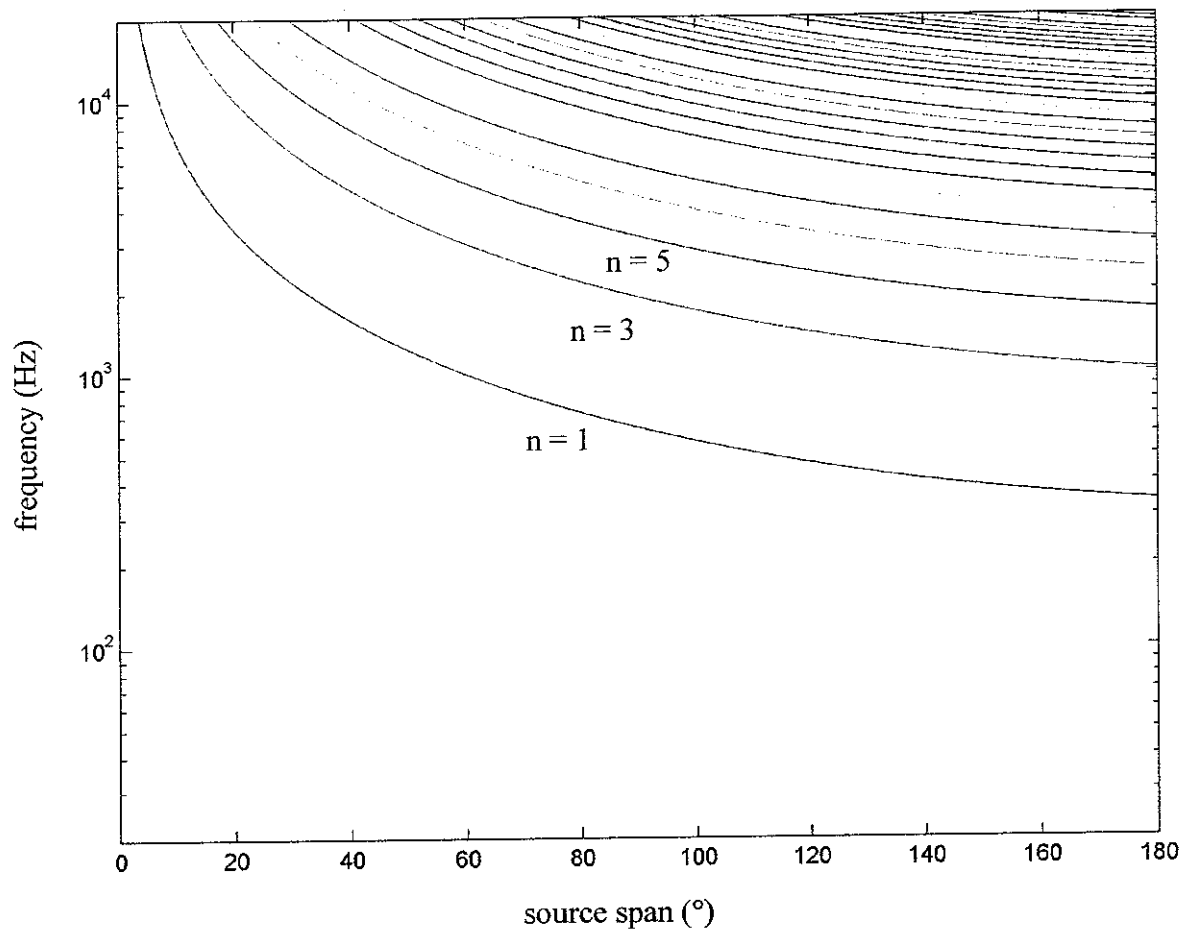


Fig. 10 The principle of the “OSD” system. The relationship between source span and frequency for different odd integer number n .

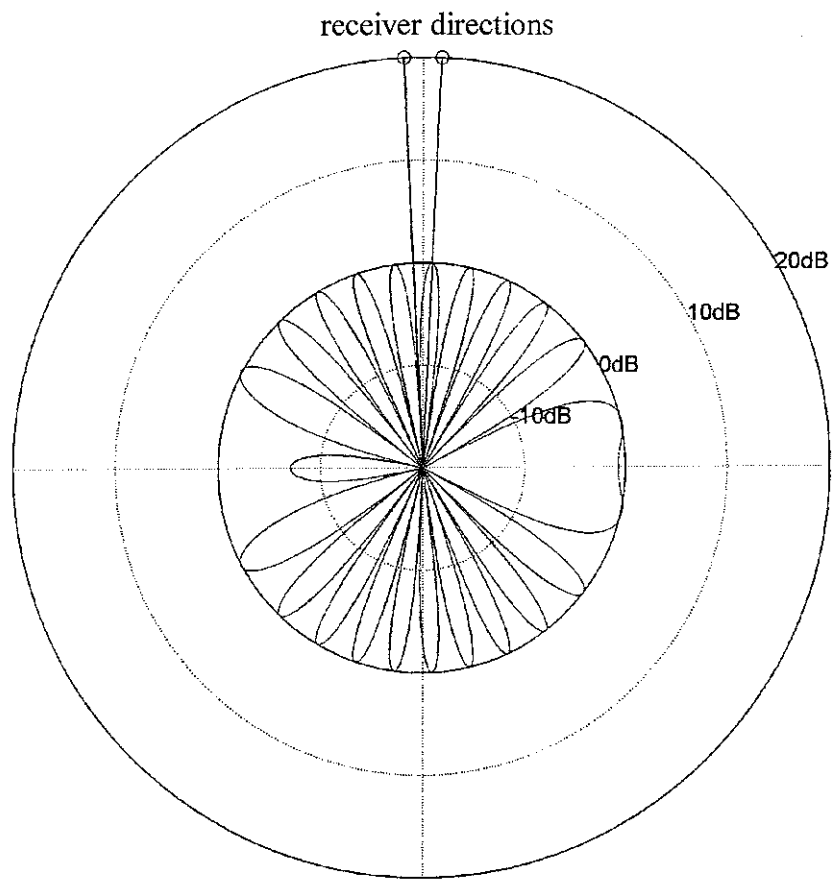


Fig. 11 Far field sound pressure level with reference to the receiver directions (0dB and $-\infty$ dB). $n = 1$.

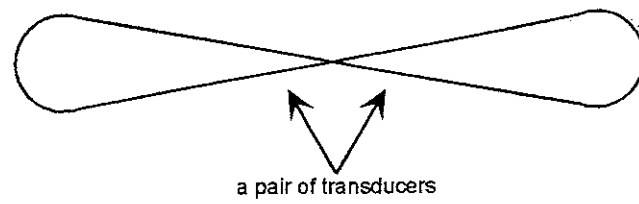


Fig. 12 Variable position (span) / frequency transducer.

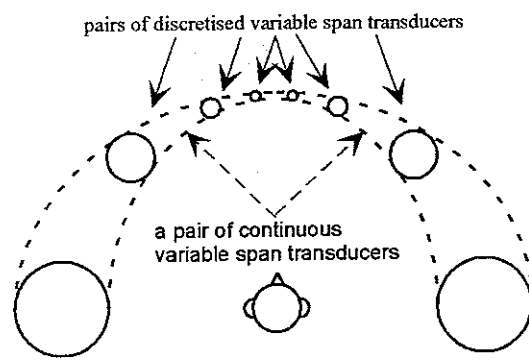


Fig. 13 Discretised variable position (span) /
frequency span transducers.

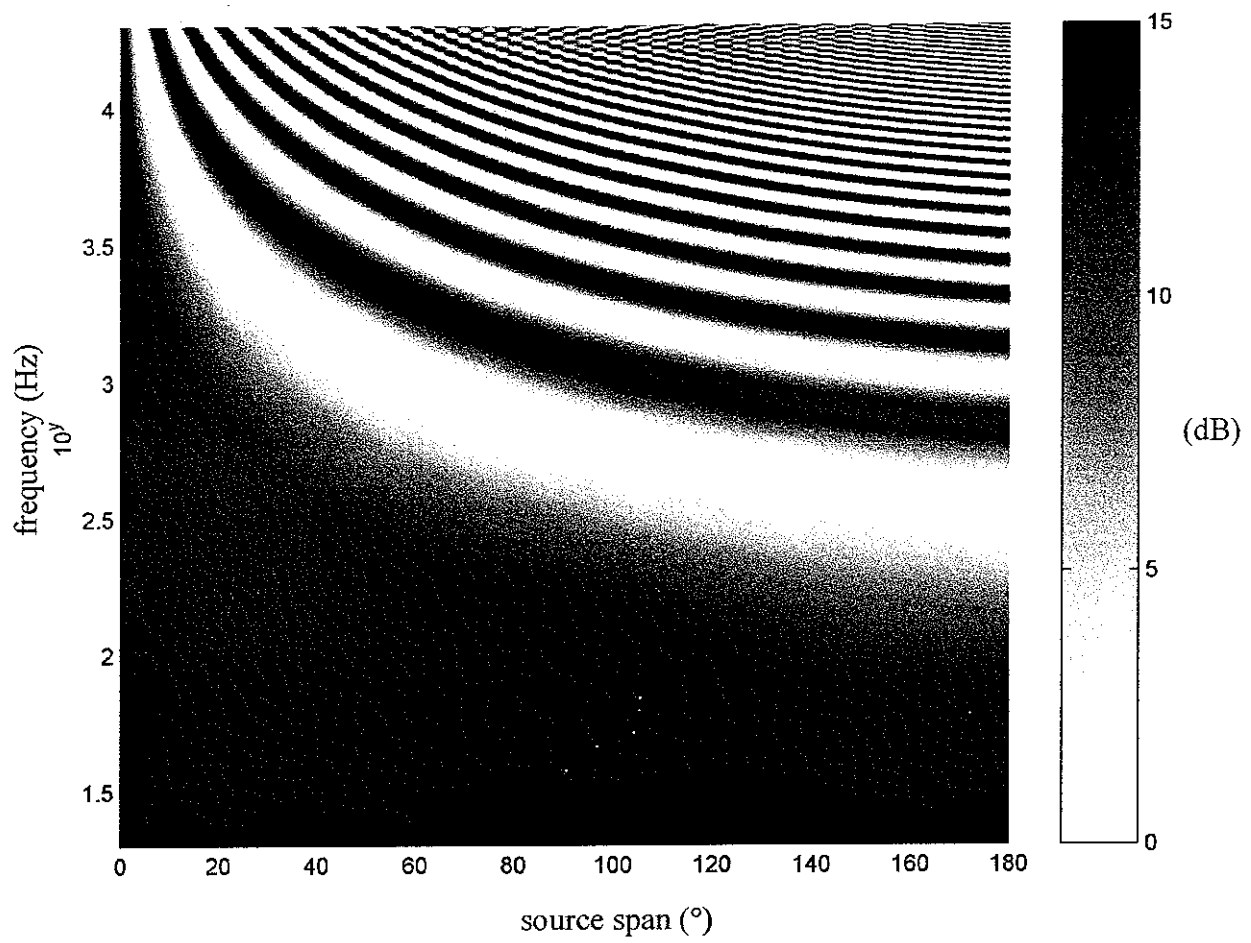


Fig. 14 Condition number $\kappa(\mathbf{C})$ of a free field plant matrix \mathbf{C} as a function of source span and frequency.

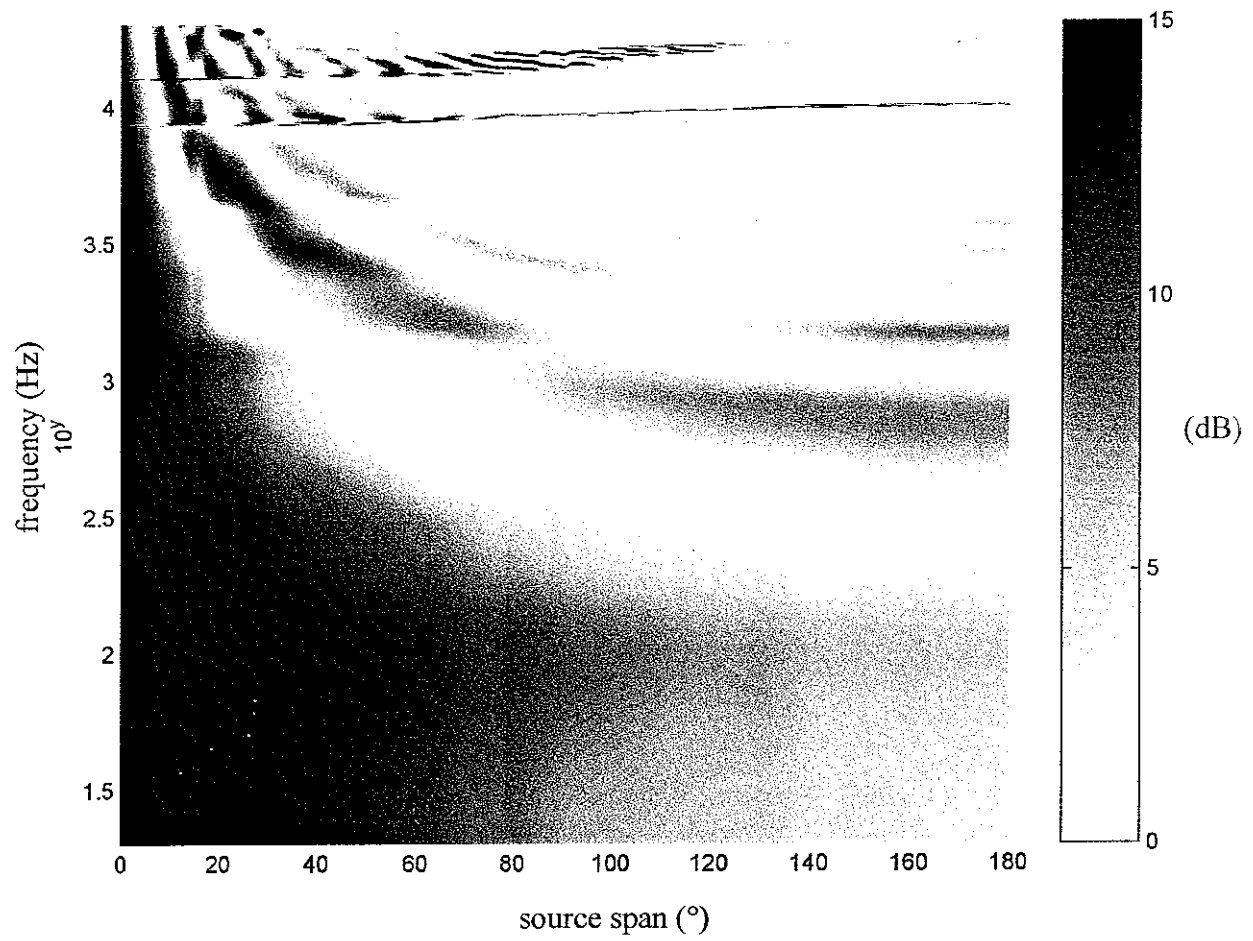


Fig. 15 Condition number $\kappa(C)$ of a plant matrix C with HRTFs as a function of source span and frequency.

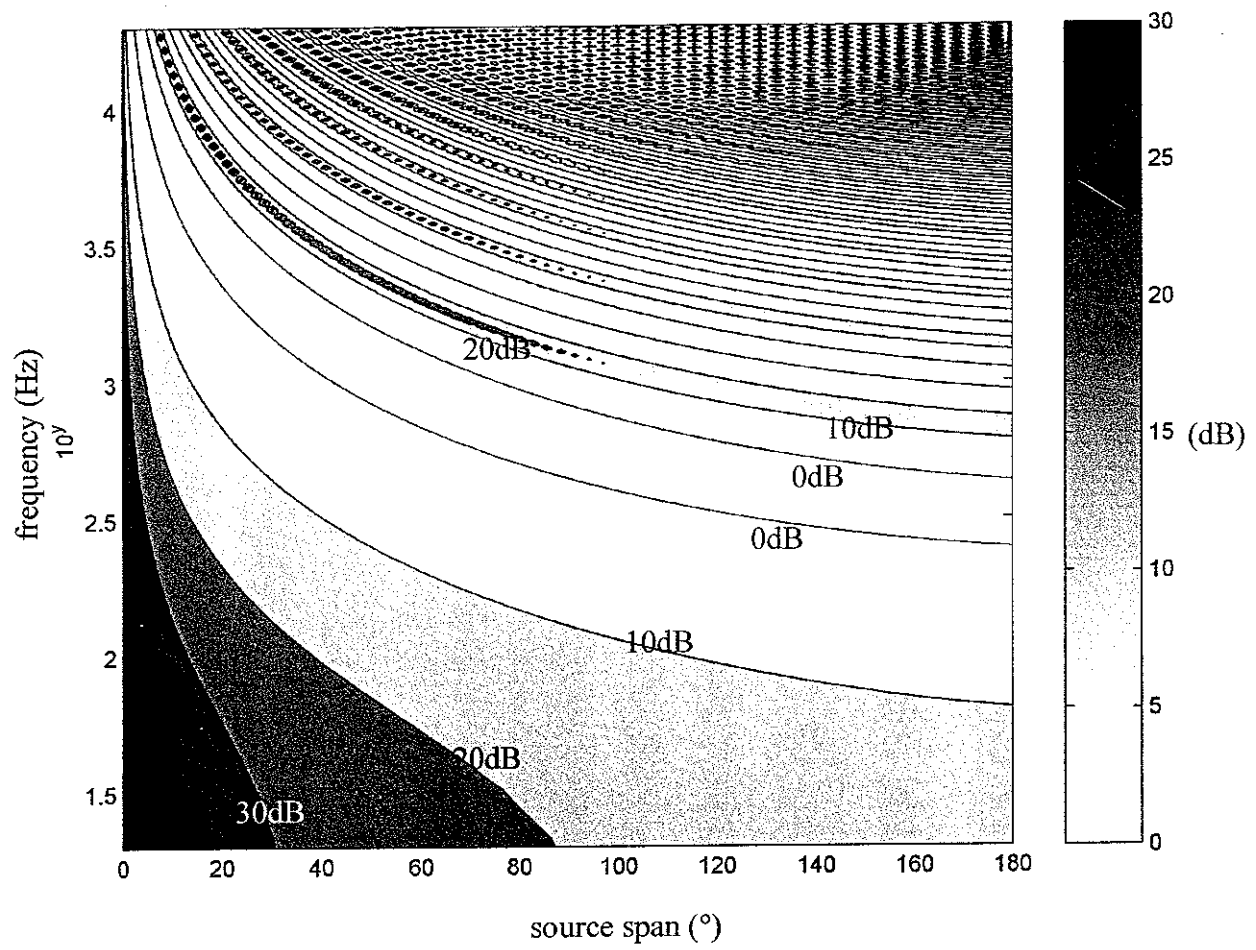


Fig. 16 Dynamic range loss as a function of source span and frequency range.

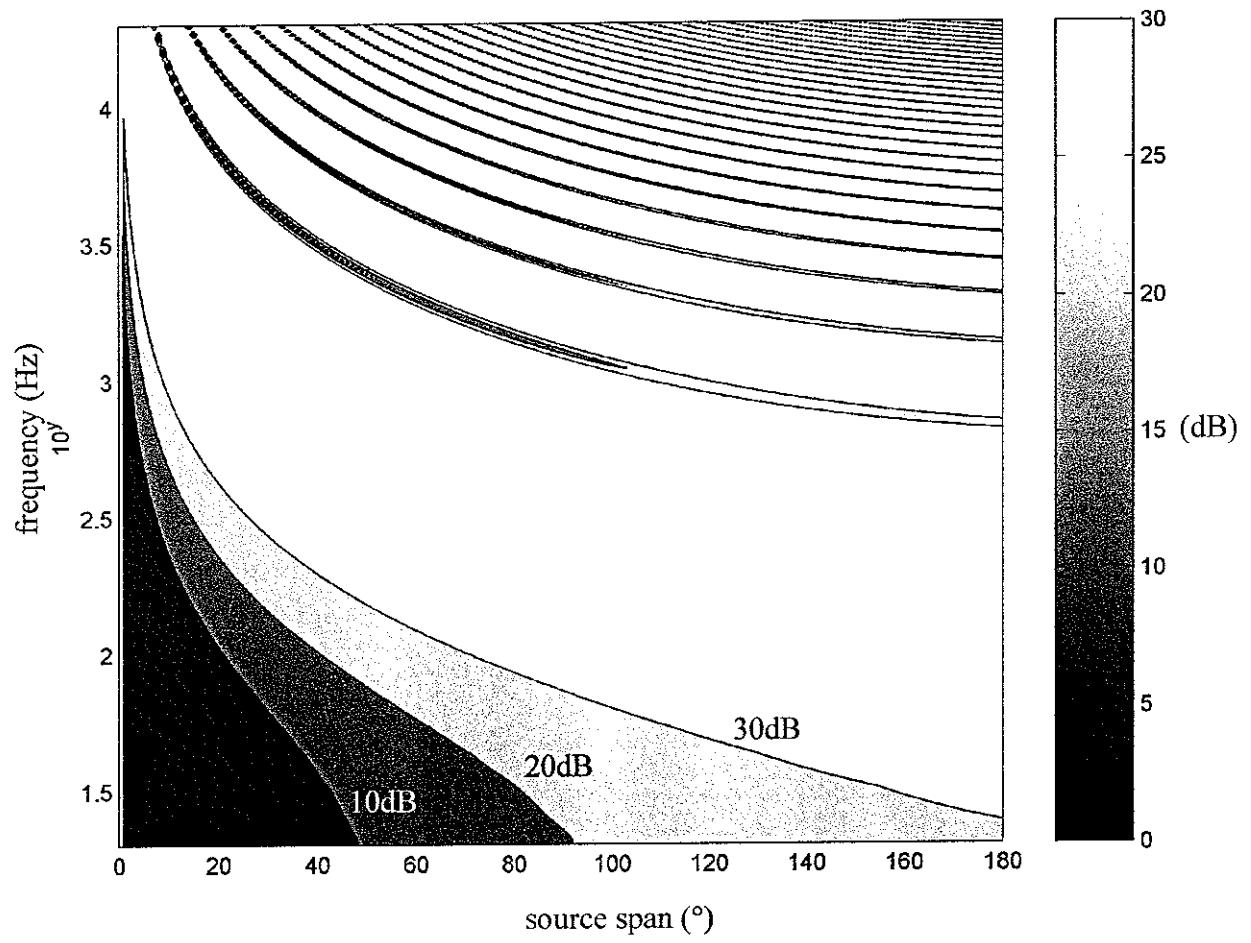


Fig. 17 Cross-talk cancellation performance as a function of source span and frequency with regularisation for 20dB dynamic range loss.

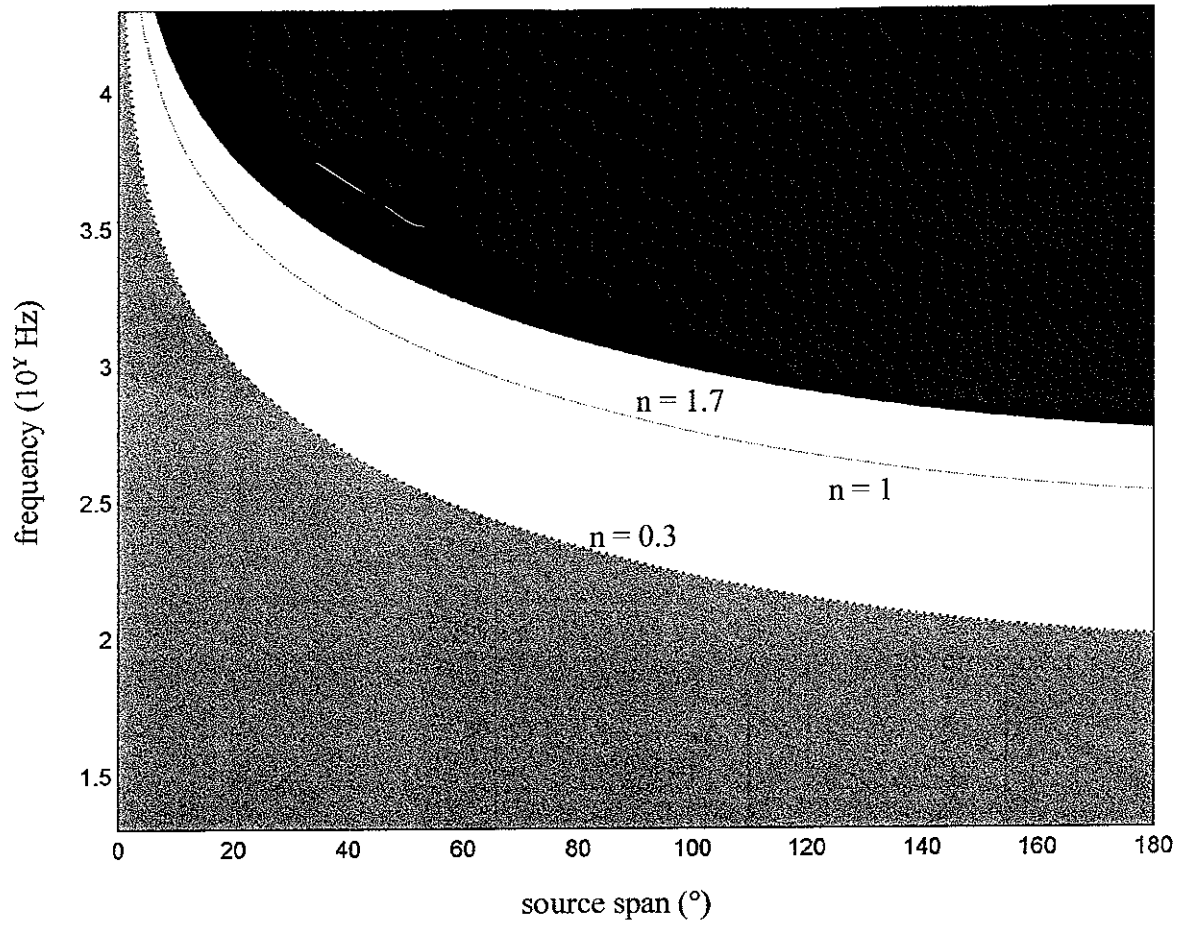


Fig. 18 The frequency/span region for systems with $n \approx 1$ and $\nu = 0.7$, and an example of discretisation for a 3-way system.

white area: $0.3 < n < 1.7$

grey area: $f < f_l$

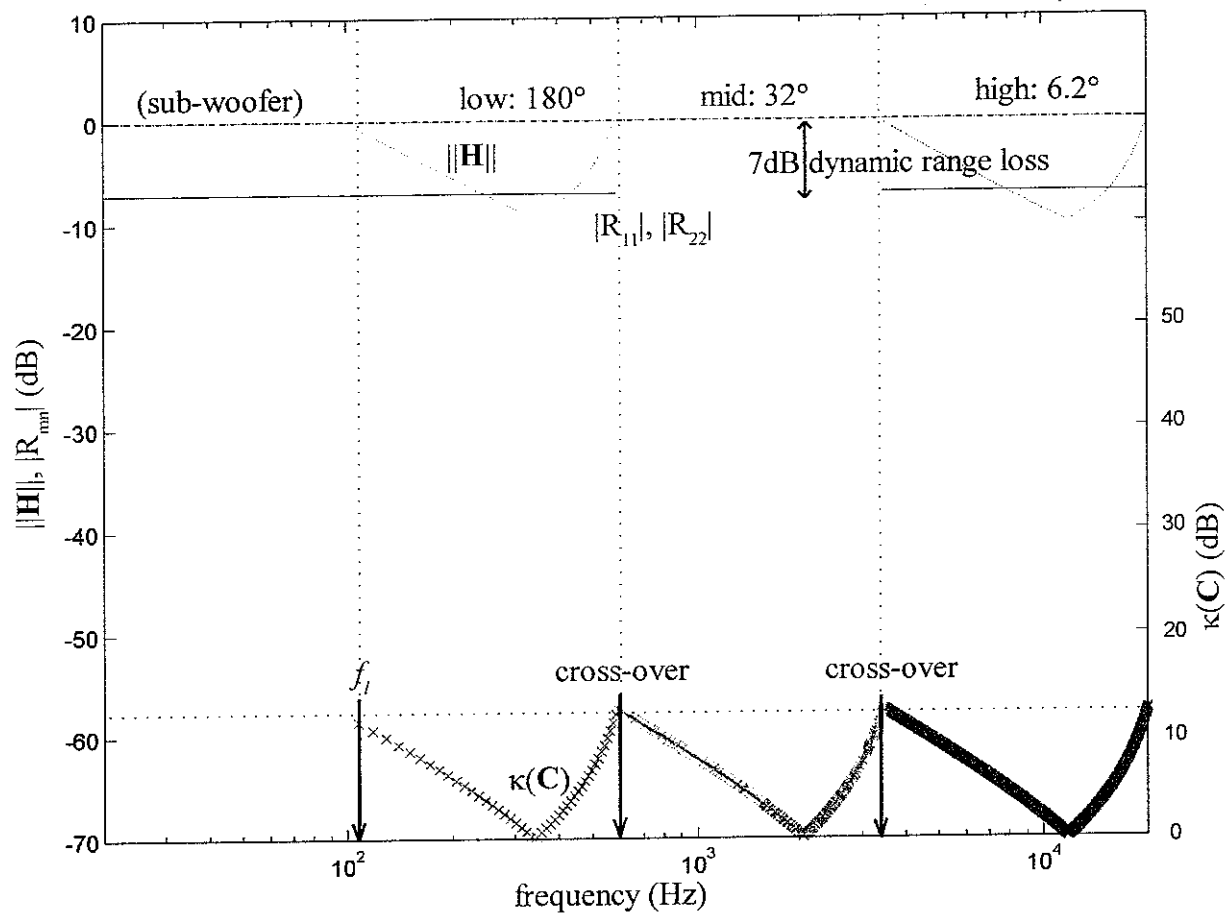
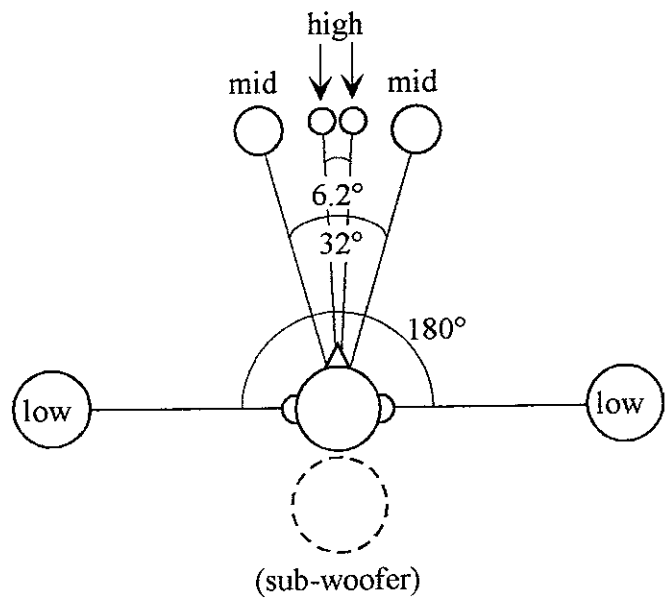


Fig. 19 An example of 3-way system with $n \approx 1$ and $\nu = 0.7$.

- source signal level with system inversion ($||\mathbf{H}||$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(\mathbf{C})$



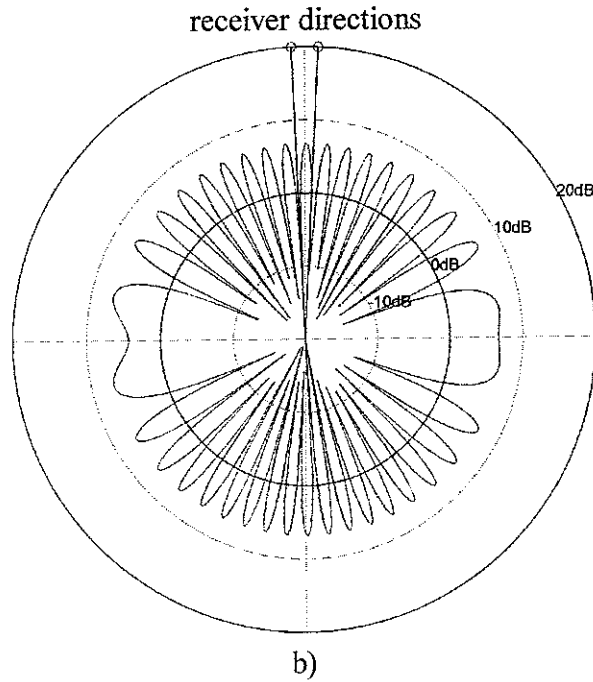
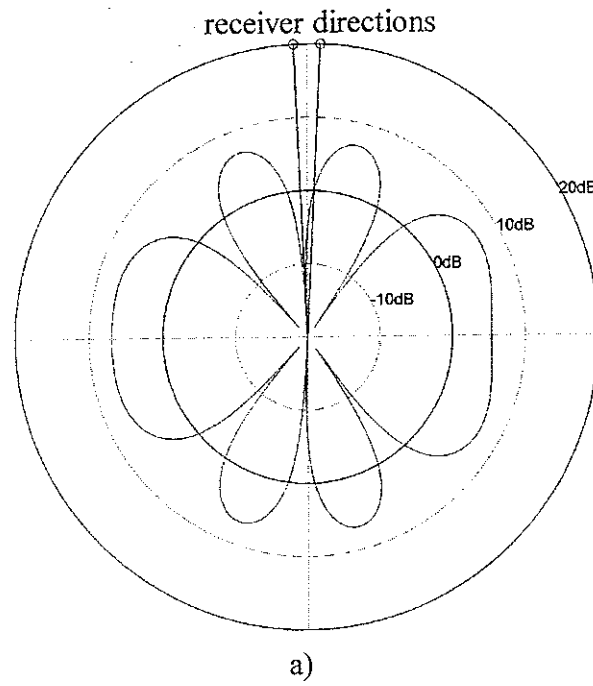


Fig. 20 Maximum far field sound pressure level produced by the transducer pairs with reference to the receiver directions (0dB and $-\infty$ dB). $\nu = 0.7$. a) $n = 0.3$ b) $n = 1.7$

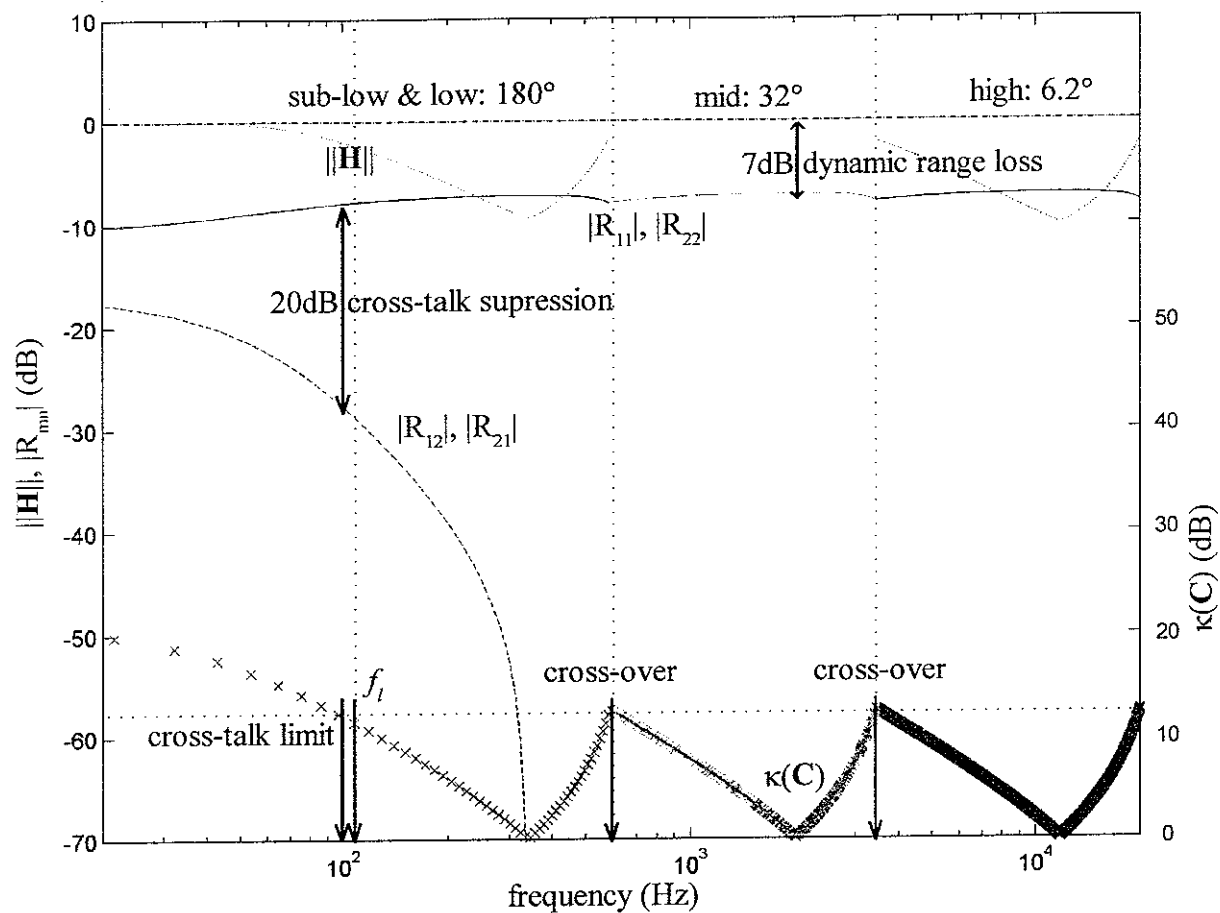
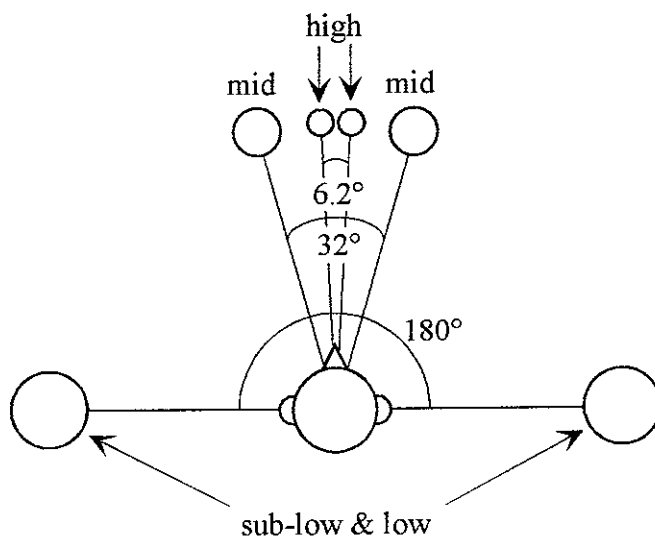


Fig. 21 An example of 3-way system with regularisation for 7dB dynamic range loss

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



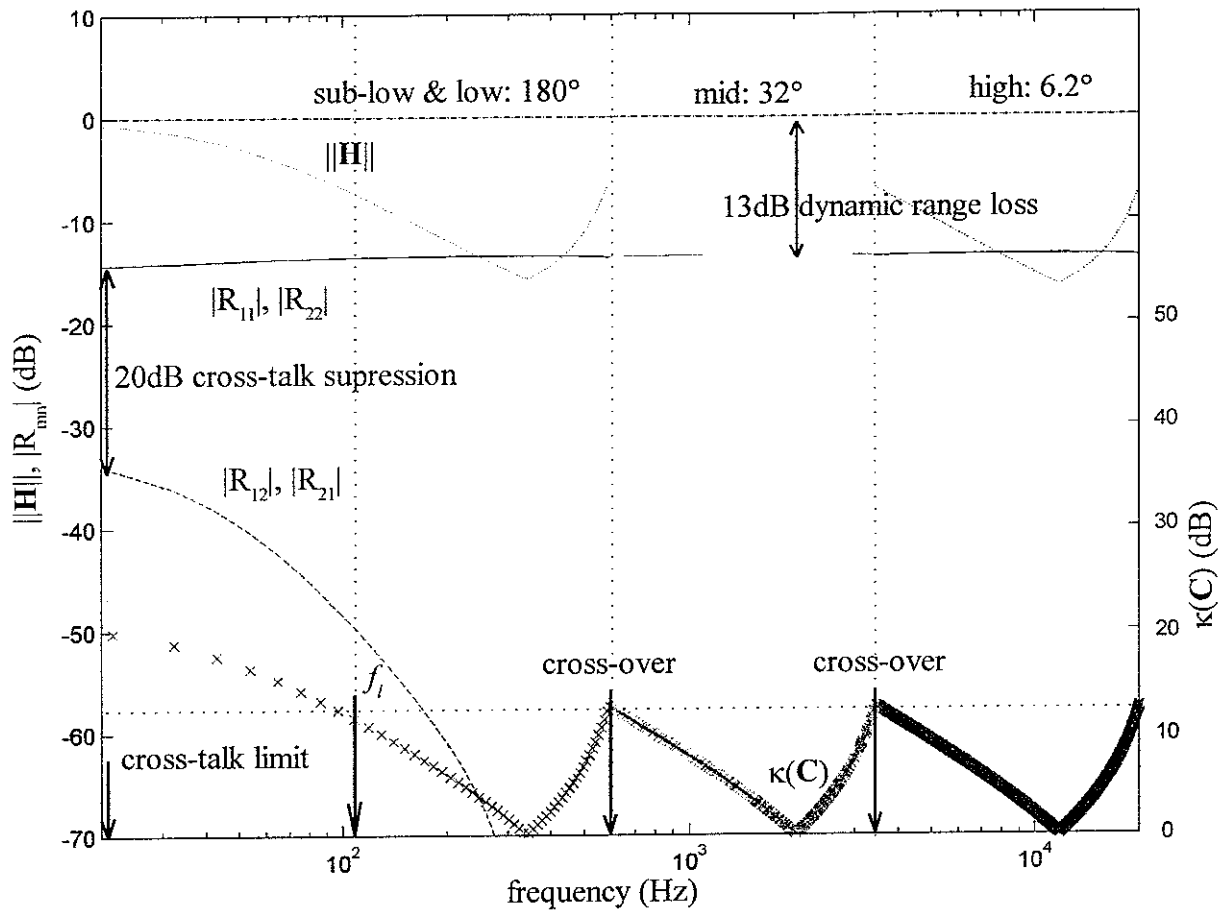
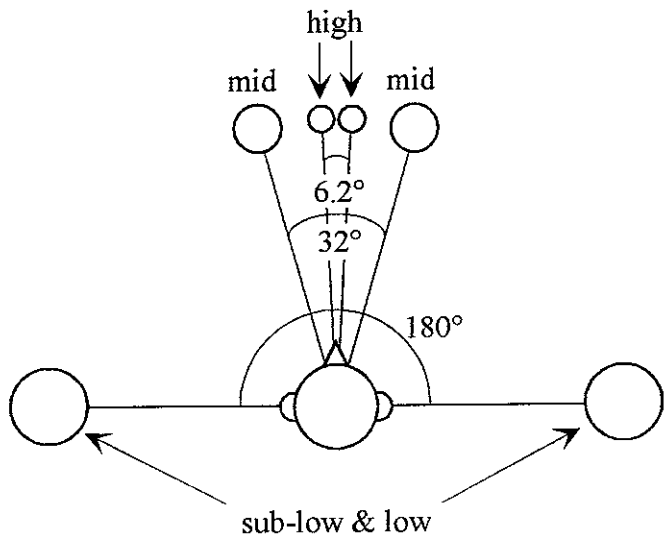


Fig. 22 An example of 3-way system with regularisation for 13dB dynamic range loss

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



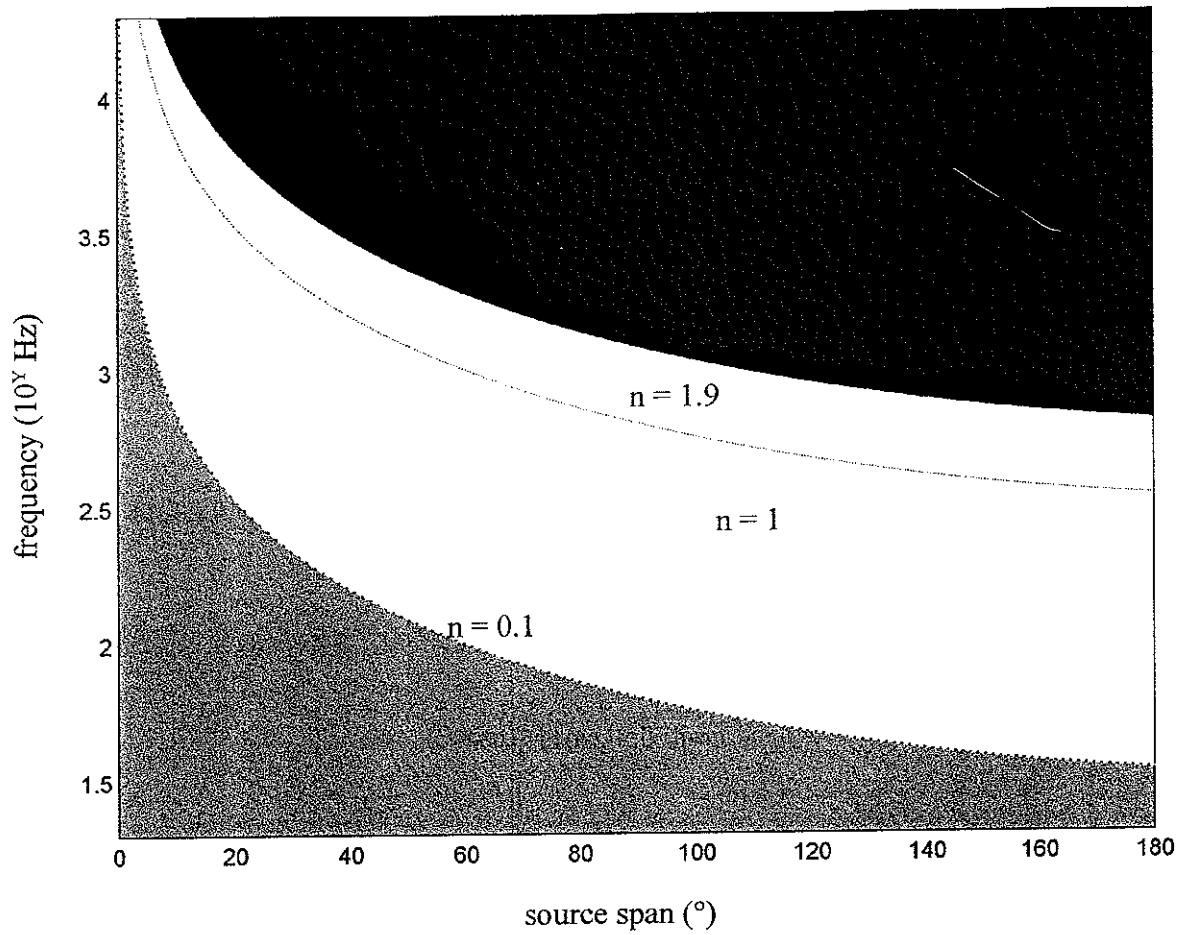


Fig. 23 The frequency/span region for systems with $n \approx 1$ and $\nu = 0.9$, and an example of discretisation for a 2-way system.

white area: $0.1 < n < 1.9$

grey area: $f < f_1$

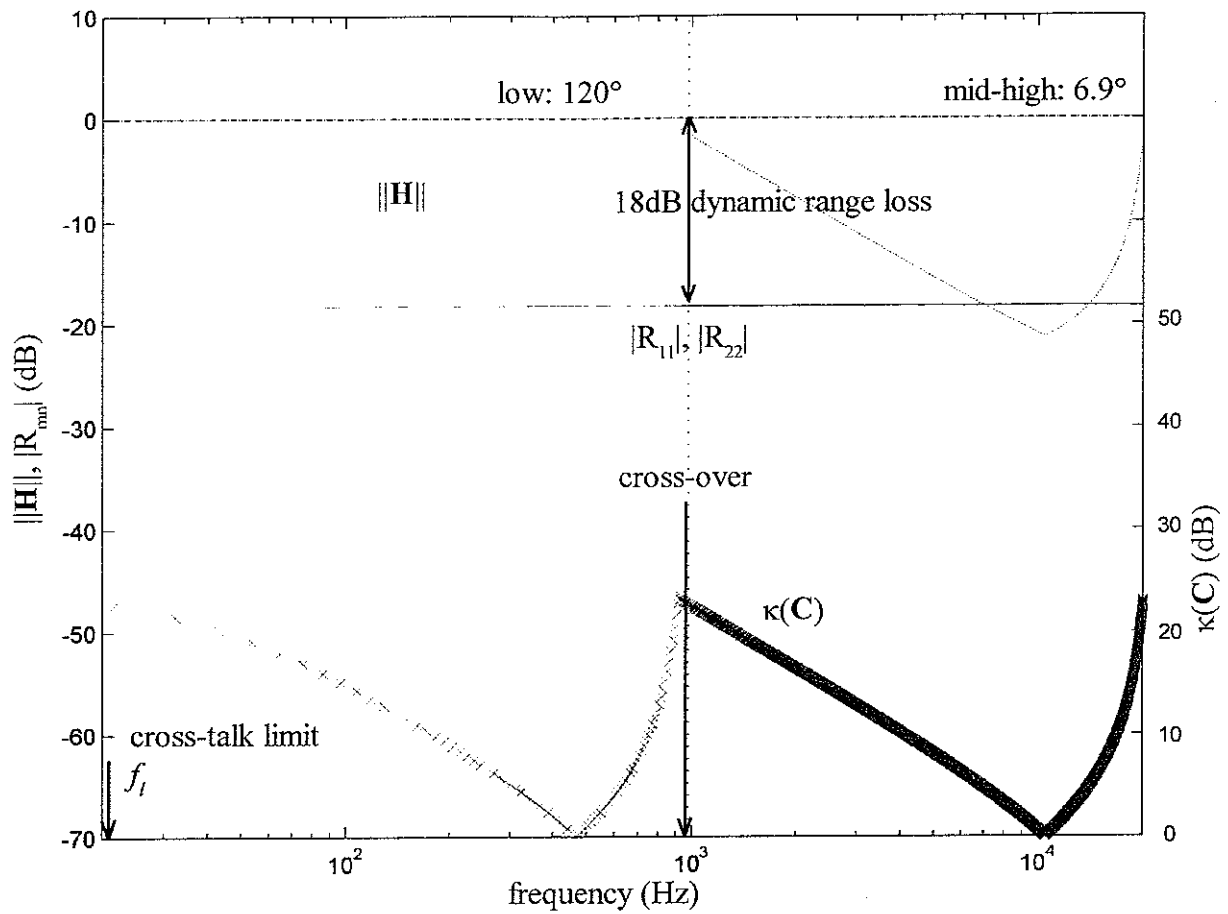
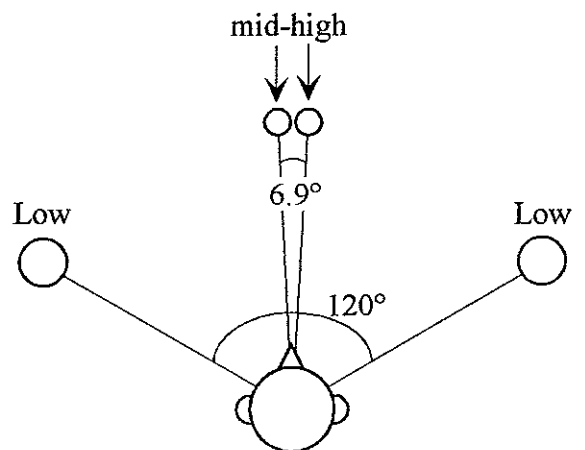


Fig. 24 An example of 2-way system with $n \approx 1$ and $\nu = 0.9$.

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



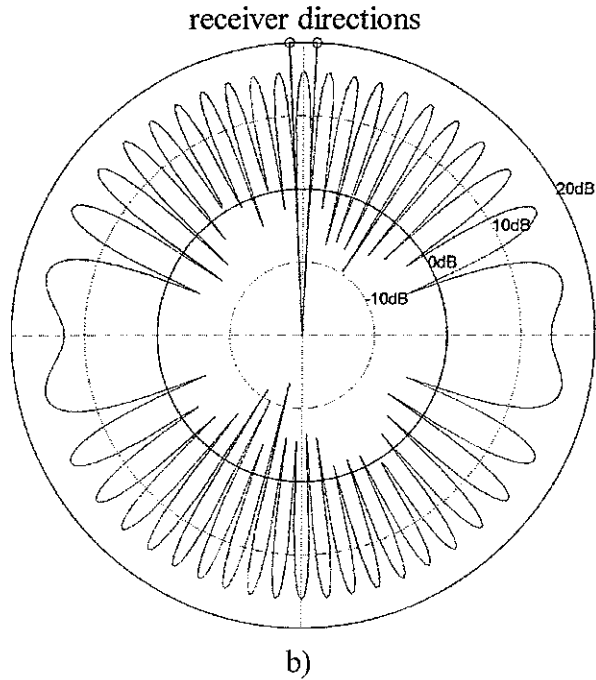
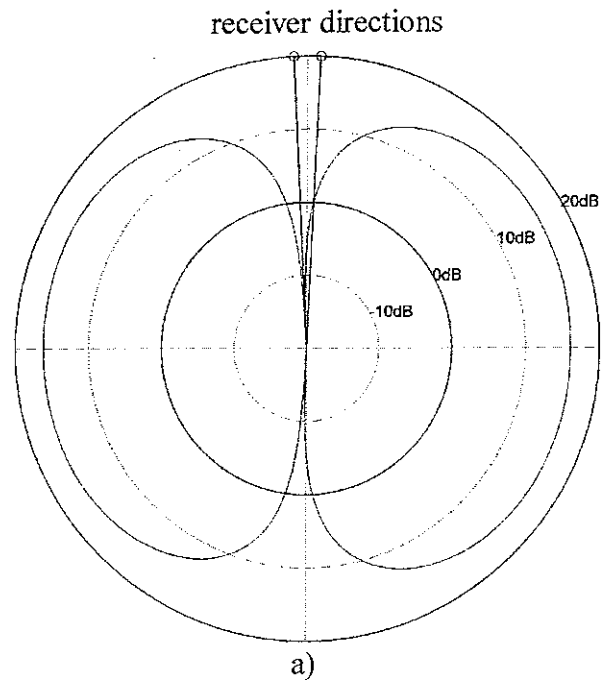


Fig. 25 Maximum far field sound pressure level produced by the transducer pairs with reference to the receiver directions (0dB and $-\infty$ dB). $\nu = 0.9$. a) $n = 0.1$ b) $n = 1.9$

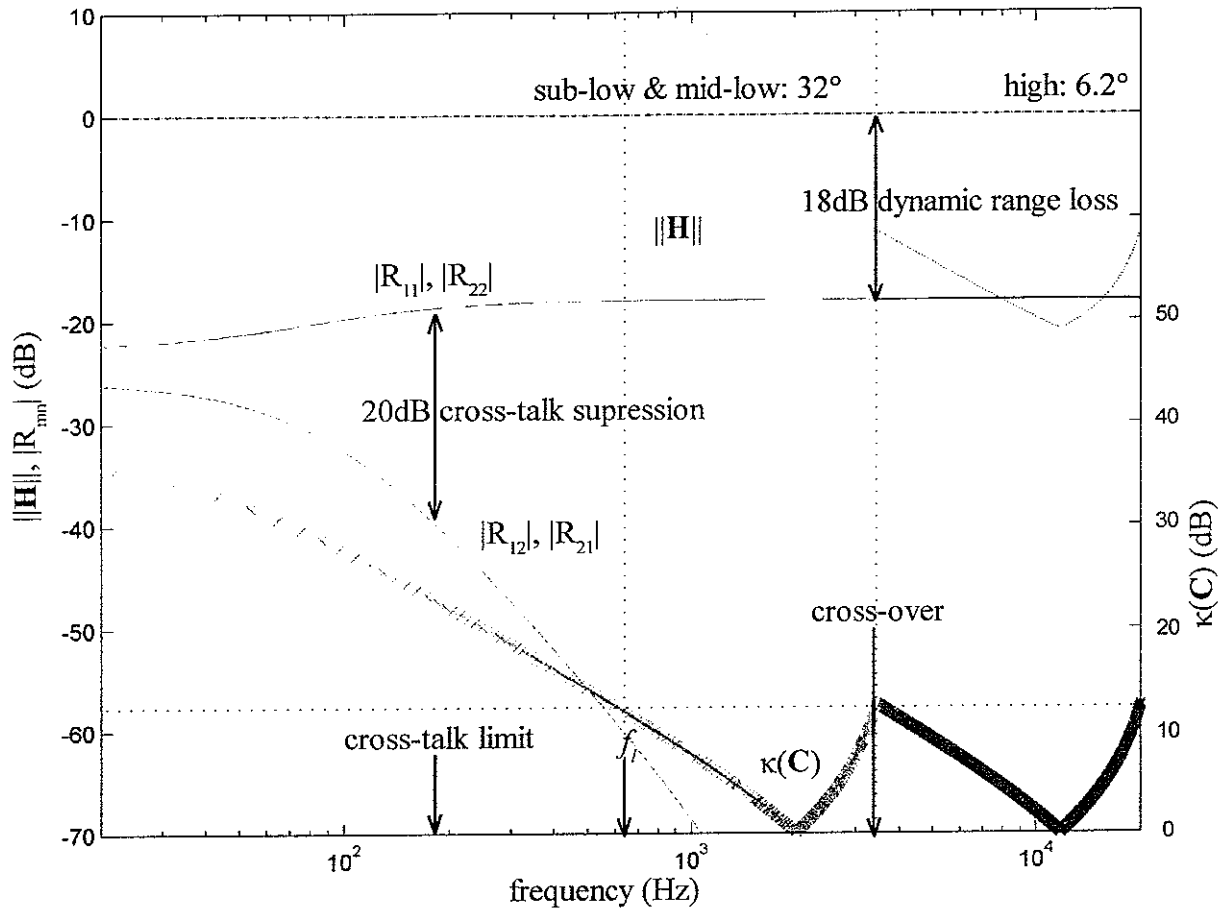
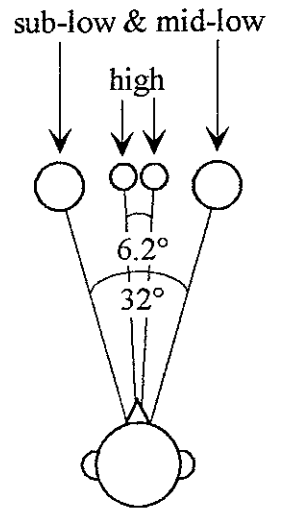


Fig. 26 An example of 2-way system with $n \approx 1$ and $\nu = 0.7$ with regularisation for 18dB dynamic range loss.

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



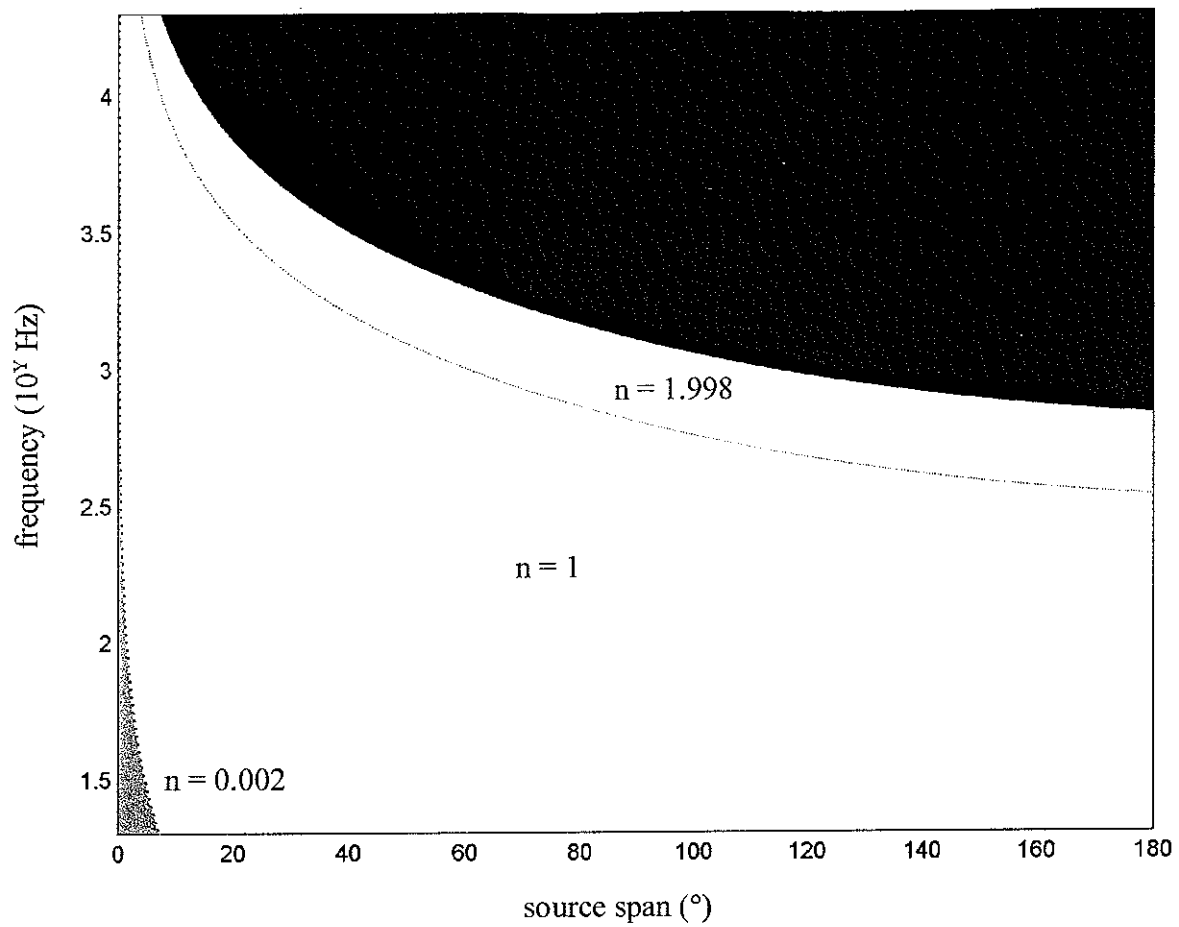


Fig. 27 The frequency/span region for systems with $n \approx 1$ and $\nu = 0.998$, and an example of discretisation for a 1-way system.

white area: $0.002 < n < 1.998$

grey area: $f < f_1$

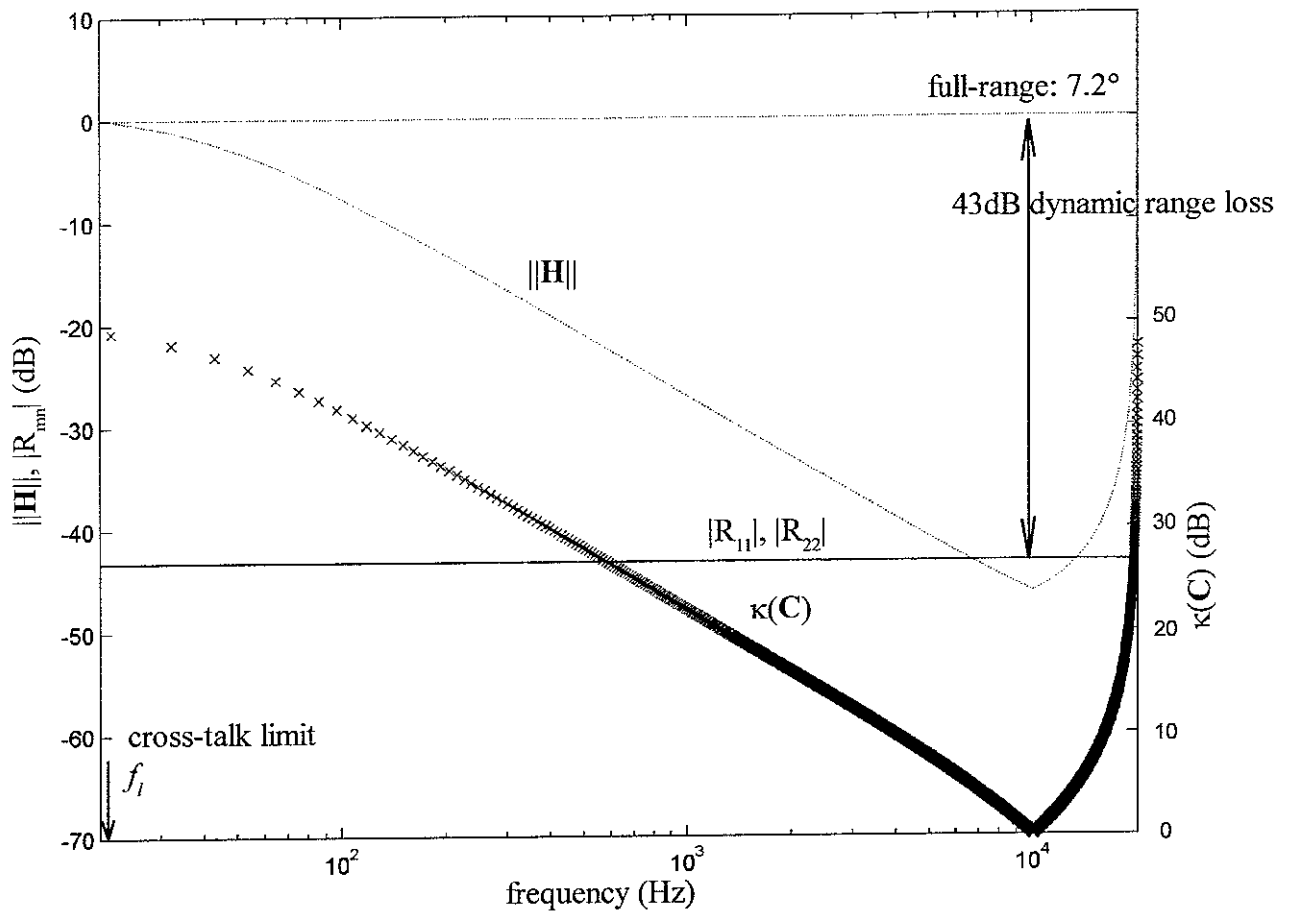
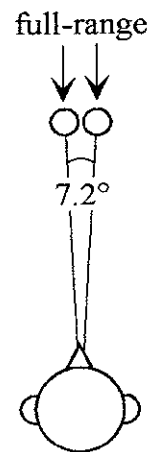


Fig. 28 An example of 1-way system with $n \approx 1$ and $\nu = 0.998$.

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



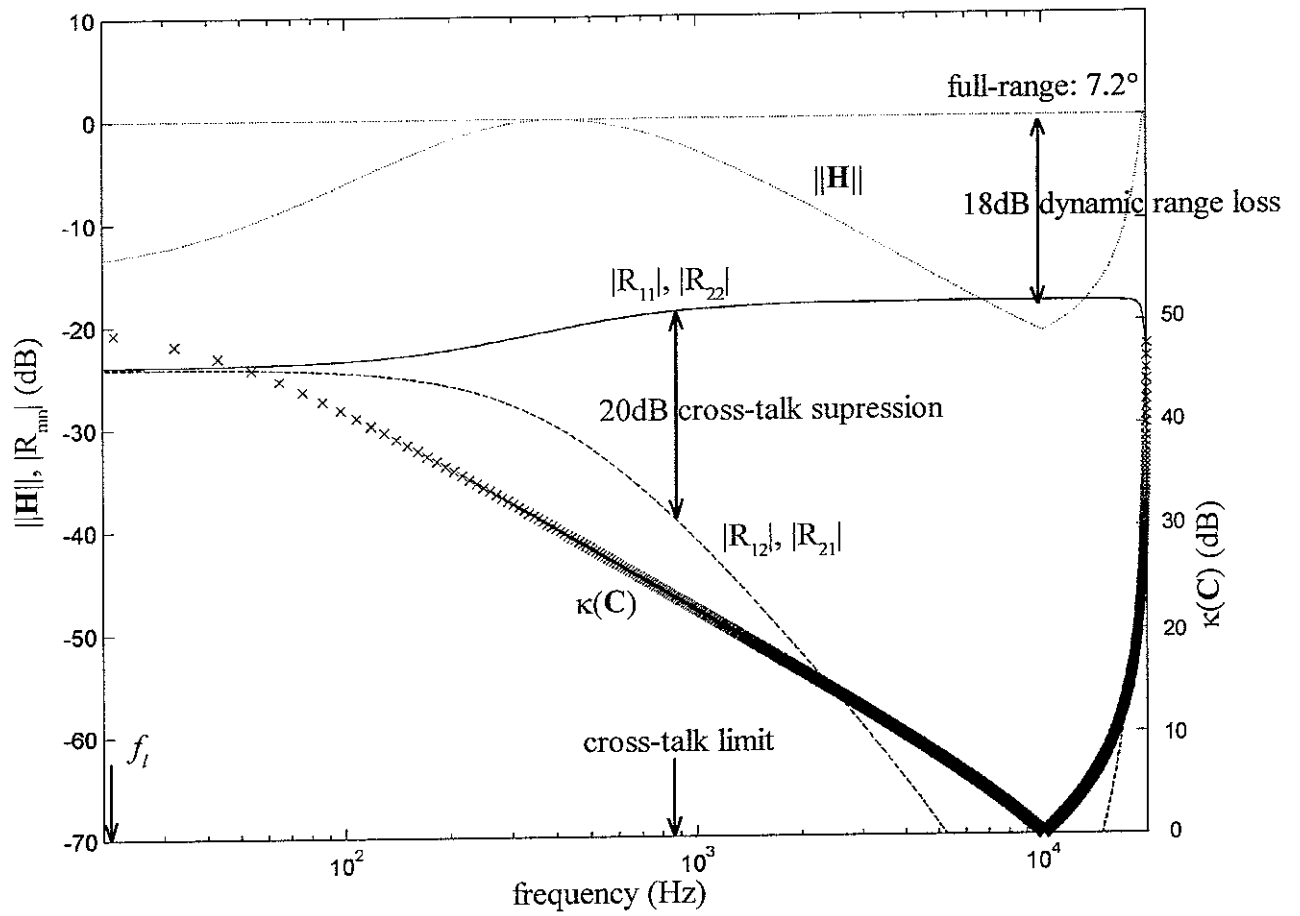
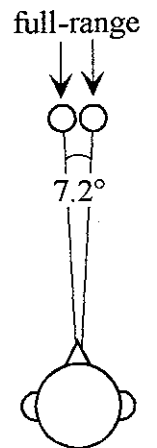


Fig. 29 An example of 1-way system with $n \approx 1$ and $\nu = 0.998$ with regularisation for 18dB dynamic range loss.

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



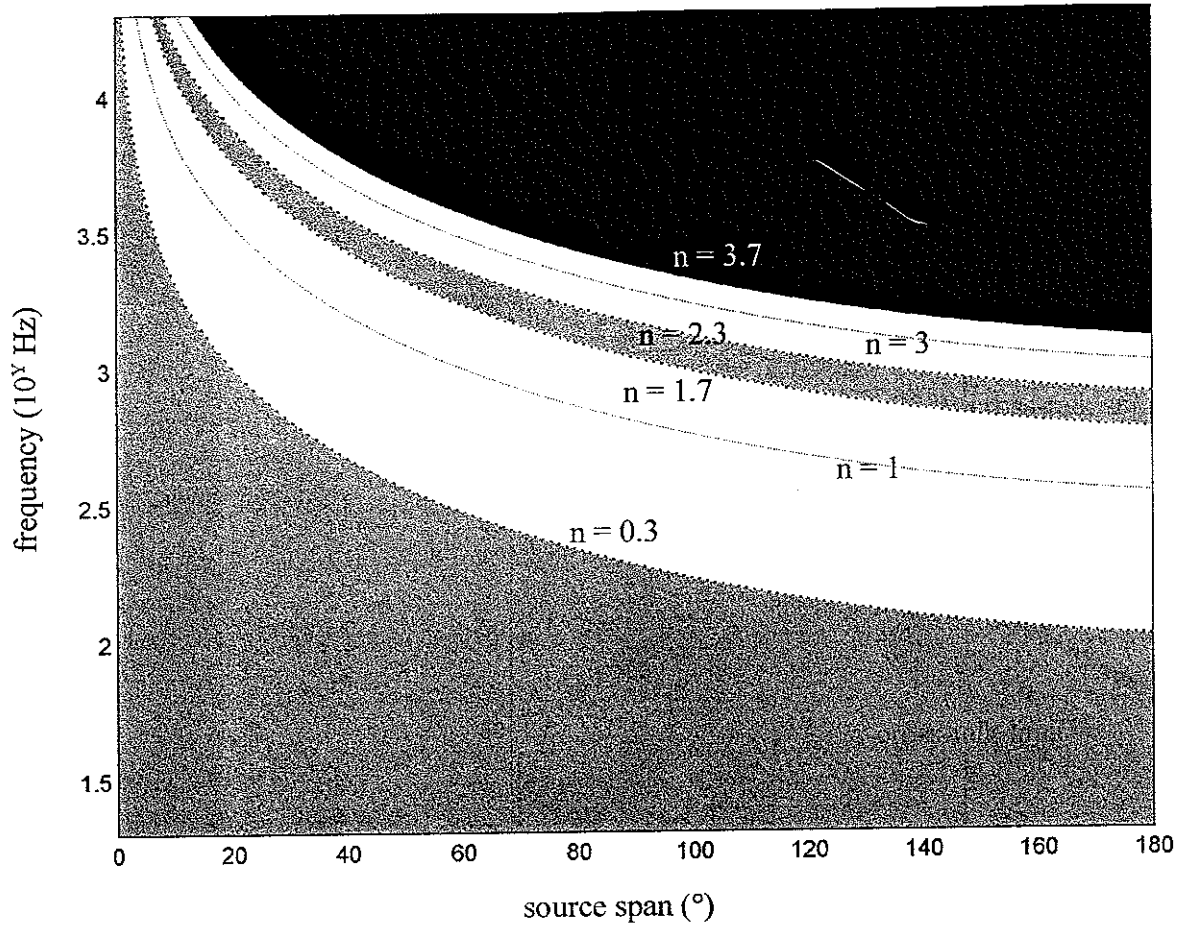


Fig. 30 The frequency/span region for a multi-region systems with $n \approx 1$ and $n \approx 3$ with $\nu = 0.7$, and an example of discretisation for a 1-way system.

white area: $0.3 < n < 1.7$

grey area: $f < f_l(n \approx 1)$ and $f_b(n \approx 1) < f < f_l(n \approx 3)$

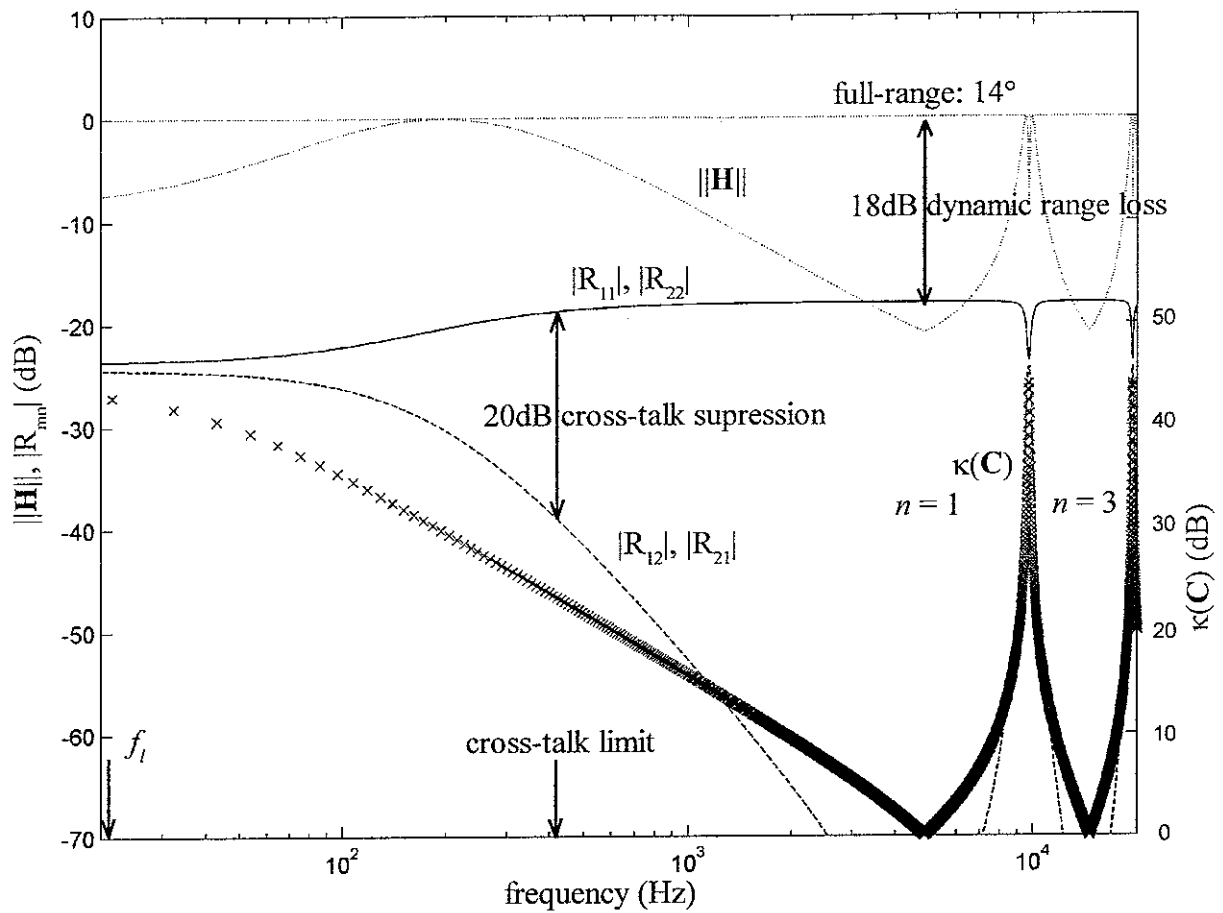
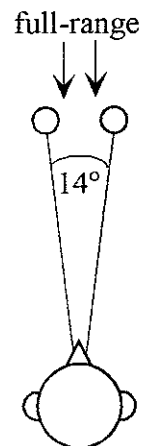


Fig. 31 An example of 1-way multi-region system with $n \approx 1$ and $n \approx 3$ with $\nu = 0.7$, with regularisation for 18dB dynamic range loss.

- source signal level with system inversion ($\|H\|$)
- reproduced signal level with the correct desired signal ($|R_{11}|, |R_{22}|$)
- - - reproduced signal level with the wrong desired signal ($|R_{12}|, |R_{21}|$, cross-talk)
- xxxx condition number $\kappa(C)$



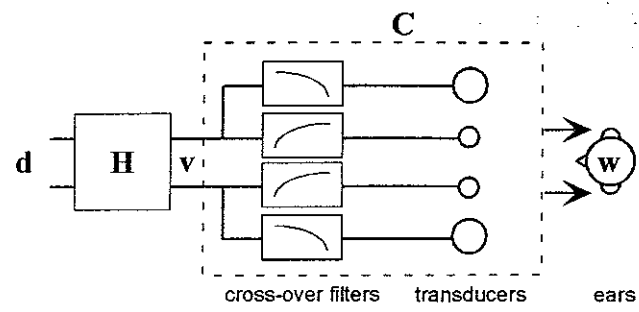
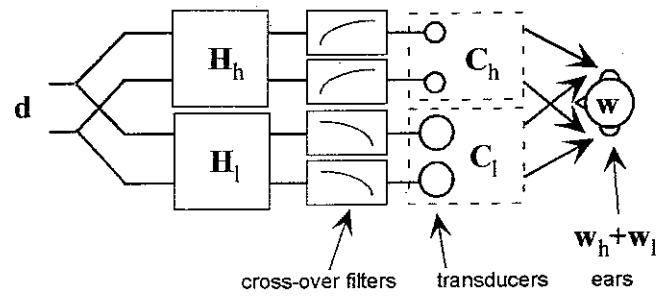
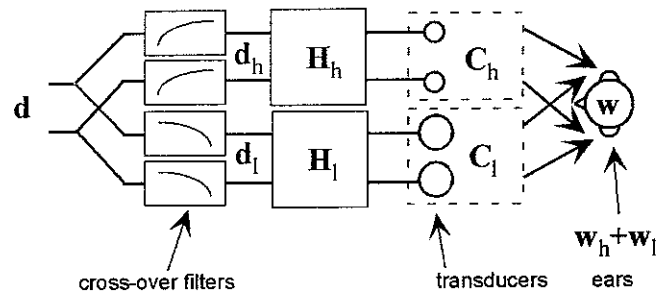


Fig. 32 Block diagrams for cross-over filters and inverse filters when a 2 by 2 plant matrix C is used to design inverse filters.



a)



b)

Fig. 33 Block diagrams for cross-over filters and inverse filters when m (number of driver pairs) of 2 by 2 plant matrices C are used separately to design m inverse filter matrices.
a) Cross-over filters after inverse filters. b) Cross-over filters prior to inverse filters.

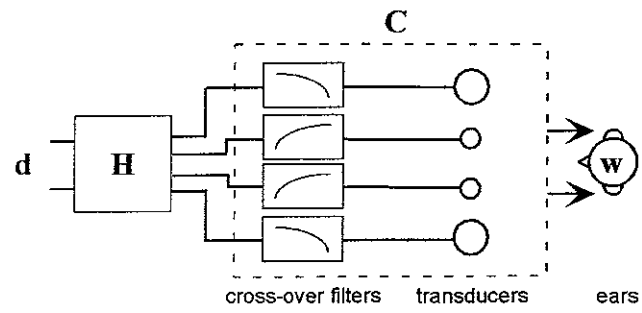


Fig. 34 Block diagrams for cross-over filters and inverse filters when a 2 by $(2 \times m)$ plant matrix **C** is used to design inverse filters.

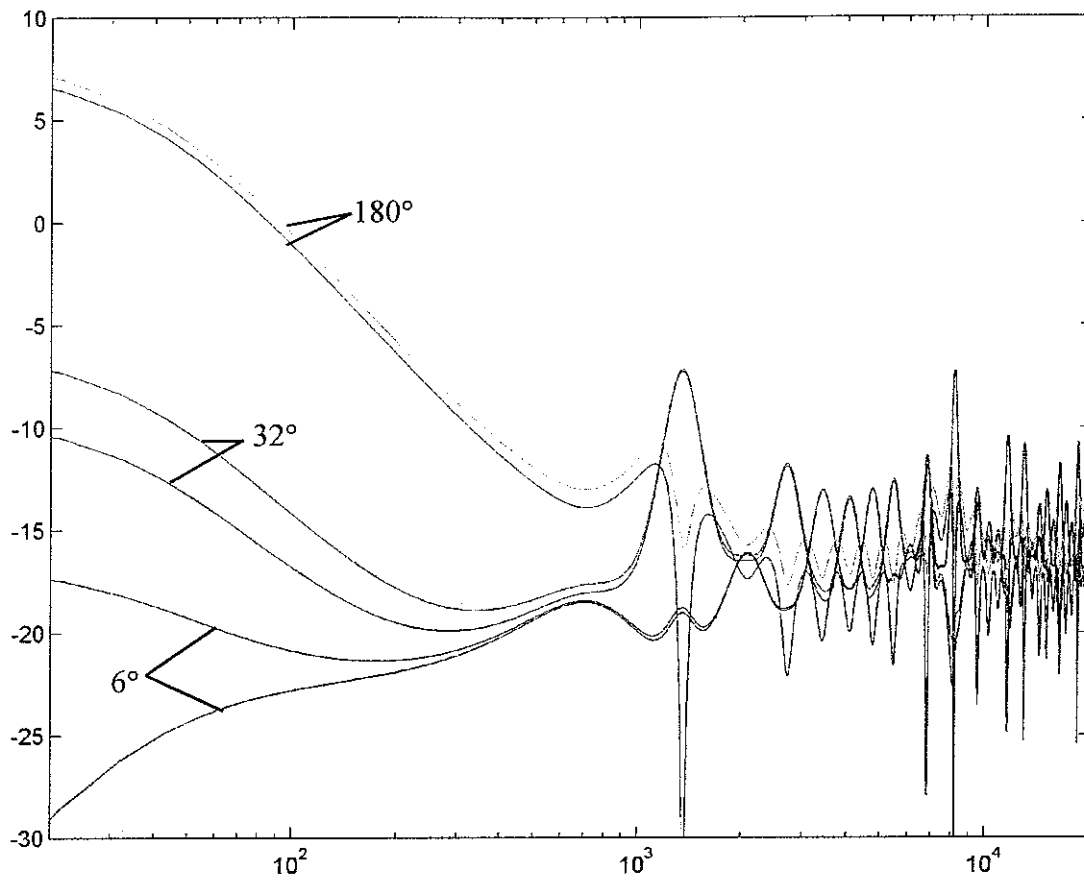


Fig. 35 An example of inverse filters for a multi-channel system (6 channels).