UNIVERSITY OF
Southampton
Institute of Sound and
Vibration Research

ISVR Technical Memorandum

Regularisation and Robustness of Personal Audio Systems

S.J. Elliott and J.Cheer

ISVR Technical Memorandum 995

December 2011

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 the 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.

UNIVERSITY OF SOUTHAMPTON

INSTITUTE OF SOUND AND VIBRATION RESEARCH

SIGNAL PROCESSING AND CONTROL GROUP

Regularisation and Robustness of Personal Audio Systems

by

S. J. Elliott and J. Cheer

ISVR Technical Memomrandum No. 995

December 2011

Authorised for issue by

Prof. P. R. White

Acknowledgements

We are grateful for discussions with Jung-Woo Choi on the results in Section 2. Section 3 also arose out of discussions with Andrew Langley.

Contents

List of Figures

Abstract

As well as being able to reproduce sound in one region of space it would be useful to reduce the level of reproduced sound in other spatial regions, with a “personal audio” system. For mobile devices this is motivated by issues of privacy for the user and the need to reduce annoyance for other people nearby. Such personal audio systems can be realised with arrays of loudspeakers that become superdirectional at low frequencies, when the array dimensions are small compared with the acoustic wavelength. The design of the array then becomes a compromise between performance and array effort, defined as the sum of the squared driving signals. Various methods of formulating this trade-off as a regularisation problem have been suggested and the connection between these formulations is discussed. Large array effort are due to strongly self-cancelling multipole arrays. A concern is then the robustness of such as array to variations in the acoustic environment and driver sensitivity and position. The design of an array that is robust to these uncertainties then leads to a generalisation of regularisation.

List of Symbols

$\mathbf{A}$	Hermitian matrix
b	constant sum of squared pressures in the bright zone
c_0	speed of sound in air
C	Acoustic contrast
d	constant sum of squared pressures in the dark zone, or distance between array elements
e	constant sum of squared source strengths, or error signal for robust filtering problem
E	expectation
f	minimising function
g	impulse response
$\mathbf{g}$	vector of real valued functions
G	Transfer response for robust filtering problem
^{H}	Harmitian, complex conjugate, transpose
k	wavenumber
L	Lagrangian cost function
$\mathbf{p}_B$	vector of complex pressure in the bright zone
$\mathbf{p}_D$	vector of complex pressure in the dark zone
p	acoustic pressure
P	Plant response for robust filtering problem
r	radius or positive real scaling factor
$\mathbf{r}$	vector of filtered reference signals
$\mathbf{R}$	matrix of filtered reference signals
R	Reverberation radius

$\mathbf{q}$	vector of complex source strengths
$\boldsymbol{s}$	Wilkinson number
^{T}	Transpose
u	Input signal
V	Volume
w	Filter coefficient
W	Robust filter for robust filtering problem
$\mathbf{x}_n$	left eigenvector
$\mathbf{y}_b$	right eigenvector
x	Reference signal
y	Output signal
$\mathbf{Z}_B$	matrix of complex transfer responses between sources and bright zone positions
$\mathbf{Z}_D$	matrix of complex transfer responses between sources and bright zone positions
Z_R	free field transfer response
Z_R	reverberant field transfer response
α	energy difference maximisation trade-off factor
β	Regularisation parameter
Δ	Random uncertainty
δ	offset distance
κ	Condition number
λ	Lagrange multiplier
$\boldsymbol{\lambda}$	Vector of Lagrange multiplier
λ_{max}	Maximum eigenvalue
ω	Angular frequency

λ_{min}	Minimum eigenvaluer
ρ_0	air density
ζ	acoustic damping ratio
%	Uncertain quantity
< >	Space average

1. Introduction

The performance of a personal audio system may be quantified by the contrast between the average mean square pressure in a bright zone and the average mean square pressure in a dark zone (Choi and Kim, 2002). An example of such a system is illustrated in Figure 1.1, in which the array of K sources, with complex source strengths, at a given frequency given by the vector $\mathbf{q}$, is used to maximise the sum of the M mean square pressures, whose complex amplitudes are given by the vector $\mathbf{p}_B$, in the bright zone in contrast to the sum of the L mean square pressures, whose complex amplitudes are given by the vector $\mathbf{p}_D$, in the dark zone.

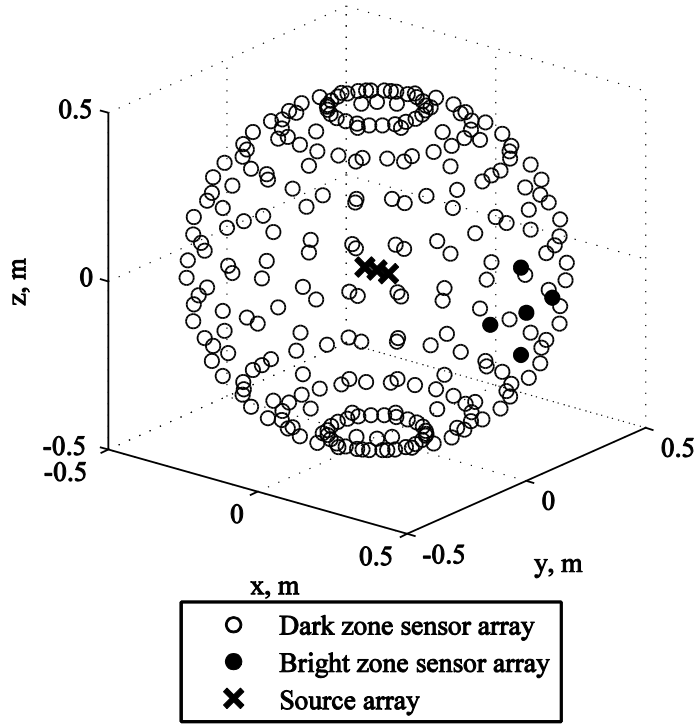


Figure 1.1 Example of a three-dimensional personal audio system, in which a source array, monopoles represented by x in this case, is used to maximise the contrast between the mean square pressure in a bright zone, measured at points with filled circles, and the mean square pressure in a dark zone, measured at points with open circles.

The matrix of complex transfer responses between the sources and the pressures in the bright zone and in the dark zone are $\mathbf{Z}_B$ and $\mathbf{Z}_D$, so that

$$\mathbf{p}_B = \mathbf{Z}_B\mathbf{q} \tag{1.1}$$

$$\mathbf{p}_D = \mathbf{Z}_D\mathbf{q} \tag{1.2}$$

The acoustic contrast, C, is defined at the given frequency by the ratio

$$\frac{\mathbf{p}_B^H\mathbf{p}_B}{\mathbf{p}_D^H\mathbf{p}_D} = \frac{\mathbf{q}^H\mathbf{Z}_B^H\mathbf{Z}_B\mathbf{q}}{\mathbf{q}^H\mathbf{Z}_D^H\mathbf{Z}_D\mathbf{q}} \tag{1.3}$$

This can be cast as a constrained optimisation problem in two separate ways, which are equivalent in this case but turn out to lead to two separate solutions when regularisation is introduced in Section 2.

The first, direct, formulation in which the contrast is maximised is if $\mathbf{p}_\mathrm{B}^H\mathbf{p}_\mathrm{B}$ is maximised with the constraint that $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D}$ is held constant with a value of d, so that the Lagrangian is

$$L = \mathbf{q}^H\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}\mathbf{q} - \lambda(\mathbf{q}^H\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}\mathbf{q} - d), \tag{1.4}$$

where λ is the Lagrange multiplier, which is positive and real in this case. The complex differential of L with respect to the real and imaginary parts of $\mathbf{q}$, $\mathrm{Re}(\mathbf{q})$ and $\mathrm{Im}(\mathbf{q})$, can be written as

$$\frac{\partial L}{\partial \mathbf{q}} = \frac{\partial L}{\partial \mathrm{Re}(\mathbf{q})} + j\frac{\partial L}{\partial \mathrm{Im}(\mathbf{q})} = 2(\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}\mathbf{q} + \lambda\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}\mathbf{q}) \tag{1.5}$$

Setting this differential to zero, and assuming that $\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}$ is invertible, yields

$$\lambda\mathbf{q} = [\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}]^{-1}\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}\mathbf{q} \tag{1.6}$$

The vector of source strengths that maximises the contrast is proportional to the eigenvector of the matrix $[\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}]^{-1}\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}$ corresponding to its largest eigenvalue (Choi and Kim, 2002). This result is illustrated for a two source example in Figure 1.2, with a separation of 4 cm at a frequency of 500 Hz, which shows the contours of constant $\mathbf{p}_\mathrm{B}^H\mathbf{p}_\mathrm{B}$ together with the curve corresponding to the constraint that $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D}$ is equal to d. The two crosses mark the pairs of source strengths corresponding to the largest values of $\mathbf{p}_\mathrm{B}^H\mathbf{p}_\mathrm{B}$ that also satisfy the constraint, which are aligned with the eigenvector of $[\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}]^{-1}\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}$ corresponding to its largest eigenvalue.

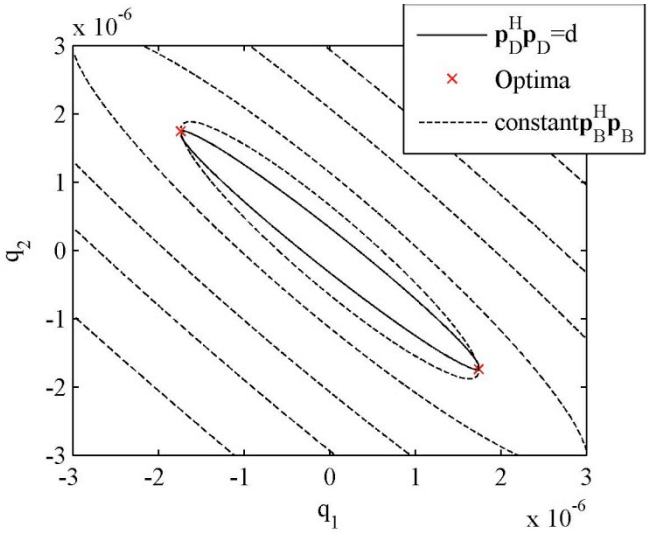


Figure 1.2 Showing contours of constant $\mathbf{p}_B^H\mathbf{p}_B$ as the dashed curves, together with the constraint that $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D} = d$, as the dark curve. The values of the two pairs of source strengths that maximise $\mathbf{p}_B^H\mathbf{p}_B$ while maintaining the constraint are shown as crosses.

The two possible sets of source strengths are equal but of opposite sign, since only the magnitude is $\mathbf{p}_B^H\mathbf{p}_B$ being maximised. The value of the Lagrange multiplier would normally have to be adjusted in such an optimisation, to meet the constraint that $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D}$ is equal to the specified value of *d*. In this case, however, where it is the ratio of $\mathbf{p}_\mathrm{B}^H\mathbf{p}_\mathrm{B}$ to $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D}$ that is being maximised, the solution is independent of the value of *d* and hence any source strength vector is valid provided that it is proportional to the eigenvector identified above.

The alternative, indirect, formulation of the problem of maximising the contrast, which turns out to be more physically meaningful when additional constraints are imposed, is to minimise the value of $\mathbf{p}_\mathrm{D}^H\mathbf{p}_\mathrm{D}$ with the constraint that $\mathbf{p}_\mathrm{B}^H\mathbf{p}_\mathrm{B}$ is held constant with a value of *b*. In this case the Lagrangian is

$$L = \mathbf{q}^H\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}\mathbf{q} + \lambda(\mathbf{q}^H\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}\mathbf{q} - b) \tag{1.7}$$

where the sign of the second term is the opposite of that in equation (1.4) so that λ is again a positive constant. Setting the differential of equation (1.7) to zero, assuming that $\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}$ is invertible, gives the equation

$$\lambda\mathbf{q} = -[\mathbf{Z}_\mathrm{B}^H\mathbf{Z}_\mathrm{B}]^{-1}\mathbf{Z}_\mathrm{D}^H\mathbf{Z}_\mathrm{D}\mathbf{q} \tag{1.8}$$

Since the Lagrangian is now being minimised with respect to $\mathbf{q}$, the solution is the eigenvector corresponding to the smallest eigenvalue of $[\mathbf{Z}_{\mathrm{B}}^H\mathbf{Z}_{\mathrm{B}}]^{-1}\mathbf{Z}_{\mathrm{D}}^H\mathbf{Z}_{\mathrm{D}}$. As $\mathbf{Z}_{\mathrm{D}}^H\mathbf{Z}_{\mathrm{D}}$ is also assumed to be invertible, however, the smallest eigenvalue of this matrix is equal to the largest eigenvalue of its inverse, $[\mathbf{Z}_{\mathrm{D}}^H\mathbf{Z}_{\mathrm{D}}]^{-1}\mathbf{Z}_{\mathrm{B}}^H\mathbf{Z}_{\mathrm{B}}$, and this formulation gives exactly the same result as the formulation above.

The example chosen to obtain the results illustrated in Figure 1.2 was of two sources spaced 4 cm apart in free space, as in Elliott *et al.* (2010), with a bright and dark zone geometry defined as depicted in Figure 1.1. The excitation frequency chosen for this example was 500 Hz, for which the sum of the modulus squared source strengths is about 6 times larger than the square of the source strength required to generate the same magnitude of pressure in the bright zone using a single monopole source. The two source strengths are then of almost equal magnitude and very nearly out of phase with one another. The directivity required to maximise the contrast in this case is generated by the interference of the pressure due to the sources acting as a dipole and the pressure due to them acting as a monopole. Since their separation is small compared with the acoustic wavelength, however, the dipole radiates far less efficiently than the monopole and it is this dipole component that dominates the source strengths. The strongly self-cancelling characteristics of this source array are similar to those seen in superdirective beamforming (Van Veen and Buckley, 1988).

The self-cancelling character of the source array at low frequencies gives rise to two problems in practice. First, the loudspeakers used to generate the source strengths have to be driven very hard, compared with a single loudspeaker, to generate the same pressure in the bright zone. This means that the individual loudspeakers have to be larger and require significantly higher levels of driving power. Secondly, the responses of the individual loudspeakers have to be extremely well matched in order that the small monopole component of their drive signals is accurately reproduced when compared with the large dipole component of these signals.

Both of these problems can be attributed to the ill-conditioning of the matrix $\mathbf{Z}_{\mathrm{D}}^H\mathbf{Z}_{\mathrm{D}}$, which has to be inverted when the largest eigenvalue of $[\mathbf{Z}_{\mathrm{D}}^H\mathbf{Z}_{\mathrm{D}}]^{-1}\mathbf{Z}_{\mathrm{B}}^H\mathbf{Z}_{\mathrm{B}}$ is being calculated. A well-known method of reducing such ill-conditioning problems is to regularise the matrix being inverted. Physically this has the effect of reducing the sum of the modulus squared source strengths in this application, as it does in active control (Elliott, 2001), and thus reducing the distance of the required source strengths from the origin in Figure 1.2, for example. The sum of the modulus squared source strengths is called the array effort.

This report first considers the regularisation of the personal audio problem of maximising the contrast by constraining the array effort. Regularisation is first considered in the frequency domain and then in a time domain formulation. The problem is then considered of making the performance of the personal audio system robust to variations in the transfer responses between the sources and the microphones in the bright and dark zones. The robustness of the system to variations of the positions and sensitivities of the individual drivers is then discussed.

2. Regularisation in the Frequency Domain

We now consider the problem of limiting the array effort while maximising the contrast. It is physically more meaningful to formulate the problem using a generalisation of the second, indirect, method discussed in Section 1. We thus seek to minimise the value of $\mathbf{p}_{\mathrm{D}}^{H}\mathbf{p}_{\mathrm{D}}$ with both the constraint that $\mathbf{p}_{\mathrm{B}}^{H}\mathbf{p}_{\mathrm{B}}$ is held constant with a value of b and that $\mathbf{q}^{H}\mathbf{q}$ is held constant with a value of e. The Lagrangian for this problem can be written as

$$L = \mathbf{q}^{H}\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \lambda_1(\mathbf{q}^{H}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - b) + \lambda_2(\mathbf{q}^{H}\mathbf{q} - e) \tag{2.1}$$

The differential of the Lagrangian with respect to the real and imaginary parts of $\mathbf{q}$, divided by two, is then

$$\frac{1}{2}\frac{\partial L}{\partial \mathbf{q}} = \mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \lambda_1\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} + \lambda_2\mathbf{q} \tag{2.2}$$

Assuming that $\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}$ is invertible, this differential is zero if

$$\lambda_1\mathbf{q} = -[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}]^{-1}[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}]\mathbf{q}. \tag{2.3}$$

So the solution for the minimum we are seeking is proportional to the eigenvector corresponding to the smallest eigenvalue of $[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}]^{-1}[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}]$, which is the largest eigenvalue of $[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}]^{-1}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}$. Although the contrast achieved is independent of the numerical value of λ_1, as discussed above, the magnitude of $\mathbf{q}$ still has to be adjusted if we need to ensure that $\mathbf{p}_{\mathrm{B}}^{H}\mathbf{p}_{\mathrm{B}}$ has the specific value b. Similarly, the value of λ_2, in the solution for $\mathbf{q}$, has to be adjusted in order to satisfy the constraint that $\mathbf{q}^{H}\mathbf{q}$ is equal to e. This constraint will not be active if e is chosen to be larger than that required in the unconstrained problem, however, in which case λ_2 will be zero.

Figure 2.1 shows the trade-off obtained between the array effort and the acoustic contrast when using this regularised solution for the vector of source strengths in the two source personal audio problem outlined in Section 1 at 100 Hz. Each point on this curve has been obtained with a different positive value of λ_2 ranging from zero, at the tip, to a very high value, at the base. For any given constraint on the array effort, the maximum possible value of the contrast can thus be read off from this curve. This trade-off between performance and effort has also been discussed by Boone *et al.* (2009) for a similar loudspeaker array problem.

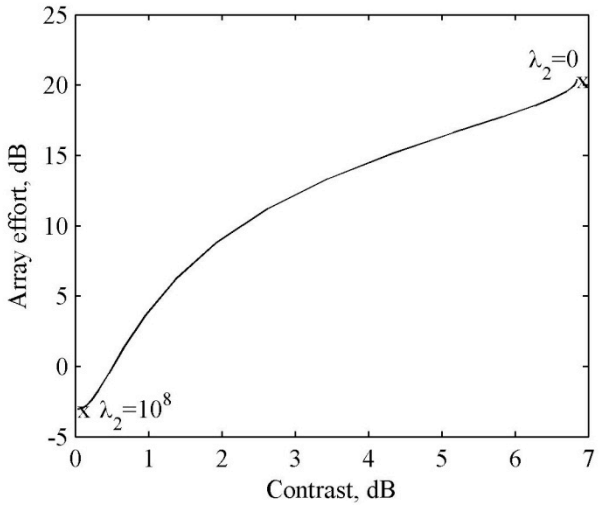


Figure 2.1 Trade-off between the array effort and the acoustic contrast achieved when using the indirect method for the regularised solution to the two source problem outlined in Section 1.

If very high values of λ_2 are used, so that the contrast is the same as a monopole source, i.e. 0 dB, the array effort is still reduced by 3 dB compared with a monopole. This is because the two source strengths are equal in this case, but their magnitudes are each ½ of the equivalent monopole. The array effort is thus $10\log_{10}$ of two times ¼, which is about -3 dB.

It is interesting to note that this solution is the same as that for the problem of minimising the combined function $\mathbf{p}_{\mathrm{D}}^{H}\mathbf{p}_{\mathrm{D}} + \beta\,\mathbf{q}^{H}\mathbf{q}$ with the single constraint that $\mathbf{p}_{\mathrm{B}}^{H}\mathbf{p}_{\mathrm{B}}$ is equal to b, for which the Lagrangian is

$$L = \mathbf{q}^{H}\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \beta\mathbf{q}^{H}\mathbf{q} + \lambda_1(\mathbf{q}^{H}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - b). \tag{2.4}$$

Half the differential of the Lagrangian is then

$$\frac{1}{2}\frac{\partial L}{\partial \mathbf{q}} = \mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \beta\mathbf{q} + \lambda_1\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q}, \tag{2.5}$$

which is zero if

$$\lambda_1\mathbf{q} = -[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}]^{-1}[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \beta\mathbf{I}]\mathbf{q}, \tag{2.6}$$

and so the solution is the eigenvector corresponding to the largest eigenvalue of $[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \beta\mathbf{I}]^{-1}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}$.

It is also interesting to consider the direct approach to this problem, which is given by maximising $\mathbf{p}_{\mathrm{B}}^{H}\mathbf{p}_{\mathrm{B}}$ with the constraints that $\mathbf{p}_{\mathrm{D}}^{H}\mathbf{p}_{\mathrm{D}}$ is held constant with a value of d and that $\mathbf{q}^{H}\mathbf{q}$ is held constant with a value of e, for which the Lagrangian is

$$L = \mathbf{q}^{H}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - \lambda_1(\mathbf{q}^{H}\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} - d) - \lambda_2(\mathbf{q}^{H}\mathbf{q} - e) \tag{2.7}$$

Half the differential of this Lagrangian is

$$\frac{1}{2}\frac{\partial L}{\partial \mathbf{q}} = \mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - \lambda_1\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} - \lambda_2\mathbf{q}, \tag{2.8}$$

which is zero if

$$\lambda_1\mathbf{q} = [\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}]^{-1}[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} - \lambda_2\mathbf{I}]\mathbf{q}. \tag{2.9}$$

The solution is thus proportional to the eigenvector corresponding to the largest eigenvalue of $[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}]^{-1}[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} - \lambda_2\mathbf{I}]$.

The trade-off using this direct approach does not give meaningful results for values of λ_2 greater than around 1850. For values of λ_2 greater than this, there is a problem in inverting the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$ matrix which, due to the ill-conditioning of the complete matrix, results in inaccurate calculation of the eigenvalues. In order to achieve meaningful results it is necessary to add regularisation to the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$ matrix and calculate the eigenvector corresponding to the largest eigenvalue of $[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + r\mathbf{I}]^{-1}[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} + \lambda_2\mathbf{I}]$ where r is a positive, real scaling factor. The trade-off for this case is shown in Figure 2.2 where the value of r has been set to zero where the direct solution is stable – i.e. for high values of contrast – and has then been increased linearly to $r = 10^6$ from the points labelled on the plot. An increasing value of r is necessary since for high levels of contrast, where λ_2=0, if a high value of r is employed the matrix to be inverted is dominated by this value and very little contrast control can be achieved. From this plot it can be seen that the solution, when stable, is identical to the indirect solution shown in Figure 2.1.

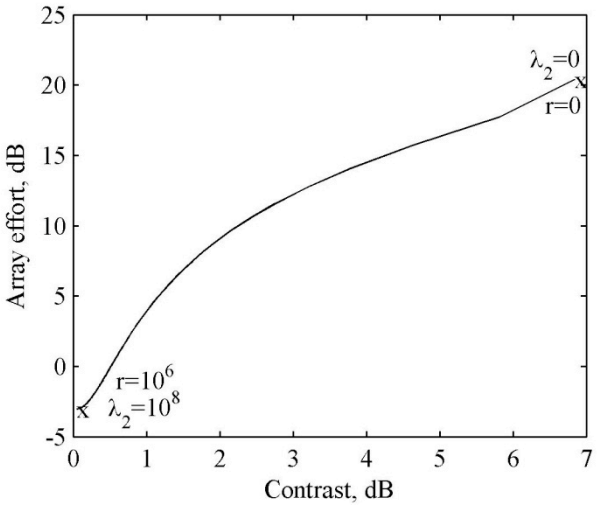


Figure 2.2 Trade-off between the array effort and the acoustic contrast achieved when using the direct method for the regularised solution to the two source problem outlined in Section 1. This trade-off is obtained with additional regularisation in the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$ matrix.

Although Figure 2.2 shows an identical solution to that obtained from the indirect regularised problem, it has been achieved through additional regularisation and does not make as much physical sense. This can be demonstrated by again considering a related problem which yields the same solution. In this case, the related problem is that of maximising the value of the combined function $\mathbf{p}_{B}^{H}\mathbf{p}_{B}-\beta\mathbf{q}^{H}\mathbf{q}$ with the single constraint that $\mathbf{p}_{\mathrm{D}}^{H}\mathbf{p}_{\mathrm{D}}$ is equal to d, for which the Lagrangian is given by

$$L=\mathbf{q}^{H}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q}-\beta\mathbf{q}^{H}\mathbf{q}+\lambda(\mathbf{q}^{H}\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q}-d). \tag{2.10}$$

Half the differential of the Lagrangian is then

$$\frac{1}{2}\frac{\partial L}{\partial\mathbf{q}}=\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q}-\beta\mathbf{q}+\lambda\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q}, \tag{2.11}$$

which is zero if

$$\lambda\mathbf{q}=-[\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}]^{-1}[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}-\beta\mathbf{I}]\mathbf{q}, \tag{2.12}$$

which, apart from the sign change in $\mathbf{q}$ that does not affect the results, is the same as equation (2.9) if β is set equal to λ_2 and λ is set equal to λ_1.

Finally, we consider the analytic solution to the problem of maximising the acoustic energy difference, as suggested by Shin *et al.* (2010), as a method of controlling the array effort. The acoustic energy

difference is defined to be $\mathbf{p}_{\mathrm{B}}^{H}\mathbf{p}_{\mathrm{B}} - \alpha\mathbf{p}_{\mathrm{D}}^{H}\mathbf{p}_{\mathrm{D}}$, which is maximised while maintaining the constraint that $\mathbf{q}^{H}\mathbf{q}$ is held equal to the value e. The Lagrangian for this formulation is

$$L = \mathbf{q}^{H}\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - \alpha\mathbf{q}^{H}\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \lambda(\mathbf{q}^{H}\mathbf{q} - e), \tag{2.13}$$

for which half the differential is

$$\frac{1}{2}\frac{\partial L}{\partial \mathbf{q}} = \mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}}\mathbf{q} - \alpha\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\mathbf{q} + \lambda\mathbf{q}. \tag{2.14}$$

This differential is zero if

$$\lambda\mathbf{q} = -[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} - \alpha\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}]\mathbf{q}, \tag{2.15}$$

and so $\mathbf{q}$ must be proportional to the eigenvector corresponding to the largest eigenvalue of $\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} - \alpha\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$.

The trade-off between array effort and contrast for this solution is shown in Figure 2.3, for which α is varied from 0, at the base, to a value of 5 at the tip. For values of α between 0 and 1 this solution is exactly the same as that shown in Figure 2.1. This is because the function that is being set to zero here, in equation (2.14), is exactly the same as the function that is being set to zero in the indirect method, that is, equation (2.2) with $\lambda_1 = -\frac{1}{\alpha}$ and $\lambda_2 = -\frac{\lambda}{\alpha}$. If α is taken to be greater than about 1, however, the contrast decreases, even though the control effort continues to rise, as shown by the curve in Figure 2.3. This is because λ_2, which is varied to produce Figure 2.1 depends on both λ and α here.

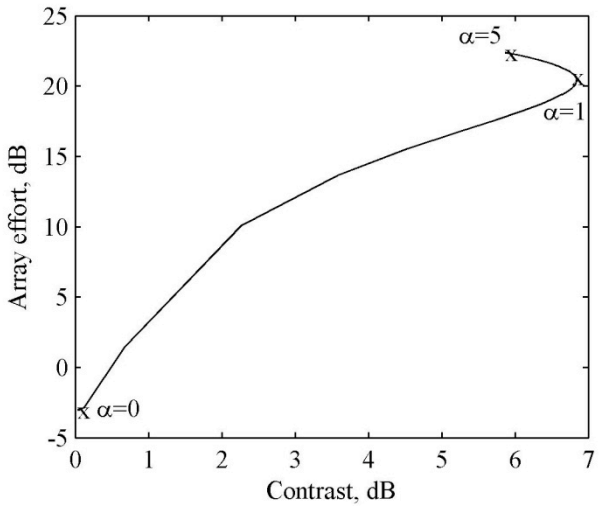


Figure 2.3 Trade-off between the array effort and the acoustic contrast achieved when using the acoustic energy difference method for the regularised solution to the two source problem outlined in Section 1.

Based on the preceding results, it is clear that there are a number of methods of solving the defined personal audio problem that provide the same trade-off between acoustic contrast and array effort. However, as shown for the direct method (equation (2.9) and Figure 2.2) the ability to compute the actual solution may be impaired by the particular method chosen. In computing problems, such as optimisation, ill-conditioning may describe any computation whose output values are very sensitive to small changes in the input data and, therefore, in the presence of small errors or perturbations an ill-conditioned problem may produce highly inaccurate results. There are two properties of a matrix that may be considered with respect to conditioning, these are the inverse of the matrix and the calculation of its eigenvalues.

The normally-used condition number of a symmetric matrix with respect to its inverse is given by the ratio of its largest to its smallest eigenvalue, which is (Wilkinson, 1965),

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}}. \tag{2.16}$$

Out of the three personal audio optimisation methods considered, both the direct and indirect methods require matrix inversion and, therefore, may be susceptible to errors due to ill-conditioning with respect to the inverse operation. For the direct method the matrix to be inverted is, according to equation (2.9), $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$, and using equation (2.16) the condition number has been calculated as $\kappa = 1.86\times10^{4}$. A large condition number compared to unity indicates that the matrix is close to singular and, therefore, conditioning may be an issue. For the indirect optimisation method the matrix to be inverted is $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}$ and, therefore, varies with the constraint on the array effort. Figure 2.4 shows the change in the condition number, κ, of the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}$ matrix for an increasing value of the Lagrange multiplier λ_2. From this plot it can be seen that the condition number of the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \lambda_2\mathbf{I}$ matrix rapidly decreases with λ_2, indicating an improvement in the conditioning of the matrix with respect to inversion. The results in Figure 2.4 suggest that the matrix inversion operation may be the cause of the problems encountered using the direct method. If this were the only cause of these problems, however, then for low values of λ_2, which correspond to high acoustic contrast, then it would be expected that computational issues due to ill-conditioning would arise in both direct and indirect methods. Sensitivity in performance due to ill-conditioning when an array is driven close to its maximum acoustic contrast performance may be expected in practice; however, in the context of simulations where errors are minimal this does not appear to be the case. Therefore, it is pertinent to investigate the conditioning with respect to the eigenvalue calculation.

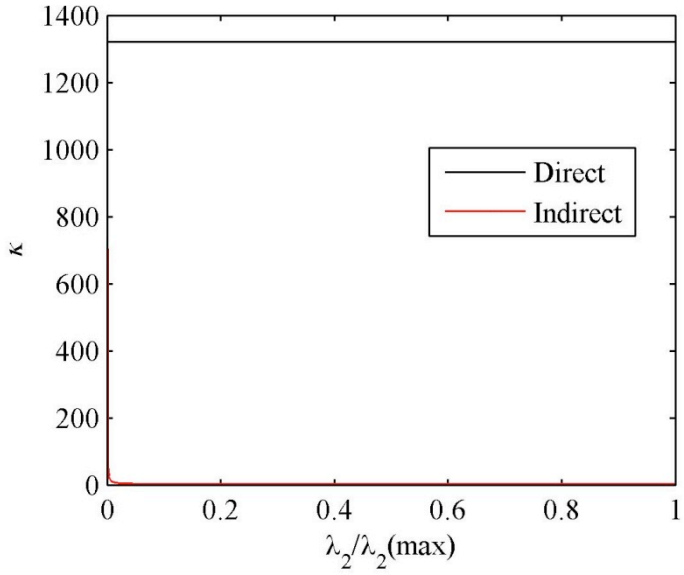


Figure 2.4 The normal condition number, κ, of the $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}$ and $\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}+\lambda_2\mathbf{I}$ matrices employed in the direct and indirect personal audio optimisation methods respectively plotted against the normalised Lagrange multiplier, λ_2.

The effect of perturbations in the elements of a matrix on its eigenvalues may be described by the Wilkinson number of the matrix (Wilkinson, 1965). The Wilkinson number is given by (Golub and Van Loan, 1996),

$$s(\lambda)=\left|\mathbf{y}^{H}\mathbf{x}\right| \tag{2.17}$$

where $\mathbf{x}_n$ and $\mathbf{y}_n$ are the left and right eigenvectors with $\|\mathbf{x}_n\|_2=\|\mathbf{y}_n\|_2=1$ and satisfy the equations,

$$\begin{aligned}\mathbf{A}\mathbf{x}_n&=\lambda_n\mathbf{x}_n\\ \mathbf{y}_n{}^{H}\mathbf{A}&=\lambda_n\mathbf{y}_n{}^{H}\end{aligned} \tag{2.18}$$

where **A** is a square Hermitian matrix and λ_n is an eigenvalue. For a non-defective matrix **A**, which has a complete set of linearly independent eigenvectors, $s(\lambda)=1$, however, for a defective matrix it is necessary to use the Jordan decomposition and $|\mathbf{y}^{\mathrm{H}}\mathbf{x}|$ is not necessarily equal to unity.

The Wilkinson number is the cosine of the angle between the left and right eigenvectors associated with the eigenvalue, λ_n, and is the reciprocal of the *eigenvalue condition number* (Golub and Van Loan, 1996); however, we shall refer to the Wilkinson number given by equation (2.17). For a perturbation of order ε the perturbation in the eigenvalue, λ_n, can be roughly related to the Wilkinson

number, $s(\lambda)$, by the value $\varepsilon / s(\lambda)$ (Golub and Van Loan, 1996). Therefore, if $s(\lambda)$ is small compared to unity the problem is ill-conditioned and small perturbations in the elements of the matrix will produce disproportionately large perturbations in the calculated eigenvalues.

Figure 2.5 shows how the Wilkinson number of the matrix from which the optimal source strengths are calculated for each optimisation method varies with the value of the Lagrange multiplier – λ_2 for the direct and indirect methods and α for the energy difference method – normalised to its maximum value for the three considered methods. Although there are two eigenvalues for the two source array, and therefore two Wilkinson numbers for each method, they both give identical values in this case. From this plot it can be seen that the Wilkinson number is very low for the direct method calculated without additional regularisation according to equation (2.9) and this can be related to the problems encountered in computing the optimal solution. That is, errors due to the numerical inaccuracies in calculating the matrix inverse lead to significant errors in the calculated eigenvalues due to the low Wilkinson number and, therefore, ill-conditioning. For the indirect solution it can be seen that the condition number is also very low for small values of λ_2 which correspond to low constraints on the array effort. However, it can also be seen that a small increase in the array effort constraint leads to a rapid increase in the eigenvalue conditioning. For the energy difference maximisation method proposed by Shin *et al* (2010) it can be seen that a Wilkinson number close to unity is achieved for all values of α. This indicates that errors in the elements of the matrix $\left[\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} - \alpha\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}}\right]$ will only result in comparable sized errors in its eigenvalues and, therefore, this method will be robust to errors, which supports the comments made by Shin *et al* (2010).

Figure 2.6 shows how the Wilkinson number varies with acoustic contrast for the indirect and energy difference methods. For the indirect solution it can be seen that the Wilkinson number only tends toward unity for low levels of acoustic contrast, which correspond to low array effort and a less self-cancelling array. For a practical implementation the errors in the elements of the matrix may be larger than those experienced due to the numerical errors in calculating the matrix inverse and, therefore, the indirect solution may also have robustness issues when high levels of contrast are attempted. Once again Figure 2.6 shows that for the energy difference method the Wilkinson number is close to unity for all levels of acoustic contrast.

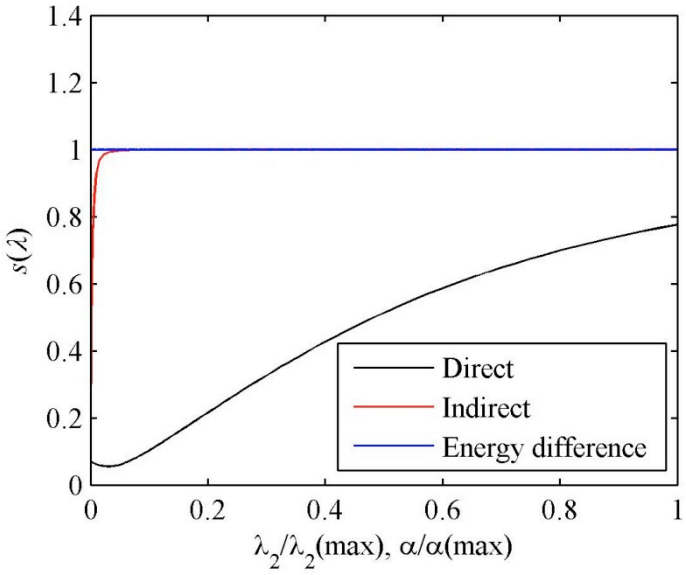


Figure 2.5 The Wilkinson number plotted as a function of λ_2 normalised to the maximum value of λ_2 for the direct and indirect methods, and plotted as a function of α normalised to the maximum value of α_{max} for the energy difference solution.

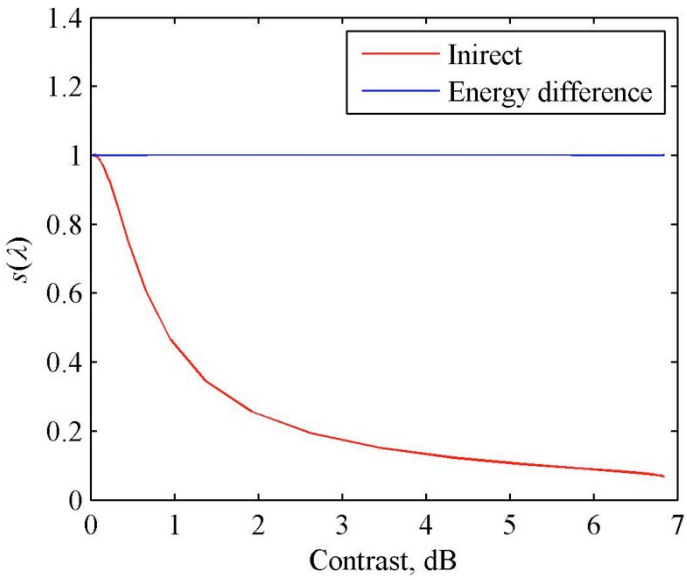


Figure 2.6 Reciprocal of the condition number plotted as a function of the acoustic contrast for the indirect and energy difference solutions.

The three formulations of the effort-constrained maximisation problem thus yield equivalent solutions. However, there may be numerical problems in the calculation using the direct method, and the acoustic energy difference method requires the judicious choice of the parameter α to obtain results in the required range. The indirect method, however, is numerically well conditioned even for small values of β in equation (2.6) and gives meaningful results for any positive value of β.

It is interesting to also consider the trade-off between the array effort and the acoustic contrast achieved for the problem outlined in Section 1, but with three sources. This is plotted in Figure 2.7, with a logarithmic scale for the array effort. Two clearly separate regions can be observed. For values of contrast below about 7 dB, the array effort rises smoothly to about 12 dB, and follows the corresponding locus for the two-source array, also plotted as the dashed line in this figure, apart from a factor of about $10\log_{10}(3/2)$. To achieve higher levels of contrast, the array effort then has to increase significantly, only reaching about 11 dB when the array effort is about 40 dB, a factor of about 1,000 larger than that required to give an acoustic contrast level of 10 dB.

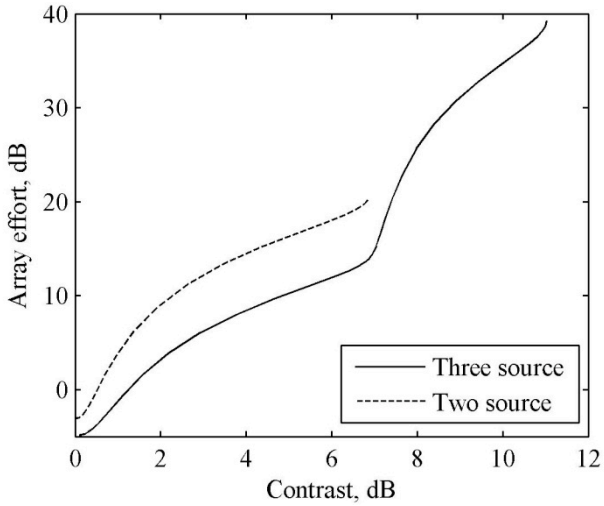


Figure 2.7 The trade-off between the level of the array effort and the contrast level for the three source array. That for the two-source array is shown dashed.

The results presented so far have all been for a single frequency, 100 Hz, chosen so that the source separation is small compared with the acoustic wavelength. Figure 2.8 shows the variation in the contrast, and the corresponding array effort, when the contrast is maximised without regularisation at different frequencies for the two and three source arrays.

Also shown in Figure 2.8 is the contrast and array effort when the indirect method is used with regularisation, using a constant value β. It is also possible to adjust the value of β at different frequencies to impose a fixed constraint on the array effort, as in Figure 2.9 for example where a maximum array effort of 20 dB, relative to a single monopole, has been imposed.

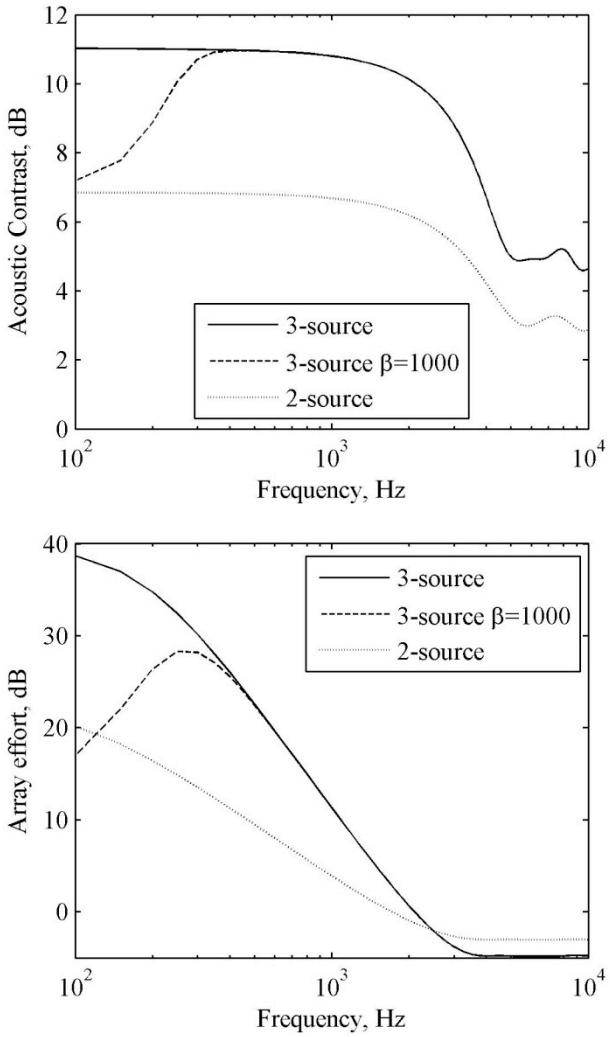


Figure 2.8 The acoustic contrast achieved and the required array effort when the three source personal audio problem is solved at different frequencies. The solution for the two source array considered in Section 1 is also shown, dashed, for reference.

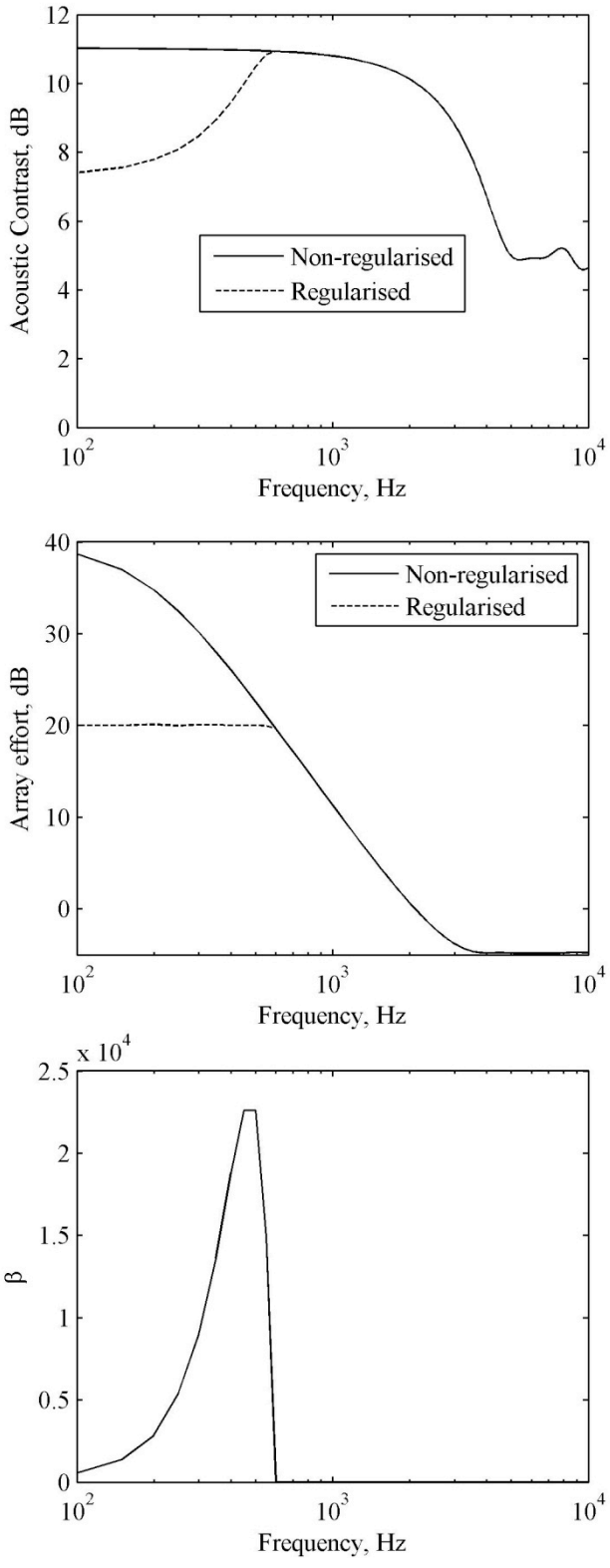


Figure 2.9 The acoustic contrast achieved and the required array effort when the three source personal audio problem is solved at different frequencies with a frequency dependent value of β, which is also shown, set to limit the array effort to a maximum of 20 dB at each frequency (dashed line).

To achieve the hypercardioid directivity the pressures from the monopole and dipole components must be of the same order of magnitude. The far-field pressure produced by a dipole source is, however, a factor of *kd* lass than that of a monopole, where *k* is the acoustic wavenumber and *d* is the separation between the sources, and so the dipole component has to be a factor of 1/*kd* larger than the monopole. This explains why the low frequency effort is proportional to the reciprocal of the frequency response squared for this array in Figure 2.8. It also provides insight into the sensitivity of the array to small changes in driver response, since the radiation from the monopole component would be much greater than the dipole component and the delicate balance between the two components would be lost.

The monopole and dipole components of the two-source array can be obtained from a simple version of the multipole expansion (Pierce, 1981). Consider the complex pressure in line and at a distance *r* away from the centre of the two-source monopole array, with source strengths q_1 and q_2, separated by a distance *d*:

$$p(r) = \frac{j\omega\rho_0 e^{-jkr}}{4\pi r}\left(q_1 e^{-jkd/2} + q_2 e^{+jkd/2}\right) \tag{2.19}$$

where ρ_0 and c_0 are the density and speed of sound and *k* is the wavenumber, equal to ω/c_0. If the array is compact so that *kd* is small compared to unity, the pressure can be approximated by taking the first order expansion of the exponentials so,

$$p(r) = \frac{j\omega\rho_0 e^{-jkr}}{4\pi r}\left[(q_1 + q_2) + \frac{jkd}{2}(q_2 - q_1)\right] \tag{2.20}$$

The two terms in square brackets are the monopole and dipole components of the pressure due to the array and are plotted in Figure 2.10. The relative amplitudes of the two components are relatively constant up to around 3 kHz, above which the array is no longer small compared with the wavelength and the series expansion of equation (2.19) is no longer valid. However, at lower frequencies, where the monopole and dipole components combine to give the optimal directivity, it can be seen that due to the $jkd/2$ weighting on the dipole component, the source strength is around three times greater than that generating the monopole component. This is a result of the self-cancelling nature of the dipole source and can be directly related to the response of a standard hypercardioid source, which is,

$$\mathrm{Resp}(\theta) = 0.25 + 0.75\cos(\theta). \tag{2.21}$$

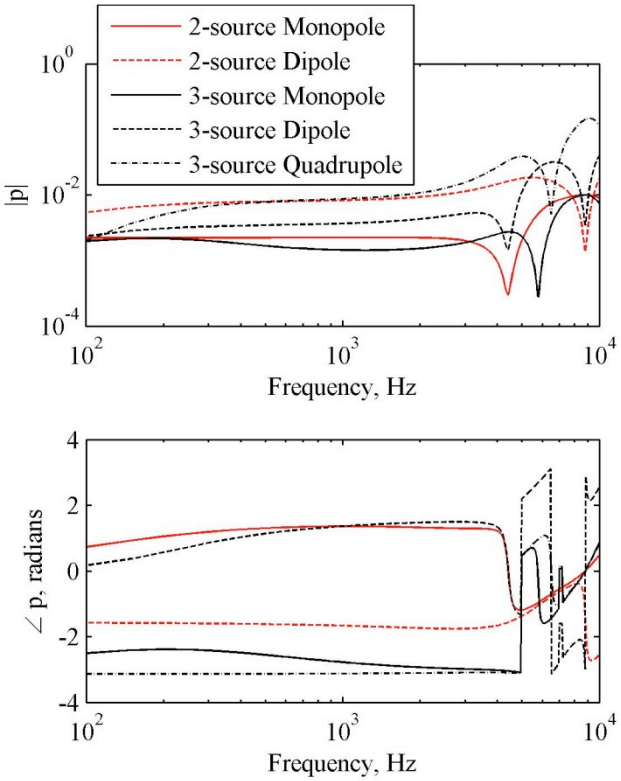


Figure 2.10 The magnitude and phase of the monopole and dipole component of the two-source optimised line array given by equation (2.20) and the monopole, dipole and quadrupole pressure components of the three-source optimised line array given by equation (2.23).

This situation is even more sensitive for the three source array, which achieves its high directivity from the interference between its monopole and dipole components and that of the longitudinal quadrupole component. Consider the pressure in line with three monopole sources, of complex source strengths q_1, q_2, and q_3 forming a line array. If the distance from the pressure measurement position to the centre source is r and the separation distance between each pair of source is d, then the complex pressure at frequency ω is equal to,

$$p(r) = \frac{j\omega\rho_0 e^{-jkr}}{4\pi r}\left[q_1 e^{-jkd} + q_2 + q_3 e^{jkd}\right] \tag{2.22}$$

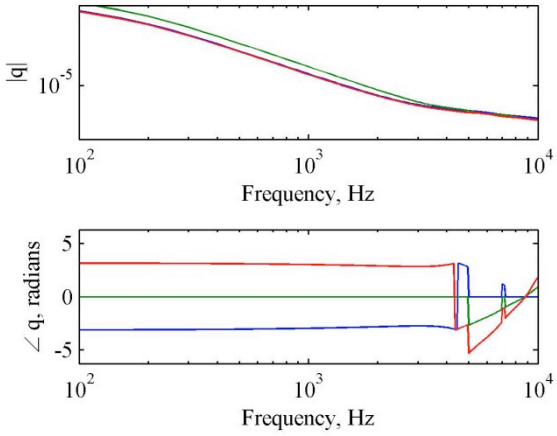
If the array is compact compared to the acoustic wavelength, so that *kd* is small compared with unity, then the pressure can be approximated by taking the two-term expansion of $e^{\pm jkd}$, to give

$$p(r) = \frac{j\omega\rho_0 e^{-jkr}}{4\pi r}\left[(q_1 + q_2 + q_3) + jkd(q_3 - q_1) - (kd)^2(q_1 + q_3)/2\right] \tag{2.23}$$

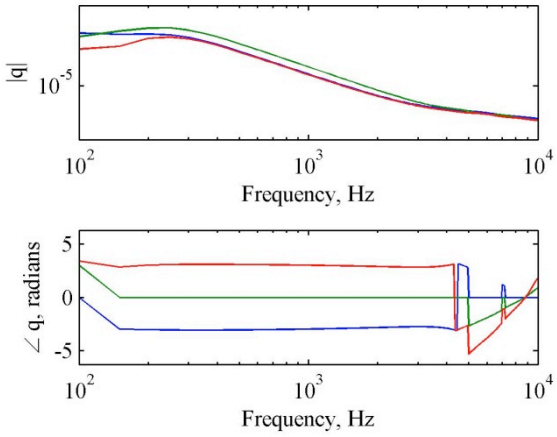
The three terms in the square brackets are the monopole, dipole and quadrupole components of the pressure due to the array (Pierce, 1981). The components of the pressure corresponding to the three terms in the square brackets of equation (2.23) have been calculated for the optimum three-source

array at a distance of $r = 0.5$ m and are shown in Figure 2.10. The relative amplitudes of the three components are fairly constant from about 300 Hz, below which the array is constrained by the limited effort, to about 3 kHz, above which the array is no longer small compared with the wavelength and the series expansion of equation (2.22) is no longer valid. Within this frequency range, however, these three multipole components of the pressure combine to give the directivity of the optimal array. Due to the jkd and $(kd)^2 / 2$ weightings on the source strengths of the dipole and quadrupole components, however, the source strengths generating these components in this frequency range are much larger than these generating the monopole component, since the sources are very self-cancelling. Very small errors in the implementation of the source strengths will then destroy this delicate balance and increase the monopole component well beyond that required for optimum performance.

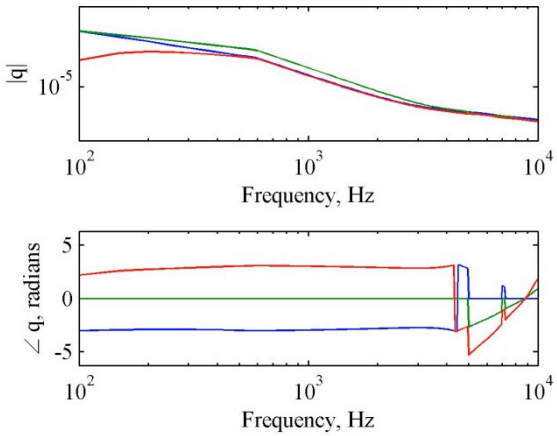
Figure 2.11 shows the variation in the magnitude and phase of the sources in the three source array, when calculated both without regularisation, with a constant value of β, and with the value of β as a function of frequency as shown in Figure 2.9. If this were to be implemented in a real-time system we would need to realise the impulse response of these filters, which are shown in Figure 2.12. From this plot it can be seen that the filters are non-causal and would in practice require a modelling delay.



(a) β=0



(b) β=1000



(c) Frequency independent β

Figure 2.11 Source strengths for three source array optimised at different frequencies for three different calues of the regularisation parameter using the indirect method. Front source (q_1) – blue line; centre source (q_2) – green line; rear source (q_3) – red line.

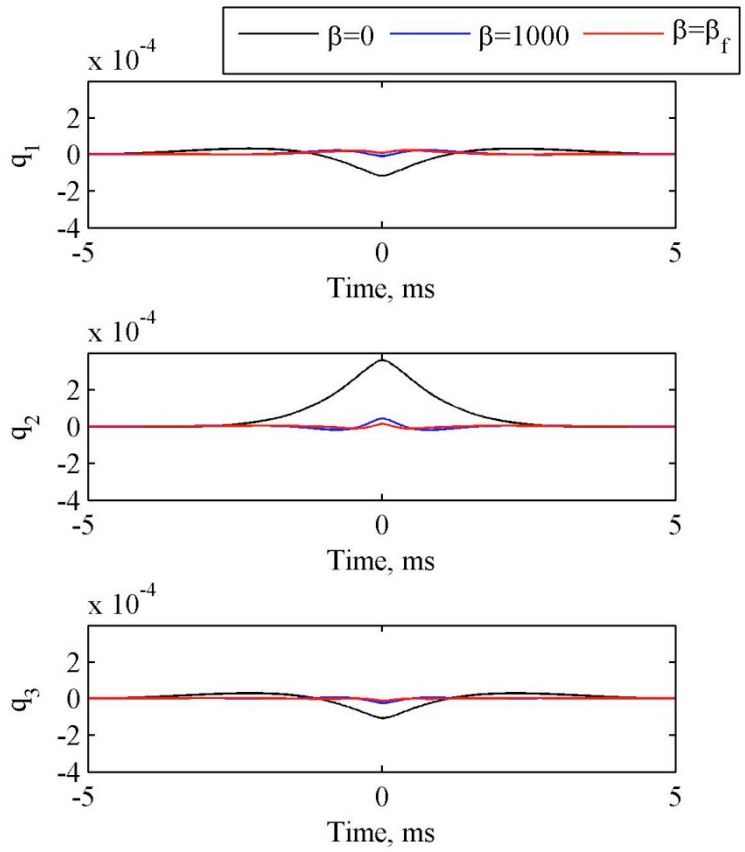


Figure 2.12 Impulse responses for the filters driving the centre and outer sources of the three source array when calculated one frequency at a time with no regularisation, with a constant value of β=1000 and with β as a function of frequency.

3. Multiple Constraints

Instead of maintaining a mean square level in a single bright zone, while minimising the mean square level in the dark zone and the array effort, it is also possible to generalise the constraints such that the mean square level is specified in a number of bright zones. This may be useful, for example if a particular level difference needs to be maintained between one region and another, while reducing the level in other regions to the greatest possible extent.

The constrained optimisation problem can be stated generally as one of minimising a function $f(\mathbf{q})$ subject to the multiple constraints that each element of the vector $\mathbf{g}(\mathbf{q})$ is equal to zero. The Lagrangian is then,

$$L(\mathbf{q},\boldsymbol{\lambda}) = f(\mathbf{q}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{q}), \tag{3.1}$$

where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers, which are real if each element of $\mathbf{g}(\mathbf{q})$ is a real-valued function, as they are here.

Specifically we assume that $f(\mathbf{q})$ is equal to the mean squared pressure in the dark zone, $\mathbf{q}^H \mathbf{Z}_\mathrm{D}^H \mathbf{Z}_\mathrm{D} \mathbf{q}$, but now $\mathbf{g}(\mathbf{q})$ is a vector with elements,

$$\mathbf{g}(\mathbf{q}) = [g_1(\mathbf{q}) + g_2(\mathbf{q}) \mathrm{K}\ g_N(\mathbf{q})], \tag{3.2}$$

where the constraints are either on multiple bright zones or on the array effort

$$\begin{aligned} g_1(\mathbf{q}) &= \mathbf{q}^H \mathbf{Z}_{B1}^H \mathbf{Z}_{B1} \mathbf{q} - b_1 \\ g_2(\mathbf{q}) &= \mathbf{q}^H \mathbf{Z}_{B2}^H \mathbf{Z}_{B2} \mathbf{q} - b_2 \\ &\mathrm{M} \\ g_N(\mathbf{q}) &= \mathbf{q}^H \mathbf{q} - e. \end{aligned} \tag{3.3}$$

If we express $\boldsymbol{\lambda}$ as $[\lambda_1, \lambda_2, \mathrm{K}\ \ \lambda_N]^T$, then half the differential of the Lagrangian with respect to $\mathbf{q}$ is equal to

$$\frac{1}{2}\frac{\partial \mathrm{L}}{\partial \mathbf{q}} = \mathbf{Z}_D^H \mathbf{Z}_D \mathbf{q} + \lambda_1 \mathbf{Z}_{B1}^H \mathbf{Z}_{B1} \mathbf{q} + \lambda_2 \mathbf{Z}_{B2}^H \mathbf{Z}_{B2} \mathbf{q} + \mathrm{K}\ \lambda_N \mathbf{q} \tag{3.4}$$

so that

$$\lambda_1 \mathbf{q} = -[\mathbf{Z}_{B1}^H \mathbf{Z}_{B1}]^{-1} [\mathbf{Z}_D^H \mathbf{Z}_D + \lambda_2 \mathbf{Z}_{B2}^H \mathbf{Z}_{B2} + \mathrm{K}\ \lambda_N \mathbf{I}] \mathbf{q} \tag{3.5}$$

A more systematic search now needs to be performed of the values of λ_2 to λ_N to both satisfy the constraints on b_1 to b_{N-1}, so that the relative mean square pressure levels in *B1* through to zone *B*(*N-1*) are maintained, and that the effort weighting is maintained.

To investigate the implementation of multiple bright zone constraints a personal audio problem has been defined with two adjacent bright zones overlapping on-axis and a single dark zone, as shown in Figure 3.1. In order to achieve a directivity that is capable of controlling the sound pressure level in the three defined zones, it is necessary to employ a second-order source array. This has been implemented here using a double row of 3 monopole sources, as shown in Figure 3.2, but could also be achieved with a single row of sources which each had more directional individual directivities. For this problem equation (3.5) becomes,

$$\lambda_1 \mathbf{q} = -[\mathbf{Z}_{B1}^H \mathbf{Z}_{B1}]^{-1}[\mathbf{Z}_D^H \mathbf{Z}_D + \lambda_2 \mathbf{Z}_{B2}^H \mathbf{Z}_{B2} + \lambda_3 \mathbf{I}]\mathbf{q} \tag{3.6}$$

The constraint on the array effort has, in the first instance, been dropped by setting $\lambda_3 = 0$, to allow a more straightforward investigation of the control strategy.

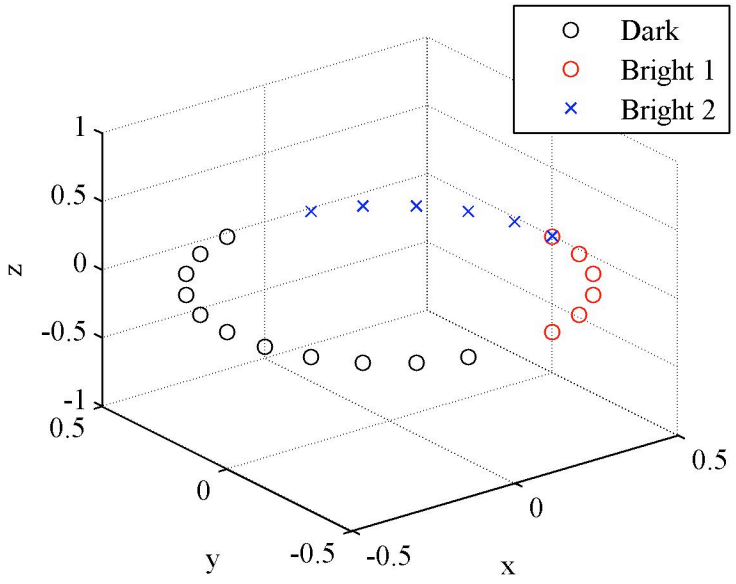


Figure 3.1 Sensor geometry with two bright zones.

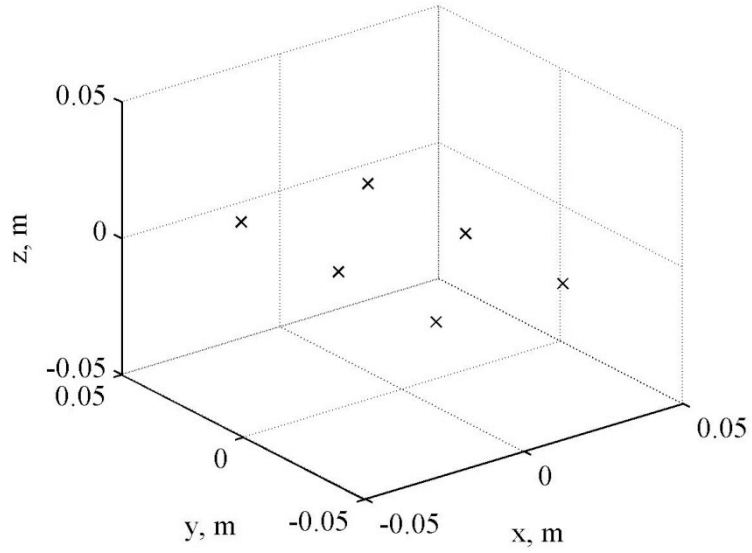


Figure 3.2 Source geometry.

Figure 3.3 shows the change in the mean squared pressure, plotted in decibels relative to the maximum mean squared pressure in bright zone 1, in the three zones defined in Figure 3.1 as λ_2, which controls the pressure in the second bright zone, is increased. From this plot it can be seen that for low values of λ_2 the pressure levels in the two bright zones are identical, whilst the level in the dark zone is 50 dB lower. As the value of λ_2 is increased the level in bright zone 2 is decreased while the level in the dark zone increases. At around $\lambda_2 = 0.25$ the mean square level in the dark zone and second bright zone are equal and beyond this point the second bright zone becomes the dominant dark zone. If the value of λ_2 is chosen to be about 7.94×10^{-2}, then the mean square pressure in bright zone 2 is 10 dB less than in bright zone 1, but the mean square pressure in the dark zone is still around 10 dB below this.

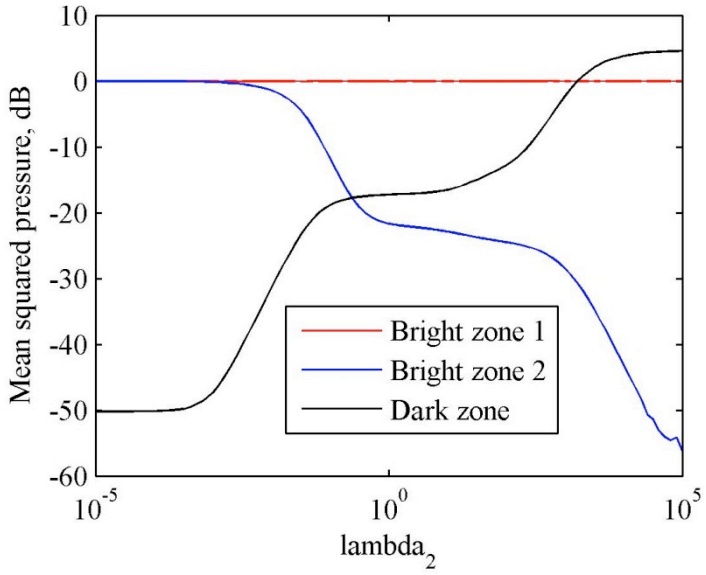


Figure 3.3 Variation in the mean squared pressure plotted in decibels relative to the maximum mean square pressure in bright zone 1 as λ_2 is increased.

Figure 3.4 shows the directivity of the optimised array for three values of λ_2 corresponding to the three regions shown in Figure 3.3. For a low value of λ_2 it can be seen that, since there is no constraint on the pressure in the second bright zone, the optimised array minimises the pressure in the dark zone which results in a broad main lobe. As λ_2 is increased a constraint is imposed that limits the pressure in bright zone 2. The blue line in Figure 3.4 shows the directivity at the point observed in Figure 3.3 where the mean squared pressure in the second bright zone is 10 dB below that in the first and 10 dB greater than that in the dark zone. At this point it can be seen that increasing the constraint on the second bright zone has caused an increase in the pressure in the dark zone. Increasing the constraint on the pressure in bright zone 2 further reduces the pressure in the second bright zone and increases the pressure in the dark zone, while the sum of the squared pressures in bright zone 1 remains constant.

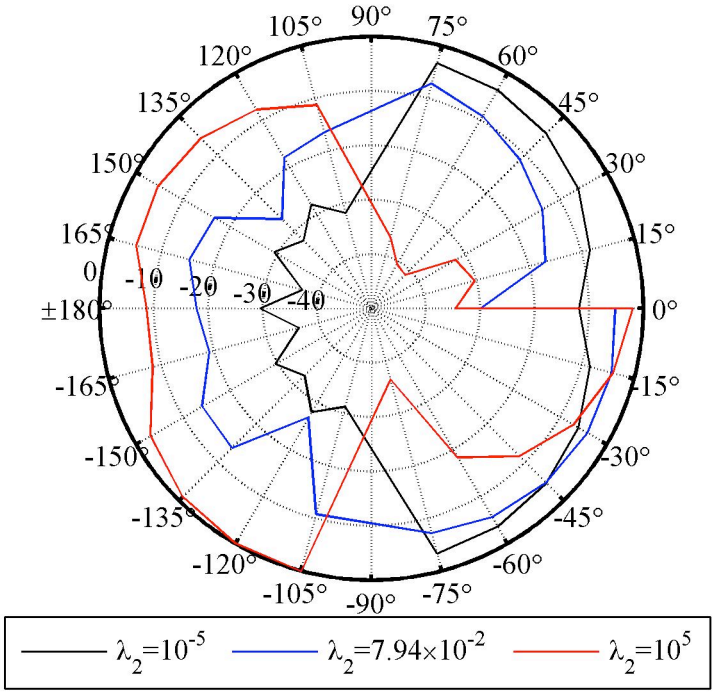


Figure 3.4 The directivity of the array for: (black) an unconstrained bright zone 2; (blue) bright zone 2 constrained to provide a mean squared pressure level 10 dB below that of bright zone 1; and (red) large constraint on the sum of squared pressures in bright zone 2, so that it becomes the dominant dark zone.

Figure 3.5 shows the trade-off between the acoustic contrast, evaluated for each bright zone individually, and the array effort. For a low value of λ_2 it can be seen that the acoustic contrast is equal for both bright zones, which corresponds to the results shown previously. As λ_2 is increased the acoustic contrast for both zones reduces which, for bright zone 1 is a result of the increase in pressure in the dark zone and for bright zone 2 is a result of both the decrease in the pressure in bright zone 2 and the increase in the dark zone pressure. There are a number of points of interest on the plot shown in Figure 3.5 for which the value of λ_2 has been indicated. The points marked with a red circle where λ_2=1580 corresponds to the point in Figure 3.3 where the mean squared pressure in the dark zone is equal to that in bright zone 1. The points marked with a black circle correspond to the point at which the mean squared pressure in the second bright zone is 10 dB below that in the first bright zone and 10 dB above that in the dark zone. The third point marked with red crosses correspond to the point on Figure 3.3 where the mean squared pressure in the dark zone is -23 dB and the mean squared pressure in bright zone 2 is -4.5 dB.

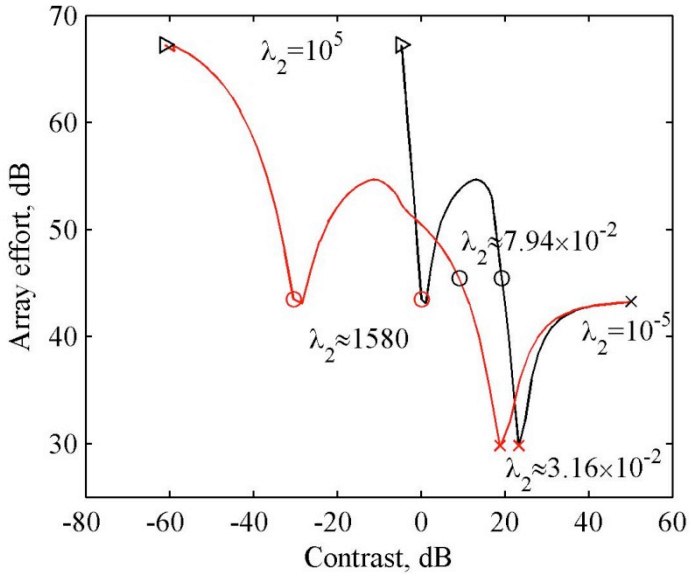


Figure 3.5 Trade-off between the array effort and the acoustic contrast achieved between bright zone 1 and the dark zone (black) and bright zone 2 and the dark zone (red) when using the acoustic contrast method with multiple constraints for the six-source 2-broadside array.

Figure 3.6 shows the acoustic contrast between the two bright zones plotted against array effort. The general trend of this plot shows that as λ_2 is increased the contrast between the two bright zones increases; however, once again the points of interest marked in Figure 3.5 have been indicated.

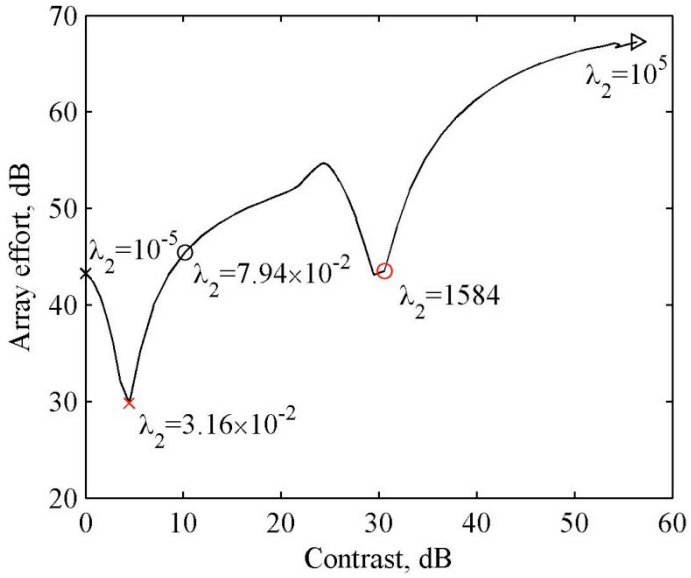


Figure 3.6 Trade-off between the array effort and the acoustic contrast achieved between bright zone 1 and bright zone 2 when using the acoustic contrast method with multiple constraints for the six-source 2-broadside array.

From the presented plots it is evident that for the optimisation process given by equation (3.6) a trade-off between the dark zone and the second bright zone is achieved. This allows the level in bright zone

1 to be maintained while the relative importance of minimising the sum of squared pressures in the dark zone and in bright zone 2 can be varied. If equation (3.6) were rearranged to give,

$$\mathbf{q} = -[\mathbf{Z}_D^H \mathbf{Z}_D]^{-1}[\lambda_1 \mathbf{Z}_{B1}^H \mathbf{Z}_{B1} + \lambda_2 \mathbf{Z}_{B2}^H \mathbf{Z}_{B2} + \lambda_3 \mathbf{I}]\mathbf{q} \tag{3.7}$$

it may be possible to facilitate a trade-off between the sum of squared pressures in the two bright zones by constraining the pressure in the dark zone whilst varying the relative values of λ_1 and λ_2 such that the desired pressures are achieved in each zone.

4. Time Domain Formulation

The acoustic contrast has been defined above at a single excitation frequency. In order to calculate the impulse response of the filters that would be required to implement a personal audio system in the time domain, the Fourier transform of a complete set of frequency responses would have to be taken. There is no guarantee that such a set of filters would be causal, however. Although the filters could be implemented with a delay to overcome non-causality, this may not be convenient in a real-time communication system and it may also lead to unwanted subjective effects, due to pre-ringing for example.

In this section, the maximisation of acoustic contrast is formulated as a time domain digital filtering problem. In this way we can directly calculate the coefficients of the digital filters required to drive each of the sources in the array for a real-time implementation. It is also possible to introduce regularisation directly into this time domain approach.

It is convenient to use the formulation developed for the calculation of the coefficients of the control filters in multichannel feedforward control systems, as described, for example, in Section 5.2.1 of Elliott (2001). The sampled output of the l-th microphone in the bright zone can be written as

$$y_{\mathrm{B}l}(n) = \sum_{k=1}^{K} \sum_{j=0}^{J-1} g_{\mathrm{B}lk}(j) u_k(n-j), \tag{4.1}$$

where $u_k(n)$ is the sampled input signal to the k-th driver and $g_{\mathrm{B}lk}(j)$ is the j-th coefficient of the impulse response for the k-th driver to the l-th microphone in the birght zone, which is modelled as an FIR filter of length J.

If the input to the k-th driver is generated by passing a signal $x(n)$ through an FIR filter with coefficients $w_k(i)$, then

$$u_k(n) = \sum_{i=0}^{I-1} w_k(i) x(n-i), \tag{4.2}$$

where all the filters are again assumed to be of the same length, for convenience.

The output of the l-th microphone in the bright zone, which is

$$y_{\mathrm{B}l}(n) = \sum_{k=1}^{K} \sum_{j=0}^{J-1} \sum_{i=0}^{I-1} g_{\mathrm{B}lk}(j) w_k(i) x(n-i-j), \tag{4.3}$$

may also be written as

$$y_{\mathrm{B}l}(n) = \sum_{k=1}^{K} \sum_{i=0}^{I-1} \mathrm{w}_k(i)\mathrm{r}_{\mathrm{B}lk}(n-i), \tag{4.4}$$

where the filtered reference signals for the bright zone are defined as

$$\mathrm{r}_{\mathrm{B}lk}(n) = \sum_{j=0}^{J-1} g_{\mathrm{B}lk}(j)x(n-j). \tag{4.5}$$

The summation over k in the expression for $y_{\mathrm{B}l}(n)$ can also be represented in vector form, to give

$$y_{\boldsymbol{B}l}(n) = \sum_{i=0}^{I-1} \boldsymbol{w}_i^T \boldsymbol{r}_{\boldsymbol{B}l}(n-i), \tag{4.6}$$

Where

$$\mathbf{w}_i = [\, w_{1i} \quad w_{2i} \quad \mathrm{L} \quad w_{Ki} \,]^T \tag{4.7}$$

and

$$\boldsymbol{r}_{\mathrm{B}l}(n) = [r_{\mathrm{B}l1}(n) \quad r_{\mathrm{B}l2}(n) \;\ldots\; r_{\mathrm{B}lK}(n)]^T. \tag{4.8}$$

The vector of all the sampled microphone outputs in the bright zone, which is

$$\boldsymbol{y}_{\mathrm{B}}(n) = [y_{\mathrm{B}1}(n) \quad y_{\mathrm{B}2}(n) \;\ldots\; y_{\mathrm{BL}}(n)]^T, \tag{4.9}$$

can now be written as

$$\boldsymbol{y}_{\mathrm{B}}(n) = \boldsymbol{R}_{\mathrm{B}}(n)\boldsymbol{w}, \tag{4.10}$$

where the matrix of filtered reference signals in the bright zone is

$$\boldsymbol{R}_{\mathrm{B}}(n) = \begin{bmatrix} \boldsymbol{r}_{\mathrm{B}1}^T(n) & \boldsymbol{r}_{\mathrm{B}1}^T(n-1) & \ldots & \boldsymbol{r}_{\mathrm{B}1}^T(n-I+1) \\ \boldsymbol{r}_{\mathrm{B}2}^T(n) & \boldsymbol{r}_{\mathrm{B}2}^T(n-1) & & \mathrm{M} \\ \mathrm{M} & & \mathrm{O} & \\ \boldsymbol{r}_{\mathrm{B}L}^T(n) & \boldsymbol{r}_{\mathrm{B}L}^T(n-1) & \ldots & \boldsymbol{r}_{\mathrm{B}L}^T(n-I+1) \end{bmatrix} \tag{4.11}$$

and the overall vector of filter coefficient is

$$\boldsymbol{w} = [\boldsymbol{w}_0^T \;\; \boldsymbol{w}_1^T \;\ldots\; \boldsymbol{w}_{I-1}^T]^T. \tag{4.12}$$

The vector of sampled outputs to the M microphones in the dark zone can be similarly written as

$$\mathbf{y}_D(n) = \mathbf{R}_D(n)\mathbf{w} \tag{4.13}$$

where $\boldsymbol{R}_{\mathrm{D}}(n)$ is the matrix of filtered reference signals in the dark zone and $\boldsymbol{w}$ is the same vector of filter coefficients as above.

If the input signal to the array, $x(n)$ is now treated as a stationary random variable, then the time-averaged acoustic contrast may be defined as being

$$C = \frac{E(\boldsymbol{y}_{\mathrm{B}}^T \boldsymbol{y}_{\mathrm{B}})}{E(\boldsymbol{y}_{\mathrm{D}}^T \boldsymbol{y}_{\mathrm{D}})} \tag{4.14}$$

where $E(\ )$ denotes the expectation operator, in the generalised sense described in Section 2.3.1 of Elliott (2001) for example. Since the filter coefficients are time invariant, this can also be written as

$$C = \frac{\boldsymbol{w}^T E(\boldsymbol{R}_{\mathrm{B}}^T(n)\boldsymbol{R}_{\mathrm{B}}(n))\boldsymbol{w}}{\boldsymbol{w}^T E(\boldsymbol{R}_{\mathrm{D}}^T(n)\boldsymbol{R}_{\mathrm{D}}(n))\boldsymbol{w}}. \tag{4.15}$$

The solution for a set of filter coefficients that maximises the time-averaged contrast can also be regularised by including a constraint on the sum of the squared value of all the filter coefficients, $\boldsymbol{w}^T\boldsymbol{w}$. More generally, the sum of the mean square signals driving the sources can be written as $\boldsymbol{w}^T\boldsymbol{R}_{xx}\boldsymbol{w}$ where $\boldsymbol{R}_{xx}$ is a matrix whose elements depend on the autocorrelation function of the input signal. The autocorrelation matrix is proportional to the unit matrix if the driving signal is white, however, and so the constraint reduces to that on $\boldsymbol{w}^T\boldsymbol{w}$. It is also possible to penalise the control effort in different frequency ranges by replacing $\boldsymbol{R}_{xx}$ with the corresponding autocorrelation matrix for a filtered reference signal, in which the filter is designed to emphasise the frequency range requiring most regularisation.

The regularised optimisation of the time-averaged contrast can now be cast as the minimisation of $E(\boldsymbol{y}_{\mathrm{D}}^T \boldsymbol{y}_{\mathrm{D}})$, subject to the constraint that $E(\boldsymbol{y}_{\mathrm{B}}^T \boldsymbol{y}_{\mathrm{B}})$ is equal to b and that $\boldsymbol{w}^T\boldsymbol{R}_{xx}\boldsymbol{w}$ is equal to e. The Lagrangian for this optimisation is thus

$$L = \boldsymbol{w}^T E(\boldsymbol{R}_{\mathrm{D}}^T(n)\boldsymbol{R}_{\mathrm{D}}(n))\boldsymbol{w} + \lambda_1(\boldsymbol{w}^T E(\boldsymbol{R}_{\mathrm{B}}^T(n)\boldsymbol{R}_{\mathrm{B}}(n))\boldsymbol{w} - b) + \lambda_2(\boldsymbol{w}^T \boldsymbol{R}_{xx}\boldsymbol{w} - e), \tag{4.16}$$

for which half the differential with respect to $\boldsymbol{w}$ is

$$\frac{1}{2}\frac{\partial L}{\partial \boldsymbol{w}} = E(\boldsymbol{R}_{\mathrm{D}}^T(n)\boldsymbol{R}_{\mathrm{D}}(n))\boldsymbol{w} + \lambda_1 E(\boldsymbol{R}_{\mathrm{B}}^T(n)\boldsymbol{R}_{\mathrm{B}}(n))\boldsymbol{w} + \lambda_2 \boldsymbol{R}_{xx}\boldsymbol{w}. \tag{4.17}$$

This differential is zero if

$$\lambda_1 \boldsymbol{w} = -[E(\boldsymbol{R}_{\mathrm{B}}^T(n)\boldsymbol{R}_{\mathrm{B}}(n))]^{-1}[E(\boldsymbol{R}_{\mathrm{D}}^T(n)\boldsymbol{R}_{\mathrm{D}}(n)) + \lambda_2 \boldsymbol{R}_{xx}]\boldsymbol{w}, \tag{4.18}$$

so that the optimum set of filter coefficients is proportional to the eigenvector corresponding to the smallest eigenvalue of the matrix above, which is the same as the largest eigenvalue of the matrix $[E(\boldsymbol{R}_{\mathrm{D}}^T(n)\boldsymbol{R}_{\mathrm{D}}(n)) + \lambda_2 \boldsymbol{R}_{xx}]^{-1} E(\boldsymbol{R}_{\mathrm{B}}^T(n)\boldsymbol{R}_{\mathrm{B}}(n))$.

5. Robustness to Uncertainty in Transfer Responses

In this section we consider how the performance of a personal audio system can be made robust to independent uncertainties in each of the transfer responses between the drivers and microphone positions. This uncertainty can arise in a number of ways, but would certainly be generated by movement of the system within a reverberant space, or the movement of other people within such a space. Each transfer response can then be considered to be the sum of a contribution from the direct acoustic field of the source and a contribution from this uncertain reverberant component. The statistical properties of the reverberant field are discussed by Schroeder (1954) and Ebeling (1984) and are summarised in Chapter 11 of Nelson and Elliott (1992). The diffuse field is made up of a large number of contributions from waves travelling in different directions. Its properties can only be usefully described in terms of their average value over frequency, or over many different points in space at a single frequency. The space-average modulus of the transfer response between a source of volume velocity q and a well separated microphone measuring a pressure p is given, for example, by Nelson and Elliott (1992) as

$$\langle |Z_r| \rangle = \left(\frac{\omega \rho_0{}^2 c_0}{8\pi V \xi} \right)^{1/2}, \tag{5.1}$$

where < > denotes space averaging, Z_r is the transfer impedance p/q, ω is the angular frequency, ρ_0 and c_0 are the density and speed of sound in the air, V is the volume of the reverberant space and ξ is the average damping ratio of the acoustic modes. The phase of Z_r is a random function of the position. This reverberant transfer response can be compared with the deterministic expression for that due to the direct field, which is equal to that in the free field and is given by

$$Z_d = \frac{j\omega\rho_0 e^{-jkr}}{4\pi r}, \tag{5.2}$$

where r is the distance from the source to the microphone.

A useful measure of the extent of the direct field around a source is given by the reverberation radius, defined to be the distance at which $\langle |Z_r| \rangle$ is equal to $|Z_d|$ and given by

$$R = \left(\frac{\omega V \xi}{2\pi c_0} \right)^{1/2}. \tag{5.3}$$

For a reverberant space having a volume of 100 m^3 with an average modal damping ratio of 0.1 the reverberant radius at 1 kHz is thus about 5 m.

To illustrate the approach used in this section, we first consider the robust solution to a conventional, single channel, filtering problem before discussing the robust solution to maximising acoustic contrast in a personal audio system. Consider the frequency-domain filtering problem shown in Figure 5.1, in which a tonal reference signal, *x*, is assumed to have unit magnitude, and the objective is to minimise the modulus squared error signal, *e*. This is the difference between the output of the robust filter we wish to design, *W*, having passed through the transfer response $\tilde{G}$ and the output of a filter $\tilde{P}$. The tildes above the responses $\tilde{G}$ and $\tilde{P}$ denote that these are uncertain quantities. The space averaged values of these quantities are equal to G and P, and their uncertain values can be written as

$$\widetilde{G} = G + \Delta G \tag{5.4}$$

$$\widetilde{P} = P + \Delta P \tag{5.5}$$

where ΔG and ΔP are the random uncertainties whose space average values are zero, but whose space average modulus squared values are denoted

$$<|\Delta G|^2>= \Delta_G \tag{5.6}$$

$$<|\Delta P|^2>= \Delta_P \tag{5.7}$$

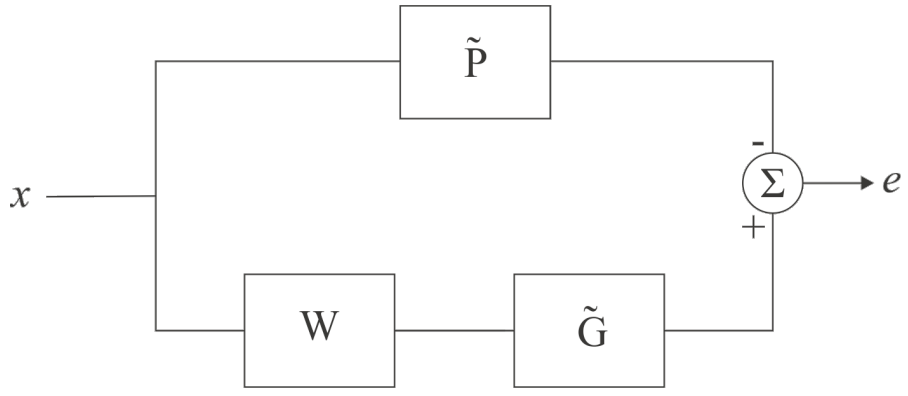


Figure 5.1 Arrangement used to illustrate a robust filtering problem in which the filter W is designed to minimise the modulus squared error, *e*, averaged over the uncertainty in the filters $\tilde{P}$ and $\tilde{G}$.

It is also assumed that the variations in G and P are independent. The complex error signal is thus

$$e = \widetilde{G}W - \widetilde{P} \tag{5.8}$$

And its space-average modulus square value is

$$<|e|^2>=<|W|^2|\widetilde{G}|^2 - W^*\widetilde{G}^*\widetilde{P} - \widetilde{P}^*\widetilde{G}W + |\widetilde{P}|^2>, \tag{5.9}$$

where the superscript * denotes conjugation. Using the properties of the uncertainties defined above, then this can be expressed as

$$<|e|^2>=<|W|^2|G|^2 + \Delta_G - W^*G^*P - P^*GW + |P|^2 + \Delta_P>, \tag{5.10}$$

The space-averaged modulus squared error is minimised if its derivative with respect to the real and imaginary parts of W are set to zero, which yields the optimum filter given by

$$W_{\text{opt}} = \frac{\mathrm{G}^*\mathrm{P}}{|\mathrm{G}|^2 + \Delta_{\mathrm{G}}}. \tag{5.11}$$

Notice that the uncertainty in P does not affect the optimum filter, but that the uncertainty in G reduces the magnitude of W, since otherwise this uncertainty would increase the space-average mean square error.

We now assume for the personal audio problem that the matrices of transfer responses from the source inputs to the pressure outputs in the bright and dark zones are also subject to uncertainties and can be written as

$$\tilde{\mathbf{Z}}_{\mathrm{B}} = \mathbf{Z}_{\mathrm{B}} + \Delta\mathbf{Z}_{\mathrm{B}}, \qquad \tilde{\mathbf{Z}}_{\mathrm{D}} = \mathbf{Z}_{\mathrm{D}} + \Delta\mathbf{Z}_{\mathrm{D}}. \tag{5.12}$$

The space average values of the uncertainties $\Delta\mathbf{Z}_{\mathrm{B}}$ and $\Delta\mathbf{Z}_{\mathrm{D}}$ are zero, so that $\mathbf{Z}_{\mathrm{B}}$ and $\mathbf{Z}_{\mathrm{D}}$ are the space-averaged mean values of the transfer response matrices. We also define the space average properties of the uncertainty matrices to be

$$<\Delta\mathbf{Z}_{\mathrm{B}}^H\ \Delta\mathbf{Z}_{\mathrm{B}}>=\mathbf{\Delta}_{\mathrm{B}}, \quad <\Delta\mathbf{Z}_{\mathrm{D}}^H\ \Delta\mathbf{Z}_{\mathrm{D}}>=\mathbf{\Delta}_{\mathrm{D}}, \tag{5.13}$$

and additionally assume that the variations in $\mathbf{Z}_{\mathrm{B}}$ and $\mathbf{Z}_{\mathrm{D}}$ are not correlated so that

$$<\Delta\mathbf{Z}_{\mathrm{B}}^H\ \Delta\mathbf{Z}_{\mathrm{D}}>=0. \tag{5.14}$$

This final assumption may be of limited validity if the bright and dark zones are close together compared with a wavelength and the uncertainty is due to the diffuse field. It does, however, greatly simplify the formulation of the robust contrast optimisation problem, however, since we can then write the space-averaged contrast as

$$C = \left\langle \frac{\mathbf{p}_{\mathrm{B}}^{\mathrm{H}}\mathbf{p}_{\mathrm{B}}}{\mathbf{p}_{\mathrm{D}}^{\mathrm{H}}\mathbf{p}_{\mathrm{D}}} \right\rangle = \frac{<\mathbf{p}_{\mathrm{B}}^{\mathrm{H}}\mathbf{p}_{\mathrm{B}}>}{<\mathbf{p}_{\mathrm{D}}^{\mathrm{H}}\mathbf{p}_{\mathrm{D}}>}. \tag{5.15}$$

In this case

$$<\mathbf{p}_{\mathrm{B}}^{\mathrm{H}}\mathbf{p}_{\mathrm{B}}> = <\mathbf{q}^{H}\tilde{\mathbf{Z}}_{\mathrm{B}}^{\mathrm{H}}\tilde{\mathbf{Z}}_{\mathrm{B}}\mathbf{q}> = \mathbf{q}^{H}(\mathbf{Z}_{\mathrm{B}}^{\mathrm{H}}\mathbf{Z}_{\mathrm{B}} + \Delta_{\mathrm{B}})\mathbf{q}, \tag{5.16}$$

and

$$<\mathbf{p}_{\mathrm{D}}^{\mathrm{H}}\mathbf{p}_{\mathrm{D}}> = <\mathbf{q}^{H}\tilde{\mathbf{Z}}_{\mathrm{D}}^{\mathrm{H}}\tilde{\mathbf{Z}}_{\mathrm{D}}\mathbf{q}> = \mathbf{q}^{H}(\mathbf{Z}_{\mathrm{D}}^{\mathrm{H}}\mathbf{Z}_{\mathrm{D}} + \Delta_{\mathrm{D}})\mathbf{q}. \tag{5.17}$$

In order to formulate the regularised maximisation of the space averaged contrast, we can thus minimise $<\mathbf{p}_{\mathrm{D}}^{\mathrm{H}}\mathbf{p}_{\mathrm{D}}>$ with the constraint that $<\mathbf{p}_{\mathrm{B}}^{\mathrm{H}}\mathbf{p}_{\mathrm{B}}>$ is equal to c and $\mathbf{q}^{H}\mathbf{q}$ is equal to e, so that the Lagrangian is

$$L = \mathbf{q}^{H}(\mathbf{Z}_{\mathrm{D}}^{\mathrm{H}}\mathbf{Z}_{\mathrm{D}} + \mathbf{\Delta}_{\mathrm{D}})\mathbf{q} + \lambda_1[\mathbf{q}^{H}(\mathbf{Z}_{\mathrm{B}}^{\mathrm{H}}\mathbf{Z}_{\mathrm{B}} + \mathbf{\Delta}_{\mathrm{B}})\mathbf{q} - c] + \lambda_2[\mathbf{q}^{H}\mathbf{q} - e]. \tag{5.18}$$

Half the differential of this Lagrangian with respect to the real and imaginary parts of $\mathbf{q}$ is thus

$$\frac{1}{2}\frac{\partial L}{\partial \mathbf{q}} = (\mathbf{Z}_{\mathrm{D}}^{H}\mathbf{Z}_{\mathrm{D}} + \mathbf{\Delta}_{\mathrm{D}})\mathbf{q} + \lambda_1(\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} + \mathbf{\Delta}_{\mathrm{B}})\mathbf{q} + \lambda_2\mathbf{q} \tag{5.19}$$

This differential will be zero if

$$\lambda_1\mathbf{q} = -(\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} + \mathbf{\Delta}_{\mathrm{B}})^{-1}(\mathbf{Z}_{D}^{H}\mathbf{Z}_{\mathrm{D}} + \mathbf{\Delta}_{\mathrm{D}} + \lambda_2\mathbf{I})\mathbf{q}, \tag{5.20}$$

and so the robustly optimum set of source strengths is proportional to the eigenvector corresponding to the smallest eigenvalue of the matrix in square brackets above, which is equal to the largest eigenvector of the matrix $(\mathbf{Z}_{D}^{H}\mathbf{Z}_{\mathrm{D}} + \mathbf{\Delta}_{\mathrm{D}} + \lambda_2\mathbf{I})^{-1}(\mathbf{Z}_{\mathrm{B}}^{H}\mathbf{Z}_{\mathrm{B}} + \mathbf{\Delta}_{\mathrm{B}})$.

If the variation in each of the elements of $\mathbf{Z}_{\mathrm{D}}$ and of $\mathbf{Z}_{\mathrm{B}}$ was independent, then $\mathbf{\Delta}_{\mathrm{D}}$ and $\mathbf{\Delta}_{\mathrm{B}}$ would be proportional to identity matrices and the solution to the robust contrast maximisation problem would be similar to a regularised form of the normal contrast maximisation problem. To demonstrate this the matrices of mean square uncertainties, $\mathbf{\Delta}_{\mathrm{D}}$ and $\mathbf{\Delta}_{\mathrm{B}}$, have been calculated as,

$$\mathbf{\Delta}_{B} = e^2\frac{\|\mathbf{Z}_{B}^{H}\mathbf{Z}_{B}\|_{F}}{K^2}\mathbf{I} \tag{5.21}$$

$$\mathbf{\Lambda}_D = e^2 \frac{\| \mathbf{Z}_D^H \mathbf{Z}_D \|_F}{K^2} \mathbf{I} \tag{5.22}$$

where $\| \ldots \|_F$ is the Frobenius norm (Golub and Van Loan, 1996) and e is the rms error. The performance of the array has been calculated for values of e of 0.05, 0.1, and 0.2 and Figure 5.2shows the change in the acoustic contrast and array effort as a result of these random variations in the acoustic transfer responses calculated according to equation (5.20). From the acoustic contrast plot it can be seen the variations in the transfer responses result in a reduction in the acoustic contrast performance achieved at low frequencies. As detailed above this is a similar effect to that observed through the implementation of regularisation.

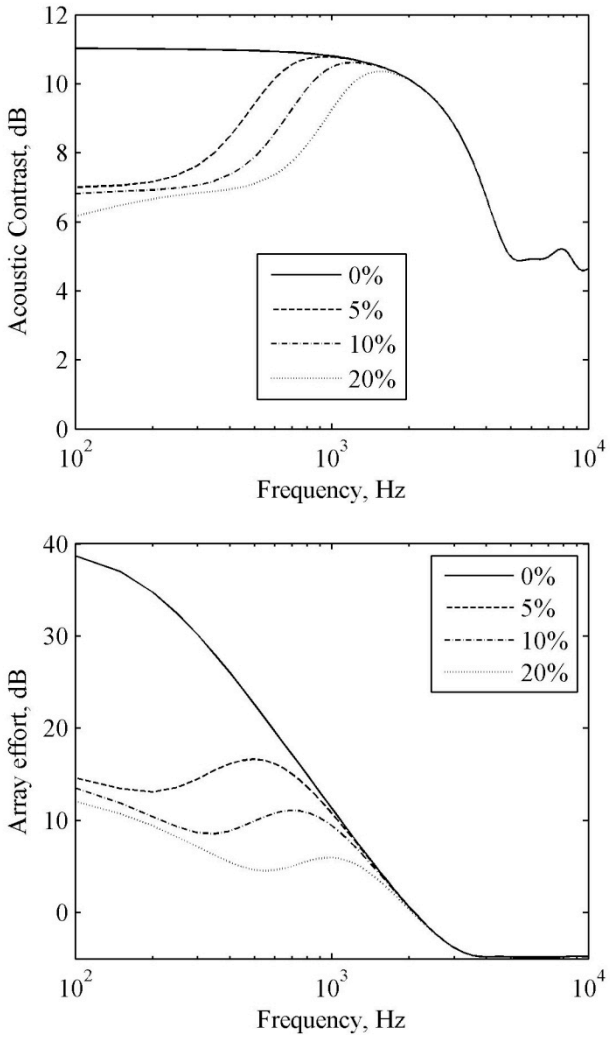


Figure 5.2 The maximised acoustic contrast and the corresponding array effort for a three source line array when there is no variation in the acoustic transfer responses, solid lines, and the acoustic contrast and array effort when there is a random variation in the acoustic transfer responses of 5% (dashed lines), 10% (dot-dash lines), and 20% (dotted lines).

To determine the relative effects of the uncertainties in the bright and dark zone transfer responses, $\mathbf{\Delta}_D$ and $\mathbf{\Delta}_B$, Figure 5.2 shows the change in the acoustic contrast and array effort as a result of the random variations given by equations (5.21) and (5.22) in the bright and dark zones (as inFigure 5.1) in the bright zone alone and in the dark zone alone. From these results it can be seen that the effect of $\mathbf{\Delta}_B$ on the solution is much less significant than that of $\mathbf{\Delta}_D$ in this case. It is also interesting to note that while the uncertainty in both the bright and dark zones and the dark zone alone cause a reduction in the performance and a corresponding reduction in the array effort, when the uncertainty is in the bright zone alone then despite the acoustic contrast being reduced the array effort is increased.

A practical situation in which the transfer responses employed in the design of an array's filters may differ from those of the operational environment is when the array is positioned in a different room. For example, an array may be optimised based on acoustic transfer responses measured in an anechoic chamber while the array is likely to be used in a reverberant environment. The acoustic response between a source and receiver in a reverberant environment is given by the summation of the direct and reverberant fields. Using equations (5.1) and (5.2) the average response in a reverberant enclosure at a specific distance from a source can be approximated as,

$$Z = Z_d + \langle |Z_r| \rangle = \frac{j\omega\rho_0 c_0 e^{-jkr}}{4\pi r} + \left(\frac{\omega {\rho_0}^2 c_0}{8\pi V \xi} \right)^{1/2} \tag{5.23}$$

Figure 5.3 shows the change in the acoustic contrast performance when the two-source line array that has been optimised in a free-field environment is positioned in two common environments:

1. A listening room with acoustic characteristics comparable to a lounge with a volume of 61 m^3 and acoustic damping of ζ=5%.
2. A small car cabin with a volume of 3.2 m^3 and acoustic damping of ζ=10%.

From Figure 5.3 it can be seen that at low frequencies both reverberant enclosures reduce the acoustic contrast performance in a manner similar to regularisation. At higher frequencies, where the free-field contrast performance begins to roll-off, the contrast performance in the enclosures is close to that the free-field. The reductions in the acoustic contrast at low frequencies are of greater magnitude in the smaller, car cabin type enclosure. This is due to the reduced volume compared to the listening room and although the increased acoustic damping reduces the effect of the enclosure on the performance the relative change in damping is much smaller than that in volume.

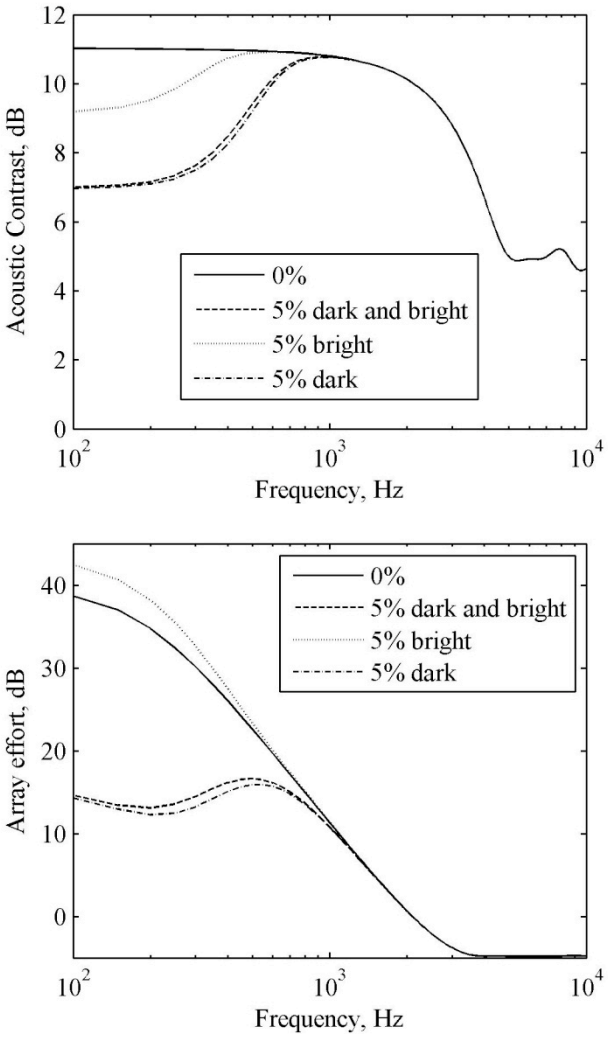


Figure 5.3 The maximised acoustic contrast and the corresponding array effort for a three source line array when there is no variation in the acoustic transfer responses, solid lines, and the acoustic contrast and array effort when there is a random variation in the acoustic transfer responses of 5% in the bright and dark zones (dashed lines), the bright zone (dotted lines), and dark zone (dot-dashed lines).

6. Robustness to Uncertainty in the Driver Position or Driver Response

It was assumed in the previous section that the variability in the transfer responses from the source array to the microphone positions in the bright and in the dark zones was independent. If the uncertainty is due to variability in the positions of the drivers, or their response at a given frequency, this assumption is not valid. It also turns out that linear arrays with more than two elements are extremely sensitive to mis-alignment in the source positions because of the way that they can then radiate sound.

Figure 6.1, for example, shows the maximised contrast and corresponding control effort when the centre source of a three-source array is offset by 5% of the separation between the sources. At low frequencies, when the separation is small compared with the wavelength, this offset significantly degrades the achievable contrast.

Consider maximising the acoustic contrast for a three source array in the geometry above, for example, when the spacing is small compared with a wavelength. This requires that the source strengths generate two, approximately equal, components of their pressure field. The first is due to the omnidirectional monopole component, which has a high radiation efficiency. The second is due to the $\cos^2\theta$ directivity of a longitudinal quadrupole, which has a low radiation efficiency. The analysis below shows that if one of the sources in the linear array is slightly offset, this generates an uncontrolled dipole component to the radiated sound, causing the highly-cancelled three source array to "spill out" sound power in another direction.

In order to quantify this observation we consider the related problem of minimising the power output of the three source array shown in Figure 6.2 by adjusting the source strengths of the two outer sources, q_1 and q_2, when the source strength of the centre source is fixed.

Following Nelson and Elliott (1992), we write the vector of source strengths to be optimised as

$$\mathbf{q} = [q_1 \quad q_2]^T \tag{6.1}$$

so that the power output can be written as

$$W = \mathbf{q}^H \mathbf{A}\mathbf{q} + \mathbf{q}^H \mathbf{b} + \mathbf{b}^H \mathbf{q} + W_2 \tag{6.2}$$

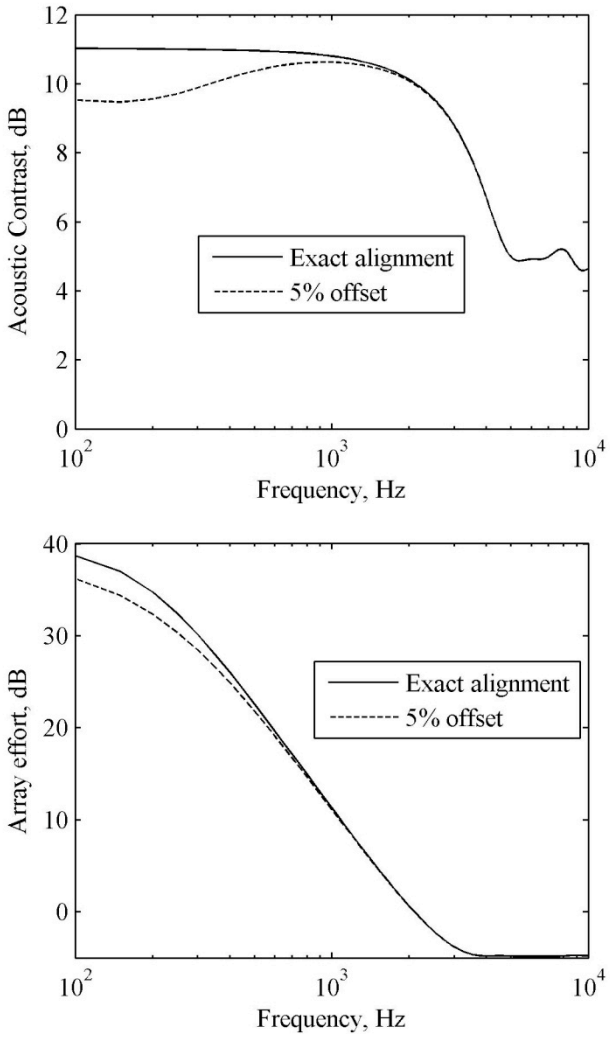


Figure 6.1 The maximised acoustic contrast and the corresponding array effort for a three source linear array when exactly aligned, solid lines, and when the centre source is offset by 5% of the separation distance (dashed line).

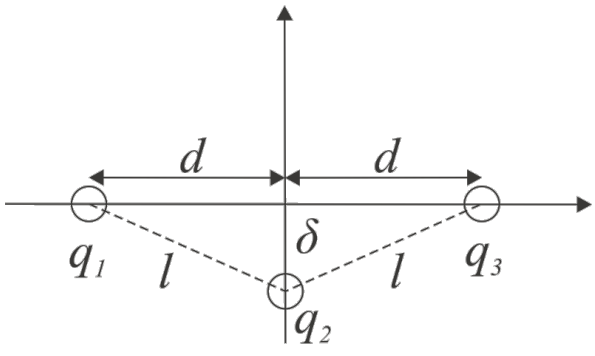


Figure 6.2 The geometry of a three source array having source strengths q_1, q_2 and q_3 with a separation distance $2d$ between q_1 and q_3 and q_2 offset by a distance δ.

where

$$\mathbf{A} = \frac{Z_0}{2}\begin{bmatrix} 1 & \text{sinc } 2kd \\ \text{sinc } 2kd & 1 \end{bmatrix} \tag{6.3}$$

$$\mathbf{b} = \frac{Z_0}{2} q_2 \begin{bmatrix} \text{sinc } kl \\ \text{sinc } kl \end{bmatrix} \tag{6.4}$$

$$W_2 = \frac{Z_0}{2} |\, q_2 \,|^2 \tag{6.5}$$

Z_0 is equal to $\omega_0 c_0 / 4\pi c_0$ and l is the distance between source 2 and sources 1 or 3, given by $\sqrt{d^2 + \delta^2}$.

The minimised power output of this array, $W_{\min}$, is given generally by $W_2 - \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b}$, which, when normalised by W_2, is given in this case by

$$\frac{W_{\min}}{W_2} = 1 - \frac{2\text{sinc}^2 kl}{1 + \text{sinc}\, 2kd} \tag{6.6}$$

At low frequencies, where the separation is small compared with the wavelength so that kd is small compared with unity, the three-term series expansion for sinc kl and sinc $2kd$ can be used to express the normalised power output as

$$\frac{W_{\min}}{W_2} = \frac{(k\delta)^2}{3} + \frac{(kd)^4}{45} \tag{6.7}$$

If the array was exactly aligned, so that δ was equal to zero, the source thus radiates with a quadrupole efficiency, proportional to $(kd)^4$. It is this component, combined with the small monopole component that can generate the high values of acoustic contrast seen in Figure 6.1 The maximised acoustic contrast and the corresponding array effort for a three source linear array when exactly aligned, solid lines, and when the centre source is offset by 5% of the separation distance (dashed line).. When the central source is offset by a non-zero distance, δ, however, there is a component to the power output proportional to $(k\delta)^2$, which is due to a dipole component radiating at right angles to the axis of the array.

If the separation between the sources is 4 cm, for example, and the array is operating at 100 Hz, the offset on the centre source, δ, required to make the dipole component of the power output in equation (6.7) equal to the quadrupole component is only about 0.76 mm, which is less than 2 % of the source

separation. Offset distances of this magnitude would easily be generated by diffraction of a practical source from nearby objects, and so it would appear to be very difficult to achieve such high levels of contrast at these very low frequencies. At higher frequencies, however, reducing the magnitude of the source strengths by regularisation will help prevent such extreme self-cancelling conditions occurring.

We now consider the effect on performance due to variability in the response of the drivers. Variation in driver response would be generated by the spread inherent in the driver manufacturing process. In principle this could be accounted for by taking individual measurements of the transfer responses from each driver to each monitoring point in the bright and the dark zone for every manufactured personal audio system, and using these to calculate individualised filters. This may even be achieved using the microphones in the mobile device itself if measured in a reproducible environment. In practice this would be an expensive process, however, and it would be better to design array geometries whose performance is not so sensitive to variations in driver response, and also to design filters that are robust to such variations. Even if individual transfer functions were used to design the filters, the driver responses would still change with time and environmental factors such as temperature, so that robust filter design is still important.

Unfortunately it is difficult to calculate the filter responses that are robust to uncertainty in driver response. Doclo and Moonen (2007), for example, considered the related problem of designing a superdirective microphone array that was robust to mis-mismatching of the microphone responses. They used a numerical method to design a two microphone array whose performance was robust to variations in microphone gain having a maximum value of ± 30%, i.e., about ± 3 dB, when these variations had a uniform distribution. They found that although the directivity factor of this array with perfectly matched microphones was 9.52 dB, this fell to an average of 4.90 dB, even after optimisation, when this uncertainty in microphone response was taken into account.

These authors also considered other ways of designing robust arrays, however, including the use of regularisation. They found that with an appropriate choice of regularisation factor, the average performance of their array, when subject to the uncertainties in microphone response as above, was 4.88 dB. The average directivity factor of the array in the face of these uncertainties was only 1.33 dB if no regularisation was used. This suggests that the performance of a regularised array can come close to that of one specifically optimised to be robust to variations in transducer sensitivity.

Figure 6.3 The contrast and array effort as a function of frequency for a two source array when the drivers are perfectly matched, solid line, when there is a mis-match of +1 dB, dashed line, or -1 dB, dot dashed line, in their responses, and the mean of the performance with ±1 dB error shows the distribution of the acoustic contrast and the array effort at different frequencies when the source strengths are those calculated for equal responses in the two drivers but there is a ± 1 dB variation in the response of one driver with respect to the other. The mean performance of the array with the ± 1

dB error in the sensitivities is very much less than the optimum performance at low frequencies, where the optimal source strengths are almost equal but opposite in phase.

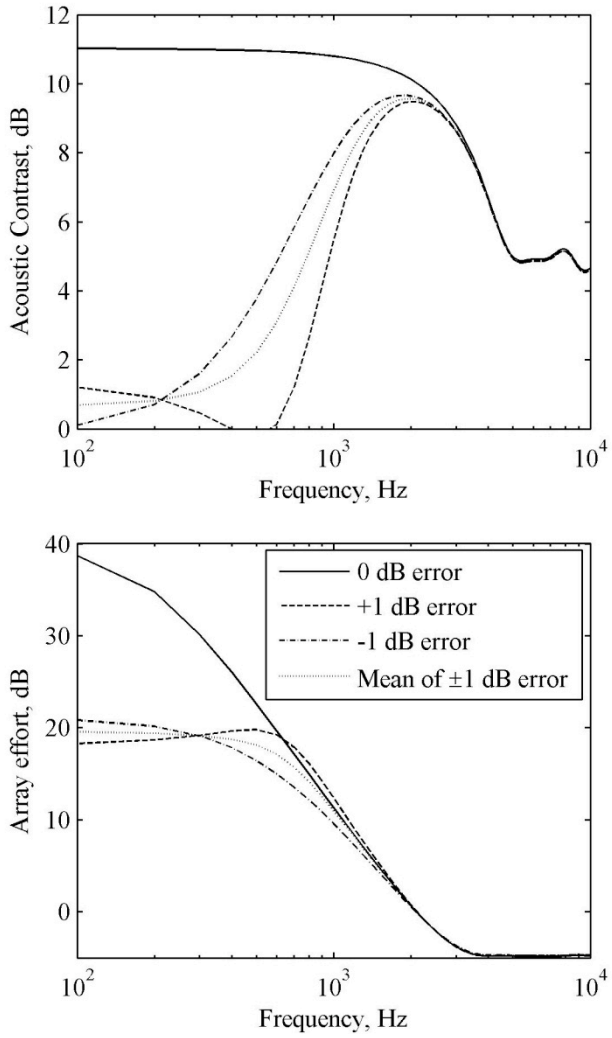


Figure 6.3 The contrast and array effort as a function of frequency for a two source array when the drivers are perfectly matched, solid line, when there is a mis-match of +1 dB, dashed line, or -1 dB, dot dashed line, in their responses, and the mean of the performance with ±1 dB error

Figure 6.4 Variation of the mean contrast for a two source array when the responses are perfectly matched (solid line) and when they are mis-matched by $\pm$ 1 dB (dashed line) as a function of the regularisation parameter for excitation frequencies of 100 Hz. shows the mean value of the contrast when the response is subject to these uncertainties as a function of the regularisation parameter for an excitation frequency of 100 Hz. The mean contrast with the uncertainties is about 4 dB with a regularisation parameter, β, of about 8×10^4, whereas it is only about 0.7 dB without. This shows that regularisation can improve the mean contrast when the sensitivities of the devices are uncertain, and has the added bonus of significantly reducing the mean array effort.

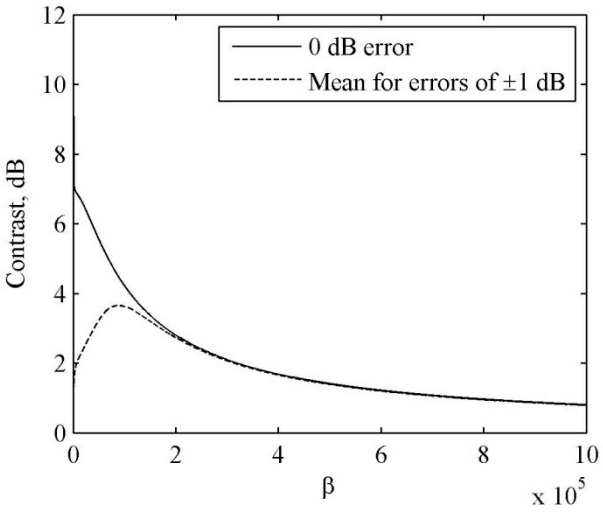


Figure 6.4 Variation of the mean contrast for a two source array when the responses are perfectly matched (solid line) and when they are mis-matched by ± 1 dB (dashed line) as a function of the regularisation parameter for excitation frequencies of 100 Hz.

Figure 6.5 shows the variation in response for the regularised array, with the regularisation parameter chosen at each frequency to maximise the acoustic contrast for errors between ± 1 dB, as shown in Figure 6.6. The results for the acoustic contrast performance support the statement that regularisation can improve the mean acoustic contrast in the presence of uncertainty in the driver responses, since the mean performance of the array with ± 1 dB error is better with regularisation than without and it is also shown that this is most significant at low frequencies. The presented array effort also indicates that the use of regularisation achieves a significant reduction in the required array effort. One of the most striking aspects of Figure 6.5 is that whereas the performance of the unregularised array drops by about 10 dB at low frequencies when ± 1 dB errors are present, the performance of the regularised array only falls by about 2 dB.

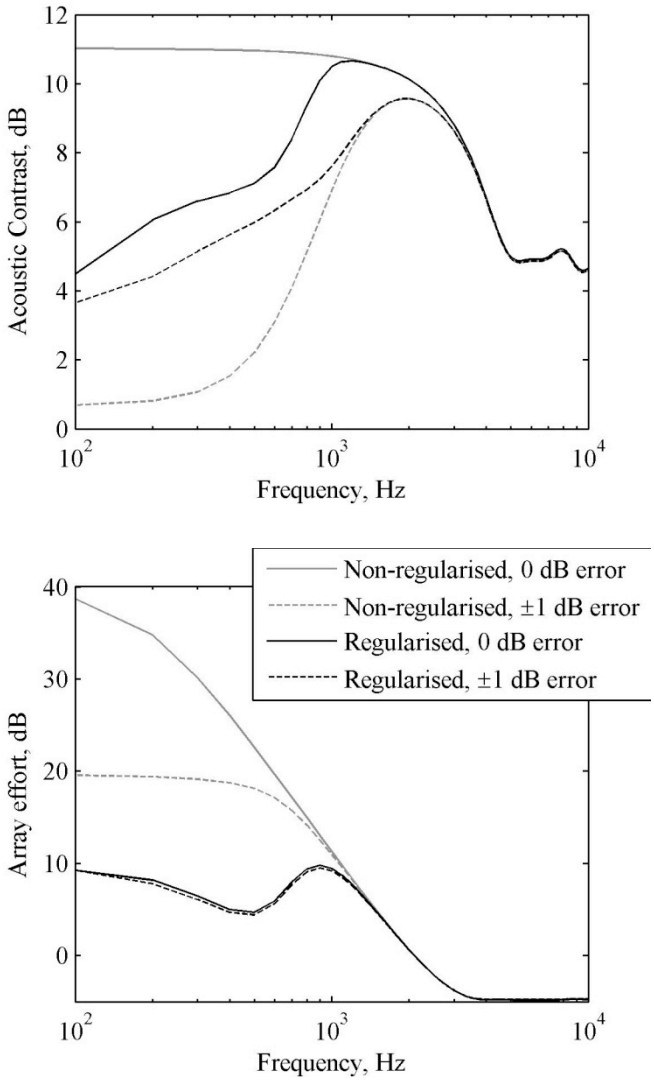


Figure 6.5 The contrast and array effort as a function of frequency for a three source array with regularisation (dark lines) and without regularisation (feint lines) when the drivers are perfectly matched (solid lines) and when there is a mis-match of +1 dB in their responses (dashed line).

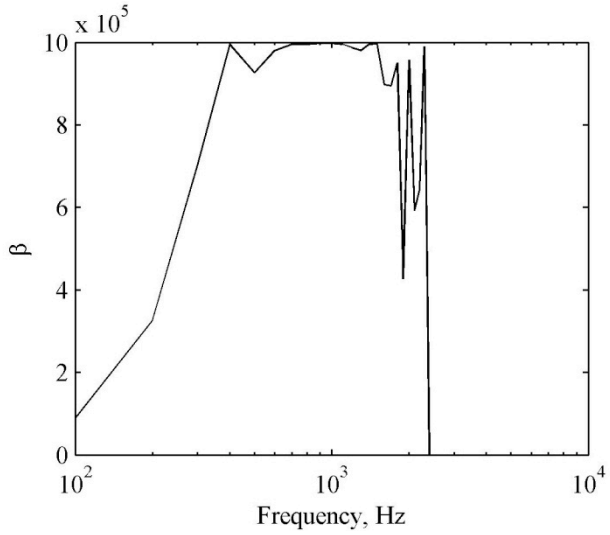


Figure 6.6 The value of β optimised to maximise the average contrast for errors in the driver response of ± 1 dB.

7. Conclusions

Compact arrays of acoustic drivers can be used to generate sound in one spatial region while minimising the sound in other spatial regions. The design of such systems is a generalisation of the conventional beamforming problem, in which the directivity is maintained in a single direction. The motivating example assumed here is of an endfire array of drivers used on a mobile device to reproduce sound for the user, but minimise sound radiation in other directions. This application imposes limits on both the size of the array and the power that is available to drive it.

This memo considered the regularisation of the optimal solution for such arrays, in order to both limit the array effort and increase the robustness of the array to changes in the acoustic environment and driver responses. It is first shown that three formulations of the regularised design of such an array behave in a similar way in theory, but have rather different numerical properties. It is shown that an indirect method provides the most robust and easily used method of imposing a constraint on the array effort in practice.

A further generalisation of the acoustic contrast control strategy is also presented which allows the mean square level to be specified in a number of bright zones. This is useful where a defined level difference is required between two zones whilst minimising the sound radiated to other regions. However, it is shown that the optimisation and trade-off between the multiple constraints is non-trivial.

Acoustic contrast control has up to this point been formulated in the frequency domain which does not ensure causal filters. A time domain formulation of the acoustic contrast strategy is presented which directly calculates the optimal filter coefficients. However, at this stage it is not clear if the presented formulation is practical.

The sensitivity of the array to changes in the acoustic environment is investigated by assuming that the transfer responses, from drivers to microphones, have both deterministic and random components. These are identified with the direct and reverberant pressures in an enclosure. At high frequencies the reverberant components become independent of each other and an analytical result for the optimum array effort can be obtained, which is very similar to the regularised result that limits the array effort. The sensitivity to uncertainties in the driver responses is more difficult to formulate analytically, but regulating the array effort is again found to improve the performance of the array when subject to these uncertainties. It is also shown that the performance of the array is very sensitive to the positions of the drivers when the size of the array is small compared with the acoustic wavelength.

The analysis of the sensitivity to position emphasises that the superdirective array can be understood by its multipole expansion. For the two source array it is the monopole and dipole components that interact to provide the optimum, hypercardioid, directivity. For the three source array it is the monopole, dipole and longitudinal quadrupole components that interact, in a rather delicate way, at low frequencies. This delicate balance is upset by changes in either the driver responses or small deviations from ideal locations.

The practical implications of this work for mobile devices are that it puts clear limits on the performance of such a driver array at low frequencies. For a given power requirement, the optimum performance can be calculated, but will be limited by the trade-off between performance and array effort. Changes in the acoustic environment are also inevitable when the device is used in practice, as are small differences in the sensitivity and acoustic position of the drivers from what they ideally would be. Although the array can be made more robust to these uncertainties, the price that must be paid is a clear limitation in their performance at low frequencies, although for the cases considered here, the performance is still a significant improvement over a delay and sum beamformer.

References

M.M. Boone and W-H Cho and J-G Ih (2009) Design of a highly directional endfire loudspeaker array, *Journal of the Audio Engineering Society,* **57**, 309-325.

J.-W. Choi and Y.-H. Kim (2002) Generation of an acoustically bright zone with an illuminated region using multiple sources. *Journal of the Acoustical Society of America* **111**, 1695-1700.

K.J. Ebeling (1984) Statistical properties of random wave fields. *Physical Acoustics* **XVII**, 233-310 (Eds. W.P. Mason and R.N. Thurston).

S.J. Elliott (2001) *Signal Processing for Active Control* (Academic Press)

S.J. Elliott, J. Cheer, H. Murfet and K.R. Holland (2010) Minimally radiating sources for personal audio, *Journal of the Acoustical Society of America* **128**(4), 1721-1278.

S. Doclo and M. Moonen (2006) Superdirective beamforming robust against microphone mismatch, *IEEE Transactions on Audio, Speech and Language Processing* **15**(2), 617-631.

G. H. Golub and C. F. Van Loan (1996) *Matrix Computations* (The John Hopkins University Press).

P.A. Nelson and S.J. Elliott (1992) *Active Control of Sound* (Academic Press).

A.D. Pierce (1981) *Acoustics: an introduction to its physical principles and applications* (McGraw-Hill Book co.)

M.R. Schroeder (1954) Die statistischen Parameter der Frequenzkurven von grossen Räumen. *Acustica* **41** 594-600. For an English translation see Schroeder, M.R. (1987), Statistical parameters of the frequency response curves of large rooms. *Journal of the Audio Engineering Society* **35**, 299-306.

M. Shin, S.Q. Lee, F.M. Fazi, P.A. Nelson, D. Kim, S. Wang, K.H. Park and J. Seo (2010) Maximization of acoustic energy difference between two spaces, *Journal of the Acoustical Society of America* **128**(1), 121-131.

B. V. Veen and K. Buckley (1988) A versatile approach to spatial filtering, *IEEE ASSP Magazine*, **5**, 4-24.

J. H. Wilkinson (1965) *The algebraic eigenvalue problem* (Clarendon Press).