OPTIMAL SOURCE DISTRIBUTION FOR MULTIPLE LISTENER VIRTUAL SOUND IMAGING

P.A. Nelson Institute of Sound and Vibration Research, University of Southampton
T. Takeuchi Kajima Corporation and Institute of Sound and Vibration Research, University of Southampton
P. Couturier Institute of Sound and Vibration Research, University of Southampton

1 INTRODUCTION

Takeuchi and Nelson [1] first proposed the Optimal Source Distribution (OSD) for achieving binaural sound reproduction for a single listener. The approach has proven to yield excellent subjective results [2] and has since been implemented in a number of products for virtual sound imaging. A remarkable property of the OSD is that the cross-talk cancellation produced at a single centrally placed listener is replicated at all frequencies at a number of other locations in the radiated sound field, a phenomenon that has since been further investigated [3, 4]. An analysis of this has recently been presented by Yairi et al [5] who also show that once a discrete approximation to the hypothetically continuous OSD is introduced, the effectiveness of the cross-talk cancellation is achieved at many but not all frequencies at the non-central positions in the radiated sound field. The work presented here provides a framework for the analysis of the multiple listener virtual sound imaging problem based on a linearly constrained least squares approach. The aim is to enable the exploitation of the fundamental property of the OSD with the objective of producing exact cross-talk cancellation for multiple listener positions at all frequencies. The background to the problem is introduced, and then the new theoretical approach is presented. Although the example given below is explained in terms of the original two-channel OSD system, it should also be emphasised that the approach presented is equally applicable to the three-channel extension [6].

2 THE OPTIMAL SOURCE DISTRIBUTION

Figure 1 shows the geometry of the two-loudspeaker/single listener problem and Figure 2 shows the equivalent block diagram, both figures replicating the notation used by Takeuchi and Nelson [1] and Yairi et al [5]. The following respectively define the desired signals for reproduction, the source signals and the reproduced signals

$$\mathbf{d}^{\mathrm{T}} = [d_R \quad d_L] \qquad \mathbf{v}^T = [v_R \quad v_L] \qquad \mathbf{w}^T = [w_R \quad w_L] \tag{1a,b,c}$$

and the inverse filter matrix and transmission path matrix are respectively defined by

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{2a,b}$$

where all variables are in the frequency domain and a harmonic time dependence of $e^{j\omega t}$ is assumed. Thus, $\mathbf{v} = \mathbf{Hd}$, $= \mathbf{Cv}$, and $\mathbf{w} = \mathbf{CHd}$. Assuming the sources are point monopoles radiating into a free field, with volume accelerations respectively given by v_R and v_L , the transmission path matrix takes the form

$$\mathbf{C} = \frac{\rho_0}{4\pi} \begin{bmatrix} \dfrac{e^{-jkl_1}}{l_1} & \dfrac{e^{-jkl_2}}{l_2} \\ \dfrac{e^{-jkl_2}}{l_2} & \dfrac{e^{-jkl_1}}{l_1} \end{bmatrix} \tag{3}$$

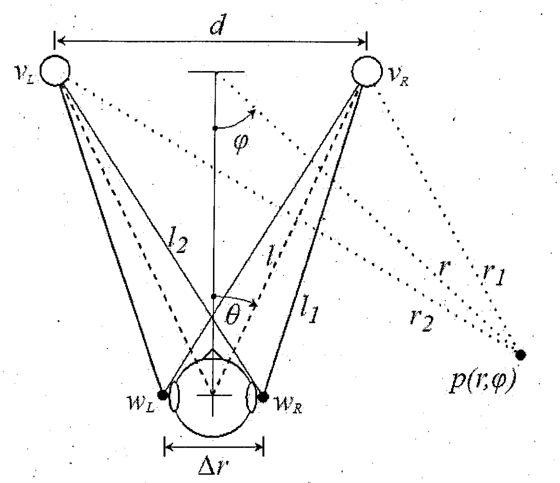

Figure 1. Geometry of the two source-single listener arrangement.

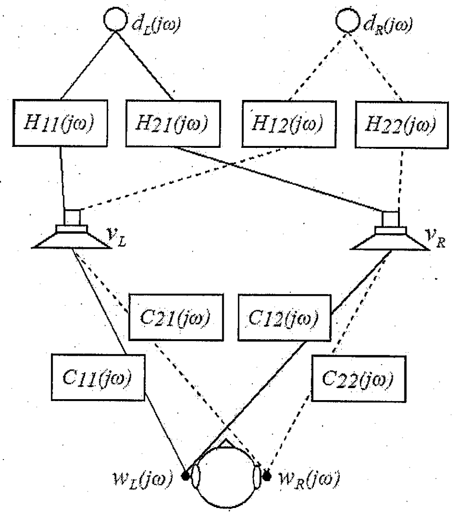

Figure 2. Equivalent block diagram

where the distances between the assumed point sources and the ears of the listener are as shown in Figure 1, $k = \omega/c_0$ and ρ_0, c_0 are respectively the density and sound speed. This matrix can be written as

$$\mathbf{C} = \frac{\rho_0 e^{-jkl_1}}{4\pi l_1}\begin{bmatrix} 1 & ge^{-jk\Delta l} \\ ge^{-jk\Delta l} & 1 \end{bmatrix} \quad (4)$$

where $g = l_1/l_2$ and $\Delta l = l_2 - l_1$. If it is assumed that the target values of the reproduced signals at the listeners ears are given by

$$\mathbf{d}^{\mathrm{T}} = \frac{\rho_0 e^{-jkl_1}}{4\pi l_1}[d_R \quad d_L] \quad (5)$$

it follows that [1] the inverse filter matrix is given by simple inversion of the elements of the matrix $\mathbf{C}$ yielding

$$\mathbf{H} = \frac{1}{1-g^2 e^{-2jk\Delta r \sin\theta}}\begin{bmatrix} 1 & -ge^{-jk\Delta r \sin\theta} \\ -ge^{-jk\Delta r \sin\theta} & 1 \end{bmatrix} \quad (6)$$

where the approximation $\Delta l \approx \Delta r \sin\theta$ has been used, assuming that $l \gg \Delta r$. Takeuchi and Nelson [1] present the singular value decomposition of the matrix $\mathbf{C}$ (and thus the matrix $\mathbf{H}$) and demonstrate that the two singular values are equal when

$$k\Delta r \sin\theta = n\pi/2 \quad (7)$$

where n is an odd integer (1,3,5...). Under these circumstances, the inversion problem is intrinsically well-conditioned and reproduction is accomplished with minimal error. Note that this condition is equivalent to the difference in path lengths Δl being equal to odd integer multiples of one quarter of an acoustic wavelength λ. Since the angle 2θ is equal to the angular span of the sources (see Figure 1), this condition also implies that the source span should vary continuously with frequency to preserve the $\lambda/4$ path length difference. The OSD is a therefore continuous distribution of pairs of sources, each radiating a different frequency, with those radiating high frequencies placed close together, and those radiating lower frequencies placed further apart. A further property of the OSD is that, provided the sources are distributed to preserve the path length difference condition $\Delta l = n\lambda/4$ where n is an odd integer, then the inverse filter matrix reduces to

$$\mathbf{H} = \frac{1}{1+g^2}\begin{bmatrix} 1 & -jg \\ -jg & 1 \end{bmatrix} \tag{8}$$

This gives a particularly simple form of filter matrix, and through the term $-jg$, involves simple inversion, 90-degree phase shift, and a small amplitude adjustment of the input signals.

3 RADIATION PROPERTIES OF THE OPTIMAL SOURCE DISTRIBUTION

As shown by Yairi et al [5] this idealised source distribution has some attractive radiation properties. This is best demonstrated by computing the far field pressure $p(r,\varphi)$ radiated by a pair of sources with inputs determined by the above optimal inverse filter matrix. Furthermore, if one assumes that the desired signals for reproduction are given by $\mathbf{d}^{\mathrm{T}} = [1 \quad 0]$, then it follows that the source signals are given by

$$\mathbf{v} = \mathbf{Hd} = \frac{1}{1+g^2}\begin{bmatrix} 1 & -jg \\ -jg & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{1+g^2}\begin{bmatrix} 1 \\ -jg \end{bmatrix} \tag{9}$$

Therefore the pressure field is given by the sum of the two source contributions such that

$$p(r,\varphi) = \frac{\rho_0}{4\pi(1+g^2)}\left[\frac{e^{-jkr_1}}{r_1} - jg\frac{e^{-jkr_2}}{r_2}\right] \tag{10}$$

which, writing $h = r_1/r_2$, can be written as

$$p(r,\varphi) = \frac{\rho_0 e^{-jkr_1}}{4\pi r_1(1+g^2)}\left[1 - jghe^{-jk(r_2-r_1)}\right] \tag{11}$$

Now note that in the far field, where $r_1, r_2 \gg d$, the distance between the sources, then it follows from the geometry of Figure 1, that

$$r_1 \approx r - \left(\frac{d}{2}\right)\sin\varphi, \qquad r_2 \approx r + \left(\frac{d}{2}\right)\sin\varphi \tag{12a,b}$$

and therefore that

$$p(r,\varphi) = \frac{\rho_0 e^{-jkr_1}}{4\pi r_1(1+g^2)}\left[1 - jghe^{-jkd\sin\varphi}\right] \tag{13}$$

The squared modulus of the term in square brackets is given by

$$\left|1 - jghe^{-jkd\sin\varphi}\right|^2 = 1 + (gh)^2 - 2gh\sin(kd\sin\varphi) \tag{14}$$

and therefore the modulus squared of the pressure field can be written as

$$|p(r,\varphi)|^2 = \left(\frac{\rho_0}{4\pi r_1}\right)^2\left(\frac{1+(gh)^2-2gh\sin(kd\sin\varphi)}{1+g^2}\right) \tag{15}$$

Now note that as r increases, such that we can make the approximation $r \approx r_1 \approx r_2$, one can also make the approximations $g \approx h \approx 1$ and this expression becomes

$$|p(r,\varphi)|^2 = \left(\frac{\rho_0}{4\pi r}\right)^2 (1 - \sin(kd \sin\varphi)) \quad (16)$$

and therefore maxima and minima are produced in the sound field when $kd\sin\varphi = n\pi/2$ where n is odd. The term $(1-\sin(kd\sin\varphi))$ becomes zero at $n = 1,5,9$ .... etc. and is equal to two when $n = 3,7,11$ ....etc. The form of the squared modulus of the sound pressure as a function of the angle φ is illustrated in Figure 3.

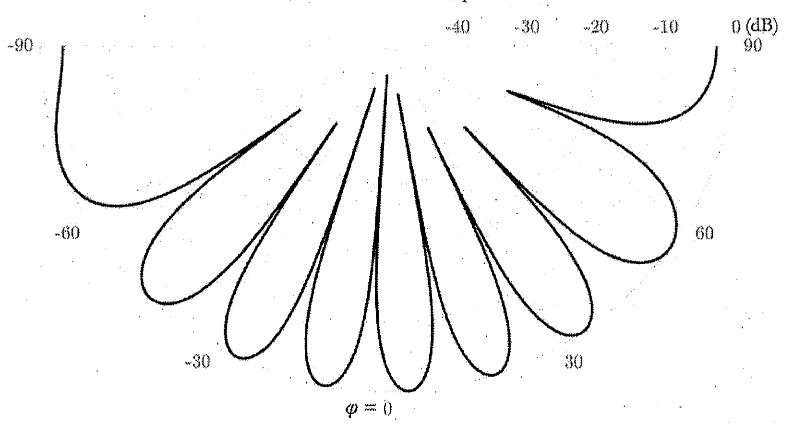

Figure 3. Directivity of the Optimal Source Distribution, showing the far field radiation pattern on a decibel scale as a function of the angle φ.

It is very important to note that this directivity pattern of the far field pressure is the **same at all frequencies**. This is because the value of $k\Delta r\sin\theta = n\pi/2$ where n is an odd integer, and since from the geometry of Figure 1, $\sin\theta = d/2l$, then $kd = n\pi l/2\Delta r$. Thus kd and therefore $\omega d/c_0$ takes a constant value (assuming $n = 1$), and is thus determined entirely by the geometrical arrangement of the two on-axis points chosen initially for cross-talk cancellation. The directivity pattern illustrated in Figure 3 demonstrates that an intrinsic property of the OSD is the production of cross-talk cancellation at multiple angular positions in the sound field. However, in any real application, any approximation to the OSD must be realised by a discrete number of loudspeakers. The work presented in the next section enables an approach to determining the signals input to a discrete number of sources in order to achieve the objective of producing cross-talk cancellation at multiple listener positions.

4 MULTIPLE DISCRETE SOURCES AND REPRODUCTION AT MULTIPLE POINTS

Now consider the case illustrated in Figure 4. The strength of a distributed array of acoustic sources is defined by a vector $\mathbf{v}$ of order M, and the pressure is defined at a number of points in the sound field by the vector $\mathbf{w}$ of order L. The curved geometry chosen here replicates that analysed by Yairi et al [7] who demonstrate a number of analytical advantages in working with such a source and sensor arrangement. In general $\mathbf{w} = \mathbf{C}\mathbf{v}$ where $\mathbf{C}$ is an $L \times M$ matrix. However, it will be helpful to partition this matrix $\mathbf{C}$ into two other matrices $\mathbf{A}$ and $\mathbf{B}$ and also partition the vector $\mathbf{w}$ of reproduced signals so that

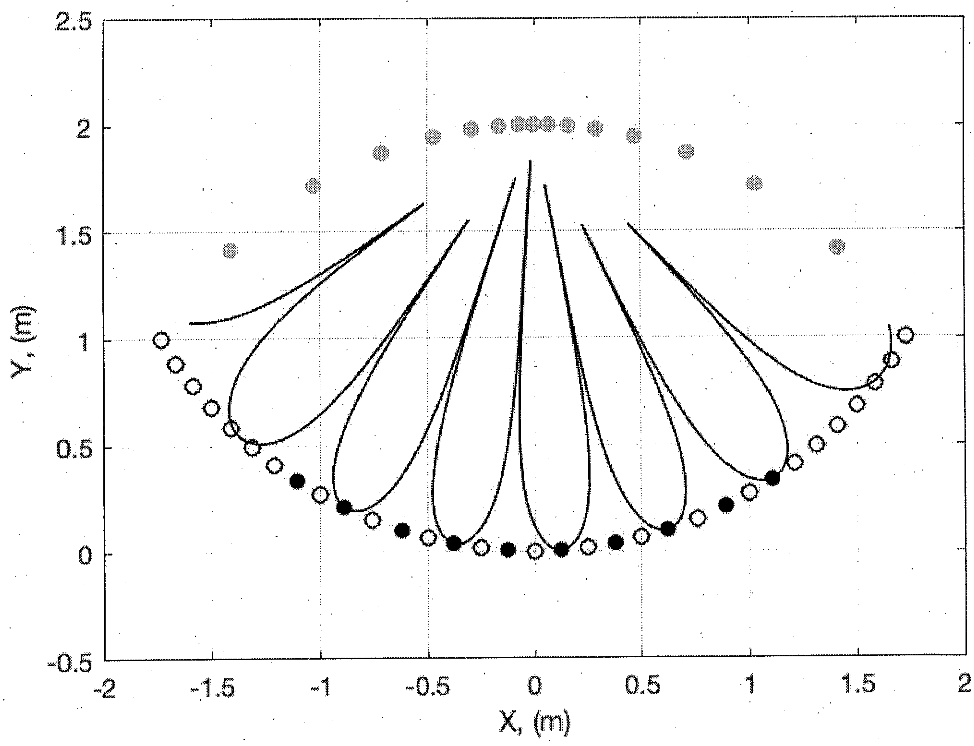

Figure 4. The arrangement of *M* point sources (grey symbols) and *L* points in the sound field at which the complex sound pressure is sampled. Cross-talk cancellation is desired at *P* points (black symbols) whilst a least squares fit to the OSD sound field is desired at $L - P = N$ points (white symbols).

$$\begin{bmatrix} \mathbf{w}_\mathrm{A} \\ \mathbf{w}_\mathrm{B} \end{bmatrix} = \mathbf{C}\mathbf{v} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathbf{v} \tag{17}$$

The vector $\mathbf{w}_\mathrm{B}$ is of order P and defines the reproduced signals at a number of pairs of points in the sound field at which cross-talk cancellation is sought. Thus $\mathbf{w}_\mathrm{B} = \mathbf{B}\mathbf{v}$ where $\mathbf{B}$ defines the $P \times M$ transmission path matrix relating the strength of the M sources to these reproduced signals. The vector $\mathbf{w}_\mathrm{A}$ is of order N and defines the reproduced signals sampled at the remaining points in the sound field. Thus $N = L - P$ and the reproduced field at these remaining points can be written as $\mathbf{w}_\mathrm{A} = \mathbf{A}\mathbf{v}$, where $\mathbf{A}$ is the $N \times M$ transmission path matrix between the sources and these points.
Now note that the desired pressure at the P points in the sound field at which cross-talk cancellation is required can be written as the vector $\hat{\mathbf{w}}_\mathrm{B}$. This vector can in turn be written as $\hat{\mathbf{w}}_\mathrm{B} = \mathbf{D}\mathbf{d}$, where the matrix $\mathbf{D}$ defines the reproduced signals required in terms of the desired signals. As a simple example, suppose that cross-talk cancellation is desired at two pairs of points in the sound field such that

$$\begin{bmatrix} \hat{w}_{B1} \\ \hat{w}_{B2} \\ \hat{w}_{B3} \\ \hat{w}_{B4} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_R \\ d_L \end{bmatrix} \tag{18}$$

The matrix $\mathbf{D}$ has elements of either zero or unity, is of order $P \times 2$, and may be extended by adding further pairs of rows if cross talk cancellation is required at further pairs of points. Similar to the analysis presented above, we assume that the inputs to the sources are determined by operating on the two desired signals defined by the vector $\mathbf{d}$ via an $M \times 2$ matrix $\mathbf{H}$ of inverse filters. The task is to find the source strength vector $\mathbf{v}$ that generates cross-talk cancellation at multiple pairs of positons in the sound field, guided by the observation of the directivity pattern illustrated in Figure 3.

5 MINIMUM NORM SOLUTION

Assume that cross-talk cancellation is required at the specific sub-set of all the points in the sound field where the number P of points in the sound field is smaller than the number of sources M available to reproduce the field. One can then seek to ensure that we make $\mathbf{w}_\mathrm{B} = \hat{\mathbf{w}}_\mathrm{B} = \mathbf{Dd}$ whilst minimising the "effort"$\|\mathbf{v}\|_2^2$ made by the acoustic sources. The problem is thus

$$min\|\mathbf{v}\|_2^2 \ \text{subject to} \ \hat{\mathbf{w}}_\mathrm{B} = \mathbf{Dd} \tag{19}$$

where $\| \ \ \|_2$ denotes the 2-norm. The well-known solution [8, 9] to this minimum norm problem is given by the optimal vector of source strengths defined by

$$\mathbf{v}_{opt} = \mathbf{B}^\mathrm{H}[\mathbf{BB}^\mathrm{H}]^{-1}\mathbf{Dd} \tag{20}$$

where the superscript H denotes the Hermitian transpose. Thus a possible solution to the problem can be found that requires only specification of the points at which cross-talk cancellation is required in the sound field. A sensible approach might be to use the directivity of the OSD as a guide to the choice of angular location these points. Note that it is also possible to include a regularisation factor α into this solution such that

$$\mathbf{v}_{opt} = \mathbf{B}^\mathrm{H}[\mathbf{BB}^\mathrm{H} + \alpha\mathbf{I}]^{-1}\mathbf{Dd} \tag{21}$$

where $\mathbf{I}$ is the identity matrix.

6 LINEARLY CONSTRAINED LEAST SQUARES SOLUTION

6.1 Extension of the unconstrained solution

A further approach to the exploitation of the known properties of the optimal source distribution is to attempt not only to achieve cross-talk cancellation at multiple pairs of points as in the case above, but also to attempt a "best fit" of the radiated sound field to the known directivity function of the OSD. In this case, the problem is a least squares minimisation with an equality constraint. Thus, as above, a sub-set of desired signals at pairs of points in the sound field is defined by $\hat{\mathbf{w}}_\mathrm{B} = \mathbf{Dd}$. The aim is also to minimise the sum of squared errors between the desired sound field and the reproduced field at the remaining $L - P = N$ points of interest (see Figure 4).

Here it is assumed that the number N of these points is greater than the number of sources M.The desired sound field at these points can be chosen to emulate the directivity of the OSD at angular positions between those at which cross-talk cancellation is sought. One therefore seeks the solution of

$$min \ \|\mathbf{Av} - \hat{\mathbf{w}}_\mathrm{A}\|_2^2 \ \text{subject to} \ \hat{\mathbf{w}}_\mathrm{B} = \mathbf{Bv} \tag{22}$$

Before moving to exact solutions, it is worth noting that Golub and Van Loan [9] point out that an approximate solution to the linearly constrained least squares problem is to solve the unconstrained least squares problem given by

$$min \left\| \begin{bmatrix} \mathbf{A} \\ \gamma\mathbf{B} \end{bmatrix} \mathbf{v} - \begin{bmatrix} \hat{\mathbf{w}}_{\mathrm{A}} \\ \gamma\hat{\mathbf{w}}_{\mathrm{B}} \end{bmatrix} \right\|_2^2 \quad (23)$$

They demonstrate that the solution to this problem, which is a standard least squares problem, converges to the constrained solution as $\gamma \to \infty$. They also point out, however, that numerical problems may arise as γ becomes large.

6.2 Solution using QR factorization

An exact solution to this problem can deduced by following Golub and Van Loan [9], but working with complex variables and replacing the matrix transpose operation by the complex conjugate transpose. The method relies on the use of the use of the QR-factorisation of the matrix $\mathbf{B}^{\mathrm{H}}$ where

$$\mathbf{B}^{\mathrm{H}} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \quad (24)$$

where $\mathbf{B}^{\mathrm{H}}$ is $M \times P$, $\mathbf{Q}$ is an $M \times M$ square matrix having the property $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{\mathrm{H}} = \mathbf{I}$ and $\mathbf{R}$ is an upper triangular matrix of order $P \times P$ and the zero matrix is of order $(M - P) \times P$. Now define

$$\mathbf{AQ} = [\mathbf{A_1 A_2}] \qquad \mathbf{Q}^{\mathrm{H}}\mathbf{v} = \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \quad (25)$$

where the matrix $\mathbf{A_1}$ is $N \times P$, the matrix $\mathbf{A_2}$ is $N \times (M - P)$, and the vectors $\mathbf{y}$ and $\mathbf{z}$ are of order P and $M - P$ respectively. As shown in Appendix 1, the optimal solution to the least squares problem can be written as

$$\mathbf{v}_{opt} = \mathbf{Q} \begin{bmatrix} \mathbf{y}_{opt} \\ \mathbf{z}_{opt} \end{bmatrix} \quad (26)$$

and partitioning the matrix $\mathbf{Q}$ such that $\mathbf{v}_{opt} = \mathbf{Q}_1\mathbf{y}_{opt} + \mathbf{Q}_2\mathbf{z}_{opt}$ enables the solution to be written as

$$\mathbf{v}_{opt} = \mathbf{Q}_2\mathbf{A}_2^{\dagger}\hat{\mathbf{w}}_{\mathrm{A}} + \left(\mathbf{Q}_1\mathbf{R}^{\mathrm{H}-1} - \mathbf{Q}_2\mathbf{A}_2^{\dagger}\mathbf{A}_1\mathbf{R}^{\mathrm{H}-1}\right)\hat{\mathbf{w}}_{\mathrm{B}} \quad (27)$$

where $\mathbf{A}_2^{\dagger} = \left[\mathbf{A}_2^{\mathrm{H}}\mathbf{A}_2\right]^{-1}\mathbf{A}_2^{\mathrm{H}}$ is the pseudo inverse of the matrix $\mathbf{A_2}$. This enables the calculation of the optimal source strengths in the discrete approximation to the OSD in terms of the signals $\hat{\mathbf{w}}_{\mathrm{B}}$ reproduced at the points of cross talk cancellation, and the remaining signals $\hat{\mathbf{w}}_{\mathrm{A}}$ specified by the directivity of the OSD.

6.3 Solution using the method of Lagrange multipliers

Whilst the above solution provides a compact and efficient means for solving the problem at hand, it is shown in Appendix 2 that it is also possible to derive a solution that is expressed explicitly in terms of the matrices $\mathbf{A}$ and $\mathbf{B}$. The solution can be derived by using the method of Lagrange multipliers where one seeks to minimise the cost function given by

$$J = (\mathbf{Av} - \hat{\mathbf{w}}_{\mathrm{A}})^{\mathrm{H}}(\mathbf{Av} - \hat{\mathbf{w}}_{\mathrm{A}}) + (\mathbf{Bv} - \hat{\mathbf{w}}_{\mathrm{B}})^{\mathrm{H}}\boldsymbol{\mu} + \boldsymbol{\mu}^{\mathrm{H}}(\mathbf{Bv} - \hat{\mathbf{w}}_{\mathrm{B}}) + \beta\mathbf{v}^{\mathrm{H}}\mathbf{v} \quad (28)$$

where $\boldsymbol{\mu}$ is a complex vector of Lagrange multipliers and the term β is used to penalise the "effort" associated with the sum of squared source strengths. As shown in Appendix 2, the minimum of this function is defined by

$$\mathbf{v}_{opt} = \left[\mathbf{I} - \widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\mathbf{B}\right]\mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} + \widetilde{\mathbf{A}}\,\mathbf{B}^{\mathrm{H}}\left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\widehat{\mathbf{w}}_{\mathrm{B}} \qquad (29)$$

where the matrices $\widetilde{\mathbf{A}}$ and $\mathbf{A}^{\dagger}$ are respectively defined by

$$\widetilde{\mathbf{A}} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1} \quad \text{and} \quad \mathbf{A}^{\dagger} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1}\mathbf{A}^{\mathrm{H}} \qquad (30)$$

This solution has also been derived by Olivieri et al [10], although the solution presented by those authors differs from that above by a factor ½ that multiplies the first term in square brackets above.

7 RESULTS OF NUMERICAL SIMULATIONS

7.1 Comparison of numerical solution techniques

Whilst full details are not presented here, extensive simulation work has been undertaken using MATLAB to evaluate the minimum norm solution and three possible methods of computing the linearly constrained least squares solution. There is inevitably a tendency for inversion problems of this type to become ill-conditioned, especially when there are a large number of sources and sensors and at low frequencies (long wavelengths). As an illustration, Figure 5 shows the condition numbers of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ for a specific arrangement of M = 15 sources with cross-talk cancellation targeted at five listener positions (P = 10) and a further number N = 23 points in the sound field at which a "best fit" to the OSD sound field is sought.

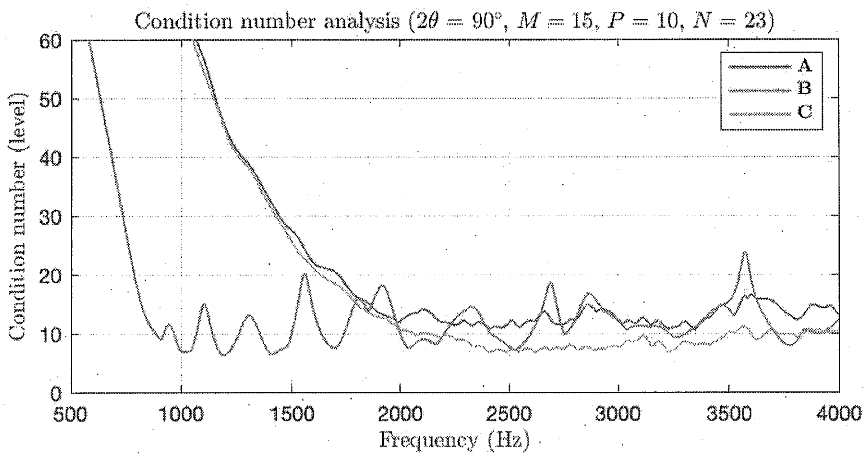

Figure 5. The condition number (on a decibel scale) of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ for M = 15, P = 10 and N = 23. Note that in this example the sources are placed on a 180 degree arc symmetrically surrounding the central listener.

It is clear that the matrix $\mathbf{A}$ is poorly conditioned at low frequencies, whereas the matrix $\mathbf{B}$, which is at the heart of the minimum norm solution, is much better conditioned at low frequencies, but suffers from a deterioration in conditioning at higher frequencies (although this has the potential to be addressed through regularisation). It is also apparent that the conditioning of the matrix $\mathbf{C}$ is dominated by the conditioning of the matrix $\mathbf{A}$ at low frequencies. In the case of the "extension to the unconstrained solution", it has been found necessary to choose the factor γ appropriately, and as suggested by Golub and Van Loan [9], this has to be made sufficiently large to ensure that the linear constraint is applied, without giving rise to numerical difficulties. Similarly, the "method of Lagrange multipliers" requires the appropriate choice of regularisation factor if a stable numerical solution is to be achieved. The method of solution using QR factorization as presented above does not require the

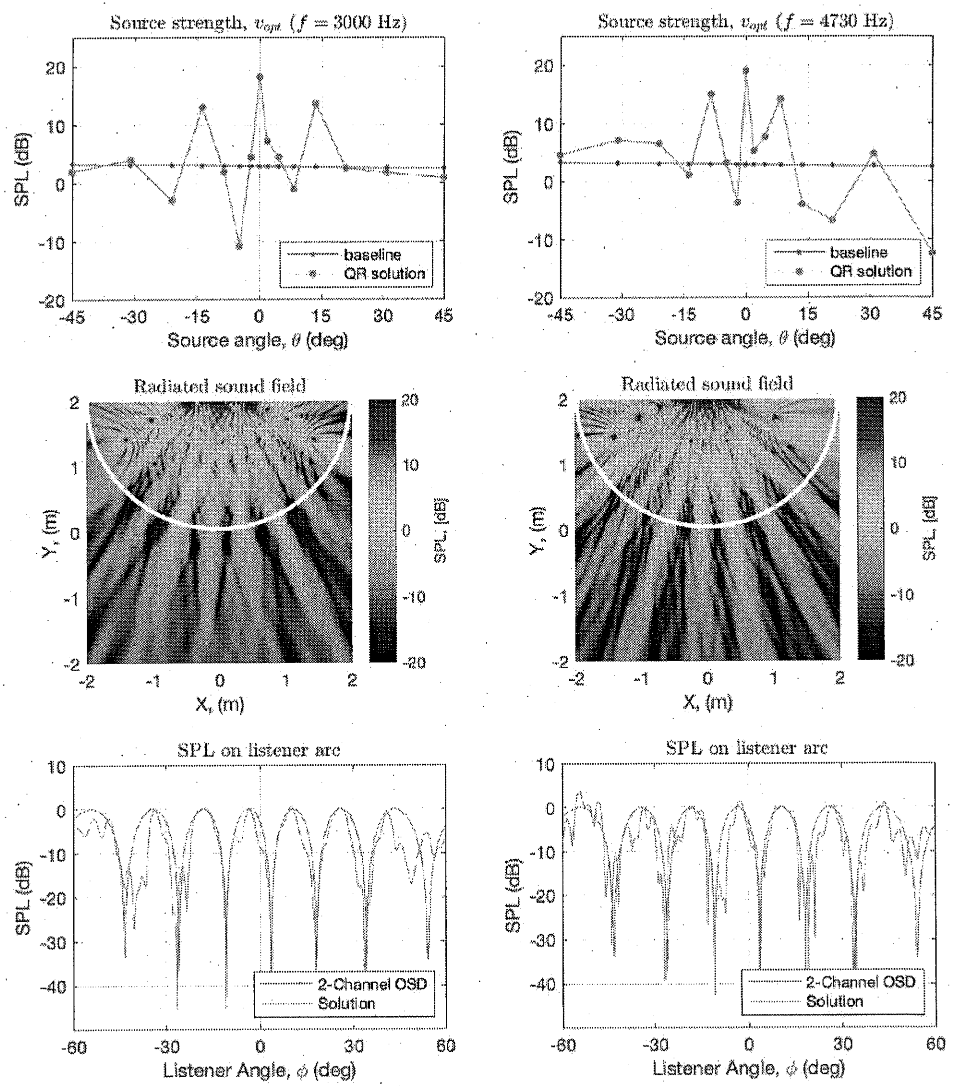

Figure 6. The sound fields at frequencies of 3000 and 4730 Hz where a good fit to the OSD sound field was achieved whilst ensuring cross-talk cancellation for five listeners.

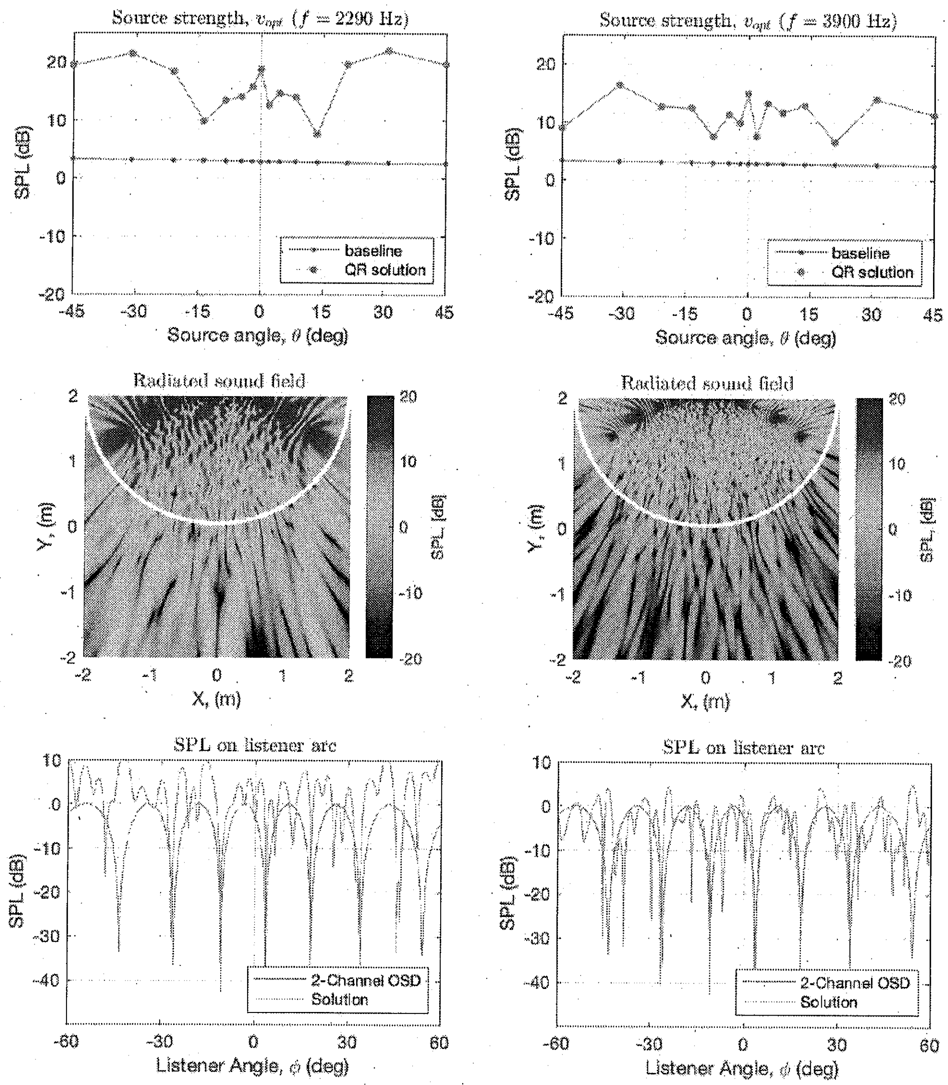

Figure 7. The sound fields at frequencies of 2290 and 3900 Hz where a good fit to the OSD sound field was not produced, despite cross-talk cancellation being achieved in narrow angular arcs around the target positions for five listeners.

choice of any free parameters, although the possibility exists of regularization of the inversion leading to matrix $\mathbf{A}_2^\dagger$ in equation (27) above. Current work is aimed at evaluating the relative merits of the solution techniques, the success of which have been found to depend on the specific range of parameters under consideration (i.e. number and specific arrangement of sources, the number of listeners, and the frequency). A more detailed comparison of the methods will be presented shortly [11].

7.2 Results of applying the constrained least squares solution

It has been found however, that the constrained least squares solution is consistently successful in ensuring cross-talk cancellation at the listener positions chosen. This is also true of the basic "minimum norm" solution aimed at ensuring cross-talk cancellation without attempting to match any particular prescribed sound field. However, for the purposes of illustration, some typical results are presented here that have been computed by using the QR factorization method (without the application of any regularisation). These results have been computed for an arrangement comprising of M = 15 sources with cross-talk cancellation enforced at P = 10 locations in the sound field (i.e. for five listeners) and a least squares fit to the OSD sound field required at N = 155 other positions.

Figure 6 shows the results at two frequencies where a very good fit was produced to the OSD sound field. In this case the target OSD sound field was that computed for the particular arrangement of sources on the arc shown in the figures and the white curve in the sound field shows the arc along which the target positions in the sound field were specified. The upper graphs show the modulus of the source strengths computed using QR factorization at the particular frequencies illustrated. These are compared with the "baseline" source strengths necessary to produce a target sound field of unity at a single point at the right ear of the central listener. The magnitude of the sound fields produced are illustrated in the middle layer of figures whilst the modulus of the sound pressure produced are shown in the lower layer of figures. Note that cross-talk cancellation is produced as prescribed.

The sound fields at a couple of frequencies at which the problem is poorly conditioned are shown in Figure 7. In these cases a good fit to the OSD sound field was not achieved, despite the enforcement of cross-talk cancellation at appropriate positions in the sound field. These results show that the application of the constraint is largely achieved, but that the angular range over which the cross-talk cancellation is effective become very small at these frequencies. These figures also show that the source strength necessary to achieve the crosstalk cancellation are of a consistently high magnitude, such source strengths being consistent with an ill-conditioned inversion problem.

8. CONCLUSIONS

The Optimal Source Distribution (OSD) is a symmetric distribution of pairs of point monopole sources whose separation varies continuously as a function of frequency in order to ensure at all frequencies a path length difference of one-quarter of an acoustic wavelength between the source pairs and the ears of a listener. The field of the OSD has a directivity function that is independent of frequency that in principle can produce cross-talk cancellation at a number of listener positions simultaneously over a wide frequency range. The problem of approximating the field of the OSD with a discrete number of transducers can be tackled using a linearly constrained least squares approach in which cross-talk cancellation is enforced at a number of listener positions in the sound field, whilst ensuring a best fit to the ideal OSD radiation. The technique shows some promise, with the OSD sound field replicated accurately at a number of frequencies whilst the field is not replicated accurately at others, even though cross-talk cancellation is nevertheless achieved to some extent at the desired positions. Investigations to date also indicate that the use of the minimum norm solution may yield better results at low frequencies due to a relative improvement in the conditioning of the problem. The implications of these findings for the subjective response of listeners in locating virtual sound images requires further investigation. It is also the intention to investigate the potential for the application of other regularisation methods that encourage sparsity in the solution for the source strengths.

APPENDIX 1. QR-FACTORISATION FOR THE DETERMINATION OF THE LINEARLY CONSTRAINED LEAST SQUARES SOLUTION

Using the definitions given in the main text, it can be shown that

$$\mathbf{Bv} = \left(\mathbf{Q}\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\right)^{\mathrm{H}}\mathbf{v} = [\mathbf{R}^{\mathrm{H}}\mathbf{0}]\mathbf{Q}^{\mathrm{H}}\mathbf{v} = [\mathbf{R}^{\mathrm{H}}\mathbf{0}]\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix} = \mathbf{R}^{\mathrm{H}}\mathbf{y} \tag{A1.1}$$

Using the property $\mathbf{QQ}^{\mathrm{H}} = \mathbf{I}$ also shows that

$$\mathbf{Av} = \mathbf{AQQ}^{\mathrm{H}}\mathbf{v} = [\mathbf{A}_1\mathbf{A}_2]\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix} = \mathbf{A}_1\mathbf{y} + \mathbf{A}_2\mathbf{z} \tag{A1.2}$$

This enables the problem to be transformed into a minimisation problem where one seeks to minimise $\|\mathbf{A}_1\mathbf{y} + \mathbf{A}_2\mathbf{z} - \widehat{\mathbf{w}}_{\mathrm{A}}\|^2$ subject to $\mathbf{R}^{\mathrm{H}}\mathbf{y} = \widehat{\mathbf{w}}_{\mathrm{B}}$. Then since $\mathbf{y} = \mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}}$ the minimisation problem reduces to

$$min\|\mathbf{A}_2\mathbf{z} - (\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{A}_1\mathbf{y})\|_2^2 = min\left\|\mathbf{A}_2\mathbf{z} - \left(\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{A}_1\mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}}\right)\right\|_2^2 \tag{A1.3}$$

This can be solved for $\mathbf{z}$ to give the following solution for the optimal source strength vector

$$\mathbf{v}_{opt} = \mathbf{Q}\begin{bmatrix}\mathbf{y}_{opt}\\ \mathbf{z}_{opt}\end{bmatrix} \tag{A1.4}$$

where the least squares solution to the minimisation problem involving the vector $\mathbf{z}$ can be written as

$$\mathbf{z}_{opt} = \left[\mathbf{A}_2^{\mathrm{H}}\mathbf{A}_2\right]^{-1}\mathbf{A}_2^{\mathrm{H}}\left(\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{A}_1\mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}}\right) \tag{A1.5}$$

and the modified constraint equation above gives

$$\mathbf{y}_{opt} = \mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}} \tag{A1.6}$$

Now note that this solution can be written explicitly in terms of the optimal vector of source strengths by partitioning of the matrix $\mathbf{Q}$ such that

$$\mathbf{v}_{opt} = \mathbf{Q}_1\mathbf{y}_{opt} + \mathbf{Q}_2\mathbf{z}_{opt} \tag{A1.7}$$

Also writing the pseudo inverse of matrix $\mathbf{A}_2$ as $\left[\mathbf{A}_2^{\mathrm{H}}\mathbf{A}_2\right]^{-1}\mathbf{A}_2^{\mathrm{H}} = \mathbf{A}_2^{\dagger}$, enables the solution to be written as

$$\mathbf{v}_{opt} = \mathbf{Q}_1\mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}} + \mathbf{Q}_2\mathbf{A}_2^{\dagger}\left(\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{A}_1\mathbf{R}^{\mathrm{H}-1}\widehat{\mathbf{w}}_{\mathrm{B}}\right) \tag{A1.8}$$

and therefore as

$$\mathbf{v}_{opt} = \mathbf{Q}_2\mathbf{A}_2^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} + \left(\mathbf{Q}_1\mathbf{R}^{\mathrm{H}-1} - \mathbf{Q}_2\mathbf{A}_2^{\dagger}\mathbf{A}_1\mathbf{R}^{\mathrm{H}-1}\right)\widehat{\mathbf{w}}_{\mathrm{B}} \tag{A1.9}$$

APPENDIX 2. THE LINEARLY CONSTRAINED LEAST SQUARES SOLUTION USING THE METHOD OF LAGRANGE MULTIPLIERS

Another method for determining the solution to

$$min\ \|\mathbf{Av}-\widehat{\mathbf{w}}_{\mathrm{A}}\|_2^2 \text{ subject to } \widehat{\mathbf{w}}_{\mathrm{B}} = \mathbf{Bv} \tag{A2.1}$$

Is to use the method of Lagrange multipliers, which is widely used in the solution of constrained optimisation problems. The analysis presented here is similar to that presented previously by Olivieri et al [10] and by Nelson and Elliott [8]. The analysis begins by defining a cost function J given by

$$J = (\mathbf{Av}-\widehat{\mathbf{w}}_{\mathrm{A}})^{\mathrm{H}}(\mathbf{Av}-\widehat{\mathbf{w}}_{\mathrm{A}}) + (\mathbf{Bv}-\widehat{\mathbf{w}}_{\mathrm{B}})^{\mathrm{H}}\boldsymbol{\mu} + \boldsymbol{\mu}^{\mathrm{H}}(\mathbf{Bv}-\widehat{\mathbf{w}}_{\mathrm{B}}) + \beta\mathbf{v}^{\mathrm{H}}\mathbf{v} \tag{A2.2}$$

where $\boldsymbol{\mu}$ is a complex vector of Lagrange multipliers and the term $\beta\mathbf{v}^{\mathrm{H}}\mathbf{v}$ is included, as will become apparent, in order to regularise the inversion of a matrix in the solution. The derivatives of this function with respect to both $\mathbf{v}$ and $\boldsymbol{\mu}$ are defined by

$$\frac{\partial J}{\partial \mathbf{v}} = \frac{\partial J}{\partial \mathbf{v}_{\mathrm{R}}} + j\frac{\partial J}{\partial \mathbf{v}_{\mathrm{I}}}, \quad \frac{\partial J}{\partial \boldsymbol{\mu}} = \frac{\partial J}{\partial \boldsymbol{\mu}_{\mathrm{R}}} + j\frac{\partial J}{\partial \boldsymbol{\mu}_{\mathrm{I}}} \tag{A2.3}$$

where $\mathbf{v} = \mathbf{v}_{\mathrm{R}} + j\mathbf{v}_{\mathrm{I}}$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_{\mathrm{R}} + j\boldsymbol{\mu}_{\mathrm{I}}$. The following identities can be deduced from the analysis presented by Nelson and Elliott [8] (see the Appendix):

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^{\mathrm{H}}\mathbf{Gx}\right) = 2\mathbf{Gx}, \ \frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^{\mathrm{H}}\mathbf{b} + \mathbf{b}^{\mathrm{H}}\mathbf{x}\right) = 2\mathbf{b} \tag{A2.4a,b}$$

Expanding the first term in the above expression for the cost function J gives

$$J = \mathbf{v}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{Av} - \mathbf{v}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} - \widehat{\mathbf{w}}_{\mathrm{A}}^{\mathrm{H}}\mathbf{Av} + \widehat{\mathbf{w}}_{\mathrm{A}}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} + (\mathbf{Bv}-\widehat{\mathbf{w}}_{\mathrm{B}})^{\mathrm{H}}\boldsymbol{\mu} + \boldsymbol{\mu}^{\mathrm{H}}(\mathbf{Bv}-\widehat{\mathbf{w}}_{\mathrm{B}}) + \beta\mathbf{v}^{\mathrm{H}}\mathbf{v} \tag{A2.5}$$

and using the above identities shows that the minimum in the cost function is given by

$$\frac{\partial J}{\partial \mathbf{v}} = 2\mathbf{A}^{\mathrm{H}}\mathbf{Av} - 2\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} + \mathbf{2B}^{\mathrm{H}}\boldsymbol{\mu} + 2\beta\mathbf{v} = 0 \tag{A2.6}$$

$$\frac{\partial J}{\partial \boldsymbol{\mu}} = \mathbf{Bv} - \widehat{\mathbf{w}}_{\mathrm{B}} = 0 \tag{A2.7}$$

Note that these equations can also be written in matrix form as

$$\begin{bmatrix}\mathbf{A}^{\mathrm{H}}\mathbf{A} & \mathbf{B}^{\mathrm{H}} \\ \mathbf{B} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{v} \\ \boldsymbol{\mu}\end{bmatrix} = \begin{bmatrix}\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} \\ \widehat{\mathbf{w}}_{\mathrm{B}}\end{bmatrix} \tag{A2.8}$$

and are sometimes known as the Karush-Kuhn-Tucker (KKT) conditions. Rearranging the first of these equations shows that

$$[\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]\mathbf{v} = \mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{B}^{\mathrm{H}}\boldsymbol{\mu} \tag{A2.9}$$

and therefore

$$\mathbf{v} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1}\left(\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{B}^{\mathrm{H}}\boldsymbol{\mu}\right) \tag{A2.10}$$

The above relationship can be solved for the vector $\boldsymbol{\mu}$ of Lagrange multipliers by using $\mathbf{Bv} = \widehat{\mathbf{w}}_{\mathrm{B}}$ so that

$$\mathbf{B}[\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1}\left(\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{B}^{\mathrm{H}}\boldsymbol{\mu}\right) = \widehat{\mathbf{w}}_{\mathrm{B}} \tag{A2.11}$$

Further manipulation can be simplified by defining the following expressions:

$$\mathbf{A}^{\dagger} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1}\mathbf{A}^{\mathrm{H}}, \quad \text{and} \quad \widetilde{\mathbf{A}} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1} \qquad \text{(A2.12a,b)}$$

These enable the above equation to be written as

$$\mathbf{B}\mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} - \mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\boldsymbol{\mu} = \widehat{\mathbf{w}}_{\mathrm{B}} \qquad \text{(A2.13)}$$

It then follows that the expression for the vector of Lagrange multipliers can be written as

$$\boldsymbol{\mu} = \left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\left(\mathbf{B}\mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} - \widehat{\mathbf{w}}_{\mathrm{B}}\right) \qquad \text{(A2.14)}$$

Substituting this into the expression for the source strength vector yields the optimal value given by

$$\mathbf{v}_{opt} = \mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} - \widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\left(\mathbf{B}\mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} - \widehat{\mathbf{w}}_{\mathrm{B}}\right) \qquad \text{(A2.15)}$$

A little rearrangement shows that this expression can also be written as

$$\mathbf{v}_{opt} = \left[\mathbf{I} - \widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\mathbf{B}\right]\mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} + \widetilde{\mathbf{A}}\,\mathbf{B}^{\mathrm{H}}\left[\mathbf{B}\,\widetilde{\mathbf{A}}\mathbf{B}^{\mathrm{H}}\right]^{-1}\widehat{\mathbf{w}}_{\mathrm{B}} \qquad \text{(A2.16)}$$

It is worth noting that in the absence of the linear constraint, such that $\mathbf{B}\widehat{\mathbf{w}}_{\mathrm{B}} = 0$, then the solution reduces to the usual regularised least squares solution

$$\mathbf{v}_{opt} = \mathbf{A}^{\dagger}\widehat{\mathbf{w}}_{\mathrm{A}} = [\mathbf{A}^{\mathrm{H}}\mathbf{A} + \beta\mathbf{I}]^{-1}\mathbf{A}^{\mathrm{H}}\widehat{\mathbf{w}}_{\mathrm{A}} \qquad \text{(A2.17)}$$

REFERENCES

1. Takeuchi, T. and P.A. Nelson, *Optimal source distribution for binaural synthesis over loudspeakers.* J Acoust Soc Am, 2002. **112**(6): p. 2786-97.
2. Takeuchi, T., M. Teschl, and P.A. Nelson, *Objective and subjective evaluation of the optimal source distribution for virtual acoustic imaging.* Journal of the Audio Engineering Society 2007. **55**(11): p. 981-987.
3. Morgan, D.G., T. Takeuchi, and K.R. Holland, *Off-axis cross-talk cancellation evaluation of 2-channel and 3-channel OPSODIS soundbars*, in *Reproduced Sound 2016.* 2016, Proceedings of the Institute of Acoustics: Southampton.
4. Haines, L.A.T., T. Takeuchi, and K.R. Holland, *Investigating multiple off-axis listening positions of an OPSODIS sound bar*, in *Reproduced Sound 2017 Nottingham.* 2017, Proceedings of the Institute of Acoustics.
5. Yairi, M., et al., *Off-axis cross-talk cancellation evaluation of Optimal Source Distribution*, in *Institute of Electronics Information and Communication Engineers (IEICE) Japan EA ASJ-H ASJ-AA.* 2018: Hokkaido University. p. 115-120.
6. Takeuchi, T. and P.A. Nelson, *Extension of the Optimal Source Distribution for Binaural Sound Reproduction.* Acta Acustica united with Acustica, 2008. **94**(6): p. 981-987.
7. Yairi, M., et al., *Binaural reproduction capability for mutiple off-axis listeners based on the 3-channel optimal source distribution principle*, in *23rd International Congress on Acoustics.* 2019: Aachen, Germany.
8. Nelson, P.A. and S.J. Elliott, *Active Control of Sound.* 1992, London: Academic Press.
9. Golub, G. and C. Van Loan, *Matrix Computations.* 1996, London: The John Hopkins University Press.
10. Olivieri, F., et al., *Comparison of strategies for accurate reproduction of a target signal with compact arrays of loudspeakers for the generation of private sound and silence.* Journal of the Audio Engineering Society, 2016. **64**(11): p. 905-917.
11. Couturier, P., *MSc Dissertation, Institute of Sound and Vibration Research.* 2019, University of Southampton.