LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

P.A. Nelson and S.J. Elliott

Institute of Sound and Vibration Research, University of Southampton, Southampton SO9 5NH, England.

1. INTRODUCTION

The purpose of this brief note is to outline a scheme for producing better approximations to the perfect reproduction of pre-recorded acoustic signals than are produced by currently available sound reproduction systems. The basic description contained in this note is applied to stereophonic sound reproduction whereby signals are recorded at two points in space in the existing sound field (at, for example, the ears of a dummy head). When these signals are replayed via two loudspeakers in a listening room the initial signals are imperfectly reproduced at the ears of a listener. With the recent advances in sound recording and reproduction technology the imperfections in the reproduction now arise from three main sources:

(1) The signal played via the right channel is reproduced at both the right and the left ears of the listener. Similarly, for the signal played via the left channel.

(2) The acoustic response of the listening room provides a reverberant field which is in addition to the reverberant field of the space in which the existing recordings were made.

(3) The frequency response of the loudspeakers used for reproduction is imperfect.

The objective of the sound reproduction system described here has been assumed to be the "perfect" reproduction of the recorded signals at the listener's ears, i.e., the signals recorded at two points in the recording space are reproduced exactly at two points in the listening space. In order to compencate for the above three factors it is necessary to introduce inverse filters which act on the inputs to the loudspeakers used for reproduction which will compensate for both the loudspeaker response and the room response. Initial attempts to design such inverse filters has been reported by Farnsworth *et al* [1] and Clarkson *et al* [2] in the single channel case. In the work reported here the inverse filtering problem is dealt with in the two channel case and, in addition, the formulation includes the effective cancellation of the undesirable "acoustical cross talk" provided by unwanted transmission from, for example, the right channel to the left ear and *vice versa*. These three problems are thus dealt with in a single formulation which involves the use of a matrix of digital filters which operates on the recorded signals prior to their transmission via the loudspeaker system.

Previous attempts have been made to employ similar techniques using analogue electronics such as those proposed by Atal and Schroeder [3], Damaske [4] and, more recently, by Sakamoto *et al* [5,6]. In the formulation presented here, however, the use of digital techniques is envisaged and the problem can be analysed using a conventional least squares approach. Two approaches to the problem are presented: one involving the design of the invese filter matrix via a "deterministic" least squares method and the other involving the use of a "statistical" least squares technique. The advantage of the second approach is that it can be made adaptive and offers the possibility of computing the filter matrix *in situ* in a given listening space. In this case the proposed system consists of a simple addition to existing stereophonic sound reproduction equipment. In simple terms the additional filter matrix can be housed in a "black box" containing the necessary micro- processors which operate on the two output signals produced by existing equipment. The black box produces two modified outputs for transmission via the loudspeaker system of the existing equipment. Additional inputs to the black box are provided by two (or more) microphones which may be placed at the two (or more) positions in the listening space at which the optimal sound reproduction is required. These microphones may be removed after adaptation of the necessary digital filters.

In addition, the approach presented has been generalised such that it can be applied to any recording technique involving sensing the sound field with any number or type of transducers. It is thus not only applicable to dummy head or coincident microphone techniques but also more recent approaches involving the effective detection of the pressure and three components of particle velocity in the sound field. In either case the approach presented enables the deduction of the inverse filter matrix enabling the production of the best approximation (in the least squares sense) to the recorded qualities in the listening space.

2. THE INVERSE FILTER MATRIX FOR STEREOPHONIC SOUND REPRODUCTION

The basic problem is illustrated in Figure 1. The recorded signals x_1 and x_2 have been produced at, for example, the ears of a dummy head. The objective is to produce precisely the same signals at the ears of a listener. These signals are represented by the outputs $\hat{d}_1$ and $\hat{d}_2$ of microphones placed at, or close to, the ears of the listener. Figure 2 shows the transmission system in block diagram form, where the transfer functions $C_{\ell m}$ represent the transmission paths from the loudspeakers to the ears of the listener. For perfect sound reproduction it is assumed that we require $\hat{d}_1 = x_1$ and $\hat{d}_2 = x_2$. In order to achieve this in principle we must operate on the inputs x_1 and x_2 with a matrix of filters having elements H_{mk} as shown in Figure 3. The presence of the diagonal elements H_{21} and H_{12} can be thought of as providing the necessary cancellation of the cross talk signals present in the listening room as a result of the diagonal elements C_{21} and C_{12}. In order to find the necessary transfer function matrix $\underline{H}$ we can work in the frequency domain and define the vector of signals $\hat{\underline{d}}$ as

LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

$$\underline{\hat{d}} = \begin{bmatrix}\hat{d}_1\\ \hat{d}_2\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \underline{C}\,\underline{y} \qquad (1)$$

where the vector of signals $\underline{y}$ is defined as

$$\underline{y} = \begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \underline{H}\,\underline{x} \qquad (2)$$

Combination of these equations yields

$$\underline{\hat{d}} = \underline{CHx} \qquad (3)$$

Thus perfect sound reproduction, i.e., such that $\underline{\hat{d}} = \underline{x}$, requires $\underline{CH} = \underline{I}$ or $\underline{H} = \underline{C}^{-1}$. Thus in frequency domain terms the necessary filter matrix $\underline{H}$ is simply the inverse of the matrix $\underline{C}$. However, we cannot compute an exact inverse of the matrix $\underline{C}$ due to the presence of non-minimum phase components in those transfer functions (it is well known, for example, that the room acoustic transfer function contains significant non-minimum phase elements [7]). It is therefore necessary to take a least squares approach to the design of the inverse filter matrix and in order to do this it helps greatly if the block diagram representation of the composite system shown in Figure 3 can be rearranged. The necessary rearrangement can be explained simply using a little matrix algebra. Equation (3) can be expanded in full as

$$\begin{bmatrix}\hat{d}_1\\ \hat{d}_2\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{bmatrix}\begin{bmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} \qquad (4)$$

and by expanding the product of the matrix $\underline{C}$ with the matrix $\underline{H}$ this can be written in the form

$$\begin{bmatrix}\hat{d}_1\\ \hat{d}_2\end{bmatrix} = \begin{bmatrix}(C_{11}H_{11} + C_{12}H_{21}) & (C_{11}H_{12} + C_{12}H_{22})\\ (C_{21}H_{11} + C_{22}H_{21}) & (C_{21}H_{12} + C_{22}H_{22})\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} \qquad (5)$$

It is also evident that this product can be written as

$$\begin{bmatrix}\hat{d}_1\\ \hat{d}_2\end{bmatrix} = \begin{bmatrix}x_1C_{11} & x_1C_{12} & x_2C_{11} & x_2C_{12}\\ x_1C_{21} & x_1C_{22} & x_2C_{21} & x_2C_{22}\end{bmatrix}\begin{bmatrix}H_{11}\\ H_{21}\\ H_{12}\\ H_{22}\end{bmatrix} \qquad (6)$$

LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

where a vector representation of the filter matrix $\underline{H}$ has now been adopted. If we now define the signals $x_k C_{\ell m}$ as the "filtered reference signals" $r_{\ell mk}$ this equation can be further reduced to

$$\begin{bmatrix} \hat{d}_1 \\ \hat{d}_2 \end{bmatrix} = \begin{bmatrix} r_{111} & r_{121} & r_{112} & r_{122} \\ r_{211} & r_{221} & r_{212} & r_{222} \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{21} \\ H_{12} \\ H_{22} \end{bmatrix} \qquad (7)$$

The corresponding block diagram of the rearranged system is shown in Figure 4, where the input signals x_1 and x_2 are first passed through the relevant elements of the transfer function matrix $\underline{C}$ before being passed through the elements of the transfer function matrix $\underline{H}$. This approach is thus essentially that taken by Elliott *et al* [7] in deriving the stochastic gradient algorithm used in problems of active control of periodic noise and vibration, and has been extended for use with multiple input signals by Nelson *et al* [8]. This block diagram arrangement can now be used to determine the optimal finite impulse response filters comprising the matrix $\underline{H}$. Two alternative least squares approaches can be taken to determine the filters which give the "best" transmission properties of the net system.

3. DETERMINISTIC LEAST SQUARES FILTER DESIGN

In this case we determine the filters comprising $\underline{H}$ which give the best least squares fit to a desired system impulse response. Firstly note that this requires an exact knowledge of the transfer functions comprising $\underline{C}$. Now observe that if the inputs x_1 and x_2 are both unit impulses (at time index n = 0 in discrete time) then the filtered reference signals $r(n)$ will be given exactly by the impulse responses of the relevant transfer functions in $\underline{C}$. Thus the net impulse response of the system (the outputs $\hat{d}_1$ and $\hat{d}_2$) are comprised of convolutions of the form described by the matrix product

$$\begin{bmatrix} c(0) & 0 & 0 & 0 \\ c(1) & c(0) & 0 & 0 \\ c(2) & c(1) & c(0) & 0 \\ c(3) & c(2) & c(1) & 0 \\ . & c(3) & c(2) & c(0) \\ . & . & c(3) & c(1) \\ c(J-1) & . & . & c(2) \\ 0 & c(J-1) & . & c(3) \\ 0 & 0 & c(J-1) & . \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & c(J-1) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ h(2) \\ h(3) \\ . \\ . \\ . \\ h(I-1) \end{bmatrix} \qquad (8)$$

where the duration of the impulse responses $c(j)$ is assumed to be J samples and the impulse response of the compensating filters $h(i)$ is assumed to be of duration I samples. Note it is also implicitly assumed that both $c(j)$ and $h(i)$ are causal impulse responses. It therefore follows from the rearranged block diagram as expressed in equation (7) that the net output sequences $\hat{d}_1(n)$ and $\hat{d}_2(n)$ due to unit impulses at the inputs at n = 0 can thus be written as a composite sequence vector defined by

$$
\begin{bmatrix} \hat{d}_1(0) \\ \hat{d}_1(1) \\ \hat{d}_1(2) \\ \vdots \\ \hat{d}_1(I+J) \\ \hline \hat{d}_2(0) \\ \hat{d}_2(1) \\ \hat{d}_2(2) \\ \vdots \\ \hat{d}_2(I+J) \end{bmatrix}
=
\left[\begin{array}{cc|cc|cc|cc}
c_{11}(0) & 0 & c_{12}(0) & 0 & c_{11}(0) & 0 & c_{12}(0) & 0 \\
c_{11}(1) & c_{11}(0) & c_{12}(1) & c_{12}(0) & c_{11}(1) & c_{11}(0) & c_{12}(1) & c_{12}(0) \\
\vdots & c_{11}(1) & \vdots & c_{12}(1) & \vdots & c_{11}(1) & \vdots & c_{12}(1) \\
c_{11}(J-1) & \vdots & c_{12}(J-1) & \vdots & c_{11}(J-1) & \vdots & c_{12}(J-1) & \vdots \\
0 & c_{11}(J-1) & 0 & c_{12}(J-1) & 0 & c_{11}(J-1) & 0 & c_{12}(J-1) \\
\hline
c_{21}(0) & 0 & c_{22}(0) & 0 & c_{21}(0) & 0 & c_{22}(0) & 0 \\
c_{21}(1) & c_{21}(0) & c_{22}(1) & c_{22}(0) & c_{21}(1) & c_{21}(0) & c_{22}(1) & c_{22}(0) \\
\cdot & c_{21}(1) & \cdot & c_{22}(1) & \cdot & c_{21}(1) & \cdot & c_{22}(1) \\
c_{21}(J-1) & \vdots & c_{22}(J-1) & \vdots & c_{21}(J-1) & \vdots & c_{22}(J-1) & \vdots \\
0 & c_{21}(J-1) & 0 & c_{22}(J-1) & 0 & c_{21}(J-1) & 0 & c_{22}(J-1)
\end{array}\right]
\begin{bmatrix} h_{11}(0) \\ \vdots \\ h_{11}(I-1) \\ \hline h_{21}(0) \\ \vdots \\ h_{21}(I-1) \\ \hline h_{12}(0) \\ \vdots \\ h_{12}(I-1) \\ \hline h_{22}(0) \\ \vdots \\ h_{22}(I-1) \end{bmatrix}
\qquad (9)
$$

This equation can be written compactly in matrix form as

$$\underline{\hat{d}}_s = \underline{C}_s \underline{h} \qquad (10)$$

where the subscript "s" has been used to denote sequence. We now require $\underline{\hat{d}}_s$ to be a least squares approximation to a desired impulse $\underline{d}$. This desired response could be, for example, simply a delayed version of the input $\underline{x}(n)$. This will allow the incorporation of an appropriate "modelling delay" in the design of the inverse filter and therefore the reduction of the minimum mean square error associated with the least squares estimate (see Widrow and Stearns [10]). Thus, for example, if a delay at n = 2 is required, the vector $\underline{d}$ is given by

$$\underline{d}^T = [0\ 0\ 1\ 0\ \ldots\ 0\quad 0\ 0\ 1\ 0\ \ldots\ 0] \qquad (11)$$

It should also be noted however that any arbitrary required definition of $\underline{d}$

could be adopted and in principle a desired reverberation of the room could be incorporated into the formulation.

We thus seek to minimise the sum of the squared errors resulting from the difference between the actual system impulse response and this desired impulse response. This cost function can be defined as

$$I = (\underline{d} - \underline{\hat{d}}_s)^T(\underline{d} - \underline{\hat{d}}_s) \tag{12}$$

which, using equation (9), can be written as

$$I = (\underline{d} - \underline{C}_s\underline{h})^T(\underline{d} - \underline{C}_s\underline{h}) \tag{13}$$

This can be expanded to give the quadratic form defined by

$$I = \underline{d}^T\underline{d} - 2\underline{d}^T\underline{C}_s\underline{h} + \underline{h}^T\underline{C}_s{}^T\underline{C}_s\underline{h} \tag{14}$$

This function is minimised by the optimal filter vector of filter coefficients defined by

$$\underline{h}_o = [\underline{C}_s{}^T\underline{C}_s]^{-1}\underline{C}_s{}^T\underline{d} \tag{15}$$

with a corresponding minimum mean square error which is given by

$$I_o = \underline{d}^T\underline{d} - \underline{d}^T\underline{C}_s\underline{h}_o \tag{16}$$

Thus the optimal filter vector h_o can be evaluated by direct inversion of the matrix $[\underline{C}_s{}^T\underline{C}_s]$. This matrix is block-Toeplitz in form and efficient algorithms exist for its inversion [11]. It should be noted that exact measurements of the element comprising $\underline{C}_s$, i.e., the transmission path impulse responses, are required for the computation.

4. STATISTICAL LEAST SQUARES FILTER DESIGN

In this approach the system is assumed to be supplied with a spectrally broad "training signal" via the inputs x_1 and x_2. The resulting outputs can be written (in discrete time) as

$$\hat{d}_1(n) = \underline{r}_1{}^T\underline{h}, \qquad \hat{d}_2(n) = \underline{r}_2{}^T\underline{h} \tag{17}$$

where the vectors of filtered reference signals are given by

$$\begin{aligned} \underline{r}_1{}^T &= [\underline{r}_{111}{}^T \quad \underline{r}_{121}{}^T \quad \underline{r}_{112}{}^T \quad \underline{r}_{122}{}^T] \\ \underline{r}_2{}^T &= [\underline{r}_{211}{}^T \quad \underline{r}_{221}{}^T \quad \underline{r}_{212}{}^T \quad \underline{r}_{222}{}^T] \end{aligned} \tag{18}$$

Each of the component vectors comprising these vectors of filtered reference signals are sequences of the form

$$\underline{r}_{\ell mk}{}^T = [r_{\ell mk}(n) \quad r_{\ell mk}(n-1) \quad r_{\ell mk}(n-2) \ \ldots \ r_{\ell mk}(n-I+1)] \tag{19}$$

The vector $\underline{h}$ is the composite tap weight vector as defined above in equation (9). The net output can thus be expressed as the composite vector

$$\hat{\underline{d}}(n) = \underline{R}\,\underline{h} \tag{20}$$

where the vector $\hat{\underline{d}}(n)$ and the matrix $\underline{R}$ are defined by

$$\hat{\underline{d}}(n) = \begin{bmatrix} \hat{d}_1(n) \\ \hat{d}_2(n) \end{bmatrix}, \qquad \underline{R} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \end{bmatrix} \tag{21}$$

We now seek the optimal filter vector to minimise the time averaged error between the actual and desired outputs. In this case the desired output is defined as the signal produced by passing the training signal through a filter having the desired impulse response. Again this could simply be a pure delay, Δ samples, or any other modified form of impulse response function. The net block diagram can thus be written in the form shown in Figure 5. In this case we seek to minimise the time averaged sum of squared errors defined by

$$J = E[(\underline{d}(n) - \hat{\underline{d}}(n))^T(\underline{d}(n) - \hat{\underline{d}}(n))] \tag{22}$$

where E denotes the expectation operator and $\underline{d}(n)$ is the vector defined by $\underline{d}^T(n) = [d_1(n)\ d_2(n)]$. Thus, J can be written as

$$J = E[(\underline{d}(n) - \underline{R}\,\underline{h})^T(\underline{d}(n) - \underline{R}\,\underline{h})] \tag{23}$$

which reduces to the quadratic form

$$J = E[\underline{d}^T(n)\underline{d}(n)] - 2E[\underline{d}^T(n)\underline{R}]\underline{h} + \underline{h}^T E[\underline{R}^T\underline{R}]\underline{h} \tag{24}$$

This function has the minimum defined by

$$\underline{h}_o = [E[\underline{R}^T\underline{R}]]^{-1}E[\underline{R}^T\underline{d}(n)] \tag{25}$$

and the corresponding minimum mean squared error given by

$$J_o = E[\underline{d}^T(n)\underline{d}(n)] - E[\underline{d}^T(n)\underline{R}]\underline{h}_o \tag{26}$$

The matrix $E[\underline{R}^T\underline{R}]$ is again of block-Toeplitz structure and can be inverted directly. An alternative technique is to use Elliott's stochastic gradient algorithm where one updates the coefficients of the composite filter vector $\underline{h}$ in accordance with

$$\underline{h}(n+1) = \underline{h}(n) - \alpha\underline{R}\,\underline{e}(n) \tag{27}$$

where α is the convergence coefficient of the algorithm and the instantaneous error vector is defined by

$$\underline{e}(n) = \underline{d}(n) - \hat{\underline{d}}(n) \qquad (28)$$

The generalisation of the use of this algorithm follows easily from the formulation presented here and the description of the algorithm presented in reference [12]. The advantage of this technique is that the implementation of the algorithm requires only an imperfect knowledge of the elements of the matrix $\underline{C}$ in order to derive the values of the filtered reference signals used in the coefficient update equation given by equation (27). In the single frequency case the accuracy with which the values of the filters $\underline{C}$ must be known is well established. More work, however, is required to establish the necessary requirements for the accuracy of the generation of the filtered reference signals in this case.

5. EXTENSION TO MULTIPLE CHANNELS

The above techniques are easily extended to deal with the case of multiple channels where, for example, the signals are recorded at K points in the recording space (producing signals x_k) and we wish to reproduce a best least squares approximation to these signals ($\hat{d}_\ell$) at L points in the listening space using M loudspeaker channels (with outputs y_m). In general it is assumed that L > M > K. Again we operate on the K recorded signals with an M × K matrix $\underline{H}$ of filters prior to their transmission via an L × M matrix $\underline{C}$ of transmission paths. This is illustrated in Figure 6. The formulation of this general problem is most easily accomplished by again using the transfer function reversal technique. The approach adopted here is identical to that presented by Nelson *et al* [9]. Thus in matrix form, again working in the frequency domain, we write the vector of M loudspeaker signals and the vector of L reproduced signals as

$$\underline{y} = \underline{H}\,\underline{x}, \qquad \hat{\underline{d}} = \underline{C}\,\underline{y} \qquad (29)$$

If these are now written in terms of the columns $\underline{h}_k$ of $\underline{H}$ and the rows $\underline{c}_\ell$ of $\underline{C}$ then we have

$$\underline{y} = [\underline{h}_1\ \underline{h}_2\ \ldots\ \underline{h}_K]\underline{x}, \qquad \hat{\underline{d}} = \begin{bmatrix} \underline{c}_1^T \\ \underline{c}_2^T \\ \vdots \\ \underline{c}_L^T \end{bmatrix}\underline{y} \qquad (30)$$

The signal produced at the ℓ'th point in the listening space is thus given by

$$\hat{d}_\ell = \underline{c}_\ell^T\underline{y} = \underline{c}_\ell^T(\underline{h}_1x_1 + \underline{h}_2x_2\ \ldots\ \underline{h}_Kx_K) \qquad (31)$$

Transposing this scalar then yields

LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

$$\hat{d}_\ell = \underline{h}_1^T \underline{c}_\ell x_1 + \underline{h}_2^T \underline{c}_\ell x_2 \ldots \underline{h}_K^T \underline{c}_\ell x_K \qquad (32)$$

This then demonstrates that the equivalent "reversed transfer function" block diagram can be expressed as shown in Figure 7. In discrete time, the signal at the ℓ'th point is the listening space can thus be expressed in terms of the convolution

$$\hat{d}_\ell(n) = \underline{r}_\ell^T \underline{h} \qquad (33)$$

where the composite tap weight vector $\underline{h}$ and the ℓ'th filtered reference signal vector $\underline{r}_\ell$ are defined by

$$\underline{r}_\ell^T = [\underline{r}_{\ell 11}^T\ \underline{r}_{\ell 21}^T \ldots \underline{r}_{\ell M1}^T\ |\underline{r}_{\ell 12}^T\ \underline{r}_{\ell 22}^T \ldots \underline{r}_{\ell M2}^T| \ldots |\underline{r}_{\ell 1K}^T\ \underline{r}_{\ell 2K}^T \ldots \underline{r}_{\ell MK}^T]$$

$$(34)$$

$$\underline{h} = [\underline{h}_{11}^T\ \underline{h}_{21}^T \ldots \underline{h}_{M1}^T\ |\underline{h}_{12}^T\ \underline{h}_{22}^T \ldots \underline{h}_{M2}^T| \ldots |\underline{h}_{1K}^T\ \underline{h}_{2K}^T \ldots \underline{h}_{MK}^T]$$

where each of the component vectors are given by the sequences

$$\underline{r}_{\ell mk}^T = [r_{\ell mk}(n)\quad r_{\ell mk}(n-1) \ldots r_{\ell mk}(n-I+1)]$$

$$(35)$$

$$\underline{h}_{mk} = [h_{mk}(0)\quad h_{mk}(1) \ldots h_{mk}(I-1)]$$

The net vector of discrete time signals produced at the ℓ points in the listening space can now be written as

$$\hat{\underline{d}}(n) = \underline{R}\ \underline{h} \qquad (36)$$

where the vector $\hat{\underline{d}}(n)$ and the matrix $\underline{R}$ are now defined as

$$\hat{\underline{d}}(n) = \begin{bmatrix} \hat{d}_1(n) \\ \hat{d}_2(n) \\ \vdots \\ \hat{d}_L(n) \end{bmatrix}, \qquad \underline{R} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_L^T \end{bmatrix} \qquad (37)$$

The optimal composite tap weight vector which minimises a cost function of the form

$$J = E[(\underline{d}(n) - \hat{\underline{d}}(n))^T(\underline{d}(n) - \hat{\underline{d}}(n))] \qquad (38)$$

(where $\underline{d}(n)$ is now the L-vector of desired impulse responses) is now deduced by following exactly the analysis presented in equations (22) to (26) above. The adaptive coefficient update equation (27) also remains applicable in this general case.

6. CONCLUDING REMARKS

The problem has been addressed of designing the inverse filter matrix necessary to enhance the sound reproduction provided by stereophic sound reproduction equipment. The scheme proposed has the advantage of providing a simple additional item of equipment which can be used with existing stereophonic reproduction systems. The principle objective of the system has been assumed to be the production of a "closest possible approximation" to the exact reproduction of the two recorded signals at two points in the listening space. This effective equalisation is likely to be highly localised spatially [1] and may well be only practically effective at the low frequency end of the audio frequency range. It may also be more effective to attempt to spread the "zone of equalisation" by sensing the sound field in the listening space using more microphones than loudspeaker channels (in which case the analysis of Section 5 is applicable). The psycho-acoustical effects of the implementation of the scheme have not been considered and in particular its effect on stereophonic image localisation has yet to be quantified.

REFERENCES

1. K.D. FARNSWORTH, P.A. NELSON and S.J. ELLIOTT 1985 *Proceedings of the Institute of Acoustics, Autumn Conference*: Reproduced Sound (Windermere). Equalisation of room acoustic responses over spatially distributed regions.

2. P.M. CLARKSON, J. MOURJOPOULOS and J.K. HAMMOND 1985 *Journal of the Audio Engineering Society* 33, 127-132. Spectral, phase and transient equalisation for audio systems.

3. B.S. ATAL and M.R. SCHROEDER 1962 *US Patent* 3,236,949. Apparent sound source translator.

4. P. DAMASKE 1971 *Journal of the Acoustical Society of America* 50. Head-related two-channel stereophony with loudspeaker reproduction.

5. N. SAKAMOTO, T. GOTOH, T. KOGURE and M. SHIMBO 1981 *Journal of the Audio Engineering Society* 29, 794-799. Controlling sound-image localisation in stereophonic sound reproduction, Part I.

6. N. SAKAMOTO, T. GOTOH, T. KOGURE, M. SHIMBO and A.H. CLEGG 1981 *Journal of the Audio Engineering Society* 30, 719-721. Controlling sound-image localisation in stereophonic sound reproduction, Part II.

7. S.T. NEELY and J.B. ALLEN 1979 *Journal of the Acoustical Society of America* 66, 165-169. Invertibility of a room impulse response.

8. S.J. ELLIOTT, I.M. STOTHERS and P.A. NELSON 1987 *IEEE Transactions on Acoustics, Speech and Signal Processing* 35, 1423-1434. A multiple error LMS algorithm and its application to the active control of sound and vibration.

9. P.A. NELSON, J.K. HAMMOND and S.J. ELLIOTT 1988 *Proceedings of the Institute of Acoustics* 10, 589-611. Linear least squares estimation problems in the active control of stationary random sound fields.

10. B. WIDROW and S.D. STEARNS 1985 *Adaptive Signal Processing* (see Chapter 10). Prentice-Hall, Englewood Cliffs.

11. S.L. MARPLE 1987 *Digital Spectral Analysis with Applications* (see Chapter 15). Prentice-Hall, Englewood Cliffs.

12. S.J. ELLIOTT and P.A. NELSON 1987 *ISVR Memorandum* no. 679.

LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

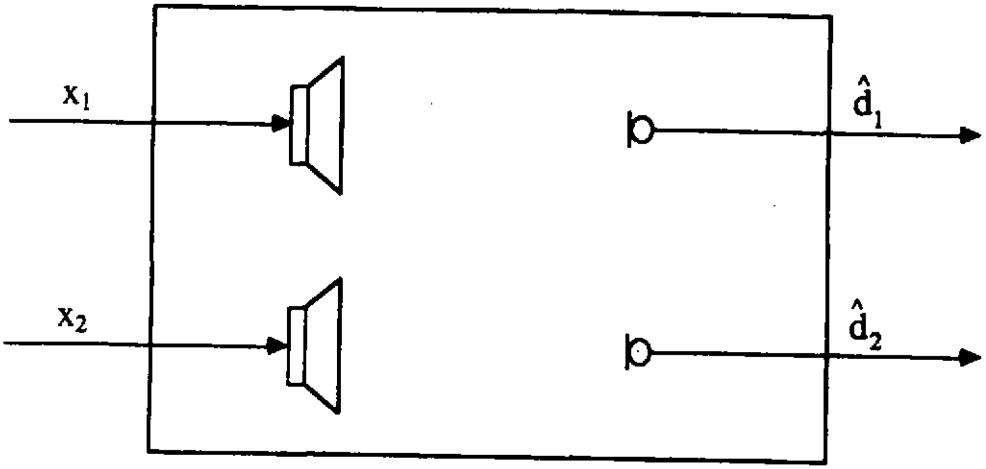

Figure 1. A stereophonic sound reproduction system ; the objective is to produce signals at the two microphones which are exactly the signals input to the loudspeakers i.e. $\hat{d}_1=\hat{x}_1$ and $\hat{d}_2=\hat{x}_2$ for "perfect" sound reproduction at the two microphone positions.

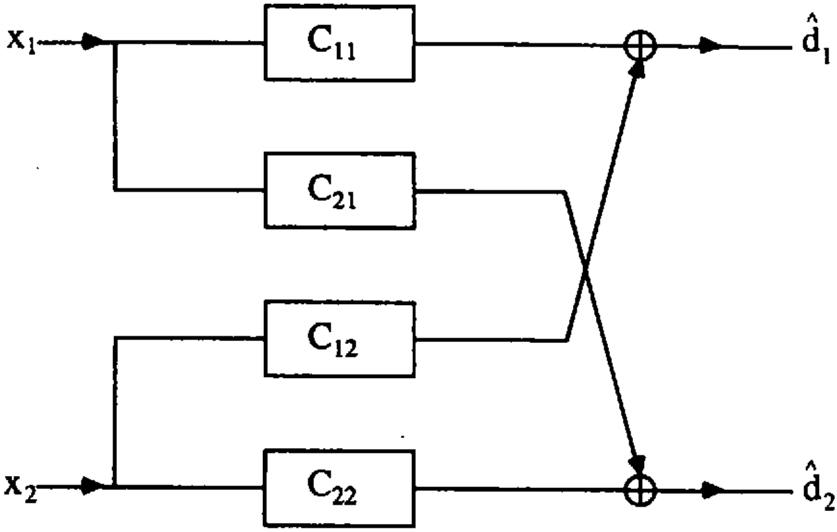

Figure 2. Block diagram representation of a stereophonic sound reproduction system prior to the introduction of compensating filters.

LEAST SQUARES APPROXIMATIONS TO EXACT MULTIPLE POINT SOUND REPRODUCTION

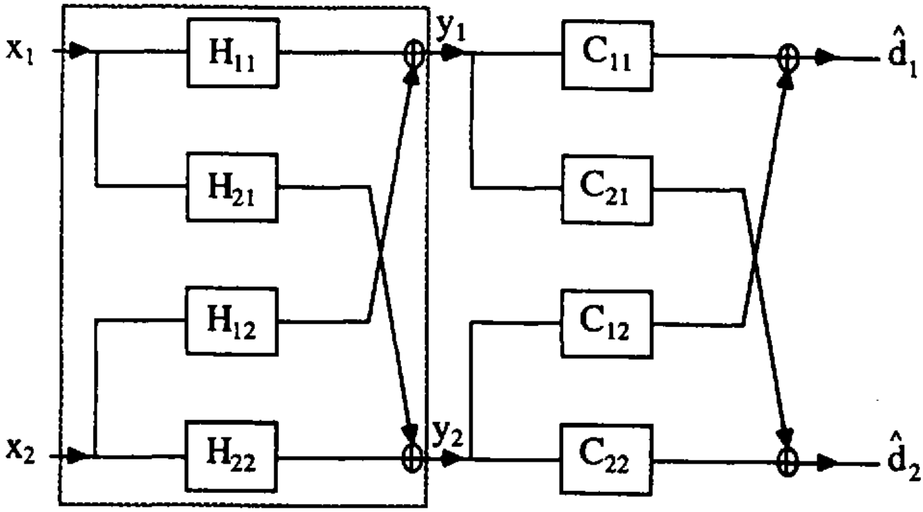

Figure 3. The block diagram of the reproduction system when the matrix **H** of inverse filters is introduced ; x_1 and x_2 are the recorded signals and y_1 and y_2 are the new input signals to the loudspeakers.

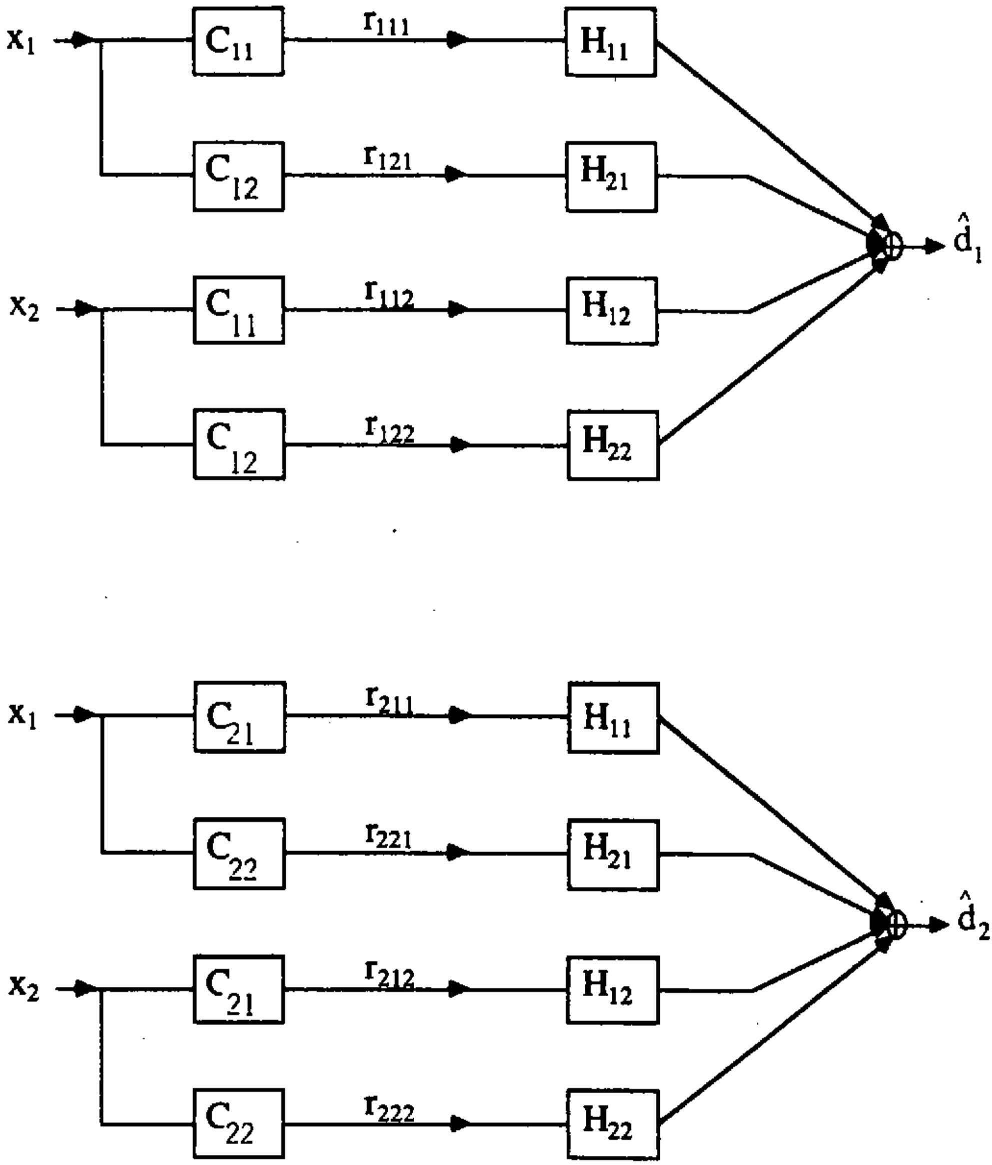

Figure 4. Equivalent block diagram when the order of operation of the elements of H and C are reversed.

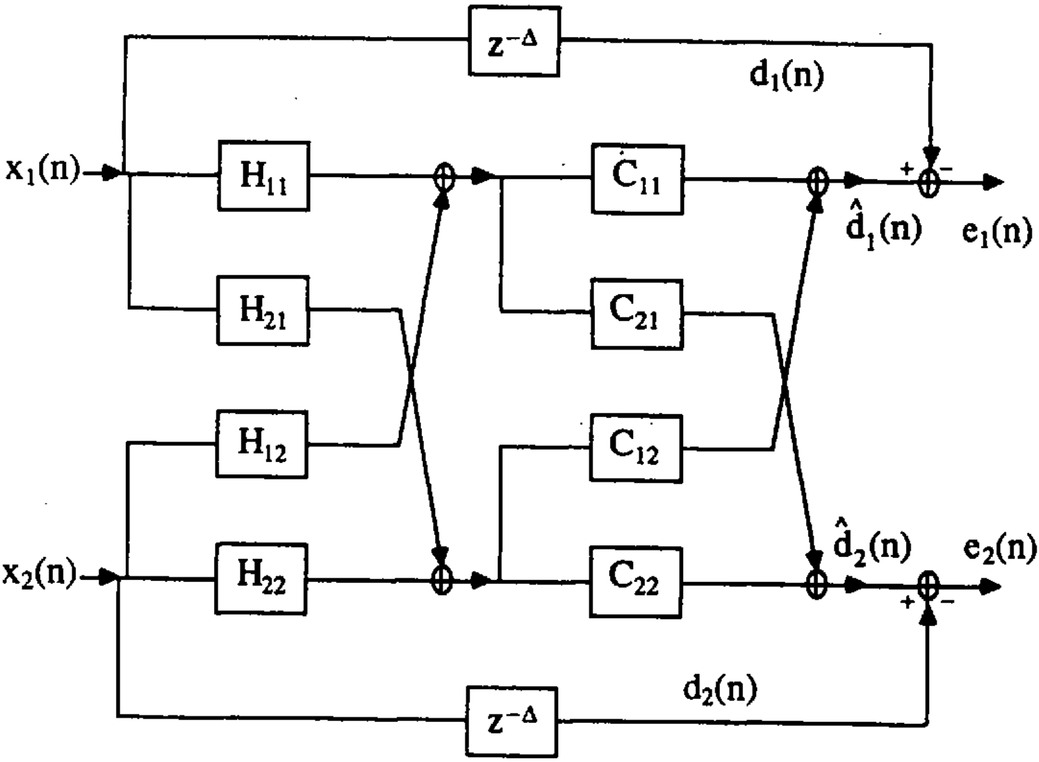

Figure 5. Block diagram of the compensated sound reproduction system when a modelling delay of Δ samples is included in order to allow for non-minimum phase components associated with the transfer functions in the matrix $\underline{C}$. The desired output signals $d_1(n)$ and $d_2(n)$ are now simply delayed versions of the recorded signals $x_1(n)$ and $x_2(n)$.

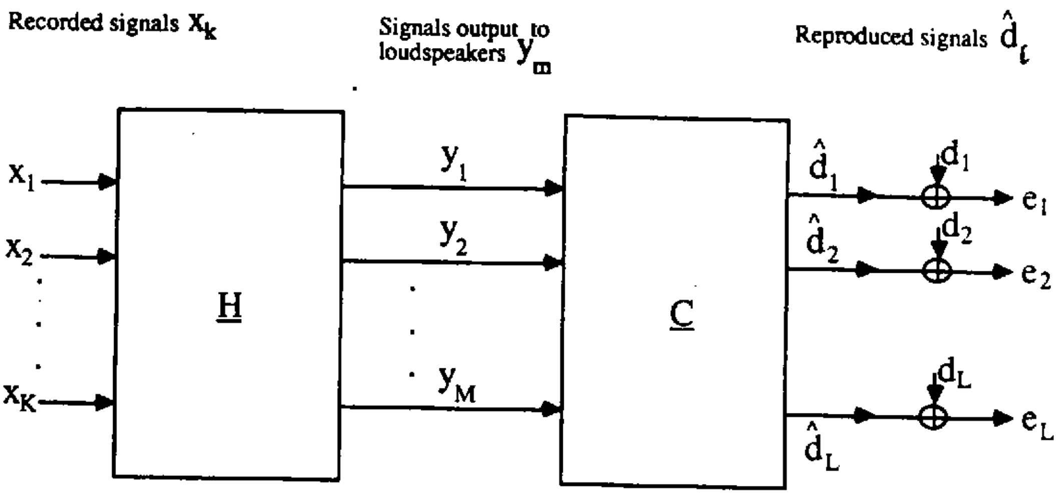

Figure 6. Generalisation of the scheme to deal with signals recorded at K points in space and the best "least sqares" reproduction of these signals at L points in a listening space, where in general L≥K.

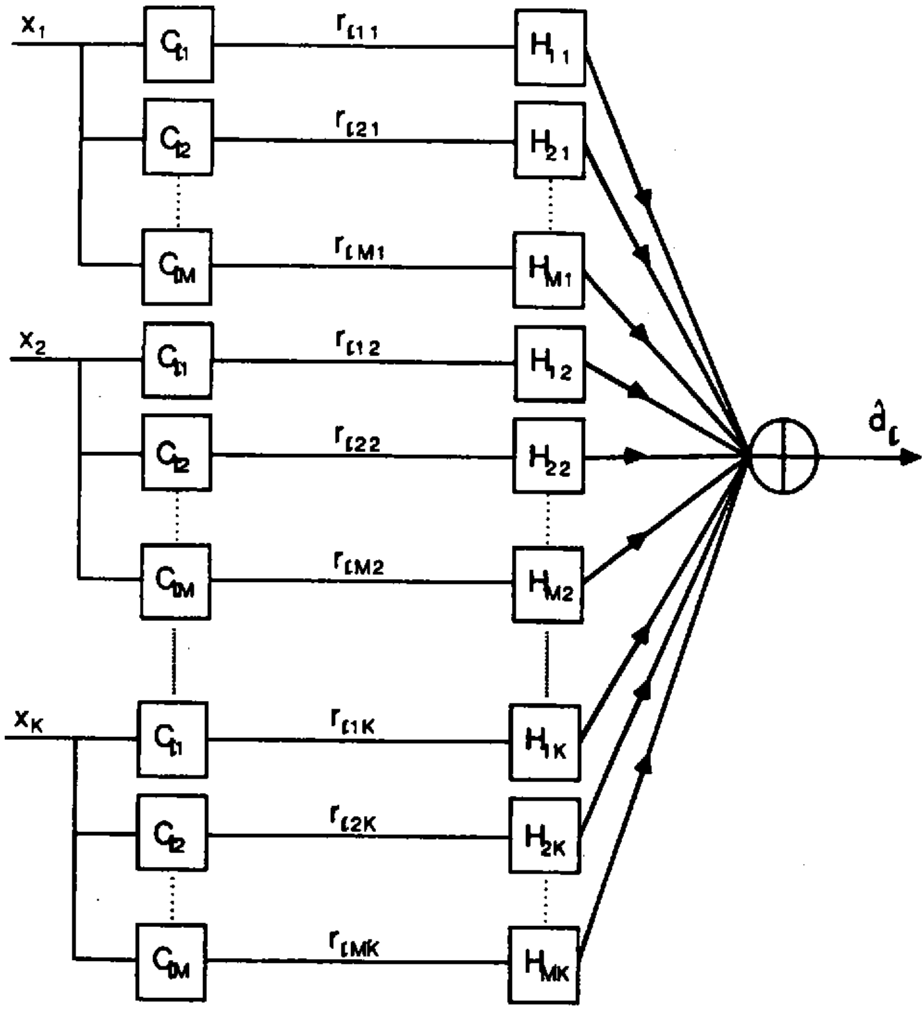

Figure 7. Equivalent block diagram representation of the generation of the signal at the ℓ'th point in the listening space.