FIFTH INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
DECEMBER 15-18, 1997
ADELAIDE, SOUTH AUSTRALIA

ACTIVE EQUALISATION OF THE SOUND FIELD IN AN EXTENDED REGION OF A ROOM

Arturo Orozco-Santillán
Department of Acoustic Technology, Technical University of Denmark
Building 352, Dk-2800 Lyngby, Denmark

ABSTRACT

A theoretical study of spatial sound equalisation in an extended region in a room has been carried out. The purpose is to reproduce sound without significant spatial fluctuations and to minimise the spectral colouration at low frequencies in a specified region of the room. The problem is first studied by means of an idealised frequency domain model. The analysis is based on the calculation of the complex source strengths that minimise the difference between the actual sound pressure and the desired sound pressure in the listening area. Results in relation to the position of the sources, the frequency range, and the size and location of the listening area are presented. However, the frequency-domain approach results in non-causal impulse responses that can be realised only at the expense of a delay. Therefore, this analysis is supplemented with a study of the equalisation carried out in the time domain. Here, a solution constrained to causality is determined. The duration of the impulse response should be minimised as well in order to avoid echos, which are undesirable in the reproduction of transient signals. This causality-constrained equalisation is compared with the optimal solution obtained in the frequency domain.

INTRODUCTION

When broadband sound is reproduced in a room, it is modified because of the resonances of the enclosure. At low frequencies the acoustic response of a normal listening room is dominated by distinct resonances and their associated mode shape functions. This means that music, speech, and other signals reproduced in a room by a conventional audio system will undergo a dramatic spectral colouration because of the room response. Since this response implies a linear distortion of the sound, this usually represents an undesirable effect. The number of acoustic resonances in a given frequency band (the modal density) increases with the square of the frequency, which means that the spectral colouration of the reproduced signal is not as severe a problem at higher frequencies.

The equalisation of the frequency response at one or multiple single points in an enclosure has been studied by several authors [1-4]. However, the resulting area of equalisation is small and it is reduced when the frequency is increased. Thus, the system is very sensitive to changes in the position of the listener's head. In the present study the local room equalisation has been desired over an extended area, large enough to allow for movements of the listener.

The general objective of the work is to study to what extent reducing the effect of the acoustic response of the room is possible at very low frequencies in an extended region of the enclosure (the listening area). Thus, a flat frequency response is sought in this area, which is a finite part of the space inside the room rather than a single point. This equalisation is considered to be carried out by passing the available audio signal through digital filters.

EQUALISATION IN THE FREQUENCY DOMAIN

The complex sound pressure amplitude p in a lightly damped rectangular enclosure at the position $\boldsymbol{r}$ and steady harmonically excited by L pistons can be expanded into series of eigenfunctions Ψ_n as [5]

$$p(\boldsymbol{r}) = \sum_n \left(\frac{\rho}{V\Lambda_n} \frac{\omega}{2\xi_n k_n k - i\left(k_n^2 - k^2\right)} \sum_{l=1}^{L} \frac{q_l}{A_l} \int_{A_l} \Psi_n(s_l)\,\mathrm{d}a \right) \Psi_n(\boldsymbol{r}) = \sum_{l=1}^{L} Z_l(\boldsymbol{r}) q_l . \tag{1}$$

Here ω is the driving angular frequency; ρ is the ambient density; ξ_n is the damping ratio of the nth mode; Λ_n is a scaling factor that depends on whether the mode is one-, two-, or three-dimensional; k is the wave number; q_l is the strength of the lth piston and A_l its area; V is the volume of the enclosure. In addition $k_n^2 = (\pi n_x/l_x)^2 + (\pi n_y/l_y)^2 + (\pi n_z/l_z)^2$, where n_x, n_y, n_z are integers and l_x, l_y, l_z are the side lengths of the enclosure in the x, y, and z directions, respectively; $Z_l(\boldsymbol{r})$ is the transfer impedance from the lth source to the point $\boldsymbol{r}$.

The objective is to equalise the sound pressure in a listening area to a desired complex sound pressure p_d. It is accomplished by minimising a cost function J defined as [6]

$$J = \frac{1}{V_T} \int_V |p(\boldsymbol{r}) - p_d(\boldsymbol{r})|^2 dV , \tag{2}$$

where V_T is the volume of the listening area. In a practical situation this integral must be approximated by a sum at a finite number M of points (sensors) in the listening area as follows

$$\hat{J} = \frac{1}{M} \sum_{m=1}^{M} |p(\boldsymbol{r}_m) - p_d(\boldsymbol{r}_m)|^2 = \frac{1}{M} (\boldsymbol{Zq} - \boldsymbol{p}_d)^H (\boldsymbol{Zq} - \boldsymbol{p}_d) . \tag{3}$$

Here $\boldsymbol{r}_m$ is the position of the mth sensor, $\boldsymbol{p} = \boldsymbol{Zq}$ according to Eq. (1), and

$$\boldsymbol{Z} = \begin{bmatrix} Z_1(\boldsymbol{r}_1) & \cdots & Z_L(\boldsymbol{r}_1) \\ \vdots & & \vdots \\ Z_1(\boldsymbol{r}_M) & \cdots & Z_L(\boldsymbol{r}_M) \end{bmatrix}, \quad \boldsymbol{p} = \begin{bmatrix} p(\boldsymbol{r}_1) \\ \vdots \\ p(\boldsymbol{r}_M) \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} q(\boldsymbol{r}_1) \\ \vdots \\ q(\boldsymbol{r}_L) \end{bmatrix}, \quad \boldsymbol{p}_d = \begin{bmatrix} p_d(\boldsymbol{r}_1) \\ \vdots \\ p_d(\boldsymbol{r}_M) \end{bmatrix} . \tag{4}$$

Therefore, the estimated cost function can be reduced to a quadratic form as

$$\hat{J} = \boldsymbol{q}^H \hat{\boldsymbol{A}} \boldsymbol{q} + \boldsymbol{q}^H \hat{\boldsymbol{b}} + \hat{\boldsymbol{b}}^H \boldsymbol{q} + \hat{d} , \tag{5}$$

where $$\hat{\boldsymbol{A}} = (1/M)\boldsymbol{Z}^H\boldsymbol{Z} , \qquad \hat{\boldsymbol{b}} = (1/M)\boldsymbol{Z}^H\boldsymbol{p}_d , \qquad \hat{d} = -(1/M)\boldsymbol{p}_d^H\boldsymbol{p}_d . \tag{6}$$

Since Eq. (5) is a quadratic function of $\boldsymbol{q}$, the optimal values of the strength of the sources that minimise $\hat{J}$, and its minimum values are given, respectively, by [6]

$$\boldsymbol{q}_o = -\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{b}}, \qquad \hat{J}_o = \hat{d} - \hat{\boldsymbol{b}}^H\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{b}} . \tag{7}$$

Simulation Results

Here, the shown results were obtained with computer simulations with programmes developed in Matlab and running on a PC. A rectangular room with dimensions 2.5 m, 3 m, and 2 m, in the directions x, y, and z, respectively, was considered. A damping ratio equal to 0.1 for all the modes was assumed. In the modal summation, 3,050 modes were taken into account, which correspond to a natural frequency less than 1200 Hz. This number of modes was sufficient to obtain reliable results. Eight sound sources modelled as squared pistons with a side length of 0.1 m were used. The listening area had a cubic shape with its sides parallel to the sides of the room. It should be mentioned that several calculations of the optimal equalisation from Eq. (2) that were carried out resulted relatively time-consuming. It was due to the integration over the listening area for each mode and the limitations of the memory of the PC. Therefore, Eq. (3) had to be used. The calculations shown in this part were carried out with 100 sensors distributed in the listening area, which provided an adequate approximation of the optimal equalisation. The desired complex sound pressure in the listening area was set to 1+j0 Pa for all the simulations that were carried out. A value of $\hat{J}_0 = 0.05$ Pa^2 was chosen as the criterion for an adequate equalisation. With this value, the fluctuations of the obtained sound pressure from the desired one are less than ± 3 dB in almost the whole listening area. The error is bigger near the boundaries of this area, specially in the corners.

Position of the Sources. The effect of several source positions in the room was first examined. A listening area with a side length of 0.4 m and its centre fixed at the point (1.10,1.00,0.85) was assumed. Fig. 1 depicts six different sets of source positions used in the simulations, and the curves of the $\hat{J}_0$ as a function of the driving frequency corresponding to these sets are shown in Fig. 2. It can be seen that the best solution with eight sources was obtained when they were placed one in each corner of the room (set 5). As a reason for this result, it can be considered that in this case all the modes can be excited by each of the eight sources, and therefore, it is possible to have a better control of every mode. This idea can be supported by comparing with the result of the first set of the source positions. Here, the degree of the equalisation is the worst shown in Fig. 2, and in this case only the x-axial modes can be controlled more properly compared to the others. Furthermore, the result was better with the set 2, and it was improved with the set 4. Although all the modes can be excited in the last case, it is not possible to excite every mode with all the sources. For comparison, the result with sixteen sources is also shown in Fig. 2. It can be seen that the improvement is not significant although the number of sources has been double. In what follows, all results have been obtained with eight sources placed one in each corner of the room.

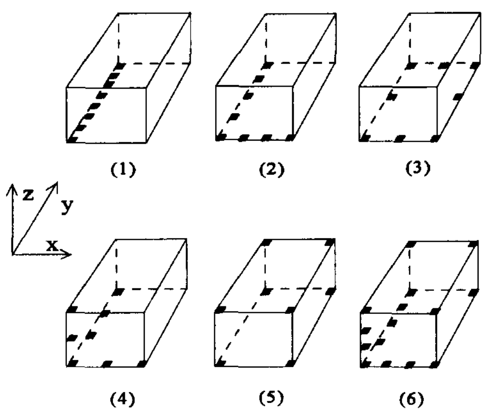


Fig. 1: Schematic diagram of the enclosure and different sets of sources positions examined in the equalisation problem.

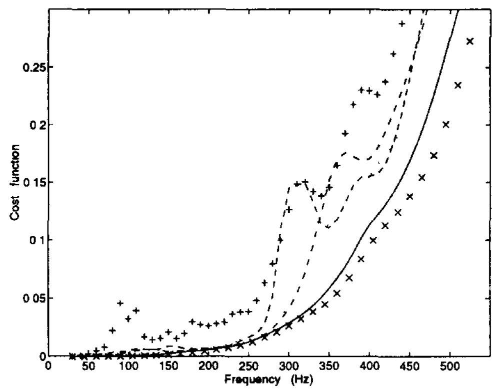


Fig. 2: Graph of $\hat{J}_0$ vs frequency corresponding to the sets of source positions that are shown in figure 1. ++ set 1, - - set 2, — · set 3, ··· set 4, — set 5, xx set 6.

Position of the Listening Area. It was found that the level of the equalisation decreases lightly when the target region is near any wall of the room. In a general way, the results can be set into three groups: (a) The listening area is in contact with two or three walls or near to them. Here the degree of the equalisation was the worst among the different positions examined. (b) The listening area is in contact or near one wall, and in this case the result was improved. (c) The listening area is relatively far from any wall of the room. With this situation, the best results were obtained. These effects are illustrated in Fig. 3. The side length of the listening area was fixed to 0.4 m. However, the differences in the performance among the three groups are relatively small in the frequency range where the cost function has a value less than 0.1, which corresponds to an adequate level of equalisation. Furthermore, this effect near the walls is less pronounced when the frequency is reduced. Consequently, if the listening area is farther from any wall of the enclosure than one quarter of the wavelength at the maximum frequency to be equalised, then the performance is practically independent of the position of this area. This is another advantage of placing the sources in each corner of the room. When the side of the listening area was less than 20 cm, the level of the equalisation was more sensitive to the position of this area for some values of the driving frequency, but the fluctuations in the cost function were not very serious.

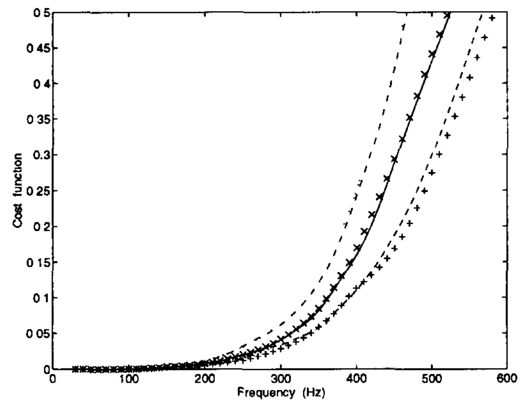


Fig. 3: Graph of $\hat{J}_0$ vs. frequency for different positions of the listening area. The positions of its centre are: ··· (0.2,0.2,0.2), —· (0.2,0.2,0.75), — (1.0,0.2,0.7), xx (0.2,1.8,0.75), ++ (1.1,1.0,0.85), -- (0.9,1.3,0.7).

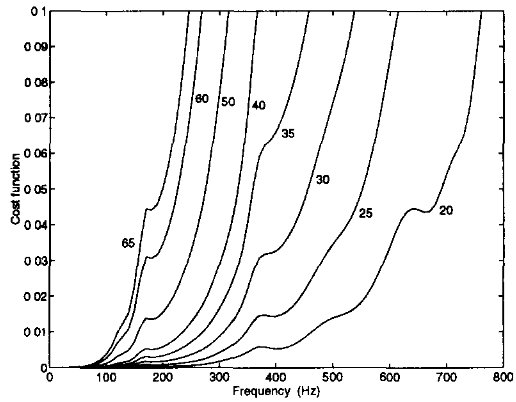


Fig. 4: Graph of $\hat{J}_0$ vs. frequency for different sizes of the listening area centred at (0.9,1.3,1.0). The numbers correspond to the side length of the listening area in cm.

Frequency Range and Size of the Listening Area. The relation between the size of the listening area and the maximum frequency that allows an acceptable level of equalisation was examined. The graphs of the cost function vs. frequency for different values of the side of the listening area are shown in Fig. 4. Here the position of the centre of the listening area was held fixed at (0.9,1.3,1.0) and its size length was varied. For each size of this area, the maximum driving frequency that allowed a value of $J_0 \leq 0.05$ Pa^2 was determined. The graph of the wavelength that corresponds to this maximum frequency vs. the side length of the listening area is depicted in Fig. 5. Here it was possible to fit a straight line in a least-squares sense; its slope was 2.58 and its *x*-intersection 0.005 m. The latter number is relatively small and it might be the result of errors in the estimation of the cost function.

When another value of the cost function is chosen instead of 0.05 Pa^2, the linear relation will be held with a different slope provided that this new value is small enough. It can be said that if the side of a cubic listening area is less than approximately one third of the smallest side of a rectangular room, then the maximum frequency that can be equalised with a good degree is inversely proportional to the side of the listening area.

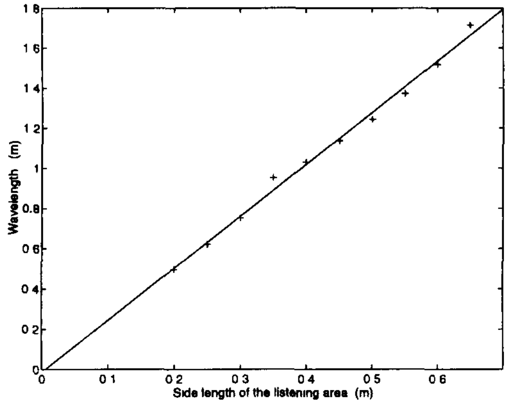


Fig. 5: Graph of the wavelength corresponding to $\hat{J}_o$ = 0.05 vs. the side length of the listening area centred at (0.9,1.3,1.0).

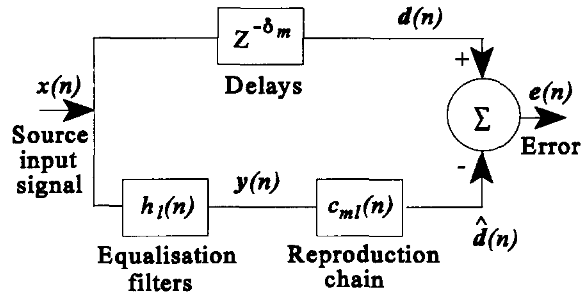


FIG. 6: Block diagram of the equalisation problem

OPTIMAL EQUALISATION IN THE TIME DOMAIN

The filters that produce the optimal equalisation in the frequency domain correspond to non-causal impulse responses, which can be implemented only at the expense of a delay. Since long delays are not acceptable for many purposes, a revised design criterion that takes account of causality and the duration of the impulse response of the equalisation filters has also been examined. The duration of these impulse responses should be minimised in order to avoid echos, which are undesirable in the reproduction of transient signals. For this case, the optimal equalisation was carried out in the time domain.

Fig. 6 depicts a block diagram of the equalisation problem in the time domain. This procedure has been discussed by Elliott and Nelson [1] for the case of one source and the equalisation at four single points. A similar implementation has been used by S. Laugesen [7] for noise cancellation in enclosures. Here M microphones are considered. The aim is to design L digital FIR filters with impulse responses $h_l(n)$, one for each sound source, such that the obtainable signal $\hat{d}_m(n)$ at the microphone m is the best approximation to the desired signal $d_m(n)$ at that microphone. In this case $d_m(n)$ is a delayed version (δ_m samples for the mth sensor) of the original input signal $x(n)$. In Fig. 6, $y(n)$ represents the vector of input signals to the L sound sources, and $c_{ml}(n)$ is the impulse response from the input of the lth source to the output of the mth sensor. Therefore, there are L x M of such impulse responses in the represented reproduction chain in Fig 6. It was assumed that these impulse responses can be modelled as FIR digital filters with J coefficients. In addition, the equalisation filters will have a FIR structure as well, with I coefficients.

The optimal solution is obtained by minimising a performance index defined by

$$\Gamma = E\left\{\sum_{m=1}^{M} |e_m(n)|^2\right\} = E\{e^T e\} \ , \tag{8}$$

where E represents the expectation operator and $e_m(n) = d_m(n) - \hat{d}_m(n)$.

If $y_l(n)$ is the input signal to the lth source, c_{mlj} is the coefficient j-1 of the reproduction chain from the lth source to the mth sensor, and a_{li} is the coefficient i-1 of the control filter corresponding to the lth source, then

$$y_l(n) = \sum_{i=0}^{I-1} a_{li} x(n-i), \qquad \hat{d}_m(n) = \sum_{L=1}^{L} \sum_{j=0}^{J-1} c_{mlj} y_l(n-j) \ . \tag{9}$$

It follows that the error signal $e_m(n)$ can be written as

$$e_m(n) = d_m(n) - \sum_{l=1}^{L} \sum_{j=0}^{J-1} a_{li} r_{lm}(n-i), \qquad \text{where} \qquad r_{lm}(n) = \sum_{j=0}^{J-1} c_{mlj} x(n-j) \ . \tag{10}$$

The error at the M microphones can be expressed as

$$\begin{bmatrix} e_1(n) \\ \vdots \\ e_M(n) \end{bmatrix} = \begin{bmatrix} d_1(n) \\ \vdots \\ d_M(n) \end{bmatrix} - \left[\begin{matrix} r_{11}(n) & \cdots & r_{1L}(n) \\ \vdots & & \vdots \\ r_{M1}(n) & \cdots & r_{ML}(n) \end{matrix} \ \middle| \ \cdots \ \middle| \ \begin{matrix} r_{11}(n-I+1) & \cdots & r_{1L}(n-I+1) \\ \vdots & & \vdots \\ r_{MI}(n-I+1) & \cdots & r_{ML}(n-I+1) \end{matrix} \right] \begin{bmatrix} \boldsymbol{a}_0 \\ \vdots \\ \boldsymbol{a}_{I-1} \end{bmatrix}, \tag{11}$$

where $\boldsymbol{a}_i = [a_{1i} \ ... \ a_{Li}]^T$. In a short form, Eq. (15) can be written as $\boldsymbol{e}(n) = \boldsymbol{d}(n) - \boldsymbol{Ra}$. Finally one can obtain the performance index as

$$\Gamma = \boldsymbol{a}^T E\{\boldsymbol{R}^T\boldsymbol{R}\}\boldsymbol{a} + \boldsymbol{a}^T E\{-\boldsymbol{R}^T\boldsymbol{d}(n)\} + E\{-[\boldsymbol{R}^T\boldsymbol{d}(n)]^T\}\boldsymbol{a} + E\{\boldsymbol{d}(n)^T\boldsymbol{d}(n)\} \ . \tag{12}$$

This performance index is a quadratic function of the coefficients in the equalisation filters, and therefore, it has a global minimum value. The optimal set of control coefficients is found to be

$$\boldsymbol{a}_{op} = \left(E\{\boldsymbol{R}^T\boldsymbol{R}\}\right)^{-1} E\{\boldsymbol{R}^T\boldsymbol{d}(n)\} \ . \tag{13}$$

Simulation Results

In the simulations described here, a source input signal with a delta function as the autocorrelation function was used. This is the worst case; therefore, a better result is expected if the signal to be reproduced is predictable. A listening area with a side length of 0.4 m and centred at the point (1.10,1.00,0.85) was considered. According to the results in the frequency domain, it was seen that an adequate optimisation with $\hat{J}_o \leq 0.06$ Pa^2 was possible with this listening area for a value of the driving frequency from zero up to 350 Hz. The same frequency interval was used here in order to compare the results in both domains. Thus, a sampling frequency of 700 Hz was assumed. The sound sources were modelled as a first-order analog high pass filter with one pole at 70 Hz to have a more realistic situation. 27 sensors uniformly distributed in the listening area were used, which produced an adequate estimation of the equalisation.

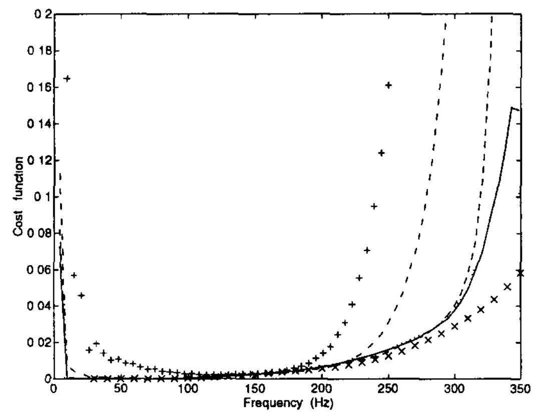
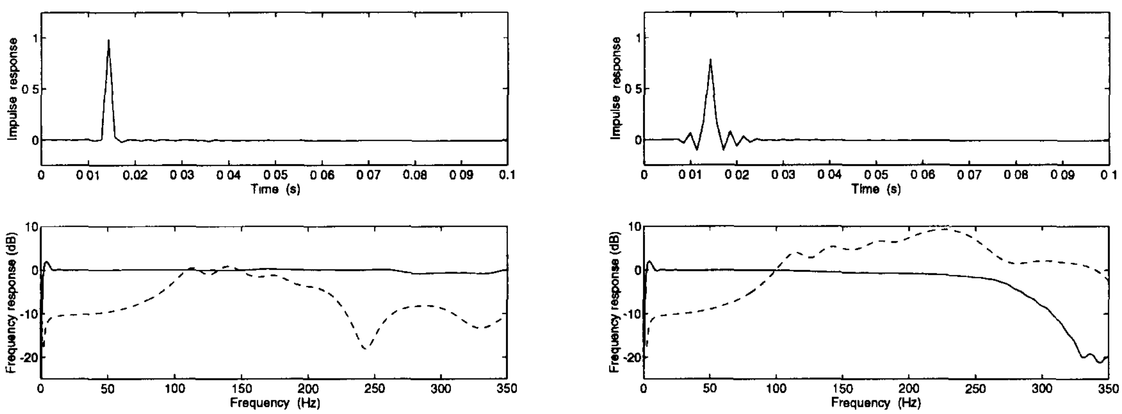


Fig. 7: $\hat{J}_o$ obtained from the optimisation in the time domain vs. frequency for different delay values: ++ 5 samples, — · 6 samples, - - 8 samples, — 10 samples, ··· 30 samples. The corresponding value for the optimal solution in the frequency domain is also shown (x x).

The discrete impulse responses of the reproduction chain were obtained from frequency response functions calculated with the modal model by using the inverse FFT. These frequency response functions had to be multiplied by the transfer function of a low pass filter, which corresponds to the action of the anti-aliasing filter used in practice anyway. A sixth-order analog lowpass Butterworth filter with a cutoff frequency of 175 Hz was applied, and 200 coefficients were used in the FIR filters of the reproduction chain. The simulations were carried out with 60 coefficients in the control filters. It was observed that practically the same result was obtained with 60 or more coefficients, and that the optimal equalisation was affected if the number was reduced to 40.

The same desired signal was used at the different M sensors. The number of delays δ_m used to generate the desired signals at the sensors turned out to have a significant effect on the optimal solution. To compare the results of the optimal equalisations corresponding to different numbers of delays (δ_m), the optimal causal filters were transformed back to the frequency domain and then used in the modal model to calculate $\hat{J}_0$ as a function of the frequency. It should be mentioned that in this case only the amplitude of the sound pressure was taken into account to obtain $\hat{J}_0$. Here J_0 was estimated with 100 sensors distributed in the listening area, and the curves are shown in Fig. 7. The best equalisation calculated before in the frequency domain is also shown for comparison. One can see that the optimal equalisation in the time domain is better when the number of delays is increased. This occurred up to a 16-sample delay. After this value the estimated cost function practically did not change, and the curves were similar to the one corresponding to 30-sample delays shown in the figure. It should be noted that a poor equalisation was obtained if the number of delays was less than 6 samples. This number of delays corresponds to the time of 5-6 samples that takes to the signals to travel from the farthest sound sources to the listening area.

In the examined simulation, a delay of 10 samples was considered adequate for an acceptable level of the equalisation up to 350 Hz. That was the minor delay that corresponds to a relatively small value of $\hat{J}_0$ at 350 Hz (less than 0.15 Pa^2). It can also be seen in Fig. 7 that the equalisation presents problems at very low frequencies, where $\hat{J}_0$ has values about 0.1 Pa^2. An explanation might be that the sound sources were modelled as high pass filters, and therefore, their response are almost zero at these frequencies. Nevertheless, this is not a serious problem since these frequencies are not audible.



(a) (b)

Fig. 8: Impulse and frequency responses at two sensors after the introduction of equalisation filters (solid lines), and the original frequency responses at these sensors (dashed lines). (a): sensor at the point (0.9,1.0,0.65), (b): sensor at the point (1.3,1.2,1.05). In 21 of the 27 sensors the equalisation was similar to (a). Here (b) corresponds to the worst result obtained.

The impulse and frequency responses at two of the 27 sensors after the equalisation was carried out are shown in Fig. 8. The original frequency responses at these sensors are also shown. In 21 of the 27 microphones a flat frequency response with fluctuation within ± 3 dB from 5 to 350 Hz was obtained. The results are similar to those presented in graphs (a) in Fig. 8. The worst result was obtained with the sensor placed at the point (1.30, 1.20,1.05), which corresponds to the graphs (b) the figure. In the other sensors, an equalisation within ± 3 dB from 5 to 275 Hz in the frequency response was obtained. As one can see, it was possible to achieve an adequate level of equalisation. However, as a consequence of the causality restriction, the optimal solution is affected and the error is increased at high frequencies. An important consequence is that an extra delay has also to be introduced in order to improve the performance of the equalisation.

CONCLUSIONS

A theoretical study of equalising the sound field at low frequencies in an extended region of a room has been carried out. Based on the results of computer simulations in the frequency domain, it can be concluded that a good optimisation can be obtained in an extended region inside a room in a given interval at low frequencies. Nevertheless, the performance is reduced with the frequency, and the size of the region depends on the maximum frequency to be equalised. According to the results, for a listening area of cubic shape, its side length cannot exceed half wavelength at the maximum frequency of the interval to be optimised if fluctuations in the obtained sound pressure less than ±3 dB from the desired value are required in the listening area.

It was also found that a better equalisation was obtained over a frequency range when the sound sources are placed in the corners of the room. In this case, it was also observed that the performance of the equalisation was lightly reduced when the listening area was placed near the walls of the enclosure. However, this effect is less pronounced when the frequency is reduced. Thus, the equalisation was found to be practically independent of the position of the listening area as long as it was approximately farther from any wall than a quarter of the wavelength at the maximum frequency to be equalised.

Unfortunately, the filters for the optimal solution obtained in the frequency domain correspond to non-causal impulse responses. Another approach is to perform the optimisation in the time domain. It was shown that the constrain of causality degrade the performance predicted by the unconstrained optimisation from the frequency domain, and that an extra delay is needed to obtain an adequate equalisation. A compromise between the length of the delay and the level of the equalisation has to be made.

REFERENCES

[1] S. J. Elliott and P. A. Nelson, "Multiple-Point Equalization in a Room Using Adaptive Digital Filters," *J. Audio Eng. Soc.*, vol. 37 (11), 1989, pp. 899-907.

[2] S. J. Elliott, L. P. Bhatia, F. S. Deghan, A. H. Fu, M. S. Stewart, and D. W. Wilson, "Practical Implementation of Low-Frequency Equalization Using Adaptive Digital Filters," *J. Audio Eng. Soc.*, vol. 42 (12), 1994, pp. 988-998.

[3] A. S. Munshi, "Equalizability of Room Acoustics," *1992 IEEE International Conference on Acoustics, Speecch ans Signal Processing,* vol. 2, 1992, pp. 217-220.

[4] F. Asano, Y. Suzuki, and T. Sone, "Sound Equalization Using Derivative Constraints," *Acustica,* 82(2), 1996, pp. 311-320.

[5] A. J. Bullmore, O. A. Nelson, A. R. D. Curtis, and S. J. Elliott, The Active Minimization of Harmonic Enclosed Sound Fields, Part II: A Computer Simulation, *Journal of Sound and Vibration*, 117(1), 1987, pp 15-33. Equations (1) and (10).

[6] F. Asano, D. C. Swason, Sound Equalization in Enclosures Using Modal Reconstruction, *J. Acoust. Soc. Am.,* 98 (4), 1995, pp 2062-2069.

[7] S. Laugesen, "Active Control of Acoustic Noise Using Adaptive Signal Processing," Ph.. D. Thesis, The Acoustics Laboratory, Technical University of Denmark, 1992.