Journal of Sound and Vibration (1991) **146**(2), 281–302

POLE AND ZERO MODELING OF ROOM TRANSFER FUNCTIONS

J. MOURJOPOULOS AND M. A. PARASKEVAS

Wire Communications Laboratory, Electrical Engineering Department, University of Patras, Patras 26500, *Greece*

(*Received* 19 *February* 1990, *and in revised form* 29 *June* 1990)

The feasibility of using all-pole and all-zero model approximations of room transfer functions was examined especially in respect to the degree that such approximations are suitable for removing room reverberation from signals. Two aspects of the above models were assessed: their success in reducing room transfer function order and their insensitivity to measurements taken for different source and receiver positions inside a room. Both these aspects are crucial to the success of practical dereverberation methods. The tests were carried out on simulated and measured data and it was found that all-pole models are most suitable than all-zero models when both of the above two aspects must be satisfied by the approximation model. It was also shown that, for a given room, an optimum all-pole model order exists which approximates the room transfer function.

1. INTRODUCTION

Over the past decades considerable advances have been achieved in the field of room acoustics by the use of signal processing techniques. Benefits have been gained mainly in the areas of measurement and analysis of room acoustics. Furthermore, the prediction of acoustical responses of auditoria and the simulation of reverberation can now be easily achieved by using such methods. However, there are still significant problems to be solved by methods allowing the manipulation and analysis of room transfer functions. Such functions exhibit high complexity and variability and for these reasons, many known signal processing methods cannot be fully exploited, especially with respect to their realization in real-time systems. Examples of this are the applications in which room reverberation must be removed from acoustic signals (e.g., speech and music), or in which the effects of transmission through rooms must be somewhat anticipated.

Applications in which dereverberation is required range from speech and music signal enhancement (when these signals have been recorded under unfavorable reverberant conditions), speech recognition by computers, noise cancellation applications, public address system equalization, etc. Some of these applications require realization of real-time processing techniques, whereas some of them require post-processing of the recorded signals. The work presented here is not concerned with a specific application, but instead is an examination of theoretical and practical aspects which may be applicable to such applications and existing dereverberation methods.

The ususal approach followed in such cases is based on the principles of signal deconvolution, or adaptive filter equalization: i.e., relying in designing linear filters to compensate for the room transfer function. In some cases such filters are designed for a single input (source) and output (receiver) configuration [1, 2] and in other cases the problem is approached by a single input multiple output (receiver) formulation, that has

0022-460X/91/080281+22 $03.00/0

often been based on empirical knowledge of the processing performed by the binaural (two-channel) hearing mechanism [3, 4], or by using multiple-point optimization [5].

Although these multiple-channel approaches offer considerable advantages in many respects, they also depend on the design of linear inverse filters of the room transfer function in the same way as the first approach does. For the design of these filters, two main limitations are imposed by the nature of room acoustics: (a) a high order filter must be designed (typically up to order of 10 000 for a finite impulse response formulation), and (b) this filter may be effective and appropriate for a very limited spatial combination of source and receiver positions within the particular enclosure [2, 5]. These limitations restrict the practical realization of equalization systems, whether these are based on single or multiple channel configuration. Multiple-point equalization by using adaptive filters [5] offers an attractive solution to this second problem, but in many aspects it may gain in efficiency by reconsidering the formulation of the filter representing the room transfer function which in turns may lead to its efficient inverse filter realization. The work presented in the following sections addresses the above problems by considering lower order approximations to room transfer functions, using all-pole or all-zero models.

Furthermore, an adaptive real-time approach was not followed in this work, but instead a systematic analysis of pole and zero modeling aspects has been attempted from simulated and measured data. The conclusions drawn from this study can be employed with single or multi-point adaptive equalization techniques, and they may also be used in more efficient realizations of such methods relying on the use of such approximations.

In more detail, the following problems have been addressed: (a) the efficient representation of simulated and measured room transfer functions in terms of all-pole or all-zero models; (b) the sensitivity of the above models to spatial variation in source/receiver configuration; (c) the derivation of an optimum all-pole order filter for a given room transfer function; (d) the results of all-pole based dereverberation, implemented in real time by using a DSP microprocessor system (Texas TMS 32020).

Section 2 of this paper introduces the basic concepts leading to all-pole or all-zero representation of room spectra. The concept of inverse filter design is also discussed. Finally, the concepts of all-pole model optimization are introduced and novel results are presented for the required approximation filter order.

Results for application of both all-pole and all-zero models to simulated acoustic enclosure spectra are presented in section 3. Two main aspects are examined: (i) the effect of filter order; and (ii) the sensitivity of these approximations for changing source/receiver spatial configuration. In the same section, results obtained in a real enclosure (the properties of which have been also simulated) are also discussed.

2. THEORETICAL CONCEPTS

2.1. INTRODUCTION

The acoustic response of any enclosure between a source and a receiver is the result of directly and multi-path transmitted sound which can be modeled by the sum of the contributions of acoustic modes. Such analysis leads to a transfer function representation of the room sound field, which can be defined for a specific source and receiver positions in the enclosure.

An impulse response (time domain) function is uniquely associated with such a transfer function. Then, by using a discrete-time representation, this function is expressed as $h(nT)$, where $n = 1, 2, \ldots$ indicates the sample value and T (s) is the sampling interval, assumed for simplicity to be $T = 1$. The impulse response function can be considered to be unique for each enclosure and for a given source/receiver configuration. However, in

practice such aspects as room temperature changes, opening a door or windows, and an audience may result in changes of these functions. Such practical aspects are not addressed here, since mainly the principles of approximating these functions by lower order models are examined.

Given that $h(n)$ is a stable and causal sequence (as is the case with all practical room response functions), then, in a z-transform sense and within an arbitrary constant scaling factor, the room transfer function can be described by a rational function of the form

$$H(z) = N(z)/D(z) = \left\{ \prod_{k=1}^{m_1} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z) \right\} \Big/ \prod_{k=1}^{p_i} (1 - c_k z^{-1}), \tag{1}$$

where the poles c_k are such that $|c_k| < 1$, $k = 1, 2, \ldots, p_i$, and the zeroes are such that $|a_k|$ and $|b_k| < 1$, $k = 1, 2, \ldots, m_0$.

From this general representation it is often useful to derive functions of all-pole or all-zero form, which in turn will lead to infinite impulse response (IIR) and finite impulse response (FIR) representations of the room filtering effects. Furthermore, it is often useful to consider approximations to the all-pole and the all-zero model which, although not exact in the form of equation (1), nevertheless give a valid representation of the room transfer function. As a general observation it can be noted that room transfer function resonances are described by the all-pole model, whereas transfer function notches are described by the all-zero model. Such models allow a more efficient implementation of room simulation and dereverberation to be achieved due to the considerable reduction of the filter order.

2.2. ALL-POLE ROOM MODELS

There is a wide use of all-pole models for approximating rational transfer functions such as the one given by equation (1) in many fields, and especially in speech analysis [6]. Such techniques have achieved considerable reduction in the order and complexity of the transfer functions, allowing efficient computational methods to be employed. The all-pole model H_p corresponding to the room transfer function (equation (1)) can be expressed as

$$H_p(z) = G \Big/ \prod_{k=1}^{p_i} (1 - c_k z^{-1}) = G \Big/ \left\{ 1 + \sum_{k=1}^{p_i} d_k z^{-k} \right\} = G/D(z), \tag{2}$$

where d_k are the all-pole model coefficients ($|d_k| < 1$, $k = 1, 2, \ldots, p_i$), and G is an arbitrary constant scaling term which, for simplicity, will be ignored in all subsequent discussion (i.e., it is assumed that $G = 1$).

The usual problem in the all-pole formulation is to determine the values of d_k, which is achieved by using the well-known least-squares techniques [6, 7]. The advantage of the all-pole model is that it allows one to determine the coefficients d_k, so that the time-domain all-pole filter response output $h_p(n)$ can be expressed as

$$h_p(n) = \delta(n) - \sum_{k=1}^{p_i} d_k h_p(n-k), \tag{3}$$

where $\delta(n)$ is the filter excitation. With the best estimate of the output denoted as $\hat{h}_p(n)$, then

$$\hat{h}_p(n) = - \sum_{k=1}^{p_i} d_k h_p(n-k). \tag{4}$$

The all-pole filter coefficients will define the optimal linear predictor. The mean-squared prediction error will be

$$e_p(n) = h_p(n) - \hat{h}_p(n) \tag{5}$$

and

$$e_p(n) = h_p(n) + \sum_{k=1}^{p_i} d_k h_p(n-k). \tag{6}$$

The squared error, R_p, will be

$$R_p = \sum_n e_p^2(n) = \sum_n \left[h_p(n) + \sum_{k=1}^{p_i} d_k h_p(n-k) \right]^2. \tag{7}$$

To solve for the predictor coefficients, equation (7) is differentiated with respect to d_i $(i = 1, 2, \ldots, p_i)$ and the result is set to zero, leading to

$$\sum_{k=1}^{p_i} d_k \left[\sum_n h_p(n-k) h_p(n-i) \right] = -\sum_n h_p(n) h_p(n-i), \qquad \text{for } 1 \leqslant i \leqslant p_i. \tag{8}$$

In a matrix formulation, following the autocorrelation approach [6, 7], one has

$$\begin{bmatrix} R_0 & R_1 & \ldots & R_{p-1} \\ R_1 & R_0 & \ldots & R_{p-2} \\ \vdots & \vdots & & \vdots \\ R_{p-1} & R_{p-2} & \ldots & R_0 \end{bmatrix} \cdot \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_p \end{bmatrix}. \tag{9}$$

The autocorrelation matrix in the left side of equation (9) is a symmetric positive Toeplitz matrix and may be efficiently solved by the Wiener-Levinson algorithm [6, 7]. When the all-pole filter is not constrained to be causal or finite, then this filter will lead to the ideal all-pole solution. By constraining the solution to be finite and causal, then the minimum-phase approximation to the transfer function is derived. Solution of the above equations were derived for measured and simulated room response data (as described in more detail in section 3.2) and for values of p in the range $50 \leqslant p \leqslant 500$. This is an untypically large order of the pole prediction model, when compared to speech analysis applications [6], when ususally $4 \leqslant p \leqslant 30$.

Typical results are shown in Figure 1, and illustrate the satisfactory representation of the room transfer function by the all-pole approximation and the effects of model order on such representations.

2.3. ALL-ZERO ROOM MODELS

To derive the all-zero representation $H_z(z)$, of the room transfer function defined by equation (1), it is possible to write

$$H_z(z) = \prod_{k=1}^{m_i} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z). \tag{10}$$

The first product term in the right side of equation (10) corresponds to the minimum-phase component, and the second product term to the maximum phase component of the room all-zero transfer function. It is now feasible to obtain a minimum phase term which includes all the $m_i + m_0$ zeros of $H_z(z)$, often called the "equivalent minimum phase" function of $H(z)$ [2, 1, 8]. In order to obtain this function, the right side of equation (10) is multiplied by the same term:

$$H_z(z) = \prod_{k=1}^{m_i} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z^{-1}) \left\{ \prod_{k=1}^{m_0} (1 - b_k z) \Big/ \prod_{k=1}^{m_0} (1 - b_k z^{-1}) \right\}. \tag{11}$$

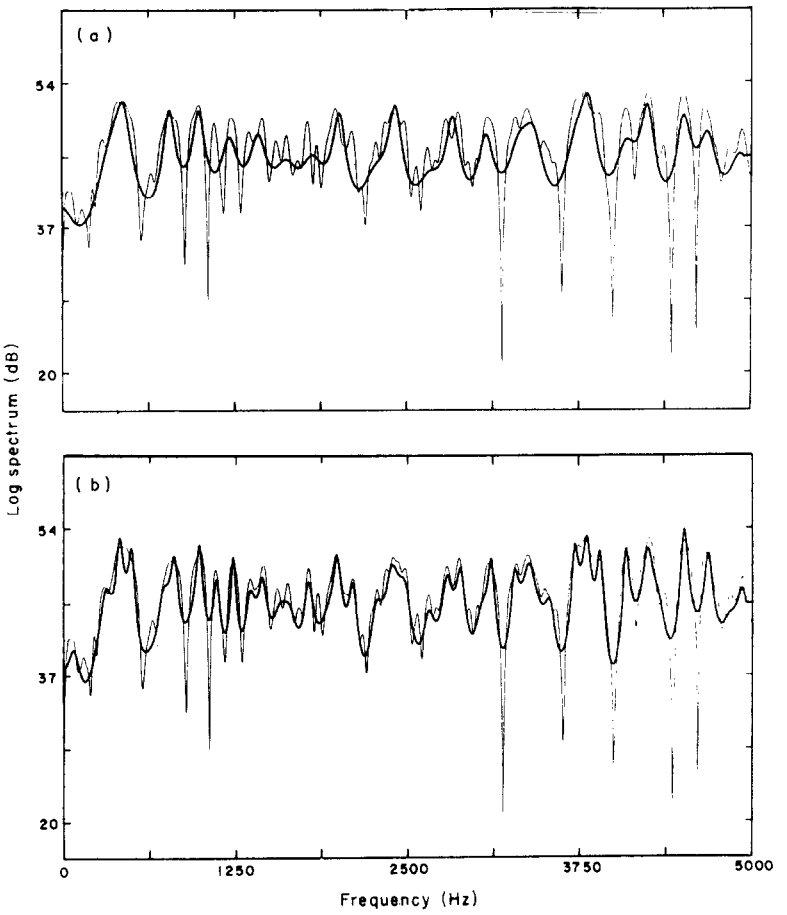

Figure 1. Room transfer function and all-pole model approximations, for predictor orders (a) $p = 50$ and (b) $p = 100$ for room 1; and (c) $p = 100$ and (d) $p = 200$ for room 2. ——, Original room spectrum; ——, all-pole room approximation.

Then, the ("equivalent") minimum-phase version of the all-zero room model will be

$$H_{ze}(z) = \prod_{k=1}^{m_i} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z^{-1}). \tag{12}$$

The function

$$H_{ap}(z) = \prod_{k=1}^{m_0} (1 - b_k z) \Big/ \prod_{k=1}^{m_0} (1 - b_k z^{-1}) \tag{13}$$

is an all-pass function, since [8] $|H_{ap}(e^{j\omega})| = 1$, for all ω.

In order to extract the all-zero approximation of the room transfer function, many alternative approaches may be adopted [8]. For the results discussed in section 3.2, the homomorphic approach was followed and the all-zero model was obtained for the full length of the room response function. As is known [1, 2], such models achieve exact approximations of the modulus of the transfer function at the expense of the time domain performance and of computational efficiency.

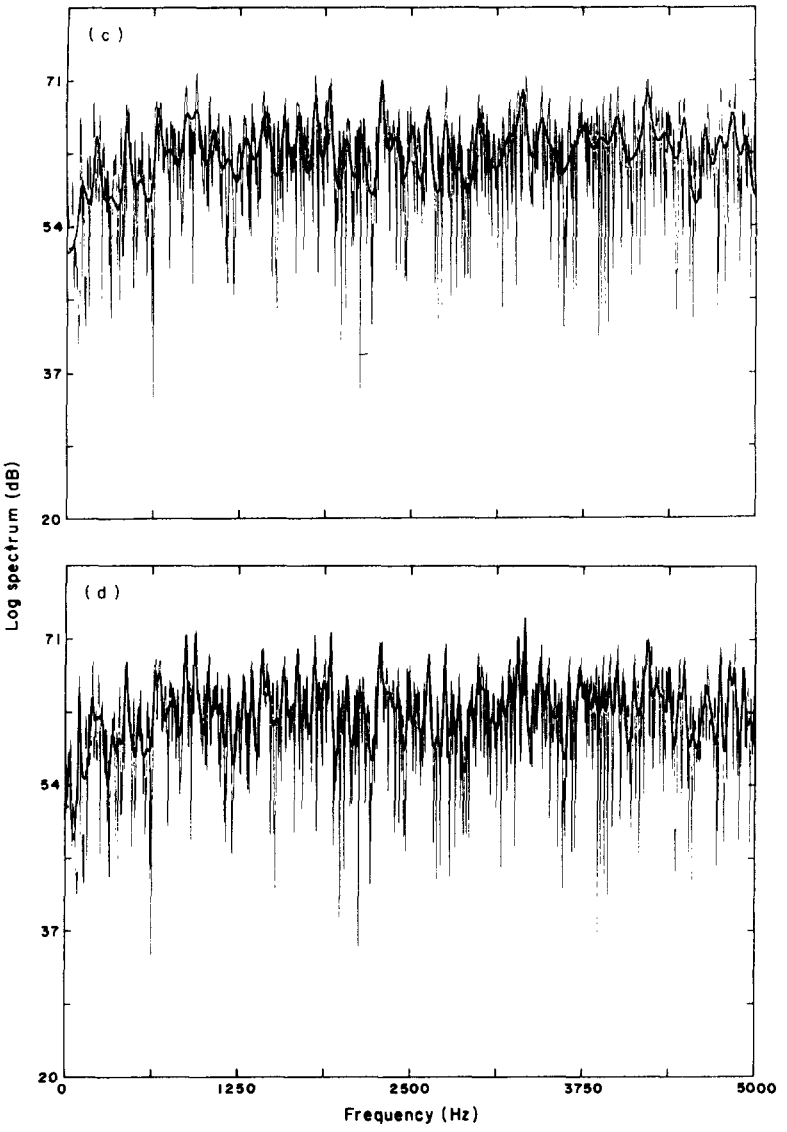

Figure 1—*continued.*

2.4. INVERSION OF THE ROOM TRANSFER FUNCTION

In many signal processing applications in the field of acoustics it is required that the effects of the signal transmission through rooms are removed, so that the source signal can be recovered or a desired acoustic signal can be reconstructed at some specific point in that enclosure. In such applications, an inverse filter to the room function must be derived, as follows. Assume that $h_i(n)$ is an inverse impulse response function, corresponding to the room impulse response, for the source/receiver placement producing $h(n)$. Then, the effect of inverse filtering on the room function will be

$$h(n) * h_i(n) = \delta(n). \tag{14}$$

In general, the design of the inverse function $h_i(\ )$ is not straightforward. The main difficulties in deriving this function from a measurement of $h(n)$ are briefly discussed

here. As given by equation (1), the room transfer function has in general mixed-phase properties, and as such it does not have a direct stable and finite inverse. For this, the minimum phase version of this function is often only inverted [1]. Alternatively, the mixed-phase properties must be compensated by the use of suitable algorithms [2, 5]. Furthermore, the inverse filter is lengthy and appropriate only for a particular measuring position [2, 5]. As was illustrated in reference [2], realization of an inverse filter at one position will in general deteriorate the response measured for other positions. With these problems in mind, possible improvements will be examined by considering the following points: (a) the efficiency of all-pole room models with respect to inverse filter length and their sensitivity to variation in measuring position; (b) the efficiency of all-zero models with respect to inverse filter length and their sensitivity to variation in measuring position.

Since both the all-pole and all-zero models are minimum phase functions (see the previous section), then the inversion of these functions (by Fourier transform methods) results in stable sequences. Such functions are derived, for example, for the all-pole model, as

$$h_{ip}(n) = Z^{-1}\{1/H_p(z)\}, \tag{15}$$

$Z^{-1}(\ )$ indicating the inverse z-transform of $H_p(z)$. Alternatively, the least-squares derivation directly yields this inverse. Compensating for the transmission delay may also achieve a stable inverse for the mixed-phase function. Furthermore, the success of the models may be estimated by considering how close their inverse functions (obtained as was explained above) can match the original function. Consider the expression

$$h(n) * h_{ip}(n) = \hat{\delta}_p(n), \tag{16}$$

where $\hat{\delta}_p(n)$ defines the approximation of the model to the ideal delta function (see Figure 2(a)). Then the all-pole model prediction error is expressed as

$$e_p(n) = \delta(n) - \hat{\delta}_p(n). \tag{17}$$

A similar expression may be also derived for the all-zero model: i.e.,

$$h(n) * h_{iz} = \hat{\delta}_z(n) \qquad \text{and} \qquad e_z(n) = \delta(n) - \hat{\delta}_z(n), \tag{18, 19}$$

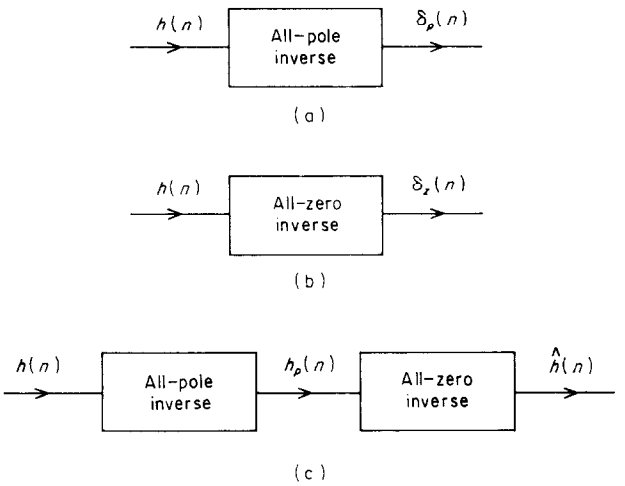

Figure 2. Block diagram of inverse filtering based on all-pole and all-zero models: (a) all-pole inversion; (b) all-zero inversion; (c) pole-zero inversion.

where $\hat{\delta}_z(n)$ defines the approximation of the model to the ideal delta function (see Figure 2(b)).

The smaller the error function e_p, or e_z, the better is the match of the particular model to the room transfer function. These criteria also will be discussed in section 3. Finally, the all-pole and all-zero models may be combined, in inverse form as is shown in Figure 2(c). Given that the ideal all-pole inverse will be (equation (2))

$$H_{ip}(z) = \prod_{k=1}^{p_i} (1 - c_k z^{-1}) \tag{20}$$

and the ideal all-zero equivalent minimum phase inverse will be

$$H_{ize}(z) = 1 \Big/ \left\{ \prod_{k=1}^{m_i} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z^{-1}) \right\}, \tag{21}$$

then the product of the combined all-pole and all-zero inverse as shown in Figure 2(c), will be a rational function of the form

$$H_{is}(z) = H_{ip}(z) H_{ize}(z) = \prod_{k=1}^{p_i} (1 - c_k z^{-1}) \Big/ \left\{ \prod_{k=1}^{m_i} (1 - a_k z^{-1}) \prod_{k=1}^{m_0} (1 - b_k z^{-1}) \right\}: \tag{22}$$

i.e., the approximate inverse to $H(z)$ within the residue all-pass function, $H_{ap}(z)$ (see equation (13)).

A typical example of the application of the all-pole inverse filter is shown in Figure 3. Figure 3(a) shows a 2048-point simulated room impulse response and Figure 3(b) shows the function $\hat{\delta}_p(n)$, after all-pole model ($p = 50$) inverse filtering. The corresponding spectra are shown in Figure 3(c) and 3(d). Corresponding results for the case of the all-zero model are shown in Figure 4. Figure 4(a) shows the function $\hat{\delta}_z(n)$ after all-zero model (2048-point) inverse filtering. The spectrum of the remaining all-pass error is shown in Figure 4(b). It is clear that this model achieves a closer approximation than the all-pole model, at the expense of computational efficiency. These results are also shown in the form of the envelope function (Figure 5), and also as reverberant decay curves corresponding to the original and processed response function (Figure 6).

2.5. ERROR MINIMIZATION AND OPTIMUM ALL-POLE MODEL ORDER

This section examines conditions relating the order of the all-pole model and the resulting approximation error. Such a relationship is useful since it predicts the success of these models, especially with respect to computational advantages gained by the lower order approximation. Let $P(e^{j\omega})$ denote the power spectrum corresponding to the original room transfer function $H(e^{j\omega})$. Also, let the all-pole model approximation power spectrum be defined as $P_p(e^{j\omega})$ [9]: i.e.,

$$P(e^{j\omega}) = |H(e^{j\omega})|^2 \qquad \text{and} \qquad P_p(e^{j\omega}) = |H_p(e^{j\omega})|^2. \tag{23, 24}$$

From equations (2) and (24), $P(e^{j\omega})$ will be

$$P_p(e^{j\omega}) = 1/|D(e^{j\omega})|^2. \tag{25}$$

From expression (6), the error in a z-transform sense, (for $z = e^{j\omega}$) will be

$$E_p(e^{j\omega}) = H(e^{j\omega}) D(e^{j\omega}), \tag{26}$$

where $E_p(e^{j\omega}) = Z\{e_p(n)\}$.

From the discussion in section 2.2 it is clear that this function represents the remaining zeroes: i.e., it represents $H_z(e^{j\omega})$ for the particular room transfer function. According to

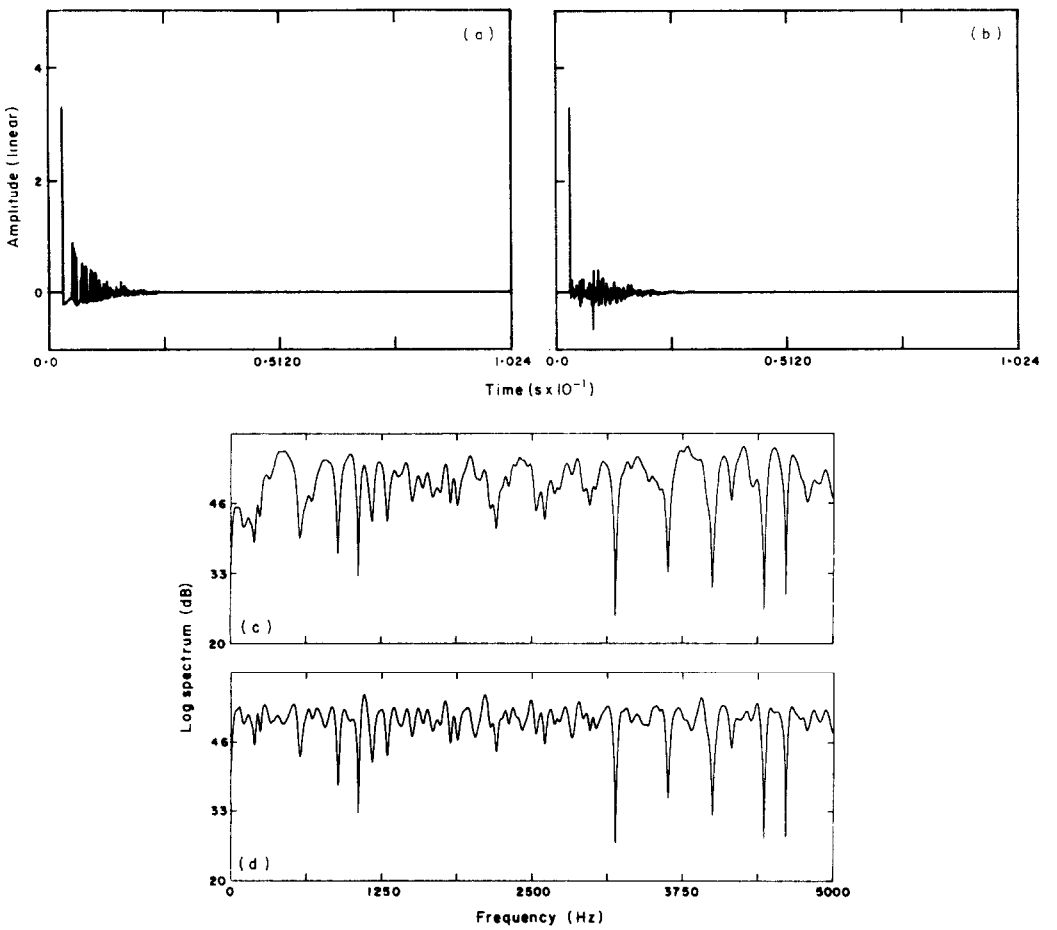

Figure 3. All-pole room response inverse filtering: (a) simulated room impulse response; (b) all-pole inverse filter residue ($p = 50$) (the first 1024 response points are shown); (c) original room response spectrum; (d) spectrum after all-pole inverse filtering.

Parseval's theorem, the overall error R_p will be

$$R_p = \sum_{n=-\infty} e_p^2(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E_p(e^{j\omega})|^2 \, d\omega. \tag{27}$$

Then, from equations (23) and (26), one has

$$P(e^{j\omega}) = |E_p(e^{j\omega})|^2 / |D(e^{j\omega})|^2. \tag{28}$$

From equations (25) and (28) and if $P(e^{j\omega}) = P_p(e^{j\omega})$, one must have that $E_P(e^{j\omega}) \to 1$: i.e., it tends to have a constant value. Then

$$E_p(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{P(e^{j\omega})}{P_p(e^{j\omega})} \, d\omega. \tag{29}$$

Therefore, for minimization of the prediction error it is required that the integrated ratio of the room spectrum to the estimated all-pole room spectrum is minimized. Ideally, this can be achieved for $p \to \infty$; i.e., when $P(e^{j\omega}) = P_p(e^{j\omega})$.

It can be shown [9] that in general the all-pole model will achieve a closer approximation to the original function at peak spectral amplitude regions, indicating a closer fit of the

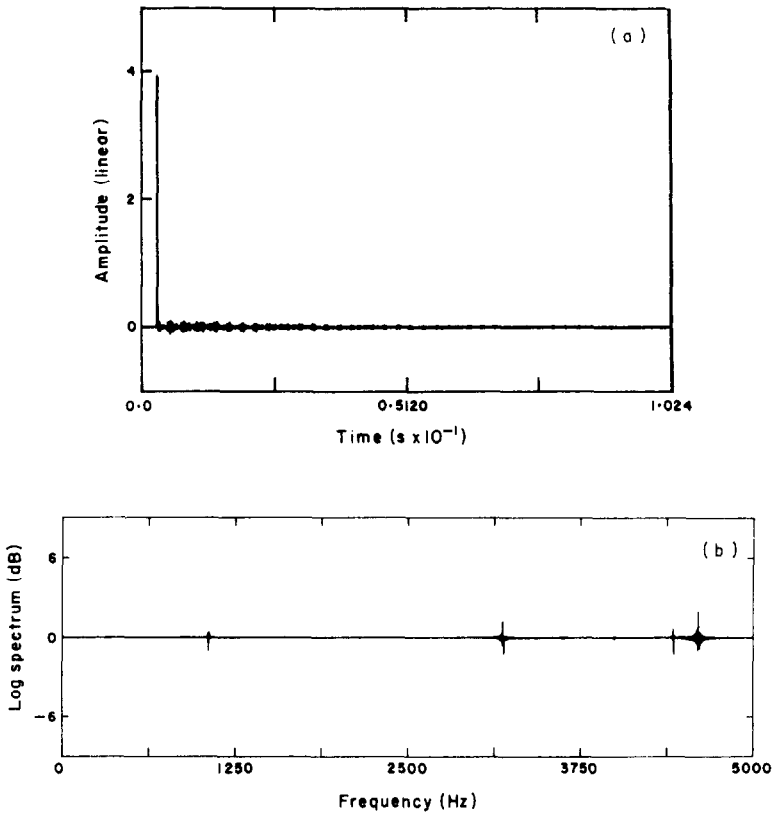

Figure 4. All-zero room response inverse filtering: (a) all-zero inverse filter residue (2048-point filter—the first 1024 response points are shown); (b) spectrum after all-zero inverse filtering.

model at the combined room resonance frequencies. It is now useful to derive a normalized expression for the minimum predictor error with respect to the overall all-pole model energy, estimated via its zeroth order autocorrelation value (i.e., $R(0)$). For this, one can define [9]

$$V_p = E_p/R(0) = \left\{\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[P(e^{j\omega})\right] d\omega\right]\right\} \Big/ R(0) \tag{30}$$

and

$$V_{min} = \left\{\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[P_p(e^{j\omega})\right] d\omega\right]\right\} \Big/ R(0), \tag{31}$$

where V_p is the normalized predictor error and V_{min} is the minimum normalized predictor error.

Ideally, the normalized error V_p is a monotonically decreasing function of p: i.e., $V_0 = 1$, for $p = 0$ and $V_{min} = V_\infty$, for $p \to \infty$. For optimum application of the all-pole model to room acoustics, it is important to relate the value of V_p to the filter order, p. As an example, this relationship (obtained via equation (30)) is shown in Figure 7 for three different source/receiver positions in one of the rooms employed for the tests. These results show that in all three positions the same optimum number of poles achieved the minimization of normalized error. The value of the normalized approximation error was

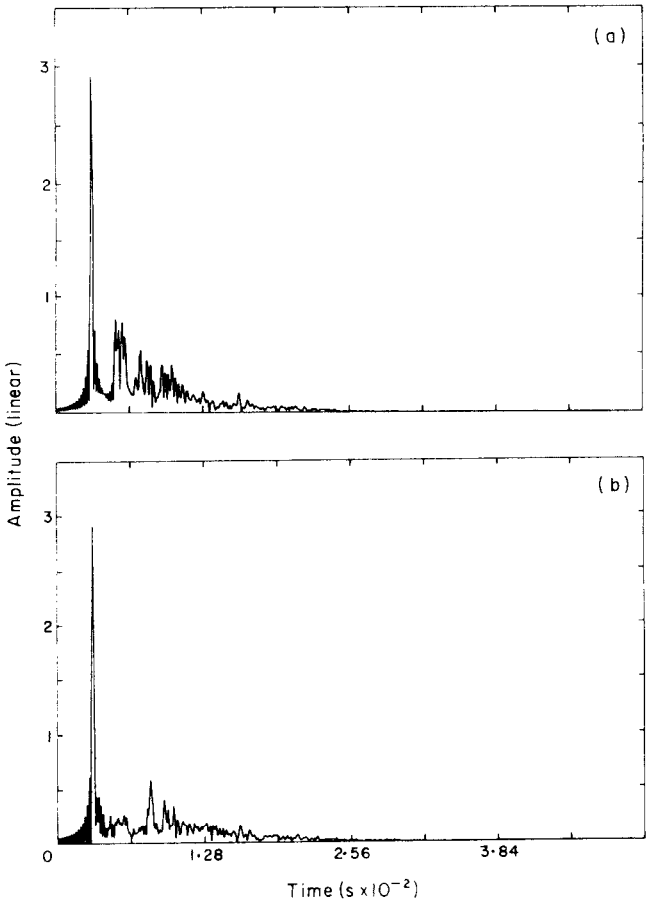

Figure 5. Effects of all-pole model inverse filter on response envelope function: (a) original response function (Room 1); (b) processed response envelope ($p = 50$).

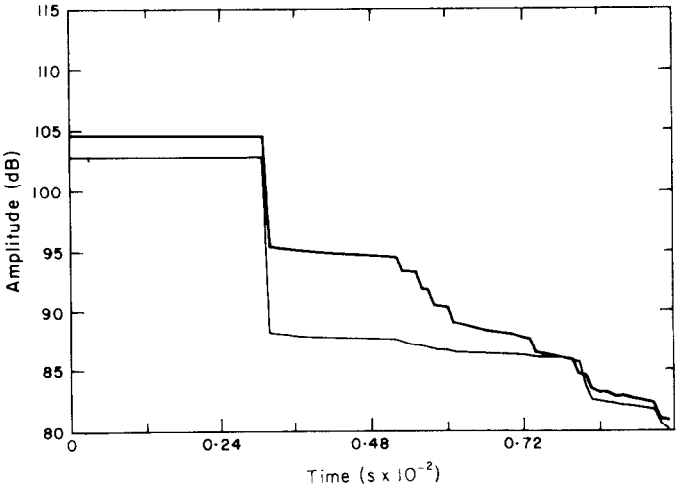

Figure 6. Comparison of reverberant decay before and after all-pole model inverse filtering (results for Room 1). ——, Original; ——, processed.

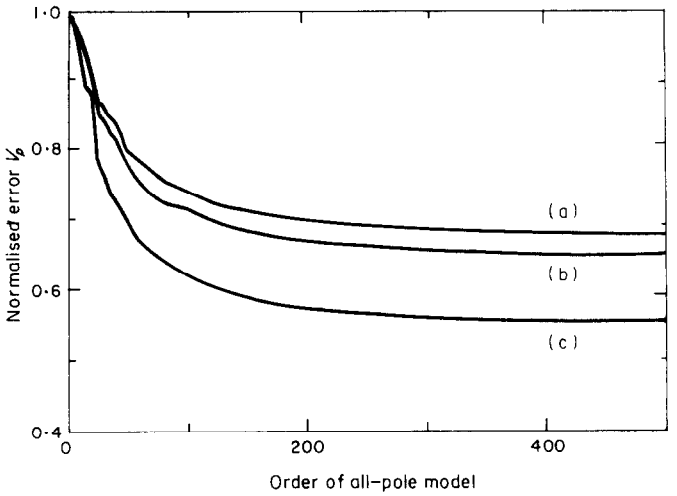

Figure 7. All-pole model normalized error evaluated as a function of the predictor order, for three different source/receiver positions in Room 1.

different in each case, since due to the source/receiver configuration fewer overall distortions were present at the measurement taken at position (a), when compared to those present for measurements at positions (b) and (c) (see section 3).

It is now feasible to obtain the theoretical limits for the optimum order of the all-pole model approximation for a given room transfer function. For practical systems such as rooms, an optimum all-pole order may be estimated when the following normalized error reduction criterion is satisfied:

$$1 - V_{p+1}/V_p < \delta. \tag{32}$$

Here δ is an arbitrary convergence threshold. Akaike [10] has suggested conditions which satisfy this criterion. This relates to the minimization of the information criterion $I(p)$

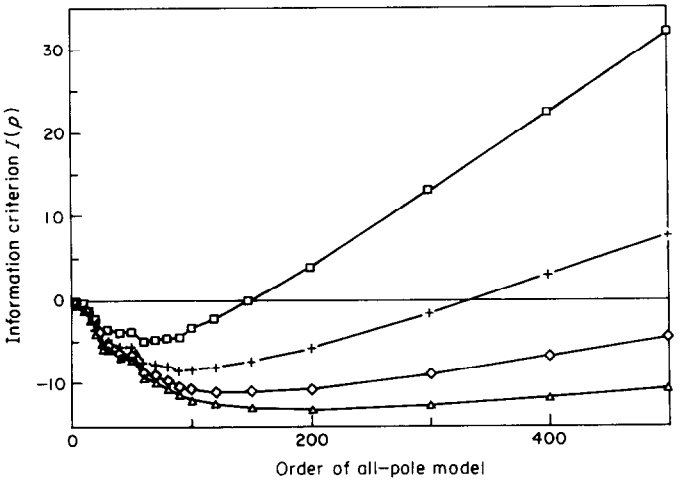

Figure 8. All-pole model information criterion evaluated as a function of the predictor order, for different lengths of room response data. — □ —, 2048 points; — + —, 4096 points; — ◇ —, 8192 points; — △ —, 16 384 points.

for the optimum value of p, so

$$I(p) = \log V_p + 2p/N_e, \tag{33}$$

where N_e is the effective number of sample values of the room impulse response data.

The results of the application of this expression to response functions corresponding to one of the rooms employed for the tests (described in section 3.2) is shown in Figure 8. This figure indicates that for shorter room impulse response functions a lower number of poles would optimize the information criterion. Clearly, therefore, larger rooms with longer reverberant decay will require a higher order of all-pole model approximation. It is also important to note that even for the case of the optimum all-pole model prediction significant gains are still achieved with respect to the order of the overall room transfer function.

In conclusion, the proposed expressions provide a useful guidance in predicting the required optimum all-pole model which represents rooms of different overall acoustic properties.

3. RESULTS

3.1. PERFORMANCE CRITERIA

The effectiveness of pole or zero models can now be assessed by performance criteria in the time and frequency domains, in a way similar to that for the ones described in reference [2]. It is assessed how close deconvolution of the room function by an inverse filter designed by the pole or zero models will be to the ideal delta function and flat spectrum: i.e., if $h_p(n)$ is the all-pole response corresponding to the room response $h(n)$, and $h_{ip}(n)$ is the inverse of the all-pole model, then

$$h(n) * h_{ip}(n) = \hat{\delta}_p(n), \tag{34}$$

where $\hat{\delta}_p(n)$ defines the approximation of the deconvolution to the ideal delta function, expressed also as an error function,

$$e_p(n) = \delta(n) - \hat{\delta}_p(n). \tag{35}$$

Similarly, if $h_z(n)$ is the zero-model corresponding to $h(n)$, with $h_{iz}(n)$ its inverse function, then

$$h(n) * h_{iz}(n) = \hat{\delta}_z(n), \tag{36}$$

and the corresponding error function will be

$$e_z(n) = \delta(n) - \hat{\delta}_z(n). \tag{37}$$

In now considering the first of the two criteria, the time-domain criterion, it is desirable to assess, in each case, the relative measure of the energy ratios between the direct and reverberant components. To achieve such a representation, the room response can be expressed as

$$h(n) = f(n) + g(n), \tag{38}$$

$f(n)$ being the direct response component occupying n_0 samples from response initialization (with the transmission path delay removed): i.e.,

$$f(n) = h(n), \qquad 0 \leq n \leq n_0 - 1. \tag{39}$$

Similarly, $g(n)$ gives the reverberant response component occupying the interval after the n_0 samples, to the end of the response: i.e.,

$$g(n) = h(n), \qquad n_0 \leq n \leq n_{max}. \tag{40}$$

n_{max} is a practical upper limit indicating the sample value at which the response has decayed to the measurement (or simulation) noise floor level. In effect, it corresponds to N_e, given in equation (33).

The time-domain criterion, expressed in decibels, is useful in order to determine the relative amount of room reverberation which can be modeled by each of the approximations. Applying this criterion to the original room response function, one has:

$$E_d/E_r = 10\log\left[\sum_{n=0}^{n_0-1} f^2(n) \Big/ \sum_{n=n_0}^{n_{max}} g^2(n)\right] \quad \text{(dB)}. \tag{41}$$

To evaluate the success (matching) of the pole and zero models, in the time domain, a modified E_d/E_r criterion was employed, giving residual reverberant energy remaining after the deconvolution with the corresponding response model. For the all-pole model this criterion is expressed as

$$E_{dip}/E_{rip} = 10\log\left[\sum_{n=0}^{n_0-1} f^2(n) \Big/ \sum_{n=n_0}^{n_{max}} e_p^2(n)\right] \quad \text{(dB)}, \tag{42}$$

and for the zero model case as

$$E_{diz}/E_{riz} = 10\log\left[\sum_{n=0}^{n_0-1} f^2(n) \Big/ \sum_{n=n_0}^{n_{max}} e_z^2(n)\right] \quad \text{(dB)}. \tag{43}$$

Here $e_p(n)$ and $e_z(n)$ are given by equations (35) and (37). For simplicity, a direct comparison of the original response and each of models was derived which is described by the time domain invertibility criterion, given, for the all-pole model as

$$I_p = E_{dip}/E_{rip} - E_d/E_r \quad \text{(dB)} \tag{44}$$

and for the zero model case as

$$I_z = E_{diz}/E_{riz} - E_d/E_r \quad \text{(dB)}. \tag{45}$$

Hence the values of I_p and I_z give the reverberant energies (in dB) which will be removed from the room response or signals after deconvolution by the response pole or zero model. Alternatively, the greater the value of this criterion (in dB) the more successful is the corresponding model in matching, the room response function. Similarly, a frequency-domain criterion can be derived indicating the spectral matching achieved by each of these models. This criterion is provided by the variance σ^2 of the modulus of the transfer function, defined for the room function as

$$\sigma = \left[\frac{1}{M}\sum_{m=0}^{M-1} [10\log|H(m)| - AV]^2\right]^{0.5} \quad \text{(dB)}, \tag{46}$$

where

$$AV = \frac{1}{M}\sum_{m=0}^{M-1} 10\log|H(m)| \quad \text{(dB)}, \tag{47}$$

and $H(m)$ is the M-point discrete fourier transform (DFT) of $h(n)$.

In the case of the pole-model, the spectral variance σ_p^2 of the "mismatch" function $\hat{\delta}_p(n)$ was taken: i.e.,

$$\sigma_p = \left[\frac{1}{M}\sum_{m=0}^{M-1} [10\log|\Delta_p(m)| - AV_p]^2\right]^{0.5} \quad \text{(dB)}, \tag{48}$$

where

$$AV_p = \frac{1}{M} \sum_{m=0}^{M-1} 10 \log |\Delta_p(m)| \quad \text{(dB)} \tag{49}$$

and $\Delta_p(m)$ is the M-point DFT of $\hat{\delta}_p(n)$. Similarly, for the zero-model case, the spectral variance was defined as

$$\sigma_z = \left[\frac{1}{M} \sum_{m=0}^{M-1} [10 \log |\Delta_z(m)| - AV_z]^2 \right]^{0.5} \quad \text{(dB)}, \tag{50}$$

where

$$AV_z = \frac{1}{M} \sum_{m=0}^{M-1} 10 \log |\Delta_z(m)| \quad \text{(dB)} \tag{51}$$

and $\Delta_z(m)$ is the M-point DFT of $\hat{\delta}_z(n)$.

For the direct comparison between the original transfer function and the models, a frequency-domain invertibility criterion was defined for the pole-model as

$$J_p = \sigma - \sigma_p \quad \text{(dB)} \tag{52}$$

and for the zero model as

$$J_z = \sigma - \sigma_z \quad \text{(dB)}. \tag{53}$$

Hence, the values of J_p and J_z give the amounts of spectral distortions (in dB) removed from the room transfer function after deconvolution by the response pole or zero model (when this value is positive). Alternatively, the larger the positive value of this criterion is (in dB), the more successful is the corresponding model in matching, in the frequency domain, the room transfer function.

In practice, better matching can be achieved for the lower-frequency regions of the room transfer function. This can be explained by the fact that for low frequencies (here the region below 500 Hz was considered), sparser room modal systems are contributing to the overall transfer function: these can be more closely described by a low order model, especially the all-pole one. To assess the effectiveness of the models in matching the low-frequency regions of room response, the variance measures of equation (48) were modified, so that the variance was taken over values of $m \leq M_{500}$, where M_{500} corresponds to the frequency of 500 Hz. These corresponding invertibility criteria are then described as J_{p500} and J_{z500}.

3.2. SIMULATIONS AND MEASUREMENT

For the results presented in the following sections, room impulse response functions were obtained by simulation or by measurement. The simulations were derived for two rectangular enclosures (described below). The physical properties of one of these enclosures corresponded to those of a real small room, the response measurements of which were also obtained. Furthermore, in this room, results were obtained for the real-time implementation of its inverse function, by using special purpose DSP hardware (Texas Instruments TMS-32020 microprocessor system).

In the two rooms used for the simulations different source/receiver positions were used, as is shown in Figures 9(a) and (b). The source/receiver co-ordinates and the dimensions of the two rooms are listed in Table 1. Room 1 is a small rectangular room, used for speech recording. Room 2 is a rectangular laboratory with fairly reflective walls, floor and ceiling.

For the room response simulations, a sound image method was employed, suggested by Allen [11]. These simulations were performed at a sampling rate of $f_s = 10$ kHz and

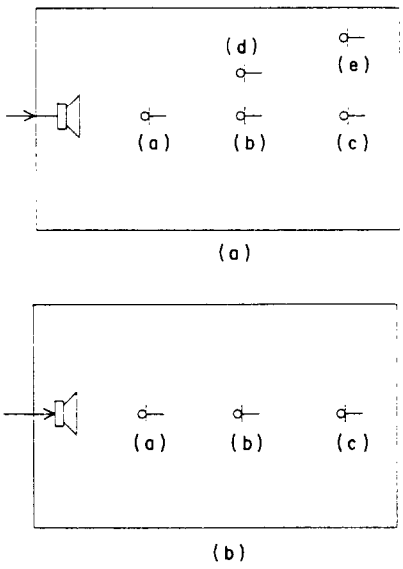

Figure 9. Configuration of source/receiver placement in (a) Room 1 and (b) Room 2.

TABLE 1

Room dimensions and different source/receiver positions for the two enclosures employed for the tests

	Room 1			Room 2		
	Height (m)	Depth (m)	Width (m)	Height (m)	Depth (m)	Width (m)
Dimensions	2·08	1·67	2·22	3·15	4·62	3·50
Source position	0·70	0·85	0·37	1·00	1·00	1·75
Receiver positions						
(a)	0·90	0·90	1·37	1·80	2·00	1·75
(b)	0·90	0·90	1·57	1·80	3·00	1·75
(c)	0·90	0·90	1·85	1·80	5·70	1·75
(d)	0·90	0·45	1·57			
(e)	0·90	0·20	1·85			

data were produced for response length of 2048 points (see Figure 3(a)). Response measurements were also carried out in Room 1 at positions corresponding to the ones employed for the simulations (see Figure 9). For these measurements a pulse generator was used for the room excitation, as is shown by the arrangement of Figure 10. The low noise and small volume of this enclosure did not necessitate any signal averaging or other

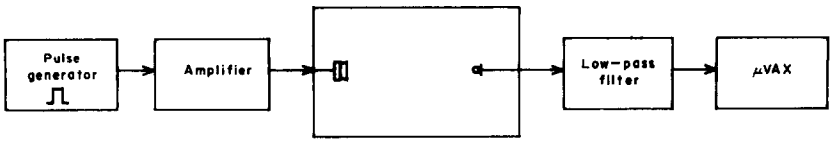

Figure 10. Block diagram of equipment arrangement for room impulse measurements.

more complex techniques to achieve responses with sufficient signal to noise ratio. The pulse duration was 20 μs, allowing an effective excitation bandwidth of approximately 50 kHz. The pulse amplitude was 2 volts. An eighth order elliptic anti-aliasing filter was employed with the cut-off frequency set at 4·5 kHz and the sampling rate was $f_s = 10$ kHz. The loudspeaker (Yamaha NS-10M), amplifier (Quad 303), and the microphone responses were in effect superimposed on the room response [2], but, given that their contribution remained constant throughout the tests, they generated only an additional linear filtering effect, also modeled by the inverse filter. Results obtained for the above tests are discussed in the next subsection.

3.3. SIMULATED ROOM RESPONSE STUDY

The success of both all-pole and all-zero modeling of the room transfer function was assessed by measurement of the effectiveness of deconvolution by inverse filters created by such models (see section 2.4). This may be interpreted as follows. It is assessed by objective measures how closely an all-pole or an all-zero model will describe the original room function. If such modeling is exact, then deconvolution by the inverse of the (pole or zero) model must yield a delta function in the time domain or a flat spectrum in the frequency domain. Deviation from these ideal conditions indicates the approximation error created by each of these models. Furthermore, the results presented in this section indicate the extent to which models constructed for one position in an enclosure can describe the room function at other positions. The first part of this study examines the effect of all-pole model order on the approximation error. As was also shown in Figure 7, it is clear that increasing the model predictor order increases its performance. This was observed in all cases: i.e., for both rooms, and in the time and frequency domains. These results are shown in Table 2. As an indication of the possible improvements obtained by the application of such a model, the reverberation time was measured from the reverberant decay before and after the application of inverse filtering, as is shown in Figure 6. Typically, for Room 2 the original RT_{60} value of 0·282 s was lowered to a value of 0·121 s.

TABLE 2

Time- and frequency-domain performance of all-pole model, for the two enclosures, and for different model orders

		Frequency				Frequency	
Room 1, order of model	Time, I_p (dB)	J_{p500} (dB)	J_p (dB)	Room 2, order of model	Time, I_p (dB)	J_{p500} (dB)	J_p (dB)
$p = 50$	3·50	1·91	0·69	$p = 100$	0·73	0·42	0·41
$p = 100$	5·60	2·05	1·10	$p = 200$	1·19	0·79	0·67
$p = 200$	8·60	2·47	1·59	$p = 400$	1·81	1·48	1·12

With reference to Figure 9, where typical source/receiver positions for the two enclosures are shown, the effect of the pole model and the zero model are given in Tables 3 and 4.

From the results of the comparison between the performance of the all-pole and all-zero models for Room 1, as shown in Table 3, the following observations can be made.

(a) Both all-pole and all-zero models improve the time-domain room response when they are employed for the exact measurement position. The improvement of the all-pole model is smaller than that of the all-zero model, but the model order is much lower

TABLE 3

Time- and frequency-domain performance of all-pole and all-zero models for Room 1; the asterisk indicates the position at which the inverse filter is estimated

	All-pole model			All-zero model		
		Frequency			Frequency	
Room 1	Time, I_p (dB)	J_{p500} (dB)	J_p (dB)	Time, I_z (dB)	J_{z500} (dB)	J_z (dB)
Position (a)*	3·50	1·91	0·69	12·03	2·91	2·28
Position (b)	−0·82	1·58	−0·15	−6·98	2·07	−0·65
Position (c)	−0·83	0·95	−0·16	−6·50	1·39	−0·60
Position (d)	−1·03	0·16	−0·28	−7·50	0·15	−0·82
Position (e)	−0·39	0·19	−0·04	−4·20	0·11	−0·49

TABLE 4

Time-and frequency-domain performance of all-pole and all-zero models for Room 2; the asterisk indicates the position at which the inverse filter is estimated

	All-pole model			All-zero model		
		Frequency			Frequency	
Room 2	Time, I_p (dB)	J_{p500} (dB)	J_p (dB)	Time, I_z (dB)	J_{z500} (dB)	J_z (dB)
Position (a)	−0·40	0·66	0·02	−16·44	−0·13	−1·14
Position (b)	−0·23	0·26	0·00	−18·40	0·16	−1·23
Position (c)*	0·73	0·42	0·42	−0·71	2·52	1·53

(typically by a factor of 40). Hence, the all-pole model is computationally more efficient than the all-zero model.

(b) Both models have removed some of the frequency-domain distortions, when employed at the measurement position. Again the all-zero model removed more of those distortions, at the expense of computational efficiency.

(c) In all cases the all-pole model proved to be much less sensitive to changes in receiver position than the all-zero model. In the time domain the position variation had a minimum effect in the all-pole model performance, whereas significant degradation was created by the all-zero model mismatch. Similarly, in the frequency domain the all-pole model appears to be much less sensitive than the all-zero model in source/receiver position variation.

(d) Both models are more efficient at removing low-frequency distortions, as indicated by the J_{500} criterion. Furthermore, the models improved the low-freqency spectrum at all measured positions. Again the all-pole model was less sensitive to variation in positions.

(e) As expected, by increasing the order of the all-pole model, a significant improvement in both time- and frequency-domain performance was achieved, as can be observed in Table 2. However, there is an optimum order for the all-pole model, as was discussed in section 2.6.

(f) It can be observed that proportionally better performance was achieved by the lower order all-pole models than by the higher order ones. Again, this conclusion applies

for orders of magnitude up to the optimum all-pole model value, as was previously discussed.

These results are confirmed, for the larger acoustic enclosure (Room 2), in Table 4. Clearly, the larger room requires higher order approximation models, and the all-zero model again appears inefficient and more sensitive to position variation. The sensitivity of the all-zero model to source/receiver placement variations can be explained by the nature of room acoustics. Transfer function zeros result from local cancellations of multi-path sound components which are easily disturbed by slight changes in source/receiver positions.

3.4. REAL ROOM RESPONSE STUDY

The procedure followed for the simulated room study was also applied to measurements of a real acoustic enclosure. For these tests, the realization of the inverse filter (see equation (16)) was achieved in real time by employing a special purpose DSP microprocessor (Texas TMS 32020) running in conjunction with A/D and D/A data conversion modules. The purpose of these tests was to verify the results of the previous study and examine the feasibility of practical application for these findings.

The real-time actual room study was carried out for the all-pole room model. This was a practical decision, since current real-time processor technology does not allow realization of the significantly longer all-zero model inverse filter [12]. This limitation highlights the efficiency of the all-pole approximation.

The procedure for the real-time tests was as follows: (a) the impulse response of an enclosure was measured (see section 3.2); (b) the all-pole mode approximation inverse filter was estimated (see section 2.4); (c) the filter coefficients were downloaded to the DSP microprocessor; (d) the measurement of the filtering operation was obtained by the computer, giving the numerical output of equations (16); (e) the results were analyzed by applying the performance criteria (section 3.1) in a way similar to that followed for the simulation tests.

For the real-time tests, a small semi-anechoic room was employed. This chamber was rectangular, of dimensions $1{\cdot}67 \times 2{\cdot}08 \times 2{\cdot}22$ m, and had an approximate reverberation time of 0·06 s. The mean typical value of ambient noise sound pressure level inside the chamber was in the order of 35 dB(A). For the impulse response measurements, the set-up shown in Figure 11 was employed, as described in section 3.2.

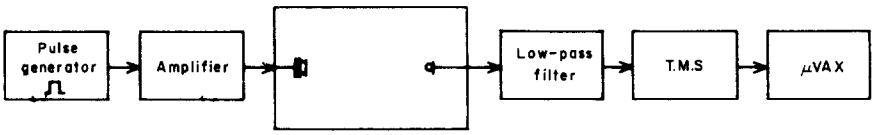

Figure 11. Block diagram of equipment arrangement for real-time inverse filter implementation.

Figure 12(a) shows the measured room impulse response and Figure 12(b) shows the all-pole model ($p = 100$) inverse filtered response, as was acquired in real time. The spectra corresponding to the above signals are shown in Figure 13. A reduction of the effects of reverberation is visible in the figures, as was expected from the simulation studies. The measured improvement for these real-time tests is given in Table 5. These results indicate that the findings of the simulated tests are valid in practical cases, and that real-time implementation of all-pole model inverse filtering can achieve the predicted improvements in the processed signal.

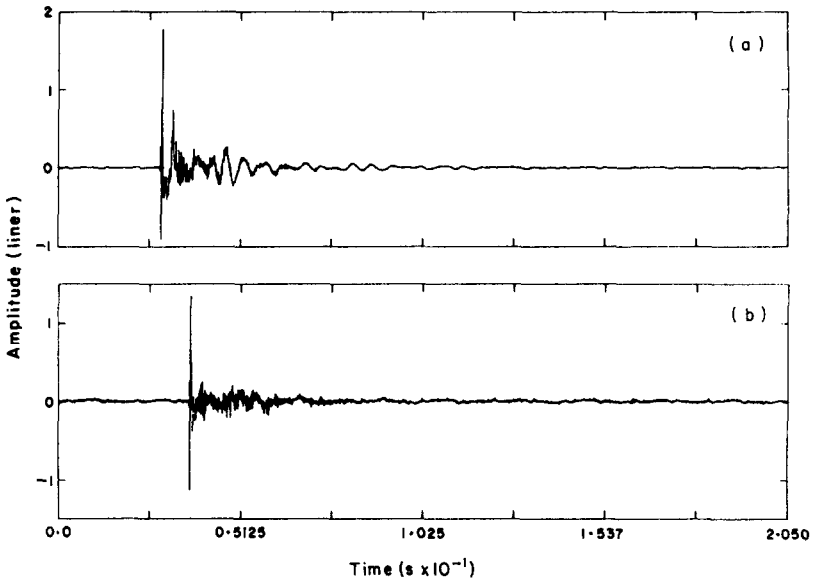

Figure 12. Results of real-time inverse filtering. (a) Impulse response measurement; (b) all-pole model inverse filter ($p = 100$) residue.

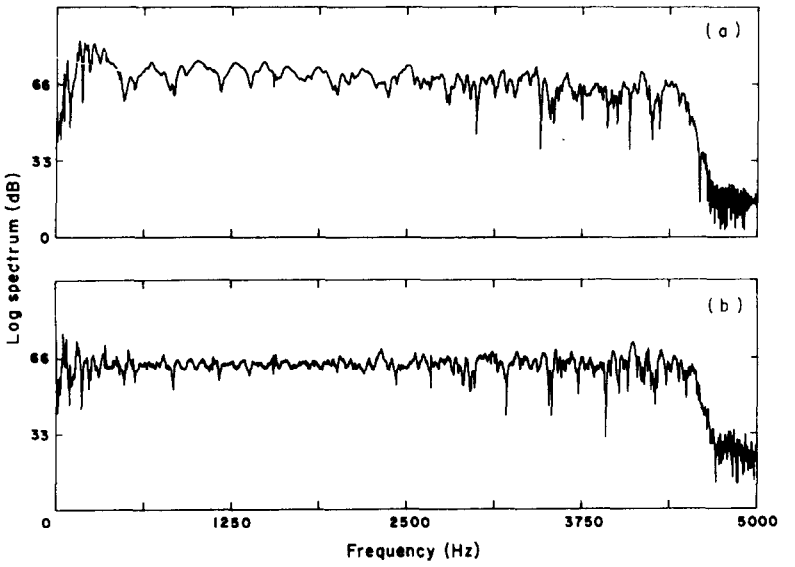

Figure 13. Frequency domain results for real-time inverse filtering. (a) Spectrum of room response; (b) spectrum after inverse filtering ($p = 100$).

4. CONCLUSIONS

The work presented in the previous sections has been addressed to specific problems arising in signal processing applications of room acoustics. The feasibility of employing the transfer function all-pole and all-zero approximations was examined in order to provide alternative approaches to two main problems associated with such functions; namely their high arithmetic order, and their sensitivity and dependence on specific source/receiver placement.

TABLE 5

Time- and frequency-domain performance of the all-pole model inverse filter, implemented in real-time

Order of model	Time, I_p (dB)	Frequency J_{p500} (dB)	Frequency J_p (dB)
$p = 100$	1·94	0·175	0·68

It was shown that such approximations can describe room transfer functions to a satisfactory degree. Furthermore, it was found that all-pole room models present significant advantages over all-zero room models, mainly because they are functions of lower order (by up to a factor of 40 approximately) and they are less sensitive to source/receiver placement variation. All-zero models can achieve, at best, an exact modeling of room spectra within an approximation error which is an all-pass component and thus generates significant time-domain model mismatch. Such all-zero models were found to be very sensitive to changes in source/receiver placement, as was expected from consideration of the physical aspects of room acoustics.

The performance of all-pole models was found to improve as the predictor model order increased. Furthermore, it was shown that an optimum order of all-pole model existed for a given room transfer function. It is also important to note that such an optimum all-pole approximation still allows a significant reduction in the order of the room transfer function. These results indicate that in many signal processing applications of room acoustics it may be both sufficient and more efficient to manipulate all-pole model coefficients rather than the raw room response data.

ACKNOWLEDGMENTS

The authors wish to thank Professor G. Kokkinakis of the Wire Communications Laboratory for his support and assistance. They also acknowledge the help of Mr E. Kyriakis-Bitzaros with the microprocessor programming.

REFERENCES

1. S. T. NEELY and J. B. ALLEN 1979 *Journal of the Acoustical Society of America* **66**, 165–169. Invertibility of a room impulse response.
2. J. MOURJOPOULOS 1985 *Journal of Sound and Vibration* **102**, 217–228. On the variation and invertibility of room impulse response functions.
3. J. B. ALLEN, D. A. BERKLEY and J. BLAUERT 1977 *Journal of the Acoustical Society of America* **62**(4), 912–915. Multimicrophone signal processing technique to remove reverberation from speech signals.
4. P. J. BLOOM and G. D. CAIN 1981 *Institute of Acoustics Spring Conference*, 22–24. Speech dereverberation: performance of signal processing algorithms and their effect on intelligibility.
5. S. J. ELLIOTT and P. A. NELSON 1989 *Journal of the Audio Engineering Society* **37**(11), 899–907. Multiple-point equalization in a room using adaptive digital filters.
6. J. MAKHOUL 1975 *Proceedings of the Institution of Electrical and Electronic Engineers* **63**, 561–580. Linear prediction: a tutorial review.
7. P. M. CLARKSON and J. K. HAMMOND 1986 *Adaptive Signal Processing Course Notes.* University of Southampton.
8. A. V. OPPENHEIM and R. W. SCHAFER 1975 *Digital Signal Processing.* Englewood Cliffs, New Jersey: Prentice-Hall.

9. J. MAKHOUL 1973 *Institution of Electrical and Electronic Engineers Transactions, Audio Electroacoustics* **AU-21**, 140-148. Spectral analysis of speech by linear prediction.
10. H. AKAIKE 1974 *Institution of Electrical and Electronic Engineers Transactions, Automatic Control* **AC-19**, 716-723. A new look at statistical model identification.
11. J. B. ALLEN and D. A. BERKLEY 1973 *Journal of the Acoustical Society of America* **66**, 943-950. Image method for efficiency simulating small-room acoustics.
12. J. MOURJOPOULOS 1988 *84th Audio Engineering Society Conference Abstracts* **36**, 384-388. Digital equalization methods for audio systems.