Proceedings of The Institute of Acoustics

# CEPSTRAL METHODS APPLIED TO THE ANALYSIS OF ROOM IMPULSE RESPONSE

J K Hammond and J Mourjopoulos*
Institute of Sound and Vibration Research, Southampton University

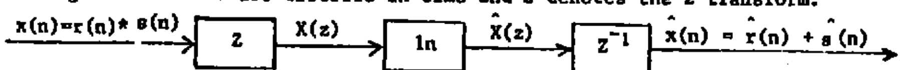
## 1. Introduction

Sounds generated in a room are distorted by reflections from boundaries, and if the room response is modelled by an impulse response function then the perceived (measured) signal is the convolution of the source signal and the room impulse response corresponding to the particular source/receiver positions. In some instances it may be desirable to remove the effects of the room from recorded data, and it is with this ultimate objective in mind that the problems of choosing appropriate models for room response are discussed.

The complex cepstrum is one approach to deconvolution and in addition offers an interesting interpretation of room characteristics. This paper will:
(a) Briefly review cepstral methods
(b) Outline the basis of the use of cepstral analysis in the treatment of room response including the problem of signal recovery.
(c) Present some results relating to a real room.

## 2. Cepstral methods - a brief review

The figure below indicates the steps in forming the complex cepstrum of a real valued signal x (assumed to be the convolution of two other signals r and s). The signals involved are discrete in time and Z denotes the z-transform.




The important feature is that the cepstrum of two convolved signals is the sum of their cepstra (if the complex logarithm is defined appropriately). The pioneering paper |1| dealing with power cepstra still remains an illuminating and informative source, though the usual description now uses the z-transform |2|. Computation is achieved using the discrete Fourier transform. Having mapped from convolution to addition, the signal $\hat{x}$ (real valued) may be processed to recover one of the components (say $\hat{s}$) by filtering followed by the appropriate inverse operations. To obtain a real valued cepstrum $\hat{x}$, its transform $\hat{X}$ must have the usual properties of the Fourier transform of a real signal. This, in turn, implies that the phase of X (note: $\hat{X}=\ell nX=\ell n|X|+j\arg X$) should be continuous and odd. This removes any ambiguity in the definition of the complex logarithm; indeed, it is the problem of ensuring that the phase satisfies these conditions that has lead to the development of 'phase unwrapping' algorithms |3|. This constitutes the main computational difficulty in cepstral analysis, though sidestepping the problem by factorising the transform has been suggested |4|.

---

* Financial support from J.S.R.U. Cheltenham is gratefully acknowledged.

Applications of the cepstrum include seismology |3|, speech processing |5|, loudspeaker assessment |6|, room response |7| and the determination of acoustic impedance |8|.

## 3. Room impulse response and the cepstrum

### 3.1 Basic principles

We emphasise that we are not concerned with estimating impulse response functions of rooms from general recorded data. Rather, signals representing room impulse response (either synthetic or experimentally measured) are taken as the starting point. The principles underlying the study are given in |10| and |2| and are summarised here.

Let h(n) denote an impulse response function with transform H(z) having mixed phase. H(z) can be written (a) H(z)=Hmin(z)Hmax(z) or (b) H(z)=Heq(z)Hap(z). In the first case Hmin(z) denotes that part of H whose poles and zeros lie within $|z|=1$, and Hmax(z) denotes that part with zeros outside $|z|=1$. Hmin is minimum phase and Hmax is maximum phase. In the second case H is rewritten as the product of an all pass filter Hap(z) and Heq(z), which is another minimum phase function that has the same spectral magnitude as H(z).

From (a) the time domain equivalent is h(n)=hmin(n)*hmax(n) so that its cepstrum $\hat{h}(n)=\hat{h}min(n)+\hat{h}max(n)$, where $\hat{h}min(n)=0$ for $n<0$ and $\hat{h}max(n)=0$ for $n>0$. By only retaining the cepstrum for positive/negative values of n the minimum/maximum phase components are retained. The minimum phase component has a causal inverse and the maximum phase component an acausal inverse.

From (b) the time domain equivalent is h(n)=heq(n)*hap(n) so that the cepstrum is $\hat{h}(n)=\hat{h}eq(n)+\hat{h}ap(n)$. The cepstrum of an all pass sequence has anti-symmetric positive and negative time components and so $\hat{h}eq(n)$ can be obtained from

$$\hat{h}eq(n) = \begin{cases} 0 & n<0 \\ \hat{h}(o) & n=0 \\ \hat{h}(n)+\hat{h}(-n) & n>0 \end{cases}$$

The sequence heq(n) is minimum phase with the same spectral magnitude as h(n) and has a causal inverse. The all pass filter has an inverse with an acausal component. In the following sections we shall discuss the application of both (a) and (b) to room impulse responses, noting that Neely and Allen |7| address the problem of invertibility of room impulse response based on the interpretation (b) above to dereverberate data processed through synthetic |9| room responses. Their work demonstrates the feasibility of removing Heq(z).

### 3.2 Data acquisition and computation methods

The impulse responses studied were either synthesised or pre-recorded. The discussion below relates to a response obtained by impulsive excitation (sampled at 10 KHz). Programs for computation of power and complex cepstra, group delay etc were developed on the ISVR PDP 11-50 computer. Currently the largest complex cepstrum calculation is limited to 1024 points. This requires the use of exponential weighting on some responses, usually corresponding to larger rooms.

### 3.3 Separation of minimum and maximum phase components

Most room impulse responses are mixed-phase signals |7| so room response h(n) is the convolution of hmin(n) and hmax(n). hmin(n) can be considered as the portion of room reflection energy that dies away in a short time after excitation; whereas hmax(n) can be interpreted as the portion of room response energy that arrives at later instances.

Figs 1 and 2 show the impulse and frequency responses of an absorbent small room (including exponential weighting). The complex cepstrum is shown in fig 3. Fig 4 shows the (negative of) group delay of the response, (positive spikes denotes zeros just inside the unit circle). Group delay indicates the relative delay introduced by the room on the various components of the signal's spectrum |11|.

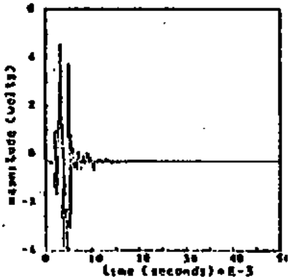



Fig 1

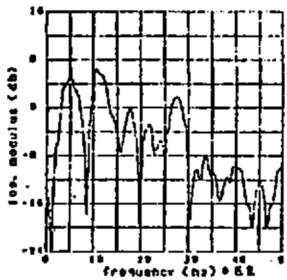



Fig 2

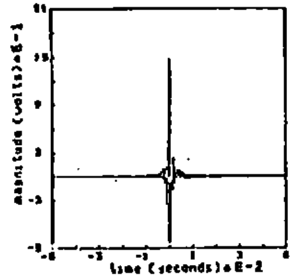



Fig 3

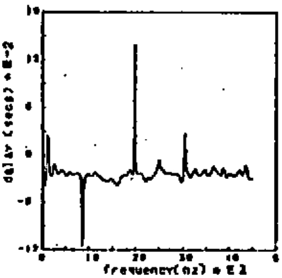



Fig 4

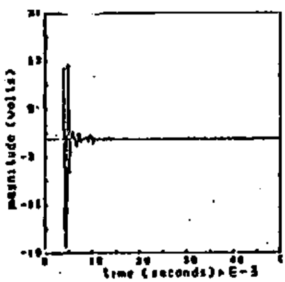



Fig 5

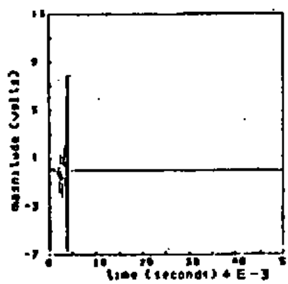



Fig 6

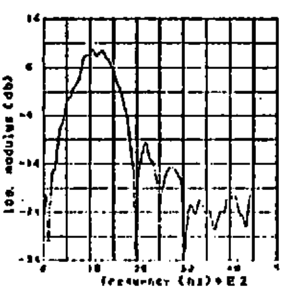



Fig 7

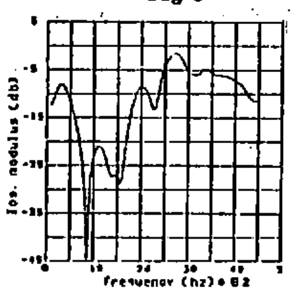



Fig 8

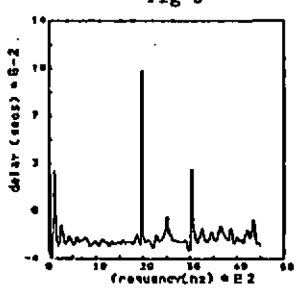



Fig 9

The room response was then separated into minimum and maximum-phase parts (figs 5,6) whose frequency responses are shown in figs 7,8. The frequency plots indicate that the transmission characteristics of the particular room were dominated by the minimum-phase response, except for a minimum near 800 Hz. The group delays of each component reinforce this conclusion. Fig 9 shows the (negative of) group delay of the minimum phase part.

### 3.4 Separation into equivalent minimum phase and all pass components

We now consider the second form i.e. h(n)=heq(n)*hap(n). The room response (fig 1 ) was separated into "equivalent" minimum-phase and all-pass components (figs 10, 11). The heq(n) component has a frequency response magnitude identical to the original room response (fig 2). The advantage of this type of separation in dereverberation work is that the resulting response heq(n) can be directly inverted and the effect of all pass component should not affect the signal's amplitude spectrum. However, subjective assessment indicates that the remaining phase distortion may still be perceived |7|.

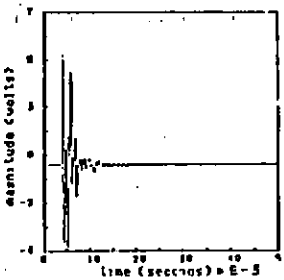



Fig 10

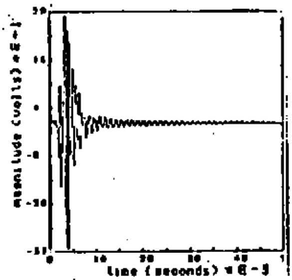



Fig 11

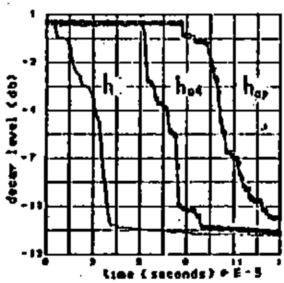



Fig 12

Finally, the energy decay rates of the original response, heq(n) and hap(n) were derived in the form of reverberation curves, using the 'integrated impulse' technique |12|. These curves (fig 12 ) confirm that the reverberation time corresponding to the "equivalent" minimum-phase response is shorter than the original room reverberation, and in addition has very high initial decay rate, whereas the reverberation time corresponding to the all-pass response component is longer.

## References


1. B P BOGERT et al. Time Series. Ed M Rosenblatt. J Wiley 1963.
2. A V OPPENHEIM & R W SCHAFER. Digital Signal Processing. Prentice Hall 1975.
3. J M TRIBOLET. Seismic Appl. of Homomorphic Signal Processing. Prentice Hall 1979.
4. K STEIGLITZ & B DICKINSON. Proc IEEE Int Conf ASSP Hartford Conn 1977.
5. A V OPPENHEIM (Ed). Appl. of Digital Signal Processing. Prentice Hall 1978.
6. J M KATES. Proc IEEE Int Conference ASSP Hartford Conn 1977.
7. S T NEELY & J B ALLEN. J. Acoust Soc Am Vol 66 No 1 July 1979.
8. S BOLTON & E GOLD. Proc IOA Melbourne July 1980.
9. J B ALLEN & D A BENTLEY. J. Acoust Soc Am Vol 65 No 4 April 1979.
10. A V OPPENHEIM, G E KOPEC & J M TRIBOLET. IEEE Trans ASSP August 1976.
11. H KUTTRUFF. Room Acoustics. Applied Science Pub. pg 69 1973.
12. M R SCHROEDER. J. Acoust Soc Am 37, 1965.