THE "BACH" EXPERIENCE
BRING A CONCERT HOME

Sattwik Basu, Saarish Kareer

University of Rochester, Department of Electrical and Computer Engineering

ABSTRACT

Inverse filtering of rooms for correcting poor frequency response is a well-researched topic in acoustics and signal processing. This paper describes the design of an audio playback system that uses room correction and then impresses the sonic qualities of a superior listening environment like a renowned concert hall to enable music lovers to have an authentic aural experience of those environments.

The first step involved characterization of rooms and using MINT [1] to find exact room inverses. Following this, binaural heads were used to measure concert halls impulse responses at various seating positions.

We observed that recordings of speech and music when filtered with the room inverses have a flat frequency response and also sound very dry as though they were being played in anechoic chambers. These anechoic sounding audio files were then convolved with the concert hall impulse responses to produce authentic experiences of listening to music in those concert halls. We were also able to validate that MINT, although simple is applicable to a wide variety of rooms and can be used for enhancing listening experiences in a cost effective way.

***Index Terms*— Impulse Response, MIMO, Inverse Filtering, Toeplitz matrix.**

1. INTRODUCTION

Attending concerts and listening to music in acoustically superior environments is something that only a handful of music enthusiasts get to experience. Audio playback through speakers in everyday scenarios is significantly affected by the room in which the audio is being played, the materials used in the room's construction, its dimensions, objects present, ambient noise and even by the position of the loudspeaker(s) itself. Thus, music often loses many of the sonic qualities that enhance the intelligibility of various sounds and our overall listening experience. Our aim is to develop a system that eliminates the effect of the room and reproduces the sonic qualities of various listening environments to enable any music lover to have an authentic aural experience of those environments, be it concert halls, studios or live music cafes at the comfort of their homes.

2. METHOD OUTLINE

1. ACOUSTICAL CHARACTERISATION OF ROOMS

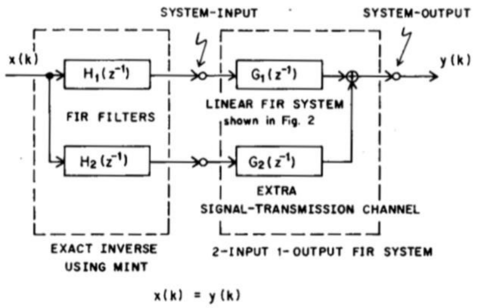
The impulse response (IR) of a room tells us about the important attributes of its acoustical behaviour [2]. We have used IRs to characterise both, the room in which the audio will be played back and the desired listening environment. For this, we have used Room EQ Wizard to compute the IRs, frequency response and Energy Decay Curves of selected spaces. The specific signals and equipment used are:

1. Logarithmic swept sines of 3/5 seconds duration in the range 20-20,000 Hz, for the room/halls respectively.
2. The Behringer ECM 8000 omnidirectional condenser microphone for rooms and a Neumann KU 100 Binaural Dummy Head for capture at various locations in the desired listening environment.
3. A pair of 75-80 dBSPL calibrated Alesis M1 Active 520 Loudspeakers, placed at desired locations.

2. INVERSE FILTERING BY MINT ALGORITHM

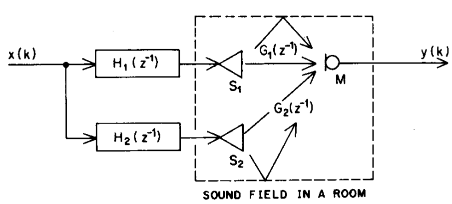
A room is known to act like a LTI filter on acoustical signals. Hence, audio playback in a room is highly affected by its acoustical properties. Rooms being non-minimum phase systems [1] make it challenging to find exact inverses that can help us reduce distortions and excessive reverberation.

This step primarily deals with finding inverse filters for the listening room. To do this we have decided to use a method known as Multiple Input/Multiple Output Inverse Theorem [1]. This method is vastly superior compared to older methods which could only produce approximate inverses of rooms from non-square impulse response matrices that resulted due to the non-minimum phase behaviour of rooms. MINT, however results in square invertible matrices and hence gives us exact inverses.



The MINT Theorem was described by Masato Miyoshi and Yutaka Kaneda in their 1988 publication titled “Inverse Filtering of Room Acoustics” [1]. This approach, relies on multiple signal transmission channels produced by two or more loudspeakers/microphones as shown in the figures above. The signal transmission channels (room response) are denoted by G1 and G2. When G1 and G2 have no common zero, i.e they are relatively prime, the system has two exact inverses (FIR filters) H1 and H2.

The practical setup: For inverse filtering, we required two identical flat response loudspeakers and a measurement microphone. The microphone is positioned at a chosen location (M) in the room of choice and the two loudspeakers are carefully aligned to face the microphone. To find G1(t) we turn on the first loudspeaker S1 while keeping S2 turned off and record the impulse response through the microphone at M. We then repeat the process to find G2(t).



We form the following Toeplitz matrix G containing the IRs G1 and G2. Mathematically in the time domain, this represents the discrete convolution matrix of G(t), which when multiplied by a matrix H(t) results in a delta matrix D(t).

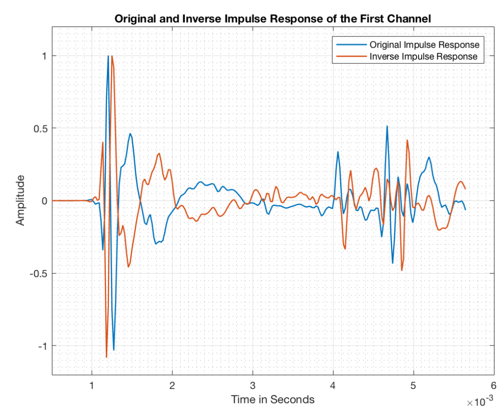
$$L+1\ \begin{bmatrix}1\\0\\0\\\vdots\\\vdots\\0\end{bmatrix} = \begin{bmatrix} g_1(0) & & g_2(0) & \\ g_1(1) & \ddots\ 0 & g_2(1) & \ddots\ 0 \\ \vdots & & \vdots & \\ & g_1(0) & g_2(n) & \\ g_1(m) & g_1(1) & & g_2(0) \\ & \vdots & & g_2(1) \\ 0 & & 0\ \ddots & \vdots \\ & g_1(m) & & g_2(n) \end{bmatrix} \begin{bmatrix} h_1(0)\\ h_1(1)\\ \vdots\\ h_1(i)\\ h_2(0)\\ \vdots\\ h_2(j)\end{bmatrix}, \quad \begin{bmatrix}H_1\\H_2\end{bmatrix} = [G_1 \quad G_2]^{-1} D.$$

$$\longleftarrow i+1 \longrightarrow \quad \longleftarrow j+1 \longrightarrow$$

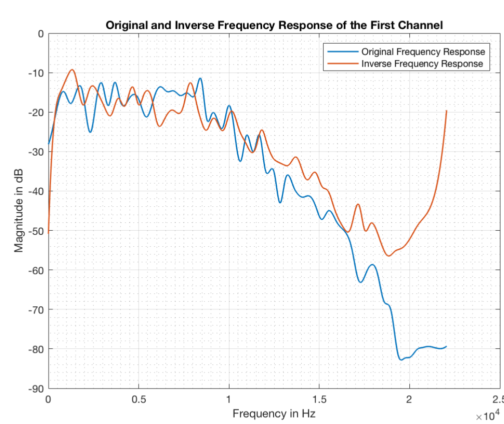
The exact inverses H1 and H2 can then be split up and computed using inv(G) and the Delta Matrix D. Now it is possible to reproduce audio at position M without any distortion, by sending each respective inverse to the source speaker in the signal flow.

Results: One signal transmission channel G1 was studied and plotted. Depending on the type of music we wanted to play back, a Band Pass filter (3rd order Butterworth)[1] was used to initially filter this impulse response, to help get rid of the ambient noise. Frequency range was selected to be 95 Hz to 3150 Hz for speech and this was not used for music at all. Next, truncation of the IR was done, so that 400-500 samples could be selected based on the room type, to account for the early reflections without entering the noise floor.

This is a comparative analysis of the impulse response G1(t) and its corresponding inverse H1(t).



To verify this result, we computed the frequency response G1(z) and H1(z), and found that the inverse 1/G1(z) is indeed equal to H1(z). The same steps were followed for the other signal transmission channel.

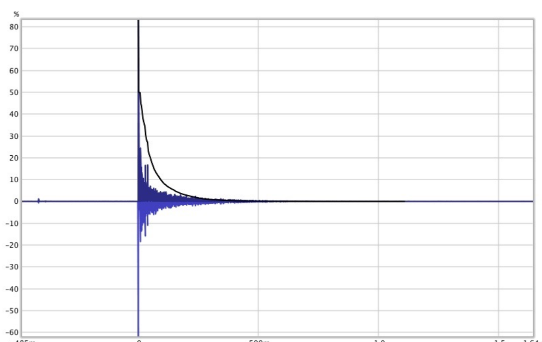


3. MODIFICATION OF INPUT USING INVERSE FILTER

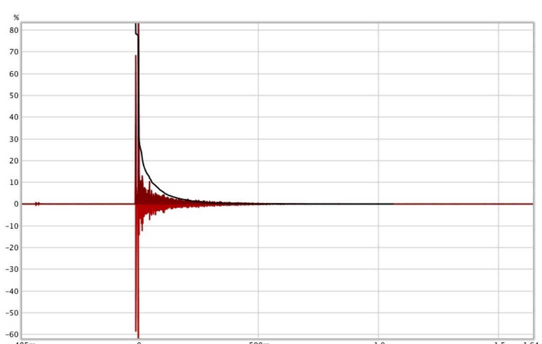
This step involved the convolution of the individual channel filter with the input audio file, be it music or speech. We split up the audio into L/R channels to facilitate effective anechoic playback.

4. REPRODUCTION OF LISTENING ENVIRONMENT

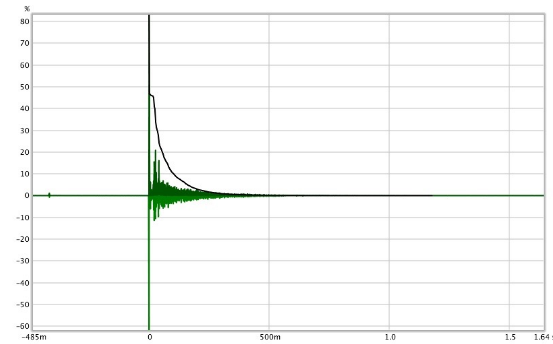
The final step was to characterise different listening environments by impulse response measurement. We chose the Strong Auditorium at the University of Rochester for our project. Again the tools used were Room EQ wizard, the binaural dummy head and one of the Alesis M1 Active 520 Loudspeaker placed at the centre of the stage. We chose 3 different locations in the audience seating area for binaural capture and following are the plots of the left ears impulse response and energy decay curves for these locations, in Seconds vs dB full scale:



LEFT BACK



CENTRE BACK



RIGHT BACK

Finally, we convolved the L/R impulse at each position with the respective channels of the inverse filtered audio file from the previous step to complete our process of authentic reproduction of the listening environment in the given room. This could be verified by listening at the chosen position (M) for all of these configurations.

3. RESULTS

The effect of the room was significantly reduced at a single listening position and the characteristics of the desired listening environment were reproduced. The results were analysed for a speech signal using appropriate filtering and truncating of the inverse filter for both channels. The waveforms of the original signal, inverted signal using one channel and the final output as it would sound at the centre back location of the auditorium are plotted below.

The output of the inverse filtered audio indeed sounded anechoic as desired. This sound can be described as dry, direct and compact. When the impulse response of the concert hall was impressed, the audio playback had an authentic feel of the chosen concert hall and was spatialized owing to the binaural impulse responses.

4. FUTURE WORK

1**.** The next step would be to characterize rooms at multiple locations using more channels and give the listener the freedom to choose from different seating positions in concert halls.

2. We would also like to research and apply more advanced room correction algorithms for better tuning of any speaker configuration and enhanced audio playback.

3. The impulse response characterises the room positionally by giving us the reverberation, absorption, reflection & frequency response parameters effectively. In fact the T_{60}, can also be found out using the energy decay curves and tested by the known values of certain rooms. We can go use this method and characterise rooms/halls at the Eastman School of Music, or larger performance areas too.

4. Using reflections and calculated room characteristics, we can also to enable superior sound from the existing speakers.

5. REFERENCES

[1] M. Miyoshi and Y. Kaneda, "Inverse Filtering of Room Acoustics" (1988), IEEE Transactions on Acoustics, Speech and Signal Processing

[2] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a swept sine technique", (2000), 108thAES Convention, Paris, France