Multipoint Equalization with the Condition Number

Bruno Korst-Fagundes, Jun Xie and Martin Snelgrove**

*Department of Electronics, Carleton University, Ottawa, Canada K1S 5B6
bkf@doe.carleton.ca
snelgar@doe.carleton.ca

Abstract

Equalizing the acoustical frequency response of a room at several points, using several loudspeakers, involves taking the inverse of a matrix frequency response. An exact inverse is generally not practical because the room transfer-function matrix is often nearly singular at some frequencies, causing high-Q equalizer peaks and making the room response unacceptable at points other than those where measurements are taken. We show how to diagnose and deal with the problem. The results presented were acquired using measurements taken in a car.

I. Introduction

A simple mathematical view of equalization is that, given a loudspeaker driving a room and an ear with an overall transfer function $H(s)$, one would put a gain $H^{-1}(s)$ in front of the loudspeaker to get an overall flat gain. This is impractical for a number of reasons:

1. Causality: the room response includes several milliseconds of pure delay, so $H^{-1}(s)$ would be non-causal.
2. Notches: if $H(s)$ has a notch, perhaps because of a hard acoustic reflection, it cannot be inverted. Even if the notch is not infinitely deep, it may not be practical to invert it.
3. Spatial variation: Correcting the response from a source to one point in a room is no guarantee that it is acceptable at other points[1]. If, for example, a deep notch occurs at one location only, correcting it will cause a very audible sharp peak elsewhere.
4. Order: Rooms have impulse responses thousands or tens of thousands of samples long, making any inverse filter a very high-order device [2]. High order brings with it a heavy computing load, high accuracy requirements, and very slow or difficult adjustment to changes in room response.

The "causality" problem can be resolved by accepting a "pseudo-inverse" H_{ps}^{-1} such that HH_{ps}^{-1} approximates a delay of some τ seconds, though now there is the practical problem of choosing a good estimate of τ. This is a simple application of psychoacoustics, in the sense that the user may not much mind a flat delay.

One way to address the "notch" problem is by putting an upper limit on the magnitude of H_{ps}^{-1} , on the principle that a small number of narrow notches are probably acceptable. This is a more dubious application of psychoacoustics, since there are obviously signals that will sound wrong.

The "spatial variation" problem can be addressed by allowing N loudspeakers, so that there are N different transfer functions available to each point. Now if the room response is measured at M microphone locations we have an MxN matrix of transfer functions $H(s)$. This may help solve the "notch" problem in that not all responses will generally have notches at the same place, but adds new ways for the inverse to misbehave.

This paper is concerned with analyzing what may happen and suggesting what do do about it. We use experimental data to make our points.

The "order" problem can be addressed with powerful, high-precision, DSP chips or by use of psychoacoustic or acoustic knowledge to find acceptable reduced-order pseudo-inverses.

In general, the idea of a mathematical inverse has to be replaced with that of a pseudo-inverse, chosen by applying knowledge of the physical problem and of the way in which sound is perceived.

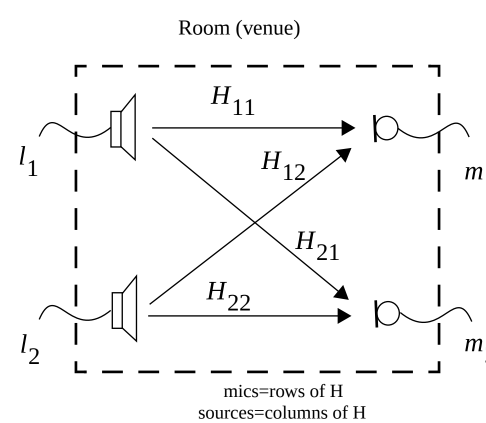
II. Notation and Assumptions

We define the room response as the vector $\boldsymbol{m}(s)$ (responses measured at the microphones), which in the frequency domain will be formed as the product of a matrix $\boldsymbol{H}(s)$ of transfer and a vector $\boldsymbol{l}(s)$ of sources (loudspeakers).

$$\boldsymbol{m} = \boldsymbol{H} \cdot \boldsymbol{l} \qquad \text{(Eq. 1)}$$

Fig. 1 shows a simple setup, illustrating the procedure for data collection.

Fig. 1: Experimental setup (example)



In order to equalize the received signal for each of the receiving positions, we could make use of a set of gains which approximate the inverse of the transfer function $H(s)$

presented at the 38th Midwest Symposium on Circuits and Systems (MWSCAS) - Rio de Janeiro - August 1995

for each frequency, avoiding the equalization with the direct inverse of the transfer function matrix. This set of gains will be referred to as “desired gains ***d***” through this paper.

The equalizer will then drive *N* loudspeakers with an approximate inverse of the room response:

$$l = \boldsymbol{H}_{ps}^{-1} \cdot \boldsymbol{d} \cdot u \qquad \text{(Eq. 2)}$$

Where *u* is a scalar representing the signal input to the equalizer.

The microphone positions are assumed to represent room response nearby, but this requires that they be placed properly and the assumption is particularly weak at high frequencies where spatial coherence is low [3]. Roughly, the signal at a point is fairly representative of that within a moderate fraction of a wavelength [4]. A practical consequence of this is that these techniques are intended for equalizing at low to medium frequencies: by 300Hz the wavelength is down to about a metre and a great many microphones would be needed to measure the behaviour of a large room.

At high frequencies it is more appropriate to use the directivity of drivers to control spatial variation, in which case DSP is used to invert the responses of drivers only rather than their interactions.

The transfer function matrix is a function of frequency, or equivalently a matrix of functions of frequency. We usually choose to sample the response at a large number of frequencies, and think of the problem as one of inverting a large number of matrices of numbers — one matrix for each frequency. Since frequency responses have magnitude and phase, we are inverting a large number of complex-valued matrices. In practice some of those will be difficult to invert, and at the corresponding frequencies we expect to have difficulties.

III. Eigenvalues and Eigenvectors

We make extensive use of the eigenvalue-eigenvector decomposition of the room transfer function. We do this separately at each frequency sample. There are several ways in which ***H*** can be difficult to invert, with different physical causes and suggesting different treatment. The eigen-decomposition can be used to identify these nicely; we’ll start by looking at the decomposition and then continue by demonstrating special cases that correspond to real problems.

This decomposition rewrites *H* in the form

$$\boldsymbol{H} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{-1} \qquad \text{(Eq. 3)}$$

where ***D*** is a diagonal matrix whose entries are “eigenvalues” and ***V*** is a matrix whose i^{th} column is the eigenvector corresponding to eigenvalue *i*. If the “input” to ***H*** is the i^{th} eigenvector, then the “output” is the same eigenvector scaled by the i^{th} eigenvalue.

If an eigenvalue is zero, the matrix can’t be inverted because there is an input (the corresponding eigenvector) which produces zero output; and obviously once the signal has gone down to zero it can’t be built back up again. There is in fact no input whatever that can produce that output.

After analyzing the eigenvalues and eigenvectors of matrix ***H***, we isolate the eigenvector associated with the smallest eigenvalue by reordering ***V*** and ***D*** so that the smallest eigenvalue comes last; then we apply the Gram-Schmidt procedure to orthogonalize and normalize ***V***. The result is a matrix that can be used to identify “good” and “bad” components of the “desired gain”. By removing the “bad” component we can remove the need for high gains.

The results of this analysis will be used to calculate the “desired gains” for the equalizer.

IV. A simple example

Suppose that a system with three loudspeakers and three microphones has, at some frequency ω_j,

$$\boldsymbol{H}(\omega_j) = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix} \qquad \text{(Eq. 4)}$$

This can’t be inverted because the rows aren’t linearly independent: row 3 is the sum of the first two. Physically, this says that no matter what loudspeaker (column of ***H***) is used, microphone 3 always receives a signal that is the sum of those at microphones 1 and 2. If the desired signal were

$$\boldsymbol{d} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \text{(Eq. 5)}$$

which means that equal gains are desired to all microphones, then there would clearly be a problem: the desired pattern does not have mic 3 receiving a signal that is the sum of those at mics 1 and 2.

On the other hand, the signal

$$\hat{\boldsymbol{d}} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \qquad \text{(Eq. 6)}$$

would be easy enough to produce and probably just as good from a practical point of view. The music will be louder at mic 3 than at the other two, but as long as that is forced to be true at all frequencies then all three frequency responses will be flat, and that is the main objective.

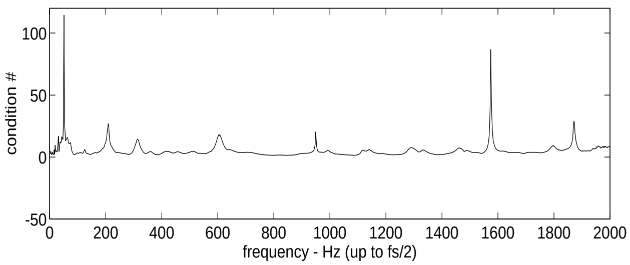
We therefore want a systematic way to identify frequencies at which there will be a problem, a way to find a “good enough” $\hat{\boldsymbol{d}}$, and a way to force the same constraints to apply at all frequencies.

V. Condition number

Numerical analysts define the “condition number” of a matrix as a measure of how close it is to not having an inverse. It is the ratio of the magnitudes of the largest and smallest eigenvalues of the matrix; a small eigenvalue makes for large condition number and large gains in the inverse, which is not feasible to implement in practical terms.

We use the condition number to identify the frequencies that will cause problems in finding the inverse of the transfer function matrix $\boldsymbol{H}$. Fig. 2 shows the peaks of the condition number to be dealt with for each frequency, from experimental data collected using four loudspeakers and four microphones.

Fig. 2: Condition Number versus frequency



This plot clearly shows that there are big problems near 50 and 1600Hz, and smaller problems at a handful of other frequencies.

The condition number plot gives a simple way to identify the most critical frequencies.

VI. Finding the “good” and “bad” components

The matrix of (Eq. 4) can be decomposed (using the MATLAB [5] “eig” command, for example) into

(Eq. 7)

$$\boldsymbol{H} = \begin{bmatrix}-0.24 & 0.74 & -0.41\\ -0.55 & -0.65 & 0.81\\ -0.79 & 0.09 & -0.41\end{bmatrix}\begin{bmatrix}15.4 & 0 & 0\\ 0 & -0.4 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}-0.41 & -0.57 & -0.73\\ 1.87 & 0.48 & -0.9\\ 1.22 & 1.22 & -1.22\end{bmatrix}$$

The first two columns of $\boldsymbol{V}$ have the property that row 3 is the sum of the first two, and represent things that can be done; the third column is something that can’t be done. As a matter of interest, the first eigenvalue also is a lot better than the second, and the corresponding eigenvalues show why: a signal that’s in-phase at all three mics is going to be easier to produce than one in which mics 1 and 2 are 180 degrees apart.

Now we want a systematic way to correct a given $\boldsymbol{d}$ to one that is easy to reproduce. A technique for doing this is to apply the Gram-Schmidt procedure to the columns of $\boldsymbol{V}$ to produce a new matrix

$$\hat{V} = \begin{bmatrix}0.75 & -0.33 & 0.58\\ -0.66 & -0.48 & 0.58\\ 0.09 & -0.81 & -0.58\end{bmatrix} \qquad \text{(Eq. 8)}$$

which converts between a representation of $\boldsymbol{d}$ in terms of what can be done with various combinations of eigenvectors and one that specifies what appears at mics. $\hat{\boldsymbol{V}}^{-1}$ in turn converts in the other direction so that if we write.

If we write then:

$$\hat{\boldsymbol{d}} = \left(\hat{\boldsymbol{V}}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\hat{\boldsymbol{V}}^{-1}\right)\boldsymbol{d} \qquad \text{(Eq. 9)}$$

we will have removed the component of $\boldsymbol{d}$ that causes trouble. In our example, a vector of ones will be changed to

$$\hat{\boldsymbol{d}} = \begin{bmatrix}0.667\\ 0.667\\ 1.333\end{bmatrix} \qquad \text{(Eq. 10)}$$

which can be reproduced. The diagonal matrix in (Eq. 9) contains zeros to null out undesirable components and ones to keep acceptable components. A similar technique could be used to select just the *un*desirable part of $\boldsymbol{d}$, if that were of interest.

Now a “pseudoinverse” of $\boldsymbol{H}$ can be obtained, which works properly provided it isn’t asked for the impossible. For an invertible matrix we could have written

$$\boldsymbol{H}^{-1} = \boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{-1} \qquad \text{(Eq. 11)}$$

where inverting $\boldsymbol{D}$ just involves taking the reciprocal of each diagonal element. For a singular matrix, though, the zero element of $\boldsymbol{D}$ blows up in $\boldsymbol{D}^{-1}$. Still, if the matrix only has to produce the right answer for combinations of the “good” eigenvectors then we could just replace the last element of $\boldsymbol{D}^{-1}$ with zero. Doing that gives us a “pseudo-inverse” of (Eq.7)

$$H_{ps}^{-1} = \boldsymbol{V}\begin{bmatrix}0.065 & 0 & 0\\ 0 & -2.5 & 0\\ 0 & 0 & 0\end{bmatrix}\boldsymbol{V}^{-1} \qquad \text{(Eq. 12)}$$

which works out to be

$$H_{ps}^{-1} = \begin{bmatrix}-0.78 & -0.96 & 1.17\\ 0.63 & 0.88 & -1.03\\ 1.67 & 2.07 & -2.53\end{bmatrix} \qquad \text{(Eq. 13)}$$

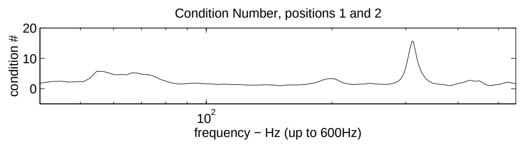
VII. Results with Experimental Data

The algorithm was applied for a set of data collected in a car, using two microphones (listening positions) and two sound sources. The graph of the condition number in figure 3 for such system shows that one major peak at around 310Hz indicates the presence of a small eigenvalue for the transfer function matrix at that frequency, which is responsible for a hard inversion for that matrix.

Figure 4 then shows each element of the direct inverse of the transfer function matrix (dotted line) plotted together with the new “pseudo” inverse (solid line). This new inverse is to be applied together with the desired vector as stated in (Eq.2).

The desired vector was taken from the eigendecomposition of the transfer function matrix at the frequency presenting a peak for the condition number, as previously described. This desired vector is a set of complex gains that will be applied to all frequencies, maintaining the overall response flat.

Fig. 3: Condition Number over frequency: two positions, two sources



Now, the two inverses of the transfer function matrix versus frequency are presented. Note the peak for the direct inverse at around 310Hz, as expected from the condition number plot, and its correction with the new "pseudo" inverse.

Fig. 4: Transfer function matrix inverses

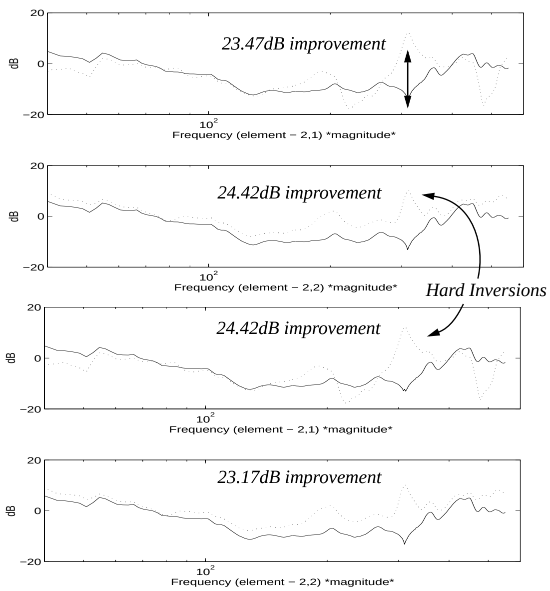
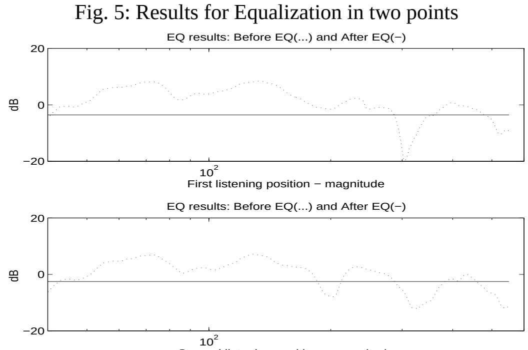


Figure 5 shows the results of the equalization, which plots the spectrum of frequencies for each microphone position, before (dotted line) and after (solid line) the desired vector and the new "pseudo" inverse have been found and applied for equalization.

Fig. 5: Results for Equalization in two points



VIII. Many constraints

Each frequency with a bad condition number implies a constraint on what can be done with the system, and that constraint must then be applied at other frequencies in order to keep the transfer functions flat. If there are too many constraints we may be in trouble, because there will be no way in which to satisfy them all.

Constraints need not contradict each other in practice, though, because they may have a common cause. For example if one microphone has lower gain than the others, getting a large signal there while having small signals elsewhere will be difficult — but difficult at all frequencies. A single constraint will do all that is required.

IX. Phase

In the practical system the gain at each frequency is complex, meaning that it has both gain and phase. We have already said that flat gain differences between microphones are acceptable, and (Section I) that simple delays are acceptable. Taken together, we are willing to allow

$$\hat{\boldsymbol{d}} = \begin{bmatrix} d_1 e^{-j\omega\tau_1} \\ d_2 e^{-j\omega\tau_2} \\ \dots \\ d_M e^{-j\omega\tau_M} \end{bmatrix} \qquad \text{(Eq. 14)}$$

and the problem is to identify acceptable gains and delays. By repeatedly "projecting out" unacceptable components at the most critical frequencies, a practical $\hat{\boldsymbol{d}}$ is obtained.

We will also be able to accept, in practice, small deviations from flat frequency responses. This means that our job is to identify an overall best fit for $\hat{\boldsymbol{d}}$, but that we can tolerate minor adustments at different frequencies.

X. Conclusions

By using matrix transfer functions it is, in principle, possible to equalize a room at several points simultaneously. Difficulties arise in inverting the matrices involved, but these can be minimized by studying the eigenstructure of the transfer function matrix and by carefully generalizing the definition of "inverse" to allow a range of acceptable solutions, for example those with flat delays and minor gain variations.

XI. References

[1] Mourjopoulos, J. - "On the Variation and Invertibility of Room Impulse Response Functions" - Journal of Sound and Vibration - vol.102(2), 1985.

[2] Neely and Allen - "Invertibility of a Room Impulse Response" - JASA, vol.66, Jul.1979

[3]Schroeder, M.R. - J.Audio Eng. Soc. vol.35, no.5, pg.299 May1987.

[4] Schroeder, M.R. - J.Audio Eng. Soc. vol35, no.5, pg.307 May1987.

[5] Matlab User's Guide- The Mathworks, Inc., 1992