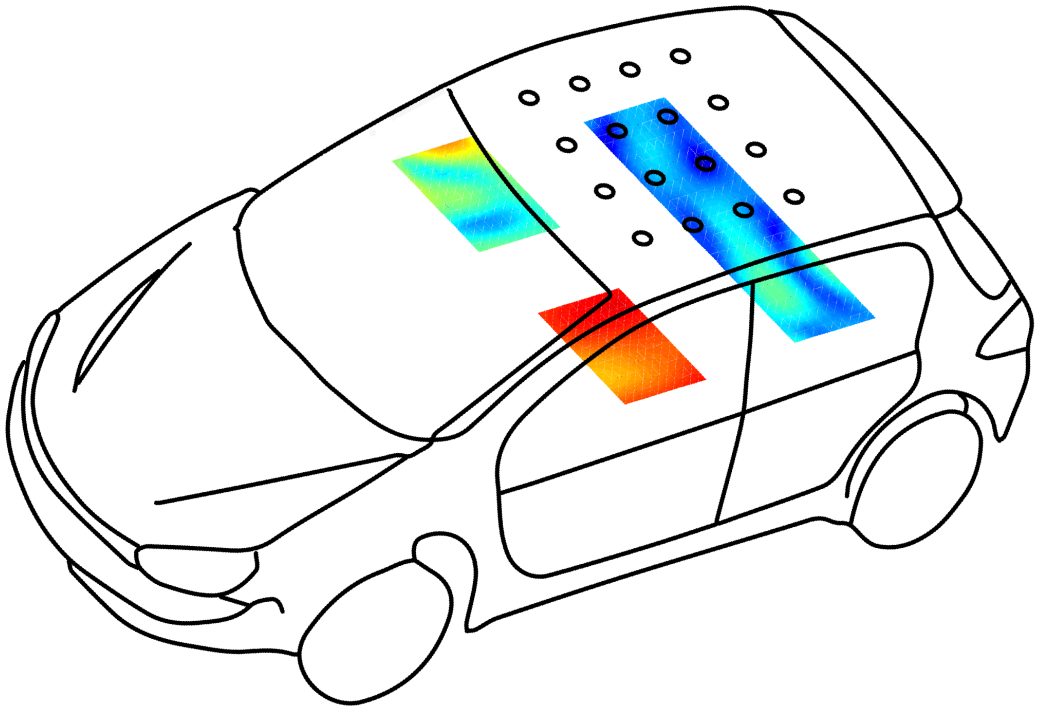
CHALMERS



Designing and Evaluating Filters for the Concept of Sound Zones

Master's Thesis in the Master's programme in Sound and Vibration

FREDRIK NORELL

Department of Civil and Environmental Engineering
Division of Applied Acoustics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2014
Master's Thesis 2014:154

Designing and Evaluating Filters for the Concept of Sound Zones

Master's Thesis 2014:154

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology
SE-41296 Göteborg
Sweden

Tel. +46-(0)31 772 1000

Reproservice / Department of Civil and Environmental Engineering
Göteborg, Sweden 2014

Designing and Evaluating Filters for the Concept of Sound Zones
Master's Thesis in the Master's programme in Sound and Vibration
Fredrik Norell
Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology

Abstract

In automotive industry, the strive to optimize the information flow directed towards the driver contributes to an increase in the use of information bearing sounds. In order to minimize distractions and maintain high comfort, there is need for a way to direct in-car audio to the intended receiver. This thesis is part of a research project aiming to investigate the use of loudspeaker arrays mounted in the ceiling of vehicles to create sound zones. The focus of the thesis is to accurately compute impulse responses (time domain), given a set of optimal complex source strengths (frequency domain) and to use the impulse responses for auralisations and objective analyses. Three frequency domain methods are evaluated: the mirroring method and two variations of homomorphic filtering. The mirroring method achieves separation between the driver seat and the passenger and back seats in the order of 17 dB and 24 dB respectively, with a variation of around 1 dB depending on which source sound is used. The delay time introduced by this method is 40 ms, i.e. small enough to be neglectable even for two-way communication. The homomorphic filtering methods achieve $7 - 13$ dB worse separation, but in return have smaller initial delay in the impulse response. The results of the thesis indicate that the current strategy for sound zones is feasible, although many challenges remain. One of these is to increase the frequency range to improve e.g. speech intelligibility.

Keywords: Digital Signal Processing, Filter Implementation, Homomorphic Filtering

Contents

Acknowledgements

I would like to thank my supervisors, Peter Mohlin at Semcon, for his genuin interest and support, and Patrik Andersson at Chalmers for the discussions and his valuable ideas. I would also like to thank Jonas Klein, the team manager at A2Zound by Semcon, for his encouraging attitude and useful input, and Stig Kleiven for his supervision during the early stages of the project, and for helpful support even after his resignation. Finally, I would like to express my gratitude to all the people at Semcon for taking me into the community and making my time during the thesis work a very memorable one.

Chapter 1

Introduction

This thesis deals with calculation of impulse responses from frequency response functions, for the concept of sound zones. Different calculation methods are evaluated subjectively, by auralisations, and objectively, by sound level differences between zones.

1.1 Background

During recent years, the automotive industry has moved towards using an increasing amount of sounds in passenger cars, conveying information about for example surrounding traffic and vehicle status [Bij 14]. With entertainment systems providing the possibility for music and movies, the potential soundscape in a modern car can be overwelming. In order to ensure safety and comfort, the in-car audio can be directed to the intended receiver. One way to do this would be to force the driver and passengers to use headphones. For obvious reasons, this is not a convenient solution. Instead, the separation of sounds can be achieved by sound zones, where sound in one zone is independent of the other zones.

Directional sound has been under study for some time, and many strategies exist. One approach is to use adjustable loudspeakers, projecting sound in the direction of the listener [Mac 06]. Other research has focused on ultrasound, using the inherent directivity of this frequency range to focus sound to the intended receiver [Wil 12], [Brb 12], [Goo 12]. Yet another way to direct sound is to use multiple sound sources and constructive and destructive interference, so called beamforming. This approach was used by Brix et al., placing multiple loudspeakers in a circle around a sound field to be controlled [Bri 12]. In another article, this specific strategy, called wave field synthesis, was compared to the simpler method of optimising con-

trasts between zones in a room [Jac 11].

Some work exist on the application of sound zones in a car compartment, to the author's knowledge mostly using beamforming, although exceptions exist [Goo 12]. Kleiven and Zinserling investigated the use of loudspeaker arrays in the ceiling, and computed optimal source strengths for the source array [Kle 12]. A real-time implementation using the the built-in sound system of a car accompanied by loudspeakers in the headrests is presented in another article [Che 13].

This master's thesis is a continuation of the work by Kleiven and Zinserling, as a collaboration between Chalmers University of Technology and industry. The aim is to use measurements and calculations from previous work to calculate finite impulse responses, i.e. FIR filter coefficients, from the optimal source strengths for each source. The filters are used for a sound zone simulation, in which virtual sound zones are created based on transfer function measurements of a real car. Evaluation of the filters are done by subjective measures, such as sound quality of auralisations, as well as objective measures, e.g. level difference between the addressed zone and a silent zone.

1.2 Objective

The objective of this thesis is to develop methods for calculation of FIR filters achieving clearly separated sound zones with a satisfactory sound quality. The frequency range of interest is in this thesis restricted to 100 Hz $-$ 800 Hz due to limited access to measurements.

1.3 Thesis Outline

In Chapter 2, a theoretical background is given, providing a foundation for the calculation methods presented in Chapter 3. The filters produced by the calculation methods are used for sound zone simulations in Chapter 4. Chapter 5 is dedicated to discussion of the methods and the results. In Chapter 6 the thesis is concluded, and suggestions for future work are given.

Chapter 2

Theory

This chapter presents the theory which the following parts of the thesis build upon. The intention is not to fully describe each topic, but to provide a sufficient theoretical background in the areas that are relevant for the thesis. The topics covered are: linear systems, the Fourier transform, the z-transform, homomorphic systems and the cepstrum. Finally, the calculation method used in previous work is described, and some references are given regarding the human perception of sound.

2.1 Linear Systems

A system H is called linear if the following relation holds [Phi 08]:

$$H(ax + by) = aH(x) + bH(y). \tag{2.1}$$

where x and y are vectors, and a and b are scalars. If H is time-invariant, it means that the output of the system only depends on the input, irrespective of when the input is given. This is described by the property below [Phi 08]:

$$x(t - t_0) \rightarrow y(t - t_0), \tag{2.2}$$

where x and y are the input and output of the system, respectively, and t_0 is a time delay. A system that is both linear and time-invariant is called a linear time-invariant system (LTI). This is an important class of systems with a very broad application. Here follows definitions and brief explanations of the concepts pole-zero representation, causality, minimum phase, group delay and spectral flatness.

2.1.1 Pole-Zero Representation

A linear system can be represented by its transfer function, for a discrete system $H(z)$ a quotient of two polynomials in the complex number z:

$$H(z) = \frac{B(z)}{A(z)}. \tag{2.3}$$

The roots to $A(z)$ and $B(z)$ play an important role in the performance of the system. For example, roots to the denominator polynomial (poles), correspond to resonances, and roots to the numerator polynomial (zeros), are antiresonances of the system. A system architecture commonly used in digital signal processing is the FIR filter, for which the denominator polynomial is unity. Therefore a FIR filter has only zeros and no poles. For a FIR filter, the coefficients of $B(z)$ constitute the impulse response of the filter.

A common way to visualise the poles and zeros of a system is by a pole-zero plot. The complex roots to the polynomials $A(z)$ and $B(z)$ are then plotted in the complex plane, usually together with a circle centered in origo and with radius one, called the unit circle.

2.1.2 Causality

When aiming to create a filter in the time domain from a frequency domain specification, it is vital that the specification is physically realisable. One criterion for this is that the system is *causal*. In a causal system the current output depends only on the current and/or past values of the input, i.e. it does not "see into the future". For a system with impulse response $h(t)$, this means that $h(t) = 0$ for $t < 0$ [Phi 08].

2.1.3 Minimum Phase

Minimum phase for a discrete-time linear time-invariant (LTI) filter $H(z)$ is given by the following definition [Smi 07]:

Definition 1 *An LTI filter $H(z) = B(z)/A(z)$ is said to be minimum phase if all its poles and zeros are inside the unit circle $|z| = 1$ (excluding the unit circle itself).*

A consequence of this definition is that both $H(z)$ and its inverse are stable as well as causal, i.e. minimum phase is a sufficient (but not necessary) condition for causality. For a time domain filter, the definitions is simply

Definition 2 *A signal $h(n)$, $n \in Z$, is said to be minimum phase if its z-transform $H(z)$ is minimum phase,*

where the signal $h(n)$ is the filter impulse response.

Another aspect of minimum phase systems is that the decay of the impulse response is fastest of all causal systems with the same magnitude response. Hence, by minimising the decay time of the impulse response counting from time zero, the initial delay will also be minimised.

2.1.4 Group Delay

Group delay is a quantity describing the phase response of a linear system. As the name suggests, it represents the time delay that a group of sinusoidal components experience when passing through the system. It can also be seen as the time delay of the amplitude envelope of a sinusoid subject to the system. The group delay $D(\omega)$ of a continuous linear system with phase response $\Theta(\omega)$ is defined by the following formula [Smi 07]:

$$D(\omega) = -\frac{d}{d\omega}\Theta(\omega). \tag{2.4}$$

2.1.5 Spectral Flatness

In frequency analysis of signals and systems, the magnitude spectrum of the signal, or magnitude response of the system, is of great use. For some applications, it is of interest to have a flat magnitude response. A quantification of the flatness is given by the *spectral flatness*, defined below [Mad 09].

Definition 3 *The spectral flatness F of a spectrum $X(k)$ is defined as the geometric mean of the power spectrum divided by the arithmetic mean of the power spectrum.*

$$F = \frac{(\prod_{k=0}^{N-1} |X(k)|)^{\frac{1}{N}}}{\frac{1}{N}\sum_{k=0}^{N-1} |X(k)|}. \tag{2.5}$$

Spectral flatness is commonly presented in decibels according to

$$F_{dB} = 20 \log F. \tag{2.6}$$

F_{dB} ranges from $-\infty$ dB for a pure sinusoid, to 0 dB for a white spectrum.

2.2 The Fourier Transform

The Fourier transform is described by

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt, \tag{2.7}$$

where $X(\omega)$ is the Fourier transform of $x(t)$. With this definition, the inverse Fourier transform becomes

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega. \tag{2.8}$$

Using Euler's formula, eq. (2.8) can be written as

$$x(t) = \frac{1}{2\pi}(\int_{-\infty}^{\infty} X(\omega)\cos(\omega t)d\omega + j\int_{-\infty}^{\infty} X(\omega)\sin(\omega t)d\omega). \tag{2.9}$$

If $X(\omega)$ is real and even, the integrand in the second term of eq. (2.9) is a product of an even function and an odd function, i.e. an odd function. Evaluated on a symmetric interval the integral becomes zero, leaving only the first term. Since $X(\omega)$ is real, the remaining term is real and $x(t)$ is a real function.

Now assume that $X(\omega)$ is imaginary and odd. In this case it is the integrand in the first term of eq. (2.9) that is an odd function, making that integral term zero. The second term remains, which is real since $jX(\omega)$ is real. Also in this case $x(t)$ is a real function.

Note that nothing is said about the causality of $x(t)$. Depending on $X(\omega)$, $x(t)$ may be causal or noncausal. However, due to linearity of the Fourier transform, if a spectrum $X(\omega)$ has an even real part and an imaginary odd part, it follows from the paragraphs above that the inverse transform $x(t)$ is real valued. This important result is used for impulse response calculation later in the thesis.

2.3 The Z-Transform

The z-transform is a useful tool when working with discrete signals and systems. Much like the Fourier transform, it transforms a time domain sequence into a complex sequence, but now defined on the complex z-plane. The z-transform $X(z)$ of a sequence $x[n]$ is defined as below [Phi 08]:

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} \tag{2.10}$$

Similar to the Fourier transform, the z-transform is linear. It also has the property that a convolution of two signals $x_1[n]$ and $x_2[n]$ in the time

domain maps to a multiplication of the transformed sequences in the z-domain [Phi 08]:

$$\mathcal{Z}\{x_1[n] * x_2[n]\} = \mathcal{Z}\{x_1[n]\} \cdot \mathcal{Z}\{x_2[n]\}. \tag{2.11}$$

2.4 Homomorphic Systems

The concept of homomorphic systems was developed by Oppenheim in the sixties [Opp 65]. The main idea of homomorphic systems is to transform a nonlinear combination of signals into an additive combination on which linear filtering can be performed.

The general homomorphic system includes a system tranformation $D_\circ$ mapping the input operation $\circ$ to addition, a linear filter L, and another system transformation D_Δ^{-1} mapping addition to the output operation Δ, shown in figure 2.1. Note that the operators $\circ$ and Δ are arbitrary in the general case.

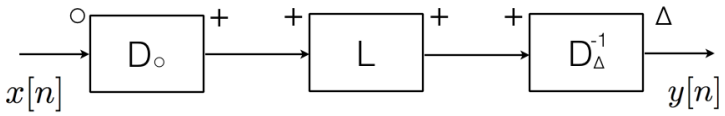


Figure 2.1: Canonic representation of a homomorphic system. $D_\circ$ and D_Δ^{-1} are the system transformations, from input operator $\circ$ to addition, and from addition to output operator Δ, respectively. L is a linear system of choice.

A special case is when the input and output operations are the same, $\Delta = \circ$, so that D_Δ^{-1} is the inverse of $D_\circ$. A trivial case is when $\circ = +$, for which the whole homomorphic system becomes linear. A more interesting case is when the operation is multiplication. It is then of interest to find the transformation $D_\circ = D_.$, mapping a multiplication of two signals to a sum of the transformed signals:

$$D_.\{x_1[n] \cdot x_2[n]\} = D_.\{x_1[n]\} + D_.\{x_2[n]\}. \tag{2.12}$$

An intuitive choice would be to take $D_.\{x[n]\} = \log(x[n])$. The expression in eq. (2.12) then becomes

$$\log(x_1[n] \cdot x_2[n]) = \log(x_1[n]) + \log(x_2[n]), \tag{2.13}$$

i.e. a linear combination of the transformed signals.

An important application of homomorphic systems arises when the operation is convolution. In this case, the transformation from convolution to addition has to be found. The transformation in eq. (2.13) maps a multiplication to addition, so if a mapping from convolution to multiplication can be found, the combination of both mappings will perform the desired transformation. A well known transformation with this property is the z-transform, defined in the previous section. Combining eq. (2.13) and eq. (2.11) and taking the inverse z-transform of the result, a transformation D_* is obtained that maps convolution to addition:

$$D_*\{x[n]\} = \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x[n]\})\}, \tag{2.14}$$

This is verified by setting $x[n] = x_1[n] * x_2[n]$ in eq. (2.14):

$$D_*\{x_1[n] * x_2[n]\} = \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x_1[n] * x_2[n]\})\}. \tag{2.15}$$

Substituting eq. (2.11) into eq. (2.15) yields

$$D_*\{x_1[n] * x_2[n]\} = \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x_1[n]\} \cdot \mathcal{Z}\{x_2[n]\})\}. \tag{2.16}$$

Since the logarithm of a product is the sum of the logarithms of each term, this can be written as:

$$D_*\{x_1[n] * x_2[n]\} = \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x_1[n]\}) + \log(\mathcal{Z}\{x_2[n]\})\}. \tag{2.17}$$

Finally, the linearity of the inverse z-transform (and the z-transform itself), gives that

$$\begin{aligned} D_*\{x_1[n] * x_2[n]\} &= \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x_1[n]\})\} + \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x_2[n]\})\} \\ &= D_*\{x_1[n]\} + D_*\{x_2[n]\}, \end{aligned} \tag{2.18}$$

which fulfills the desired property, i.e. mapping convolution to addition. While the inverse z-transform is unnecessary for getting the additive property, it gives rise to important properties which are discussed in the following section.

2.5 The Cepstrum

In the previous section it was described how to transform two nonlinearly combined signals into an additive combination of the signal transforms.

The special case when the nonlinear combination was achieved by convolution was described, and a transformation for this case was given in eq. (2.14). In fact, this equation is identical to the definition of the complex cepstrum [Chi 77]:

Definition 4 *The complex cepstrum $\hat{x}[n]$ of a discrete signal $x[n]$ is defined as*

$$\hat{x}[n] = \mathcal{Z}^{-1}\{\log(\mathcal{Z}\{x[n]\})\}. \tag{2.19}$$

Sometimes the power cepstrum is used instead of the complex cepstrum. It was originally defined as the power spectrum of the logarithm of the power spectrum, a definition which lives on in some applications, but was changed to the following definition [Chi 77]:

Definition 5 *The power cepstrum $x_{pc}[n]$ of a discrete signal $x[n]$ is defined as*

$$x_{pc}[n] = (\mathcal{Z}^{-1}\{\log(|\mathcal{Z}\{x[n]\}|^2)\})^2. \tag{2.20}$$

An important property of the complex cepstrum is how poles and zeros of a system are mapped into the cepstral domain when the complex cepstrum of the system is calculated. Poles and zeros inside the unit circle are mapped to the positive x-axis (or the positive *quefrency* axis), while poles and zeros outside the unit circle are mapped to the negative quefrency axis [Wol 09].

2.6 Miscellaneous

2.6.1 Calculation Method in Previous Work

In order to understand the difference between the work presented in this thesis and the previous work, it is of interest to study the method used there. The foundation of the method is the relation described in the following expression [Kle 12]:

$$\begin{bmatrix} p_A \\ p_S \end{bmatrix} = \begin{bmatrix} \Psi(A,L) \\ \Psi(S,L) \end{bmatrix} [Q_L], \tag{2.21}$$

where A and S stands for addressed and silent zone, L denotes source number, p_A and p_S are the sound pressures in the respective zones, Ψ is the transfer function from source to zone, and Q_L is the source strength for source L. Setting the sound pressures to $p_A = p_T$ and $p_S = 0$, where p_T is the target pressure, and solving for Q_L yields a set of complex source

strengths. Due to the nature of the transfer function matrix Ψ, the equation must be solved numerically with an error tolerance. In the previous work this was done by the Moore-Penrose pseudoinverse [Kle 12], [Bar 11]. Therefore, when inserting the calculated complex source strengths back into eq. (2.21), the target pressures are not reproduced perfectly. For more details on the calculation process, the reader is refered to the article by Kleiven and Zinserling [Kle 12]. Results from this method and the methods presented in this thesis are compared in Chapter 4.

2.6.2 Time Delay in Communication

Digital signal processing inevitably introduces time delays in the signal chain. In real-time applications it is therefore important to know how users react to various time delays. In a future car featuring a sound zone system, it is reasonable to assume that two-way communication is one of the applications. For this case, the International Telecommunication Union recommends an ear-to-mouth delay, i.e. the total delay experienced by the user, of maximum 150 ms [ITU 03].

2.6.3 Separation of Interfering Radio Programs

Another factor to take into account for a sound zone application is the separation needed between the addressed zone and a silent zone or, in a future version, between two addressed zones with different audio content. For the latter case a minimum separation of 11 dB is recommended for interfering radio programs [Dru 94]. This figure is probably highly dependent of the type of audio content, and the users' personal preferences. For example, an audiophile might demand higher audio integrity than the average radio listener.

Chapter 3

Filter Realisation Methods

In this chapter the methods for calculating finite impulse responses, i.e. FIR filters coefficients, from frequency response functions are described, building upon the theory in Chapter 2. Three methods are explained and evaluated: one method based on inverse transformation of an altered spectrum, called the mirroring method, and two methods based on filtering in the cepstral domain, so called homomorphic filtering. Firstly it is described how the data obtained from previous work was prepared for impulse response calculation.

3.1 Preparation of Data from Previous Work

The frequency data from previous work was given in third-octave band frequencies. Since the discrete fourier transform relies on constant frequency steps, there was a need to find the points in between the third-octave bands, i.e. interpolating the frequency data [Phi 08]. There are different ways to interpolate a signal, one of the simplest being linear interpolation, in which a straight line is drawn between each of the original data points. The new points are found on the line, with a resolution that is decided by the user. Other more sophisticated interpolation methods are e.g. quadratic curve fitting and spline interpolation [Cro 83]. Since the measured transfer functions, and the complex source strenghts calculated from these, are unlikely to resemble mathematical functions it was decided to use linear interpolation for its simplicity.

3.2 The Mirroring Method

The mirroring method is a way to alter the frequency response so that its inverse transform is real valued. It is based on the properties of the Fourier transform presented in Section 2.2, i.e. that a frequency response with even real part and odd imaginary part transforms to a real valued impulse response.

For discrete signals, the discrete Fourier transform (DFT) has to be used. The properties of the Fourier transform given in Section 2.2 are basically the same for the DFT. However, instead of a symmetry around $\omega = 0$, the DFT of a real signal is symmetric around the Nyquist frequency N [Sel 01].

Say that the impulse response $h[n]$ of a single sided frequency response $H(k)$ is to be calculated. A double sided frequency response is then created by mirroring $H(k)$ in $k = N$. The real and imaginary parts of the resulting function are both even around N. The imaginary part is then made odd by conjugating the mirror image, i.e. the frequency components above N. Finally, $h[n]$ is obtained by taking the inverse DFT of the double sided frequency response.

As mentioned in Section 2.2, $h[n]$ is not guaranteed to be causal. In order to ensure causality, a linear delay t_d corresponding to half the impulse response length was added to the desired frequency response by multiplying with the factor $e^{-j\omega t_d}$. The delay size was motivated by the fact that the undelayed impulse response contained roughly the same amount of energy for negative time as for positive time. A delay of half the impulse response length places the main peak in the middle of the signal, and thus includes maximal energy from both tails.

3.3 Homomorphic Filtering

In Section 2.5 it was discussed how poles and zeros map into the cepstral domain when a system is transformed to its complex cepstrum, i.e. poles and zeros outside the unit circle map to the negative quefrency axis, and poles and zeros inside the unit circle map to the positive quefrency axis. In Section 2.1.3 it was concluded that a filter with all its poles and zeros inside the unit circle is minimum phase and therefore causal. For FIR filters such as the impulse responses considered in this thesis, stability is not an issue since this type of filters have no poles.

Thus, if a homomorphic system as the one shown in figure 2.1 is designed such that the linear filter L alters the cepstrum to remove all negative quefrency content, a minimum phase filter is obtained. The two methods

described below are similar, and both produce minimum phase FIR filters, but with different magnitude responses. This way both filters can be minimum phase, even though their phases are not the same. More on this is seen when the methods are evaluated in Section 3.4.

3.3.1 Homomorphic Mirroring

One way to alter the cepstrum is to mirror the negative quefrency content onto the positive axis, and add it to the positive quefrency content. This way no zeros of the system are removed, since the zeros with magnitude larger than one are mirrored across the perifery of the unit circle. This method leaves the magnitude response unchanged.

3.3.2 Homomorphic Windowing

Another approach is to apply a rectangular window to the cepstrum, setting all negative quefrency components to zero. Since a rectangular window in the time domain transforms to a sinc function in the frequency domain, multiplying the cepstrum with a rectangular window corresponds to convolving the log spectrum with a sinc function. This in turn leads to a frequency smearing of the log spectrum, smoothing out the fastest changes, resulting in a filter of lower complexity.

3.4 Evaluation

In this section the different methods are evaluated and compared. The objective of each method is to calculate an impulse response from a given frequency response. The impulse responses of the methods are compared among each other. They are then evaluated in the frequency domain and compared with the original frequency response. The frequency response used for evaluation is one of the complex source strengths calculated in previous work. Finally the zeros of the filters are plotted, illustrating how different window widths in the cepstral domain affect the zeros of the homomorphic windowing filter.

Important performance measures are how well the impulse responses represent the desired frequency response, and how long initial delay they have. A short initial delay is advantageous in real time applications, and vital for two-way communication as discussed in Section 2.6.

All results in this chapter are based on filters with 128 coefficients. As mentioned in Section 3.3.2, the filter complexity changes when the cep-

strum is windowed. This section includes results from the homomorphic windowing method with windows of widths 8, 16 and 32 samples.

3.4.1 Impulse Response

Figure 3.1 shows the impulse responses calculated using the mirroring method and the homomorphic mirroring method. Since the mirroring method requires a causal system, a delay of 40 ms, i.e. half the impulse response length at a sample rate of 1600 Hz, was included in the frequency domain, see Section 3.2. This shows in the mirroring method impulse response as a time shift of the same amount.

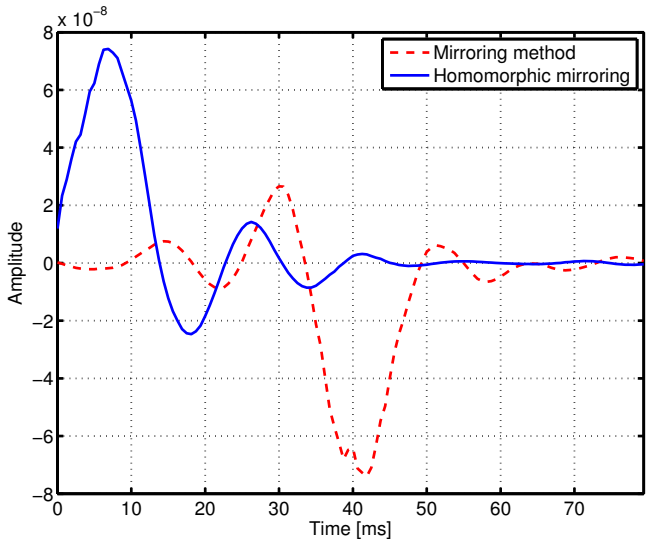


Figure 3.1: Impulse response of the mirroring method and the homomorphic mirroring method.

As described in Section 3.3.2, the homomorphic windowing method can be used with varying window sizes, creating filters of different complexity. It is thus of interest to compare the performance as a function of the filter complexity. Figure 3.2 illustrates the effect of different filter complexities on the impulse response of the homomorphic windowing method. Window widths of 8, 16, and 32 samples are evaluated. As expected, a lower filter complexity gave a faster decay. Compared to the impulse responses in Figure 3.1, these had shorter initial delay, less than 5 ms, and lower amplitude by approximately a factor 6.

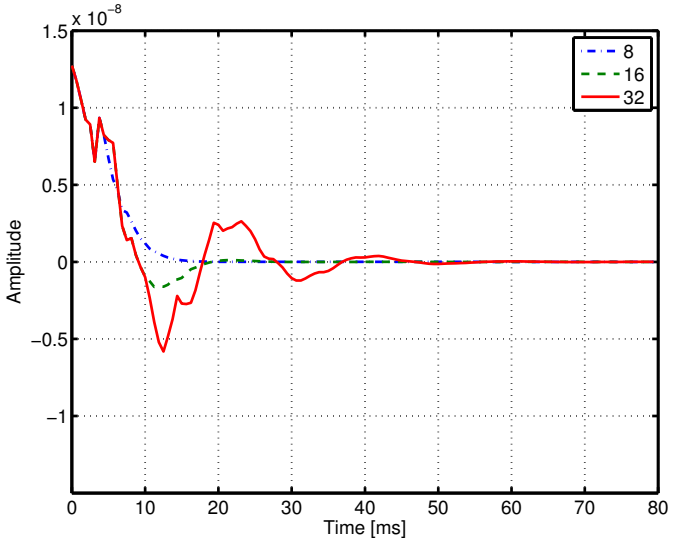


Figure 3.2: Impulse response of the homomorphic windowing method for varying effective filter orders.

3.4.2 Frequency Response

A frequency domain evaluation of the calculated impulse responses is given in figures 3.3 to 3.6, starting with the magnitude response for the mirroring method and the homomorphic mirroring method. Both methods achieved excellent magnitude following for frequencies below 700 Hz. The deviations above 700 Hz can be attributed to truncation errors as well as bad frequency resolution. This conclusion is drawn since the frequency responses based on filters of length 2048 did not have these deviations, see figure A.3 in Appendix A.

The phase comparison in figure 3.4 shows that the mirroring method followed the desired phase reasonably well. The phase response produced by the homomorphic mirroring method was considerably smaller than that for the mirroring method, as is expected from a minimum phase filter. In theory, a 40 ms delay should add a linear slope to the phase. However, the desired phase in figure 3.4 is only piecewise linear. The explanation for this can be found in how the calculation software handles delay. In general a linear, positive delay shifts the signal in time, to the right on the positive time-axis. A linear negative delay on the other hand shifts the signal to the left. Since the time-axis extends infinitely in both directions the signal is never "wrapped around". In MATLAB however, the signal is composed of a vector with a finite number of elements. When the signal is time shifted in

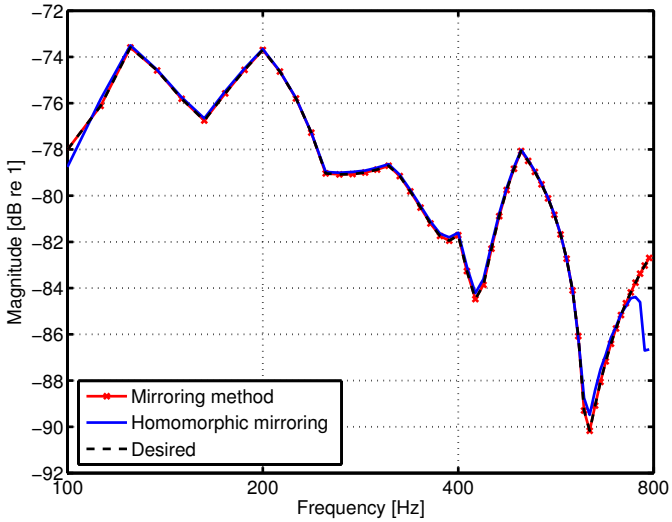


Figure 3.3: Magnitude response of the mirroring method and the homomorphic mirroring method compared to the desired magnitude response.

the positive direction, the elements closest to the right end of the signal are shifted out, appearing in the left end. Thus, a positive delay corresponding to exactly half the signal length can not be distinguished from a negative delay of the same amount. Since positive and negative time delays correspond to linear decrease and increase in the phase, respectively, the above reasoning may explain why the phase in figure 3.4 varied between a linear increase and a linear decrease.

The magnitude responses of the homomorphic windowing method for window widths of 8, 16, and 32 samples as well as the desired source strength are given in figure 3.5. It is clear that the filter with a 32 samples homomorphic window better followed quick changes in the desired response. It was also more prone to ripple. The filters of lower complexity had smoother responses, further from the desired response.

In figure 3.6 the phase responses of the homomorphic windowing method are plotted for window widths of 8, 16, and 32 samples. They are compared to the undelayed phase response of the desired source strength, as well as the homomorphic mirroring method. For the magnitude responses, it was seen that the most complex filter achieved the best approximation. For the phase reponse no such relation can be seen. While the higher complexity filters had faster phase variations, they did not follow the desired phase better. The minimum phase response produced by the homomorphic mir-

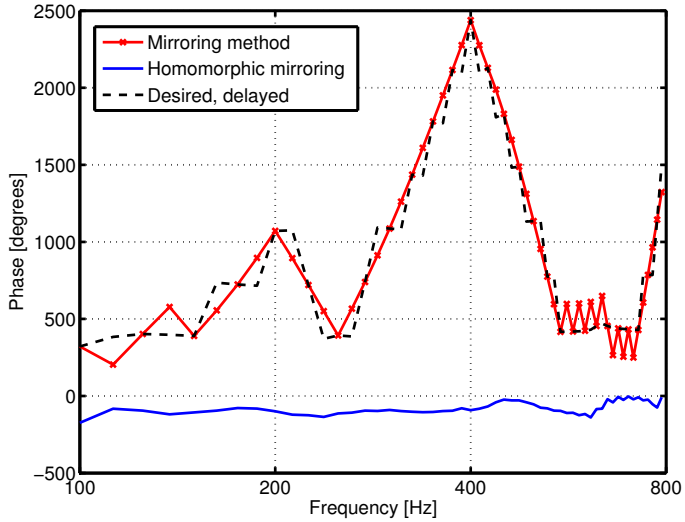


Figure 3.4: Phase response of the mirroring method and the homomorphic mirroring method compared to the desired magnitude response delayed by 40 ms.

roring method seems to be farther from the desired phase as compared to the homomorphic windowing method. A possible interpretation of this may be that the excellent magnitude following of the homomorphic mirroring method comes at the price of worse phase following. This may then motivate the use of the homomorphic windowing method, which exhibits worse magnitude following but slightly better phase following.

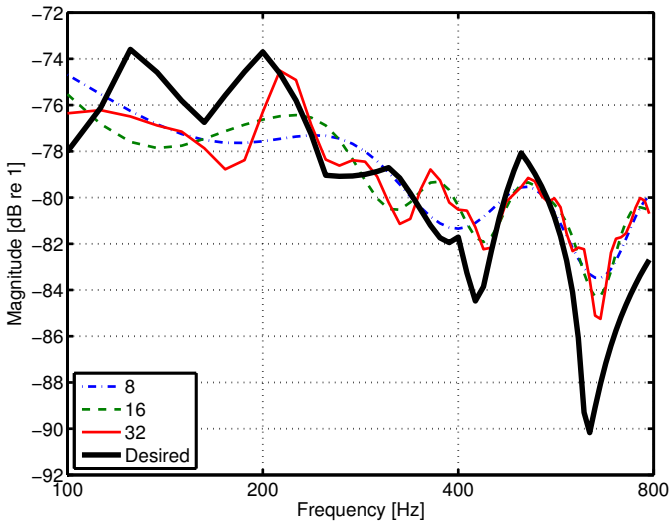


Figure 3.5: Magnitude response of the homomorphic windowing method compared to the desired magnitude response.

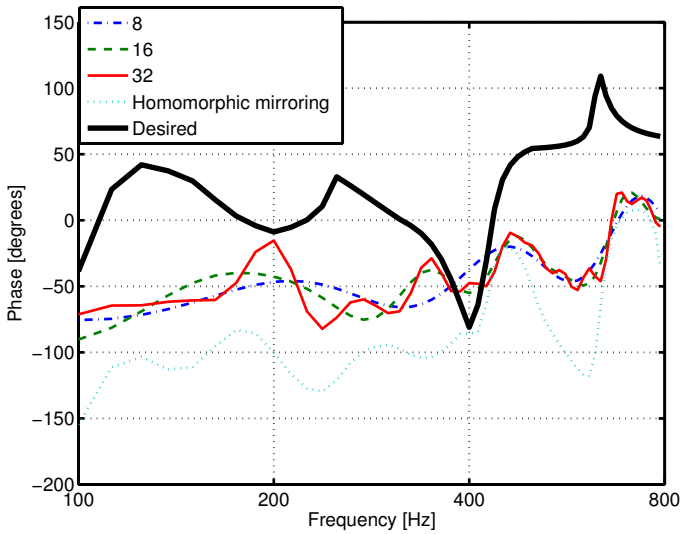


Figure 3.6: Phase response of the homomorphic windowing method and the homomorphic mirroring method compared to the desired phase response, with no delay included.

3.4.3 Filter Zeros

When varying the window size in the homomorphic windowing method, the zeros of the filter were changed. This is evident in figure 3.7, for filters with homomorphic window widths of 8, 16 and 32 samples. Most notable is how the maximum zero radius increased with the window width, effectively decreasing the causality margin. However, all zeros lie within the unit circle, so the filters are causal as well as minimum phase, according to the definitions in Section 2.1.

For reference, it is interesting to look at the zero plots of the other two methods. Figures 3.8 and 3.9 show the zeros of the filters from the mirroring method and the homomorphic mirroring method, respectively. It is clear that the mirroring method produced zeros outside the unit circle, so it can be confirmed that this filter is non-minimum (or mixed) phase, as was indicated by the impulse response in figure 3.1.

The zeros of the homomorphic mirroring method are plotted in figure 3.9. As for the homomorphic windowing method, all zeros lie inside the unit circle. The claim in Section 3.3 that both methods produce minimum phase filters is thus confirmed for the complex source strength used for evaluation in this section.

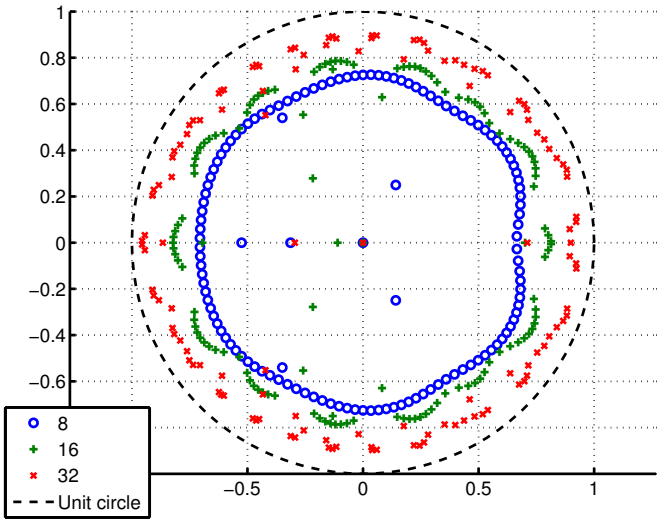


Figure 3.7: Zero plot for the homomorphic windowing method, with filter orders 8, 16 and 32.

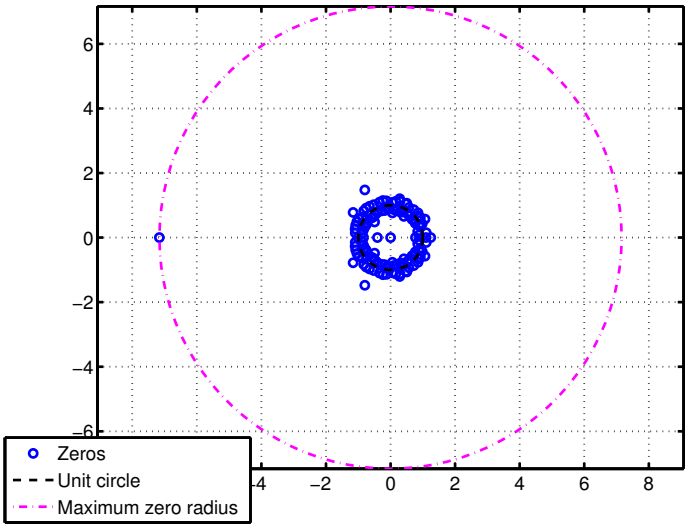


Figure 3.8: Zero plot for the mirroring method.

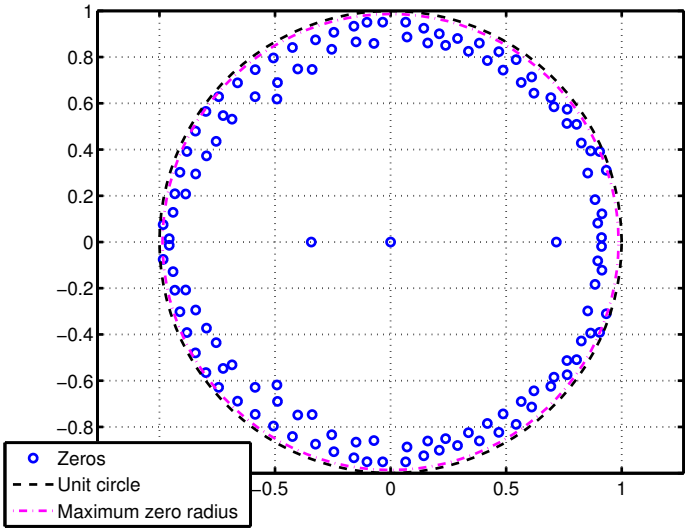


Figure 3.9: Zero plot for the homomorphic mirroring method.

3.5 Comments

As stated in Section 3.4, all results presented in this chapter are based on one of the complex source strengths calculated in previous work. The intention was to illustrate the differences in performance between the methods, and for this purpose one single source strength sufficed. However, an interesting trend is seen when looking at the impulse responses calculated from all sources. Both the homomorphic filtering methods, i.e. homomorphic mirroring and homomorphic windowing, produced impulse responses with only positive initial peaks, although varying in amplitude and phase. This can be seen in figures A.8 and A.9 in Appendix A, in which impulse responses calculated by the homomorphic filtering methods, for all complex source strengths, are plotted. The reason for this behaviour is believed to be found in the minimum phase property of the filters. Due to the linearity of the Fourier transform, a negative sign in the impulse response translates to a negative sign in the frequency response, i.e. a negation of each complex bin in the DFT. This corresponds to a phase shift of 180 degrees. Adding a phase shift of 180 degrees violates the minimum phase property since the phase is then no longer minimum. Thus, since the homomorphic filtering methods give minimum phase filters, all impulse responses have positive peaks.

Another comment to be made based on the impulse response plots in Appendix A is that the difference in amplitude of a factor 6 seen in figures 3.1 and 3.2 is not representative for all impulse responses of the homomorphic windowing method. In average, the homomorphic windowing method seems to produce impulse reponses of around half the amplitude compared to the two other methods.

Chapter 4

Sound Zone Implementation

Chapter 3 introduced three methods for calculating impulse responses from a given frequency response. The methods were tested on a complex source strength given by Kleiven and Zinserling [Kle 12]. In this chapter it is described how impulse responses for all complex source strengths were used to set up a sound zone simulation, with transfer function data from a real car and sound files as input. Resulting sound pressure levels and zone separation are discussed and compared to results in previous work. Additionally, auralisations of the sounds in the listening zones are subjectively evaluated.

4.1 Simulation Setup

The simulation setup had 16 sources, placed in a quadratic array in the ceiling. Three zones were defined: one for the driver seat, one for the front passenger seat, and one for the back seats. Each zone was divided into a measurement grid with 143, 142 and 330 points (driver, passenger and back zones, respectively). Figure 4.1 gives an overview of the source and receiver configuration.

Transfer function data from each source to each measurement position was provided by previous work, based on FEM modeling of the car compartment, from 100 Hz to 500 Hz, and measurements, from 500 Hz to 800 Hz. This also included loudspeaker responses [Kle 12].

Three different kinds of input signals are considered in this chapter: an ideal impulse, a speech recording, and a guitar recording. The recordings were done with a sampling frequency of 48 kHz, and were about 3 seconds long. Since the transfer function data only reaches up to 800 Hz, the sounds were decimated by a factor 30, so that the new sampling frequency became

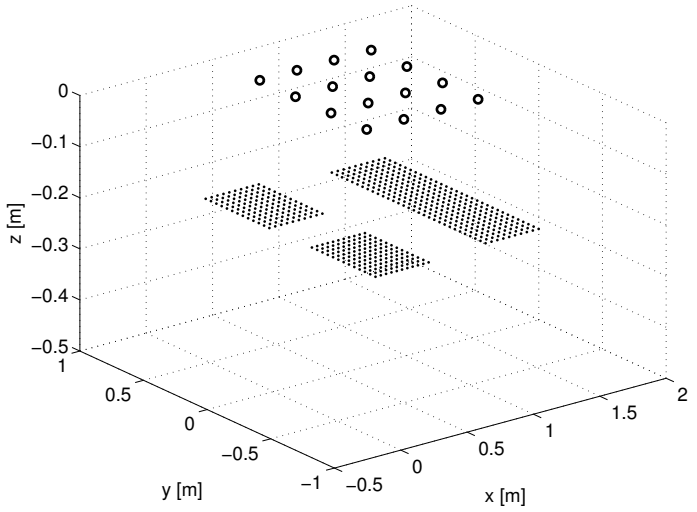


Figure 4.1: Source and receiver configuration. Sources are marked as circles and measurement positions are marked as dots. The driver zone is the one closest to the point of view.

1600 Hz.

Finally, impulse responses of length 128 were used as input filters, calculated from the complex source strengths of all 16 sources, using the methods described in Chapter 3.

Figure 4.2 shows a schematic of the simulation setup. Arrows represent signals, blocks represent filtering stages and circles represent summation. Signal dimensions, i.e. the number of channels, are given for each signal.

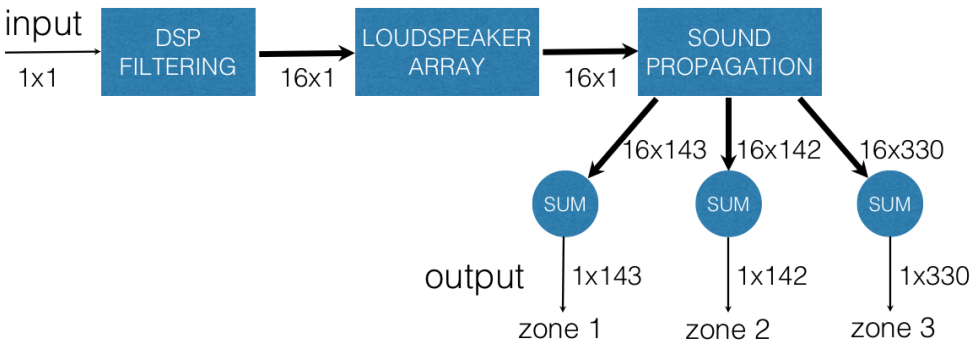


Figure 4.2: Overview of the signals and filtering stages in the sound zone simulation. Zone 1 is the driver seat, zone 2 is the passenger seat and zone 3 is the back seats.

4.2 Objective Results

A measure for evaluation of the sound zone performance is the separation, i.e. the difference in sound level, between the addressed zone (the driver zone in this implementation), and the silent zones (the passenger and back zones). The separation is seen when the spectra of the sounds are plotted over frequency in the section below. In the next section the sound pressure is plotted as a function of measurement position in each zone. Then a comparison to previous work is made, illustrating how the results of the methods used in this thesis differ from the frequency domain calculations by Kleiven and Zinserling. Finally, a summary provides the total separation levels between the addressed zone and the silent zones for all methods, including the method from previous work.

In Section 3.5 it was mentioned that both homomorphic filtering methods produced impulse responses with positive initial peaks for all sources, and a possible explanation was given. Tests not included here showed that other combination of signs performed better than having only upward peaks. At this point, no analytical procedure has been found to obtain the optimal signs. The approach taken was to empirically determine the combination of signs that produced the best results in terms of zone separation, by trying all combinations of the 16 impulse responses, multiplied with either 1 or -1. All results presented for the homomorphic filtering methods in this thesis are taken from the sign combination with the best result.

4.2.1 Magnitude Spectra of Output Signals

The spectra of the sounds in each zone are plotted for the three methods in figures 4.3 to 4.6, starting with the mirroring method, then the homomorphic mirroring method and lastly the homomorphic windowing method for a window width of 16 samples. The input signal to the sound zone system was an impulse.

The greatest level differences between the addressed zone and the silent zones are found for the mirroring method, shown in figure 4.3. The addressed zone exhibits a rather flat spectrum, while the silent zones have dips coinciding with the third-octave band frequencies. The latter observation is interesting, and may be related to the fact that the complex source strengths as well as the frequency reponse measurements of the car compartment only existed in third-octave band frequencies, given from previous work. As mentioned in Section 3.1, a finer resolution of the frequency data was obtained by linear interpolation. The error introduced by the in-

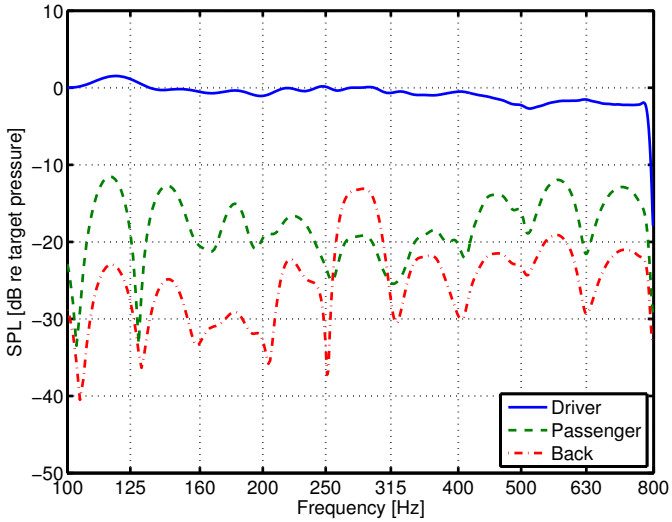


Figure 4.3: Narrow band levels for the mirroring method, with an impulse as input signal. The spectral flatness for the driver level is -0.13 dB.

terpolation may explain the decrease in separation for frequencies between the original frequency points. If this is the case, there is great potential for improvement in future work if measurements and calculations of higher intrinsic frequency resolution can be provided. A crude approximation of the results that may be possible for this case is obtained by connecting the dots from each dip in figure 4.3, plotted in figure 4.4. The mean separation for this approximation becomes 22.7 dB for the passenger zone and 30.6 dB for the back seat zone.

The two homomorphic filtering methods give smaller level differences, but smoother spectra in the silent zones, as seen in figures 4.5 and 4.6. For sound reproduction purposes, it is preferable if the sound zone system alters the input spectrum as little as possible. Thus, for an impulse input, a spectral flatness of 0 dB in the driver zone is desired. Of the three methods, the mirroring method produced the flattest spectrum in the driver zone, with a spectral flatness of -0.13 dB between 100 Hz and 800 Hz. Next came the homomorphic windowing method with window width 16, with -0.34 dB, and last the homomorphic mirroring method with -0.73 dB. All methods seem to produce a prominent dip in the output spectrum for all zones above 775 Hz. A definitive explanation for this has not been found. However, similar tendencies can be seen in the results given in the article by Kleiven and Zinserling, although not as pronounced [Kle 12]. Between

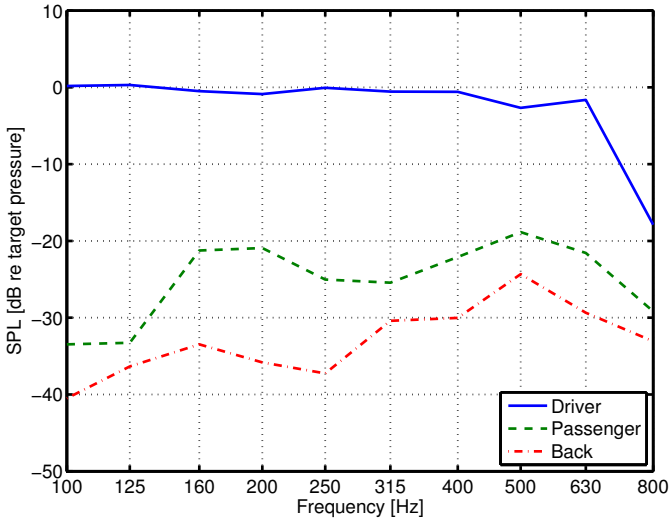


Figure 4.4: Normalised sound levels obtained from the mirroring method, evaluated in frequencies given by the dips in figure 4.3.

100 Hz and 775 Hz, the magnitude spectrum in the addressed zone produced by the mirroring method stays within ± 2 dB.

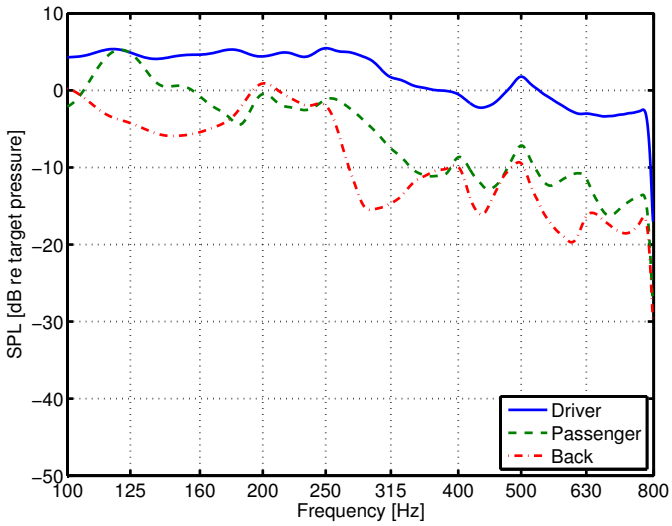


Figure 4.5: Narrow band levels for the homomorphic mirroring method, with an impulse as input signal. The spectral flatness for the driver level is -0.73 dB.

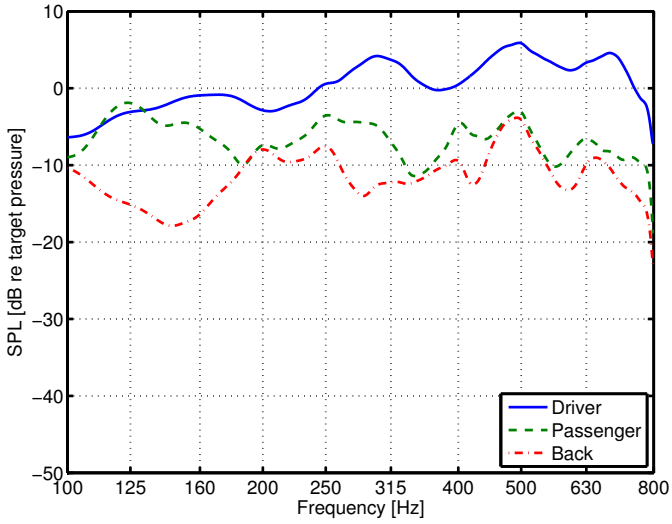


Figure 4.6: Narrow band levels for the homomorphic windowing method, with a window width of 16 samples. The input signal was an impulse. The spectral flatness for the driver level is -0.34 dB.

4.2.2 Sound Levels for Individual Measurement Positions

The results in the previous section were obtained by averaging sound pressures over measurement positions, producing one sound pressure per zone. This section presents surface plots, illustrating how the sound level varied over measurement positions. Figures 4.7 to 4.9 show the total sound levels for each measurement position for the mirroring method, the homomorphic mirroring method and the homomorphic windowing method with window width 16.

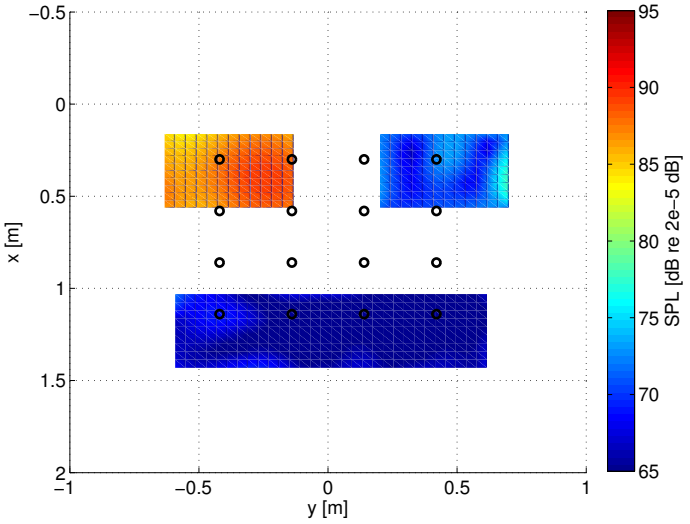


Figure 4.7: Sound levels as a function of measurement position. Mirroring method, with an impulse as input signal. The standard deviations in driver, passenger and back zones are 1.1 dB, 1.2 dB and 1.7 dB, respectively.

It can be seen that the driver zone had the highest sound levels for all methods, around 90 dB. It also had the smallest variation for most methods, as is seen in Table 4.1. In order to allow head movements without perceived change in sound level, it is desirable to have as small variation as possible between measurement positions, primarily in the addressed zone. In this sense the mirroring method was the best method, with standard deviation of 1.1 dB in the driver zone.

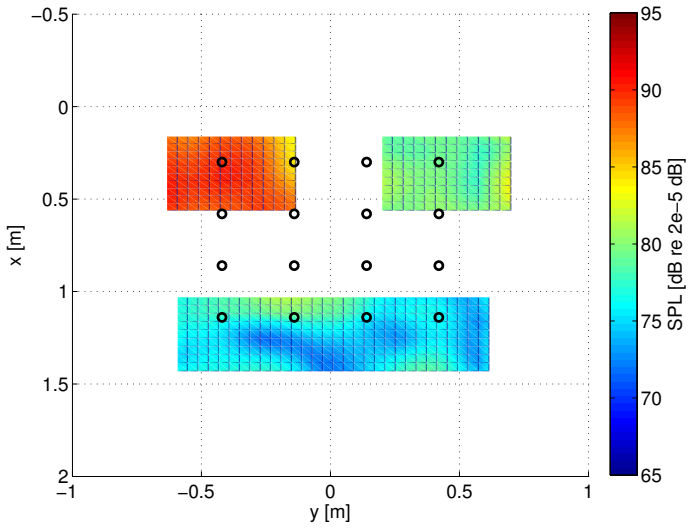


Figure 4.8: Sound levels as a function of measurement position. Homomorphic mirroring method, with an impulse as input signal. The standard deviations in driver, passenger and back zones are 1.5 dB, 1.1 dB and 1.6 dB, respectively.

Table 4.1: Standard deviation of sound levels for different measurement positions.

Standard deviation [dB]	Driver	Passenger	Back
Frequency domain [Kle 12]	0.7	3.4	2.7
Mirroring method	1.1	1.2	1.7
Hom. mirroring	1.5	1.1	1.6
Hom. windowing (16)	1.5	1.9	1.4

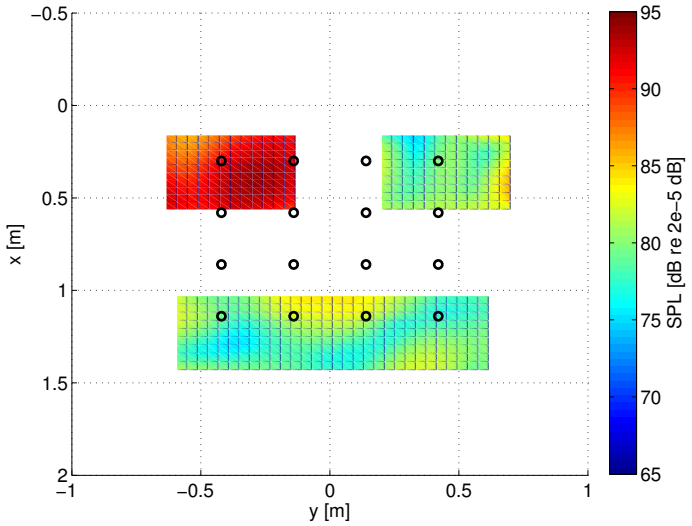


Figure 4.9: Sound levels as a function of measurement position. Homomorphic windowing method with a window width of 16 samples, with an impulse as input signal. The standard deviations in driver, passenger and back zones are 1.5 dB, 1.9 dB and 1.4 dB, respectively.

4.2.3 Comparison to Previous Work

As explained in Section 2.6, the method used in previous work is strictly based on frequency domain calculations, using target sound pressures to calculate complex source strengths by solving eq. (2.21). These are inserted in the same equation, so in the ideal case the target pressures would be reproduced exactly, but since the solution to eq. (2.21) is only approximate, the resulting sound pressures are not equal to the target pressures.

Figure 4.10 shows the sound pressure levels over frequency obtained from the mirroring method, smoothed to third-octave band frequencies and normalised to 0 dB. These may be compared to the sound pressure levels in figure 4.11, coming from the method from previous work, with the same source configuration as in this thesis. It is observed that the level curves for the driver zone are similar for both methods, although slightly smoother in figure 4.11. The single value separation levels for the previous method are 22.5 dB between driver and passenger zone, and 31.5 dB between driver and back zone, as compared to 17.0 dB and 24.2 dB for the mirroring method with impulse input. These differences illustrate the performance loss occuring when going from a frequency domain calculation, to time domain signals and convolution filters. Looking back at the mirroring method results in figure 4.4, in which it was attempted to approximate the sound pressure levels unaffected by interpolation errors, more similarities to the results in previous work are seen. Both the shape of the curves and the separation levels, 22.7 dB and 30.6 dB, match previous work better.

In figure 4.12, the sound levels from previous work are plotted for individual measurement positions, using the same source and receiver configuration as in this thesis. Apart from the larger separation, which was also seen in figure 4.11, the most notable difference from the results in the previous section is in the variation over measurement positions. The driver zone level has a standard deviation of 0.7 dB, as compared to 1.1 dB for the mirroring method. The silent zones have larger variations, 3.4 dB and 2.7 dB. However, since the sound level in these zones are so low, the variation is of less importance.

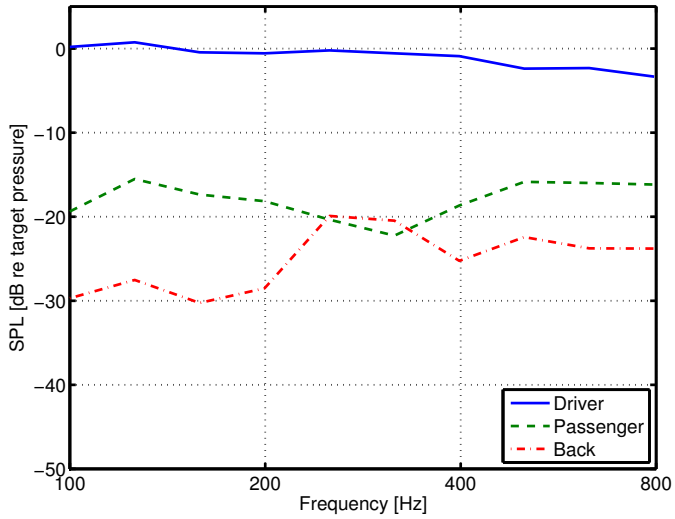


Figure 4.10: Normalised sound pressure levels produced by the mirroring method, with an impulse as input signal.

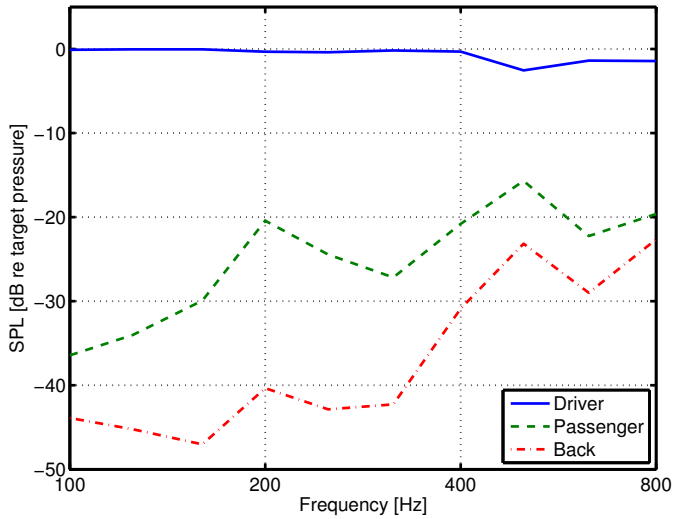


Figure 4.11: Normalised sound pressure levels produced by the method in previous work.

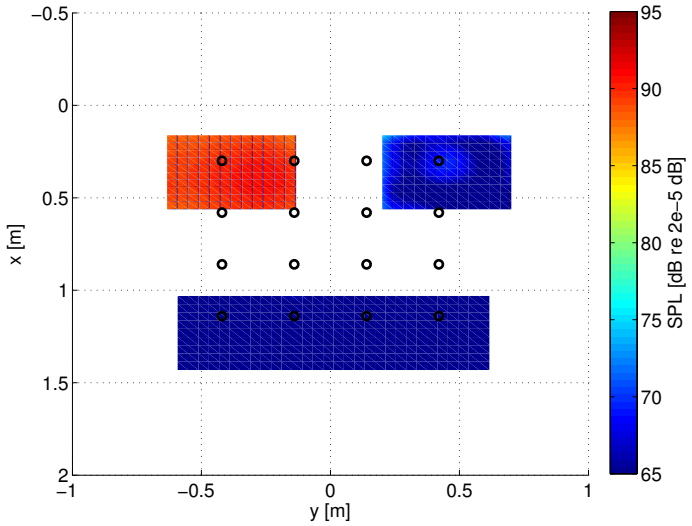


Figure 4.12: Sound levels as a function of measurement position, by the method from previous work. The standard deviations in driver, passenger and back zones are 0.7 dB, 3.4 dB and 2.7 dB, respectively.

4.2.4 Summary

Table 4.2 summarises the difference in sound level between the addressed zone and the silent zones, for all calculation methods proposed in this thesis. Results from the method in previous work are included for comparison. Of the proposed methods, the filters calculated by the mirroring method achieved the best separation. Compared to the results in previous work for impulse input, the separation for the mirroring method was around 5 dB lower for the passenger zone, and 7 dB lower for the back seat zone. However, for the speech and guitar inputs the mirroring method gave 7 - 15 dB lower separation, mostly due to the fact that the method in previous work performed better for those inputs. In terms of variation over measurement positions in the addressed zone, the mirroring method had the best performance of the methods in this thesis. The standard deviation of the driver zone level for this method was 0.3 dB higher than that for the method proposed by Kleiven and Zinserling.

Table 4.2: Separation for different sound input, calculated in the frequency range 100 Hz to 800 Hz. The window width used in the homomorphic windowing method is displayed in parenthesis.

Separation from driver [dB]	Impulse		Speech		Guitar	
	pass.	back	pass.	back	pass.	back
Frequency domain [Kle 12]	22.4	31.5	24.9	36.0	25.4	38.7
Mirroring method	17.0	24.2	16.9	24.5	18.3	24.9
Hom. mirroring	8.9	14.7	8.6	14.6	9.1	14.8
Hom. windowing (8)	10.3	11.3	10.2	11.2	10.1	11.2
Hom. windowing (16)	9.8	11.2	9.8	11.1	9.6	11.0
Hom. windowing (32)	9.6	11.2	9.6	11.0	9.5	11.2

4.3 Subjective Results

The signals produced by the sound zone simulation presented in this chapter were auralised and evaluated subjectively. This section aims at describing the sound quality as perceived by the author during informal listening tests. The sounds considered are the same as in Section 4.2, i.e. an impulse, a speech recording and a guitar recording. For the guitar recording, comments on the musical quality of the sound are included. The focus is directed to the quality of the sound reproduction in the addressed zone, since the sound in this zone is the most audible, and intended for critical listening. As stated in Section 4.1, the frequency spectra of all sounds were limited to frequencies below 800 Hz, making speech more dull and impulses less distinct for example. This should be kept in mind when listening to and evaluating the sounds. Further, the different complexities of the homomorphic windowing method are here treated as one, using window width 16, since the sounds were very similar.

4.3.1 Impulse Sound

The sound in the driver zone produced by the mirroring method was not as sharp as the original impulse, and had a smeared out feel to it. This was true also for the homomorphic mirroring method, and here the sound was even less sharp. The homomorphic windowing method, on the other hand, produced sharper and more distinct sounds although not as sharp and distinct as the original sound. An explanation for these characteristics may be found in the spectra of the sounds, seen in figures 4.3 to 4.6. There it shows that the homomorphic windowing method produced more gain in the high frequency region, above 400 Hz, than the other methods, giving rise to a sharper sound. The more distinct feel is likely caused by the more compact impulse responses of this method, as shown in figures A.7 to A.9 in Appendix A.

4.3.2 Speech Recording

The mirroring method reproduced the speech signal in the driver zone with no audible difference to the original sound. The homomorphic mirroring method gave a slightly more booming sound in the driver zone, discriminating it from the original sound. The homomorphic windowing method produced a thinner sound, with considerably less low frequency content as compared to the original sound. Again, looking at figures 4.3, 4.5 and

4.6, it is confirmed that while the mirroring method gave a rather flat response for low frequencies, the homomorphic mirroring method produced more gain in that range. The homomorphic windowing method, however, resulted in a spectrum with an attenuated low frequency region.

4.3.3 Guitar Recording

The reproduction of the guitar recording as performed by the mirroring method was fairly consistent with the original sound, although slightly more blurry. The homomorphic mirroring method gave the recording some resonance, and made it sound less full than the original. The homomorphic windowing method gave a brighter sound than the other methods, which the author found pleasing. However, since the frequency spectra of the sounds were very limited, making the sounds inherently dull, it is reasonable to assume that brightness was valued more than if full-range sounds had been used.

4.3.4 Summary

Of the three methods evaluated in this thesis, the mirroring method reproduced the input sound most accurately in terms of frequency content and overall impression. However, the most distinct sound was obtained from the homomorphic windowing method. This was most apparent for the impulse, which was less smeared out than for the other methods.

Chapter 5

Discussion

One of the most interesting aspects of this thesis was to see how much of the performance from previous work that could be preserved when implementing the complex source strengths in FIR filters, as this would give a first indication of how a real-world implementation would perform. In Chapter 2 the method used by Kleiven and Zinserling was briefly described. A multiplication of source strengths and propagation transfer functions in the frequency domain produced sound pressures in each measurement position. Theoretically, there is of course no difference between doing a multiplication in the frequency domain and a convolution in the time domain. However, when transforming the complex source strengths and transfer functions into the time domain, care needs to be taken to ensure e.g. causal and real valued time signals. The mirroring method succeeded well in realising the magnitude and phase responses of the complex source strengths, although a time delay was necessary to achieve causal impulse responses. In the end, the mirroring method performed $5 - 7$ dB worse than the method in previous work for impulse input. Since a time delay does not change the phase relations of a filter, this is not believed to be the reason. The most likely cause is the interpolation of the frequency data given from previous work. It was done in order to get uniform frequency spacing, to allow the ordinary FFT algorithm to work. It was seen that the sound pressure levels in the silent zones were fluctuating, with the lowest levels close to the original, not yet interpolated, frequency points. This can be interpreted as an indication that the interpolation error makes the sound zones less efficient for frequencies between the original points. Thus, to improve performance, uniformly spaced frequency data from the research of Kleiven and Zinserling should be acquired. An estimation of the improvement was made by only considering the data points in which the lowest

sound pressure levels occured, i.e. near the original frequency points. The separations for that case, 22.7 dB and 30.6 dB, were indeed very close to the separations achieved by the method in previous work, 22.4 dB and 31.5 dB.

The performance of the two homomorphic filtering methods was less satisfying, with separations $10 - 20$ dB lower than the method in previous work, for impulse input. This is most likely explained by the large deviations from the phase responses of the desired complex source strengths. While the delay time introduced by the filtering was decreased from 40 ms to under 5 ms, the phase responses needed for correct interference patterns in the zones were destroyed. Hence, even though the filters' magnitude responses were unchanged for the homomorphic mirroring method, the separation was made worse. The incorrect interference patterns may also explain why the output magnitude spectra, with impulse input, deviated so much from a flat response, leading to audible differences in sound quality. The homomorphic windowing method was an attempt at a compromise between phase following and magnitude following, aiming for better phase following by allowing more deviations from the desired magnitude response. While tendencies of this was seen, the phase was still far off from the desired phase response. One way to improve the phase following could be to window the cepstrum differently, for example allowing some noncausal zeros. As for the mirroring method, the filter would then need delay in order to be causal. However, the necessary delay time might be reduced compared to the delay needed in the mirroring method.

In terms of sound quality, the more compact impulse responses of the homomorphic windowing method compared to the mirroring method caused the sounds to have a more distinct feel. This together with the lower initial delay comprise the only apparent benefits of using the presented homomorphic filtering techniques for a sound zone application.

Chapter 6

Conclusion

This thesis presents three methods for calculating impulse responses from frequency responses. These are evaluated subjectively and objectively by a sound zone simulation. The mirroring method, which is based on inverse transformation of an altered version of the frequency response, performed a separation between the addressed zone (driver seat) and the silent zones (passenger and back seats) of 17.0 dB and 24.2 dB respectively, or 5 dB and 7 dB less than for the method used by Kleiven and Zinserling. The time delay introduced by the filtering was 40 ms, and the flatness of the frequency response in the driver zone was ± 2 dB in the range 100 Hz − 775 Hz. Two homomorphic filtering methods were developed for the purpose of reducing initial delay by creating minimum phase versions of the desired magnitude response. However, this gave too large deviations from the desired phase responses, and the separations achieved by these methods are significantly worse than those from the mirroring method. The sound quality of the auralisations in this thesis suffer from a small frequency range. However, the mirroring method reproduced sound in the addressed zone which for some signals was indistinguishable from the input sound.

Chapter 7

Future Work

For future work in this project, it is recommended that transfer functions and source strengths with uniform frequency spacing are generated, to eliminate the interpolation errors believed to cause the performance difference between the results in this thesis and those from previous work. Further, it is of great interest to adapt the current work to the full audio range, as well as to study how different source configurations affect the sound zone performance. When a full range simulation has been done successfully, the work may be focused on implementing the filters in a digital signal processor, and investigating the possibilities and limitations of the current sound zone strategy in a real-world application.

References

[Bij 14] Bijsterveld, K., Cleophas, E., Krebs, S., and Mom, G.: *Sound and Safe: A History of Listening Behind the Wheel*, New York, 2014

[Goo 12] Goose, S., and Arman, F.: *System and Method for Creating Personalized Sound Zones*, US 8,126,159B2, United States Patent, 2012

[Wil 12] Wilder, R. L., Wells, W. R., Mattice, H, and Griswold, C. W.: *Gaming machines with directed sound*, US 6,638,169B2, United States Patent, 2003

[Brb 12] Barbagallo, M.: *Modulation and demodulation of audio sound beams using ultrasonic parametric array loudspeakers*, Chalmers University of Technology, 2008

[Jac 11] Jacobsen, F., Olsen, M., Møller, M., and Agerkvist, F.: *A Comparison of Two Strategies for Generating Sound Zones in a Room*, ICSV18, Rio de Janeiro, Brazil, 2011

[Bri 12] Brix, S., Sladeczek, C., Franck, A., Zhykhar, A., Clausen, C., and Gleim, P.: *Wave Field Synthesis Based Concept Car for High-Quality Automotive Sound*, AES 48TH International Conference, Munich, Germany, 2012

[Che 13] Cheer, J., and Elliott, S.: *Design and Implementation of a Personal Audio System in a Car Cabin*, Proceedings of Meetings on Acoustics, Vol. 19, 2013

[Mac 06] Machida, A.: *Eye-Based Control of Directed Sound Generation*, US 2006/0140420, Patent Application Publication, 2006

[Kle 12] Kleiven, S., and Zinserling, B.: *Local Sound Separation in Cars*, AES 48TH International Conference, Munich, 2012

[Opp 65] Oppenheim, A. V.: *Superposition in a Class of Nonlinear Systems*, Massachusetts Institute of Technology, 1965

[ITU 03] ITU-T: *G.114: One-way Transmission Time*, International Telecommunication Union, 2003

[Sel 01] Selesnick, I. W., and Schuller, G.: *The Discrete Fourier Transform*, The Transform and Data Compression Handbook, Boca Raton, 2001

[Dru 94] Druyvesteyn, W. F., Aarts, R. M., Asbury, A. J., Gelat, P., and Ruxton, A.: *Personal Sound*, Proceedings of the Institute of Acoustics, Vol. 16, pp. 571, 1994

[Wol 09] Wolfe, P. J.: *Homomorphic Signal Processing. ES 157/257: Speech and Audio Processing*, Harvard Engineering and Applied Sciences, 2009

[Chi 77] Childers, D. G., Skinner, D. P. and Kemerait, R. C.: *The Cepstrum: A Guide to Processing*, Proceedings of the IEEE, Vol. 65, No. 10, 1977

[Phi 08] Phillips, C. L., Parr, J. M. and Riskin, E. A.: *Signals, Systems, and Transforms*, Pearson Education Ltd., London, 2008

[Smi 07] Smith, J.o.: *Introduction to Digital Filters with Audio Applications*, http://ccrma.stanford.edu/ jos/filters/, online book, 2007 edition, accessed 140416

[Mad 09] Madhu, N.: *Note on Measures for Spectral Flatness*, Electronics Letters , Vol. 45, No. 23, pp. 1195-1196, November 2009

[Bar 11] Barata, J. C. A., and Hussein, M. S.: *The Moore-Penrose Pseudoinverse. A Tutorial Review of the Theory*, Instituto de Física, Universidade de Saõ Paulo, 2011

[Cro 83] Crochiere, R. E., and L.R. Rabiner.: *Multirate Digital Signal Processing*, Englewood Cliffs, 1983

Appendix A

Figures

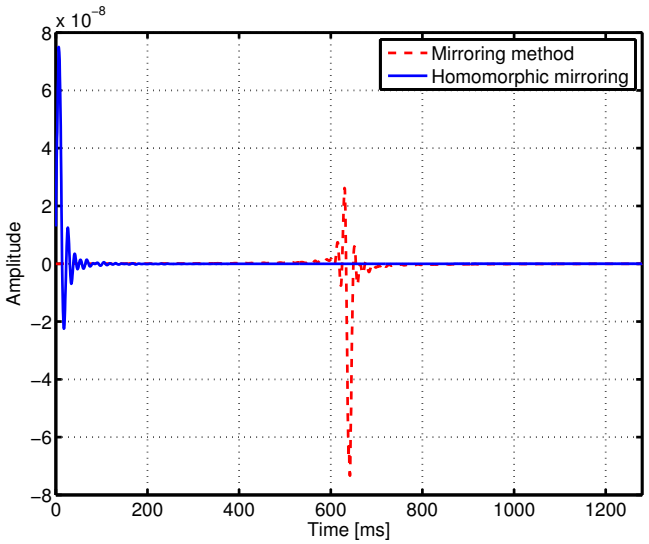


Figure A.1: Impulse response of the mirroring method and the homomorphic mirroring method. 2048 filter coefficients.

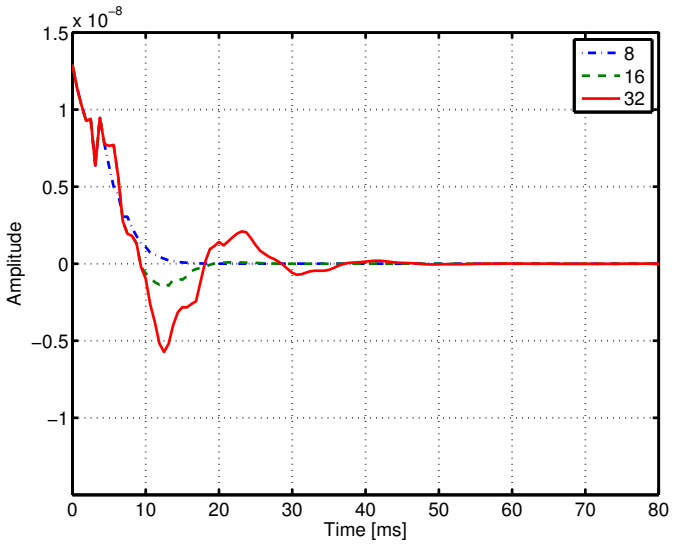


Figure A.2: Impulse response of the homomorphic windowing method for varying effective filter orders. 2048 filter coefficients.

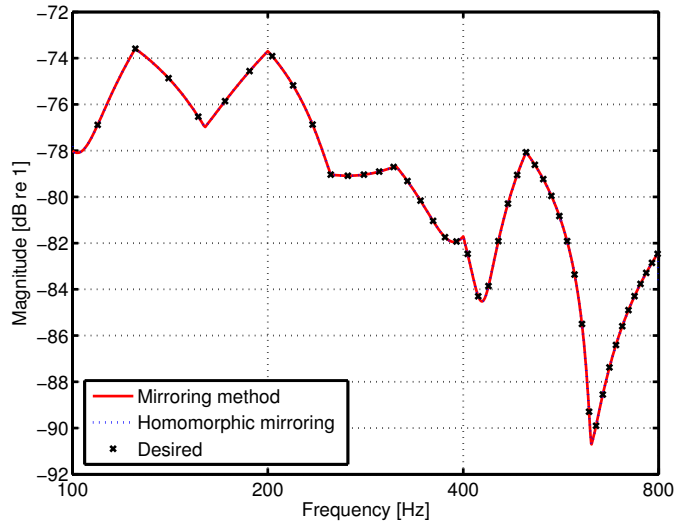


Figure A.3: Magnitude response of the mirroring method and the homomorphic mirroring method compared to the desired magnitude response. 2048 filter coefficients.

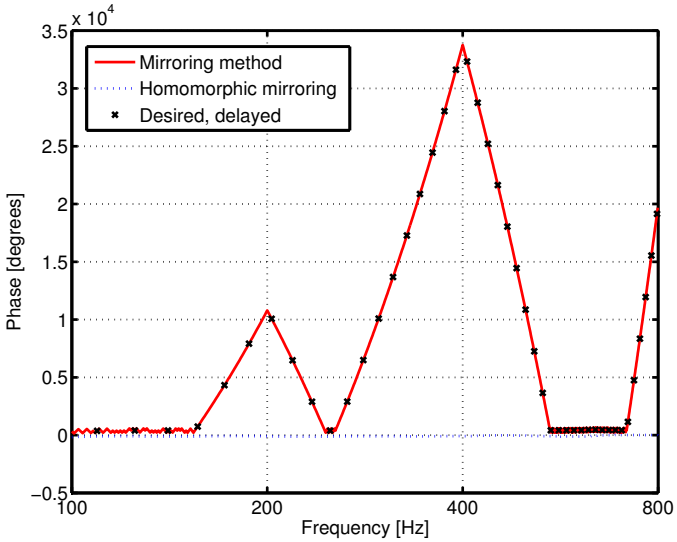


Figure A.4: Phase response of the mirroring method and the homomorphic mirroring method compared to the desired magnitude response delayed by 640 ms. 2048 filter coefficients.

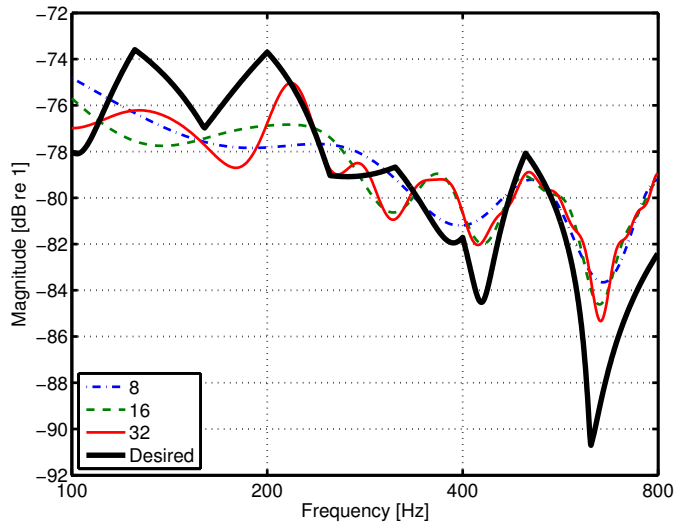


Figure A.5: Magnitude response of the homomorphic windowing method compared to the desired magnitude response. 2048 filter coefficients.

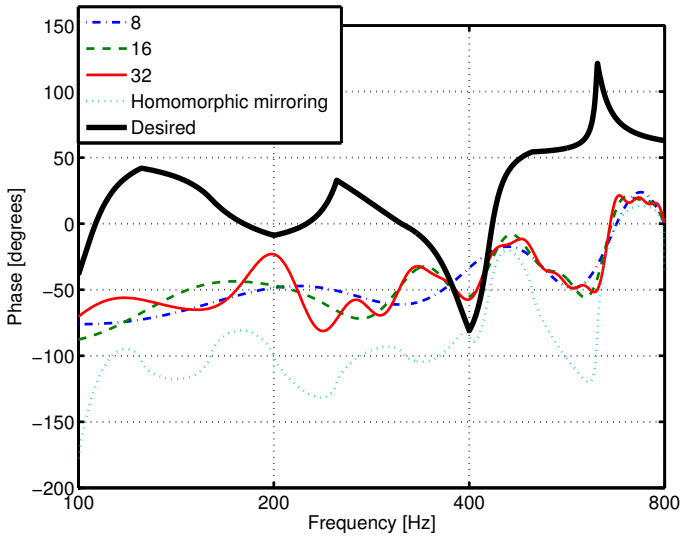


Figure A.6: Phase response of the homomorphic windowing method and the homomorphic mirroring method compared to the desired phase response, with no delay included. 2048 filter coefficients.

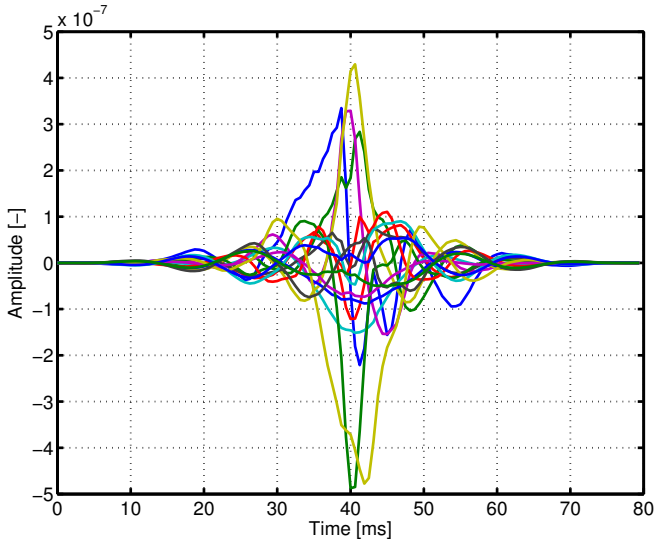


Figure A.7: Impulse responses for all sources, calculated by the mirroring method. 128 filter coefficients.

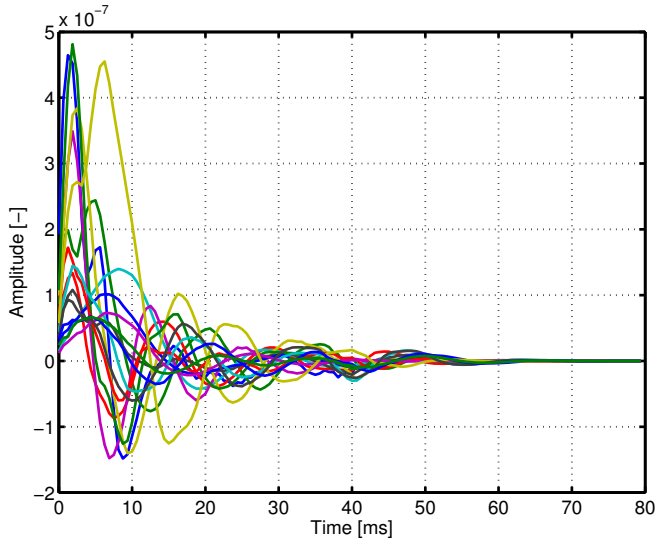


Figure A.8: Impulse responses for all sources, calculated by the homomorphic mirroring method. 128 filter coefficients.

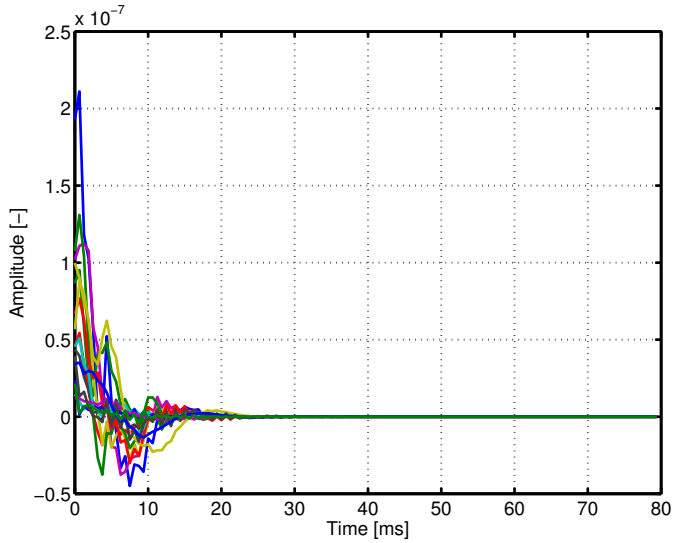


Figure A.9: Impulse responses for all sources, calculated by the homomorphic windowing method with a window width of 16. 128 filter coefficients.