HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering

Riitta Väänänen

Efficient Modeling and Simulation of Room Reverberation

This Master's thesis has been submitted for official examination for the degree of Master of Science in Espoo on June 3, 1997

Supervisor of the Thesis	Instructor of the Thesis
Professor Matti Karjalainen	MSc Jyri Huopaniemi

Table of Contents

List of Abbreviations

BEM	boundary element method
DFT	discrete Fourier transform
DSP	digital signal processing
EDR	energy decay relief
FEM	finite element method
FFT	finite element method
FIR	finite impulse response
HRTF	head related transfer function
IIR	infinite impulse response
ITD	interaural time difference
MFDN	multiple feedback delay network
MLS	maximum length sequence
STFT	sshort-time Fourier transform

List of Symbols

c	velocity of cound
C	modulus of compression
δ	Dirac's delta function
D_f	frequency density of modes
D_t	time density of reflections
f	frequency
f_S	sampling frequency
ϕ_{lr}	cross correlation function of l and r
ϕ_{ll}	autocorrelation function of l
$h(n)$	impulse response of a discrete time system
$H(z)$	transfer function of a discrete time system
L_p	sound pressure level
p	sound pressure
$_0$	reference sound pressure level
Ψ	velocity potential of sound
θ	angle of azimuth
r	radius
ϱ	mass density
ρ	density of reflections
t	time
T_{60}	reverberation time
τ	time delay
Θ	phase response
u	particle velocity
V	volume of room
ω	angular frequency

1. Introduction

Modeling of *room reverberation* by means of audio signal processing can be used to add spatial properties to sound. Its applications areas are, for example, commercial reverberators which can be used in recording studios, home sound reproduction systems and in film industry. It can also be a part of *auralization* process where it can be understood as means to simulate reverberation of a defined room. Modeling of room reverberation can also be used for architectural design of rooms or concert halls or evaluating the acoustical properties of existing halls.

The main aim of this master's thesis was to find algorithms by which room reverberation can efficiently be simulated using digital signal processing. The reverberation modeling was mainly dealt with as being a way to implement *spatialization* by creating artificial reverberation to sound in order to improve the listening experience of recorded or synthesized sound. Also commonly used computational methods for modeling of room acoustics were reviewed.

1.1 Sound Spatialization

Fig. 1.1 illustrates the whole field of spatialization which covers both artificial room reverberation modeling and auralization. Auralization means that the binaural and spatial information present in a room when sound is performed in it, e.g., music or speech, is simulated and reproduced to give an impression of listening to the sound inside a same kind of a room. In auralization the directional information of the sound source is rendered with the aid of *head-related transfer functions* (HRTFs) which are simulated to give three-dimensional information of the direct sound and first reflections (Kleiner, Dalenbäck and Svensson 1993). The use of HRTFs in Fig. 1.1 is referred to as *directionalizing*.

Auralization is based on physical modeling of a binaural room response either by means of geometrical room acoustics or mathematical modeling of a sound field. In artificial reverberation modeling the task is to find algorithms which produce a perceptively good reverberation effect to sound. In both auralization and artificial reverberation simulation the result of modeling is an impulse response which can be implemented by means of digital signal processing. In artificial reverberation

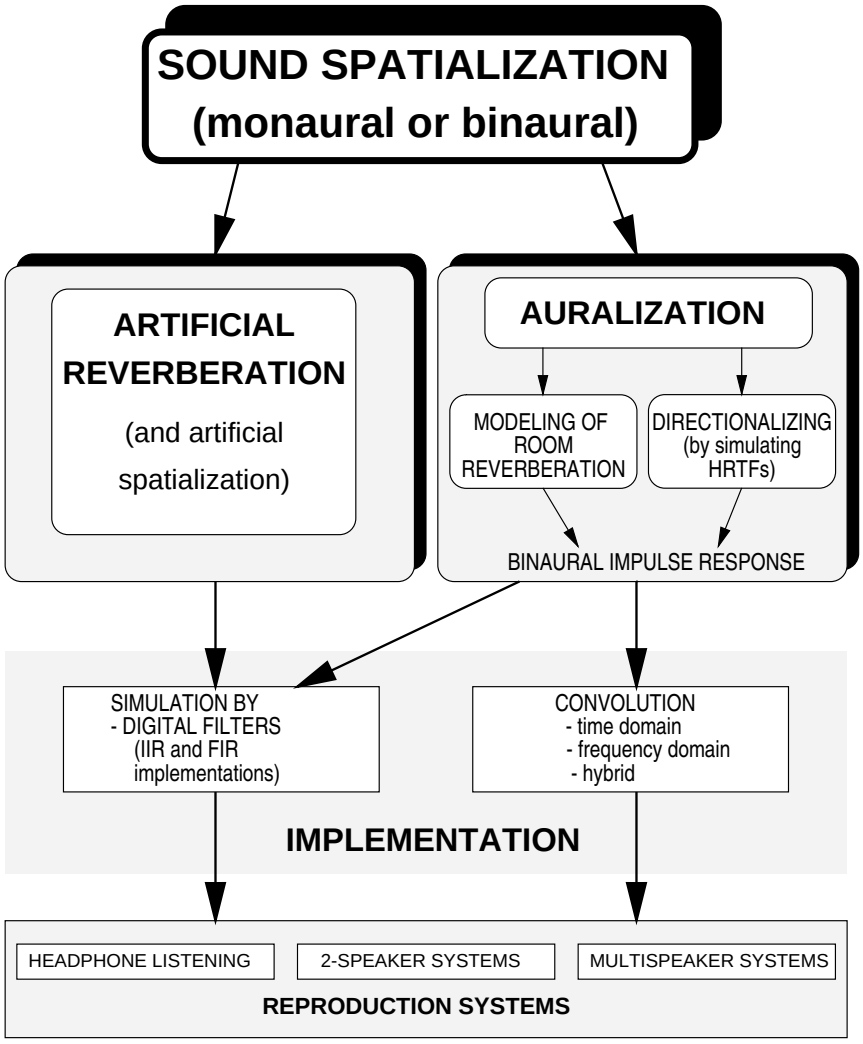


Figure 1.1: Different areas of sound spatialization, their implementation methods, and reproduction systems where they can be applied to.

modeling, different *FIR* and *IIR filters* are used to implement the desired algorithm. In auralization, the binaural response is used for convolving it with the source signal, which can be implemented in time or frequency domain or as hybrid methods which combines time- and frequency-domain convolutions. In auralization, same kind of algorithms are often used for late reverberation modeling as in the case of totally artificial reverberation.

1.2 Room Response Modeling

The impulse response between two points in a room is normally divided into two parts, one consisting of the direct sound and early reflections and the other of late reverberation (Fig. 1.2). In room acoustics modeling, these two parts are often im-

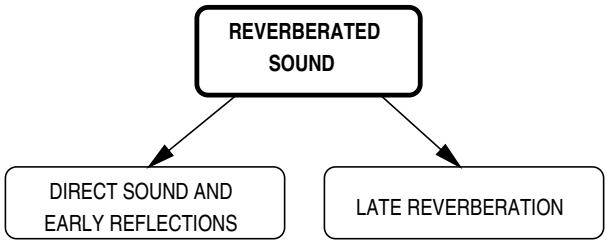


Figure 1.2: Division of the artificially reverberated sound to early reflections and late reverberation.

plemented in different ways. This is done to make the simulation computationally efficient in order to render real-time processing of sound (Schroeder 1962), (Schroeder 1970).

The early reflections part of the room impulse response is often implemented with an FIR filter as a tapped delay line (Schroeder 1970). The late reverberation is most often simulated by recursive filters since they provide longer impulse responses with smaller amount of computation than the FIR filters.

The late reverberant sound field is normally considered nearly *diffuse* which means that the magnitude and the direction of the sound pressure are randomly distributed. Also the reflection density is so high that the individual reflections cannot be perceived (Barron 1993). In a good sounding room the late reverberation is normally characterized by an exponentially decaying envelope curve as a function of time, and a nearly flat frequency response which leads to a task of producing an impulse response which resembles an exponentially decaying white noise sequence (Jot 1992b), (Jot and Chaigne 1991). Although this requirement for late reverberation seems like a simple aim to obtain it is still the most difficult part in artificial simulation of room reverberation when the processing power is restricted, like in the case of real-time applications. Despite the great amount of work various researchers have done on simulating natural sounding late reverberation since the 1960's the main problem remains to produce reverberation which does not contain flutter echo nor has a metallic sound or which has a smooth response to transient-like sounds.

1.3 Outline of Chapters

This thesis consists of two parts, the first of which (Chapters 1-4) is about theory of room acoustics and its modeling methods. The second part of this thesis is covered by Chapter 5 which concentrates on the practical side of this project, namely, simulation of different reverberation algorithms and their evaluation based on the qualitative properties of their impulse responses.

In Chapter 2 the fundamentals of room acoustics are presented starting from the basic acoustical quantities and the wave equation in room acoustics. The definition and characteristics, as well as measuring methods of the room impulse response

are discussed. The most popular methods for computational modeling of room acoustics are dealt with, and the standardized room acoustic parameters are shortly reviewed. Chapter 3 begins with the basics of digital signal processing. The basic digital filter elements and structures used in audio signal processing, and the most popular audio effects and their digital implementations are studied. Chapter 4 deals with different methods used for producing a reverberation effect to sound. The requirements of good quality reverberation are presented, as well as the theoretical limitations of the reverberator structures for producing certain reflection and modal densities of the reverberator response. Means to implement artificial spatialization, i.e., an impression of a broader sound field by adding incoherence to difference reproduction channels are also considered. In Chapter 5 the reverberators which have been simulated in this work are presented and evaluated according to objective criteria, and results from different reverberators are compared. Chapter 6 concludes the aims and results of this project, and presents the application areas where the results of reverberation modeling can be applied to. Future work and improvements are also considered.

2. Room Acoustics

Acoustics of an enclosed space, such as a room or a hall, depends on the architectural configuration of the space, and on the absorbing properties of the materials covering the surfaces inside that space. The sound that is perceived by a listener is a combination of the direct sound, and the sound which is reflected off the walls, the floor, the ceiling and other obstacles which the sound interferes with before arriving at the listener. The behavior of sound and thus the acoustic quality of a room can be defined from the impulse response between a sound source and a receiver in the room, since the response includes all the information that is needed to predict how the sound is perceived by the listener.

This chapter begins with definitions of sound and the basic acoustical quantities. The wave equation, and its significance in room acoustics, as well as the concept of normal modes of a room are presented. The propagation of sound in an enclosed space and the different parts of a room impulse response are dealt with in detail, and also different methods for measuring the impulse response are briefly explained. The most popular room acoustic modeling methods are described, containing different geometric and mathematical modeling methods. The last topic of this chapter is the room acoustic parameters, which can be computed from an impulse response of a room to be evaluated, and thus they compose an objective criterion for quality of room acoustics.

2.1 Basic Acoustical Quantities

2.1.1 Sound Pressure and the Particle Velocity of Sound

Sound is defined as change of pressure which is propagating as waves in the transmitting medium and can cause a hearing sensation. The medium can be either gas, liquid or solid material. In gas and liquid the sound travels as longitudinal pressure waves but in solid material it can also cause transversal motion of the medium, i.e., motion which is perpendicular to the direction of the sound propagation. Only the longitudinal component of sound can be perceived by the human ear, since the transversal motion does not necessarily cause changes of pressure in the medium.

Sound can be heard assuming that the frequency of sound is within the frequency range audible to humans (approximately 16 Hz - 16 kHz) and that the amplitude, i.e., the magnitude of pressure fluctuation of the sound exceeds the listener's threshold of audibility which is the minimum pressure amplitude which can cause a hearing sensation. This leads to a definition of *sound pressure level* which is a logarithmic ratio between the actual sound pressure and a reference level which is an average of the threshold of audibility of different people. The reference level is chosen to be $p_0 = 20$ μPa which is approximately the threshold pressure of sound at 1000 Hz frequency and corresponds to 0 dB. Thus the sound pressure level in decibels is expressed as

$$L_p = 20 \log \mid \frac{p}{p_0} \mid \quad . \tag{2.1}$$

Sound pressure can be solved from Euler's equation:

$$\nabla p(t) = -\varrho \frac{\partial u(t)}{\partial t}, \tag{2.2}$$

where $u(t)$ is *particle velocity* and ϱ is *mass density* of the medium. Particle velocity of sound means the velocity at which the particles of the medium move about their state of equilibrium. It can be described by the following equation

$$\nabla \cdot u(t) = -C \frac{\partial \rho(t)}{\partial t}, \tag{2.3}$$

where $u(t)$ is the particle velocity, C is the modulus of compression, t is time and ϱ is the mass-density of the medium.

2.1.2 Velocity of Sound

Velocity of sound in gas is given by equation

$$c = \sqrt{\frac{nRT_0}{M}} \ , \tag{2.4}$$

where M is the weight of n mols of the gas, R is a gas constant and T_0 is the temperature of the gas. Thus the velocity of sound depends only on the square root of the temperature of the gas. An approximation of velocity of sound in the air is given by equation

$$c \approx 331.4 \frac{\mathrm{m}}{\mathrm{s}} \cdot \sqrt{1 + \frac{t}{273}}, \tag{2.5}$$

where t is the temperature of the air in centigrades.

2.2 Wave Equation in Room Acoustics

Equations 2.2 and 2.3 can be used to derive *the wave equation* of the velocity potential which can be used to mathematically solve the sound field in a room. The

three-dimensional wave equation is:

$$\nabla^2\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}, \tag{2.6}$$

where the velocity potential can be expressed as

$$\psi = -\nabla u. \tag{2.7}$$

The solutions of the wave equation for the velocity potential in a room with parallel walls and containing lossless medium are

$$\psi = D \cdot \cos(\frac{l\pi}{a}x)\cos(\frac{m\pi}{b}y)\cos(\frac{n\pi}{c}z), \tag{2.8}$$

where D is a constant gain, a, b and c are the dimensions of a rectangular room, and l, m and n are arbitrary integers. The solutions of Equation 2.8 with different values of l, m and n are called *eigenfrequencies of a room*. The values of l, m, and n of which one is 1 and the other two are zeros correspond to the lowest eigenfrequencies in each of the three directions. These are frequencies the wavelengths of which are twice the length of the dimensions a, b and c of the room. The next eigenfrequencies (l, m or n equals to two) correspond to wavelengths 3/2 times the room dimension etc. In practice this means that a rectangular room with smooth, reflective walls does not enforce the sound at low frequencies with a high frequency density which is necessary for pleasant sounding room acoustics. To obtain high eigenfrequency density, the walls should not be parallel and the surfaces of the walls should be irregularly shaped.

2.3 Room Impulse Response

Sound propagates in air as longitudinal pressure waves which are characterized by their frequency, amplitude and direction. In free space sound spreads to the environment without interfering with any objects, and the $1/r$ law can approximately be applied to the sound pressure. This means that the sound pressure is inversely proportional to the distance from the sound source, thus every time the distance is doubled the sound pressure level is decreased by 6 dB. In free space a listener can localize the sound source according to the sound level at the receiving point and to the directional information which is obtained from the different received signals at the two ears of the listener.

When a sound source and a receiver are in an enclosed space such as a room or a concert hall, the received sound is affected by the impulse response between the source and the receiver. An impulse response is the time-domain function $h(t)$ by which the sound signal $x(t)$ is affected when it travels from the source to the receiver, yielding the output signal $y(t)$ of the system, i.e., the signal at the receiver. The output signal $y(t)$ can be expressed as a time-domain convolution of the sound signal and the impulse response:

$$y(t) = x(t) * h(t), \tag{2.9}$$

where $*$ is used to denote convolution which is defined as

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau. \tag{2.10}$$

The impulse response is defined as a sound signal at the receiver when the excitation signal is Dirac's delta function $\delta(t)$, which ideally has a length zero in the time domain, and an infinite magnitude at time zero, and the energy over the whole time-domain $= 1$:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1 \tag{2.11}$$

The impulse response of an enclosed space characterizes its acoustical properties and depends on the size and the geometric shape of the room, and on the reflective properties of all the surfaces inside the room.

Since an impulse response of a room can only be defined between the source and the receiver, it always depends on their locations inside the room. Nevertheless, when the receiver is a person he/she receives a different signal at each ear. This leads to a definition of a binaural impulse response which is audible to the listener and actually consists of two separate impulse responses, which contains directional information of the direct sound and the early reflections, and gives an impression of the size and the shape of the space, and the position of the listener inside the room.

A response of a room to sound is often treated as two separate parts, the first comprising of direct sound and early reflections and the second of late reverberated sound. The division is done since often different room reverberation modeling methods and different filter types are used to implement these two parts. An impulse response between two points inside a room defines what the acoustics of a particular room sounds like. In Fig. 2.1 an impulse response of an imaginary room is illustrated in a form of an echogram which indicates the time delays and the magnitudes of the reflections proportional to the excitation sound.

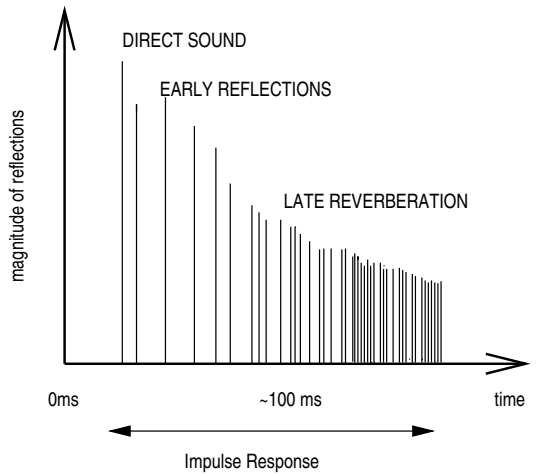


Figure 2.1: Impulse at the sound source, direct sound, early reflections and late reverberation at the receiver.

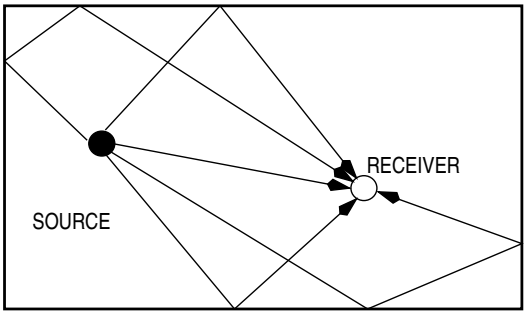


Figure 2.2: Direct sound and first reflections.

It can be seen that the envelope curve of an impulse response of a room is decaying as a function of time and that the reflection density of the late reverberation is greater than that of the first reflections.

2.3.1 Direct Sound and Early Reflections

Direct sound is the sound coming from the source to the receiver directly without interfering with any objects. Its amplitude is determined by the $1/r$ law, and it depends on the distance between the source and the receiver as well as on the properties of the transmitting medium. Early reflections are reflected from the side walls, the ceiling, and the floor and arrive at the listener within approximately 100 ms after the direct sound (Barron 1993) (Fig. 2.2). Together with the direct sound the early reflections are important for localization of the sound source and for giving the listener an impression about the size and shape of the room, and about the place and the orientation of the listener inside the room (Kuttruff 1993).

The early reflections are often simulated with the aid of *geometrical room acoustics* where the sound is considered propagating like rays perpendicular to the wave front and that the wave behavior of the sound is neglected. Thus phenomena which are typical for waves, such as diffraction and dispersion are ignored in geometrical room acoustics. There are several methods developed to model the propagation of sound using information about room geometry, like the geomrtrical modeling methods called image source method and ray-tracing. These and some other methods for simulating sound fields in rooms are explained in Chapter 2.4.

2.3.2 Late Reverberation

Late reverberation part of the room response is characterized by an increased reflection density and often by an exponentially decaying envelope curve (Schroeder 1962). A typical assumption to the late reverberant sound field is that it is nearly diffuse which means that the direction and the amplitude of the sound are randomly distributed (Barron 1993). The late reverberant sound field does not provide directional cues about the sound, and therefore it is often modeled statistically without

a need to know the exact sound pressure distribution at each point in a room at any moment. This decreases the requirements for the computational performance of the simulation system compared to exact modeling of the sound field or the impulse response, since information about the listener's position does not have to be involved in the late reverberation modeling. Also, statistical room acoustical parameters, such as reverberation time, are based on analyzing the late reverberation decay of a room. These parameters will be studied in Chapter 2.5.

The quality of room acoustics is complex to evaluate. The desired acoustical properties of a room, a hall or an auditorium largely depend on what purpose it is designed for and what kind of sound material is meant to be performed in it. Many room acoustic parameters have been defined to measure the objective quality of the room acoustics and will be presented in Chapter 2.5. One important criterion for a pleasant sounding room is that its late reverberant sound field is nearly diffuse, the reflection density is high enough and that the magnitude response has a high modal density to prevent coloration of sound (Schroeder 1962). These properties have always been aimed at when designing digital reverberators. When sound fields are simulated, the characteristics of the late reverberation define how the sound is enforced and how natural it sounds.

2.3.3 Measuring of Room Impulse Response

An impulse response between a source and a receiver can approximately be measured using an impulse-like sound, such as a gun shot or a wideband noise burst, as an excitation signal. The problem when using this kind of signals is a poor signal-to-noise ratio achieved, and also that the measurements are not completely repeatable due to the radiation pattern of the sound source. It is difficult to create an impulse sound with a sufficiently high total energy compared to the level of the background noise produced by the measuring equipment and other sounds which are present in the room.

A method which provides a considerably better signal-to-noise ratio when determining an impulse response of a room is to use a pseudo-random noise sequence such as in *the maximum-length sequence* (MLS) method (Schroeder 1965),(Borish and Angell 1983). In this method the excitation signal is a stationary digital pseudo-random noise sequence. The impulse response is rendered by calculating the cross correlation function between a periodical pseudo-random sequence and the output signal of the system, i.e., the signal at the receiver. A disadvantage of the MLS method is that it can only be applied when a system is time-invariant which is not necessarily a characteristic to an impulse response of a room at low frequencies (Vorländer and Bietz 1994).

2.4 Computational Modeling of Room Acoustics

In the simulation of room acoustics the aim is to model a physical space, either an existing or an imaginary room, and its response to sound at the places of interest. The room response functions resulting from the models are used for reproducing the spatial impression of the room by convolving it with sound, or they can be utilized for predicting the acoustic behavior and quality of a room (Kuttruff 1993).

In this section, methods of geometrical and mathematical modeling of room acoustics are reviewed. In geometrical modeling the aim is to obtain an impulse response which gives an impression about the size and shape of the modeled space, or directional information about the sound source to the listener. The propagation of sound is simplified by neglecting the wave behavior of it to avoid massive computation. Mathematical modeling, on the other hand, is based on solving the wave equation in a room, and phenomena typical for waves, such as, dispersion and diffraction are also modeled.

2.4.1 Geometric Modeling of Room Acoustics

Two common methods for geometric modeling of room acoustics are *ray-tracing* (Krokstad, Strom and Sorsdal 1968) and *image-source method* (Gibbs and Jones 1972). They both take room geometry and the spatial configuration of absorbing materials into account. The room geometry is most often modeled as a combination of flat surfaces. The behavior of sound waves is modeled as rays which are perpendicular to the wave front of a propagating sound. The free field attenuation of sound pressure occurs according to the $1/r$ law as the pressure is propagating in the medium and the reflections at the surfaces occur by the Snell's law. The total distance-dependent attenuation of an individual sound ray from the source to the receiver is

$$a = \frac{k^m}{r}, \tag{2.12}$$

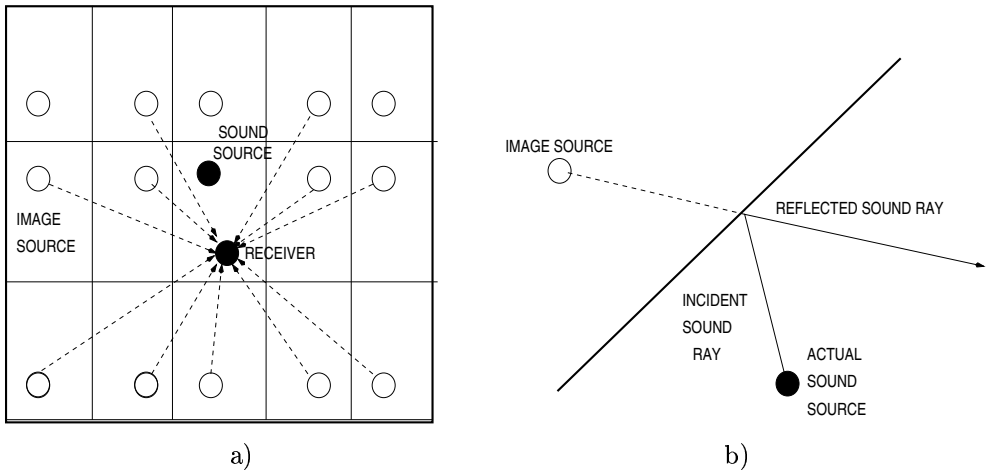
where k is the reflection coefficient of the reflecting surfaces, m is the amount of reflections on the path of the propagating sound, and r is the distance the sound travels. The time delay corresponding to a sound ray is

$$\tau = \frac{r}{c}, \tag{2.13}$$

where c is the velocity of sound. A more natural impression of the reflections can be obtained by adding lowpass filtering to the individual reflections to simulate frequency dependent air absorption to sound.

Image-Source Method

Image-source method has been explained, e.g., in (Gibbs and Jones 1972), (Allen and Berkley 1979) and (Borish 1984). When the image-source method is used for room acoustics modeling the surfaces inside the room are presented as combinations of flat surfaces. The room walls are considered as reflectors which absorb part of



a) b)

Figure 2.3: a) Actual sound source in a room and image sources. b) Sound reflecting from a surface.

the sound energy and reflect the rest of it. The angle of reflection is the same as the angle of incidence and thus the reflected wave can be replaced by a ray coming from an image-source which is a virtual mirror source behind the reflective wall. The principles of forming the image sources and the reflections are illustrated in Fig. 2.3

Image-source method simplifies strongly sound propagation in a room because it does not involve sound diffusion when the sound reflects from irregular surfaces. This is a defect of the image-source method since it is most often desirable to have rough surfaces in a room to increase the diffusion of the sound field. Nevertheless it is a popular method for computer simulation of room acoustics and especially for simulation of the early reflections because when only low-order reflections are modeled. When the order of the image sources is increased the amount of reflections is increased exponentially with the order. This makes the model impractical for modeling other parts of the impulse response than just the few lowest order reflections.

Ray-Tracing

Ray-tracing method has been explained, e.g., in (Krokstad et al. 1968). In ray-tracing the sound source is modeled by a sound radiation pattern which is composed of vectors pointing outward from the source the vector length indicating the intensity of the radiation to that direction. The geometrical shape of the room is presented as a combination of flat surfaces, like in the case of image-source method. The sound pressure is defined at the places of interest by summing together the sound rays passing through finite volume elements (Fig. 2.4). The same rules for the sound attenuation and reflections are applied to the ray-tracing method as for the image-source method.

Often the diffusion at the reflective surfaces is also modeled to yield a more flutter free and natural reverberation. Compared to the image-source method ray-tracing requires more calculation in the beginning of the impulse response but it is more

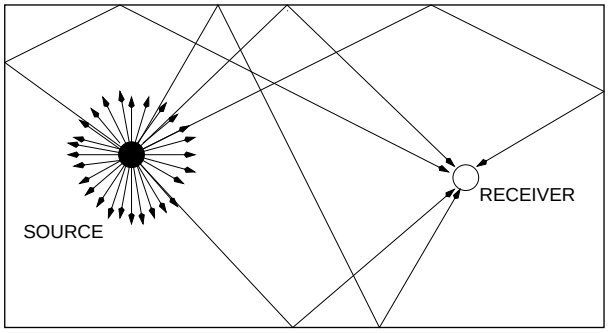


Figure 2.4: The ray-tracing method: The sound pressure at the receiver is a superposition of the rays passing the receiver.

efficient when the latter part of the impulse response is modeled. That is because the amount of computation increases linearly whereas that of an image-source method increases exponentially. Since the number of the rays passing the observed volume decreases with time the accuracy of ray-tracing modeling decreases as well.

2.4.2 Element Methods and Finite Difference Methods

Element methods and *difference methods* in room acoustic modeling are based on solving the wave equation, in other words the aim is to define a discrete sound field in a room. *Finite element methods* (FEM) and *boundary element methods* (BEM) are methods which can be used to model the propagation of low frequency sound in a room since they take the wave behavior of sound into account. In these methods the whole room (in FEM) or its boundaries (in BEM) is divided into elements which are the smallest spatial units where the sound pressure is computed. The size of the elements defines the smallest wavelength and thus the highest frequency for which the model gives accurate results. The computational complexity of these systems depends on the amount of elements used and this is why it is most practical for modeling the sound field in a room at low frequencies. In (Choi and Tachibana 1989) an impulse response in a sound field is obtained by a two-dimensional finite element method and in (SenGupta 1991) the resonances of an enclosure are studied by a three-dimensional finite element method. An example of BEM is given in (BaiMingsian 1992) where the method is used for finding acoustic resonances of enclosures.

Difference methods are based on finite-difference approximation of the time and space derivatives of the wave equation, as explained in (Botteldooren 1995). An example of a difference method is a *three-dimensional waveguide mesh* where the room is modeled as a 3-D grid of junctions which are connected to the neighboring junctions by waveguides (Van Duyne and Smith 1993). The sound pressure at each junction is calculated as a sum of the sound pressures coming along the waveguides leading to that junction. The boundary conditions are modeled with specific termination nodes which implement the properties of the walls and other surfaces

influencing the propagation of sound (Savioja, Rinne and Takala 1994), (Savioja, Backman, Järvinen and Takala 1995).

2.4.3 Hybrid Methods

In hybrid methods different ways to simulate the distribution of sound are combined in order to optimize the accuracy of the model and the computational requirements. In (Vorländer 1989) it is suggested that the ray-tracing is used for finding the visible image sound sources from which the actual reflections are calculated. This is also a principle used in Odeon room acoustics modeling program (Naylor and Rindel 1994). A hybrid of time- and frequency-domain methods is considered in (Savioja et al. 1995), for modeling the propagation of low-frequencies by mathematical methods and high frequencies by methods based on geometrical modeling. The problem in these kind of hybrids is how to combine the time and frequency domain responses together.

2.5 Room Acoustical Parameters

Room acoustical parameters describe the properties of a room in terms of measurable quantities. They can be divided to monaural (single-channel) and to binaural (two-channel) parameters which also take the directional information and binaural aspects of the sound into consideration. Monaural parameters are calculated from a room impulse response measured with one omnidirectional microphone. Binaural parameters, on the other hand, are calculated from two separate impulse responses either with the aid of directional microphone measurements or by dummy head or real head measurements.

The room acoustical parameters which are reported in the ISO 3382 standard (ISO/DIS-3382 1995) for measurement of the reverberation time are shortly presented below. Monaural parameters reported in the standard are

- *Reverberation time (T_{60})*
- *Early Decay Time (EDT)*
- *Strength (G)*
- *Clarity (C)*
- *Deutlichkeit (D50)*
- *Center time (TS)*

And the binaural parameters are

- *Lateral Energy Fraction (LEF)*

- *Inter-aural Cross-correlation Coefficient (IACC)*

According to the standard above all room acoustical parameters should be calculated for bandpass filtered impulse responses using either octave band or third octave band filtering.

2.5.1 Monaural Room Acoustical Parameters

Reverberation time T_{60}

Reverberation time T_{60} is the time during which the sound pressure level decreases 60 dB from the steady-state level after the sound has abruptly stopped. It can be approximated by using *Sabine's formula*

$$T_{60} = \frac{0.163V}{A}, \tag{2.14}$$

where V is the volume of the room and A the total absorption area $\sum \alpha_i S_i$, where α_i are the absorption coefficients of the areas S_i. Another way to define reverberation time is using *Eyring's reverberation formula*

$$T_{60} = 0.161 \frac{V}{-ln(1-\bar{\alpha})S} \tag{2.15}$$

Eyring's formula is a better approximation of reverberation time when the absorption coefficient approaches unity. This happens for example in an anechoic chamber where Sabine's formula still gives a finite reverberation time whereas the value derived from Eyring's formula approaches zero which is closer to the real situation.

Reverberation time cannot be solved very reliably from these equations. When defining the reverberation time of a particular room it is calculated from the impulse response of the room which is normally obtained by integrated impulse response method using maximum length sequences for measuring the impulse response (Schroeder 1965), (Schroeder 1979). Depending on the level of the background noise reverberation time is calculated from a backward integrated impulse response from -5 dB level to either -25 dB level or -35 dB level, zero level being the level of the direct sound at the receiver. The resulting reverberation time is labeled either T_{20} or T_{30}, respectively.

In Table 2.1 some preferable reverberation times for different types of performed sounds are listed according to Barron (1993). It should be shortest for speech which contains fast variations and therefore a long reverberation time would reduce the intelligibility of speech. On the other hand for music it has to be longer to reinforce the performance sufficiently.

Early decay time EDT

Early decay time EDT is obtained in the same manner as T_{60}, except that the time is measured from the initial 10 dB decay and that time is multiplied by six. According

Type of sound material	Recommended T_{60}
Organ music	> 2.5
Romantic classical music	$1.8 - 2.2$
Early classical music	1.6-1.8
Opera	1.3-1.8
Chamber music	1.4-1.7
Drama theatre	0.7-1.0

Table 2.1: Recommended reverberation times for different sound material (Barron 1993).

to (ISO/DIS-3382 1995) both the T_{60} and EDT should be measured since EDT is subjectively more important to the perceived reverberance, and T_{60} describes more the physical properties of the room (Barron and Lee 1988).

Strength G

The sound Strength G is the logarithmic ratio of the squared and integrated sound pressure of the measured impulse response to the that of the response measured 10 meters from the same sound source in a free field:

$$\mathrm{G} = 10 \log_{10} \frac{\int_0^\infty p^2(t) dt}{\int_0^\infty p_{10}^2(t) dt}. \tag{2.16}$$

Clarity C and C_{50}

Many room acoustical parameters are based on the balance between early and late arriving sound energy, as explained in (Barron and Lee 1988). A common formula for these parameters is

$$\mathrm{C_{te}} = 10 \log_{10} \frac{\int_0^{te} \mid h(t) \mid^2 dt}{\int_{te}^\infty \mid h(t) \mid^2 dt}, \tag{2.17}$$

where time te in milliseconds is varied according to the parameter which is to be calculated. Normally the upper limit of the integral in the denominator is the reverberation time. Below the C_{te} values mentioned in the ISO 3382 are shortly described. C_{50} and C_{80} describe the ratio between the useful early reflections for speech and music, and the reflections which are heard as late reverberation. C_{80} is usually called clarity (for music) and it is also denoted by C or K.

Deutlichkeit D

Another nearly same form of an integral formula as above is used to describe clarity for speech where the denominator is the whole energy of the impulse response. This parameter is called *deutlichkeit* D and it is calculated from the equation:

$$\mathrm{D} = 10 \log_{10} \frac{\int_0^{50} \mid h(t) \mid^2 dt}{\int_0^\infty \mid h(t) \mid^2 dt}. \tag{2.18}$$

Center time TS

Another monaural integral parameter of room acoustics is center time TS:

$$\mathrm{TS} = 10 \log_{10} \frac{\int_0^\infty t \mid h(t) \mid^2 dt}{\int_0^\infty \mid h(t) \mid^2 dt}. \tag{2.19}$$

2.5.2 Binaural Parameters of Room Acoustics

Spatial parameters of room acoustics are based either on measurements done with a directional microphone or on *inter-aural cross correlation* (IACC) which is derived from dummy head or real head recordings. The definitions of the parameters from directional microphone measurements resemble many of the monaural room acoustic parameters because they are based on the energy of the impulse response. The IACC on the other hand is based on measurements in the two ears of a listener and thus it does not depend only on the acoustics of the room but also on the structure of the head and the outer ears of the listener.

Lateral Energy Fraction (LEF)

Lateral Energy Fraction is defined as the ratio between the energy of the lateral reflections measured with a figure-of-eight microphone, to reflections measured with an omnidirectional microphone:

$$\mathrm{LEF} = \frac{\int_{5\mathrm{ms}}^{80\mathrm{ms}} h_L^2(t) \cos\theta dt}{\int_0^{80\mathrm{ms}} h^2(t) dt}, \tag{2.20}$$

where $h_L(t)$ is the impulse response from the figure-of-eight microphone and $h(t)$ from the omnidirectional microphone. The first 5 ms from the figure-of-eight microphone has been left out from the computation to eliminate the effect of the direct sound, and the time limit 80 ms is to yield the lateral energy fraction only to the early reflections.

The need for calculating the lateral energy arises from the fact that halls where there is more energy in the reflections from the side walls than from the front and back walls are most often preferable (Barron 1971). The reason for this is that the side wall reflections cause incoherence between the two ears of the listener whereas front-back reflections cause same sound to both ears which is not considered a pleasant experience.

Inter-aural Cross-correlation Coefficient (IACC)

Inter-Aural Cross-correlation Coefficient (IACC) is based on the cross-correlation function of the signals measured at both ears of a listener, as explained in (Ando 1985). These signals also include the influence of the outer ear and the head on the response to the source signal and reflected sounds. *Interaural cross-correlation function* (ICF) is a cross-correlation of the binaural signals measured either with a

dummy head or a real head and it is defined as:

$$\phi_{lr}(\tau) = \lim_{x\to\infty} \frac{1}{2T} \int_{-T}^{T} f_l(t) f_r(t+\tau) dt \tag{2.21}$$

where $f_l(t)$ and $f_r(t)$ are the signals at the left and right ears, respectively.

The value of ICF indicates the direction of the sound relative to the head of the listener. The correlation is strongest for the directicf sound when the listener is facing towards the sound source, i.e., the signals at both ear entrances are similar for $t = 0$. It is weakest in a diffuse sound field where there should be no correlation between the two points according to the directions and the phases of the sound pressures. This can be seen from the normalized interaural cross-correlation for N first reflections.

The autocorrelation functions of $f_l(t)$ and $f_r(t)$ are

$$\phi_{ll}(\tau) = \lim_{x\to\infty} \frac{1}{2T} \int_{-T}^{T} f_t(t) f_l(t+\tau) dt \tag{2.22}$$

and

$$\phi_{rr}(\tau) = \lim_{x\to\infty} \frac{1}{2T} \int_{-T}^{T} f_r(t) f_r(t+\tau) dt. \tag{2.23}$$

The normalized interaural cross-correlation function ICF for N first reflections can be derived from those:

$$\phi_{lr}^{(N)}(\tau) = \frac{\sum_{n=0}^{N} A_n^2 \phi_{lr}^{(n)}(\tau)}{\sqrt{\sum_{m=0}^{N} A_m^2 \phi_{ll}^{(m)}(0) \sum_{k=0}^{N} A_k^2 \phi_{rr}^{(k)}(0)}}. \tag{2.24}$$

Normally IACC is defined as the maximum value of the ICF for interaural delays τ which are shorter than 1 ms:

$$\text{IACC} = \max \mid \phi_{lr}(\tau) \mid, \quad \mid \tau \mid \leq 1\text{ms} \tag{2.25}$$

In general the value of IACC indicates the diffuseness of a sound field and can be used for describing the late reverberation properties of a room.

2.6 Summary

In this chapter the fundamentals of room acoustics were presented in the extent which is useful for further understanding of this thesis. First the basic acoustical quantities were defined, and the significance of wave equation in room acoustics was dealt with. The concept of room impulse response was explained and methods for room response and room acoustics modeling were discussed. Finally the room acoustical parameters which are important for objective evaluation of the quality of room acoustics were dealt with.

3. Digital Audio Signal Processing

This chapter begins with a short introduction to the operation and the basic elements of digital filters, after which the principles of producing audio effects, such as, echo and reverberation, flange, chorus and phasing are explained.

3.1 Digital Filters

Digital filters can be implemented by computational algorithms, e.g., on *digital signal processors* (DSP). They provide effective computation on signals which in the case of audio processing are sampled and quantized data of some sound material. Digital filters implement time delays, and multiplication and summing operations derived from mathematical algorithms designed for filtering the data (Oppenheim and Schafer 1975). Digital filters can be illustrated as block diagrams (Fig. 3.1). Typically elements of these diagrams are *delays* and *summing* and *multiplication* operators. The time delays are represented with rectangular boxes containing information about the amount of unit delays each of which has a duration of one sampling period. Multiplication operators are marked triangles indicating the direction of the signal flow, and the summing operations with circles containing a + -sign. The direction of signal flow within a filter is marked with arrows, the input $x(n)$ normally coming from the left and the output $y(n)$ of the filter going to the right. Fig. 3.1 is an illustration of a simple digital filter which takes $x(n)$ as an input and sums an unchanged version of that signal with a delayed and gained version of

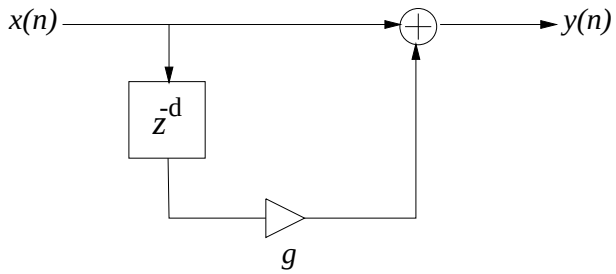


Figure 3.1: A block diagram of a digital filter containing a delay of length d samples, one summing and one multiplication operation.

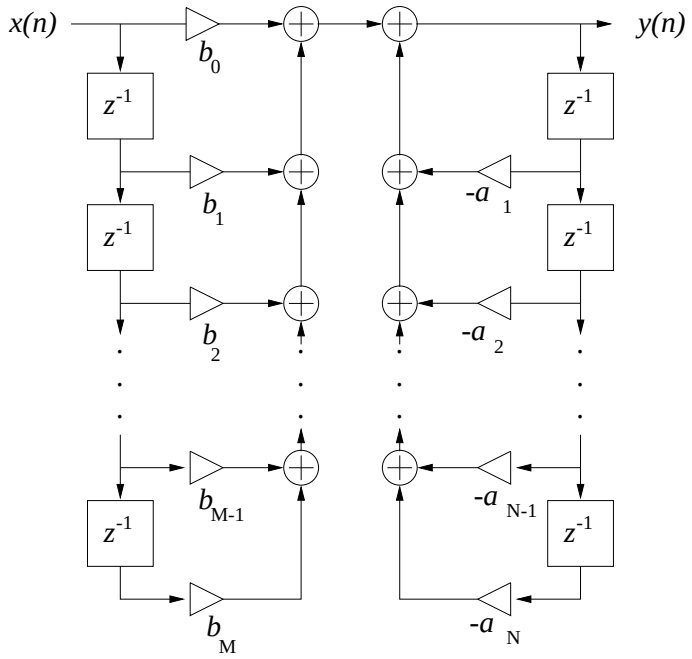


Figure 3.2: Structure of an IIR digital filter with multiple feedforward and feedback loops.

itself. The output of this filter as a function of its input is $y(n) = x(n) + gx(n-d)$. An IIR digital filter which is composed of multiple *feedforward* and *feedback* loops is illustrated in Fig.3.2. The feedback loops with coefficients $a_1 ... a_N$ cause the output of the filter to recirculate in the filter. This kind of a filter can be expressed by a difference equation

$$y(n) = \sum_{k=0}^{M} b_k x(n-k) - \sum_{k=1}^{N} a_k y(n-k), \tag{3.1}$$

where output $y(n)$ is the *response* of a filter to the input $x(n)$. An impulse response of a discrete time system is defined as the response of a system for a discrete unit impulse signal $\delta(n)$, i.e.,

$$\delta(n) = \begin{cases} 1 & \text{when } n = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

A digital filter containing only feedforward loops (with gains $b_1, ..., b_M$) is called a *finite impulse response* filter (*FIR* filter) since its output for an impulse becomes zero after the delay $k = N$. An *infinite impulse response* filter (*IIR* filter) has an infinite impulse response because of its feedback loops and thus the signal recirculates in the filter theoretically an infinite amount of time producing an infinitely long impulse response.

The z-domain transfer function of an IIR filter derived from equation 3.1 is

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 + \sum_{m=1}^{N} a_m z^{-m}}. \tag{3.3}$$

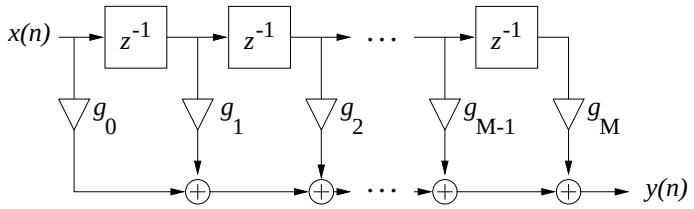


Figure 3.3: A tapped delay line structure of an FIR filter.

The frequency response of an IIR filter is characterized by the *zeros* and the *poles* of the transfer function. These are the values of z which cause the numerator or denominator to become zero. The filter has resonances at frequencies that correspond to the poles of a filter. An IIR filter is stable when any bounded input produces a bounded output. A linear and time invariant system is stable if and only if

$$\sum_{k=-\infty}^{\infty} |h(k)| < \infty. \tag{3.4}$$

A digital filter is stable if the poles of the transfer function are inside the unit circle in the z-domain.

When digital signal processing is used for room reverberation modeling an important task is to optimize the amount of processing and the needed amount of delay memory especially at high sampling rates. With FIR filters an impulse response of any form could be implemented as a *tapped delay line*, where for each sample of the impulse response there is a gained output in a corresponding position of the delay line (Fig 3.3) (Schroeder 1970). In room response modeling this means discrete convolution of the room impulse response and the source signal and it requires a large amount of computations depending on the sampling rate and the length of the impulse response, and this is a reason to look for IIR implementations for reverberation modeling to obtain longer impulse responses with smaller amount of computation and memory than in the case of FIR filters.

3.2 Digital Audio Effects

In the area of sound reproduction digital filters can be used to implement different effects on sound, e.g., reverberation and echoes, flange, chorus and phasing, which take advantage of delays by mixing the direct sound with an attenuated and delayed version of itself. In multi-channel systems also directional information can be added to sound by amplitude or phase panning between the loudspeakers, or by using HRTFs to modify the frequency response of the sound sources to give directional cues to the listener. By digital filtering it is possible to create effects which cannot occur in natural environments, such as, *reverse reverberation* or changing the pitch of sound (t.c n.d.), but they will not be discussed here. Some of the digital audio effects will be shortly reviewed like they are explained in (Orfanidis 1996). Also

Smith (1982) has reported about these digital effects and Levine (1996) has done research on effects processing on subband audio. In this section most emphasis is put on the common filter structures used for producing reverberation since it is the main topic of this thesis.

3.2.1 Echo and Reverberation

Echoes are sounds which are reflected off the walls and other objects the propagating sound interferes with. They arrive to a listening point delayed and attenuated relative to the direct sound, which is the sound coming straight from the sound source to the listener. Thus a single echo can be recreated with a simple FIR filter already illustrated in Fig. 3.1. Echoes, which are incident in a room with many reflecting surfaces, can be simulated with a tapped delay line which is an FIR filter containing unit delays in cascade (Fig. 3.3). The outputs from the line are taken after delays of different durations, multiplied with different coefficients and summed together. The lengths of the delays are proportional to the lengths of the paths the sound travels from the source to the listener and the gains are inversely proportional to the corresponding path lengths and they also depend on the reflection coefficients of the materials on the reflecting surfaces.

A single delay filter of Fig. 3.1 has the lowest fundamental resonance frequency which corresponds to the delay of the filter. In other words the filter amplifies sound when the length of the delay is a multiple of the wavelength of the sound in delay units. The resonance frequencies can be expressed as

$$f_n = nf_sD, \qquad n = 1, 2, ...D - 1, \tag{3.5}$$

where f_s is the sampling frequency and D is the delay of the filter. This filter structure is called an *FIR comb filter* since it produces peaks in the frequency domain with even intervals, and maximum attenuation in the middle of adjacent resonance frequencies. These resonances can be considered to simulate the normal modes of a room with parallel walls (Orfanidis 1996). FIR reflections are often used as basic elements to create early reflections and they are implemented in a form of tapped delay lines the delay lengths and the taps of which are obtained from an image-source model (Ch. 2.4.1).

A comb filter can also be implemented as an IIR filter by changing the feedforward delay branch to a feedback loop. The impulse response of an IIR comb filter of Fig. 3.4 is $y(n) = gy(n - D) + x(n)$, and it contains several peaks in the time-domain impulse response having intervals of d samples, as can be seen in Fig. 3.5.

The transfer function of this filter is

$$H(z) = \frac{1}{1 - gz^{-d}}, \tag{3.6}$$

and to fulfill the condition of stability g has to be smaller than 1. The impulse response, the frequency response, the pole-zero plot and the group delay of an IIR

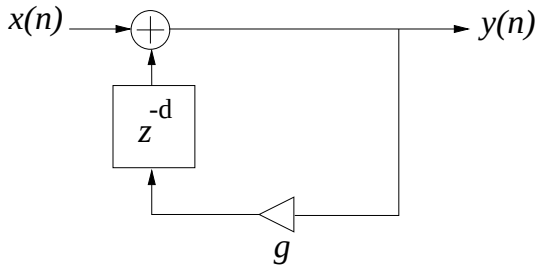


Figure 3.4: IIR implementation of a comb filter.

comb filter with a feedback loop gain 0.5 and delay of 20 ms are illustrated in Fig. 3.5. By setting $z = e^{j\omega}$ we get the group delay as a function of ω:

$$\tau_g(\omega) = -\frac{\partial\Theta(\omega)}{\partial(\omega)}, \tag{3.7}$$

where $\Theta = arg[H(e^{j\omega})]$ is the phase response of the system and $\omega = 2\pi f$ is the angular frequency. Group delay describes *dispersion*, i.e., how the different frequency components pass through the filter with different speeds. The impulse response of a comb filter is a decaying sequence of reflections with gains $1, g, g^2, g^3, ...,$ and it can easily be seen that if the absolute values of the coefficient g is less than one the impulse response decays as a function of time, and if it is greater than 1 the filter becomes unstable and the magnitude of the impulse response increases with time.

Whereas FIR comb filters can be used to simulate discrete reflections in a room IIR realizations are often used to produce late reverberation because of the long impulse response that can be produced by them with smaller amount of computation than in the case of FIR filters. When connecting IIR comb filters in parallel they provide a physical model for *normal modes* of a room which means the standing plane waves between parallel walls, gained at the frequencies defined by Eq. 2.8.

From the delay line lengths of comb filters, and the number of the filters the reflection density and the modal density, the latter being the theoretical average number of resonances per hertz, can be defined (Jot and Chaigne 1991), (Jot 1992b). By adjusting the gains in the feedback loops the length of the impulse response, and thus the reverberation time of the parallel comb filters can be modified.

With a small number of parallel comb filters it is not possible to obtain reflection density which is sufficient to give an impression about diffuse late reverberation, where the reflections cannot be heard as discrete echoes anymore. Also the modal density should be high enough to avoid *coloration* of sound, i.e., emphasis of certain frequencies, and this is obtained by keeping the delay lengths close to each other and by increasing the total delay length of the comb filters (Jot and Chaigne 1991). The reflection and modal densities could be increased by increasing the number of comb filters in the parallel connection. Nevertheless, in real-time implementations the processing and the memory capacity of the processor sets limits to the number of parallel comb filters and often they are not used alone to produce reverberation of sufficient reflection and modal densities.

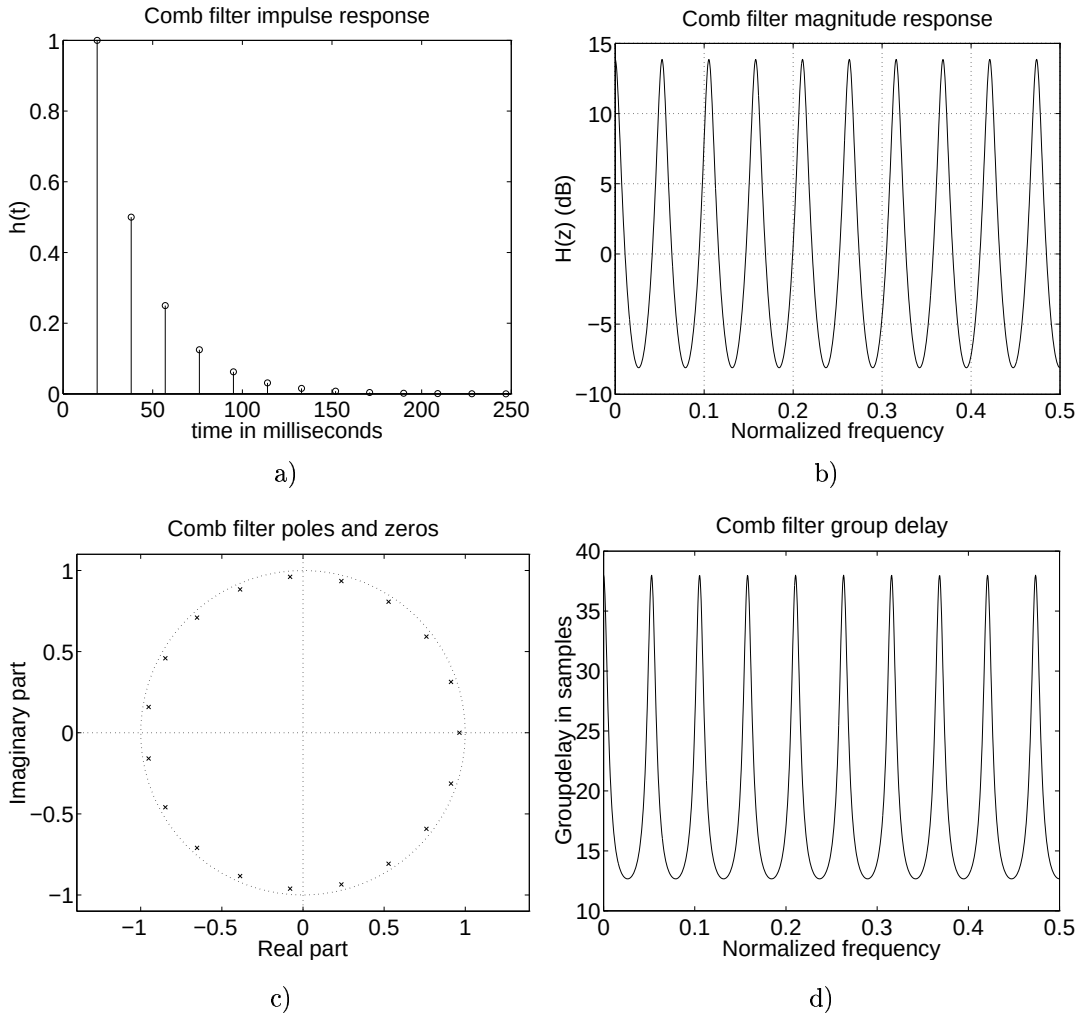


Figure 3.5: a) Impulse response, b) magnitude response, c) pole-zero plot and d) group delay of an IIR comb filter with a feedback loop gain $g = 0.5$ and delay of 20ms (sampling frequency f_s=1000Hz)

One solution for increasing the reflection density of a parallel comb filter structure suggested by Schroeder (1962) is to use allpass filters which have a flat magnitude response but produce a decaying pulse train as an impulse response as well. An IIR comb filter can be changed to a comb-allpass filter by adding a feedforward loop from the input of the comb filter to the output of the delay line (Fig. 3.6).

This filter structure has an impulse response which resembles that of a comb filter with the difference that the first pulse is negative because of the negative feedforward connection and that the gains of the pulses are $-g, 1-g^2, (1-g^2)g, (1-g^2)g^2, ...$ The frequency response of a comb-allpass filter of Fig. 3.6 is

$$H(z) = \frac{z^{-d} - g}{1 - gz^{-d}}, \tag{3.8}$$

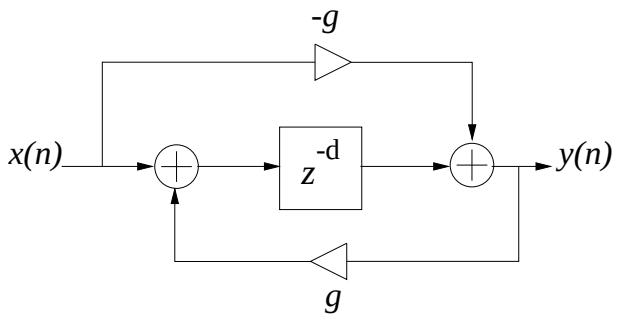


Figure 3.6: A block diagram of an comb-allpass filter which is obtained from a comb filter by adding a feedforward loop from the input to the output of the filter.

and when the numerator is divided by $z^{-}d$ the transfer function becomes

$$H(z) = z^{-d}\frac{1 - gz^{d}}{1 - gz^{-d}}, \tag{3.9}$$

where it can be seen that the absolute value of $H(z)$ is 1, i.e., the filter has a unity gain at all frequencies.

The impulse response, the frequency response, the pole-zero plot and the group delay of this kind of an allpass filter is illustrated in Fig. 3.7. In reverberation simulation, comb-allpass filters are often cascaded with parallel comb filters since they pass through all frequencies equally. Because of the exponentially decaying pulse train they produce as an impulse response, they have a diffusive effect on individual peaks in the impulse response of the comb filters, thus increasing the reflection density of the reverberation. Also they cause a dispersive effect to the filtered signal because of the variation in the group delay as a function of frequency (Fig. 3.7). The decay of modes of the responses of both the comb and the comb allpass filters are plotted in Fig. 3.8. The plots have been achieved by making a *Short Time Fourier Transform (STFT)* to the impulse responses and by backward integrating the individual frequency bins. This procedure can be used for illustrating and evaluating the time-frequency structure of an impulse response of a reverberator as J.-M. Jot has explained in (Jot 1992a) and it will be discussed in more detail in Ch 4. In these plots the resonances in the impulse response are detectable as well as their continuation. It can be seen that even if the long term magnitude response of the comb allpass filter is flat its short time spectrum contains resonances which resemble those of a comb filter. Thus it can be assumed that also the comb allpass filters can cause perceptual coloration to sound.

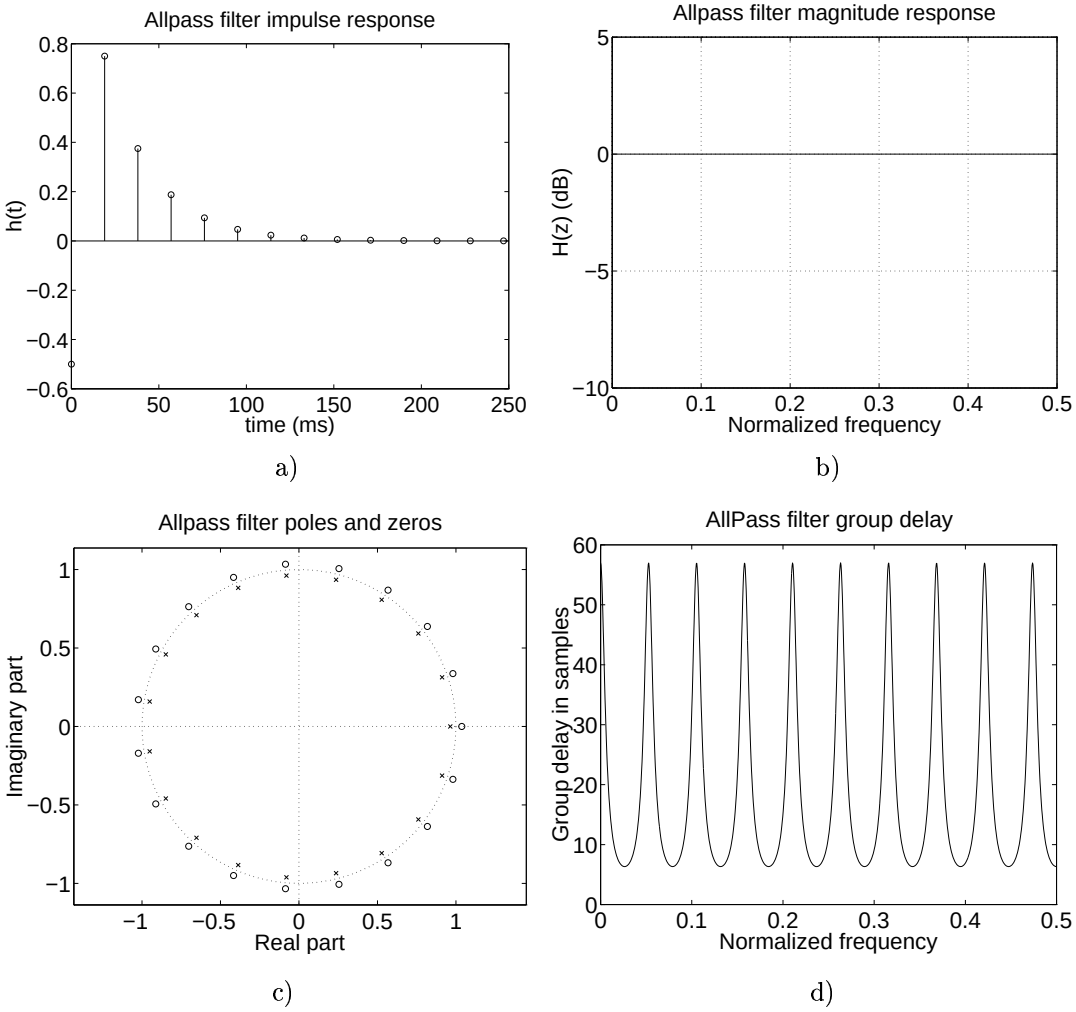
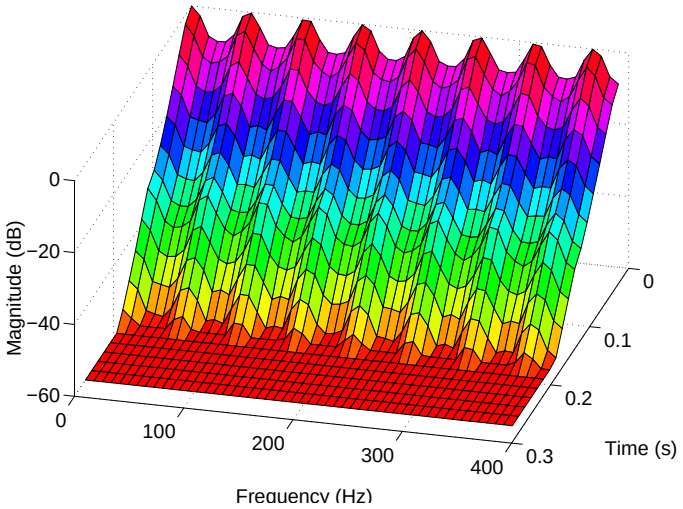
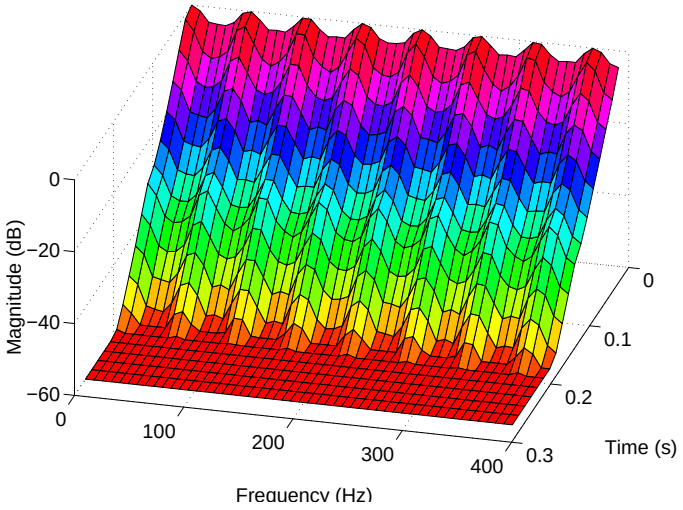


Figure 3.7: a) Impulse response, b) magnitude response, c) pole-zero plot and d) group delay of a comb-allpass filter with a gain $g = 0.5$, delay 20 ms and (sampling frequency=1000 Hz)



a)



b)

Figure 3.8: A backward integrated STFT plot of a) a comb filter and b) of a comb allpass filter.

3.2.2 Flange, Chorus and Phasing

A *flanging* effect can be implemented with a single FIR comb filter similar to that in Fig. 3.1 with the difference that a fixed length delay d is replaced with a time varying delay. This kind of a filter can be described with an equation

$$y(n) = x(n) + ax(n - d(n)), \tag{3.10}$$

where a is a gain of the feedforward loop and $d(n)$ is a sinusoidally varying delay:

$$d(n) = \frac{D}{2}(1 - cos(2\pi F_d n)), \tag{3.11}$$

where F_d is a low frequency around 0.2 Hz (Levine 1996) causing the minima of the magnitude response to slightly move back and forth on the frequency axis. Gain

a controls the depth of the notches of the magnitude response and the minimum and the maximum gains appear at intervals which are multiples of f_s/d and odd multiples of $f_s/2d$, respectively.

Chorusing is an effect which makes a single musician sound like a group of musicians performing the same piece at the same time, but slightly non-synchronously (Levine 1996). This is obtained with a parallel FIR comb filter structure where in each feedforward branch there is a time-varying delay of a length about $25 - 30$ ms (Levine 1996). For example in a case of two delayed versions of the direct sound thus imitating three musicians the filter is characterized by an equation

$$y(n) = x(n) + a_1(n)x(n - d_1(n)) + a_2(n)x(n - d_2(n)). \tag{3.12}$$

In this case the delays are varied randomly with a low frequency below 1 Hz.

Phasing is an effect implemented with a *notch filter* and it has the same principle as a filter used to produce flanging with the difference that instead of having multiple notches in the magnitude response a notch filter has a single time-varying attenuation at a notch frequency. This kind of a filter causes a phase shift to the input signal at the notch frequency which is controlled with a low-frequency oscillator. It is characterized in addition to its notch frequency also by a Q-factor which describes the narrowness of the notch. A notch filter can be implemented by designing a *notch polynomial* starting with placing zeros and poles of the filter inside the unit circle. First the zeros are placed like in the flanging filter and then the poles are placed at the same frequencies as the zeros so that they cancel the effect of zeros at other frequencies except at the desired notch frequency. A more detailed description of the design and implementation of a notch filter is presented in (Orfanidis 1996).

3.3 Summary

In this chapter the fundamentals of digital signal processing were dealt with, to the extent that is necessary for understanding the digital filter structures presented in the next chapter for reverberation modeling. Some of the digital audio effects, which reverberation modeling is a part of, were also explained.

4. Digital Reverberators

Digital reverberators are used to add late reverberation properties of room acoustics to sound at low processing cost. The most straightforward way to produce reverberation to sound is to convolve it with a sampled impulse response of a real room. This would result in a great amount of computation since each sample of the source signal would require a number of operations which is equal to the number of samples in the measured impulse response, i.e., the sampling frequency multiplied by the duration of the response in seconds. For a two second response and a sampling frequency of 44.1 kHz this would yield $2 \cdot 44100 = 88200$ operations per input sample for each channel. A frequency domain convolution requires less operations but the processing delay makes it impractical for real time implementations. Jot, Larcher and Warusfel (1995) have divided the reverberation rendering into three approaches which are 1. *hybrid time-domain and frequency-domain convolution*, 2. *a hybrid approach which combines multirate filter banks with parametric modeling* and 3. *artificial reverberation using feedback delay networks.* The first and the second approach are shortly explained and more emphasis is put to the recursive filtering approach which is the most researched one in reverberation modeling.

4.1 Different Approaches in Reverberation Modeling

4.1.1 Hybrid Time-Domain and Frequency-domain convolution

Gardner (1995) has introduced a method for eliminating the input-output delay of the frequency-domain convolution which is about twice the length of the convolved signal and caused by the transformations between the time and the frequency domains. In this method the impulse response is divided into blocks of different sizes which are convolved with the input signal separately in the frequency domain. The results are delayed according to the position of the block in the filter response, and summed together. The first block is not delayed and thus it is convolved in the time domain by a direct form FIR filter. The response is divided into blocks as follows:

The first response block after the time domain convolution block is of length N. This block is delayed $2N$ because of the input-output delay in the transformation

and thus the time-domain convolution has to be of length $2N$. The following blocks have to be started at a delay of at least twice the block size to ensure a uniform demand on the processor. When the block sizes are kept as powers of two, they become $2N$, N, N, $2N$, $2N$, $4N$, $4N$ etc.

4.1.2 Filter Banks and Parametric Modeling of Room Impulse Response

Schoenle, Fliege and Zoelzer (1993) have proposed a method for real-time simulation of room impulse response by subband processing. In multirate analysis a filter bank is used for wavelet decomposition to obtain a time-frequency analysis of the impulse response signal. The decomposition gives high frequency resolution and low time resolution to low frequency signals and vice versa for the high frequency signals.

In this method the subband approximation is done on both the input signal $x(n)$ and the response of the system to that signal $y(n)$ yielding the subband signals $x_1(n_1)...x_p(n_p)$ and $y_1(n_1)...y_p(n_p)$. The subband responses are approximated by parametric transfer functions $H_i(z) = B_i(z)/A_i(z)$ producing output signals $\hat{y}_1(n_1)...\hat{y}_p(n_p)$. These signals are combined in a synthesis filter bank and an approximation $\hat{y}(n)$ is obtained. When the input signal $x(n) = \delta(n)$ the output is $h(n)$ which is the impulse response of the system and $\hat{h}(n)$ is the approximation of the impulse response. The parametric transfer functions are implemented by a cascade of a comb filter ($H_{C,i}(z) = 1/(1 - g_i z^{-N_i})$) and an FIR filter approximation $H_{M,i} = b_0 + b_1 z^{-1} + ... + b_{M_i} z^{-M_i}$ of the first reflections:

$$H_i(z) = \frac{b_0 + ...b_{M_i} z^{-M_i}}{1 - g_i z^{-N_i}}. \tag{4.1}$$

The gains g_i are obtained by fitting a straight line to the logarithm of the exponentially decaying envelope of the subband impulse response:

$$\log(g) = \frac{\log(h_e(l_1)) - \log(h_e(l_2))}{l_1 - l_2}, \qquad l_1 < l_2, \tag{4.2}$$

where h_e is the envelope curve of the subband impulse response and l_1 and l_2 are the coordinates on the time axis at which the line is fitted to the logarithmic h_e curve.

4.1.3 Artificial Reverberation by Recursive Digital Filters

Most of the digital reverberators are based on Schroeder's work on artificial reverberation. In article (Schroeder 1962) he has analyzed the properties of natural reverberation that occurs in rooms and suggests a filter which artificially reproduces those properties. After his pioneer work on that area many researchers have referred to his work when trying to find efficient algorithms for implementing artificial reverberation which better imitates natural and good sounding room acoustics.

Most digital reverberators are combined structures of comb and comb-allpass filters. Comb filters produce resonances to the sound at even frequency intervals.

Allpass filters on the other hand have complex phase responses and pass all the frequencies through with the same gain in the magnitude response. The impulse response of a comb-allpass filter is a decaying pulse sequence and thus they have a multiplicative effect on the echo density of a comb filter when connected in series. They are often used to increase the reflection density of the comb filter structures. They have also been used alone cascaded, like in (Schroeder 1962) and (Moorer 1979) or they can be nested, i.e., connected inside one another to produce late reverberation (Gardner 1992b, 1992b).

In this chapter the properties which make artificial reverberation sound natural and pleasant are dealt with. Methods and digital filter structures used for producing reverberation effect on sound are reviewed, and also results from measurements of two commercial reverberators are presented.

4.2 Requirements for Digital Reverberators

A common task in producing pleasant and natural sounding digital reverberation is to design a filter structure having an impulse response which reaches its maximum level fast in the beginning of the response and has an exponentially decaying envelope curve (Fig. 4.1). To make the filter simulate a diffuse sound field the echo density must have a fast buildup and it must be high enough so that the individual reflections cannot be perceived (Schroeder 1962), (Moorer 1979), (Jot and Chaigne 1991), (Gardner 1992b). In real acoustic environments the reflection density grows exponentially as a function of time, and this is approximated by an equation

$$\rho = \frac{4\pi c^3}{V} t^2, \tag{4.3}$$

where ρ is the number of reflections per second, c is the velocity of sound, V is the volume of the room and t is the time in seconds (Moore 1990).

The required reflection density depends on the nature of the reverberated sound. The reverberation is perceived as flutter echo when the individual reflections do not overlap in time. If the duration of sound is shorter than the time intervals between adjacent reflections the resulting sound contains fluttering (Moore 1990). As a requirement for a sufficient echo density Schroeder (1962) suggests a minimum of $1,000$ echoes per second to achieve flutter-free reverberation, but later Jot and Chaigne (1991) have written that $10,000$ echoes per second is needed to avoid fluttering for short transient sounds.

Another critical issue in producing pleasant reverberation is to avoid coloration of sound which in the frequency domain means that certain frequencies in the magnitude response of the filter are emphasized. This causes the reverberation sound "metallic", as though the sound, for example music or speech, was performed in a room with highly reflective parallel walls, thus gaining too much the normal modes of the room. In the time domain this corresponds to fluttering in the response and

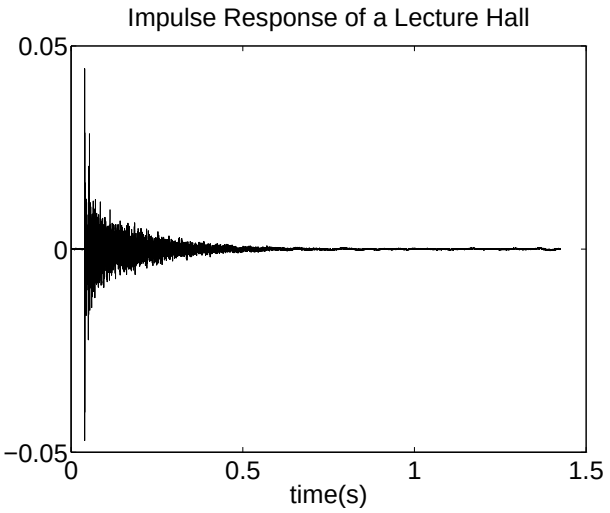


Figure 4.1: An impulse response of a lecture hall.

lack of diffuseness of the sound field. This leads to a requirement of high frequency density, which means that the resonance peaks in the frequency domain are so close to each other that the human ear cannot distinguish between them (Jot and Chaigne 1991),(Schroeder 1962),(Moorer 1979).

A third property of natural room reverberation is a frequency dependent reverberation time. This is normally rendered by adding low-pass filters to the feedback loops of the reverberators which yields reverberation times shorter for high than for low frequencies.

The requirements mentioned above are general properties of one-channel, i.e., mono reverberators. Nevertheless, to give a more spacious impression of sound, the digital reverberators should produce *incoherence* to sound when applied to multi-channel systems. This is based on the fact that the human ear is very sensitive in perceiving correlation of sounds which arrive at both ears, and strong correlation is generally considered unpleasant. For example, halls which produce lateral reflections with a large energy compared to the energy of the front-back reflections are preferred since the lateral reflections produce different signals at both ears (Ando 1985). All the above requirements depend largely on people's subjective opinions and that is the reason why different values for the requirements have been reported in different references.

4.3 Reverberators Based on Recursive Digital Filters

4.3.1 Schroeder Reverberator

The block diagram of the reverberator suggested by Schroeder (1962) is illustrated in Fig. 4.2. It contains four comb filters in parallel and two comb-allpass filters

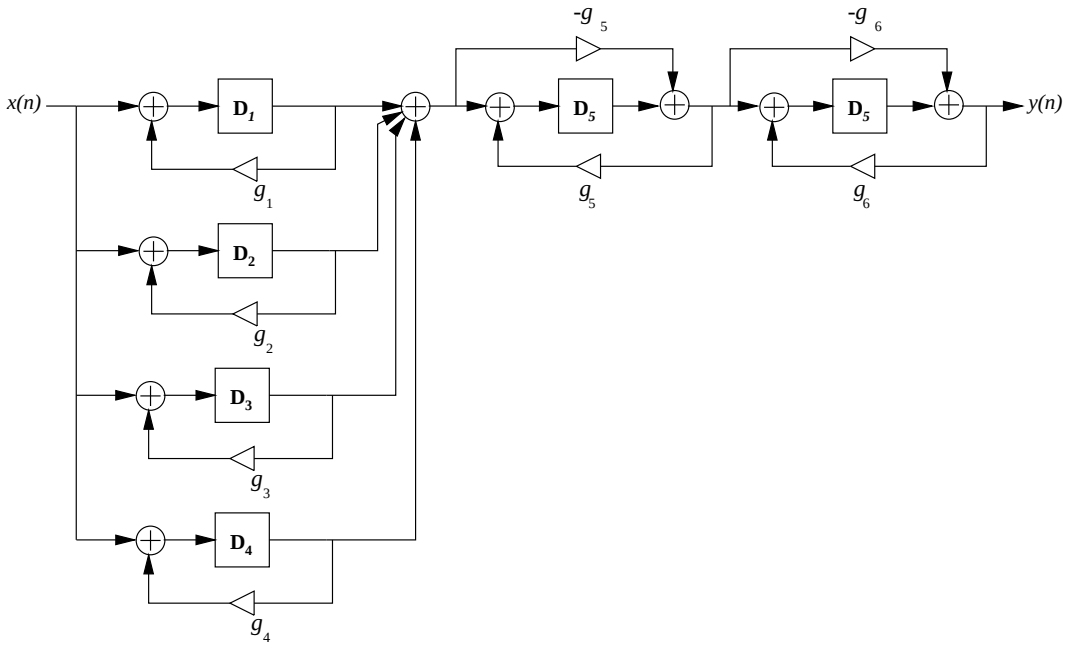


Figure 4.2: Schroeder reverberator.

in series with the comb filters. The comb and comb-allpass filters are of the same structure as was explained in Chapter 3.

Schroeder also derived a relationship between the parameters of a comb filter, i.e., the gains and the delay lengths, and the reverberation time. Since for every trip around the feedback loop the sound level decreases $-20\log|g|$ decibels, and the reverberation time is defined as the time the sound decays 60 dB, this relationship can be expressed as

$$T_{60} = \frac{60}{-20\log|g|} \cdot D = \frac{3}{\log|\frac{1}{g}|} \cdot D, \tag{4.4}$$

where g is the gain of the feedback loop, T_{60} is the reverberation time in seconds and D is the length of the delay in seconds. Schroeder (1962) presented that the delay line lengths should be kept within a close range, and that the delay lengths should be mutually incommensurate to obtain large modal density at low frequencies, and to avoid strong clustering of the modes and superposition of reflections.

Schroeder (1962) also studied reverberators consisting only of series of comb-allpass filters. He suggested a ratio of 1:3 between the delay lengths of consecutive comb-allpass filters, which would lead to a reflection density of $3(n-1)/d_1$ after the n^{th} allpass unit, d_1 being the delay of the first allpass filter. One of Schroeder's suggestions for a reverberator was a series of five allpass filters with the delay lengths defined like was explained above, and the direct sound fed straight to the output multiplied by some gain. This structure would lead to an reflection density of 810/s for $d_1 = 100$ ms. This kind of a construction has a defect that the reverberated sound arrives late, a few hundred milliseconds after the direct sound and that the echo density is not high in the beginning of the decay (Moorer 1979).

A further improvement to a reverberator in Fig. 4.2 is to take frequency-dependent reverberation time into account when modeling late reverberation. This is based on the fact that the air absorbs high frequencies more than low frequencies thus making the reverberation time longer for low than for high frequencies, and that the reflecting surfaces also have frequency dependent absorption coefficients. This can be simulated by adding lowpass filters to the feedback loops of the comb filters which leads to relatively longer reverberation times for low frequencies. Schroeder already suggested placing a first order IIR lowpass filter in the end of each delay line to make the reverberation sound more realistic. The same lowpass filter structure has also been used in the later reverberators.

4.3.2 Other Comb and Allpass Filter Reverberators

Moorer (1979) has suggested further refinements on Schroeder's four comb - two allpass -filter reverberator. He has also done research on different kinds of allpass structures and evaluated their properties, like the reflection density, smoothness of the decay and the response to impulse-like sounds. Moorer (1979) refers to (Schroeder 1962) and claims that Schroeder's reverberator does not give smooth enough a response for impulse-like sounds such as drum strokes but produces reverberation where flutter echo is perceivable, instead of smooth diffuse reverberation, and also that it has a metallic sound in the end of the decay. As an improvement to that structure Moorer suggests a parallel of six comb filters followed by one allpass filter. Each comb filter would contain a lowpass filter in the feedback loop (Fig. 4.3).

Lowpass filters used in this structure are simple first order IIR filters. The block diagram of this kind of a filter is illustrated in Fig. 4.4. The coefficients of the lowpass filter, according to Moorer (1979) are defined by an equation

$$g_2 = g(1 - g_1), \tag{4.5}$$

where g_2 and g_1 refer to the coefficients of a comb filter and the lowpass filter in the end of the delay line, and g is between 0 and 1 for stability, defined by the Equation 4.35. Moorer (1979) said that this filter gives the best results when the lengths of the comb delays are linearly distributed over a ratio of 1:1.5 and all the g_2 are $(1 - g_1)$ multiplied with a same constant g. This would give a possibility to control the reverberation time only by varying the value of g. Moorer suggests replacing the comb- and allpass filters with corresponding oscillatory filters. This means making the coefficients of the filters oscillate sinusoidally. He says this is an improvement to the comb-allpass reverberator but that the response of the filter for impulse-like sounds is still not good.

According to Jot and Chaigne (1991, 1992b) the frequency density D_f and the time density D_t of a parallel comb filter can be expressed as

$$D_f = \sum_{p=0}^{P-1} \tau_p, \tag{4.6}$$

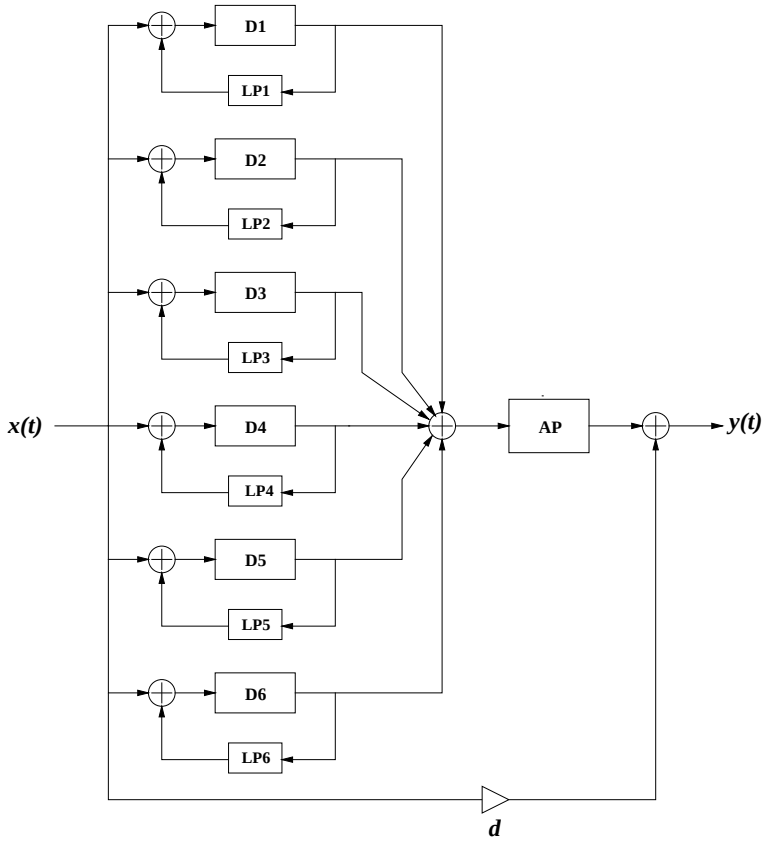


Figure 4.3: A block diagram of Moorer's reverberator. Blocks denoted with LP1...LP6 are the low pass filters in the comb feedback loops and AP is the allpass filter used for diffusing the reflections from the comb filter outputs.

and

$$D_t = \sum_{p=0}^{P-1} \frac{1}{\tau_p}, \tag{4.7}$$

where τ_p is the length of the delay p and P is the number of delay lines. The required amount of delay lines P and the average delay length τ can be derived from the two previous equations:

$$P = \sqrt{D_f \cdot D_t} \tag{4.8}$$

and

$$\tau = \sqrt{D_f / D_t}. \tag{4.9}$$

Gardner (1992b, 1992b) has done research about different combinations of allpass filters to simulate late reverberation. He suggests nested allpass filters in series to produce a fast buildup of echo density. A nested allpass filter means that the delay line of the allpass filter is replaced with another allpass filter and an additional delay to increase the reflection density (Fig. 4.5). Three different reverberators for rooms of different sizes are presented in (Gardner 1992b) and (Gardner 1992a), each

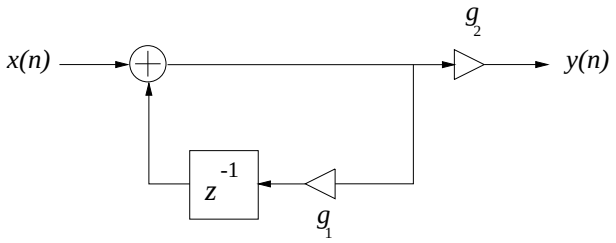
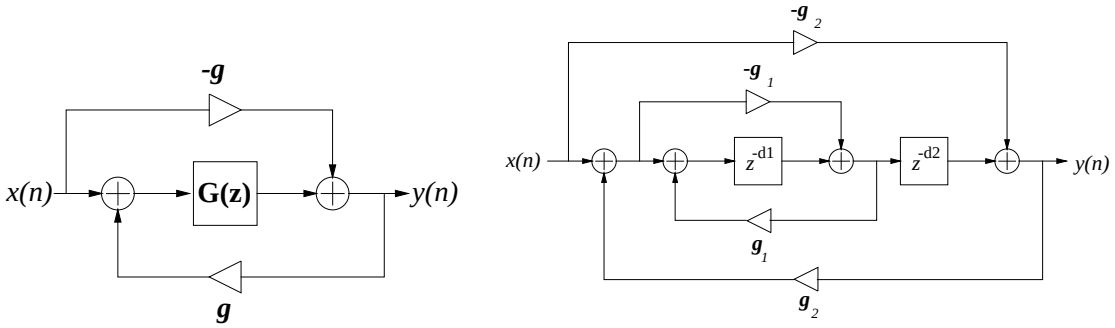


Figure 4.4: A block diagram of a first order IIR low-pass filter used in Moorer's reverberator.



a) b)

Figure 4.5: a) A common block diagram of nested allpass filter. b) A nested allpass filter with one inner filter.

containing single and nested allpass filters in series and varying amount of outputs taken from between the allpass filters and summed together with the output of the last allpass filter. Also a feedback connection from the output of the to the beginning of the delay line containing a constant gain and a lowpass filter is included in all the three structures.

4.3.3 Multiple Feedback Delay Networks

Multiple feedback delay networks (MFDN) are used as a generic method to simulate late reverberation properties of a room impulse response. MFDNs are recursive digital networks where several delay lines are feedback-connected through a square *feedback matrix*, which sums the outputs of the delay lines in proportions defined by the row elements of the matrix, and feeds it back to a delay line defined by the row number of the matrix. The advantage of this kind of a filter is the increased time density of reflections (Jot and Chaigne 1991), (Jot 1992b),(Smith and Rocchesso 1994),(Stautner and Puckette 1982). Jot (1992b) has done extensive research about MFDNs in his doctoral thesis, and Rocchesso (1993, 1996, 1995a) has also studied them not only as a room reverberation model but also as a general physical model of a resonator for sound processing, e.g., for modeling musical instruments. Rocchesso and Smith (1997) have also done a comprehensive summary of the theory and applications of MFDNs.

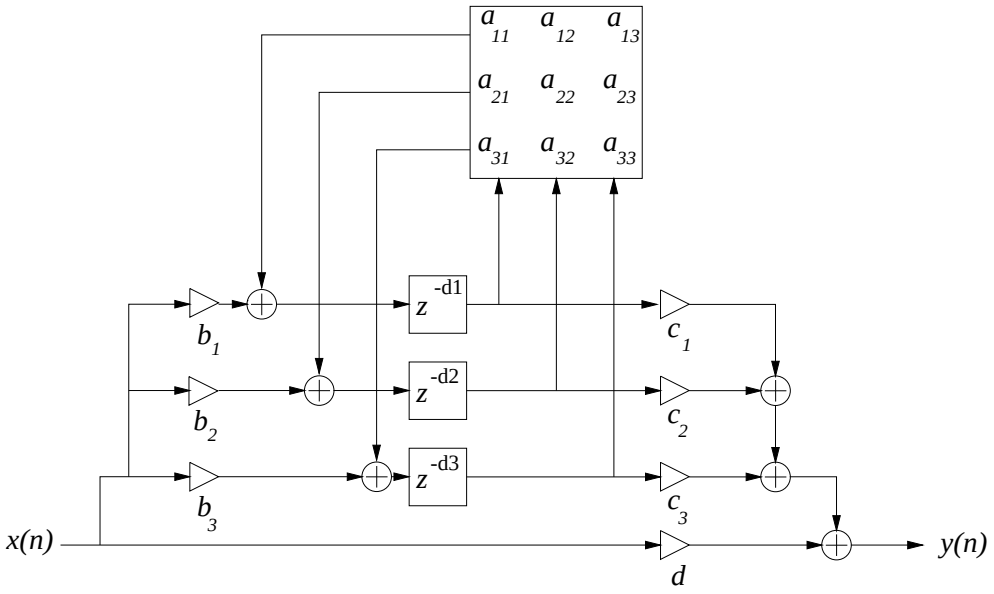


Figure 4.6: A general model of a 3x3 multiple feedback delay network, as defined in (Jot and Chaigne 1991).

A physical correspondence to this structure is given by Rocchesso (1995a). He explains the response of an MFDN by diffusion which occurs at walls of a room. Diffusion causes sounds of wavelengths comparable to the roughness of the surface to scatter in different directions. Since the normal modes of a room correspond to standing plane waves between the walls (Chapter 2.2), and the individual comb filters represent these normal modes 3.2.1, the scattering can be thought of as transferring energy from one normal mode to another. This is actually what happens in an MFDN system, where parallel delay lines divide the sound energy among all the delay lines.

A general model of an MFDN, as defined in (Jot and Chaigne 1991), is illustrated in Fig. 4.6. The relation between the inputs and the outputs of a network containing an NxN matrix is expressed as

$$y(n) = \sum_{i=1}^{N} c_i \cdot q_i(t) + d \cdot x(t), \tag{4.10}$$

where q_i is the output of the i^{th} delay line and is expressed as

$$q_j(t + d_j) = \sum_{i=1}^{N} a_{ij}(t) + b_j \cdot x(t), \tag{4.11}$$

where a_{ij} is an element of the feedback matrix. The transfer function for this system can be written using matrix notation for the feedback coefficients and delay lengths, and expressing q, b and c as N-dimensional vectors. The transfer function for this system is

$$H(z) = \frac{Y(z)}{X(z)} = \mathbf{c}^T[\mathbf{D}(z^{-1}) - \mathbf{A}]^{-1}\mathbf{b} + d, \tag{4.12}$$

where $\mathbf{D}(\mathbf{z^{-1}})$ is a diagonal matrix with elements z^{-d_i}, and $\mathbf{c}$ and $\mathbf{b}$ are column vectors consisting of coefficients c_i and b_i, and d is the constant gain of the direct sound to the output. The poles of this system are the solutions of the characteristic equation

$$det[\mathbf{A} - \mathbf{D}(z^{-1})] = 0, \tag{4.13}$$

and the zeros can be solved from the equation

$$det[\mathbf{A} - \mathbf{b}\frac{1}{d}\mathbf{c}^T - \mathbf{D}(z^{-1})] = 0. \tag{4.14}$$

The transfer function cannot be solved analytically in a general case but when it is chosen to be either unitary, i.e., $\mathbf{A^T} = \mathbf{A^{-1}}$, or a triangular matrix, the poles and zeros can be found (Jot and Chaigne 1991), (Jot 1992b), (Stautner and Puckette 1982). Especially interesting is the case of unitary matrices because they cause the system poles to lie on the unit circle. Stability of the system can be ensured by multiplying the unitary feedback matrix by a diagonal matrix, the elements of which are of magnitude less than one. Practically this means multiplying the output of each delay line with the gains in the diagonal matrix. Jot (1992b) suggests the gains to be

$$k_i = \alpha^{d_i}, \tag{4.15}$$

where d_i is the length of the delay line i in delay units and α is a number smaller than 1. This ensures the poles to have equal magnitude and thus minimizes the coloration of the reverberation, since responses from all the delay lines decay at the same rate.

Rocchesso and Smith (1997) have presented how the values of b and c should be chosen to yield a maximally flat magnitude response of the MFDN in the case of circulant matrices. They have proved that by setting vectors $\mathbf{b^T} = [1, 1..., 1]$ and $\mathbf{c^T}$ in Equation 4.12 so that n elements are equal to -1, n elements equal to 1 and the remaining elements equal to zero the response becomes flat in the case that the lengths of the delay lines d_i are equal and the direct sound gain $d = 1$. This choice of b and c causes the poles an zeros to cancel each other, because they are reciprocals. The response becomes an impulse and therefore the delay lengths have to be different to produce practical responses. The positions of the frequency peaks, on the other hand, depend on the eigenvalues of the feedback matrix and on the delay line lengths. The eigenvalues of a circular matrix can be calculated by taking an FFT of the first row of the matrix. With the above choices of $\mathbf{b}$, $\mathbf{c}$, d and d_i the poles and the zeros of the Equation 4.12 are the m^{th} roots of the eigenvalues of $\mathbf{A} - z^m\mathbf{I}$ and $\mathbf{A} - \mathbf{bc}^T - z^m\mathbf{I}$, respectively.

Two special cases of feedback matrices are triangular and diagonal matrices. The characteristic equation of the feedback matrix in both cases is (Jot 1992b)

$$\prod_{i=1} N(a_{ii} - z^{-d_i}) = 0, \tag{4.16}$$

thus reducing the response of the system to that of parallel IIR comb filters, where the output of each delay line is fed only to the input of the same delay line gained

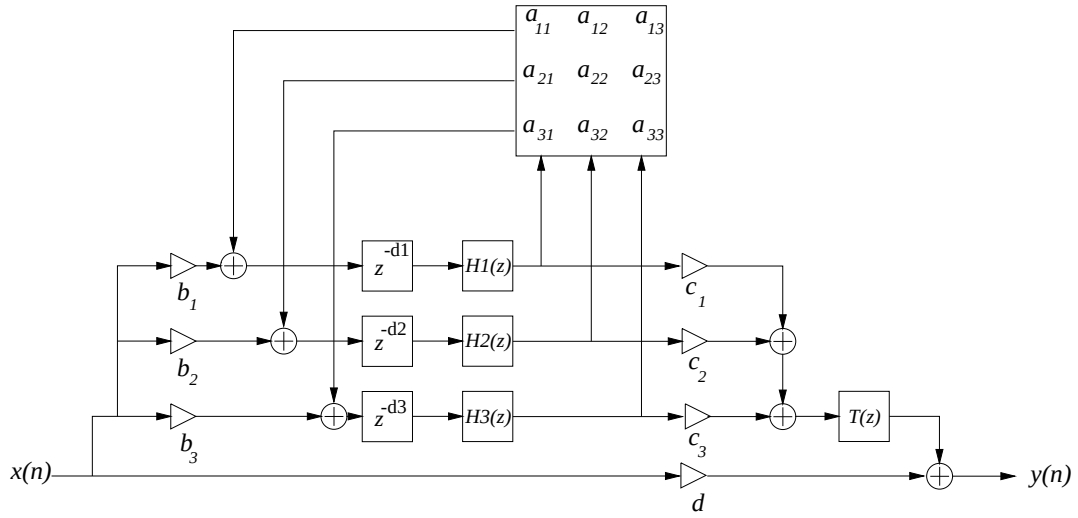


Figure 4.7: A $3x3$ MFDN with lowpass filters $H_i(z)$ and a tone corrector $T(z)$.

with coefficients a_{ii}. When using sparse matrices, the advantage of the maximum reflection density is lost and therefore unitary matrices are most preferable for reverberation modeling.

The feedback matrix structure of Fig. 4.6 can be further improved to better simulate the late reverberation of rooms by adding lowpass filters to the ends of the delay lines and a *tone corrector* after the summation of all the outputs of the delay lines. The general model with the lowpass filters and the tone corrector is illustrated in Fig. 4.8 (Jot 1992b). The purpose of the lowpass filter is to keep the reverberation time shorter for low than for high frequencies. The tone corrector is added to correct the frequency response envelope which is modified by the absorbent lowpass filters.

If the loop gains of comb filters (or an MFDN) are replaced with absorbent filters which contain an attenuating coefficient k_i and a lowpass filter the transfer functions of the absorbent filters can be derived from the Eq. 4.4 (Jot 1992a). For each delay line in the $j\omega$ domain

$$T_{60}(\omega) = -\frac{60}{20\log_{10}|H_i(e^{j\omega})|} \cdot \frac{d_i + \tau_{pi}}{f_s}, \tag{4.17}$$

where $\omega = 2\pi f_s \cdot f$ is the normalized angular frequency ($0 \leq \omega \leq \pi$), $T_{60}(\omega)$ is the frequency-dependent reverberation time, $H_i(e^{j\omega})$ is the transfer function of the absorbent filter, d is the delay in unit delays, f_s is the sampling frequency and $\tau_{pi} = -\frac{arg[H_i(e^{j\omega})]}{\omega}$ is the phase delay of the lowpass filter. From this equation the transfer function of $H_i(e^{j\omega})$ can be derived:

$$|H_i(e^{j\omega})| = 10^{-\frac{3 \cdot d_i}{T_{60}(\omega) \cdot f_s}}. \tag{4.18}$$

To yield the parameters of the lowpass filter, the wished T_{60} curve has to be defined. This curve can be approximated from an impulse response of a real room, for example.

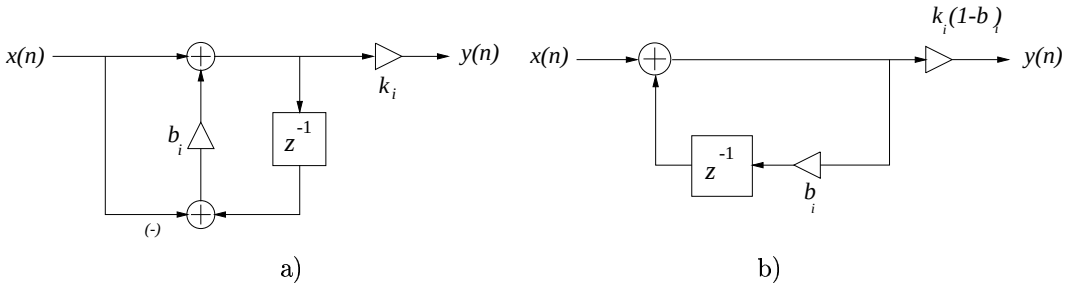


Figure 4.8: a) An IIR low-pass filter used by Jot (1992a). b) equivalent structure as Moorer (1979) presented the lowpass filter.

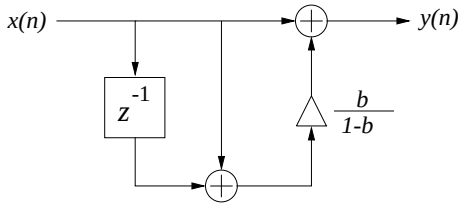


Figure 4.9: An FIR implementation of the tone corrector of Fig. 4.7.

The transfer function of the tone corrector can be derived from the equation

$$|T(e^{j\omega})|^2 = \frac{1}{T_{60}(\omega)}. \tag{4.19}$$

Jot has suggested first order filters for both the absorbent filters and the tone corrector. The former would have form of an IIR lowpass filter of Fig. 4.8 and its transfer function in z-domain is

$$H_i(z) = k_i \frac{1 - b_i}{1 - b_i z^{-1}} \tag{4.20}$$

and the latter would be a first order FIR filter as in Fig. 4.9 with transfer function

$$T(z) = g \frac{1 - bz^{-1}}{1 - b}. \tag{4.21}$$

Jot (1992a) has presented a method for modeling a real room impulse response by using a 12^{th} order IIR filter for the tone correction, to modify the initial frequency response envelope of the room impulse response.

The coefficients k_i, b_i, g and b of these filters are derived as follows: The transfer function of the absorbent filter $H_i(z) = k_i \frac{1-b_i}{1-b_i z^{-1}}$ is written:

$$H_i(z) = k_i \cdot \delta k_i(z), \tag{4.22}$$

where k_i is the constant gain which defines the zero frequency reverberation time and $\delta k_i(z)$ is the first order lowpass filter defining the frequency dependent gain with a unity gain at the zero frequency. In $j\omega$ domain these components of the absorbent filter are in decibels:

$$K_i = 20 \log_{10} k_i = -\frac{60}{T_{60}(0)} \cdot \frac{d_i}{f_s} \tag{4.23}$$

$$\Delta K_i(\omega) = 10 \log_{10} |\delta k_i(e^{j\omega})|^2, \tag{4.24}$$

The complete transfer function in decibels is the sum of K_i and δK_i:

$$K_i + \Delta K_i(\omega) = -\frac{60 \cdot d_i}{T_{60}(\omega) \cdot f_s} \tag{4.25}$$

Now it can be written (if b_i of Eq. 4.20 is $<< 1$)

$$T_{60}(\omega) = T_{60}(0) \cdot \delta T_{60}(\omega), \tag{4.26}$$

which is analogous to the Equation 4.20, and $\delta T_{60}(\omega)$ describes the variation of the frequency envelope. Combining the equations 4.25 and 4.26 we get

$$\frac{1}{\delta T_{60}(\omega)} = 1 + \frac{\Delta K_i(\omega)}{K_i} = 1 - \frac{10}{K_i} \log_{10}(\frac{1}{|\delta k_i(e^{j\omega})|^2}), \tag{4.27}$$

and when

$$\frac{1}{|\delta k_i(e^{j\omega})|^2} = 1 + \frac{2b_i}{(1-b_i)^2}(1 - cos(\omega)) \tag{4.28}$$

is inserted to Equation 4.27, we get

$$\frac{1}{\delta T_{60}(\omega)} \approx 1 - \frac{20}{ln(10)} \frac{b_i}{K_i}(1 - cos(\omega)) \tag{4.29}$$

which yields

$$b_i = K_i \frac{ln(10)}{20(1 - cos(\omega))}(1 - \frac{T_{60}(0)}{T_{60}(\omega)}). \tag{4.30}$$

Jot (1992b) has suggested a parameter to modify the frequency dependent reverberation time:

$$\varepsilon = T_{60}(\pi)/T_{60}(0), \tag{4.31}$$

i.e., the ratio between the reverberation times at the Nyquist frequency and at the zero frequency, when Eq 4.30 would yield

$$b_i = K_i \frac{ln(10)}{40}(1 - \frac{1}{\varepsilon}). \tag{4.32}$$

Nevertheless an exact solution for b_i when $\omega = \pi$ can be derived from equation 4.27:

$$\frac{1}{\delta T_{60}(\pi)} = 1 - \frac{20}{K_i} \log_{10}(\frac{1}{(1-b_i)(1+b_i)}) \tag{4.33}$$

which yields:

$$b_i = 1 - \frac{2}{1 + k_i^{(1-1/\varepsilon)}}. \tag{4.34}$$

and from Eq. 4.23

$$k_i = 10^{-\frac{3d_i}{f_s T_{60}(0)}}. \tag{4.35}$$

$$g = \sqrt{\frac{\frac{1}{f_s}\sum d_i}{Tr(0)}}, \quad b = \frac{1 - \sqrt{\varepsilon}}{1 + \sqrt{\varepsilon}}, \tag{4.36}$$

where $\varepsilon = Tr(\pi)/Tr(0)$, i.e., the ratio between the reverberation time at the Nyquist frequency and the zero frequency (Jot and Chaigne 1991). The reverberation characteristics of the filter can be defined by modifying the coefficients of the lowpass and tone corrector filters independently of the coefficients in the feedback matrix.

Jot (1992a) has used MFDNs for simulating an impulse response of an existing room. This is done via an analysis/synthesis method where the impulse response of a real room is first measured. Then a time-frequency analysis is done to the impulse response by a *short time Fourier transform (STFT)* and the resulting frequency bins are reverse-time integrated to produce an *energy decay relief (EDR)*, where the reverberation times of individual modes can be detected. By adjusting the filter parameters of an MFDN the response of the filter can be modified to simulate the time-frequency characteristics of a room impulse response.

4.3.4 Waveguide Networks

In (Smith and Rocchesso 1994) a connection between MFDN and *digital waveguide networks (DWN)* is presented. The waves which arrive at each junction of the DWN correspond to a delay line output of an MFDN. In this model the *scattering junctions*, as Smith and Rocchesso (1994) refers to the junctions of the FDN, connect N branches of length $\frac{m_i}{2}$ where m_i refer to the corresponding length of of a delay line in an MFDN, and N to the dimension of the feedback matrix of the MFDN. This is based on the fact that the waves along the waveguides travel twice the distance of the waveguide to complete one round trip. Each waveguide is terminated by a non-inverting termination node. An outgoing traveling wave along each waveguide is expressed as

$$s_i(n + m_i) = \sum_{j=1}^{N} a_{ij} s_j(n), \tag{4.37}$$

where $a_{i,j}$ are elements of *scattering matrix* which correspond to the feedback matrix of the MFDN.

DWN provides a more physical reality-based model for digital reverberation via sound scattering. They can be made lossless by choosing the scattering matrix so that its eigenvalues lie on the unit circle. In the same way as in the case of MFDN the reverberation produced by DWN can be made frequency dependent by adding lowpass filters in the structures.

4.3.5 Production of Incoherence to Digital Reverberators

Up to this point only mono reverberation has been discussed in this chapter. In reverberators uncorrelated signals should be produced to different channels to create a broader impression about the sound field.

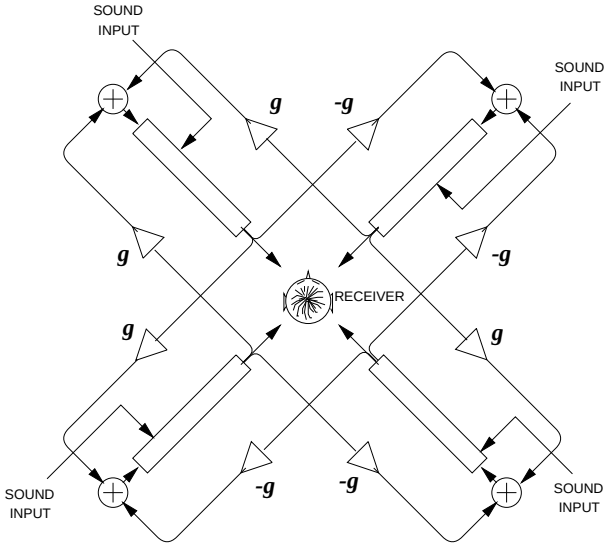


Figure 4.10: The structure of Stautner's multichannel reverberator (Stautner and Puckette 1982).

Stautner and Puckette (1982) have solved this problem in the case of the MFDN by feeding different combinations of delay line outputs to different channels. In their four-channel system each channel is assigned to one delay line output of an $4X4$ MFDN. A source signal which is originally fed to one channel is first fed to the inputs of the delay lines of the adjacent channels. Then from the outputs of those delay lines it is fed further to the diagonally placed channel and back to the original channel etc (Fig. 4.10). This construction corresponds to a feedback matrix

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \cdot g, \tag{4.38}$$

where $|g| \leq \frac{1}{\sqrt{2}}$. The resulting signals at the outputs of the four channels are mutually incoherent. An additional advantage of feeding the signal only to the adjacent channels is that the timing of the source signal and the following reflections occurs in a natural order, i.e., a signal originally excited in one channel arrives last in the opposite channel.

In the nested allpass filter structures by Gardner (1992b, 1992b) incoherence is obtained by varying slightly the delay lengths of the allpass filters between different channels. In this case the processing has to be done separately to each channel. In the case of mono sound material this seems like wasting the processing time and delay memory, and incoherence producing should be implemented in a more efficient way such as in the case of the MFDN. Nevertheless when the reproduced sound is already multichannel material each channel may have to be processed separately and then this kind of a solution is practical.

4.3.6 Commercial Products

Two commercial audio effect processors were measured, namely, m5000 Audio Mainframe (t.c n.d.) and Korg reverberator (**?**). From the analysis of the impulse responses some implementation principles can be concluded. I will present some of the results from the reverberation algorithms of m5000 to give an impression about the implementation methods. With m5000 it is possible to produce both time invariant and time variant reverberation. According to its manual the time variance is implemented by variating lengths of some delay lines of the late reverberation unit. This affects the position of the poles of the filter transfer function and thus it breaks the steady harmonic structure of the response and prevents strong coloration. The time invariant responses were measured using a pseudorandom sequence with flat spectrum and random phase, and the time-variant responses were measured using an impulse signal as an excitation.

In the manual of the m5000 (t.c n.d.) different parameters for modifying the reverberation are explained. They are described in the following:

1. *Mix* defines defines the amount of reverberation relative to the direct sound. A reverberation free sound is referred to as "dry" whereas "wet" is used to describe sound with a great proportion of reverberated sound.
2. *Decay* defines the decay time of the reverberation.
3. *x low* is a relative decay time multiplier for low frequencies.
4. *x high* is a relative decay time multiplier for the high frequencies.
5. *lo-xovr* and *hi-xovr* define the crossover frequencies of the parameters *x low* and x high, respectively.
6. *Diffuse* defines how each reflection is spread over time. It affects the time density of the reverberation. From the measured responses it can be seen that the "diffuseness" is implemented with a comb-allpass filter of a short delay (less than a millisecond).
7. *Shape* defines the pattern of the early reflections used in the beginning of the response. The four possibilities are *hall, fan, prism* and *horse shoe.*
8. *x size* scales the dimensions of the simulated hall, i.e., the delays of the early reflections (defined by the shape-parameter) and thus the effective size of the space to be simulated.
9. Predelay sets the initial time gap between the direct sound and the early reflections.
10. *Revfeed* sets the time before the late reverberation starts relative to the early reflections.
11. *Hicut* defines the cutoff frequency (in 1/3 octave steps) of a lowpass filter (frequency rolloff 6dB/octave) which results in a "warmer" reverberation.

12. *Attenuation* sets the high frequency rolloff for *hicut*.

13. *Initlevel* sets the level of the early reflections relative to the late reverberation.

14. *Revlevel* sets the level of the late reverberation relative to the early reflections.

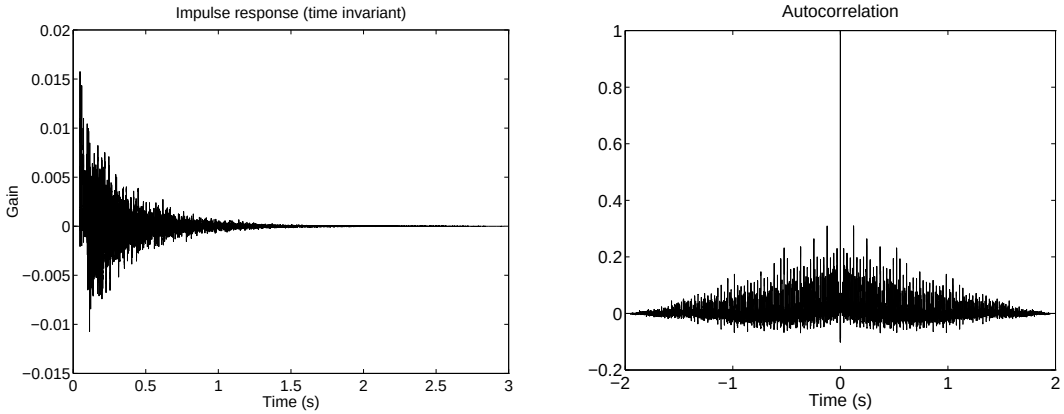
15. *Rwidth* defines the amount of stereo in the reverberation, zero corresponding to mono reverberation and 1 to totally independent processing of two channels.

16. *Modulation rate* and *Modulation depth* are available for time varying reverberation algorithms and they mean the rate and amplitude at which the length of a delay line is modulated.

In Figures 4.11, 4.12 and 4.13 impulse responses of different reverberation algorithms of m5000. In Fig. 4.11 a) is an artificial reverberation response of a church and in Fig. 4.11 b) its autocorrelation function which implies the periodicity of the signal. In Fig. 4.12 a) is the response of a concrete space reverberation and its autocorrelation function Fig. 4.12 b). The early reflections are distinctive and the late reverberation part of the response has a very clear parallel comb filter -like structure. An impulse response, and the EDRs of the responses measured from different channels of a time variant reverberator are plotted in Fig. 4.13. The variation of resonances as a function of time can be seen in Fig. 4.13 b) and c) where the EDRs of a time variant response are plotted up to 600 Hz.

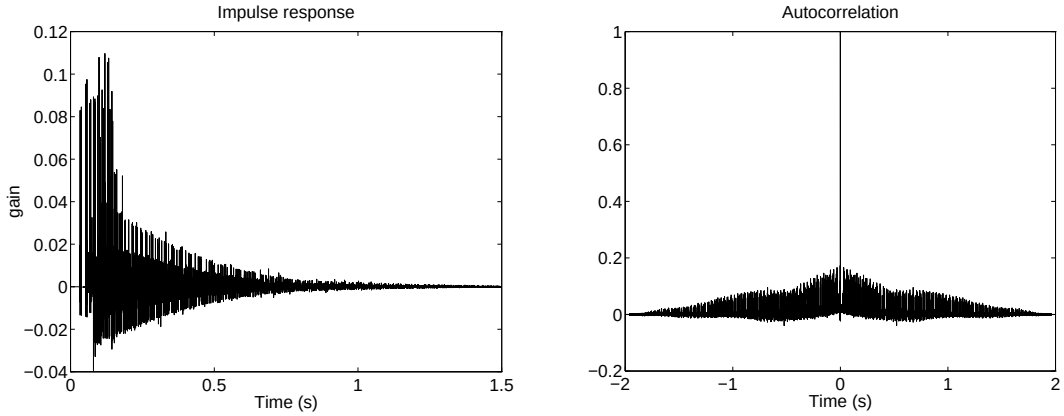
4.4 Summary

In this chapter the methods for reverberation modeling found in the literature have been presented. The main emphasis was put on recursive digital filters which are used for late reverberation modeling. The filter structures were presented as well as the means to control the characteristics of the reverberation, such as, the frequency dependent reverberation time. The concept of creating artificial spatialization, i.e., uncorrelated signals to different channels was discussed. Finally, results from an analysis of a commercial reverberator were presented, and the adjustable parameters were explained according to the manual of the reverberator, to give an impression about how the reverberation characteristics can be defined.



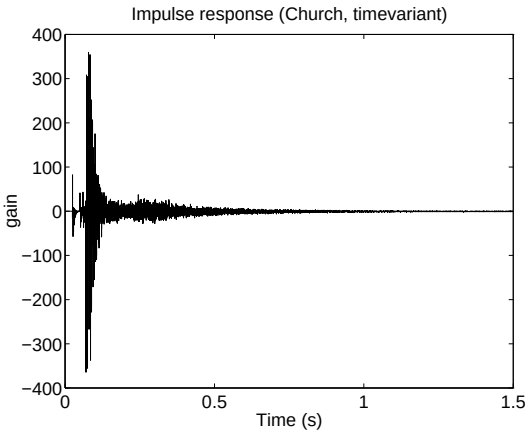
a) b)

Figure 4.11: Impulse response of church reverberation effect of m5000 a) and its autocorrelation function.

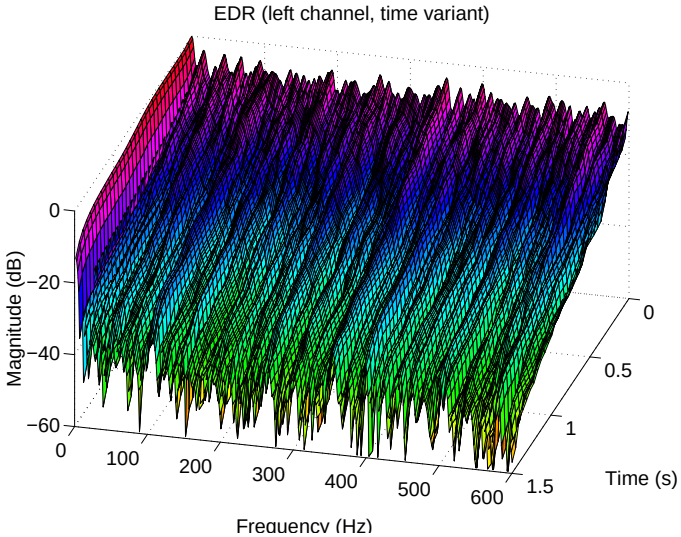


a) b)

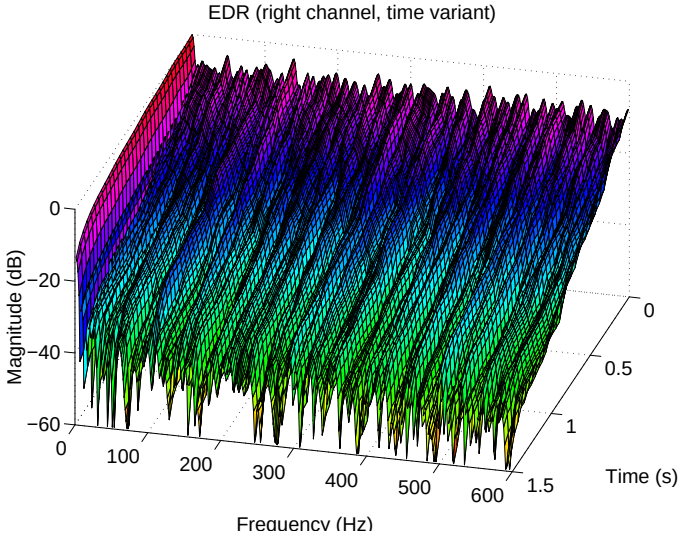
Figure 4.12: Impulse response of reverberation effect of m5000 simulating a concrete room a) its autocorrelation function.



a)



b)



c)

Figure 4.13: An impulse response (a) and EDRs from two different channels up to 600 Hz (b) and (c) of a time-variant reverberation effect.

5. Simulation and Evaluation of Digital Reverberators

In this chapter the simulated reverberator structures and their evaluation are described. First, the simulation methods of the algorithms will be explained as well as the evaluation methods applied to the responses of the reverberators. Then, the modeling of early reflections is discussed, how their delays and gains are obtained from a computational model of a hall, and how they can be combined to a late reverberation unit. In Section 5.4, a simple structure of parallel comb filters and its time- and frequency properties is studied since nearly all reverberators are based on some modification of that structure. In Section 5.5 the structures of the simulated reverberators are shown, and objective evaluation based on the time and frequency domain properties of their impulse responses is dealt with. Finally, approximations about the computational requirements of the reverberators is presented.

5.1 Simulation of Reverberators

As a tool for simulation and evaluation of the digital reveberators Matlab was used. The time-domain convolution of the digital reverberation algorithms was implemented by C programming language. The C-functions were compiled to mex functions to be called from Matlab. The reason for using C for the implementation of the algorithms is that the compiled mex functions are faster than Matlab functions in executing the computationally heavy algorithms, which are required for reverberation modeling.

The parameters which modify the reverberator response are, for example, the gains of the filter and the lengths of the delay lines. These parameters can be passed to a mex function together with the signal to be filtered by that reverberator, and the variable the mex function returns is the filtered signal. In the reverberators I have simulated the input arguments to the function are:

- Reverberation time at zero frequency, from which the gains k_i in the ends of the delay lines are calculated according to Equation 4.35.

- Ratio between the reverberation times at the Nyquist frequency and zero frequency, from which the lowpass filter coefficients b_i are calculated from Equation 4.34.
- Delay lengths which are passed to the function in a single vector to be used as lengths of parallel delay lines, and also the delay lengths of comb-allpass filters used in the structures.
- Gains associated to the filter, e.g., the vectors $\mathbf{b}^T$ and $\mathbf{c}^T$ of the MFDN (Eq. 4.12), and the gains which define the proportions of the direct sound, early reflections, and the reverberated sound, and the gains of the comb-allpass filters.
- Early reflections, i.e., their delays and gains proportional to the direct sound.
- Input signal which can be either and impulse function (the first element 1 and the other elements zeros, having a length which is equal to the wished impulse response length), or a sound sample of speech or music to be filtered to give and impression about the subjective quality of the algorithm.

5.2 Evaluation Methods of Reverberators

The objective evaluation of the responses is based on the following tests done on the impulse response of the reverberator, and which help to criticize and predict the quality of the reverberation. Most of them are based on the reverberation quality tests that Griesinger (1989) has presented:

- Plots of reflection density as a function of time
- Plots of energy as a function of time (integrated impulse response)
- Plots of the autocorrelation function
- Sonograms, i.e., plots of the frequency contents as a function of time
- Plots of the reverberation time as a function of frequency
- Plots of the steady state amplitude response as a function of frequency
- Plots of backwards integrated short time Fourier transform
- Listening to the response to different excitation signals

These evaluation methods are shortly explained in the following:

1. *Plots of reflection density as a function of time*
   The same method as in (Griesinger 1989) has been used to determine the reflection density: The number of reflections within 20 dB of the largest reflection inside a sliding window of 20 ms are counted. The hop size used for the sliding window is half of the window length, i.e., 10 ms. This test indicates the buildup of the reflection density and the maximum level of it.

2. *Plots of energy as a function of time*
   The reversed time integrated curve (Schroeder 1965) of an unfiltered impulse response is plotted to illustrate the shape of the decay.

3. *Plots of the autocorrelation function*
   Autocorrelation function of an impulse response brings out the repetitive behavior in the response (Jot 1992b) which produces detectable peaks in the autocorrelation sequence. It also indicates the "mixup time", i.e., the time after which the response resembles random noise. Griesinger (1989) suggests multiplying the impulse response by an exponentially increasing function to keep the level of the response constant before computing the autocorrelation function. Autocorrelation function is defined by the Equation 2.22 or 2.23, and it is normalized by dividing the function by the value of the autocorrelation function at time zero:
   $$\phi(\tau)_{norm} = \frac{\Phi(\tau)}{\Phi(0)}. \tag{5.1}$$

4. *Sonograms*
   Griesinger (1989) has mentioned sonograms for detecting the modes which have longest reverberation times. They are plots of frequency contents as a function of time.

5. *Plots of reverberation time as a function of frequency*
   This test is good for evaluating whether the third octave, or octave band reverberation times correspond, for example, to those of existing halls.

6. *Plots of the steady state amplitude response as a function of frequency*
   The magnitude response of the reverberator helps to determine the modal density, and whether there are large regular valleys between strong peaks in the magnitude response of the reverberation.

7. *Plots of backwards integrated short time Fourier transforms*
   In this test a time-frequency analysis is done on the impulse response by Short Time Fourier Transform (STFT). The resulting frequency bins are reverse-time integrated to create a three dimensional plot where the modes are monotonously decreasing, and the continuity of each mode can be detected. These plots reveal the modes which have the longest reverberation times. If the modes decay at very different rates, a thin and metallic sound can be expected. Jot (1992a) has used this method to visualize the energy decay of the reverberation, and has used a term *Energy Decay Relief (EDR)* to describe these curves.

   In STFT the response is divided into overlapping segments in the time domain. The segments are windowed, and the frequency contents of the signal in each window are analyzed by *Discrete Fourier Transform (DFT)* using a *Fast Fourier Transform (FFT)* algorithm (Oppenheim and Schafer 1975). The window has to be select according to the needed resolution in the time and frequency domain. The length of the window in the time domain defines the lowest frequency component that can be analyzed with the window. On the

other hand lengthening the window reduces the time resolution of the analysis, which is a direct consequence of Heisenberg's uncertainty principle. Thus a compromise between the frequency resolution and the accuracy of time development of the frequency contents has to be done. The type of the window chosen depends on its time and frequency properties. The properties of a Kaiser window in the frequency domain are the width of the main lobe which determines the frequency resolution, and the attenuation of the sidelobes, which determines the cross-talk between different channels. Harris (1978) has written about various window types and their time and frequency domain properties.

In the STFT analysis a *kaiser window* has been used, since the frequency resolution of this window, i.e., the main lobe width can easily be defined by a single parameter. Kaiser window is defined by

$$w(n) = \frac{I_0[\beta\sqrt{(\frac{N-1}{2})^2 - [N - (\frac{N-1}{2})]^2}]}{I_0[\beta(\frac{N-1}{2})]}, \tag{5.2}$$

where

$$I_0(x) = \sum_{k=0}^{\infty}[\frac{(\frac{x}{2})^k}{k!}]^2. \tag{5.3}$$

β is a parameter used for controlling the trade-off between the sidelobe level and the main lobe width. Increasing β makes the main lobe narrower but decreases the attenuation of the side lobes. The parameter $\beta = 1$ is chosen to yield a high frequency resolution. In Fig. 5.1 the time- and the frequency domain presentations of the Kaiser window used in the analysis of the responses in a short frequency range are plotted. The length of the window is 8192 samples, which corresponds to 186ms with 44.1kHz sampling frequency. The analyzed frame of the signal is zero padded to 4 times the window length, i.e., $4 \cdot 8192 = 32768$ which corresponds to a resolution of $44100/32768 = 1.34$Hz in the FFT. To visualize the decay of the whole frequency range I have used a shorter window (4096 samples and the same amount of the FFT points) and thus the smallest details of the frequency structure cannot be detected.

8. *Listening*
   Listening is a good test for judging the reverberation quality for different signals. Clicks and other impulse-like sounds reveal whether the reverberation causes fluttering. Reverberation with too few modes gives metallic sound to speech, music and other sounds people are familiar with.

For evaluation of the simulated reverberator structures tests number 1., 3., 6., 7., and 8. were used. Additionally, a normalized cross-correlation function of the responses from different channels when evaluating two-channel responses was applied. The normalized cross correlation function is defined by the Equation 2.24. It implies whether there exists incoherence between different channels and thus whether the two-channel response gives a good spatial impression. Correlation is detected as peaks in the response in the vicinity of zero time lag, i.e., the origin of the plot.

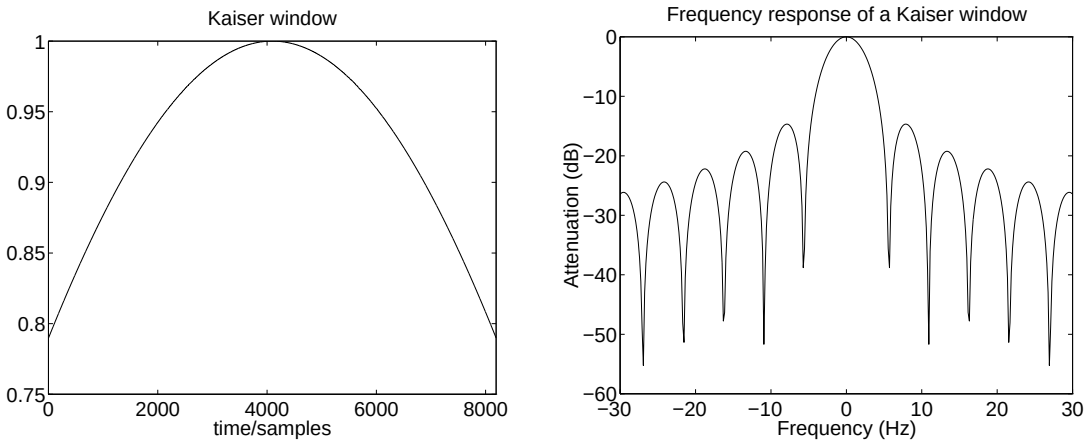


Figure 5.1: a) The time domain function of a kaiser window of length 8192, $\beta = 1$ b) its magnitude response for 44.1 kHz sampling frequency and the length of the FFT is 32768.

By listening to the response coherence is easily detected. If the two channels correlate considerably the sound source is localized inside the listener's head in both loudspeaker and headphone listening. If there is enough incoherence between the channels the sound field seems to spread to the surrounding space.

5.3 Modeling of Early Reflections

The tests mentioned above were mostly done on the responses of the filters used for late reverberation modeling. Since the early reflections have to be modeled to give a pleasant and natural impression about the reverberation, they were added to the models, when listening to their response to sound. The early reflections were also used for calculating the reflection density in the time domain, since they affect the density of the reflections. Some examples are also given about the influence of the early reflections to the decaying modes of the reverberator response in the backwards integrated short-time spectra.

The early reflections used in all the reverberators are obtained from a ray-tracing model of the Sigyn Hall in Turku. The early reflections during the first 55 ms are reduced to 20 reflections and they are implemented as a tapped delay line and each output of the delay line is lowpass filtered to simulate air absorption. The early reflections are rendered separately for two channels to produce incoherence between the channels. The incoherence is implemented by applying a small time delay between the channels, according to the direction of the sound. The time delays are calculated according to the following equations (Blauert 1983):

$$d = a \sin \theta + a\theta, \tag{5.4}$$

and the interaural time difference (ITD) is

$$ITD = \frac{d}{c} = \frac{a(\sin\theta + \theta)}{c} \tag{5.5}$$

where d is the distance difference between the ears, c is the velocity of sound, and θ and a are derived according to Fig. 5.2. Reflections with incident angles greater than 90°, i.e., the reflections coming from behind the listener, are rendered by placing them symmetrically according to the B-axis in Fig. 5.2, to the front. As the value of c, 350 m/s has been used and 17.5 cm as the diameter of the head, i.e., $a = 17.5/2 = 8.75$ cm.

The reduction to 20 reflections is done in the same way as in (Gardner 1992b). First, the reflections with amplitude smaller than 0.01 are deleted and then the reflections within 1 ms are merged. The delay d and the amplitude a of the resulting reflection from two combined reflections are defined by the following equations:

$$d = \frac{d_1 a_1^2 + d_2 a_2^2}{a_1^2 + a_2^2} \tag{5.6}$$

and

$$a = \sqrt{a_1^2 + a_2^2}, \tag{5.7}$$

where d_1, d_2, a_1 and a_2 are the delays and amplitudes of the original reflections from the ray-tracing model. Additionally, the listening space is laterally divided into two sectors so that reflections from different half planes are not merged together. This method could be extended to more channels so that reflections coming from between adjacent two channels are fed only to those channels.

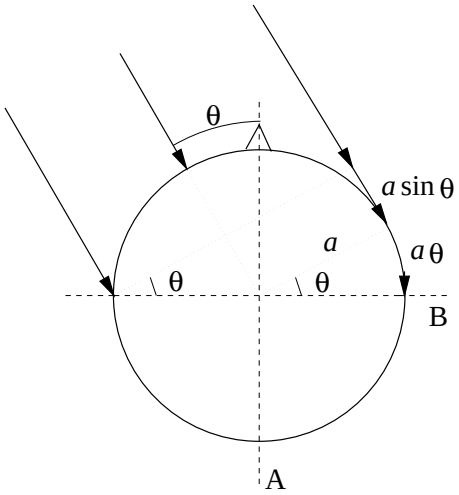


Figure 5.2: Approximation of the distance difference at the two ears of a listener

The delays and amplitudes of the early reflections used for evaluation are listed in Table 5.1. In Fig. 5.3 the schematic of late reverberation rendering is illustrated. The source sound is fed to both output channels, and to the delay line which implements the delays of the early reflections for both channels. In Fig. 5.3 a) sound from

the late reverberator is divided into two channels depending on the implementation of the the late reverberation. To produce incoherent signals from the late reverberation unit, for example, different delay line output combinations can be taken from the reverberator to different channels. Incoherence between the channels can also be produced by feeding the incoherent early reflections to similar late reverberation blocks which are connected to different output channels (Fig. 5.3 b). The spatial impression rendered this way is considerable but on the other hand a separate late reverberator is needed for each channel, and the needed amount of delay memory and processing is multiplied by the number of the channels.

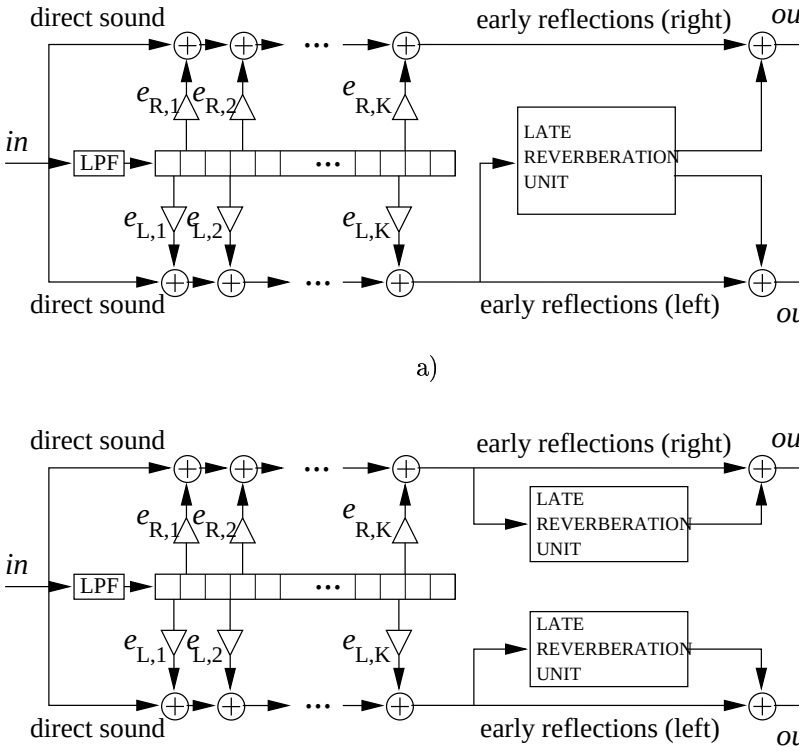


Figure 5.3: A simplified diagram of the reverberation modeling. a) Same processing used for two channels. b) Separate processing of two channels.

	R(ms)	L(ms)	R	L	R	L	R	L	amplitude
			($f_s = 48kHz$)		($f_s = 44.1kHz$)		($f_s = 32kHz$)		
1	0.00	0	0	0	0		0	0	1.193
2	1.73	1.73	83	83	76	76	55	55	0.628
3	25.08	25.38	1204	1218	1106	1119	803	812	0.8142
4	26.89	27.64	1290	1326	1186	1219	860	884	0.5163
5	30.05	29.68	1443	1425	1325	1309	962	950	0.6610
6	31.01	30.65	1488	1471	1368	1352	992	981	0.3544
7	32.75	33.49	1572	1607	1444	1477	1048	1072	0.3236
8	36.02	35.66	1729	1712	1588	1573	1152	1141	0.5083
9	38.04	37.28	1826	1789	1676	1644	1217	1193	0.575
10	37.77	37.94	1813	1821	1666	1673	1209	1214	0.3384
11	39.89	39.16	1915	1880	1759	1727	1277	1253	0.3004
12	42.85	42.13	2057	2022	1890	1858	1371	1348	0.2940
13	46.12	45.88	2214	2202	2034	2023	1476	1468	0.2656
14	47.51	46.75	2280	2244	2095	2062	1520	1496	0.2997
15	47.28	47.51	2269	2281	2085	2095	1513	1520	0.1244
16	49.03	48.75	2353	2340	2162	2150	1569	1560	0.1696
17	48.84	49.15	2344	2359	2154	2168	1563	1573	0.1044
18	49.81	49.53	2391	2377	2197	2184	1594	1585	0.219
19	54.65	55.23	2623	2651	2410	2436	1749	1767	0.5720
20	56.03	55.75	2689	2676	2471	2459	1793	1784	0.4182

Table 5.1: 20 early reflections of the Sigyn Hall. First two columns are delays in ms relative to the direct sound (0ms delay), the following four columns are the delays in samples for 48, 44.1, and 32 kHz sampling frequency, and in the last column are the amplitudes of each reflection.

5.4 Parallel Comb Filters

With parallel comb filters alone it is not possible to obtain reflection density which is high enough to avoid fluttering and coloration in the reverberation. Nevertheless, parallel comb filters are a part of most digital reverberators in some form, and therefore the time and the frequency responses of that structure are studied to give an impression about the effect of the adjustable parameters on the response.

In Figure 5.4 the general form of parallel comb filters is presented. This is a modification of the structure used by Schroeder (1962) and Moorer (1979). The difference is that the gains of the feedback loops are in the ends of the delay lines before the feedback connection, instead of placing them in the feedback loops. This is done to obtain an exponential decay right from the beginning of the response, so that already the first outputs from the delay lines are multiplied by the attenuating gain. I have applied the same modification to all reverberation algorithms containing comb filters or an MFDN structure.

The impulse response of the three parallel comb filters is plotted in Fig. 5.5 a), and in Fig. 5.5 b) the 60 dB decay of the same response is plotted to reveal the

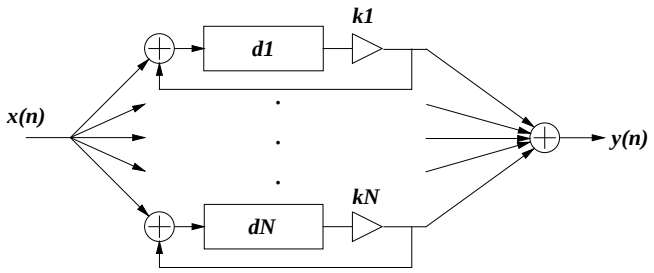


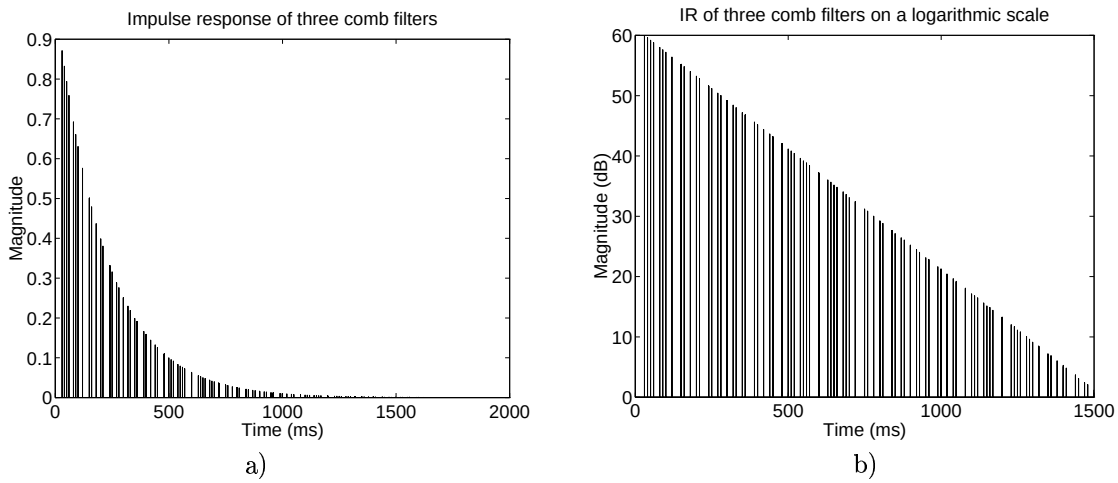
Figure 5.4: A general block diagram of parallel comb filters.

exponential decay of the response. In Fig. 5.5 c), the magnitude responses of each comb filter are plotted and in Fig. 5.5 d) is the magnitude response of the parallel connection. It can be seen that the resonances coincide around 100 Hz and cause a high peak in the response and also that there are regular gaps in the magnitude response. These are features to be avoided in the reverberation modeling since they cause strong coloration to sound. Therefore an important phase in the reverberator design is choosing the delay lengths so that the frequency response does not contain a considerable comb structure or regular peaks. The comb structure can be seen also in Fig. 5.6 a) where the energy decay is plotted as a function of frequency up to 400Hz. This kind of plots clearly reveal the longer reverberation times of the resonance frequencies. Fig. 5.6 b) contains the EDR up to 20kHz for otherwise a same kind of a comb filter connection except that there are lowpass filters associated to each delay line, which yield shorter reverberation times for the high than for the low frequencies. The delay line gains k_i, and the lowpass filter gains b_i were calculated from the Equations 4.34 and 4.35 to give the same reverberation time to each filter and also same frequency-dependent decays. The reverberation times for which each k_i was calculated was 1.5 s and the $\varepsilon = T_{60}(\pi)/T_{60}(0)$ value used for calculating the gains b_i is 0.3. The delay line lengths, the gains k_i and b_i are listed in Table 5.2

i	length in ms	length (f_s=44.1 kHz)	g_i	b_i
1	30 ms	1323	0.87	0.16
2	40 ms	1764	0.83	0.21
3	50 ms	2205	0.79	0.26

Table 5.2: The delay lengths and the loop gains of the three parallel comb filters used to render the impulse response in Fig. 5.5 and the EDR in Fig. 5.6 a). The last column contains the gains of lowpass filters (Eq. 4.34) added to the delay line ends to obtain frequency-dependent reverberation time (the EDR in Fig. 5.6 b))

.



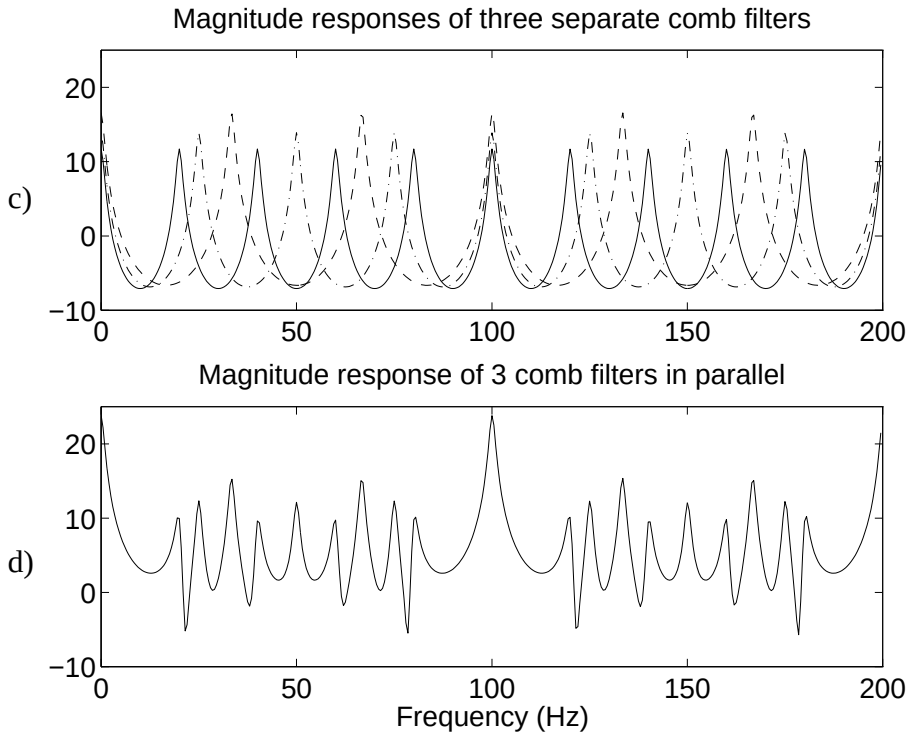
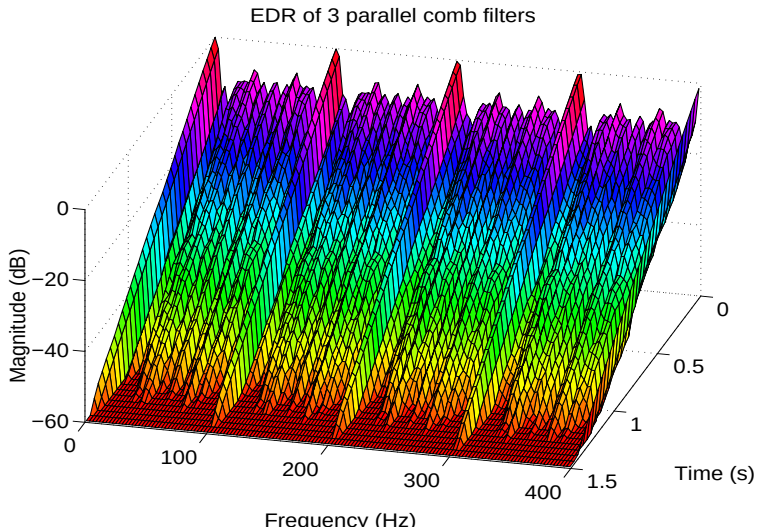
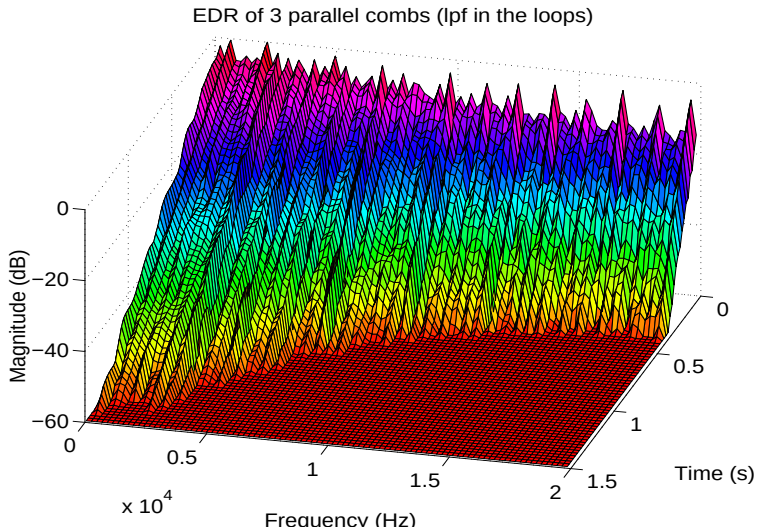


Figure 5.5: An impulse response of three comb filters in parallel a) on a linear scale and b) on a logarithmic scale. c) The magnitude responses of three individual comb filters and d) the response of the same comb filters connected in parallel. The delay lengths and gains are listed in Table 5.2.



a)



b)

Figure 5.6: a) An EDR of three parallel comb filters up to 400 Hz. b) Same as in a) but up to 20 kHz (f_s=44.1 kHz) and lowpass filters added to the ends of the delay lines.

5.5 Evaluation results

Reverberators containing various amounts of delay lines and with different delay lengths were studied. The structures and the evaluation results according to the criteria explained in Section 5.2 are presented in this section. In all the reverberators containing parallel delay lines, same amount of delay lines have been used to make the results comparable. Also the lengths of the delay lines in different reverberators are the same. For all the simulations a sampling frequency of 44.1 kHz was used. The reverberation time and the ratio $T_{60}(\pi)/T_{60}(0)$ were also set the same for all the simulations, and the loop gains k_i and the lowpass filter coefficients b_i are calculated from the Equations 4.35 and 4.34. The values common for all the measurements are listed in Table 5.5.

The delay line lengths are chosen to be between 50ms and 75ms to overlap slightly with the early reflections (listed in Table 5.1). The delay lengths were chosen to be prime numbers to avoid superposition of individual reflections and also their mutual lengths were tried to be set incommensurate to avoid clustering of the resonances. All the measurements were done with two sets of delay lines one consisting of 6 delay lines and the other of 8 delay lines. This is done to bring out the effect of increasing the number of delay lines.

i	delays (ms)	delays (in samples)	k_i	b_i
6 delay lines				
1	49.7	2191	0.795490	0.260765
2	57.9	2552	0.766060	0.301265
3	64.3	2835	0.743753	0.332279
4	69.5	3067	0.725951	0.357182
5	73.0	3221	0.714369	0.373440
6	75.0	3309	0.707983	0.382420
8 delay lines				
1	49.7	2191	0.795490	0.260765
2	57.8	2549	0.766300	0.300932
3	64.2	2833	0.743908	0.332062
4	69.0	3041	0.727924	0.354416
5	71.7	3163	0.718709	0.367343
6	73.0	3221	0.714369	0.373440
7	74.8	3297	0.708722	0.381380
8	75.0	3309	0.707983	0.382420

Table 5.3: The delay lengths, loop gains and lowpass filter coefficients of the simulated reverberators.

5.5.1 Modifications of the Moorer Reverberator

In this section a reverberator structure similar to Moorer's reverberator is studied. The reverberator consists of 6 or 8 parallel comb filters followed by a single comb-allpass filter for increasing the reflection density of the reverberator response. Three different lengths of the comb-allpass filter delay has been applied, to reveal its affect on the reflection density. A general block diagram of this structure is shown in Fig. 5.7.

The theoretical reflection densities (Eq. 4.7) for parallel comb filters with the delay line sets listed in Table 5.5 are $D_t = \sum 1/d_i = 94.4$/s for 6 comb filters and $D_t = 121.8$/s for 8 comb filters. Thus they are far away from the value that, for example, Griesinger (1989) has suggested for the reflection density, to avoid fluttering in the reverberation (10000 reflections/s). By adding the comb-allpass filter in series with the comb filters, the reflection density can be increased considerably.

In Fig. 5.8 the impulse responses of this reverberator are plotted in the cases of 6 and 8 comb filters. The length of the comb-allpass filter delay was 6 ms (265 samples). Fig. 5.10 shows the reflection densities of these structures, resulting from different lengths of the comb-allpass filters. As the comb-allpass filter delay length, 2 ms, 3 ms and 6 ms were tried, and the value of the allpass filter gain g (see Fig. 3.6) was 0.7 in each case. It can be seen that in all of the responses the average reflection density does not increase as a function of time, even if strong fluctuation in its magnitude appears, and that the maximum average level of the reflection density is about 3000 reflections/s. Also the time domain structure does not look very dense and therefore fluttering can be expected. The magnitude spectra between 0-100 Hz of the responses resulting from 6 and 8 parallel comb filters are shown in Fig. 5.9 a) and b), respectively. The comb-allpass filter following the parallel comb filters does not affect the frequency contents of the response, and therefore the magnitude response is the same, regardless of the delay length or the gain of the comb-allpass filter. In Fig. 5.11 a) and b) the EDRs of the structure with 6 and 8 comb filters between 0-200 Hz is shown, which reveals that the different modes within a short frequency range decay with approximately same rates. Fig. 5.11 c) shows the EDR of 8 parallel comb filters in the frequencies up to 20 kHz, and it can be seen that the reverberation time decreases monotonously as a function of frequency.

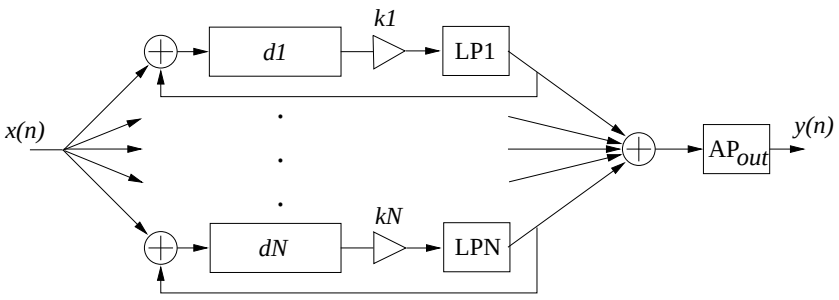


Figure 5.7: Parallel comb filters followed by a single comb-allpass filter.

In Fig. 5.12 a) an impulse response where the early reflections are added and also fed to the late reverberation unit, is illustrated. In Fig. 5.12 b) is the reflection density of the same response. It can be seen when comparing with the time density plot in Fig. 5.8 that the reflection density is significantly increased, which can be expected because now the late reverberation actually contains the late reverberation for each individual early reflection. In Fig. 5.13 the EDR of the same structure is plotted, which shows the effect of the early reflections to the spectrum. Obviously feeding the early reflections to the late reverberation unit they produce strong peaks which can be assumed to cause perceivable coloration in the perceived response. Thus it is important how the early reflections are selected and how many of them are fed to the late reverberation unit. For example, Jot (1992b) suggests that only the first few reflections should be used as an input to the late reverberation unit.

In Fig. 5.14 The autocorrelation plots of the impulse responses reveal the regular structure of the responses. Figures 5.14 a) and c) show the autocorrelation plots of the Moorer's reverberator response containing 6 and 8 parallel comb filters, and In Fig. b) and d) are the autocorrelations of the responses where the early reflections are added. It can be seen that the early reflections make the autocorrelation plots denser and the peaks appear more irregularly and thus the very harmonic structure of the pure late reverberation is broken. In Fig. 5.15 the crosscorrelation function is computed between responses where incoherent early reflections of the right and the left channels are fed to late reverberation units which are similar for both channels, i.e., the late reverberation block does not create the incoherence but the different timing of the early reflections alone (see Fig. 5.3). The crosscorrelation reveals whether there is incoherence between the channels, by giving a more flat crosscorrelation when the coherence between the responses decreases.

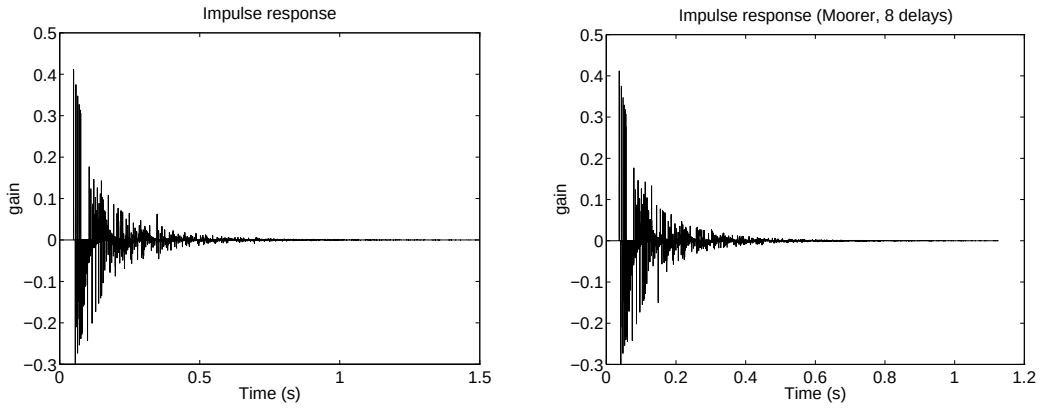


Figure 5.8: Impulse responses of a) 6 parallel comb filters and b) 8 parallel comb filters, with one cascaded comb-allpass filter containing a 6 ms delay.

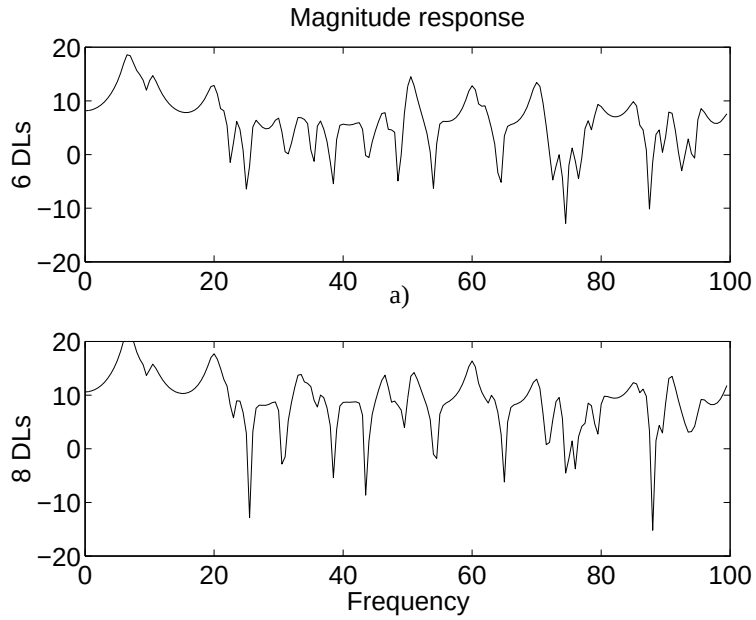


Figure 5.9: The magnitude spectra between 0-100 Hz of reverberators containing a) 6 comb filters b) 8 comb filters.

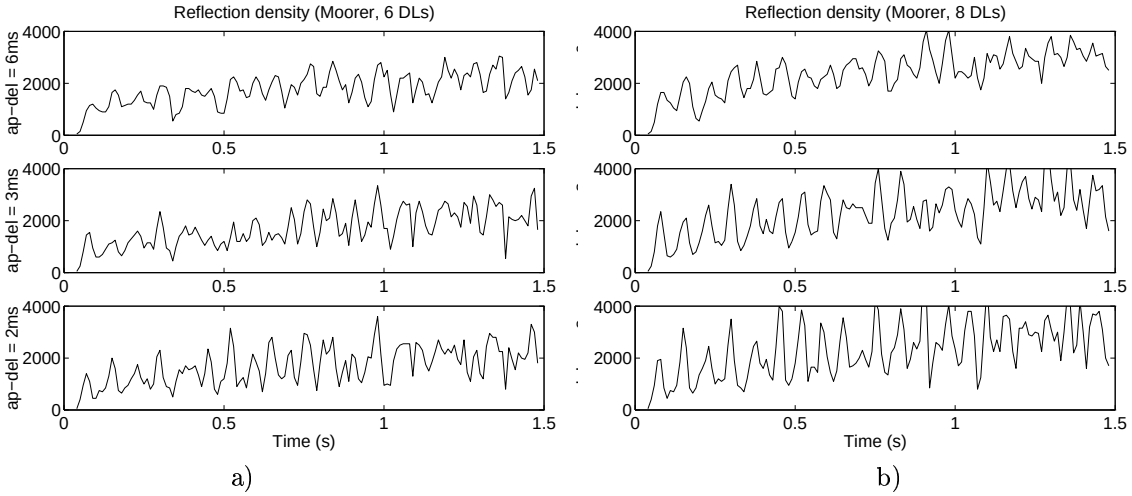
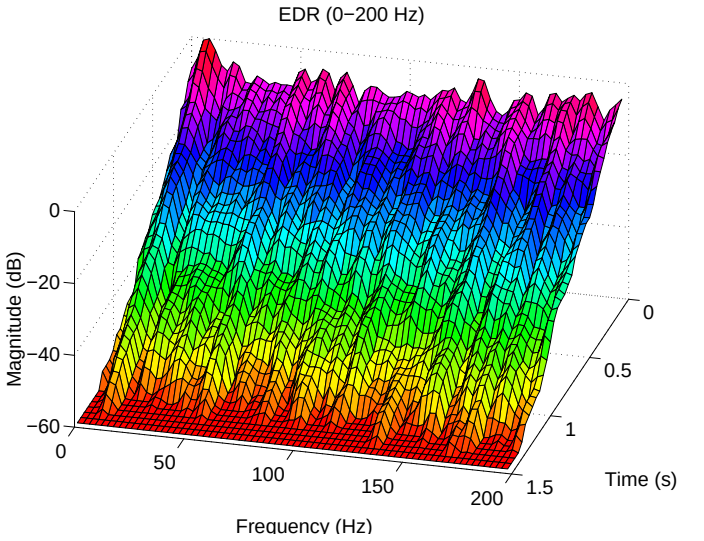
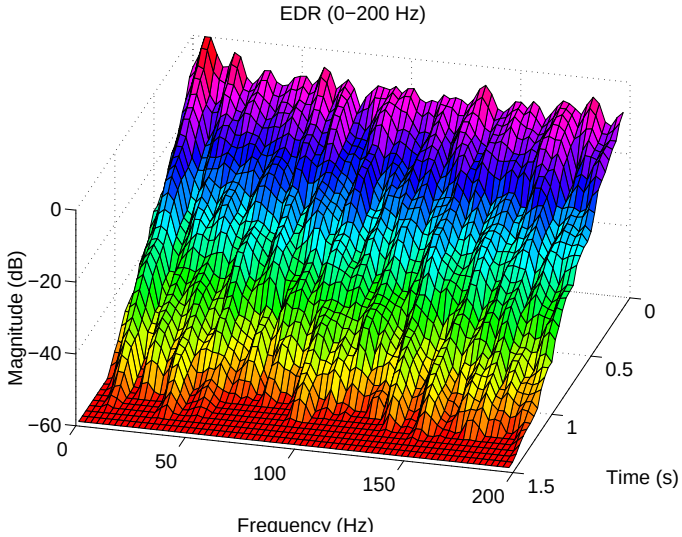


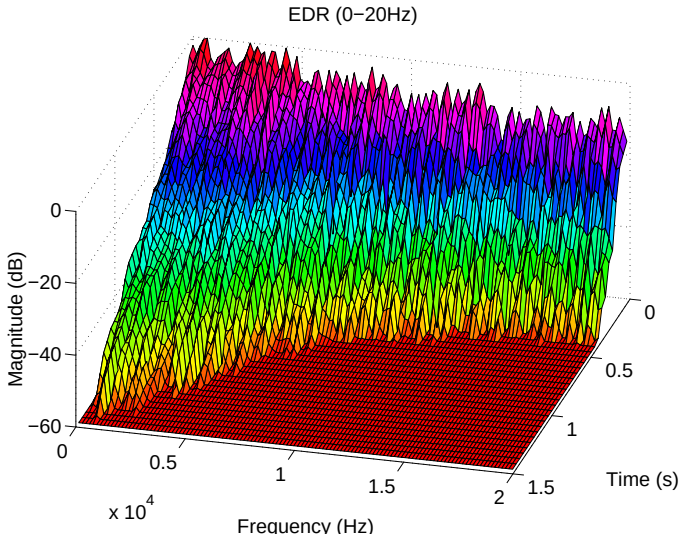
Figure 5.10: Reflection densities of a) 6 parallel comb filters and b) 8 parallel comb filters, with one cascaded comb-allpass filter of delays 2 ms, 3 ms, and 6 ms. The gain g of the comb-allpass filter was 0.7.



a)

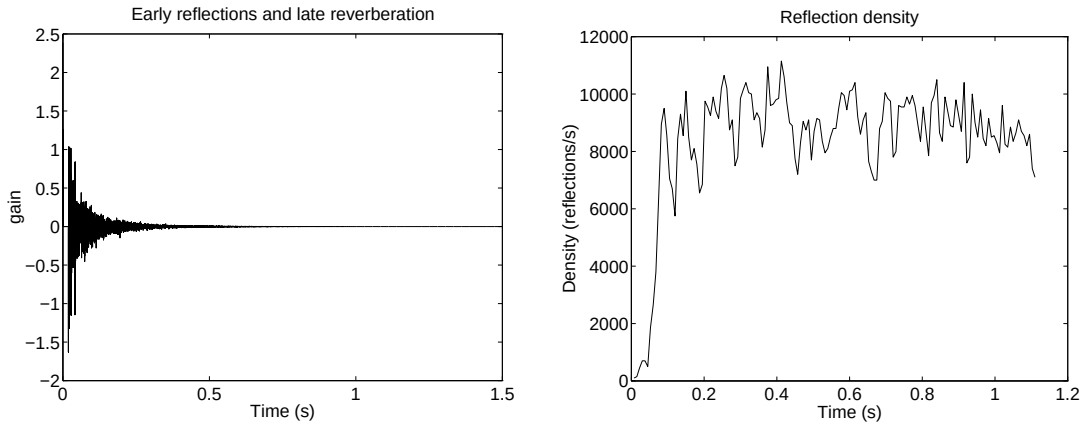


b)



c)

Figure 5.11: Energy Decay Reliefs a) with 6 and b) with 8 comb filters ($0-200$ Hz) and c) with 8 parallel comb filters ($0-20$ kHz).



a) b)

Figure 5.12: a) Impulse response of 8 parallel comb filters with one cascaded allpass filter (allpass delay 6ms) and 20 early reflections fed to the reverberator and directly to the output. b) The reflection density as a function of time of the same structure.

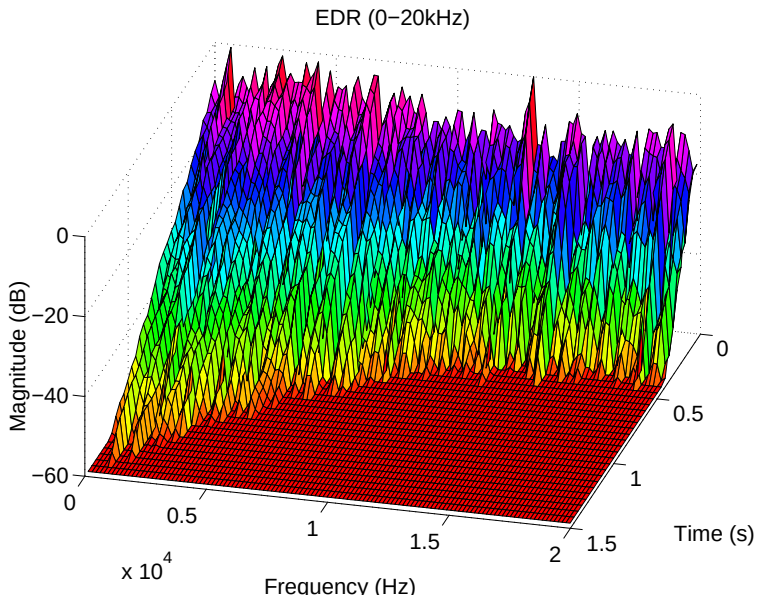


Figure 5.13: An EDR of a response from 8 parallel comb filters and early reflections up to 20 kHz.

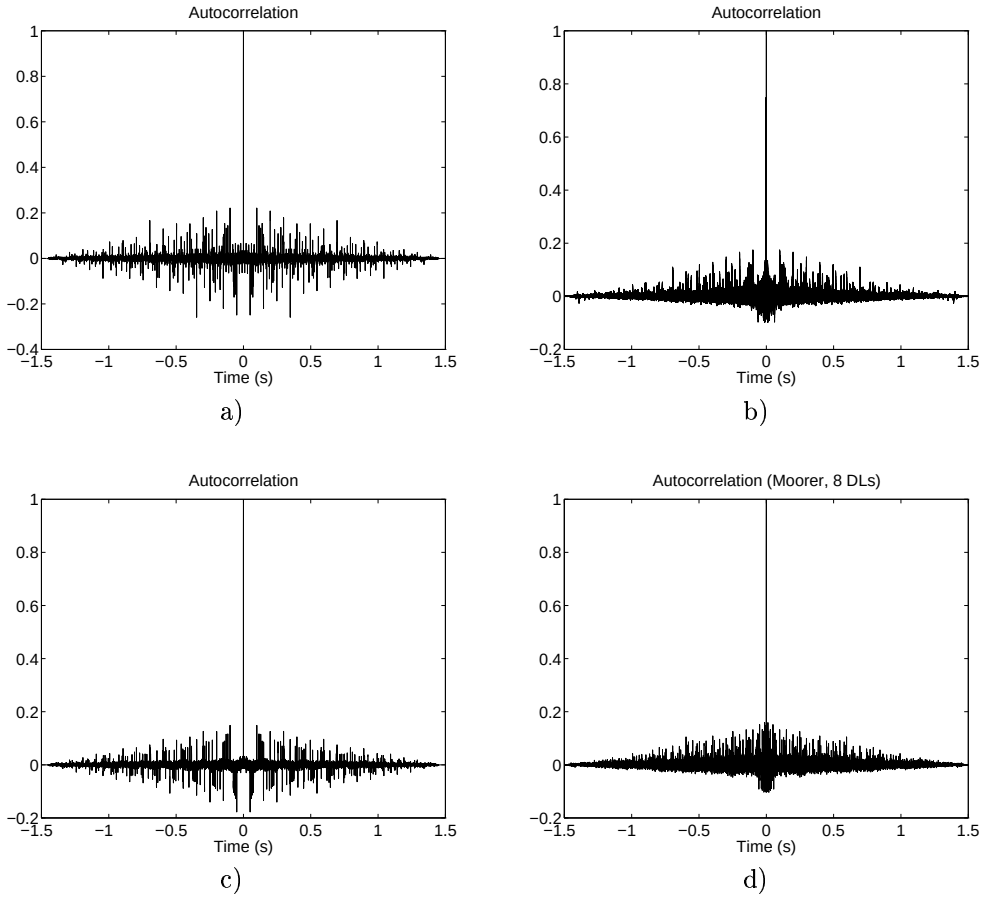


Figure 5.14: Autocorrelation functions of impulse responses from reverberators with a) 6 comb filters, b) 6 comb filters with early reflections, c) 8 comb filters and d) 8 comb filters with early reflections.

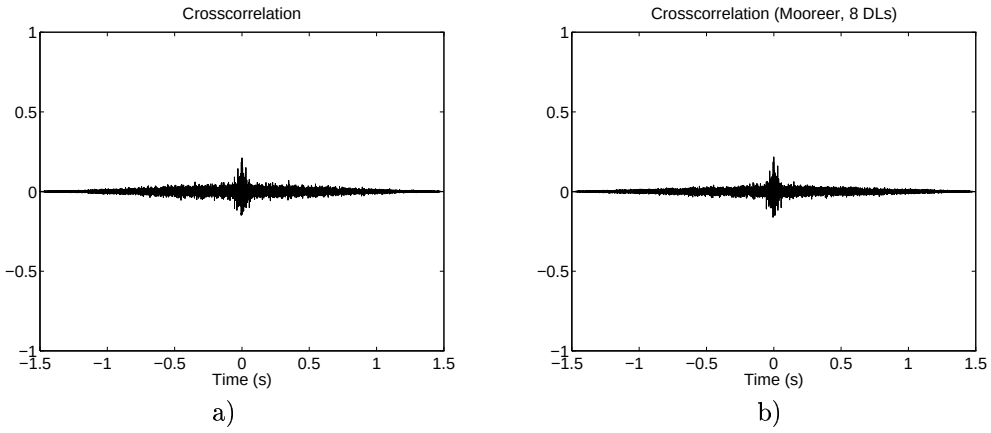


Figure 5.15: Normalized crosscorrelation functions between two channels which are achieved by filtering right and left channel early reflections separately by a reverberator in Fig. 5.7. a) 6 parallel comb filters. b) 8 parallel comb filters

5.5.2 MFDN

In the simulation of MFDNs, unitary and circular feedback matrices were used, as Jot (1992b) suggests. The simulated structure is the same as in Fig. 4.6 with absorbent filters ($H_i(z)$ in Fig. 4.7) added after the delay lines. The coefficients k_i and b_i of the absorbent filter are the same as listed in Table 5.5. The matrices are circulated versions of the following form (Jot 1992b):

$$\mathbf{A_N} = \begin{bmatrix} \alpha & \beta & \cdots & \beta \\ \beta & \alpha & \cdots & \beta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \cdots & \alpha \end{bmatrix} = \beta \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} + (\alpha - \beta) \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{5.8}$$

To yield a unitary and circular NxN matrix the following conditions have to be fulfilled:

$$\alpha^2 + (N-1)\beta^2 = 1 \tag{5.9}$$

and

$$2\alpha\beta + (N-2)\beta^2 = 0, \tag{5.10}$$

i.e., the norm of each column has to be one and the scalar product of two columns has to be equal to zero. For the sizes $N = 6$ and $N = 8$ the following matrices were used.

$$\mathbf{A_6} = \begin{bmatrix} 1 & -2 & 1 & 1 & 1 & 1 \\ 1 & 1 & -2 & 1 & 1 & 1 \\ 1 & 1 & 1 & -2 & 1 & 1 \\ 1 & 1 & 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & 1 & 1 & -2 \\ -2 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} /3 \tag{5.11}$$

and

$$\mathbf{A_8} = \begin{bmatrix} 1 & -3 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -3 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -3 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -3 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -3 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & -3 \\ -3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} /4. \tag{5.12}$$

Also versions of the matrices where the different elements are on the diagonal of the matrix have been tried. In that case the feedback matrix system can be replaced by parallel comb filters, where the sum of the comb filter outputs are fed to the inputs of the comb filters multiplied by a single coefficient. This corresponds to a situation where the diagonal elements $(\alpha - \beta)$ from Equation 5.8 are feedback coefficients of the comb filters and the sum of the comb filter outputs is multiplied by the coefficient β. The block diagram of this system is illustrated in Fig. 5.16. From Equations 5.9 and 5.10 it can be derived that with all matrix sizes the value of $\alpha - \beta$ can be chosen

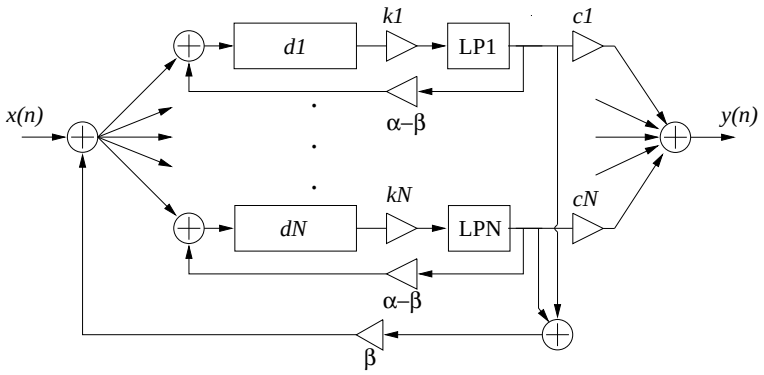


Figure 5.16: A modified MFDN reverberator where the delay lines are replaced by comb filters and the feedback matrix by a single feedback gain.

to be 1 and the feedback coefficient for the summed feedback β becomes a negative number whose absolute value is smaller than one. The advantage of this reverberator when compared to the MFDN is the reduced amount of computation. It increases the reflection density as a function of time like the MFDN but a disadvantage is that because of the greater magnitude of the feedback coefficients $(\alpha - \beta)$ of the comb filters as opposed to the total feedback gain β (e.g., for $N = 6$: $\alpha - \beta = 1$, $\beta = -1/3$ and for $N = 8$: $\alpha - \beta = 1$, $\beta = -1/4$), the modes of the comb filters themselves are emphasized causing stronger comb structure to the frequency response, than in the case of circulated versions of the same matrix. When the elements $(\alpha - \beta)$ are not on the diagonal of the matrix the energy of the modes are more efficiently distributed between different modes. The coefficients preceding the filter output (gains cn in Fig. 5.16) have been chosen so that half of the gains cn are -1 and the rest are 1, to yield a maximally flat response like was explained in Section 4.3.3.

In reverberators where the outputs of all the delay lines are fed to all the inputs of the delay lines, the reflection density of the delay line outputs increases as a function of time, and the signals they produce are mutually uncorrelated. This gives a good possibility for creating a pseudo-stereo effect by feeding different delay line outputs to different channels. This has been applied so that every second delay line output is fed to one channel and the other half of the delay lines to the other. This seems to create a good impression of a more spacious sound field.

In Figures 5.17-5.20 the impulse responses, the reflection density functions, EDRs, the magnitude spectra and autocorrelation functions, of the response an MFDN produces, are shown. From Fig. 5.17 c) and d) it is obvious that the reflection density increases as a function of time, and in the case of an 8x8 MFDN faster than in the case of the 6x6 MFDN. The reflection density of the 8x8 MFDN quickly builds up to the level of approximately 10000 reflections/s which was the minimum requirement Griesinger (1989) suggested to avoid fluttering for short transient like sounds. The magnitude spectra in Fig. 5.18 show that even if the responses do not contain more resonances compared to the Moorer's reverberator structure during the first 100 Hz, the modes seem to be less sharp as a result of the multiple feedback

connection. Thus it can be expected that the MFDN produces less coloration than parallel comb filters. Also the autocorrelation plots in Fig. 5.20 show that there is not same kind of regular periodicity in the response as from that of the comb filters (Fig. 5.14).

In Figures 5.21-5.23 response data from separate channels is presented. In Fig. 5.22 a) and b) are the reflection density plots of the same reverberator as those in Fig. 5.17 but for half of the delay lines (i.e., for two separate channels). It can be seen that the time density is not significantly reduced compared to the total response, i.e., the response where the channels are added together. In Fig. 5.23 the EDRs of two channels of an MFDN are plotted. It can be seen that the modes are placed differently in them, so they produce uncorrelated spectra to each channel. In Fig. 5.24 the same spectra are presented for reverberator of Fig. 5.16, and deeper valleys in the spectra can be noticed even if uncorrelated modes are still present. This implies that the energy of the modes is not as efficiently divided among the delay lines, as in the case of the circulated feedback matrix and therefore in each channels energy from some modes is missing. The crosscorrelation functions of the right and left channels of an MFDN for 6 and 8 delay lines, without and with the early reflections are plotted in Figures 5.25 and 5.26, respectively.

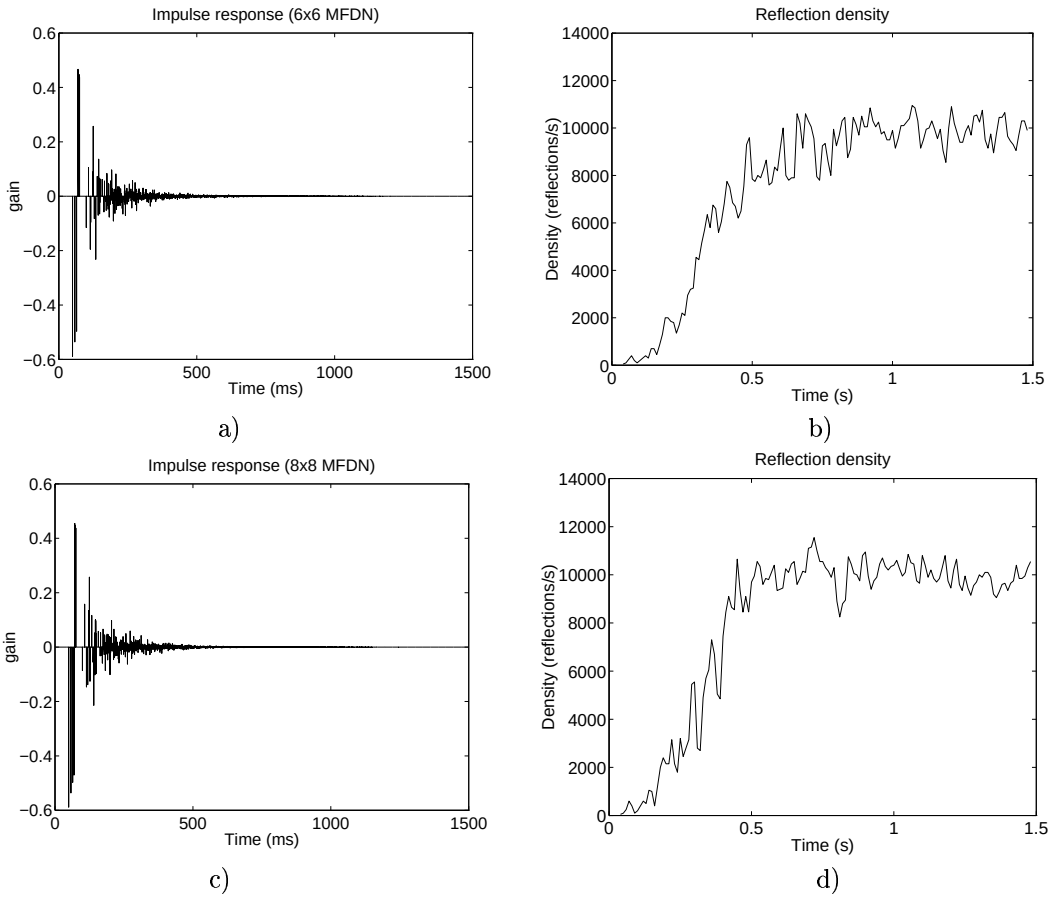


Figure 5.17: a) Impulse responses and the reflection densities of a), b) $6x6$ and c), d) $8x8$ MFDN reverberator.

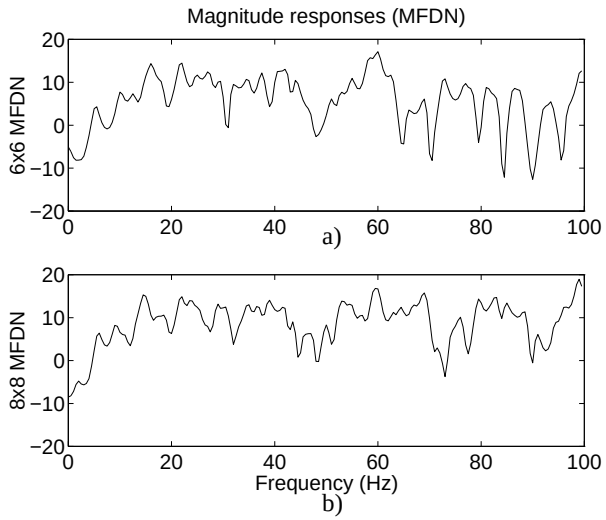


Figure 5.18: A magnitude response between 0-100 Hz a) of an $6x6$ MFDN matrix and b) of an $8x8$ MFDN system.

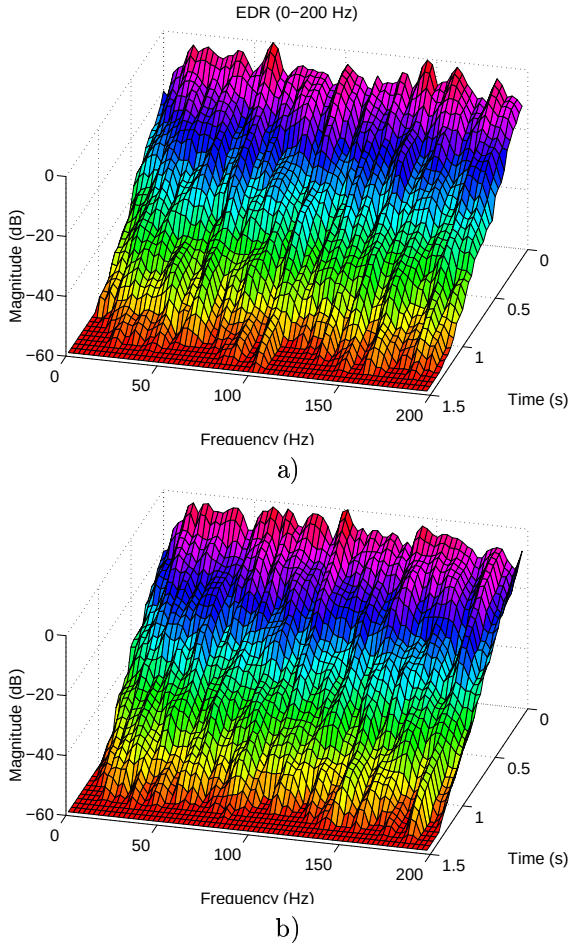
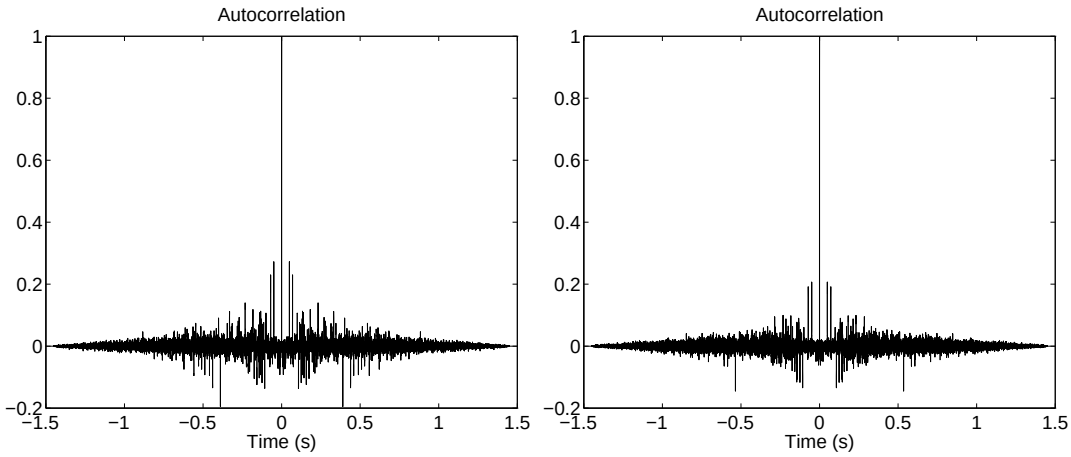
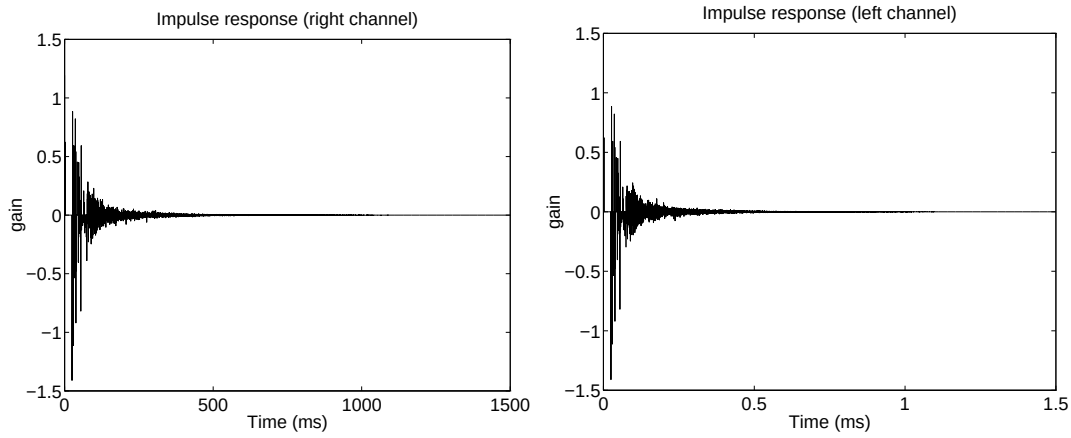


Figure 5.19: Energy Decay Reliefs between 0...200 Hz of a) 6x6 and b) 8x8 MFDN reverberators.



a) b)

Figure 5.20: Autocorrelation functions of a) 6x6 and b) 8x8 MFDN matrix system.



a) b)

Figure 5.21: Responses of one channel, i.e., from the outputs of half of the delay lines from an MFDN reverberator 6x6 (a) and 8x8 (b) MFDNs with early reflections.

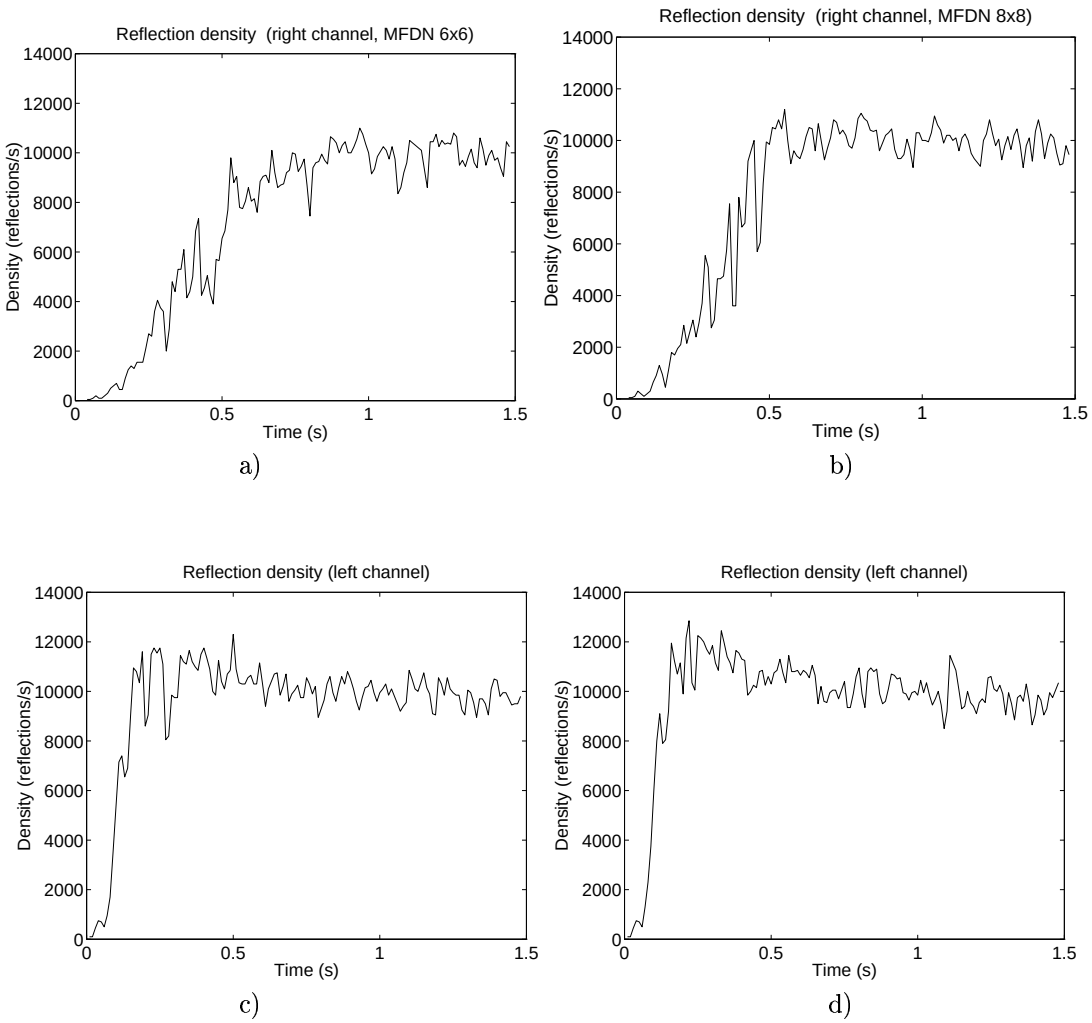
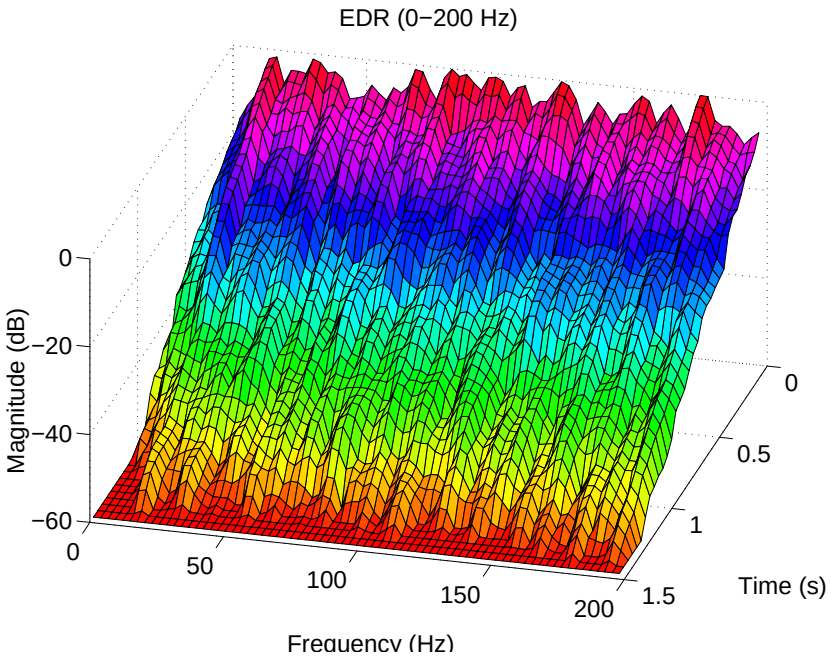
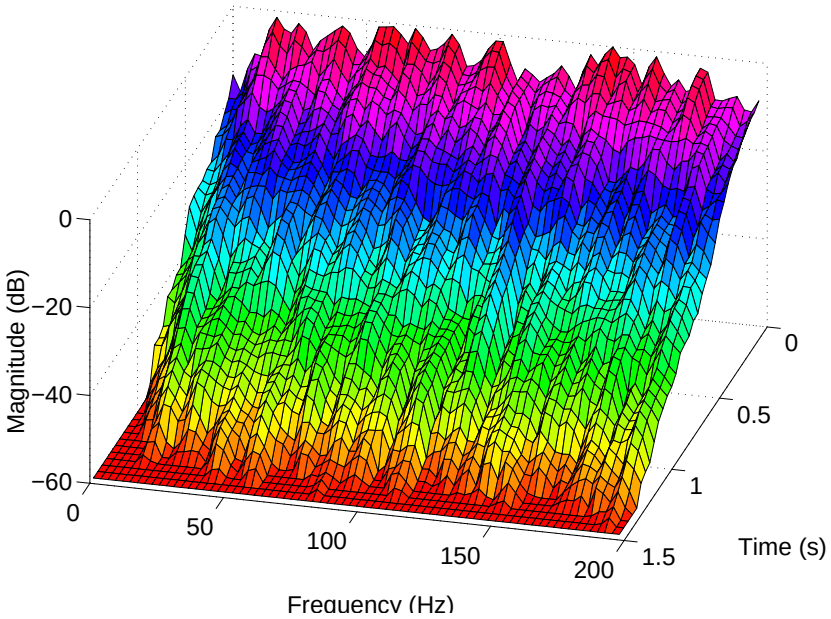


Figure 5.22: Reflection density functions of one channel responses of a) 6x6 and b) 8x8 MFDN. In c) and d) 20 early reflections are added and fed to the late reverberation unit (responses in Fig. 5.21).

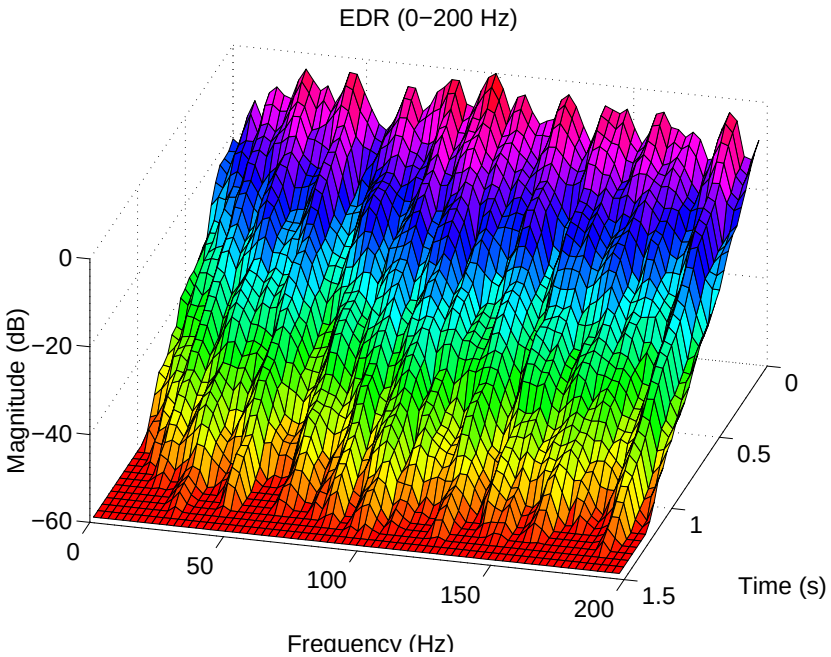


a)

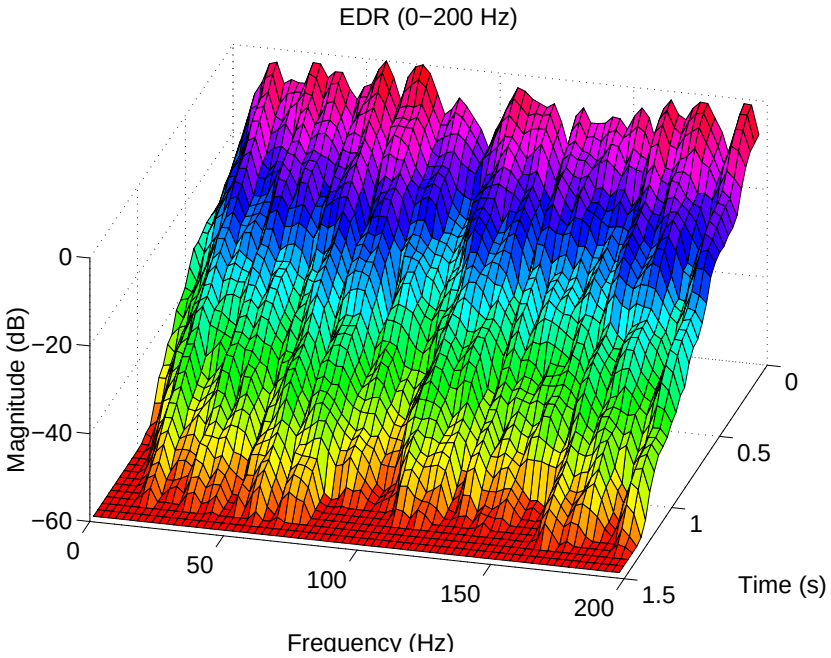


b)

Figure 5.23: EDRs of the right and the left channels of an 8x8 MFDN.

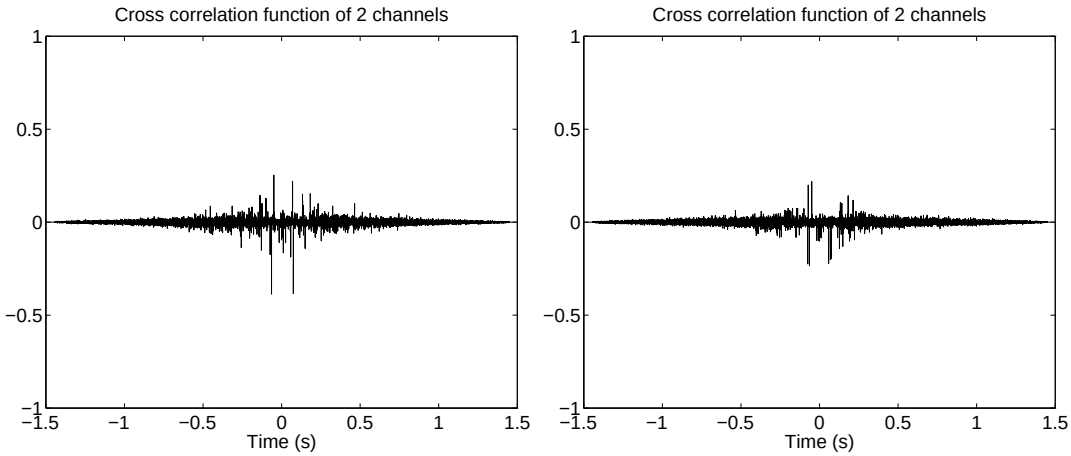


a)



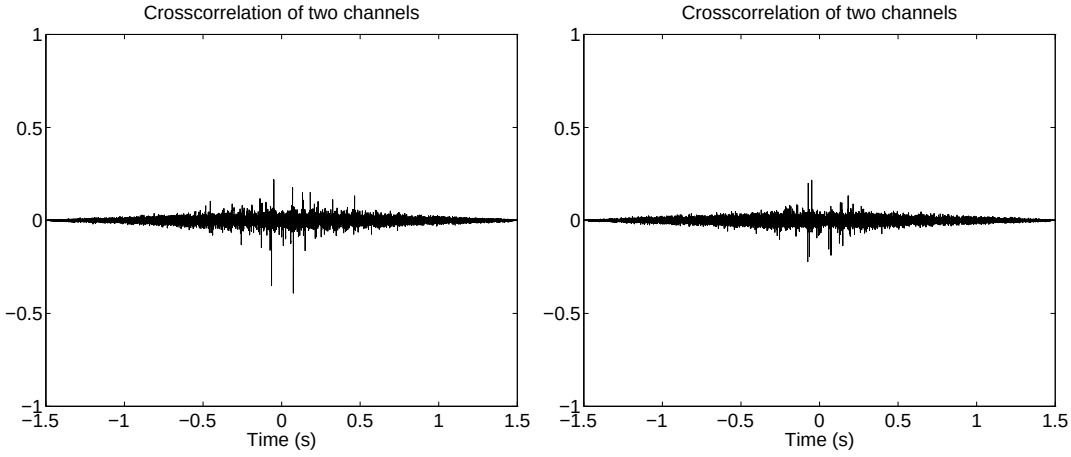
b)

Figure 5.24: EDRs of two channels of a system of Fig. 5.16



a) b)

Figure 5.25: Crosscorrelation functions of the right and the left channels a) MFDN with 6 delay lines and b) 8 delay lines



a) b)

Figure 5.26: Cross correlation functions of the two channels with early reflections a) 6 delay lines b) 8 delay lines

5.5.3 Feedback Through a Comb-Allpass Filter

A modification of the reverberator in Fig. 5.16 is to feed the sum of the comb filter outputs through a comb-allpass filter back to the inputs of the comb filters (Huopaniemi, Karjalainen, Välimäki and Huotilainen 1994). The improvement achieved by this modification is that the allpass filter in the feedback loop acts dispersively, i.e., causes the different frequencies to proceed through the filter at different speeds. This breaks the strict harmonic structure of the comb filter responses. Also the recirculating reflections spread in time resulting in increase of the reflection density. Care must be taken when choosing the parameters related to the comb-allpass filter in the feedback loop, i.e., the delay line length and the gain of the filter, and also the total feedback gain g. If the gain g of the feedback loop in Fig. 5.27 or the delay length of the allpass filter are too large the filter produces strong frequency peaks in the magnitude response or becomes even unstable at some frequencies.

In Fig. 5.28 a) an impulse response of the filter structure in Fig. 5.27 is plotted. The delay length of the filter AP_{fb} was 2 samples and the gain was 0.4. The total feedback gain g in the feedback loop was 0.6. Fig. 5.28 b) shows the reflection density of the same response. It can be seen that the reflection density grows as a function of time as in the case of the MFDN. In Fig. 5.29 a) is the EDR of the same response, and it can be seen that there are modes around about 10 kHz whch have longer reverberation times than the frequencies nearby those frequencies. The longer decay of those modes, as well as their frequency range, is caused by the filter AP_{fb} can be affected by the gain g, and the comb-allpass filter delay length and gain. Nevertheless, when decreasing the gain g too much the advantage of the growing reflection density is lost. Fig. 5.29 shows the autocorrelation of the response, from which some high-frequency correlation can be detected. This correlation is caused by the above mentioned high-frequency modes which live longer than the others.

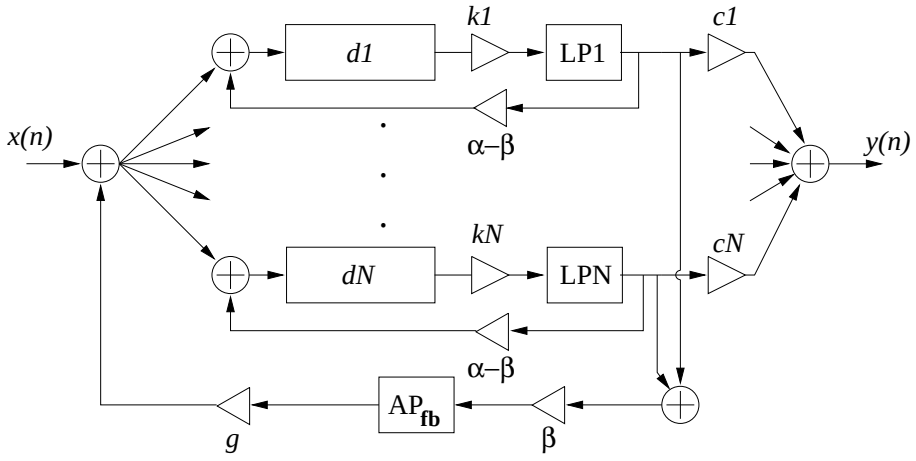


Figure 5.27: Reverberator with feedback through a comb-allpass filter.

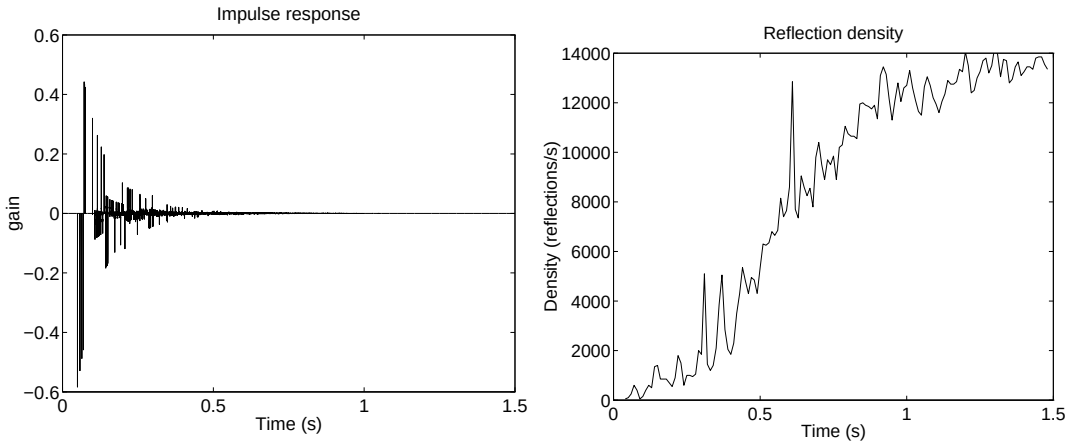


Figure 5.28: a) Impulse response of a reverberator where a summed feedback is done through a comb-allpass filter with a delay length of 2 samples and the gain is 0.4. The total feedback gain g in Fig. 5.27 is 0.6. b) The reflection density of the same response.

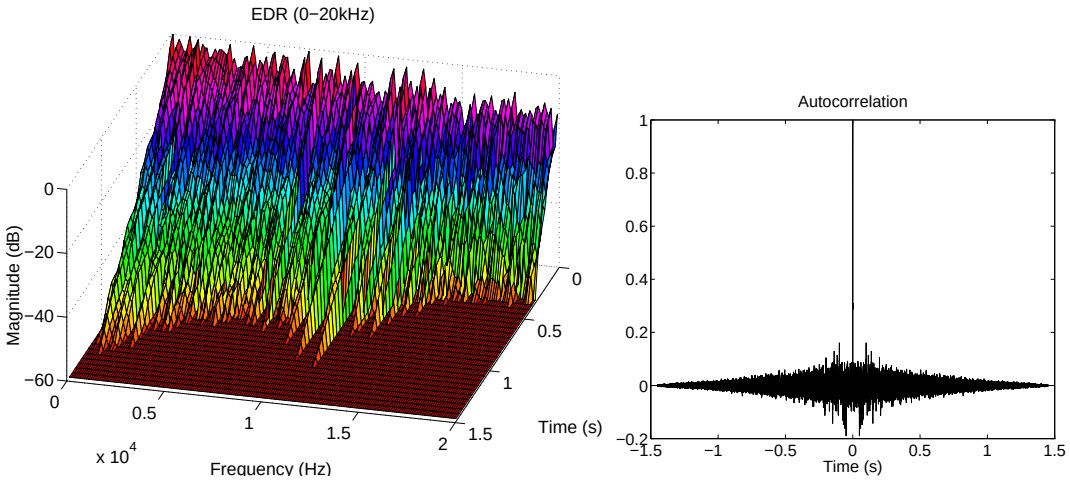


Figure 5.29: a) An EDR between 0-20 kHz of a reverberator where a summed feedback is done through a comb-allpass filter (Fig. 5.27). b) Autocorrelation function of the same response.

5.5.4 Comb-Allpass Filters After Each Delay Line

In this section a reverberator structure where comb-allpass filters are inserted inside combfilter loops is studied. This structure has been explained in (Väänänen, Välimäki and Huopaniemi 1997), and its main advantage is an increased reflection density as a function of time. The structure of the simulated reverberator is illustrated in Fig. 5.30. The values of the coefficients α and β and cn are defined similarly as in the reverberators explained in the previous two sections. The comb-allpass filters inside the combfilter loops do not affect the overall stability of the response like in the case where a single comb-allpass filter is inserted only to the common feedback loop, and thus the gain g of Fig. 5.27 can be left out.

In Fig. 5.31 a) is an impulse response of a single combfilter with a comb-allpass filter in the loop and in Fig. 5.31 b) is the magnitude response of the same structure between $400 - 1200$ Hz. The length of the combfilter delay line is 1000 samples (22.7 ms when f_s =44100), the feedback gain = 0.8 and the delay and the gain of the allpass filter are 100 units and 0.7, respectively. The advantage of an increased reflection density can be seen already in the response of a single combfilter with the comb-allpass filter in the loop. Another effect of the comb-allpass filter can be seen in the frequency-domain response where dispersion causes the frequencies to appear at uneven intervals. The dispersion is a result of different value of the phase delay at different frequencies which causes the length of the total loop delay to be different at different frequencies. The phase delay of the total delay of the combfilter delay and the comb-allpass filter is plotted in Fig. 5.31 c), where it can be detected that the affect of the comb-allpass filter of the delay is strongest at the low frequencies. When choosing the comb-allpass filter delays and gains it is important that they are chosen so that the the frequencies at which the density of the resonances is increased because of dispersion, do not occur at the same frequencies in the case of all comb-allpass filters.

In Fig. 5.32 is the impulse response and the reflection density of this reverberator with 6 delay loops. It can be seen that the reflection density increases faster and up to a higher level than for example that of an 6x6 MFDN reverberator (see Fig. 5.17 b)). Thus it can be considered that with this structure the number of delay lines could be reduced as opposed to the MFDN structure, to obtain the same reflection density, and thus the computation could be reduced. The lengths of the comb-allpass delays were chosen to be approximately 10% of the total delay length of the loop. In other words the sum of the comb loop delay and the comb-allpass delay was now the same as the delay lengths in the previous three discussed reverberator structures. The gains of each comb-allpass filter was 0.5. The magnitude response and the EDR of the reverberator response in Fig. 5.33 a) and b) show that the comb-allpass filters do cause any uncontrollable behavior in the frequency-domain response of the structure.

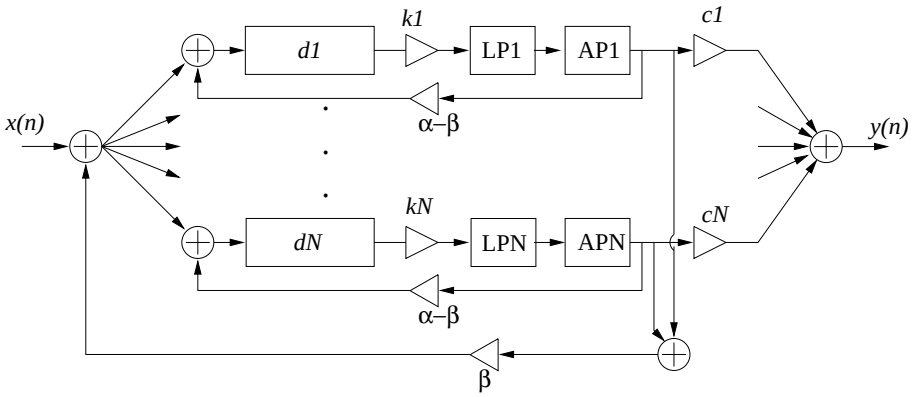


Figure 5.30: Reverberator with comb-allpass filters after the delay lines.

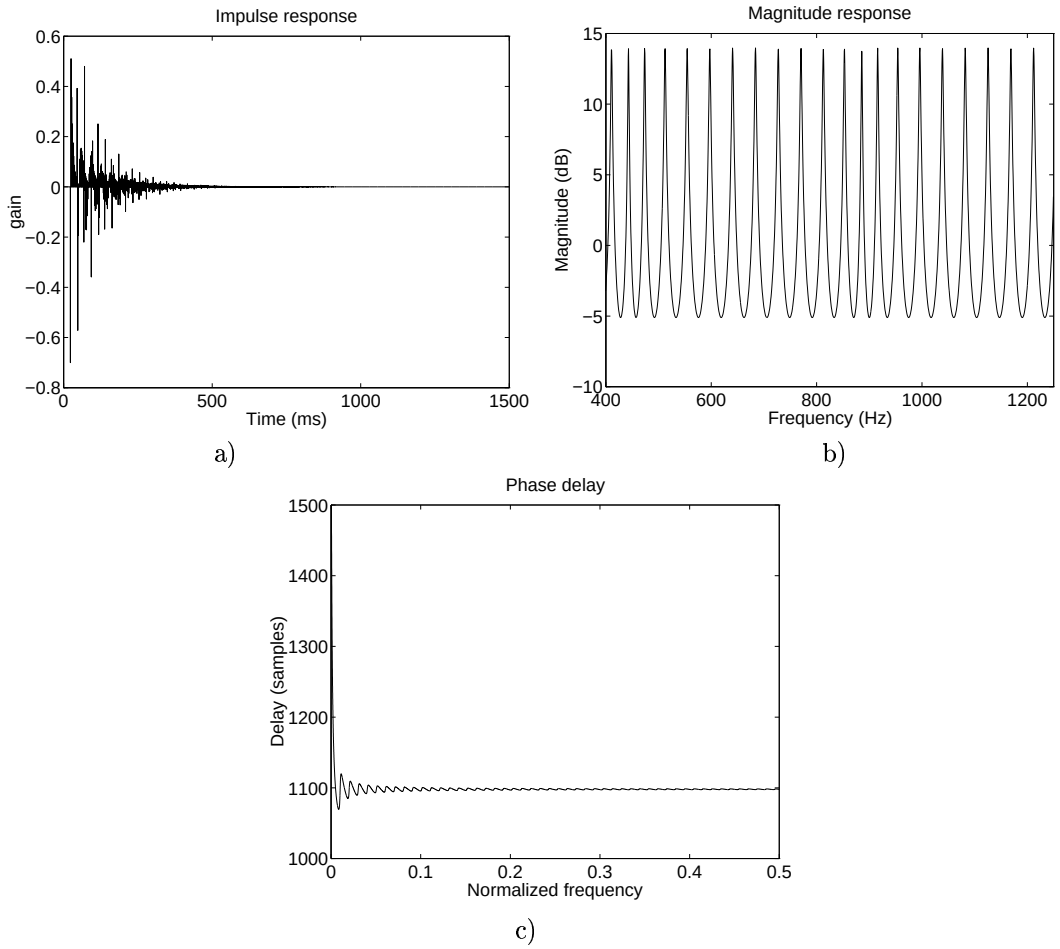
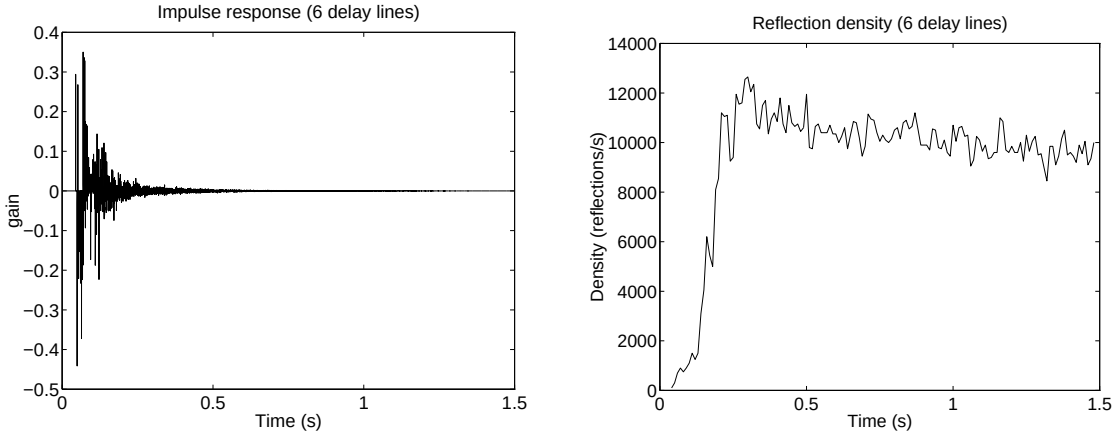
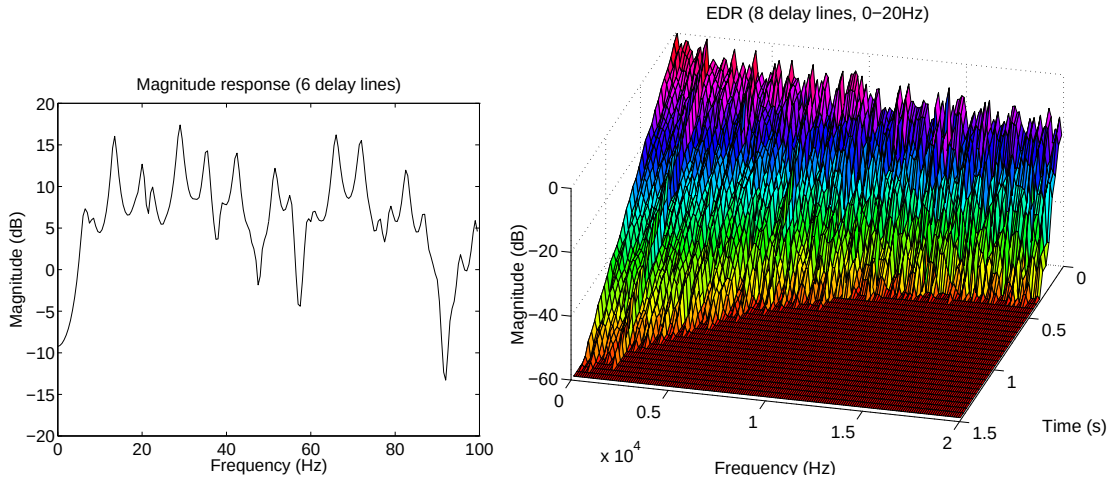


Figure 5.31: Impulse response (a) and magnitude response (b) of a comb filter where a comb-allpass filter is added in the end of the delay line. The length of the comb delay line is 22.7ms and gain = 0.8. The delay of the allpass filter is 2.3ms and the gain = 0.7. c) The phase delay of the combfilter delay and the comb-allpass filter together.



a) b)

Figure 5.32: a) The impulse response and b) the reflection density function of the reverberator where comb-allpass filters are added inside the combfilter loops (see Fig. 5.30). The lengths of the comb-allpass filters was 219, 255, 284, 307, 322 and 331 samples which are approximately 10 % of each total delay loop length. The gains of the comb-allpass filters were 0.5.



a) b)

Figure 5.33: a) The magnitude response ($0 - 100$Hz) and b) the EDR of the response in Fig. 5.32.

5.6 Computational Requirements for the Simulated Reverberators

In this section the computational requirements of each simulated reverberator structure will be presented. The approximations of the requirements are based on the reverberation algorithms and are independent of the environment where they could be implemented. This means that the following assumptions are made:

1. The amount of floating point operations per sample are counted straight from the amount of the summing and multiplication operations in the algorithm, i.e., the possibility of parallel operations is not considered.
2. The amount of memory operations is derived from the times a stored variable or a coefficient is used, or a delay line is updated. Thus it is assumed that a load from or a store to memory has to be done each time a variable or a coefficient is accessed. In other words no registers are used for storing the coefficients of the filter or the values of the delay lines.
3. The intermediate results of the floating point operations can be stored in registers, i.e., every time a result from an operation is used for the following operation it is not stored to the memory between the operations and thus the amount of memory operations is not increased.
4. The constant coefficients reserve an amount of memory in words which is equal to the amount of the coefficients, and the delay lines take an amount of memory in words which is equal to the total length of the delay lines in samples.

By making these assumptions a rough idea about the amount of operations per sample and and the amount of the needed memory is given. Floating point operations are not analyzed separately, but the summing and multiplication operations and the delay line pointer updates are counted together. Memory stores and fetches are under the category of memory operations.

In Table 5.6 the amount of operations per sample of a comb-allpass filter, an IIR lowpass filter and a combfilter with a lowpass filter in the end of the delay line are calculated. The amount of the needed memory is counted for 67.0ms delay line (2950 samples at $f_s = 44.1$kHz) in the case of the comb filter, which is the average length used in the simulation examples in Sections 5.5.1-5.5.4. The length of the delay line in the allpass filter is 6.0ms (265 samples at $f_s = 44.1$kHz). The table gives an impression about how adding a comb or an allpass filter affects on the computational and memory requirements of a reverberator. In Table 5.6 the same specifications are given to the simulated reverberators. The formulas for counting the requirements for each structure containing n delay lines are given and also specifically for $n = 8$. All the approximations for the computational complexity is done for direct form realizations of the algorithms.

	floating point operations	memory operations	amount of memory (words)
comb-allpass filter	5	4	266
first order IIR lowpass filter	3	4	3
combfilter (with lowpass filter in the end of the delay line)	7	7	2954

Table 5.4: The computational requirements of the basic filter structures used in reverberators per sample and per second at the sample rate of 44.1kHz (the first and the second column) and the memory requirements.

	FLOPs/sample	memory operations	needed amount of memory (words)
Parallel combfilters (with IIR lowpass filters in the loops)			
n delay lines	$6n + (n-1)$	$7n$	$\sum d_i + 4n$
6 delay lines	41	42	17197
8 delay lines	55	56	23634
MFDN (Fig. 4.7)			
n delay lines	$2n^2 + 7n - 2$	$n^2 + 6n$	$\sum d_i + n^2 + 3n$
6 delay lines	112	72	17227
8 delay lines	182	112	23690
simplified MFDN (Fig. 5.16)			
n delay lines	$10n$	$8n + 1$	$\sum d_i + 5n + 1$
6 delay lines	60	49	17204
8 delay lines	80	57	23643
allpass filter in the feedback loop (Fig. 5.27)			
n delay lines	$10n + 5$	$8n + 4$	$\sum d_i + 5n + 1 + d_{allp}$
6 delay lines	65	52	17205 ($d_{allp} = 1$)
8 delay lines	85	60	23644 ($d_{allp} = 1$)
allpass filters after each delay line (Fig . 5.30)			
n delay lines	$11.5n + 3$	$12n + 1$	$\sum d_i + 5n$
6 delay lines	72	73	17203
8 delay lines	97	97	23642

Table 5.5: The floating point operations and memory operations per sample and the amount of memory needed in the simulated reverberator structures. The delay line lengths (d_i) are the same as listed in Table 5.5 for 6 and 8 delay lines

5.7 Summary

In this chapter the simulated reverberator structures have been presented, and evaluation based on the time and frequency properties of their responses, as well as on the computational requirements of the algorithms.

The evaluation of the simulated late reverberation implies whether the time and frequency responses are dense enough, and whether the reverberation time is monotonously decreasing as a function of frequency. Although the evaluation is mostly done on the pure late reverberation algorithm, i.e., to the reverberation without the early reflections, it is important to consider the effect of the early reflections to the total response of the reverberation. The early reflections affect increasingly to the reflection density but when they are badly chosen they may cause unpleasant timbre or coloration to sound. Also their level or timing relative to the direct sound and late reverberation affects the subjective quality of the reverberation (Czyzewski 1990).

Nearly all recursive digital reverberators are based on parallel delay lines which are feedback connected either so that the outputs are connected only to the input of the same delay line, as in the case of parallel comb filters, or they can be multiply connected, in which case each delay line output is connected to more than just to the input of the same delay line, like in the case of the MFDNs. The advantage of the latter choice of the feedback connections is that it better simulates natural room impulse response by increasing the reflection density as a function of time, and still the frequency dependent reverberation time can be controlled with as few parameters as in the case of parallel comb filters. A further advantage of MFDNs and its modifications (reverberators in Sections 5.5.2-5.5.4) is that uncorrelated signals can easily be derived from the different delay lines. This is especially advantageous for reverberators where a pseudo-stereo response is wished for spatialization of mono source signal.

6. Conclusions, Application Fields and Further Work

In this chapter are reviewed the application fields where sound spatialization can be applied to. The main topics of this thesis, as well as the results of simulation and evaluation of the artificial reverberation algorithms are summarized, and further improvements on digital reverberation algorithms is considered.

6.1 Application Fields of Sound Spatialization

As was already reviewed in Chapter 1 sound spatialization can be considered either as auralization, when the binaural response between the source and a listener in a room is modeled, or as artificial spatialization where a reverberation and pseudo stereophonic effect to sound is created to give the reproduced sound a more spacious impression. In this section the different sound reproduction systems, where the results could be applied to, are shortly reviewed.

6.1.1 Multi-Channel Sound Reproduction

In conventional two channel stereo systems the different channels can be reproduced by two or more loudspeakers, or by headphones, and the stereo effect is a result of different signals at the listener's ears. Further improvement to that is to use more than two channels to make the sound field broader around the listener. By adding reverberation effect to the channels an impression that the sound is listened in a larger space can be created. If the responses of the different channels are also uncorrelated, the sound field appears even wider. Commercial multichannel systems by Dolby Technologies, are reviewed in the following as well as the Ambisonic Surround Sound system in which a true three dimensional sound field is produced:

1. *Dolby Stereo* and *Dolby Digital* are multichannel encoding systems for film sound, where the separate channels are matrix-encoded to the film track. Dolby Stereo encodes four channels which are the left and right stereo channels, a center channel to keep the dialogue in the middle of the screen, and a surround

channel for the film sound effects. Dolby Digital is an encoding system which matrix-encodes six digital channels to the film track. The two additional channels compared to Dolby Stereo are a second surround channel and a bass track which can be added to all other channels.

2. *Dolby Surround* and *Dolby Surround Pro Logic* are decoding systems for Dolby Stereo sound material for home equipment. Dolby surround decodes the four tracks to three channels, the left, right and the surround channels. Dolby Surround Pro Logic recovers all the four channels encoded by Dolby Stereo.

3. *Dolby Surround Digital* is a home equipment decoding system to recover the six channels encoded by Dolby Digital.

4. *Ambisonic Surround Sound* aims to recreate a three dimensional sound field of the recording situation. It encodes a three dimensional sound field to four channels. The encoded channels are X, Y, Z, and W, from which the first three ones give the coordinates of a monophonic sound source and the W channel is a scaling factor which gives a more even distribution between the four channels. Four channels are required for reproducing the sound if the sound field is created in the lateral plane and eight if a three dimensional sound field is to be created (Malham and Myatt 1995).

6.1.2 Virtual Environments

The above commercial techniques are meant for reproducing sound material where the multichannel information is already recorded. Nevertheless, in modern applications such as computer games, virtual reality and multimedia environments the spatial information to sound can be created by auralization, in which case stereophonic sound is achieved by processing a monophonic natural or synthesized sound. In this case room acoustic modeling techniques can be used for real time simulation of a virtual space.

6.2 Summary, Conclusions and Further work

In this thesis for the degree of Master of Science the fundamentals of room acoustics, the room acoustic modeling techniques, spatialization and room reverberation modeling by means of digital signal processing were studied. The main aim of this research was to find efficient algorithms for creating artificial room reverberation effects as well as artificial spatialization, i.e., stereophonic effect to sound, and compare and evaluate the algorithms according to objective criteria. In this research the reverberation algorithms were simulated as C functions to be called from Matlab. They were convolved in time domain with the input signal, and various methods to evaluate the objective quality were implemented as Matlab functions.

To conclude, the multiply connected feedback delay reverberators presented in Sections 5.5.2-5.5.4 give the most promising results when naturalness and computational efficiency, as well as the artificial spatialization are the main issues in late reverberation simulation. The reverberator structure described in Section 5.5.4, where comb-allpass filters are inserted in the parallel combfilter loops seems especially promising when the reflection density of the reverberation and the computational efficiency are of importance.

As further improvements to reverberation modeling time-varying delay lines in the recursive filter structures can be considered, since it seems to be the solution to avoid coloration in commercial reverberators, such as m5000 analyzed in Section 4.3.6. In this case fractional delays could be considered since changing the delay length a whole sample at a time causes detectable distortion to sound. Another improvement and extension to the models simulated in Chapter 5 is to include more modifiable parameters to the reverberation algorithms, e.g., the ones explained in Section 4.3.6 in connection with the m5000 audio effect processor. In this work in the simulated structures only the proportions of the direct sound, early reflections and the late reverberation, and the frequency-dependent reverberation time were able to be adjusted. An important task would be to arrange subjective listening tests where different features like sufficient reflection and modal densities would be evaluated, since there seems to be no information available about any results from such listening tests or whether this kind of tests have been done at all.

The artificial reverberation modeling could also be extended to modeling of real existing acoustic spaces. In this case the reverberation could be designed with the aid of parameters and early reflections derived from a measured impulse response of a room, or from a response computed by some of the room acoustics modeling methods explained in Chapter 2.

Bibliography

Allen, J. B. and Berkley, D. A. (1979). Image method for efficiently simulating small-room acoustics, *Journal of the Acoustical Society of America* **65**(4): 943–950.

Ando, Y. (1985). *Concert Hall Acoustics*, Springer Verlag, Berlin Heidelberg.

BaiMingsian, R. (1992). Study of acoustic resonances in enclosures using eigenanalysis based on boundary element methods, *Journal of the Acoustical Society of America* **91**(5): 2529–2538.

Barron, M. (1971). The subjective effects of first reflections in concert halls - need for lateral reflections, *Journal of Sound and Vibration* **15**: 211–232.

Barron, M. (1993). *Auditorium Acoustics and Architectural Design*, E & FN Spon.

Barron, M. and Lee, L. (1988). Energy relations in concert hall auditoriums, *JASA* **84**(2): 618–628.

Blauert, J. (1983). *Spatial Hearing, the Psychophysics of Human Sound Localization*, The MIT Press, Cambridge, Massachussets.

Borish, J. (1984). An extension of the image model to arbitrary polyhedra, *Journal of the Acoustical Society of America*.

Borish, J. and Angell, J. B. (1983). An efficient algorithm for measuring the impulse response using pseudorandom noise, *Journal of the Acoustical Society of America* **31**(7): 478–487.

Botteldooren, D. (1995). Finite-difference time-domain simulation of low frequency room acoustic problems, *Journal of the Acoustical Society of America* **98**(6): 3302–3308.

Choi, S. and Tachibana, H. (1989). Estimation of impulse response in a sound field by the finite difference method, *Proceedings of the 13th International Congress on Acoustics (I.C.A.)*, Vol. 2, Belgrad, pp. 129–132.

Czyzewski, A. (1990). A method for artificial reverberation quality testing, *Journal of the Acoustical Society of America* **38**(3): 129–141.

Gardner, B. (1992a). A real-time multichannel room simulator, *Technical report*, MIT.

Gardner, W. (1992b). *Virtual acoustic room*, Master's thesis, MIT.

Gardner, W. (1995). Efficient convolution without input-output delay, *Journal of the Audio Engineering Society* **43**(3): 127–136.

Gibbs, B. M. and Jones, D. K. (1972). A simple image method for calculating the distribution of sound pressure levels within an enclosure, *Acoustica* **26**: 24–32.

Griesinger, D. (1989). Practical processors and programs for digital reverberation, *AES 7th International Conference*, pp. 187–195.

Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete-fourier transform, *Proceedings of the IEEE*, Vol. 66.

Huopaniemi, J., Karjalainen, M., Välimäki, V. and Huotilainen, T. (1994). Virtual instruments in virtual rooms - a real-time binaural room simulation environment for physical models of musical instruments, *Proceedings of the International Computer Music Conference*, Århus, Denmark, pp. 455–462.

ISO/DIS-3382 (1995). Acoustics - measurement of the reverberation time of rooms with reference to other acoustical parameters, *Technical report*, International Organization for Standardization.

Jot, J.-M. (1992a). An analysis/synthesis approach to real-time artificial reverberation, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 1992*, San Francisco, pp. II 221–II 224.

Jot, J.-M. (1992b). *Etude et realisation d'un spatialisateur de sons par modeles physique et perceptifs*, PhD thesis, l'Ecole Nationale Superieure des Telecommunications, Telecom Paris 92 E 019.

Jot, J.-M. and Chaigne, A. (1991). Digital delay networks for designing artificial reverberators, *An Audio Engineering Society preprint 3030 (E-2). Presented at the 90th AES Convention*, Paris, pp. 1–14.

Jot, J.-M., Larcher, V. and Warusfel, O. (1995). Digital signal processing issues in the context of binaural and transaural stereophpny, *Presented at the 98th AES convention. An Audio Engineering Society preprint 3980 (I6)*, Paris.

Kleiner, M., Dalenbäck, B.-I. and Svensson, P. (1993). Auralization - an overview, *Journal of the Audio Engineering Society* **41**(11): 861–875.

Krokstad, A., Strom, S. and Sorsdal, S. (1968). Calculating the acoustical room response by the use of a ray racing method, *Journal of Sound and Vibration* **8**(1): 118–125.

Kuttruff, K. H. (1993). Auralization of impulse responses modeled on the basis of ray-tracing results, *Journal of the Audio Engineering Society* **41**(11): 876–880.

Levine, S. N. (1996). Effects processing on audio subband data, *Proceedings of the International Computer Music Conference*, pp. 328–331.

Malham, D. and Myatt, A. (1995). Sound spatialization using ambisonic techniques, *Computer Music Journal* **19**(4): 58–70.

Moore, F. R. (1990). *Elements of Computer Music*, Prentice-Hall, Inc., chapter 4.

Moorer, J. A. (1979). About this reverberation business, *Computer Music Journal* **3**(2): 13–28.

Naylor, G. and Rindel, J. H. (1994). *Odeon Room Acoustics Program, Version 2.5 User Manual*, The Acoustics Laboratory, Technical University of Denmark.

Oppenheim, A. V. and Schafer, R. W. (1975). *Discrete-Time Signal Processing*, Precntice-Hall, Engelwood Cliffs, NJ.

Orfanidis, S. J. (1996). *Introduction to Signal Processing*, Prentice Hall International. Inc.

Rocchesso, D. (1993). Multiple feedback delay networks for sound processing, *X Colloquium on Musical Informatics*, Milano.

Rocchesso, D. (1995a). The ball within the box: A sound-processing metaphor, *Computer Music Journal* **19**(4): 47–57.

Rocchesso, D. and Smith, J. O. (1997). Circulant and elliptic feedback delay networks for artificial reverebration, *IEEE Transactions on Speech and Audio Processing* **5**(1): 51–63.

Savioja, L., Backman, J., Järvinen, A. and Takala, T. (1995). Waveguide mesh method for low-frequency simulation of room acoustics, *15th International Congress on Acoustics*, Trondheim, Norway, pp. 637–640.

Savioja, L., Rinne, T. and Takala, T. (1994). Simulation of room acoustics with a 3-d finite difference mesh., *Proceedings of the International Computer Music Conference*, pp. 463–466.

Schoenle, M., Fliege, N. and Zoelzer, U. (1993). Parametric approximation of room impulse responses by multirate systems, *International Conference on Aoustics, Speech and Signal Processing (ICASSP) Proceedings*, Minneapolis, pp. I–153 – I–156.

Schroeder, M. R. (1962). Natural sounding artificial reverberation, *Journal of the Audio Engineering Society* **10**(3): 219–223.

Schroeder, M. R. (1965). A new method of measuring reverberation time, *The Journal of the Acoustical Society of America* **37**: 409–412.

Schroeder, M. R. (1970). Digital simulation of sound transmission in reverberant spaces, *Journal of the Acoustical Society of America* **47**(2): 424–431.

Schroeder, M. R. (1979). Integreated impulse method measuring sound decay without using impulse, *Journal of the Acoustical Society of America* **66**(2): 497–500.

SenGupta, G. (1991). Finite element analysis of natural frequencies of acoustic enclosures with periodic properties, *Journal of Sound and Vibration* **145**(3): 528–532.

Smith, J. and Rocchesso, D. (1994). Connections between feedback delay networks and waveguide networks for digital reverberation, *Proceedings of the International Computer Music Conference*, pp. 376–377.

Smith, J. O. (1982). An allpass approach to digital phasing and flanging, *Technical report*, CCRMA Department of Music, Stanford University.

Stautner, J. and Puckette, M. (1982). Designing multi-channel reverberators, *Computer Music Journal* **6**(1): 569–579.

t.c (n.d.). *m5000 Audio Mainframe manual*.

Van Duyne, S. A. and Smith, J. O. (1993). Physical modeling with the 2-d digital waveguide mesh, *Proceedings of the International Computer Music Conference*.

Väänänen, R., Välimäki, V. and Huopaniemi, J. (1997). Efficient and parametric reverberator for room acoustics modeling, *To be published in the International Computer Music Conference (ICMC)*, Thessaloniki, Greece.

Vorländer, M. (1989). Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm, *Journal of the Acoustical Society of America* **86**(1): 172–178.

Vorländer, M. and Bietz, H. (1994). Comparison of methods for measuring reverberation time, *Acustica* **80**(1): 205–215.