Diploma Thesis

Room Response Equalization and Loudspeaker Crossover Networks

(Implementation of the software tools RoLoSpEQ and D-MLCNC)

Author:
Benjamin Mathias Dietze

Technical University of Graz
Institute of Broadband Communications

Director:
Univ.-Prof. Dipl.-Ing. Dr. techn. Gernot Kubin

Assessor and supervisor:
Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Gerhard Graber

Graz, July 21, 2010

STATUTORY DECLARATION:

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, July 21, 2010

Benjamin Mathias Dietze

Acknowledgments

I like to thank my advisor Gerhard Graber for his enthusiasm and for his extensive support during this thesis. I especially wish to thank Franz Zotter for interesting conversations and inputs concerning the sweep measurement method. I also wish to thank Alois Sontacchi for providing valuable suggestions.

Furthermore, I wish to thank all friends and fellow students accompanying me during my years of study.

Especially I like to thank my parents for their support, enabling this study and the marvelous time in Graz. I also want to thank my grandmother for her support and care.

I like to dedicate this work to my partner Birgit. I wish to thank you for your marvelous support, for the patience of many hours of proof reading, and for being there for me every time.

Abstract

In this thesis a software tool for room plus loudspeaker equalization within small rooms is developed. To achieve an equalization a filter generation method, based on an accurate measurement method using exponential sweeps and a target response generation approach, is proposed. A considerable effort has been extended on the implementation of a quasi real-time simulation of the generated equalizing filters, which is based on the introduction of a block-processing approach. This block-processing approach is enabled by using the MATLAB utility "Playrec". As supplementation a further software tool for a digital simulation of multichannel loudspeaker crossover networks, which is also based on the developed block-processing approach, is developed. Thus with both tools a combination of speaker development and subsequent speaker room adaption is supported. The outcomes of this thesis are the two described ready-to-use software tools RoLoSpEQ (**Ro**om plus **Lo**ud**Sp**eaker **EQ**ualizer) and D-MLCNC (**D**igital – **M**ultichannel **L**oudspeaker **C**rossover **N**etwork **C**ontroller).

Kurzfassung

Diese Arbeit befasst sich mit die Entwicklung eines Softwareprogramms zur Anpassung des Schallübertragungsverhaltens von Lautsprechern an die akustischen Gegebenheiten in kleinen Hörräumen. Diese Anpassung geschieht mit Hilfe von digitalen Filtern. Die Filtererstellung basiert auf einer präzisen Messmethodik mit exponentiellen Sweeps und einem Verfahren zur automatischen Generierung eines Zielübertragungsverhaltens. Eine beträchtlicher Aufwand wird für die Implementierung der Simulation der generierten Anpassungsfilter in quasi Echtzeit betrieben. Diese Simulation basiert auf den Einsatz einer blockweisen Signalverarbeitungsstrategie, welche durch Nutzung des MATLAB Zusatzprogramms "Playrec" ermöglicht wird. Als Ergänzung wird ein weiteres Softwareprogramm zur digitalen Simulation von mehrkanaligen Lautsprecherfrequenzweichen entwickelt, welches ebenso auf dieser blockweisen Signalverarbeitungsstrategie basiert. Unter Nutzung beider Softwareprogramme wird eine Kombination der Entwicklung von Lautsprechern und anschließender Lautsprecheranpassung an akustische Gegebenheiten ermöglicht. Die Resultate dieser Arbeit sind die beiden beschriebenen direkt einsatzbereiten Softwareprogramme RoLoSpEQ (**Ro**om plus **Lo**ud**Sp**eaker **EQ**ualizer) und D-MLCNC (**D**igital – **M**ultichannel **L**oudspeaker **C**rossover **N**etwork **C**ontroller).

Contents

1. Introduction

The interaction between loudspeakers and rooms has been a subject of several studies [SW82, PHR94, Too08]. Mainly in the case of small rooms (e.g. living rooms) the impact of room acoustics on the sound reproduction quality of loudspeakers is hard to predict. This is especially due to the effects of the loudspeaker placement in the room and the strong influences of isolated room modes at low frequencies [Too08]. The intention of alternative quality improvement strategies is mainly based on limited possibilities, which are the optimization of the loudspeaker placement and of the room acoustics. According to Toole [Too08] one possibility is the use of digital equalization filters, which allow a correction of adjacent boundary effects caused by the loudspeaker placement, whereas room modes normally cannot be avoided. Pedersen et al. [Ped06, PM07] have developed "RoomPerfect", an approach based on a digital equalization in order to avoid the negative impacts on the loudspeaker sound reproduction behavior in arbitrary shaped rooms. The specialty of this approach is the consideration of two different listening positions. For this purpose two different equalizing filters are derived, one for the listening sweet-spot (FOCUS) and one for the whole room (GLOBAL). With "RoomPerfect", in the case of an equalization in the listening sweet-spot, both the correction of adjacent boundary effects and room mode influences is enabled.

The motivation in this thesis is the room plus loudspeaker equalization in small rooms in order to improve the loudspeaker sound reproduction quality. The developed approach is based on the "RoomPerfect" system, which is implemented on several DSP hardware devices (e.g. on the KRK device ERGO [KS08]). "RoomPerfect" provides a simple measurement routine with fully automatic target response and equalization filter derivation. However an intervention possibility in the target and filter generation process would be desirable in order to achieve an adaption to specific human hearing habits.

Thus in the course of this thesis the software tool RoLoSpEQ is implemented, which enables an automatic generation but also flexible configuration of target responses and equalization filters. RoLoSpEQ enables an accurate measurement of room plus loudspeaker responses in multiple room positions with target response and room equalization filter generation for the listening sweet-spot and for a global listening position within the room. As special feature RoLoSpEQ enables a quasi real-time signal processing, that is the target responses and therefore the equalization filters can be adapted by the user during audio playback.

An additional contribution of this thesis the software tool D-MLCNC, which is a digital multichannel crossover network simulator. D-MLCNC provides an adaptable speaker development based on analog filter theory. The crossover network can also be simulated in quasi real-time and thus the filter configuration can be variated during audio playback. With both software tools a system supporting loudspeaker development and subsequent room adaption is realizable.

This thesis is subdivided into six parts, whereas in part I-III the theoretical background is discussed and in part IV-VI the practical contributions are explained. Part I deals with the theory of room acoustics, whereas the focus is put on the acoustics of small rooms, and

optimization techniques concerning room response equalization. Part II provides an introduction of the filter theory and the practical application in loudspeaker crossover networks. In part III two measurement techniques based on stepped sines and exponential sweeps are discussed and evaluated.

Part IV describes the practical preparations for the implementation of the software tools RoLoSpEQ and D-MLCNC in MATLAB [Mat08]. In part V the structure and functionality of RoLoSpEQ and the developed approach for the target response generation is described and respectively in part VI the software tool D-MLCNC is described.

Part I.

Theoretical Background – Room and Loudspeaker Acoustics

This part deals with the interaction between loudspeakers and rooms. In order to understand this relationship the fundamentals of room acoustics are essential and are described in chapter 2. The basic theories and corresponding fields of application are explained in more detail.

However, the basic room acoustical theory is sometimes not appropriate to explain several effects, which appear in the case of a loudspeaker sound reproduction in small listening rooms. Thus in chapter 3 several impacts on the loudspeaker sound reproduction, which are important in this thesis, are explained.

Finally in chapter 4 the basics of DSP (digital signal processing) based loudspeaker improvement strategies are explained.

2. "Traditional" Room Acoustics

2.1. Wave Acoustic Theory – Modal Region

Under special conditions *wave acoustic theory* provides an adequate description of the propagation of acoustic waves. Thus an exact description of the sound pressure p and the sound velocity $\vec{v}$ at certain positions is possible. In room acoustics the wave acoustic theory is limited to the low-frequency region, where wavelengths[1] are similar to the room dimensions [GW05, Too08].

The wave acoustic theory is based on the wave equation. The basic wave equation of the sound pressure p can be expressed according to [GW05]:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \cdot \Delta p. \tag{2.1}$$

Solving this equation is only possible in the case of simple basic conditions (e.g. rigid wall). E.g. for a plane wave propagation the wave equation changes to:

$$\frac{\partial^2 p\,(x,t)}{\partial t^2} = c^2 \cdot \Delta p\,(x,t) = c^2 \cdot \frac{\partial^2 p\,(x,t)}{\partial x^2}, \tag{2.2}$$

where x is the position in $[m]$ and t is the time in $[s]$.

The wave equation enables a description of several acoustic wave phenomena in the low-frequency region. One important phenomenon are *resonances* (also known as *room modes* – see chapter 3.2.5), which appear more or less in any room because of interfering reflected waves on different room boundaries [GW05, Too08].

2.2. Geometric Acoustic Theory – Ray Acoustics

Geometric acoustics is an adequate description of reflections in rooms. The theory is based on the straight-lined propagation of sound waves. Thus the sound waves are considered to

[1] The *wavelength* is defined as $\lambda = \frac{c}{f}$ in $[m]$, where c is the *speed of sound* in $\left[\frac{m}{s}\right]$ and f is the frequency of the sound wave in Hz.

be rays, which act according to the *laws of geometrical optics* [GW05]. In correspondence to a light beam in the case of optics, in the case of geometric acoustics the term *sound beam* is used instead.

The sound beam can be interfered if there is a certain barrier (e.g. a wall) or a change in the medium [GW05]. One important interference is caused by a reflection of the sound beam at an object. Such reflections are responsible for several effects, which have a major impact on the sound reproduction behavior of loudspeakers in a small room (e.g. adjacent boundary interference – see chapter 3.2.4).

The basic laws of optics also define the restrictions for the usage of the geometric acoustic theory. Hence for sound reflections important basic conditions have to be met. E.g. if a reflection appears on a diffusive object surface the sound beam will be scattered in several directions and the basic law of reflection[2] won't be met. Thus the wavelength has to be large compared to the shape of the reflection surface. In this case the reflected sound waves can be calculated using virtual image sources [GW05, Kut09, EP09]

2.3. Statistical Room Acoustics – Diffusion of Sound

At higher frequencies an increase in the density of room resonances can be expected (see chapter 3.2.5). Therefore the significance of the *wave acoustic theory* decreases [GW05].

For higher frequencies *statistical room acoustics* theory is one possibility to get an accurate description of the acoustical behavior of **large rooms** and is based on *energy balance considerations*. Moreover a completely *diffuse sound field* is assumed in the whole room.

2.3.1. Diffuse Sound Field Theory

A *diffuse sound field* offers constant characteristic within the whole space (e.g. in a room).

The *sound energy* has to be equal within the room [GW05, Too08, Kut09]. Due to this assumption no *isolated room modes or resonances* will exist (cp. to chapter 3.2.5).

Such sound fields are termed *homogeneous sound fields*.

Consequently the *room geometry* will also be an irrelevant factor. The reason is that the sound energy in a certain room point has to be a combination of sound events of several directions to meet the *condition of homogeneity*.

Such sound fields are termed *isotropic sound field*.

But such an *ideal sound field* does not exist in reality. Hence an accurate (and true) description of the (real) room acoustical behavior with statistical room acoustics is restricted by the theory of diffuse sound. Anyway the *Diffuse Sound Field Theory* can be considered as reliable for an evaluation of room acoustical behavior, if certain restrictions are regarded (e.g. the critical distance, see chapter 2.3.4).

[2] The angle of incidence is equal to the angle of reflection.

2.3.2. Room Reverberation – Measurement of the RT_{60}

The *room reverberation* might be one of the most important and representative measures of the acoustical behavior of rooms. In short the reverberation time is commonly known as RT_{60}.

The reverberation time was introduced by W.C. Sabine. Sabine discovered the independency of the power of the sound field excitation and the reverberation time. The definition **can be derived as follows**:
A steady-state sound field is a sound field which is excited by a *constant energy sound-source*. If the sound source is switched on, the sound field is built up to a sound level where the absorbed energy (absorption by room boundaries and the medium) equals the emitted energy of the sound source. If the sound source is switched off, the sound field energy decays. In the case of an *ideal diffuse sound field* a perfect *linear decay* appears for a logarithmic scaling in dB_{SPL}. The time elapsing during a $60dB_{\mathrm{SPL}}$ decay is termed reverberation time RT_{60} [Kut09].

Thereby the power of the excitation signal has no influence on the resulting revererberation time. Merely a sufficient signal-to-noise ratio (SNR) has to be reached to ensure an adequate extrapolation of the RT_{60}. In minimum the excitation signal has to be $20dB$ above the *room noise level*.

The room reverberation time is a property of the room itself. A correct measurement of the room reverberation implies an omnidirectional source (e.g. a special loudspeaker), which offers a uniform sound excitation in all room directions [Too08]. Consequently, also an omnidirectional microphone has to be used for the measurement process. The microphone dimensions should be as small as possible, avoiding certain interruptions to the sound-field for high frequencies where the wavelength is small compared to the dimensions of the microphone.

2.3.3. Sabine's Equation – Calculation of the RT_{60}

For large and highly reflective rooms a high diffuse sound field can be assumed. In this case the Sabine formula leads to an accurate prediction of the reverberation time in a room. The Sabine equation can be expressed according to:

$$RT_{60} = 0.161 \left[\frac{s}{m}\right] \cdot \frac{V}{A}, \tag{2.3}$$

where V is the *room volume* in $\left[m^3\right]$ and A is the total absorption in "sabins" [Too08]. The total absorption A is calculated as sum of N different absorbent areas, located on the room boundaries, multiplied by the corresponding absorption coefficient:

$$A = (S_1 \cdot \alpha_1 + S_2 \cdot \alpha_2 + S_3 \cdot \alpha_3 + ...) = \sum_{i=1}^{N} S_i \cdot \alpha_i, \tag{2.4}$$

where S_i is the i^{th} area in $\left[m^2\right]$ and α_i is the i^{th} absorption coefficient, which has no dimension. A is also known as *equivalent absorption area* [GW05, Kut09], as multiplication of S_i and α_i yields to a certain absorption area with a absorption coefficient $\alpha = 1$. Hence this "new" absorption area offers an *equivalent* absorption.

In smaller cavities a strong change of the basic conditions in the room has to be expected

(e.g. a room volume reduction). Therefore the high diffusive sound field may not be assumed anymore. As a result the Sabine equation becomes less reliable for rooms, which are even smaller and more absorptive. Such rooms are also termed non-Sabine rooms [Too08].

2.3.4. Critical Distance

If a constant sound source in a room is switched on, a steady state sound field is built up (see chapter 2.3.2). This steady state sound field offers a steady state energy density E_{ss} starting from a minimum distance of the sound source. This minimum distance is also termed *critical distance*.

The *critical distance* is a distance at which the energy density of the emitted sound E_{dir} (the direct path signal) is equal to the steady state energy density E_{ss}.

Without derivation the critical distance for an ideal point source with omnidirectional radiation is as follows [GW05, Kut09]:

$$r_c = 0.057 \left[\sqrt{\frac{s}{m}}\right] \cdot \sqrt{\frac{V}{T}}, \tag{2.5}$$

where V is the *room volume* in $\left[m^3\right]$ and T is the reverberation time RT_{60} in $[s]$. Among others (e.g. Eyring's reverberation formula) this equation also depends on the Sabine equation (see chapter 2.3.3).

In the case of a sound source with non-omnidirectional radiation the equation changes to [GW05, Kut09]:

$$r_c = 0.057 \left[\sqrt{\frac{s}{m}}\right] \cdot \sqrt{\gamma \cdot \frac{V}{T}}, \tag{2.6}$$

where γ is the *sound power concentration* (or directivity) of the source (in case of an ideal omnidirectional source $\gamma = 1$). γ is the ratio between sound intensity[3] in main radiation direction and the average sound intensity in all radiation directions:

$$\gamma = \frac{I_{\mathrm{main}}}{I_{\mathrm{average}}}. \tag{2.7}$$

In practice γ is dependent on frequency, that is the critical distance is dependent on frequency as well:

$$r_c\,(f) = 0.057 \left[\sqrt{\frac{s}{m}}\right] \cdot \sqrt{\gamma\,(f) \cdot \frac{V}{T}}. \tag{2.8}$$

2.3.5. Schroeder (Transition) Frequency

As describe in chapter 2.1 *wave acoustic theory* was found to deliver accurate descriptions of the room acoustical behavior in the low-frequency region.

For the acoustical *high-frequency region statistical room acoustics* is used instead. Accurate descriptions of the acoustical behavior can be expected for high reverberant rooms with a high diffusivity (see chapter 2.3). Nevertheless this situation cannot be expected for small listening rooms [Too08], which are of main importance in this thesis.

[3] The sound intensity (or acoustic intensity) $\vec{I} = p \cdot \vec{v}$ is a directive/vectorial quantity, where p is the sound pressure in $[Pa]$ and $\vec{v}$ is the sound velocity in $\left[\frac{m}{s}\right]$.

In between these two acoustical regions a *transition zone* can be found, which is illustrated in figure 2.1. In the center of this transition zone a certain transition frequency can be found.

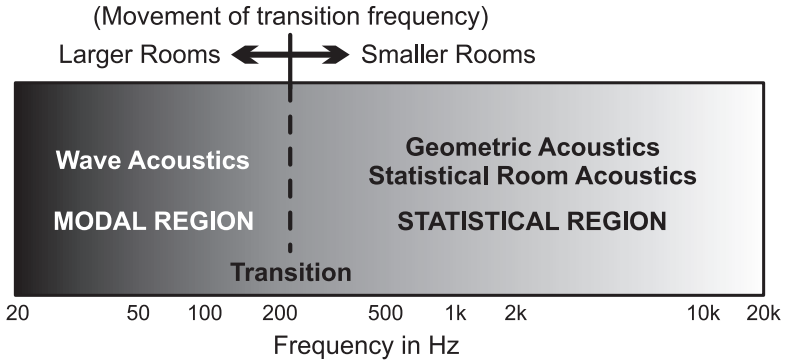

Figure 2.1.: "Transition-zone" between the acoustical LF- and HF-region [Too08].

In this figure another very important effect can be observed. The position of the transition zone is dependent on the room size. Thus the transition frequency increases in the case of smaller rooms. Nevertheless the statistical prediction of the transition frequency in small rooms has not yet been sufficiently examined [Too08].

On the contrary for large or high reverberant rooms – offering high diffuse sound fields – the transition frequency is well defined, and known as *Schroeder Transition Frequency* [Too08, Kut09]. Calculating the reverberation time RT_{60} with Sabine's equation, the Schroeder transition frequency can be calculated according to the following equation:

$$f_c = 2000\sqrt{\frac{T}{V}} \tag{2.9}$$

where T is the reverberation time in $[s]$, an V is the volume of the room in $\left[m^3\right]$.

In small listening rooms, this calculation results in a wrong prediction of the transition frequency, as high diffusive sound fields cannot be expected (see chapter 3). Thus in the case of non-Sabine rooms equation (2.9) yields a transition frequency, which is too low. This is, because the Schroeder transition frequency f_c is based on statistical room acoustics theory.

3. The Acoustics of Loudspeakers in Small Listening Rooms

The main intention of the software based tool RoLoSpEQ (Room and LoudSpeaker EQualizer), which was developed during this thesis (see part V), aims at an improvement of the sound reproduction quality of loudspeakers, which is assumed to be corrupted by the acoustics in *small or domestic listening rooms*. This limitation is defined because home-listeners mainly use small rooms for the sound reproduction with loudspeakers (e.g. loudspeakers in a living room) [Too08]. Nevertheless small rooms are another acoustical challenge, as especially the low frequency behavior can be completely different to the behavior of large rooms. Hence it is obvious to limit the usage of RoLoSpEQ to small and domestic listening rooms, as complete different requirements are needed for the room plus loudspeaker response equalization in small or large rooms.

As described in chapter 2 *statistical room acoustic* can not be directly used for non-high reverberant rooms like small listening rooms. Furthermore the *acoustics of domestic listening rooms* is not studied as extensively as the acoustics of large listening areas (e.g. concert halls) [Too08]. However, many well studied acoustical effects (e.g. standing waves or reflections), which are important in large rooms, are even more important in small rooms.

As described in chapter 2.3.5 a certain transition frequency between the acoustical high-frequency and low-frequency range cannot be calculated directly in the case of small listening rooms, because the *diffuse field theory* is not accurate. Nevertheless it would be desirable to separate the frequency regions according to the acoustical impacts on the loudspeaker sound reproduction. In practice, however, the transition frequency exists in any room and is dependent on the room size. For small listening rooms the transition frequency can reach frequencies up to $500Hz$ [Too08].

During this thesis the statistical room acoustic theory is not used for the evaluation of the acoustical behavior of loudspeakers in small listening rooms. This is because the *main timbral influences* are caused by the effects which appear below the transition frequency, in the acoustical low-frequency region (see chapter 3.2.5 and 3.2.4). Hence several effects can be described by the *wave acoustic theory*.

In the following paragraphs the main timbral influences on the sound reproduction behavior of loudspeakers in small rooms are presented. In addition, in the case of negative influences several proposals can be found which can lead to an improvement of the sound reproduction quality with loudspeakers.

3.1. What is a Small Listening Room?

In general a small listening room is a *non-high reverberant room*, which offers a *non-diffuse sound field*. This is because in domestic rooms a *high absorption* can be expected for a wide frequency region. This high absorption is caused e.g. by different room furnishing [Too08].

In contrast to large listening rooms (e.g. concert halls), a more specific sound field has to be expected. In large listening rooms no dominant reflections are expected, as the sound field is assumed to be homogeneous.

The sound field in small listening rooms consists of the *direct-sound signal* (e.g. from the loudspeaker), several *strong early reflections* (e.g. from the back-side wall of the loudspeaker) and a *much-diminished late reflected sound field* [Too08].

Hence the sound field in small listening rooms will be *non-homogeneous*.

3.2. Source of Acoustic Distortions and Perception of Sound

In figure 3.1 different effects caused by the sound reproduction behavior of loudspeakers in small rooms are illustrated.

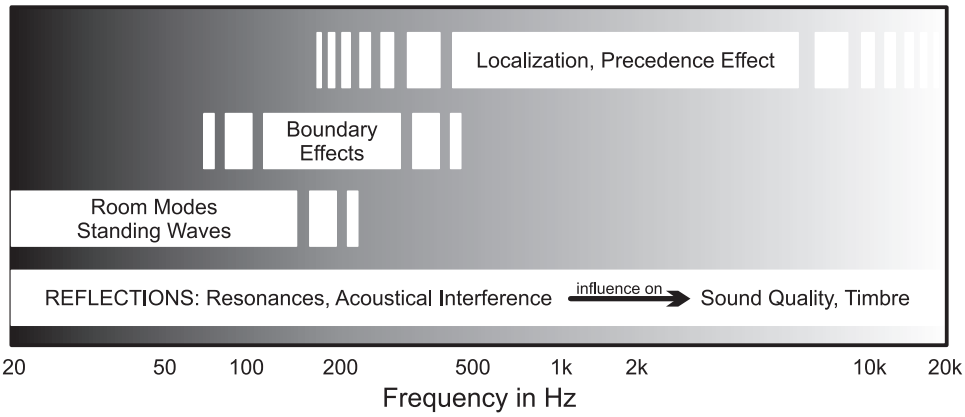

Figure 3.1.: Acoustical effects/distortions at different frequency regions (adapted version, original by Toole [Too08]).

One can see, that certain acoustical effects are limited to certain frequency regions. However each effect is based on the occurrence of *sound reflections*, what can be seen in the very bottom of this figure. In summary *timbral influences* can be expected for the whole human hearing frequency range.

But some effects have more influence on the sound quality then others. A short description of each effect and its corresponding impact on the sound reproduction behavior of loudspeakers is discussed in the following paragraphs.

3.2.1. Summing Localization

The term *localization* describes the phenomenon of *image sources* [Too08]. Image sources are virtual sources which are perceived in directions where no physical sound sources exist in reality – this is a psychoacoustical effect.

If the same signal is applied without a delay on two loudspeakers positioned with a reasonable distance from each other, the perceived result is a phantom image source in between the two loudspeakers. If a delay is introduced in the signal path of one loudspeaker (the radiated signals are still coherent) the phantom image is moved toward the loudspeaker, which radiates the earlier signal. This effect is termed *summing localization*. The summing localization effect is limited to a maximum delay of approximately $1ms$ and for a listening position in the loudspeaker sweet spot[1].

The effect of summing localization is important for the correction of certain displacement errors of loudspeakers in the room. If one loudspeaker is closer to the listener than the other one the phantom image source will be moved toward the loudspeaker which is closer to the listener. Therefore this loudspeaker should be delayed to avoid the image shifting described above. This correction method is also enabled in the software tool RoLoSpEQ (see part V).

3.2.2. Precedence Effect

The *precedence effect* is also known as *Haas effect* or *law of the first wavefront* [Wei08]. This term describes the phenomenon, that the *perceived direction of a sound source* is dependent on the first arriving sound wave.

The effect occurs if early reflections of the direct sound arrive within a certain time interval (e.g. within $1 \rightarrow 30ms$ for speech signals [Too08]). Within this so called "fusion zone" humans are not aware of sound reflections from other directions[2].

Sound reflections that are even later in time (delay $\geq 30ms$), may be perceived as spatially separated source images or even as *echo signals*.

3.2.3. First-Order Reflections

First-order reflections from the walls directly behind the loudspeakers are reported to have negative effects on the sound reproduction quality and the image source localization [Too08]. These reflections are highly coherent to the direct sound radiated by the loudspeakers and additionally from the same direction. A high sound coloration and an impaired image localization might appear.

An improvement can be achieved by adding sound absorbing material behind the loudspeakers.

Lateral first-order reflections are reported to be preferable for a stereo listening situation, because the direction of lateral reflections is completely different to the direct sound radiated by the loudspeakers. Thus the lateral first-order reflections result in a "broadening of the sound image" [Too08], because human perception is capable to separate these two sound events.

[1] A listening position that is central perpendicular on the line between the two loudspeakers in a certain distance. For stereo positioning the listener and loudspeaker placement should be arranged like an equilateral triangle, where the aperture angle is $60°$.

[2] With the exception of early first-order reflections from the same direction like the direct sound – see chapter 3.2.3.

3.2.4. Loudspeaker Placement – Adjacent Boundary Effects

The *loudspeaker placement* has a major impact on the reproduction behavior of a loudspeaker or even what this loudspeaker sounds like [Too08, EP09], especially for lower frequencies. The influences for different loudspeaker placements are caused by a combination of the effects of *room modes* (see chapter 3.2.5) and the interference effects of *adjacent boundaries*. In this paragraph the main topic is the second one (adjacent boundary effect).

Adjacent boundary effects occur at frequencies, where the loudspeaker distance from adjacent boundaries is less then one wavelength [Too08, EP09]. Hence the lower frequency region is affected. Depending on the distance from the boundary, delayed reflections interfere with the direct radiated sound and cause fluctuations in the sound reproduction behavior of the loudspeaker.

Two different methods are known to minimize the effect of adjacent boundaries:

1. An optimized loudspeaker position can be found for each loudspeaker within the space, where adjacent boundary effects are minimized. Here, different mounting methods of loudspeakers (e.g. in-wall) can be used to get an optimized radiation behavior as well [Too08].

2. A room plus loudspeaker equalization can be used to adapt the loudspeaker frequency response to room acoustical basic conditions.

The first method is often merely realizable with very much effort. Basically living-rooms are not optimized listening rooms at all. Therefore a free choice of the perfect loudspeaker place is often not possible. Additionally the perfect place for the loudspeaker might not be the right place according to other reproduction criteria (e.g. Stereo or 5.1 Surround).

The second method is the main intention in the software tool RoLoSpEQ (see part V). This is because this method is considered to be an interesting option for average home-listeners. In addition the loudspeaker placement according to used reproduction methods (e.g. Stereo) can be freely chosen within the listening room.

An objective evaluation:

To illustrate the impact of the loudspeaker placement, two different loudspeaker positions are measured in a typical small living-room (see figure 3.2). The room is rectangular and the size is $522cm$ for the length, $391cm$ for the width and $260cm$ for the height. Hence the room has a volume of approximately $53m^3$. Inside the living-room is a considerable number of furnishing. Therefore the sound absorption is high and no diffuse sound field could be assumed – isolated modes can be expected for lower frequencies (see chapter 3.2.5).

The measurement microphone is an omnidirectional DPA 4006-TL microphone. The sound pressure level (SPL) is calibrated and the loudspeaker sound level is chosen according to a normal living-room level. The measurement signal is a swept-sine with $3s$ duration, which is generated and processed with a sampling rate of $f_s = 48kHz$. The impulse responses are calculated with a length of $2s$. This impulse responses are analyzed in full length in order to consider all room reflections, which interfere the loudspeaker response (room plus loudspeaker response). Therefore a 96000 samples FFT is used to analyze the data with a frequency resolution according to $1Hz$. Further details about the measurement method are discussed in part III.

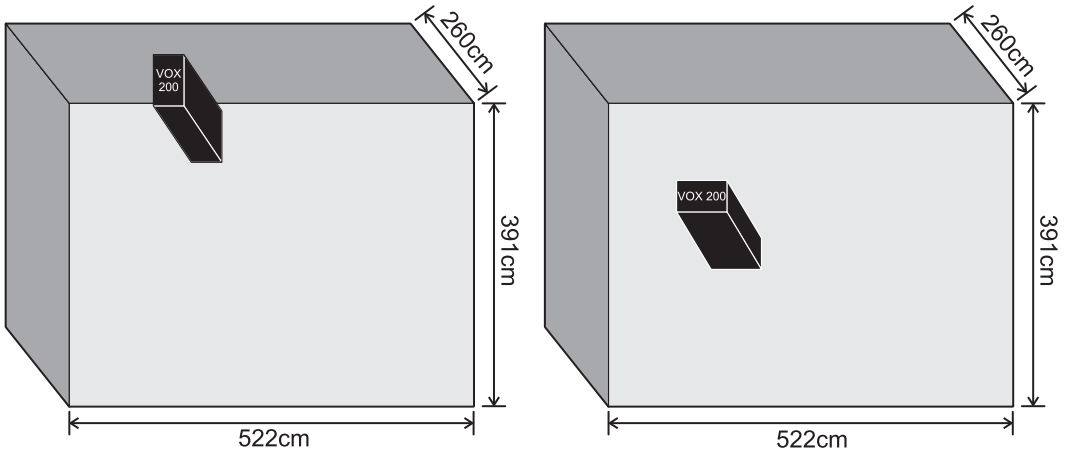

Figure 3.2.: Two different loudspeaker placements in a small living room. Border placement with π steradian radiation (left figure) and ground placement with 2π steradian radiaton (right figure).

At **the first measurement** situation a pillar loudspeaker (Visaton VOX200) is placed close to a room border on the floor[3], well away from any room corners (see left figure in figure 3.2). This *border placement* is also termed according to its resulting sound radiation aperture angle in steradians. Hence this placement yields a π steradian radiator and a quarter sphere radiation can be expected for lower frequencies [Dic05, Too08]. In total nine SPL measurements with different microphone positions are performed within the space.

In figure 3.3 these nine different measurements can be seen. All measured responses are post processed by a smoothing function[4] with $1/24th$ octave resolution. Additionally the root mean square (RMS) of the sound pressure P in $[Pa]$ for each frequency bin of several measurement is calculated, to get an accurate information about the sound reproduction behavior of this loudspeaker in this room:

$$P_{\mathrm{rms}}(f) = \sqrt{\frac{1}{N}\sum_{N=1}^{9} P_N^2(f)}, \tag{3.1}$$

where f is a certain frequency in $[Hz]$, and N is the measurement number. This method is appropriate to get the shape of the room-boundary effects [Too08], and is also used in the software tool RoLoSpEQ (see part V).

At **the second measurement** position the same pillar loudspeaker is placed well away from any side wall (see right figure in figure 3.2). This *ground placement* yields a 2π steradian radiator. A half sphere radiation can be expected at lower frequencies [Dic05, Too08]. The nine SPL measurements are repeated.

In figure 3.4 the measured responses are illustrated. Several measurements are again post processed by the smoothing function with $1/24th$ octave resolution. The root mean square (RMS) is also calculated, in order to get an accurate information about the sound reproduction behavior of this loudspeaker at this new position.

[3]This is the standard position of this loudspeaker in this living-room.
[4]The process of smoothing is discussed in more detail in chapter 13.

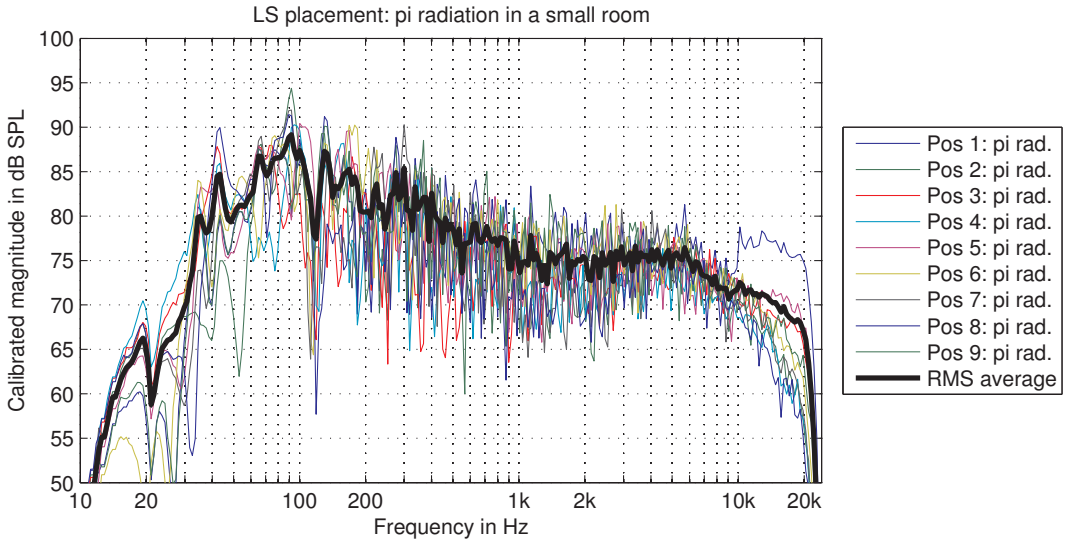

Figure 3.3.: Measurements of a pillar loudspeaker in a small room; *SPL measurements* in 9 Positions with *SPL RMS average* (black curve); π steradian radiation placement.

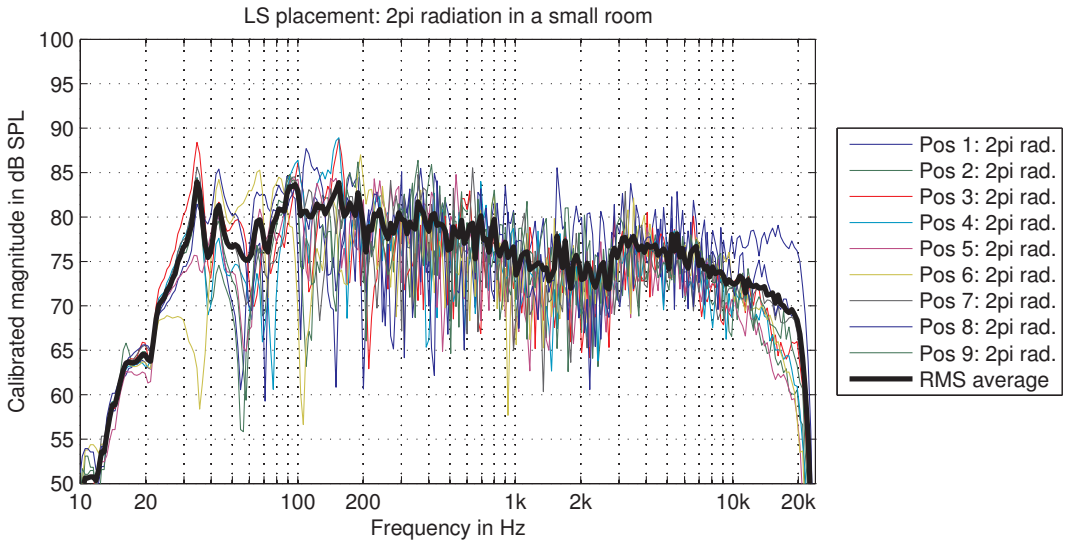

Figure 3.4.: Measurements of a pillar loudspeaker in a small room; *SPL measurements* in 9 Positions with *SPL RMS average* (black curve); 2π steradian radiation placement.

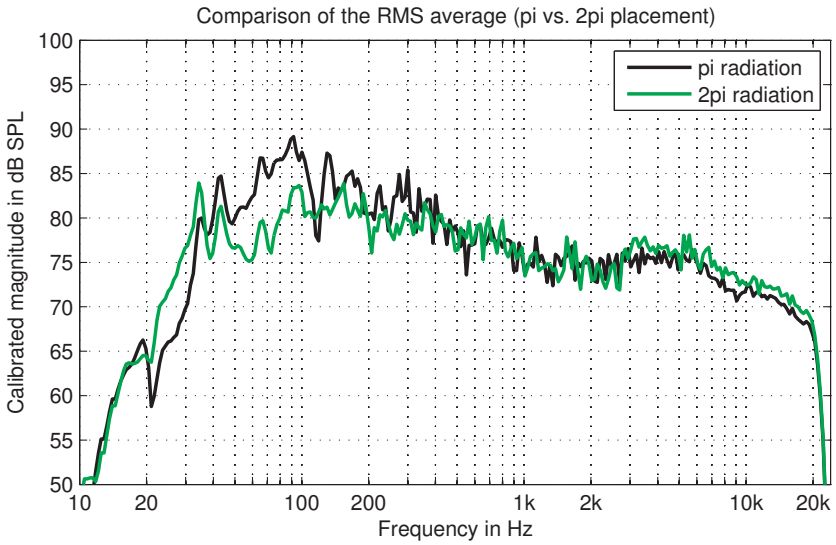

Figure 3.5.: Measurements of a pillar loudspeaker in a small room; *SPL RMS mean* direct comparison; π (black curve) vs. 2π (green/gray curve) steradian radiation placement.

In figure 3.5 the direct comparison of the RMS calculations of both loudspeaker placements is illustrated. One can see, that the influences of the placements decrease for higher frequencies, as the two measurements are mainly similar for frequencies above $400Hz$. This might be caused by the decrease of the wavelength. As a result the 2π or the π radiation cannot be assumed for higher frequencies, as the loudspeaker directivity increases [Too08].

However for frequencies below $400Hz$ the impact of different placements highly increases. This is induced by the different radiated sound power for distinct loudspeaker placements. For an omnidirectional source (with 4π steradian radiation) the radiated sound power is the same in each direction [Too08].

In the case of a 2π steradian radiation the complete sound intensity is radiated in a half sphere. As a result an up to $+3dB$ higher sound power[5] radiation can be expected compared to an omnidirectional source.

In the case of a π steradian radiation the complete sound intensity is radiated in a quarter sphere. As a result an up to $+6dB$ higher sound power radiation can be expected compared to an omnidirectional source.

Conclusion:
For frequencies above approximately $400 - 500Hz$ (for small listening rooms) the effect of adjacent boundaries is strongly diminished [Too08], mainly due to the increasing loudspeaker directivity for higher frequencies. Consequently, the adjacent boundary effect is a *low frequency broad-band effect* and is not focused on specific frequencies (cp. room modes – see chapter 3.2.5).

The adjacent boundary effect has considerable influences on the sound reproduction behavior of loudspeakers [Too08, EP09]. These influences are highly dependent on the acoustical

[5]The acoustical power is defined as $P_{\text{ac}} = \int^S \vec{I} \cdot \mathrm{d}\vec{\mathrm{S}} = \int^S p \cdot \vec{v} \cdot \mathrm{d}\vec{\mathrm{S}}$, where $\vec{I}$ is the sound intensity in $\left[\frac{W}{m^2}\right]$ which is integrated over the complete radiation surface S in $\left[m^2\right]$.

behavior of the room. For a very high absorbent room (e.g. the control room of a recording studio) the direct sound will dominate the reflected sound field (reflections will be strongly diminished), thus the impacts will be small. Even higher influences can be expected for less absorbent rooms, like average living-rooms. Sound reflections from room boundaries cause a timbral change in the sound reproduction at the lower frequency region. A room plus loudspeaker equalizer can be used to reduce these timbral impacts.

3.2.5. Room Modes – The Audibility of Resonances

Room Modes are one important acoustical phenomenon, which can have a significant impact on the sound reproduction behavior of loudspeakers in rooms in specific listening positions. In general room modes are modal frequencies (so called eigenfrequencies) of a room with any shape. Hence any room has got room modes at certain frequencies, where a *standing wave* appears.

Standing waves are caused by a perfect constructive interferences of sound waves, which are traveling between two or more room boundaries [Too08]. The standing waves appear for several modal frequencies of a room and can be derived with the basic wave equation (see chapter 2.1).

Eigenfrequencies in a rectangular room:
The sound field within a rectangular room can be described by solving the acoustic wave equation [GW05, Kut09]. For the solution of the wave equation several conditions have to be met:

- The room has rigid boundaries. Therefore the sound velocity $\vec{v}$, being the normal component on the surfaces of the walls, has to be zero [GW05, Kut09].
- The room is excited by a harmonic oscillation [GW05].

For the condition of rigid walls the basic interference of the incident and the reflected sound wave is [GW05]:

$$\underline{p}\,(x,t) = \underline{p}_i\,(x,t) + \underline{p}_r\,(x,t) = \left(\underline{\hat{p}}_i \cdot e^{-jkx} + \underline{\hat{p}}_r e^{jkx}\right) \cdot e^{j\omega t} = \underline{\hat{p}}_i \left(e^{-jkx} + e^{jkx}\right) \cdot e^{j\omega t}, \quad (3.2)$$

where $\underline{p}$ is the sound pressure in $[Pa]$, x is the position in x-direction in $[m]$, t is the time in $[s]$ and k is the wavenumber in $\left[\frac{1}{m}\right]$.

The expansion of equation (3.2) for the two additional directions y and z for a 3D space in Cartesian coordinates yields:

$$\underline{p}\,(x,y,z,t) = \underline{\hat{p}}_i \left(e^{-jkx} + e^{jkx}\right) \cdot \left(e^{-jky} + e^{jky}\right) \cdot \left(e^{-jkz} + e^{jkz}\right) \cdot e^{j\omega t}. \quad (3.3)$$

With Euler's formula[6] the complex exponential function can be exchanged by a simple cosine function. The expansion of equation (3.2) for two additional directions y and z in a 3D space in Cartesian coordinates yields:

$$\underline{p}\,(x,y,z,t) = 8\underline{\hat{p}}_i \cos(kx) \cdot \cos(ky) \cdot \cos(kz) \cdot e^{j\omega t}. \quad (3.4)$$

[6] Euler's formula describes an exponential function with a trigonometric function according to: $\cos(kx) = \frac{e^{-jkx}+e^{jkx}}{2}$.

The relationship between the room dimension and the eigenfreuencies is [GW05]:

$$l_{x/y/z} = n_{x/y/z}\frac{\lambda_{\mathrm{res},x/y/z}}{2} \quad \Rightarrow \quad f_{\mathrm{res},x/y/z} = \frac{c_0 n_{x/y/z}}{2l_{x/y/z}}, \tag{3.5}$$

where n is an integer value and represents the order of the according room mode.

With equation (3.5) it is possible to calculate several eigenfrequencies within an ideal shaped rectangular room. Another representation can be found if specific eigenfrequencies are evaluated on the axes of a Cartesian coordinate system, that is the axes are representing the eigenfrequency in a certain direction (x,y or z). The distance from the origin represents the eigenfrequency of a standing wave which can appear within the room [GW05]:

$$f_{\mathrm{res}} = \sqrt{\left(\frac{c_0 n_x}{2l_x}\right)^2 + \left(\frac{c_0 n_y}{2l_y}\right)^2 + \left(\frac{c_0 n_z}{2l_z}\right)^2}. \tag{3.6}$$

The eigenfrequencies can be subdivided into three basic groups [GW05]:

1. **Axial modes:** The sound wave propagation is parallel to one room boundary and is perpendicular on two opposite walls. Two of the three order numbers n_x, n_y, n_z have to be zero. Axial modes have the major impact on the resonant behavior in a room, because of the low density of room modes to very low frequencies, which belong mainly to axial modes.

2. **Tangential modes:** The sound wave propagation is perpendicular on one room boundary and is tangential to a pair of walls. The sound wave hits another pair of walls with oblique incidence. One of the three order numbers n_x, n_y, n_z has to be zero.

3. **Oblique modes:** The sound wave hits each wall with oblique incidence. All three order numbers n_x, n_y, n_z have to be non-zero.

An objective evaluation:

Equation (3.6) is used to calculate room modes of a typical small living-room, which is subsequently measured. The room is rectangular and the size is $l_x = 522cm$ for the length, $l_y = 391cm$ for the width and $l_z = 260cm$ for the height.

In figure 3.6 the calculated axial modes are illustrated. A considerable number of isolated

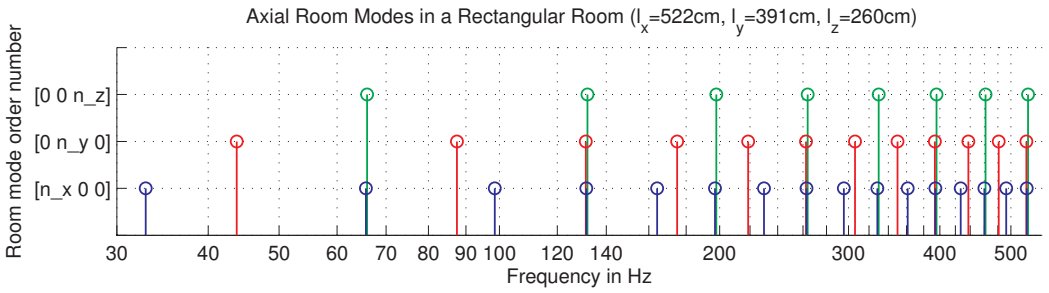

Figure 3.6.: Axial room modes in a small listening room.

axial modes have to be assumed for low frequencies, which might have an enormous impact on

the loudspeaker sound reproduction in this room at certain listening positions. This impact might be caused by a huge amount of specific (multiple) isolated modes[7].

In figure 3.7 the calculated tangential modes are illustrated. A higher mode-density can be

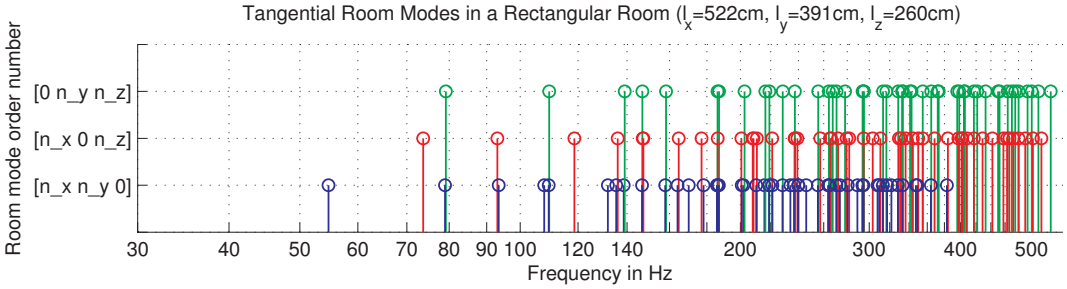

Figure 3.7.: Tangential room modes in a small listening room.

observed. Therefore the influence of tangential modes on the loudspeaker sound reproduction will be diminished. Nevertheless, (multiple) isolated tangential modes can be observed for frequencies below approximately $120Hz$. These (multiple) isolated modes might superpose with each other, or additionally with isolated axial or oblique modes with the same frequency.

In figure 3.8 the calculated oblique modes are illustrated. The oblique modes are shifted

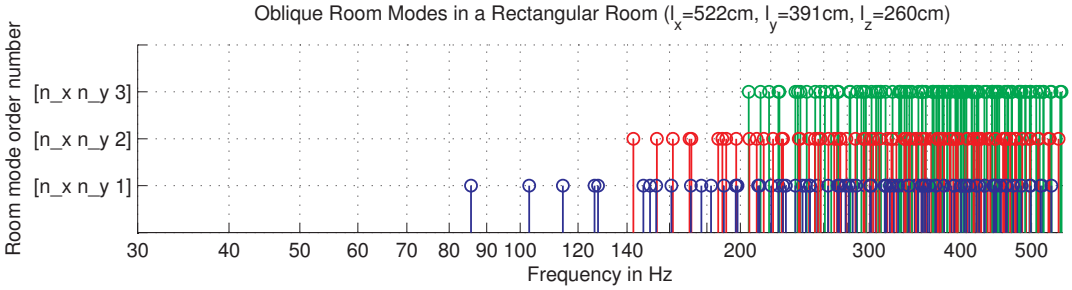

Figure 3.8.: Oblique room modes in a small listening room.

to even higher frequency regions. As a result an increased mode-density can be observed and the influence on the loudspeaker sound reproduction can be assumed to be very low. Nevertheless, isolated oblique modes can be observed for frequencies below approximately $140Hz$. These isolated modes might superpose with axial or tangential modes with the same frequency.

A **frequency response measurement** of a pillar loudspeaker (Visaton VOX200) in a typical living-room is performed. The pillar loudspeaker is positioned close to a wall in the listening-room (this equals a 2π steradian radiation – see figure 3.2). The measurement microphone is placed in three different positions within the room. The measurement process and the basic conditions of the measurement are discussed in more detail in chapter 3.2.4.

[7] In the case of isolated modes no other modes are in the near vicinity. Such modes can be superposed by further modes with the same frequency and are known as multiple isolated modes. Hence multiple isolated modes have got an increased impact on the loudspeaker sound reproduction at certain listening positions.

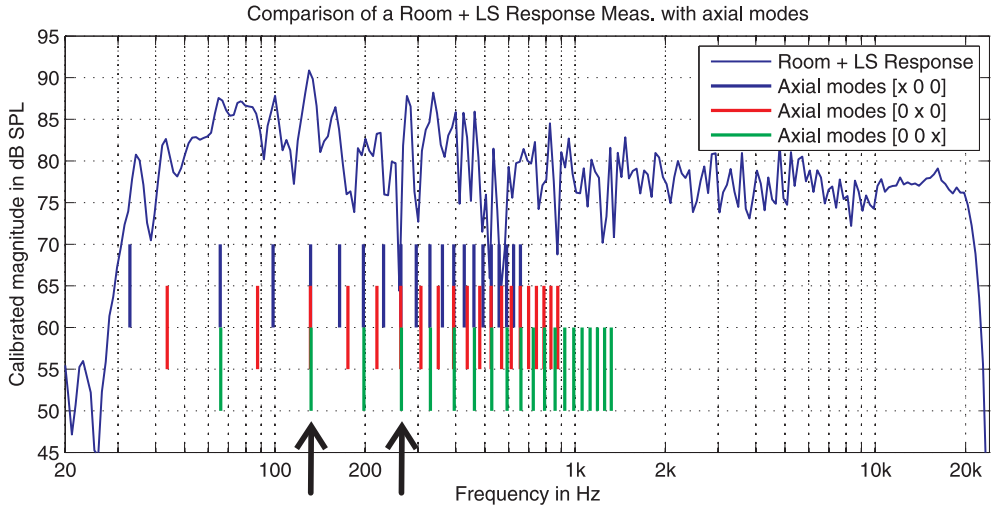

Figure 3.9.: Room/Loudspeaker Response and Axial Modes (Position 1: on axis).

In figure 3.9 the frequency response measurement of the pillar loudspeaker is shown in conjunction with the illustration of possible axial modes (see figure 3.6). The measurement microphone is placed on-axis with the loudspeaker baffle.

Certain correlations can be observed between the illustrated axial modes and the frequency response. E.g. at a frequency of approximately $130Hz$ a high peak in the frequency response can be observed. This peak might be caused by three isolated axial modes which superpose at certain room positions.

At a frequency of approximately $260Hz$ a large dip can be seen in the frequency response. The microphone could have been in a node of the standing wave, which appears at this frequency.

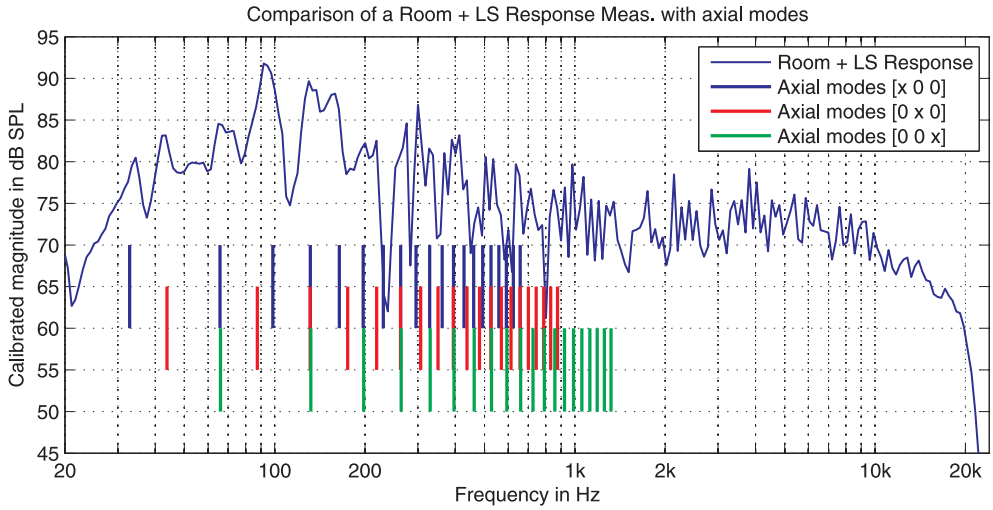

Figure 3.10.: Room/Loudspeaker Response and Axial Modes (Position 2).

In figure 3.10 and 3.11 the frequency response measurement of the same pillar loudspeaker

is again illustrated, but the measurement microphone is placed off-axis with the loudspeaker baffle and the microphone positions are also different from each other. Due to the changes of

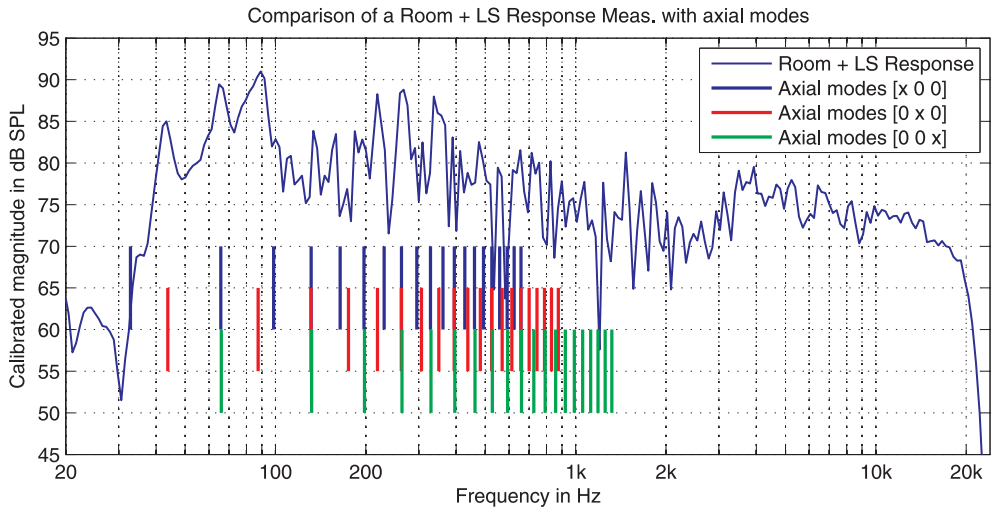

Figure 3.11.: Room/Loudspeaker Response and Axial Modes (Position 3).

the measurement positions significant differences can also be observed between the frequency response measurements.

However, the influence of certain isolated room modes can again be observed, but in fact at different frequencies.

Conclusion:
The three presented frequency responses (figure 3.9, 3.10 and 3.11) are very different from each other.

At lower frequencies the differences are caused mainly by the room modes. The impact on the frequency response is dependent on the position in the room due to standing wave characteristics [Too08]. Thus an equalization of modal peaks in the frequency response is only possible for one single position within the room (e.g. the listening sweet spot).

At higher frequencies the differences are caused by the increasing directivity behavior of the loudspeaker [Dic05, Too08].

4. Improving the Loudspeaker Sound Reproduction in Small Rooms

The main intention in this thesis is the improvement of the loudspeaker sound reproduction in small listening rooms. One possibility is the usage of a digital equalizer to avoid timbral influences by the room's acoustics. The advantage is a low effort, which is necessary using an equalization filter instead of other conventional acoustical optimization methods (e.g. complex absorbers).

In this chapter principles of room response equalization are discussed. Special focus is put on the generation of minimum-phase equalization filters and the basics of the target response generation. Furthermore one room equalizing approach is introduced, which provides the basis for the development of the software tool RoLoSpEQ (see part V).

4.1. Room Response Equalization with Digital Signal Processing

In more general the term *room response equalization* describes the following task:

- Equalize a certain room response, which is a combination of the loudspeaker and the room response. This equalization aims at diminishing negative room acoustical effects on the loudspeaker sound reproduction quality (see chapter 3).

Thus a room response measurement[1], which is measured with a loudspeaker in a room yields a room plus loudspeaker response in reality. This is due to the certain directivity behavior of a standard loudspeaker [Too08]. Thus in this thesis mainly the term room plus loudspeaker response is used.

Digital signal processing (DSP) is used for the task of room plus loudspeaker response equalization in this thesis. In general room plus loudspeaker response equalization aims at removing signal distortions by applying an inverse filter $H_{\text{room,inv}}$ which removes or diminishes *undesired acoustical effects*.

Two basic methods can be used for room response equalization:

1. Correcting the magnitude response of a measured acoustic response. Thus no correction of the phase response is requested [KP07].

2. The magnitude response and phase response are corrected. This task is often termed dereverberation [KP07].

If merely the magnitude response of the acoustic frequency response is affected *minimum phase equalization* is used [KP07]. If the phase response behavior shall be corrected as well, non-minimum phase equalization is preferred.

[1] Measurement methods are discussed in more detail in part III.

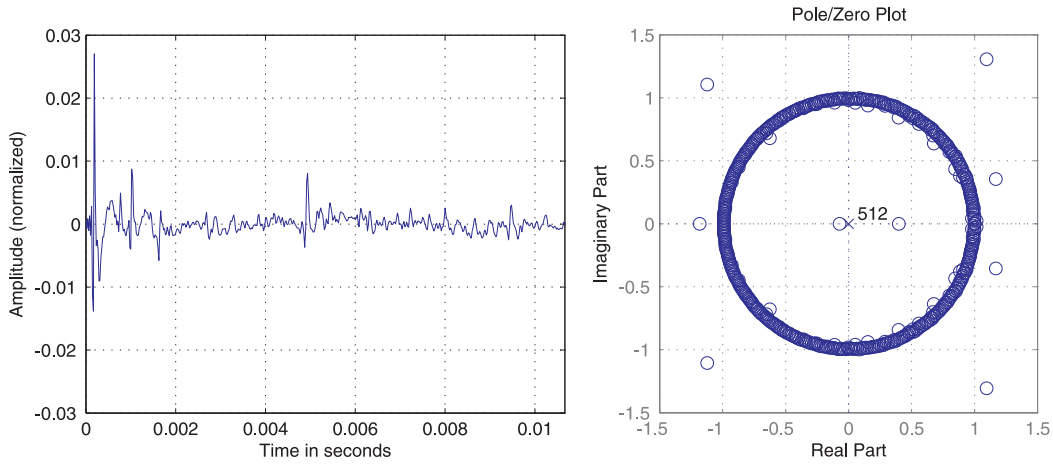

Figure 4.1.: Typical Room Impulse Response with corresponding Zero-Pole Plot.

4.1.1. Non-minimum phase behavior of Room Impulse Responses

Neely and Allen [NA79] have published a paper about the *invertibility of room impulse responses*. One important goal of this publication is the transformation of common maximum-phase impulse responses (IRs) into minimum phase IRs. This is necessary, as this process enables the possibility to invert only the minimum phase part of a maximum-phase IR. The inversion of a minimum phase system $H_{\mathrm{mp}}(z)$ always results in a *causal and stable inverse* $H_{\mathrm{mp,inv}}(z) = \frac{1}{H_{\mathrm{mp}}(z)}$, which is discussed in more detail in chapter 7.2.3.

Any finite energy time signal[2], including common room impulse response (RIR) measurements, can be characterized by the magnitude and the phase of its Fourier transform according to:

$$H(\omega) = |H(\omega)| \exp[j\phi(\omega)]. \tag{4.1}$$

In figure 4.1 a typical room impulse response with corresponding zero-pole plot is illustrated. This RIR is measured with an exponential sweep at a sampling frequency of $f_s = 48kHz$. In the figure the first 512 samples of the RIR can be observed.

A considerable number of zeros is positioned outside the unit circle which implies that it is in fact a non-minimum phase RIR. However all poles are inside the unit circle, so the signal is causal and stable (see chapter 7.2.1).

4.1.2. Minimum phase Room Impulse Responses

As proposed by Neely and Allen [NA79] the phase term of the Fourier transform of a non-minimum phase system can be subdivided into a minimum phase and a non-minimum phase part as follows:

$$\phi(\omega) = \phi_{\mathrm{mp}}(\omega) + \phi_{\mathrm{ap}}(\omega) \tag{4.2}$$

[2]To identify a energy signal two conditions have to be met [Mey09]:

- the total signal energy has to be finite, so $W = \int_{-\infty}^{\infty} x^2(t)\,dt < \infty$
- and the mean signal power has to decrease, so $P = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t)\,dt \to 0$,

where x is the time signal and T the observation time.

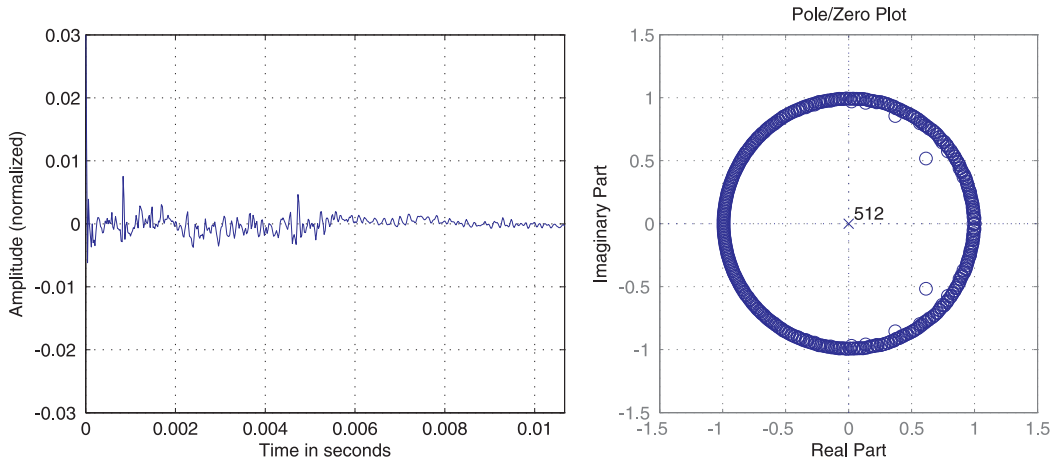

Figure 4.2.: Typical Room Impulse Response (minimum phase part) with corresponding Zero-Pole.

where index "mp" denotes the minimum phase part and "ap" denotes the allpass part, which adds the non-minimum phase behavior. This procedure is well known as "Minimum phase Allpass Decomposition" and was established by Oppenheim et al. [OSB99]. As a result $H(\omega)$ is divided into two subsystems:

$$H(\omega) = H_{\mathrm{mp}}(\omega) \cdot H_{\mathrm{ap}}(\omega). \tag{4.3}$$

The allpass system H_{ap} only applies phase modifications as the magnitude response is $|H_{\mathrm{ap}}| =$ const. [Mey09]. Thus the minimum phase system H_{mp} shows exactly the same magnitude response as the original non-minimum phase system:

$$|H(\omega)| = |H_{\mathrm{mp}}(\omega)|. \tag{4.4}$$

In figure 4.2 the RIR, which is already known from figure 4.1, is illustrated, though merely the minimum phase part is depicted.

As a result all zeros and poles are inside the unit circle, thus a stable and causal inverse signal exists.

Comment:
In order to calculate the minimum phase part of a non-minimum phase RIR, a special relation between real part and imaginary part of minimum phase systems is used for this thesis. The basis of this relation is the Hilbert transform. Further details of this procedure are discussed in more detail in chapter 9.1.1.

4.1.3. Minimum Phase Filtering

As mentioned by Karjaleinen and Paatero [KP07] minimum phase equalization will be effectual enough for basic room correction tasks. The phase response errors[3], which may be produced by *high quality loudspeakers*, are reported to be not perceivable beside common

[3]The phase response errors are the group delay deviations [KP07].

room acoustic effects (see chapter 3). Hence in general there is no need for non-minimum phase filtering aiming at a phase response correction to get an ideal constant group delay.

Minimum phase filters are also used in the software tool RoLoSpEQ (see part V). This is due to the intention of keeping influences on the native loudspeaker sound reproduction behavior as small as possible.

Another important motivation is the choice of the basic measurement procedure. Because a set of measurements is used to calculate a global RMS average aiming at a global representation of the sound reproduction behavior in the listening room, the individual phase responses of these measurements cannot be respected anymore. This is due to different phase behavior in different room positions. An averaging of the phase response might not lead to significant information.

4.1.4. What's the Target Frequency Response?

The creation of an adequate target frequency response is a major challenge of a successful room plus loudspeaker response equalization. The target frequency response is the desired frequency response after applying a specific equalizer.

Basically two approaches can be used to generate a suitable target frequency response:

1. The target response is an *ideal system response*, which is flat with a constant group delay. Therefore the ideal target impulse response is an ideal Dirac function.

   In order to avoid an overload of the audio equipment (especially of used loudspeakers) a roll-off at lower frequencies can be inserted to consider the loudspeaker high-pass behavior.

2. Or the target response aims at minimizing impacts of the listening room acoustics, hence the influences on a certain loudspeaker frequency response is minimized as well. This effect is welcome in order to preserve the timbral specifications of the used loudspeakers.

   A certain prediction of acoustical impacts on the loudspeaker sound reproduction can be achieved e.g. by the measurement of adjacent boundary effects, which cause effectual distortions if the loudspeakers are placed in the vicinity of room boundaries (see chapter 3.2.4). Up from this measurement a target response can be estimated (e.g. see chapter 21), either manually or automatically.

During this thesis it has turned out that a fixed target response definition does not yield satisfactory room equalization results. Therefore the second approach is preferred in the software tool RoLoSpEQ (see chapter 21).

4.2. Lyngdorf "RoomPerfect"

Lyngdorf audio has developed the "RoomPerfect" system [Ped03a, Ped03b, Ped06, PM07, EAP07, PT07, PHR94]. The approach aims at an improvement of the sound reproduction behavior of loudspeakers in arbitrary listening rooms. "RoomPerfect" is used in the KRK "ERGO" system as well [KS08], which is presented in chapter 4.3.

The basic ideas of the "RoomPerfect" system also provide a basis for the software tool RoLoSpEQ, which was developed during this thesis (see chapter V). Important strategies of the "RoomPerfect" system are discussed within this paragraph.

4.2.1. Measuring the Sound Reproduction Behavior – Radiation Resistance

One of the first reported strategies for measuring the influences of listening room acoustics on the loudspeaker reproduction behavior in the bass frequency range, was published by Pedersen in 2003 [Ped03a, Ped03b] and is termed "ABC" (Adaptive Bass Control). This approach is based on the fact, that the *acoustic power output of a loudspeaker* in a room highly differs from that one measured in the *free field situation*. This is also known as effect of adjacent boundaries and is discussed in mored detail in chapter 3.2.4. In order to get a general description of this effect a special measurement method was developed. With this measurement method the radiated acoustic power is derived by measuring the *radiation resistance* of the loudspeaker. This is possible, because the radiation resistance is directly proportional to the acoustic power of a loudspeaker.

The equalization filter calculation in the ABC system is based on two measurements [Ped03a]. First the radiation resistance of a certain loudspeaker model is measured in a reference room (e.g. by the loudspeaker developer). Then the measurement has to be accomplished in the listening room and further filter calculations can be done.

The radiation resistance represents the *mechanical load* caused by the ambient medium (e.g. air), which has an influence on the acoustic response of the loudspeaker chassis [Dic05, GW05]. Especially the bass performance of a loudspeaker is highly dependent on this effect.

The main advantage of this approach is the fact, that the influences of the whole listening room are considered with one measurement.

The main disadvantage is, that the measurement of one general acoustic response fails to address any local phenomena, which can be found in different listening positions [Ped07] (e.g. high peaks in the room response at one certain listening position caused by room modes – see chapter 3.2.5).

4.2.2. Measuring the Sound Reproduction Behavior – RMS

A central problem of authentic loudspeaker reproduction is the dependence of timbre on the loudspeaker placement. This effect is known as adjacent boundary effect (see chapter 3). Furthermore due to local acoustical phenomena at certain room positions (e.g. room modes – see chapter 3.2.5) it is not possible to get an accurate representation of the loudspeaker reproduction behavior by evaluating one measurement performed at one single room position. Hence with this measurement data an accurate target frequency response cannot be generated as well. However this is an approach, which is commonly used in practice [Ped03a, Ped03b, Ped06, PM07, EAP07, PT07, PHR94].

To avoid the problems caused by a single position measurement, the "RoomPerfect" approach does not merely perform one single point to point measurement of the acoustic response, as one point to point measurement cannot represent the global sound reproduction behavior of the loudspeaker in the room. Thus a few measurements at multiple (random) room positions are performed and used to derive a global representation of the loudspeaker sound reproduction behavior by calculating the root mean square (RMS) of the sound pressure in several room positions (see equation (3.1)).

Performing in maximum nine measurements in different room positions is reported to yield an accurate average of the sound reproduction behavior [Ped07].

This method provides an accurate description of the adjacent boundary effects [Too08]. Therefore this approach is also used in the software tool RoLoSpEQ for the generation of target frequency responses and room plus loudspeaker response equalizing filters (see chapter 21).

4.2.3. Filter Gain Limitations

The "RoomPerfect" filters are based on a measurement set of at least three randomly selected measurement positions within a listening room [PT07] – one measurmenet position is the listening sweet spot. Based on this measurement set, two different listening situation are considered:

1. Room response correction in the entire room.

   Merely a correction of adjacent boundary effects can be achieved. This correction filter is further termed GLOBAL room equalizer.

2. Room response correction in the listening sweet spot.

   A correction of adjacent boundary effects and standing waves (room modes) can be achieved. This correction filter is further termed FOCUS equalizer.

If a correction filter is created for the FOCUS position, the sweet spot measurement provides the basis for the filter derivation. At first a correction filter magnitude response is derived according to a suitable FOCUS target response. However, it is a problem, that single point measurements offer a very unsteady behavior due to standing wave phenomena, which can be observed in figure 4.3. Thus a non-optimized FOCUS filter might have similar characteristics.

A FOCUS equalizer can be derived e.g. with the illustrated target frequency response (red/grey curve) according to:

$$\left|H_{\text{Filter,FOCUS}}\left(e^{j\omega}\right)\right| = \frac{\left|H_{\text{FOCUS}}\left(e^{j\omega}\right)\right|}{\left|H_{\text{Target}}\left(e^{j\omega}\right)\right|}. \tag{4.5}$$

This calculation process results in the red/gray curve, which is illustrated in figure 4.4.

As expected, this non-optimized correction filter calculation results in a very unsteady behavior of the filter magnitude response as well. The major problem of this equalizing filter are very high peaks which are not practical in the case of any equalization, because high signal distortions can be induced. These high peaks are caused by standing waves, which are caused by strong dips in the room plus loudspeaker response. E.g. if the listening position

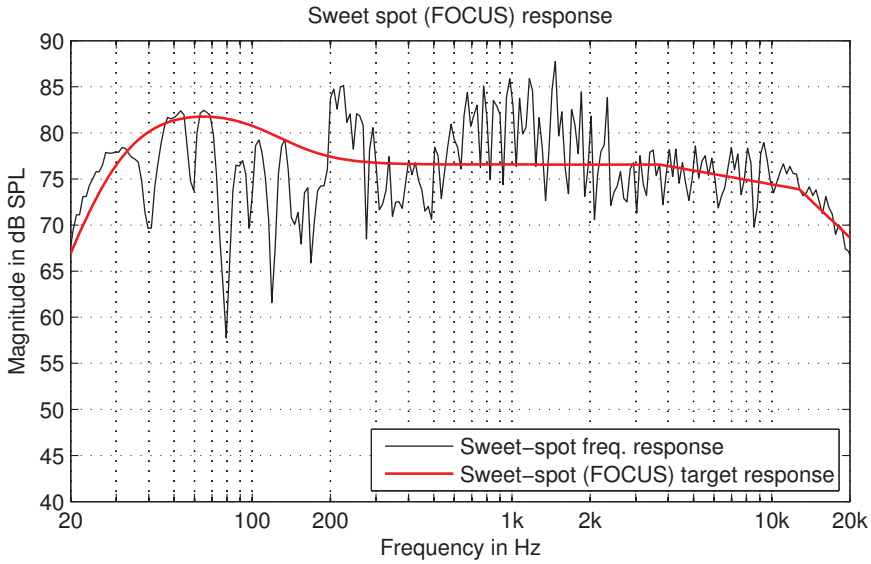

Figure 4.3.: Frequency response in the listening sweet spot (black) and desired target frequency response (red/gray) – calculated with RoLoSpEQ (see part V).

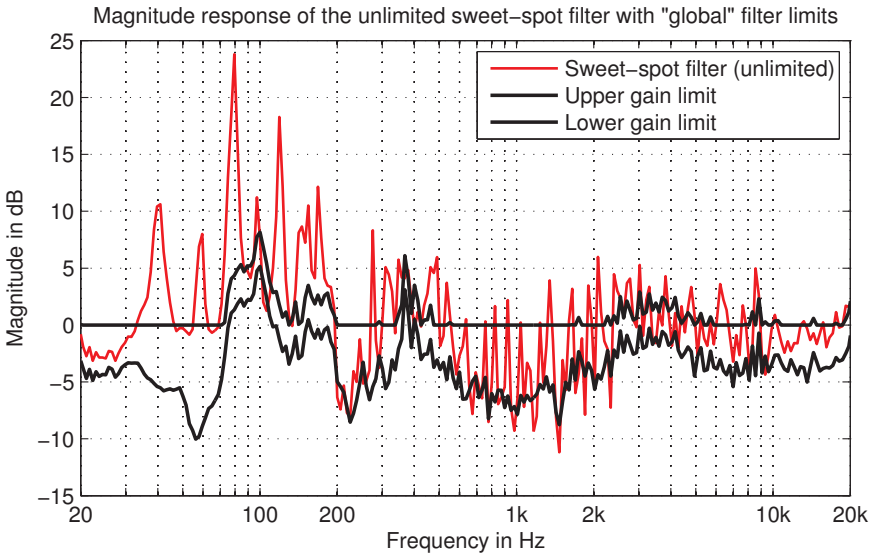

Figure 4.4.: Unlimited sweet spot (FOCUS) filter (red/gray) and gain limits calculated with the global frequency response (black).

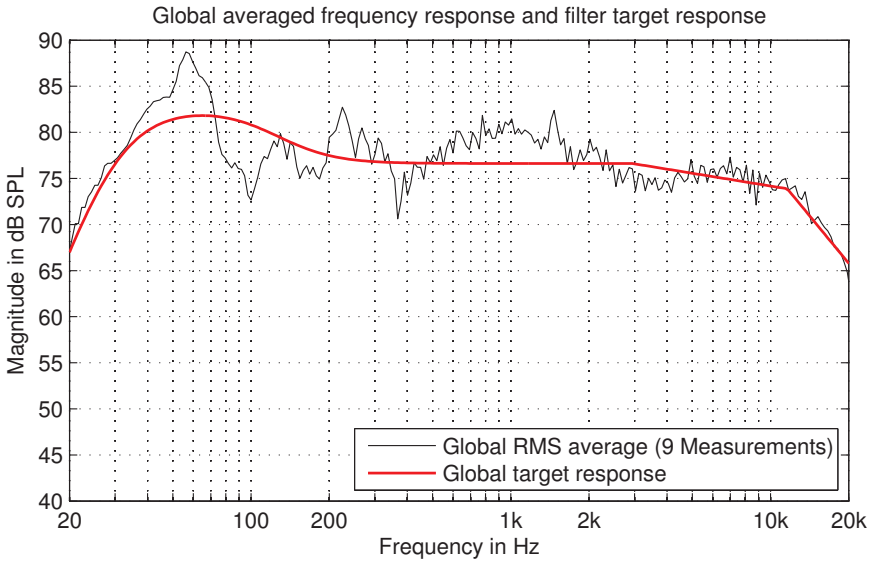

Figure 4.5.: Global averaged (RMS) frequency response (black) and desired global target frequency response (red) – calculated with RoLoSpEQ (see part V).

is in a node of a room mode the correction filter aims at an adjustment of the dip. However, these peaks and dips in the sweet spot frequency response measurement are caused by room modes, which are a characteristic of the listening room and the listening position (see chapter 3.2.5) – therefore dips in the room response cannot be avoided by an equalization. Dominant peaks, which are caused by room modes can, in a single room position, be equalized by attenuating the radiated energy at this certain peak frequency.

Two possibilities are conceivable to avoid peaks in the equalizing filter response:

1. Smoothing of the filter response.
2. Reducing the high filter gains with a gain limitation. Hence an adaption of the FOCUS equalizer to the room mode behavior is made – high peaks are detected with further measurement information.

The second approach is used in the "RoomPerfect" system and in the software tool RoLoSpEQ (see part V).

For this purpose a global average measurement is calculated and used for the generation of a global filter response according to a desired target frequency response. This global averaged frequency response and a suitable target frequency response are illustrated in figure 4.5.

The GLOBAL equalizer magnitude response (a filter for the entire room) can be derived according to:

$$\left|H_{\text{Filter,GLOBAL}}\left(e^{j\omega}\right)\right| = \frac{\left|H_{\text{RMS}}\left(e^{j\omega}\right)\right|}{\left|H_{\text{Target}}\left(e^{j\omega}\right)\right|}. \tag{4.6}$$

This GLOBAL correction filter magnitude response is used for the derivation of filter limits for an optimization of the FOCUS equalizer. The limits are illustrated in figure 4.4 as well (black curves). The derivation of the filter gain limits can be performed according to:

$$G_{\text{lower}} = \left|H_{\text{Filter,GLOBAL}}\left(e^{j\omega}\right)\right| / \sqrt{2}, \tag{4.7}$$

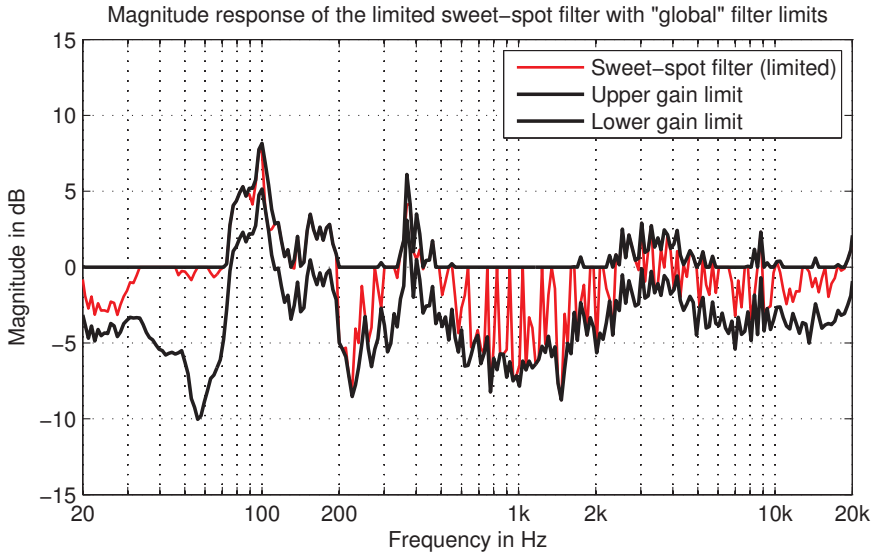

Figure 4.6.: Optimized FOCUS equalizer magnitude response (red/gray), which is limited by upper and lower gain limits (black).

$$G_{\text{upper}} = \begin{cases} 1 & \text{if } \left|H_{\text{Filter,GLOBAL}}\left(e^{j\omega}\right)\right| < 1, \\ \left|H_{\text{Filter,GLOBAL}}\left(e^{j\omega}\right)\right| & \text{if } \left|H_{\text{Filter,GLOBAL}}\left(e^{j\omega}\right)\right| \geq 1. \end{cases} \tag{4.8}$$

In figure 4.6 the optimized (limited) FOCUS filter can be seen in context with the filter gain limits.

Conclusion:

- The non-optimized FOCUS equalizing filter frequency response offers a very unsteady behavior, due to local room mode phenomena.
- The GLOBAL equalizer is effectual for the reduction of adjacent boundary effects. Thus this filter enables an improvement of the loudspeaker sound reproduction behavior at arbitrary listening positions. However local room mode phenomena are not considered.
- Local room mode phenomena can be considered by a combination of both, the non-optimized FOCUS equalizer and the GLOBAL equalizer. For this purpose gain limits are calculated for a limitation of the FOCUS equalizing filter. In the case of an attenuation in the GLOBAL equalizer, the upper gain limit is set to zero, because too much energy is assumed within the sound field (see equation (4.8)). Otherwise a maximum gain according to the GLOBAL equalizer is allowed.

  Applying these gain limits to the non-optimized FOCUS equalizer improves the unsteady behavior. The result is the optimized FOCUS equalizer. Especially high gains are avoided, as these are mainly caused by room mode phenomena.

Comment:
This approach is also realized in the software tool RoLoSpEQ and is termed “Adaptive Room Mode Detection” (ARD).

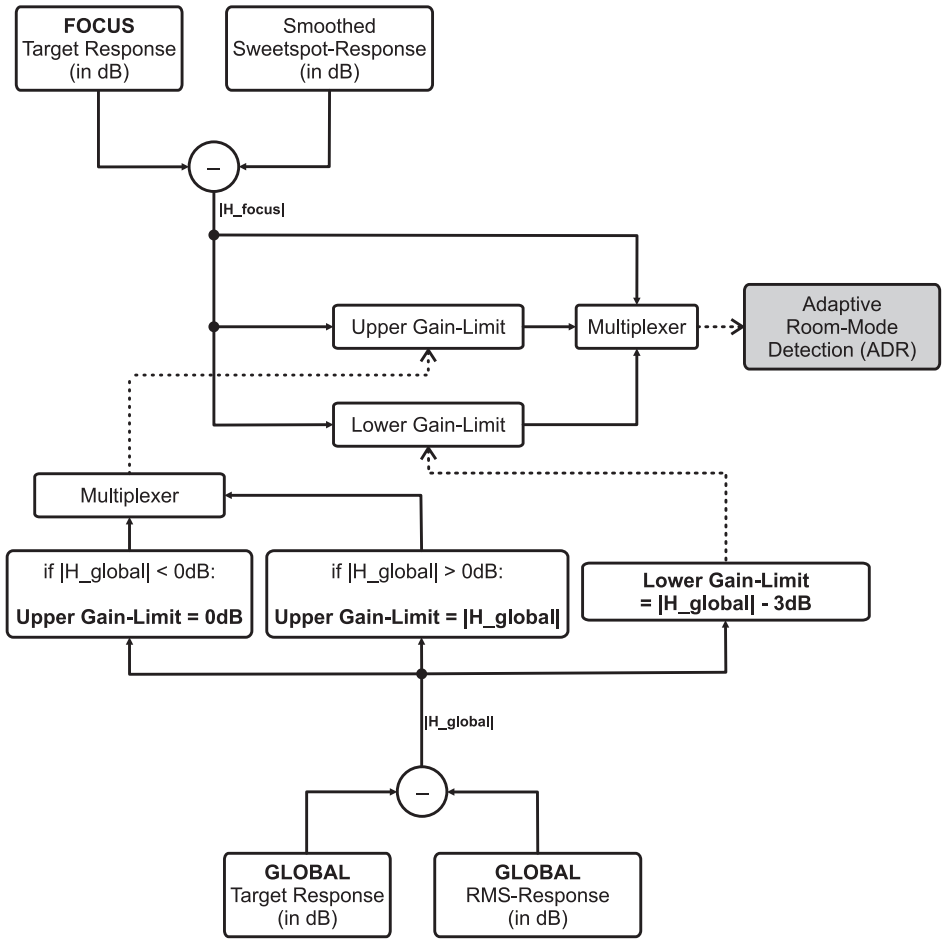

Figure 4.7.: Focus Equalizer: Adaptive Room Mode Detection (ARD).

Summarizing in figure 4.7 the discussed *FOCUS equalizer gain limitation approach* is illustrated, according to the generation procedure in RoLoSpEQ. The ARD can be activated in the *Filter Selector* in the main window of RoLoSpEQ (see chapter 20.4).

4.2.4. Preserving Natural Timbre

Pedersen et al. [PM07, EAP07] have reported a phenomenon known as "Room Gain", which appears in rooms at lower frequencies. The term "Room Gain" describes the characteristic increasing of the sound pressure level towards lower frequencies, which is also discussed in chapter 3.2.4. Thus this "Room Gain" is also a result of adjacent boundary effects (e.g. see figure 3.3).

The "Room Gain" is perceived as a natural "sound effect" by the human ear. Thus, in the "RoomPerfect" system it is considered in the automatic target frequency response derivation in order to preserve a natural timbre of the loudspeaker sound reproduction. The "Room Gain" configuration is set to a fixed value, which is derived in a standard listening room according to the criteria of IEC 60286-13 [EAP07].

A low-shelving filter with a gain of $V_0 = 6dB$ and a cut-off frequency of $f_c = 120Hz$ was

found to be the optimal solution to approximate the "Room Gain" in this standard listening room [PM07, EAP07].

Nevertheless, during this thesis it has turned out that the "Room Gain" cannot be defined to a fixed value. The "Room Gain" is also considered in the software tool RoLoSpEQ. However the configuration can be modified to the user's preferences and can be compared to other "Room Gain" adjustments, in order to find the most suitable configuration. This "natural" configuration varies enormously for different rooms, loudspeakers and placements (e.g. see figure 3.5). A subjective evaluation would be desirable in order to find a general "Room Gain" setting for different small rooms.

4.3. A Conventional Hardware Solution – KRK "Ergo"

KRK Systems offers a device called "ERGO" (**E**nhanced **R**oom **G**eometry **O**ptimization). In ERGO a digital signal processor (DSP) is used in order to apply special correction filters on an audio signal in real-time to overcome the problems of loudspeaker sound reproduction in conventional listening areas. The approach is based on the "RoomPerfect" algorithm, which was developed by Lyngdorf Audio and is discussed in more detail in chapter 4.2.

Figure 4.8.: KRK "ERGO": the base unit.

The complete system consists of three parts:

1. The base unit (see figure 4.8) includes all the hardware (DSP, connectors, switches etc.), which is necessary for the implementation in an existing audio hardware setup.

2. A calibrated microphone, which is attached one time for the measurement process.

3. A software application (for Mac and PC), which guides the user through the whole measurement process and automatically calculates the filter configuration which can be saved in the base unit finally. The filters can not be manipulated or seen by the user.

4.3.1. Hardware Specifications

Hardware specifics are available from the ERGO user manual [KS08], which provides the information within this paragraph.

ERGO can be used either as a *stand-alone filter device* including a loudspeaker *volume controller*, or in addition as *Firewire Audio Interface*. Therefore four input channels and six output channels are available.

The signal processing is managed by a "Blackfin DSP" from Analog Devices, which has a clock-frequency of $400MHz$. The "RoomPerfect" algorithm is applied on measurement-sets, which are calculated with $96kHz$ sampling frequency. Though correction filters are calculated simply for a frequency band between $20 - 500Hz$. Thus merely low frequencies are affected by the filters. FIR filters are generated with a length of merely 1024 filter taps (see chapter 9).

Either two separate 2.0 stereo loudspeaker setups or one 2.1 stereo loudspeaker setup with subwoofer can be applied to the analog outputs of ERGO. Input signals can be attached either analog by the stereo input or digitally by SPDIF. The audio performance is denoted by a signal to noise ratio (SNR) of $118dB$ and total harmonic distortions with noise (THD+N) of 0.003%.

4.3.2. Methodology

ERGO measures a stereo loudspeaker setup inside the listening room. The measurement data is analyzed to get specifics about the room-loudspeaker relationship. Using the power-averaged sound pressure (RMS) in randomly chosen room positions leads to an accurate description of the global loudspeaker sound reproduction behavior. Special filters are generated to overcome problems concerning the sound reproduction quality, which are induced e.g. by the loudspeaker placement in the room (room acoustical behavior – see chapter 2 and 3).

Two different filter sets are created according to different listening positions:

1. Room Equalizing in the sweet spot (FOCUS) position, which shall include both, frequency response and phase corrections [KS08].

2. Room Equalizing for the whole room (GLOBAL), which shall affect merely the frequency response based on the power-averaged sound pressure (RMS) response.

Part II.

Theoretical Background – Equalizers

5. Basics

In audio signal processing the term *equalizer* describes the application of *audio filters*. In general filters are frequency dependent systems with certain properties. Such filters offer e.g.:

1. frequency regions where the original signal is blocked (so called stop-band)
2. frequency regions where the original signal is passed through (so called pass-band)
3. certain phase shifts, which are applied to the original signal.

In audio processing commonly used filters are mostly linear systems. Linear systems keep the law of superposition, including additivity and scalability [Mey09, OSB99]. Furthermore if the system is *time-invariant*, that is the system properties do not change over a certain time interval, a system with the following properties is received:

- *linear time-invariant* or short **LTI-system**, in the case of analog filters (continuous-time filters)
- *linear time-invariant discrete* or short **LTD-system**, in the case of digital filters (discrete-time filters).

Such filters are mostly considered as to be time invariant, though small time variances can be expected, e.g. because of component aging. Digital systems are less affected by such kind of influences.

Possible application:
Filters can be used to reach a separation of desired signal parts from undesired signal parts (e.g. any distortions). Also a splitting of different frequency regions is conceivable. This technique is broadly used in loudspeaker crossover networks to feed separate loudspeaker drivers with adequate signals (see chapter 10).

5.1. Filter types

Four different basic filter types (see figure 5.1) are important for audio filter applications and are listed below:

- **Low-pass and high-pass filters** have a pass-band either in the low frequency range (low-pass) or in the high frequency range (high-pass). A certain cut-off frequency f_c denotes the beginning of the filter slope.
- **Band-pass and band-stop filters** have a pass-band (in the case of a band-pass filter) or a stop-band (in the case of band-stop filter) in a mid frequency range. The lower slope is denoted by the lower cut-off frequency f_l and the upper slope is denoted by the upper cut-off frequency f_u. The center frequency f_c is located in the middle of f_l and

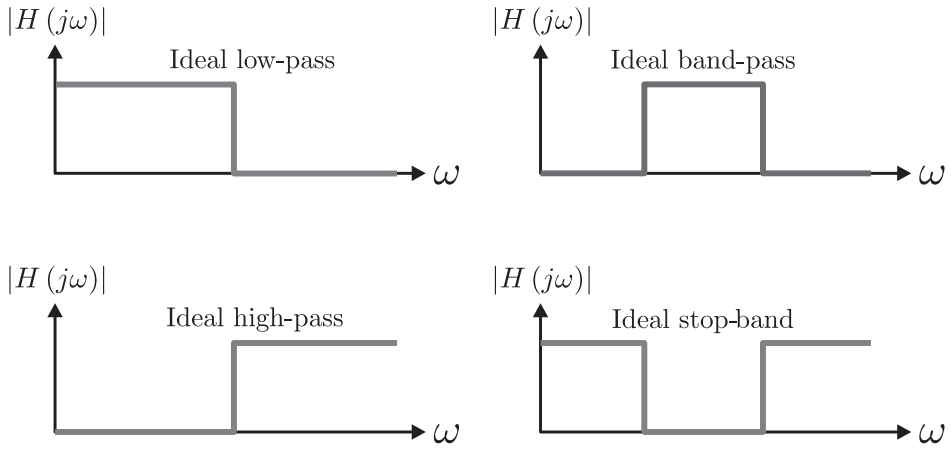

Figure 5.1.: Ideal amplitude responses (single side view) of the basic filters [Mey09].

f_u and is proportional to f_l and f_u [Zö08]. Furthermore the bandwidth can be calculated by:

$$f_b = f_u - f_l. \tag{5.1}$$

Other special filter types can be constructed with a combination of the above mentioned basic filter types. Further details and exact derivations can be found in literature [Zö08]. Special filter types are e.g.:

- **Octave filters**, which are band-pass filters with a constant bandwidth of exactly one octave.
- **Shelving filters**, which are special weighting filters [Zö08] without pass-band or stop-band. Either the lower or the higher frequency range can be amplified or attenuated. Thus for low-frequency modifications a *Low-Shelving filter* is used and respectively for high-frequency modifications a *High-Shelving filter* is used. For example a low-shelving filter is realized by a parallel connection of a direct path with a low-pass filter [Zö08]:

$$H_{\text{shelving,low}}(s) = 1 + H_{\text{lp}}. \tag{5.2}$$

- **Peak filters**, which are as well special weighting filters are used as modifier for the mid frequency range. Therefore peak filters are used to complete the characteristics of shelving filters. For example a peak filter is realized by a parallel connection of a direct path with a band-pass filter [Zö08]:

$$H_{\text{peak}}(s) = 1 + H_{\text{bp}}. \tag{5.3}$$

6. Analog Filters

Though each used filter during this thesis is designed in digital manner (see part V and VI), a short overview of analog filter design can be found within this chapter. This is due to the fact, that several digital filters depend on the theory of their analog counterparts (e.g. Butterworth approximation). Moreover analog filters are an essential part in the realization of digital networks (e.g. the anti-aliasing filter to keep the conditions of the Nyquist criterion in the case of analog to digital signal conversions) [Z08]. Further details about analog filters can be found in literature [Mey09, TS09].

6.1. System Representation

In system theory [Mey09] analog audio filters are considered as conventional LTI-systems due to their linear and time-invariant (LTI) characteristics. In general LTI-systems are described in the *Laplace transform space*. LTI-system can be described in the time domain with the following *real-valued linear differential equation*:

$$a_0 \cdot y(t) + a_1 \cdot \dot{y}(t) + a_2 \cdot \ddot{y}(t) + ... = b_0 \cdot x(t) + b_1 \cdot \dot{x}(t) + b_2 \cdot \ddot{x}(t) \tag{6.1}$$

Applying the *Laplace transform*, equation (6.1) can be transformed to a *complex-valued algebraic equation*:

$$a_0 \cdot Y(s) + a_1 \cdot s \cdot Y(s) + a_2 \cdot s^2 \cdot Y(s) + ... = b_0 \cdot X(s) + b_1 \cdot s \cdot X(s) + b_2 \cdot s^2 \cdot X(s) + ... \tag{6.2}$$

Using further transpositions the algebraic equation results in a rational function. This rational function is an adequate system description in the Laplace transform space[1]:

$$H(s) = \frac{Y(s)}{X(s)} = \frac{b_0 + b_1 \cdot s + b_2 \cdot s^2 + ...}{a_0 + a_1 \cdot s + a_2 \cdot s^2 + ...} = \frac{\sum_{i=0}^{m} b_i \cdot s^i}{\sum_{i=0}^{n} a_i \cdot s^i}. \tag{6.3}$$

Equation (6.3) is the standard representation of a LTI-system. With the method of factorization another representation with separated *zeros and poles* can be achieved:

$$H(s) = \frac{b_m}{a_n} \cdot \frac{(s - s_{N_1})(s - s_{N_2})...(s - s_{N_m})}{(s - s_{P_1})(s - s_{P_2})...(s - s_{P_n})} = \frac{b_m}{a_n} \cdot \frac{\prod_{i=0}^{m} (s - s_{N_i})}{\prod_{i=0}^{n} (s - s_{P_i})}. \tag{6.4}$$

[1] Replacing the factor s with $j\omega$ in equation (6.3) results in the corresponding *Fourier transform* of the LTI-system, in the case of a stable system [Mey09]. Hence the amplitude spectrum $|H(j\omega)|$ can be directly derived via $H(s)$.

The degree n of the denominator of $H(s)$ denotes the corresponding *system order*. In the case of audio-filters the system order is called *filter order* instead.

Without derivation:

- The higher the *filter order* n, the more precise the approximation of the ideal filter (see chapter 6.4). Therefore the filter slope is dependent on the filter order n.
- The higher the filter order n, the higher the realization effort.

6.2. All-pole-filter

Filters (LTI-systems) with a filter transfer function containing no zeros are called *all-pole-filters*[2] [Mey09]. Hence the basic filter description of an all-pole-filter in the Laplace-transform space is:

$$H(s) = \frac{b_0}{a_0 + a_1 s + a_2 s^2 + ... + a_n s^n}. \tag{6.5}$$

All-pole-filters are of special interest, because several low-pass approximations are based on an all-pole-structure (see chapter 6.4). These approximations are known as *Butterworth*, *Tschebyscheff I*, and *Bessel approximation*.

A typical *pole-zero plot* of a second order low-pass filter is illustrated in figure 6.1, where the all-pole behavior can be seen by the exclusive presence of two conjugate-complex poles. This

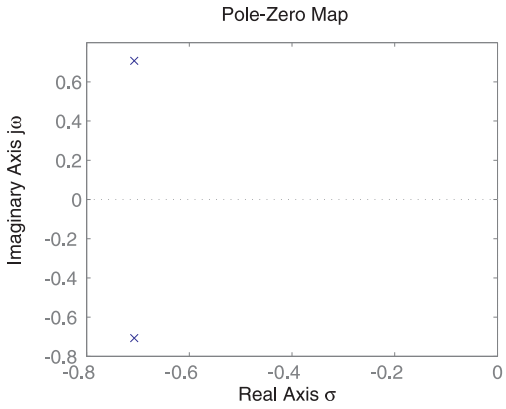

Figure 6.1.: Pole-zero plot of a second order (Butterworth) low-pass in the "s-plane".

is the pole-zero plot of a second order Butterworth low-pass filter with a cut-off frequency of $\omega_c = 1\ [rad/s]$. The cut-off frequency is chosen, as it yields a *normalized* low-pass filter, which can be used for further filter calculations in the following paragraphs (see e.g. chapter 6.5).

The corresponding transfer function is:

$$H_{\mathrm{lp}}(s) = \frac{1}{1 + \sqrt{2}s + s^2}. \tag{6.6}$$

[2] In german literature this type of transfer function is also known as "Polynomfilter".

6.3. System Stability

For stable LTI-Systems, each pole of the system transfer function has to be positioned in the left half-plane in the *complex s-plane*[3], hence $\mathrm{Re}\{\underline{P}_i\} < 0$. Such polynomials are also known as *Hurwitz polynomials* [Mey09].

6.4. Approximation of the Ideal Low-pass

The ideal low-pass filter provides the basis in designing every other idealized basic filter type (e.g. ideal high-pass filter). The general filter calculation method is discussed in chapter 6.5. But why filter-approximations are used, if an ideal solution seems to exist.

To answer this question the single side amplitude spectrum of the ideal low-pass filter, shown in figure 5.1, has to be modified. Due to the character of Fourier-transform the double side character of *continuous time Fourier-spectra* has to be considered, which leads to figure 6.2. As a result of Fourier-transform's nature, a backward transformation to an equivalent

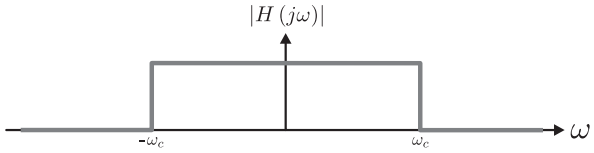

Figure 6.2.: Amplitude response of an ideal low-pass (double side view).

time signal results in the *sinc-function*:

$$h(t) = \frac{\sin(\omega_c t)}{\omega_c t} \quad \multimap \quad H(\omega) = \frac{\pi}{\omega_c} \cdot \mathrm{rect}\left(\frac{\omega}{2\omega_c}\right). \tag{6.7}$$

This sinc-function can be determined by any Fourier transformation table and describes an *acausal* continuous-time filter impulse response, which is not realizable in practice [OSB99, Mey09]. Instead, approximated versions of the ideal low-pass filter have to be used. These approximations are well defined according to different criterions [Mey09].

Common approximations are e.g. *Butterworth*, *Tschebyscheff*, *Cauer* and *Bessel approximation*. Further specifics and detailed derivations of each filter approximation can be found in literature [Mey09, OSB99, TS09].

A short example of the approximation process:
An appropriate approach is used (without derivation) containing a selectable function, which is also known as *characteristic function*. The characteristic function, in turn, is selected according to the desired filter approximation (e.g. Butterworth).

In equation (6.8) the approach can be seen [Mey09], where the polynomial F denotes the characteristic function:

$$|H(j\omega)|^2 = \frac{1}{1 + F(\omega^2)} \tag{6.8}$$

In the case of F being a polynomial, $H(s)$ will be an all-pole-filter. Otherwise if F is a rational function, $H(s)$ will be a rational function as well.

[3]The complex s-plane is the corresponding mapping plane for the Laplace transform space.

For example, to achieve a Butterworth-approximation the characteristic function has to be $F\left(\omega^2\right) = \left(\omega^n\right)^2 = \omega^{2n}$ [Mey09], where n denotes the filter order. Thus equation (6.8) is modified to:

$$|H\left(j\omega\right)| = \frac{1}{\sqrt{1+\omega^{2n}}}. \tag{6.9}$$

For the normalized cut-off frequency $\omega = \omega_c = 1$ the filter response results in $|H\left(j\omega\right)| = \frac{1}{\sqrt{2}}$. This equals to an attenuation of $20 \cdot \log_{10}\left(\frac{1}{\sqrt{2}}\right) \approx -3\ dB$, which is one of the characteristic behaviors of the Butterworth-approximation [OSB99, TS09, Mey09].

Furthermore, for $\omega \gg 1$ the filter response results in $|H\left(j\omega\right)| = \frac{1}{\sqrt{\omega^{2n}}} = \frac{1}{\omega^n}$. As a result the filter-slope equals to $-n \cdot 20 \cdot \log_{10}\left(\omega\right)\ dB = -n \cdot 20\ dB$ **per decade**, which equals to $-n \cdot 6\ dB$ **per octave**. Actually this is, because $H\left(s\right)$ is an all-pole-filter.

6.5. Filter calculation

In order to minimize the calculation effort, the same calculation process is performed for the generation of any basic filter type (see chapter 5.1). This calculation process is as follows [Mey09]:

1. calculating a normalized (cut-off frequency $\omega_c = 1$) low-pass filter (e.g. equation (6.6))
2. denormalizing the low-pass filter to achieve a selectable cut-off frequency
3. using the frequency transformation to transform the denormalized low-pass filter into the desired filter type (high-pass, band-pass or band-stop).

6.5.1. Normalized low-pass filters

The calculation of the required low-pass filter, which fits desired specifications (e.g. filter approximation with specific filter order) is performed just one time – for one specific frequency. This specific cut-off frequency is called *normalized cut-off frequency*. Therefore the cut-off frequency is always set to $\omega_c = 1\ \ [rad/s]$.

A second order Butterworth approximation, which is one example of a normalized low-pass filter, can be seen in chapter 6.2.

Filter coefficients are required for the computation of the filter transfer function of the normalized low-pass filter with any specifications. These do not have to be calculated manually. Several normalized low-pass filters are available in filter tables which are also stored in popular software tools (e.g. MATLAB [Mat08]) by default. Such filter tables can be found in literature [TS09, Mey09] as well.

6.5.2. Denormalization

Usually the basic low-pass filter is of "normalized type", that is $\omega_c = 1\ [rad/s]$ (see chapter 6.5.1). As in practical experience various cut-off frequencies are used, the normalized low-pass filter has to be denormalized. Therefor the *Laplace varible* s is replaced by $\frac{s}{\omega_c}$ in the normalized transfer function [ZÖ8].

For example:
The transfer function of a normalized second order Butterworth low-pass filter is as follows (see chapter 6.2):

$$H_{\mathrm{lp}}(s) = \frac{1}{1 + \sqrt{2}s + s^2}. \tag{6.10}$$

The denormalized version is:

$$H_{\mathrm{lp}}(s) = \frac{1}{1 + \sqrt{2}\left(\frac{s}{\omega_c}\right) + \left(\frac{s}{\omega_c}\right)^2} = \frac{\omega_c^2}{\omega_c^2 + \sqrt{2}\omega_c s + s^2}. \tag{6.11}$$

As can be seen the cut-off frequency ω_c can be set individually in the denormalized transfer function.

6.5.3. Frequency Transformation

Basic filter types, such as high-pass filters, band-pass filters and band-stop filters are obtained by using a *Frequency transformation*. That is each filter type (see chapter 5.1) can be obtained by transforming the prepared low-pass filter (see chapter 6.5). The transformation varies depending on the filter type. Thus a *low-pass to high-pass transformation* and a *low-pass to band-pass transformation* is used for the computation of basic filter types. Basically the *Laplace variable* s is exchanged by corresponding transformation variables.

In the case of *low-pass to high-pass transformation* the Laplace variable s is exchanged by:

$$s \rightarrow \frac{\omega_c^2}{s} \tag{6.12}$$

In the case of *low-pass to band-pass transformation* the Laplace variable s is exchanged by:

$$s \rightarrow s + \frac{\omega_c^2}{s} \tag{6.13}$$

The *band-stop filter* is obtained by a low-pass to band-pass transformation, as a parallel connection of a band-pass with a band-stop filter will lead to a system with a constant frequency response $H_{\mathrm{bp}}(s) + H_{\mathrm{bs}}(s) = 1$. Thus a band-stop filter is calculated as follows:

$$H_{\mathrm{bs}}(s) = 1 - H_{\mathrm{bp}}(s). \tag{6.14}$$

Further details and exact derivations of several transformation processes can be found in literature [Mey09].

For example a second order Butterworth high-pass filter is obtained by the corresponding *denormalized* low-pass filter. Exchanging the Laplace variable s by $\frac{\omega_c^2}{s}$ in formula (6.11) leads to:

$$H_{\mathrm{hp}}(s) = \frac{\omega_c^2}{\omega_c^2 + \sqrt{2}\omega_c\left(\frac{\omega_c^2}{s}\right) + \left(\frac{\omega_c^2}{s}\right)^2} = \frac{s^2}{\omega_c^2 + \sqrt{2}\omega_c s + s^2}. \tag{6.15}$$

7. Digital Filters

In *Digital Signal Processing* (DSP) *digital systems* are used to process *digital signals*, which often are analog signals before. Instead of the term digital systems the term *digital filter* is commonly used. But in digital domain the term filter is used in more general sense than in analog domain. A digital filter can offer an arbitrary transfer-function behavior, while an analog filter always describes a frequency selective characteristic (e.g. low-pass behavior).

One important advantage of digital filters is the simple and flexible generation and modification.

Digital filters are used in the developed software tools RoLoSpEQ and D-MLCNC (see part V and VI). Thus the necessary theory, which is important in this thesis, is discussed in the following paragraphs.

7.1. System Representation

The intention of finding an adequate system description for analog filters leads to the *Laplace-transform*. The Laplace-transform offers a simple and complete description of any LTI-system (see chapter 6.1).

In digital domain – where discrete systems/signals are present – such linear time-invariant systems are known as *LTD-systems* (linear time-invariant discrete). The corresponding transform for simple and complete system descriptions for discrete systems/signals is known as *z-transform* [Mey09].

The z-transform is used to calculate transfer functions, containing poles and zeros which can lead to an adequate prediction of the resulting frequency response or even the system stability. The z-transform of a real valued discrete-time system/signal is *complex-valued*.

Furthermore the z-transform is an essential tool to describe *recursive systems* – known as IIR-systems (infinite impulse response). On the one hand using the z-transform enables the description of *sequences of infinite length* in a *closed form*. On the other hand due to the characteristics of IIR-system transfer-functions it is possible to choose the position of poles and zeros, which conforms to the analog filter behavior.

Detailed derivations of the z-transform can be found in literature [Sch08, Mey09].

7.2. LTD Systems

Any LTI-system can be described by a *linear differential equation* in the continuous time-domain (see equation (6.1)).

LTD-systems are described with a *difference equation* in the *discrete time-domain* [Mey09]

as follows:

$$y[n] + a_1 \cdot y[n-1] + a_2 \cdot y[n-2] + ... + a_M \cdot y[n-M] = ...$$
$$... = b_0 \cdot x[n] + b_1 \cdot x[n-1] + b_2 \cdot x[n-2] + ... + b_N \cdot x[n-N]$$

or in short:

$$y[n] = \sum_{i=0}^{N} b_i \cdot x[n-i] - \sum_{i=1}^{M} a_i \cdot y[n-i], \tag{7.1}$$

where y is the output and x the input of the system. The variable n is the current "time-step", which is increased gradually according to the sampling interval. $a_{1...M}$ and $b_{0...N}$ are constant weighting factors.

As can be seen from this equation, in LTD-systems the three basic operations – *addition*, *multiplication with a constant value* and *time delaying* – are used. Thus digital systems are realized with the corresponding basic signal processing operators – *adders*, *multipliers* and *delayers* [Mey09].

In accordance with LTI-systems, LTD-systems can be described either in the *discrete time-domain* (e.g. with the *difference equation* or the *impulse-response*) or in *discrete z-transform space*[1] (e.g. with the transfer-function, the frequency response or the pole-zero plot). The z-transform can be regarded as pendant of the *Laplace-transform* for discrete signals [Mey09] and leads to a conversion of the difference equation to a *rational function*. The z-transform of equation (7.1) is as follows:

$$y[n] \quad \circ\!\!-\!\!\bullet \quad Y(z) = \sum_{i=0}^{N} b_i \cdot X(z) \cdot z^{-i} - \sum_{i=1}^{M} a_i \cdot Y(z) \cdot z^{-i}. \tag{7.2}$$

Dividing Y by X the *transfer-function* of the LTD-system can be calculated by:

$$H(z) = \frac{\sum_{i=0}^{N} b_i \cdot z^{-i}}{1 + \sum_{i=1}^{M} a_i \cdot z^{-i}} = \frac{b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2} + ... + b_N \cdot z^{-N}}{1 + a_0 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} + ... + a_M \cdot z^{-M}}. \tag{7.3}$$

Further system analysis (e.g. stability analysis) can be made by determining the poles and zeros of this transfer function (see chapter 7.2.1).

Exchanging the transform variable z in equation (7.3) with $e^{j\Omega}$, the corresponding *frequency response of the LTD-system* is achieved directly:

$$H\left(e^{j\Omega}\right) = \frac{\sum_{i=0}^{N} b_i \cdot e^{-ij\Omega}}{1 + \sum_{i=1}^{M} a_i \cdot e^{-ij\Omega}}, \tag{7.4}$$

where Ω denotes the *normalized frequency* for discrete-time signals $\Omega = \omega T = \omega / f_s = 2\pi f / f_s$. T [in s] is the sampling interval, thus $f_s = 1/T$ is the sampling frequency of the discrete-time signal.

[1] As pendant of the Fourier transform for continuous-signals, another type of *Fourier transform for discrete-signals* is available, which is in short commonly known as DTFT (Discrete-Time Fourier Transform) [Mey09]. The DTFT is a continuous function.

7.2.1. System Stability

In order to derive stable LTD-systems, each pole of the system transfer function has to be positioned inside the unit-circle in the *complex z-plane*[2], hence $|\underline{P}_i| < 1$.

7.2.2. Types of LTD-systems

Two different types of LTD-systems are known [Sch08, Mey09]:

1. **Recursive LTD-systems**, which are **IIR-systems** (infinite impulse response). Current values of IIR-systems depend on a finite amount of previous input values and a finite amount of previous output values. Thus IIR-systems contain a feedback of the output to the input, that is *infinite length* of corresponding impulse-responses might be expected.

   The Following important specifics are considered:

   - IIR-systems *might get unstable* due to the feedback characteristics (see chapter 7.2.1).
   - *Digital simulation* (see chapter 8.1) enables to simulate the behavior of a corresponding analog LTI-systems.
   - The *calculation effort* of IIR-filters is less than of FIR filters due to lower filter orders.
   - *Frequency dependent phase distortions* have to be expected for any causal IIR-filter (certainly zero phase filtering can be performed with acausal IIR filtering as well [OSB99, Mey09]).

2. **Non-recursive LTD-systems**, which are known as **FIR-systems** (finite impulse response). Current values of FIR-systems just depend on a finite amount of previous input values. The impulse-response is of *finite length*.

   The following important specifics are considered:

   - FIR-systems *are always stable* due to the fact that each pole of the transfer function is positioned in the origin of the complex z-plane (see chapter 7.2.1).
   - FIR-systems can be designed with *linear phase behavior*, which implies constant phase and group delay.
   - Filter transfer functions with arbitrary shape of the corresponding frequency response can be calculated (e.g. via Hilbert-Transform, see chapter 9.1.1).

7.2.3. Minimum Phase LTD-System

LTD-systems (linear time-invariant discrete) are termed minimum phase if they are causal and stable and if there exists a causal and stable inverse [OSB99], so that:

$$H_{\mathrm{mp,inv}}(z) = \frac{1}{H_{\mathrm{mp}}(z)}. \tag{7.5}$$

[2] the corresponding mapping plane for the z-transform space

In general the condition of *stability and causality* is met if each pole of the transfer function is positioned inside the unit circle (see chapter 7.2.1). Certainly this condition does not restrict the positions of existing zeros. For several applications, which require a system inversion it can be necessary to use an additional restriction – **the inverse of the system has to be causal and stable, as well**.

For this purpose each zero and pole has to be inside the unit circle. Such systems are commonly known as *minimum phase systems*.

In minimum phase systems, the decrease in phase is less compared to non-minimum phase systems due to the special position of poles and zeros [Mey09].

As one result the signal delay is less compared to non-minimum phase systems, as well. The signal delay can be expressed by the group delay, which is the negative derivative of the phase function:

$$\tau_{\mathrm{Gr}}\left(j\omega\right) = -\frac{d\phi\left(j\omega\right)}{d\omega}. \tag{7.6}$$

8. Recursive/IIR Filters

In figure 8.1 the signal flow graph of an N_{th} order IIR-filter is shown. The structure of this

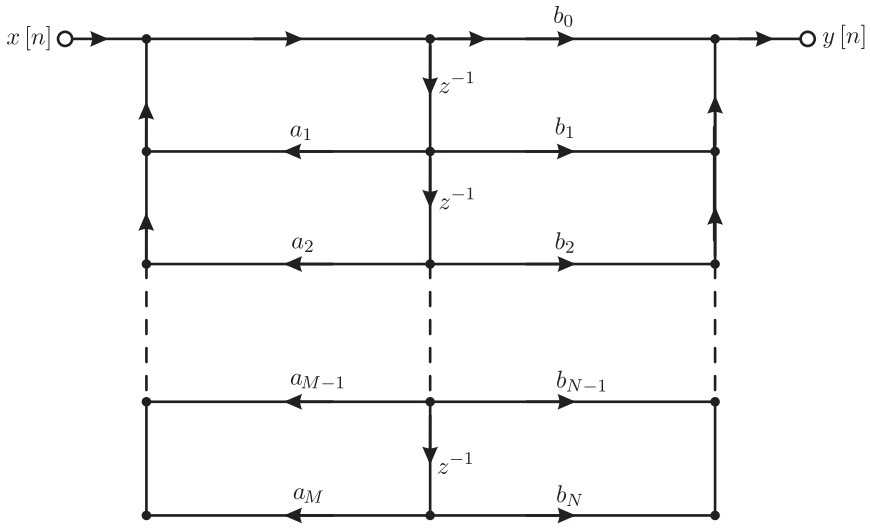

Figure 8.1.: Signal flow graph: N_{th} order IIR-filter in direct-form II realization [OSB99].

filter is termed *direct-form*[1] *II realization*, which is one possible realization of an IIR-filter. As can be seen in this figure IIR filters contain both, feedback and feedforward paths. Hence IIR filters offer a transfer function, which conforms equation (7.3). The filtered output signal $y\,[n]$ can be directly derived in the discrete time-domain according to equation (7.1).

IIR filters contain feedback paths, which can be seen in figure 8.1 and offer a transfer function, which conforms equation (7.3). The free manipulation of the filter-coefficients $b_{0\ldots N}$ and $a_{1\ldots M}$ enables independent positioning of poles and zeros (this property is only supported by IIR-filters).

In contrast, analog filter transfer-functions contain both, poles and zeros. In special cases (e.g. All-pole-filter – see chapter 6.2) merely poles are existing.

Hence IIR-filters are a suitable choice for the *digital simulation* of analog filters (see chapter 8.1).

Another important filter generation possibility which should be mentioned in this context is the direct *filter approximation in the z-domain*. The approximation process is based on special algorithms, which carry out an independent variation of the filter coefficients and the filter order, to receive a predefined system behavior. Most famous algorithms are *Fletcher-Powell* and *Yule-Walker* [Mey09].

[1] Several direct-form realizations of an LTD system are a direct realization of the difference equation (7.1). Alternative filter realizations are the cascade form (see chapter 14.2.2) and the parallel form [OSB99, Sch08, Mey09].

8.1. Digital Simulation of Analog Filters

If an analog LTI-system $H_a(s)$ is exchanged by a discrete-time LTD-system $H_d(z)$ with approximately identical behavior, the term *digital simulation* is used. Digital simulation is only used in case of IIR-filters [Mey09].

In short, digital simulation is achieved by mapping the pole-zero constellation from the complex *s-plane* to the complex *z-plane*. That is, IIR-filters can be designed by a model of an analog-filter [Mey09]. For this purpose different methods are available [Mey09]:

- The *bilinear z-transform* (see chapter 8.1.1), which aims at an optimal mapping of the corresponding system frequency-response.
- The *impulse-invariant z-transform*, which aims at an exact mapping of the corresponding system impulse-response.
- The *step-invariant z-transform*, which aims at an exact mapping of the corresponding system step-response.

In the following paragraph the focus is put on the *bilinear z-transform*, as it is used for the generation of several filters within this thesis (see part IV).

The basic filter design:
The *analog filter model* is based on a *normalized reference low-pass filter*. Subsequently a *denormalization* of this filter leads to a reference low-pass with *arbitrary cut-off frequency*.

The *frequency transformation*, as an optional processing, performs a transformation of the low-pass to the desired filter type. This methodology is explained in more detail in chapter 6. But one specialty should be mentioned in the case of digital filters:
The *frequency transformation* can either be performed (before the transformation) in the analog s-domain or (after the transformation) in the z-domain [Mey09].

8.1.1. Bilinear Transform

The *bilinear transform* (also termed bilinear z-transform) aims at an almost accurate simulation of the frequency-response of a given analog filter-model.

Three important conditions have to be met [Mey09]:

1. The frequency-response of the digital system has to be *periodical* in $2\pi f_s$, as the discrete spectrum is of periodical nature. This is necessary in order to prevent aliasing effects.
2. A rational function for $H_a(s)$ should result in a rational function for $H_d(z)$.
3. A stable analog system has to be transformed to a stable digital system.

In simple terms the bilinear transform maps the corresponding points of the frequency-response of the analog filter (or analog system) $|H_a(j\omega)|$ to corresponding frequency-bins of the digital filter $\left|H_d\left(e^{j\Omega}\right)\right|$. Due to the frequency limitations for discrete-time systems[2], specific "frequency-points" of the analog frequency-response are not mapped exactly to corresponding "frequency-points" of the digital frequency-response and vice versa. Hence a

[2]Explained by the *Nyquist criterion* and the periodicity of the DTFT (Discrete-Time Fourier Transform).

warping of the frequency response occurs [OSB99, Mey09].

In the following section a *short derivation*[3] of the bilinear z-transform can be found:

The basic theory:
Basically, the z-transform, which is the basis of the bilinear z-transform, is an exact mapping of the discrete-time domain to the complex continuous z-domain [Mey09]. This can be explained with **a short derivation** of the z-transform:

The spectrum of a discrete-time signal is obtained by the DTFT [OSB99, Mey09]. The DTFT is defined as follows:

$$x[n] \quad \circ\!\!-\!\!\bullet \quad X\left(e^{j\Omega}\right) = \sum_{n=-\infty}^{\infty} x[n] \cdot e^{-jn\Omega}, \tag{8.1}$$

where Ω denotes the *normalized frequency* for discrete-time signals $\Omega = \omega T = \omega / f_s = 2\pi f / f_s$. T in $[s]$ is the sampling interval, thus $f_s = 1/T$ is the sampling frequency of the discrete-time signal. Furthermore the unit-delay is important in case of LTD-systems (see chapter 7.2).

For this process, a unit time-shift yields:

$$x[nT - T] = x[n-1] \quad \circ\!\!-\!\!\bullet \quad X\left(e^{j\Omega}\right) e^{-j\omega T}. \tag{8.2}$$

Moreover the *Laplace-transform* can be used for a *discrete-time signal* if $j\omega$ is exchanged by $s = \sigma + j\omega$, where σ is a damping factor, which is used to force convergence[4]. As the condition of absolute summability is met for any *stable signal* the damping factor can be set to $\sigma = 0$ [Mey09]. Hence the substitution $j\omega = s$ yields:

$$x[n] \quad \circ\!\!-\!\!\bullet \quad X(s) = \sum_{n=-\infty}^{\infty} x[n] \cdot e^{-nsT} \tag{8.3}$$

$$x[n-1] \quad \circ\!\!-\!\!\bullet \quad X(s) \cdot e^{-sT}. \tag{8.4}$$

In summary a description of a stable LTD-signal (see equation (8.1) and (8.2)) can be found in the Laplace-transform space (see equation (8.3) and (8.4)). Starting from there, the mapping of the s-plane to the z-plane can be seen.

Moreover a substitution of the unit-delay in equation (8.4) is introduced:

$$z = e^{sT}. \tag{8.5}$$

This is the central relationship between the s-plane and z-plane. Equation (8.5) leads to the general definition of the z-transform:

$$x[n] \quad \circ\!\!-\!\!\bullet \quad X(z) = \sum_{n=-\infty}^{\infty} x[n] \cdot z^{-n}. \tag{8.6}$$

[3] Further details and derivations are well described in literature [OSB99, Sch08, Mey09].

[4] "Absolute summability is a *sufficient* condition for the existence of a Fourier transform representation" [OSB99]

Transformation process of the bilinear transform:
An expression for the direct substitution of the variable s in any LTI-system $H_a(s)$ is based on equation (8.5). As $z = e^{sT}$ for any stable signal/system, this equation can be transformed to:

$$s = \frac{1}{T} \ln(z), \tag{8.7}$$

which represents the mapping of the s-plane to the z-plane. As equation (8.7) is a *transcendental function*[5] which implies the problem of infinite behavior, a power series is used [Mey09]:

$$a = \frac{1}{T} \ln(z) = \frac{1}{T} \cdot \left[\left(\frac{z-1}{z+1} \right) + \frac{1}{3} \left(\frac{z-1}{z+1} \right)^3 + \frac{1}{5} \left(\frac{z-1}{z+1} \right)^5 + \dots \right]. \tag{8.8}$$

The exclusive consideration of the first term in equation (8.8) is known as **bilinear z-transform**:

$$s \approx \frac{2}{T} \cdot \frac{z-1}{z+1}. \tag{8.9}$$

Thus the bilinear z-transform is a *non-linear mapping*. Without derivation the direct relation between the "analog frequency" and the "digital frequency" is [Mey09]:

$$\omega_a = \frac{2}{T} \tan \frac{\omega_d T}{2}. \tag{8.10}$$

The transformation is accomplished by the following steps:

1. pre-warping of the cut-off frequency according to equation (8.10). So if a digital low-pass should have a certain cut-off frequency ω_d, a analog model with a cut-off frequency $\omega_a = \frac{2}{T} \tan \frac{\omega_d T}{2}$ has to be transformed.
2. design of an analog filter model with desired properties.
3. bilinear transform via substitution of the variable s.

Fortunately these steps don't have to be calculated manually due to modern software solutions (e.g. MATLAB [Mat08]), where the bilinear transform is performed automatically for the process of filter design and simulation.

8.2. Filter Charateristics

As IIR-filters can be designed according to analog models, several filter characteristics known from analog filter design can also be derived (see chapter 6.4). A few common approximations are *Butterworth-*, *Tschebyscheff-*, *Cauer-* and *Bessel-approximation*.

[5] In the case of transcendental functions no analytical solutions can be found.

9. Non-recursive/FIR Filters

In figure 9.1 the signal flow graph of an N^{th} order[1] FIR filter is shown. The structure of this

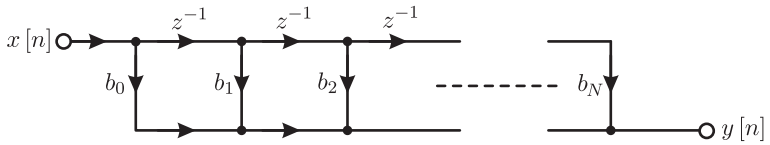

Figure 9.1.: Signal flow graph: N^{th} order FIR filter in *direct-form realization*.

filter is termed *direct-form realization* [OSB99] or rather *transversal realization*[2] for an FIR filter, which is one possible filter structure.

As can be seen in this flow graph, a FIR filter contains exclusively *feed-forward paths*. Thus merely the filter-coefficients $b_{0\ldots N}$ can be influenced. Due to the missing feedback-paths each filter-coefficient $a_{1\ldots M}$ can be assumed to be zero. Hence equation (7.3) changes to:

$$H(z) = \sum_{i=0}^{N} b_i \cdot z^{-i} = b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2} + \ldots + b_N \cdot z^{-N}. \tag{9.1}$$

Furthermore, it is obvious that the filter-coefficients $b_{0\ldots N}$ directly represent the impulse-response $h[n]$ of the FIR filter [Mey09]:

$$h = [b_0 \quad b_1 \quad b_2 \quad \ldots \quad b_{N-1} \quad b_N]. \tag{9.2}$$

Because in case of FIR-system only zero-positions can be defined (each pole is fixed in the origin of the complex z-plane), the *digital simulation* of analog systems is not possible. Therefore alternative methods are used for the FIR filter design.

One important advantage of FIR filters is the possible linear-phase behavior, which can be achieved by a causal FIR filter (for IIR-filters linear-phase behavior can be achieved merely by an acausal filtering). However linear-phase filtering is not discussed, because linear-phase filters are not considered in this thesis. According to recent literature [KP07, PT07] minimum-phase filters are assumed to be sufficient for basic acoustical equalization tasks (e.g. room plus loudspeaker response equalization).

An easy and efficient method of minimum-phase FIR filter design is based on the *Hilbert transform*, which is explained in detail in chapter 9.1.1.

[1] Instead of the filter order, the term filter taps is often used in case of FIR filters.
[2] Other filter structures (e.g. *Cascade-structure*) are well described in literature [OSB99, Sch08, Mey09]

9.1. Filter Design

FIR filters are used for *room response equalization* in the tool RoLoSpEQ (see part IV). For this purpose it is necessary to find an adequate filter design method in order to derive a minimum-phase filter, enabling any desired filter magnitude responses. As merely corrections of magnitude responses are performed minimum phase filtering is used for the equalization (see chapter 4.1.3).

Several algorithms in the software tool RoLoSpEQ merely consider the magnitude response of the measurement data. Therefore the final calculated filter response contains no phase information at all. However in the case of minimum phase systems the phase response is directly related to the magnitude response. The basis of this relationship is provided by the *Hilbert transform* and the *complex Cepstrum*, which can be used for the derivation of a minimum phase filter impulse response. This filter impulse response is directly applicable as FIR filter.

9.1.1. Hilbert Transform

The Hilbert transform defines a relationship between the real and imaginary parts of the Fourier transform of a *causal* sequence $x[n]$ [OSB99]. As a direct relationship between real and imaginary part of the Fourier transform of a *causal* sequence $x[n]$ exists, it might appear that this relationship automatically implies a relationship between the magnitude and the phase of the Fourier transform of a *causal* sequence $x[n]$ as well. However this is not true [OSB99].

Nevertheless it is possible to achieve a relationship between the magnitude and the phase with the Hilbert transform if causality is imposed to the complex cepstrum $\hat{x}[n]$, which is derived from the causal sequence $x[n]$ [OSB99]. The complex cepstrum $\hat{x}[n]$ can be derived according to:

$$\hat{x}[n] \quad \circ\!\!-\!\!\bullet \quad \hat{X}\left(e^{j\omega}\right) = \ln\left[X\left(e^{j\omega}\right)\right] = \ln\left|X\left(e^{j\omega}\right)\right| + j\arg\left[X\left(e^{j\omega}\right)\right], \tag{9.3}$$

where:

$$x[n] \quad \circ\!\!-\!\!\bullet \quad X\left(e^{j\omega}\right) = \left|X\left(e^{j\omega}\right)\right| e^{j\arg\left[X\left(e^{j\omega}\right)\right]}. \tag{9.4}$$

Due to the causality of $\hat{x}[n]$, the real and the imaginary parts of $\hat{X}\left(e^{j\omega}\right)$ are directly related with the Hilbert transform. Hence the magnitude and the phase of $\hat{X}\left(e^{j\omega}\right)$ are related, because the term $\arg\left[X\left(e^{j\omega}\right)\right]$ in equation 9.3 denotes the phase.

The further condition of a minimum phase behavior of $X(z)$ (all poles and zeros are inside the unit circle – see chapter 7.2.3) also implies causality of the complex cepstrum $\hat{x}[n]$ [OSB99].

Calculation of a minimum-phase filter IR:
According to the theory described above, the causal impulse response $h_{\mathrm{mp}}[n]$ of the magnitude response of a given minimum phase system $H_{\mathrm{mp}}(z)$ can be derived with the following steps:

1. Calculation of the complex cepstrum

$$\hat{h}[n] = \mathrm{IFFT}\left\{\ln\left(\left|H_{\mathrm{mp}}(z)\right|\right)\right\},$$

where $\left|H_{\mathrm{mp}}(z)\right|$ is the desired magnitude response of the minimum phase FIR filter.

2. As the complex cepstrum $\hat{h}[n]$ will be acausal the minimum phase condition is not met. Thus the complex cepstrum has to be zeroed for negative frequencies by setting the second half of $\hat{h}[n]$ to zero [NA79]. This has to be done according to:

$$\hat{h}_{\text{causal}}[n] = \begin{cases} \hat{h}[n], & \text{if } n = 0, N/2, \\ 2\hat{h}[n], & \text{if } 1 \leq n < N/2, \\ 0, & \text{if } N/2 < n \leq N-1. \end{cases}$$

3. Calculation of the Fourier transform of the $\hat{h}_{\text{causal}}[n]$ according to:

$$\text{FFT}\left\{\hat{h}_{\text{causal}}[n]\right\} = \ln\left[H_{\text{mp}}\left(e^{j\omega}\right)\right].$$

4. Elimination of the ln operator according to:

$$e^{\ln\left[H_{\text{mp}}\left(e^{j\omega}\right)\right]} = H_{\text{mp}}\left(e^{j\omega}\right).$$

5. Finally an IFFT leads to the desired IR of the minimum-phase FIR filter:

$$h_{\text{mp}}[n] = \text{IFFT}\left\{H_{\text{mp}}\left(e^{j\omega}\right)\right\}.$$

This calculation process is used in the software tool RoLoSpEQ and is also illustrated in figure 16.4. A detailed practical example can be found in chapter 16.2.1.

10. A Practical Application – Loudspeaker Crossover Networks

Loudspeaker crossover networks are one practical application of the basic filter theory. Furthermore a software based multichannel loudspeaker crossover network was developed during this thesis and is discussed in more detail in part VI. This software tool is called D-MLCNC (**D**igital – **M**ultichannel **L**oudspeaker **C**rossover **N**etwork **C**ontroller) and is completely implemented in MATLAB [Mat08].

Basically a loudspeaker crossover network is used in order to split an audio signal into separate frequency bands, which are suitable for the used loudspeaker drivers in a multi-way loudspeaker. E.g. in figure 10.1 a two-way loudspeaker with a simple loudspeaker crossover network is illustrated. The frequency bands are split with a low-pass filter for the woofer and with a high-pass filter for the tweeter. The intersection point of both filter frequency responses is called crossover frequency.

Beside this basic application more special filter applications are used in practice (e.g. signal delaying). Several filter types and applications which are enabled in the software tool D-MLCNC are discussed in the following paragraphs.

10.1. Analog Crossover Networks

Analog filters are often used in loudspeaker crossover networks. For this purpose two different possibilities are known to realize an analog crossover network.

On the one hand passive crossover networks can be used. Passive crossover networks make use of passive analog filters, which are directly applied in between an amplifier and the loudspeaker drivers [Dic05, GW06, Wei08]. The main disadvantage of this approach is, that the filter behavior is coupled to the load resistance of the loudspeaker driver [Dic05]. But the load resistance is dependent on frequency, thus the filter calculation can be merely performed for one specific frequency (mainly the crossover frequency). Load resistance compensation

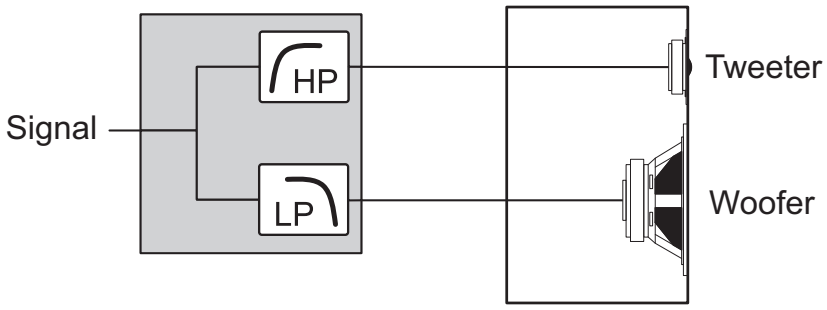

Figure 10.1.: Basic application of a loudspeaker crossover network.

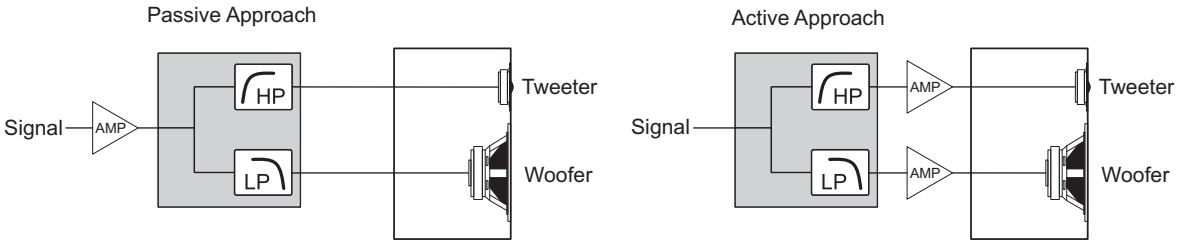

Figure 10.2.: Analog crossover networks: Passive crossover network approach (left figure) and active crossover network approach (right figure).

networks are used in practice to diminish this effect [Dic05, GW06].

On the other hand active crossover networks can be used, where active analog filters are used instead of passive analog filters. Each loudspeaker driver is supplied by one distinct amplifier. Furthermore the active analog filters are placed before the amplifiers in the signal path. Hence active filter behavior is decoupled of the load resistance of a loudspeaker driver.

Summarizing, in figure 10.2 both approaches for analog crossover networks are illustrated.

10.2. Digital Crossover Networks

Digital crossover networks are based on digital signal processing (DSP) and are often implemented directly on DSP-Hardware (e.g. dbx DriveRack[1]). The basic principle of a digital crossover network is comparable with the active analog crossover network approach. As supplementation, in digital crossover networks any conceivable filter approach can be realized using digital signal processing (e.g. adaptive filters).

The software tool D-MLCNC is a digital crossover network simulator. Because the common loudspeaker crossover filter theory is based on analog filters (see e.g. in [Dic05]) mainly IIR filters, which enable a digital simulation of analog filter counterparts are used in the software tool D-MLCNC. The main advantage of D-MLCNC is the flexibility of the filter generation, which enables a fast modification of the filters in order to get an optimization of the transmission behavior of a loudspeaker.

Summarizing, the signal flow-graph of a digital crossover network is illustrated in figure 10.3.

Important filter types and signal modifications in the case of digital crossover networks and the software tool D-MLCNC are discussed in the following paragraphs.

10.2.1. Splitting of the Audio Signal

The splitting of the audio signal aims at an optimization of the transmission behavior of a multi-way loudspeaker, where different types of loudspeaker drivers are used (see figure 10.1). As a certain loudspeaker driver (e.g. a tweeter) offers a good transmission behavior in a specific frequency region [Dic05], a signal splitting is essential in order to avoid distortions in non-optimal frequency regions.

[1]Further information can be found on the dbx webpage: http://www.dbxpro.com/4800/index.php.

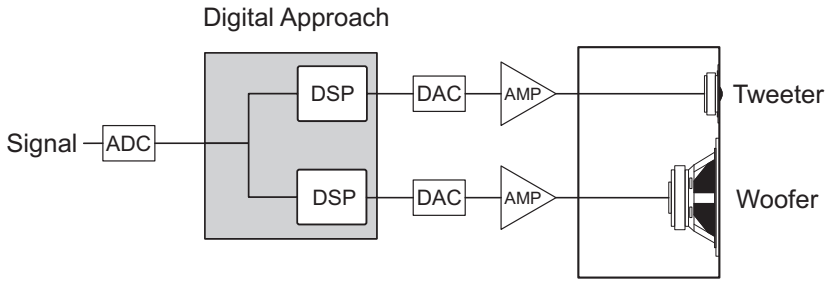

Figure 10.3.: Digital crossover network: two signal conversions are necessary to enable a digital signal processing (DSP). Certainly DSP enables a realization of any conceivable filters, with very complex specifics.

In the case of digital crossover networks the splitting of the audio signal into separate frequency bands is achieved using the basic filter types low-pass, high-pass and band-pass (see chapter 5.1). Though a band-pass filter can also be achieved by a combination of one low-pass filter and one high-pass filter. The intersection point between overlapping crossover filters is also called crossover frequency.

Basically in DSP the filters can be implemented either as IIR filters or FIR filters. However, as much research is conducted on analog loudspeaker crossover network filter design (e.g. in [Dic05]) it is evident to simulate the analog filter behavior with DSP. For this purpose in the software tool D-MLCNC a digitally simulated Butterworth approximation is used (see chapter 15.1).

10.2.2. Signal Inversion

A splitting of the audio signal into suitable frequency bands is important in multi-way loudspeakers. For this purpose high-pass and low-pass filters can be used.

However for several filters a completely different phase behavior might appear at the crossover frequency of a multi-way loudspeaker. If e.g. the phase behavior of both filters is inverted at the crossover frequency, a sound cancellation appears [Dic05, GW06]. Several second order filters offer a 180° phase relation between the high-pass and the low-pass filter at the crossover frequency [Dic05].

To avoid the sound cancellation a signal inversion is used. In DSP the signal inversion can be easily achieved by a multiplication of the input signal with -1:

$$y\,[n] = -x\,[n]\,. \tag{10.1}$$

10.2.3. Gain Controlling

Gain controlling is necessary to get a sound level adjustment of different loudspeaker drivers. In general, tweeters offer a higher efficiency in comparison to woofers [Dic05]. Hence a gain adjustment is required in order to ensure an optimized transmission behavior in a multi-way loudspeaker. In DSP the gain controlling can be easily achieved by a multiplication of the input signal with a certain constant value:

$$y\,[n] = A \cdot x\,[n]\,, \tag{10.2}$$

where $20\log_{10} A$ is the level adjustment in $[dB]$.

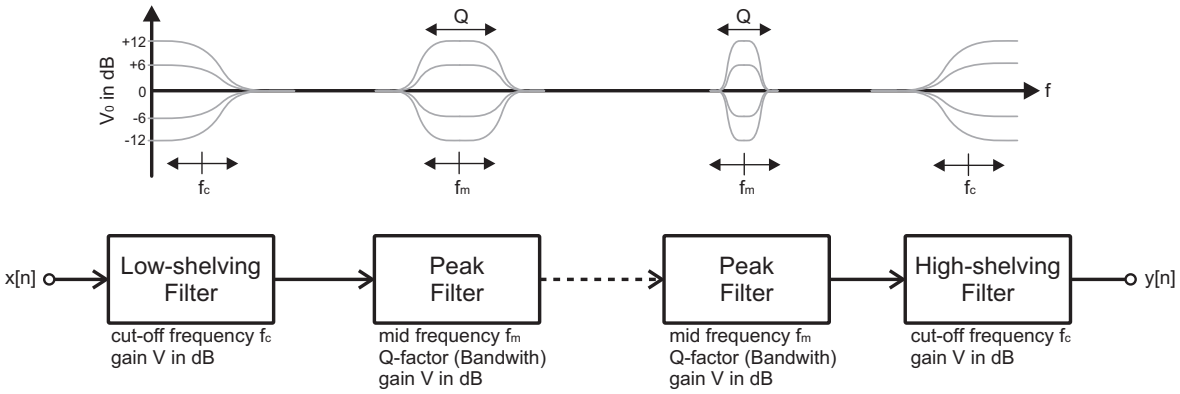

Figure 10.4.: Series connection of parametric filters for equalization [ZÖ8].

10.2.4. Signal Delaying

Signal delaying is a very important task to ensure a optimal radiation behavior of non-coincident[2] multi-way loudspeakers at the crossover frequencies of different loudspeaker drivers [Dic05].

The acoustic centers of non-coincident multi-way loudspeakers are displaced from each other and are additionally frequency dependent [Dic05]. Allocating a certain delay Δt to a signal of one loudspeaker driver yields a virtual displacement of the acoustic center. A correct alignment of the acoustic centers is very important in order to achieve an accurate operation of a crossover network at the crossover frequency (see Linkwitz Riley filter in chapter 10.2.6).

In DSP the signal delaying can be achieved by FIR filters. This approach is discussed in more detail in chapter 16.

10.2.5. Equalization

Equalization can be used for the linearization of the transmission behavior. Due to certain specifics of a loudspeaker driver the linearity of the loudspeaker frequency response is often not satisfiable. Parametric filter structures enable an accurate correction of frequency response non-linearities.

Parametric filter structures enable direct access to the parameters of a certain filter transfer function [ZÖ8]. Hence the variation of associated parameters yields a variation of e.g. the cut-off frequency of a shelving filter. The benefit is that the variation of certain filter specifics (e.g. the cut-off frequency) does not induce a recalculation of several filter coefficients.

Commonly used parametric filters are shelving filters and peak filter (see chapter 5.1).

In figure 10.4 the frequency response behavior and the adjustment possibilities of shelving and peak filters are illustrated.

[2] If coincident loudspeaker drivers are used the radiation origins are equal. Almost coincident drivers are coaxial loudspeakers, where merely a time lag appears [Dic05]. In the case of non coincident loudspeaker drivers the radiating origins are separated and lobing appears at the crossover frequency of two loudspeaker drivers (see figure 10.5).

Shelving Filters

A simple denormalized (see chapter 6.5.2) transfer function of an analog first order low-shelving filter can be achieved by a parallel connection of an all-pass and a low-pass filter, based on Zölzer [Z08] according to:

$$H_{\text{ls}}(s) = 1 + H_{\text{lp}}(s) = 1 + \frac{s + V_0\omega_c}{s + \omega_c}, \tag{10.3}$$

where ω_c is the angular cut-off frequency and V_0 is the gain in dB.

By using an all-pass decomposition in equation (10.3) and in addition the bilinear z-transform the transfer function of a digital first order low-shelving filter is achieved [Z08]:

$$H_{\text{low-shelving}}(z) = \frac{1 + \left(1 + a_{\text{B/C}}\right)\frac{H_0}{2} + \left(a_{\text{B/C}} + \left(1 + a_{\text{B/C}}\right)\frac{H_0}{2}\right)z^{-1}}{1 + a_{\text{B/C}}z^{-1}}, \tag{10.4}$$

where $H_0 = 1 - V_0$ and:

$$a_{\text{B}} = a_{\text{Boost}} = \frac{\tan(\omega_c T/2) - 1}{\tan(\omega_c T/2) + 1}, \tag{10.5}$$

$$a_{\text{C}} = a_{\text{Cut}} = \frac{\tan(\omega_c T/2) - V_0}{\tan(\omega_c T/2) + V_0}. \tag{10.6}$$

Hence the boost case ($V_0 \geq 0$) and the cut case $V_0 < 0$ have to be considered separately in order to achieve a symmetric filter response for several gain factors [Z08].

Furthermore a digital first order parametric low-shelving filter can be derived, based on Zölzer [Z08] according to:

$$H_{\text{high-shelving}}(z) = \frac{1 + \left(1 - a_{\text{B/C}}\right)\frac{H_0}{2} + \left(a_{\text{B/C}} + \left(a_{\text{B/C}} - 1\right)\frac{H_0}{2}\right)z^{-1}}{1 + a_{\text{B/C}}z^{-1}}, \tag{10.7}$$

where $H_0 = 1 - V_0$ and:

$$a_{\text{B}} = a_{\text{Boost}} = \frac{\tan(\omega_c T/2) - 1}{\tan(\omega_c T/2) + 1}, \tag{10.8}$$

$$a_{\text{C}} = a_{\text{Cut}} = \frac{V_0\tan(\omega_c T/2) - 1}{V_0\tan(\omega_c T/2) + 1}. \tag{10.9}$$

Peak Filters

Peak filters can be used to diminish certain peaks and dips in the loudspeaker frequency response, as a peak filter can be used for boosting or cutting of desired frequency regions [Z08].

A simple denormalized transfer function of an analog second order peak filter can be achieved by a parallel connection of an all-pass and a band-pass filter, based on Zölzer [Z08] according to:

$$H_{\text{peak}}(s) = 1 + H_{\text{bp}}(s) = 1 + \frac{s^2 + \frac{V_0}{Q}\omega_m s + \omega_m}{s^2 + \frac{1}{Q}\omega_m s + \omega_m}, \tag{10.10}$$

where ω_m is the angular mid frequency, V_0 is the gain in dB and $Q = \omega_m/B$ is the Q-factor, where B is the bandwidth in $[Hz]$.

By using an all-pass decomposition in equation (10.10) and additionally the bilinear z-transform the transfer function of a digital first order low-shelving filter is achieved [Z08]:

$$H_{\text{peak}}(z) = \frac{1 + \left(1 + a_{\text{B/C}}\right)\frac{H_0}{2} + d\left(1 - a_{\text{B/C}}\right)z^{-1} - \left(a_{\text{B/C}} + \left(1 + a_{\text{B/C}}\right)\frac{H_0}{2}\right)z^{-2}}{1 + d\left(1 - a_{\text{B/C}}\right)z^{-1} - a_{\text{B/C}}z^{-2}}. \quad (10.11)$$

The boost case ($V_0 \geq 0$) and the cut case $V_0 < 0$ have to be considered separately [Z08]:

$$a_{\text{B}} = a_{\text{Boost}} = \frac{\tan\left(\frac{\omega_m}{Q}T/2\right) - 1}{\tan\left(\frac{\omega_m}{Q}T/2\right) + 1}, \quad (10.12)$$

$$a_{\text{C}} = a_{\text{Cut}} = \frac{\tan\left(\frac{\omega_m}{Q}T/2\right) - V_0}{\tan\left(\frac{\omega_m}{Q}T/2\right) + V_0}. \quad (10.13)$$

Additionally the frequency parameter d and the coefficient H_0 are derived according to:

$$d = -\cos(\Omega_m), \quad (10.14)$$

$$H_0 = 1 - V_0, \quad (10.15)$$

where $\Omega_m = \omega_m \cdot T = \omega_m \cdot 1/f_s$.

10.2.6. Linkwitz-Riley Filter

The *Linkwitz-Riley filter* (L-R filter) is often preferred for the use in loudspeaker crossover networks [Dic05]. In fact the L-R filter is not a further filter approximation, but it depends on the basic Butterworth-approximation [Lin78, Dic05, Boh05].

The main intention of this filter-type is the optimization of the reproduction behavior at the crossover frequencies of multi-way loudspeakers for non-coincident loudspeaker configurations. Special attention is paid to the *directivity behavior* at the crossover frequency of different loudspeaker drivers in the vertical plane in front of the baffle. This is, because for non-coincident loudspeakers a radiation tilt can appear at the crossover frequency [Dic05] (see figure 10.5). Moreover, *interference effects* due to the displacements of the radiation origins cause problems in the directivity behavior, because at least two non-coincident loudspeakers radiate coherent signals with different phase relations. Therefore the basic filter approximations (e.g. Butterworth-approximation) may not result in a flat overall frequency response in the range of the crossover frequency.

In figure 10.5 the radiation pattern of a two-way loudspeaker with different crossover design is illustrated. A radiation tilt in vertical direction can be observed in the case of a second order Butterworth crossover network (left figure). To avoid this radiation tilt a L-R crossover network can be used instead (right figure). Hence in the case of L-R crossover networks a flat overall frequency response can be achieved on axis with the loudspeaker baffle.

The L-R filter is achieved by a serial connection of two basic Butterworth-filters. As any Butterworth low-pass filter with filter-order n shows a magnitude of $\left|H\left(e^{j\Omega}\right)\right| = -3dB$

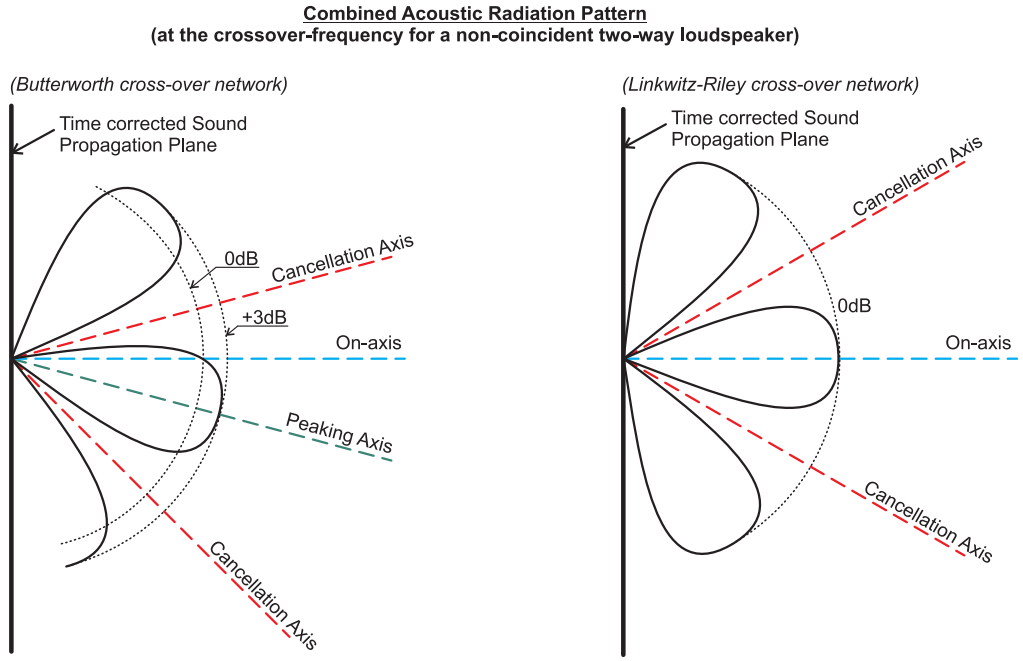

Figure 10.5.: Vertical Radiation Pattern of a Non-Coincident two-way loudspeaker with different loudspeaker crossover networks [Boh05].

and a phase-shift of $\phi = n \cdot (-45°)$ at the cut-off frequency, a serial connection of two identical Butterworth-filters results in a magnitude of $\left|H\left(e^{j\Omega}\right)\right| = -6dB$ and a phase-shift of $\phi = 2n \cdot (-45°) = n \cdot (-90°)$ at the cut-off frequency. Thus resulting Linkwitz-Riley filter order is $2n$.

As a basic principle the acoustic centers of different loudspeakers have to be equally displaced from the loudspeaker baffle in order to enable the discussed optimization behavior of L-R crossover networks [Dic05]. The adjustment of acoustic centers can be achieved with a signal delay (see chapter 10.2.4). Attention has to be paid to the frequency dependency of the acoustic center displacement [Dic05].

Part III.

Measurement Techniques

Measuring accurate transfer functions with appendant impulse responses (IR) are important tasks in this thesis.

For this purpose two different measurement methods are discussed in the following chapters namely the stepped sine method and the exponential sweep method.

The exponential sweep method is used for several measurement tasks in the developed software tools RoLoSpEQ and D-MLCNC (see part V and VI).

In the case of the exponential sweep method a non-unsubstantial disadvantage of a commonly used analysis method is revealed and a modification is reported, which leads to an improvement of the negative effects.

Furthermore a smoothing technique is introduced namely the moving average technique, which allows an accurate smoothing of measured magnitude responses.

11. Stepped Sine Method

The *stepped sine method* is still a popular measurement method for highly accurate frequency response measurements or distortion measurements [MÖ1], however this method is very time consuming. The excitation signal is a single pure sine with a specific frequency. The frequency is gradually increased to get a set of measurements at certain frequencies within the desired frequency range. Very often a logarithmic spacing of the sine frequencies is used, because a high frequency resolution of high frequencies, which is e.g. a result of the frequency linear resolution for FFT (Fast Fourier Transform) based broadband measurement methods (see chapter 12), is undesired for common audio measurement tasks [MÖ1].

The major advantage of the stepped sine method is an enormous SNR (signal to noise ratio), which can be achieved for one single measurement [SPR99, MÖ1]. The high SNR is a consequence of the low crest factor[1] of the sine signal. Hence the achievable measurement accuracy and therefore the reproducibility of the measurement for one single frequency is very high [SPR99, MÖ1].

11.1. Basic Methodology

The device under test (DUT) is excited by a single pure sine with a certain frequency. The response of the DUT (e.g. a room) is recorded simultaneously (e.g. with a omnidirectional microphone). It needs to be pointed out, that an accurate recording of the DUT is only possible if the DUT has already settled to a steady state. Hence for resonant frequencies of the DUT (e.g. at room modes – see chapter 3.2.5) a sufficient length of the excitation signals and a certain waiting time after each frequency step has to be ensured [SPR99, MÖ1].

[1] The crest factor is the ratio between the peak value and the root mean square (RMS) for any wave. In the case of a sine with amplitude $A = 1$ the crest factor is $C = 1/\sqrt{2}$ or $-3dB$.

11.2. Signal Calculation

In the case of a stepped sine the excitation signal is a pure sine:

$$y(t) = A \cdot \sin(\omega t + \phi), \tag{11.1}$$

where t is the current time in $[s]$, A is the amplitude, ω is the angular frequency and ϕ is the phase shift in rad.

As $\omega = 2\pi \cdot f$, equation (11.1) can be transposed to:

$$y(t) = A \cdot \sin(2\pi f t + \phi), \tag{11.2}$$

where f is the sine frequency in Hz.

The sine frequency has to be gradually increased. In the case of a logarithmic spacing [MÖ1] it is common to increase the frequency in steps of $1/n$ octaves. The increased frequency is used to calculate the pure sine of the next frequency step.

An octave equals a doubling of the frequency f. Thus the increasing of $1/n$ octave can be calculated according to:

$$f_2 = f_1 \cdot 2^{(1/n)}, \tag{11.3}$$

where f_1 is the lower frequency, f_2 is the increased frequency and n is the increasing step in $1/n$ octaves.

An arbitrary amount of repetitions of the gradual frequency increasing can be achieved by a multiplication with a constant integer i according to:

$$f_2 = f_1 \cdot 2^{(1/n)} \cdot 2^{(1/n)} \cdot \ldots \cdot 2^{(1/n)} = f_1 \cdot 2^{(i/n)}, \tag{11.4}$$

where i denotes the i_{th} increasing step.

11.3. Signal Analysis

Two methods for signal analysis are commonly used [MÖ1]:

1. **Analysis in the time domain:** at first the recorded response is filtered [MÖ1]. Then a rectifying of the filtered response data yields the level of the response at this certain frequency.

2. **Analysis in the frequency domain:** A FFT is performed to transform and analyze the response data in the frequency domain [MÖ1]. The major advantage of the FFT analysis is the complete suppression of any other frequencies, as the excitation frequency can be determined separately in the frequency domain. Attention has to be paid to the correct size of the FFT, because to avoid spectral leakage effects the FFT size N has to be exactly a multiple of the signal period [ZAA+08, OSB99, MÖ1].

In order to filter the response according to the first method in practice, a 1/3-octave pass-band filter might be a good choice (see chapter 5.1). This pass-band filter should have a mid frequency f_m, which equals the current excitation frequency: $f_m = f_{\text{signal}}$. This filtering should ensure the independency of the analysis of several influences (e.g. harmonic distortions or harmonic modal frequencies).

11.4. Discrete Implementation

In the case of a discrete implementation in MATLAB [Mat08] the generated sine has to be a discrete-time signal.

The first challenge is the calculation of a discrete-time sequence $t_d[n]$, which represents the continuous-time t in equation (11.2). This time sequence offers a certain length in samples according to the duration of the pure sine. The step-size Δt of the time sequence is constant and equals the sampling time. Hence the step-size Δt is dependent on the sampling frequency f_s and is calculated as follows:

$$\Delta t = T_s = \frac{1}{f_s}, \tag{11.5}$$

where Δt is the step-size in $[s]$, T_s is the sampling time in $[s]$ and f_s is the sampling frequency in $[Hz]$. Thus the discrete-time sequence will be:

$$t_d[n] = (n-1) \cdot \Delta t, \qquad \text{if } (n-1) \leq D - \Delta t, \tag{11.6}$$

where $n \geq 1$ denotes the discrete time.

Furthermore for a stepped sine measurement a start and stop frequency f_{start} and f_{stop} has to be chosen. The complete amount of possible increasing steps within the frequency range $f_{\text{start}} \leq f \leq f_{\text{stop}}$ for an increasing in $1/n$ octaves can be calculated as follows:

$$I = \log_{2^{(1/n)}} \left(\frac{f_{\text{stop}}}{f_{\text{start}}} \right) = \frac{\log\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)}{log\left(2^{1/n}\right)}. \tag{11.7}$$

Hence a number of I certain excitation frequencies can be calculated according to equation (11.4):

$$f[i] = f_{\text{start}} \cdot 2^{(i/(n-1))}, \qquad \text{if } i < I, \tag{11.8}$$

where i is an integer value an denotes the i^{th} frequency step.

The final discrete-time sine excitation signal of the i^{th} frequency step can be calculated and expressed according to:

$$y_i[n] = A \cdot \sin(2\pi f[i]\, t_d[n] + \phi). \tag{11.9}$$

In standard cases a sine signal with an amplitude level of $0dB$ ($A = 1$) and with no phase shift ($\phi = 0$) can be used:

$$y_i[n] = \sin(2\pi f[i]\, t_d[n]). \tag{11.10}$$

11.5. An Objective Evaluation

In order to evaluate the stepped sine method, a small living-room with a length of $522cm$, a width of $391cm$ and a height of $260cm$ is measured. A pillar loudspeaker (Visaton VOX200) is used for the excitation of the room. The room response is recorded with an omnidirectional microphone (DPA 4006-TL) in an arbitrary room position. The excitation level is empirically chosen to a common living-room level.

The measurement is performed in MATLAB with the following properties:

- A sampling rate of $f_s = 44.1kHz$ is defined.

- Signal length of discrete sine signals: $D = 2s$.
- The response is recorded for $1s$ after a certain waiting time of $t_{\text{wait}} = 1s$ in order to ensure a steady state sound field measurement.
- The measurement frequency region is $10 \leq f\,[i] \leq f_s/2$.
- The frequency resolution is defined by the step-size of 1/24 octave.
- Excitation signals according to equation (11.10).
- The signal analysis is performed in the time-domain according to the first method, which is discussed in chapter 11.3.

In figure 11.1 the resulting frequency response (FR) of the measurement can be seen.

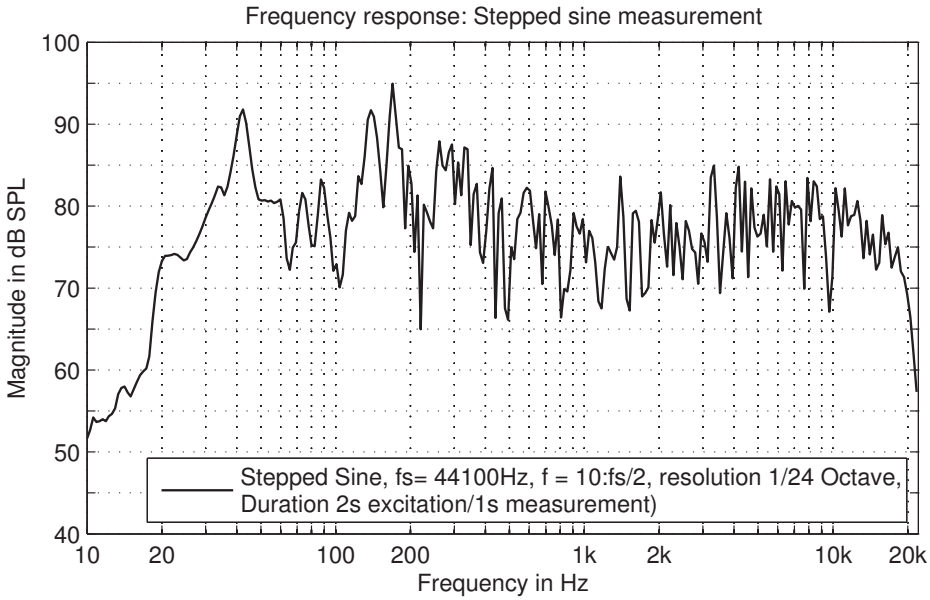

Figure 11.1.: Room + loudspeaker frequency response: Stepped sine method.

For this measurement with the step-size of 1/24 octave and $2s$ excitation signal duration, a total time of approximately $D_{\text{step}} = I \cdot 2s \approx 480s = 8$ minutes for one frequency response measurement of the whole human frequency range between $20Hz - 20kHz$ is required. Due to the reported accuracy of stepped sine measurements [MÖ1] this FR measurement provides a basis for the evaluation of a further measurement method (the swept sine method – see chapter 12).

12. Exponential Sweep

Although the FR measurement with the stepped sine method yields accurate measurement solutions (see chapter 11.5), this method is not often used in practice (e.g. for loudspeaker response measurements).

On the one hand the stepped sine method requires very much time for one measurement of the whole human frequency range between $20Hz - 20kHz$. The total measurement duration can be reduced to a fraction by using appropriate broadband measurement methods (e.g. MLS method, exponential sweep method).

On the other hand common tasks of digital signal processing often require certain discrete-time system impulse responses[1] (IR), which directly provide complete system descriptions [ZAA+08, Z08, Mey09]. Thus a FFT can be used for further analysis of the discrete system (e.g. frequency response, phase response, group delay). Certainly the stepped sine method does not directly provide the calculation of the system IR. Hence alternative measurement methods, which provide a similar measurement accuracy are commonly used (e.g. exponential sweep method) (see chapter 12.4).

All measurements within this thesis are performed with the exponential sweep method, which is discussed in the following paragraphs.

12.1. Basic Methodology

The exponential sweep method is based on a chirp signal $x(t)$. The length of this signal is defined by a duration D.

For the calculation of a system response an inverse signal $x_{\text{inv}}(t)$ has to be determined in order to meet the following condition [Z08]:

$$x(t) * x_{\text{inv}}(t) = \delta(t - D), \tag{12.1}$$

where t is the time in $[s]$, and δ is the ideal Dirac delta function. Thus the convolution of the excitation signal x with the inverse signal[2] x_{inv} yields a Dirac delta function which is delayed according to the signal duration D.

The signal x is used for the excitation of the device under test (DUT). The response of the DUT $y(t) = x(t) * h(t)$ is recorded, where $h(t)$ is the IR of the DUT. Using the inverse signal x_{inv}, a deconvolution of the recorded response $y(t)$ is enabled [Z08]:

$$y(t) * x_{\text{inv}}(t) = x(t) * h(t) * x_{\text{inv}}(t) = h(t - D), \tag{12.2}$$

which directly yields the delayed IR $h(t - D)$ of the measured system.

In practice the conditions are not that perfect. On the one hand an ideal Dirac delta function

[1] Traditional impulse response (IR) measurement methods are based on an impulsive excitation (e.g. shot of a pistol or bouncing ballons) [M01, Wei08].

[2] In the case of the exponential sweep method, the inverse signal is not the reversed signal.

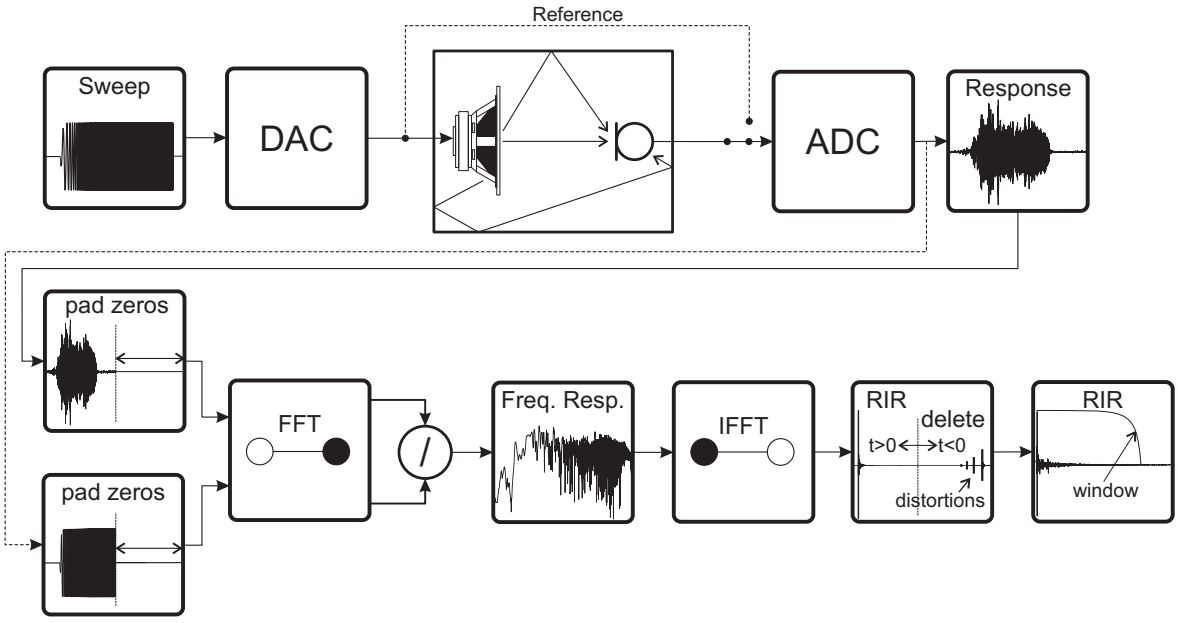

Figure 12.1.: Signal processing stages of the sweep measurement method described by Müller [MÖ1, Wei08].

can not be produced in practice due to its definition [Mey09]. On the other hand the calculation of an ideal inverse x_{inv} is another difficult challenge, which might be complicated by discrete-time signal processing theory and/or numerical inaccuracies in the practical inversion process.

In general two methods are reported for the process of signal inversion:

1. Majdak et al. [MBL07] report the signal inversion according to $X_{\text{inv}}(\omega) = \frac{X(-\omega)}{\left|X(\omega)^2\right|}$.

2. According to Oppenheim et al. [OSB99] the inverse of a discrete signal (or system) can be calculated according to $X_{\text{inv}}(j\omega) = \frac{1}{X(j\omega)}$. This method is proposed by Müller [MÖ1, Wei08] as well and is used for further calculations within this thesis.

It is not necessary to transform the inverse signals X_{inv} back to time-domain with an inverse fast Fourier transform (IFFT), if a correct circular deconvolution is performed in the frequency domain (see chapter 12.3). That is the deconvolution of the complex system FR can be performed according to:

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)}. \tag{12.3}$$

In figure 12.1 the signal processing stages of the exponential sweep method according to Müller [MÖ1] is illustrated. A discrete-time sweep with exponential increasing frequency is generated and used for the excitation of the DUT (in this case a room). The response is recorded and zero-padded to the double length of the excitation signal to perform a circular deconvolution in the following stages (see chapter 12.3). After FFT the room response is deconvolved with the reference excitation signal, which was converted twice to correct the influences of the signal path as well. The result is the FR of the room. Due to the characteristics of the exponential sweep method the FR can be "cleaned" from influences caused by any

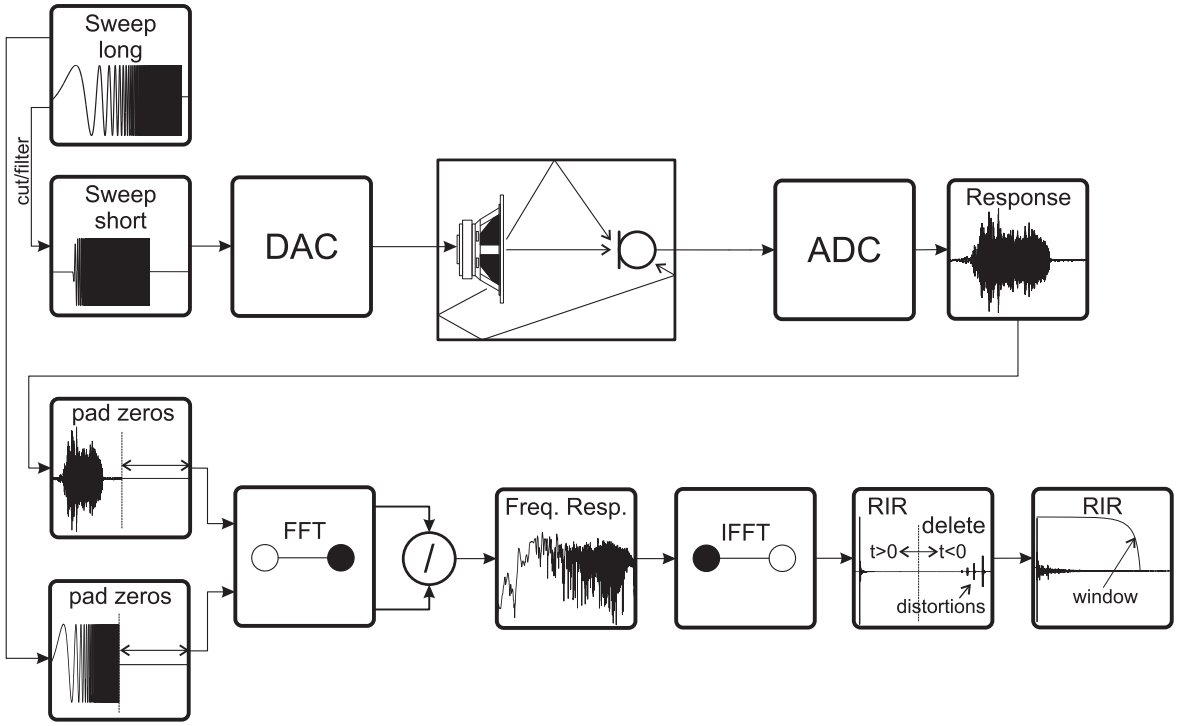

Figure 12.2.: Signal processing stages of the sweep measurement method, which was developed during this thesis.

distortions, as these can be individually separated in the room impulse response (RIR) [MÖ1]. One can see that distortions are related to negative times (see figure 12.1), however these are located on the right side after deconvolution. This is a result of the used deconvolution method. After IFFT the distortions can be deleted. The result is a clean IR measurement.

In figure 12.2 the signal processing stages of the exponential sweep method which was established during this thesis is illustrated. A basic signal processing according to figure 12.1 can be observed. The main difference of the developed method is the usage of two different signals for the excitation of the DUT and the deconvolution of the recorded response. A long sweep is calculated which has the same frequency increasing (sweep rate) as the short excitation signal, but a wider frequency range and therefore a longer duration.

If a small frequency range is used for excitation (e.g. $100Hz \leq f \leq 5000Hz$) the deconvolution according to figure 12.1 yields an increasing of non-excited frequencies outside the excitation frequency range (see figure 12.3 black curve). In contrast, if the deconvolution is performed with a long reference sweep according to figure 12.2 this problem will not appear and only excited frequencies can be observed in the FR (see figure 12.3 red/gray curve). Thus, with the developed method a detection of the excitation bandwidth is enabled without a priori knowledge.

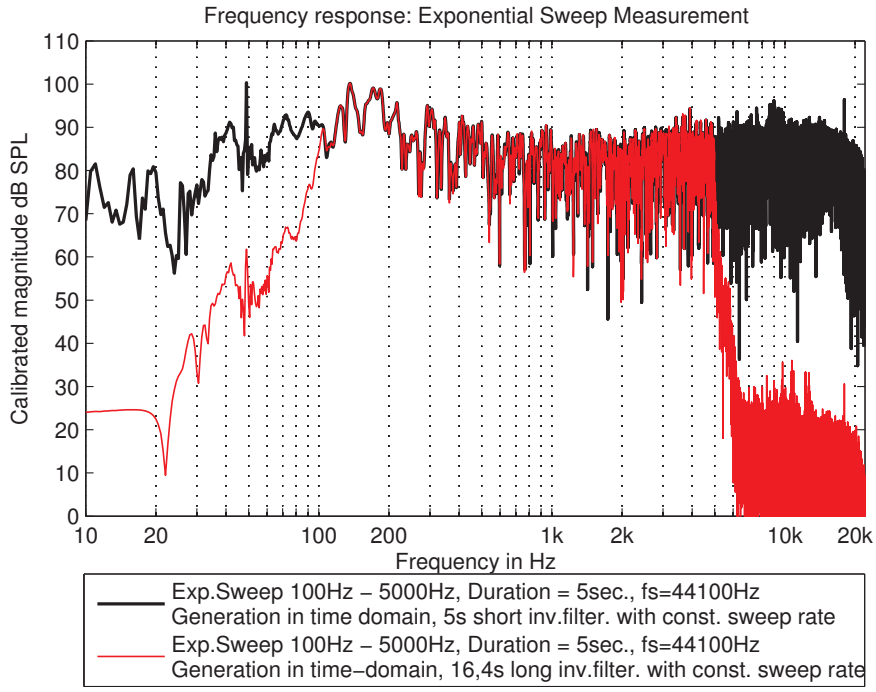

Figure 12.3.: Comparison of the method illustrated in figure 12.1 (black curve) and the method illustrated in figure 12.2 (red/gray curve), in which a long sweep is used for a deconvolution. Exponential sweep measurement: $f_{\text{start}} = 100Hz$; $f_{\text{stop}} = 5000Hz$; $D = 5s$; $f_s = 44.1kHz$.

12.2. Discrete Implementation

Exponential sine sweeps can be derived either in time-domain, or in frequency-domain [MÖ1, Wes09]. Even though both methods are implemented and evaluated during this thesis it is not necessary to use the frequency-domain calculation in the developed software tools RoLoSpEQ and D-MLCNC.

The main advantage of the time-domain approach is the simple derivation and calculation of the exponential sweep signal.

The main advantage of the frequency-domain approach is the possibility to calculate special sweeps with arbitrary magnitude spectrum [MÖ1, Wes09]. Thereby a constant temporal envelope of the excitation is still provided, as the variation of the magnitude spectrum affects only the variation of the sweep rate[3]. However the resulting sweep might not be an exponential sweep anymore, as the frequency increasing might no longer have exponential character due to the variation of the sweep rate. Even though an exponential sweep can be calculated with this method as well [MÖ1, Wes09].

[3]The sweep rate describes the increasing of the sine frequency. Hence in the case of sweeps the frequency is dependent on time.

12.2.1. Calculation in Time-Domain

The exponential sweep method is based on a special sine signal $x(t)$ with time-dependent frequency increasing in a defined frequency band $f_{\text{start}} \rightarrow f_{\text{stop}}$ which is passed through within a certain sweep length D. Thus this is a chirp signal [Z08].

In this paragraph one possible derivation of the exponential sweep signal is discussed.

The basic calculation of a discrete-time sine signal can be seen in chapter 11.2 (equation (11.1)). More general a sine signal can be described according to the following equation [Wes09]:

$$x(t) = \sin[\phi(t)], \qquad \phi(t) = \int \omega(t)\, dt, \tag{12.4}$$

where t is the time in $[s]$ and ω is a time dependent angular frequency.

In the case of an exponential discrete-time sweep the exponential frequency increasing can be derived by a general exponential function:

$$\omega(t) = k \cdot a^{\frac{t}{\tau}}, \tag{12.5}$$

where t is the current time in $[s]$, and k and τ are unknown constants, which have to be calculated. Thus two basic conditions have to be found.

As the two frequencies f_{start} and f_{stop} are known, two basic conditions can be formulated. The first basic condition for $t = 0$ yields:

$$k = \omega_{\text{start}} \tag{12.6}$$

and second basic condition for $t = D$ yields:

$$k \cdot a^{\frac{D}{\tau}} = \omega_{\text{start}} \quad \Rightarrow \quad \tau = \frac{D}{\log_a\left(\frac{\omega_{\text{stop}}}{\omega_{\text{start}}}\right)} = \frac{D}{\log_a\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)}, \tag{12.7}$$

where D is the the length of the exponential sweep in $[s]$.

The phase term $\phi(t)$ can be derived using equation 12.4. Integration of $\omega(t)$ and further transpositions yield [Wes09]:

$$\phi(t) = \frac{k \cdot \tau}{\ln a} \cdot \left[a^{\frac{t}{\tau} \cdot \log_a\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)} - 1\right]. \tag{12.8}$$

Finally the equations are combined and according to equation 12.4 the exponential sweep signal can be calculated:

$$x(t) = sin\left(2\pi \cdot \frac{f_{\text{start}} \cdot D}{\ln\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)} \cdot \left[\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)^{\frac{t}{D}} - 1\right]\right). \tag{12.9}$$

For a discrete implementation (e.g. in MATLAB [Mat08]) a discrete-time sequence $t_d[n]$ is needed, which is used in equation (12.9) instead of t. The definition of the discrete-time sequence $t_d[n]$ can be seen in chapter 11.2 (equation (11.6)). Thus equation (12.9) will be changed to:

$$x[n] = sin\left(2\pi \cdot \frac{f_{\text{start}} \cdot D}{\ln\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)} \cdot \left[\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)^{\frac{t_d[n]}{D}} - 1\right]\right). \tag{12.10}$$

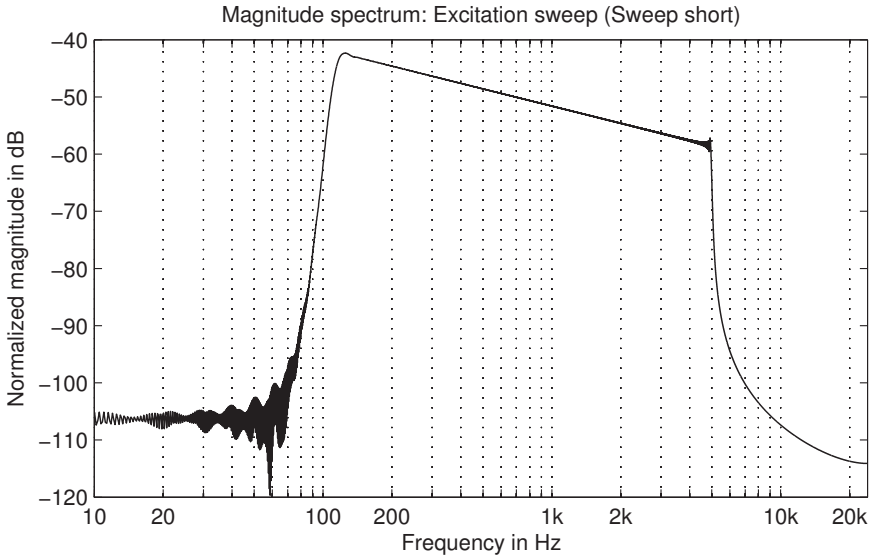

Figure 12.4.: Magnitude spectrum of the excitation sweep.

12.2.2. Derivation of the "New Method"

In chapter 12.1 the exponential sweep method by Müller [MÖ1, Wei08] and the alternative method which was developed during this thesis are introduced. The main difference of these methods is the usage of two different signals for the deconvolution of recorded responses. The developed method makes use of two separate signals – one for the excitation of the DUT and another longer one for the deconvolution of the recorded response.

The derivation of these signals is discussed within this paragraph.

An exponential sweep is calculated according to equation (12.10) with the following specifications: $f_s = 48kHz$, $f_{\text{start}} = 100Hz$, $f_{\text{stop}} = 500Hz$ and $D = 2s$. The magnitude spectrum of the excitation sweep $|X_{\text{short}}(j\omega)|$ is illustrated in figure 12.4. If a deconvolution is calculated according to equation (12.3) with the illustrated short excitation sweep, an incorrect weighting of non-excited frequency regions will appear. The result can be seen in figure 12.3 (black curve).

To avoid the incorrect weighting of non-excited frequency regions in the deconvolution process, a long sweep with identical properties is used instead. The magnitude spectrum $|X_{\text{long}}(j\omega)|$ of the long reference sweep is illustrated in figure 12.5 in context with the magnitude spectrum $|X_{\text{short}}(j\omega)|$ of the excitation sweep. If a deconvolution is calculated according to equation (12.3) with the long reference sweep, a correct weighting of non-excited frequency regions can be observed. This is also illustrated in figure 12.3 (red/gray curve).

Calculation:

In general a discrete-time sweep can be calculated according to equation (12.10). Thus the long reference sweep $x_{\text{long}}[n]$ is calculated according to this equation as well.

Due to the specifications of the excitation sweep, a long reference sweep has to be found with one identical basic condition. This characteristic basic condition is the sweep rate, which

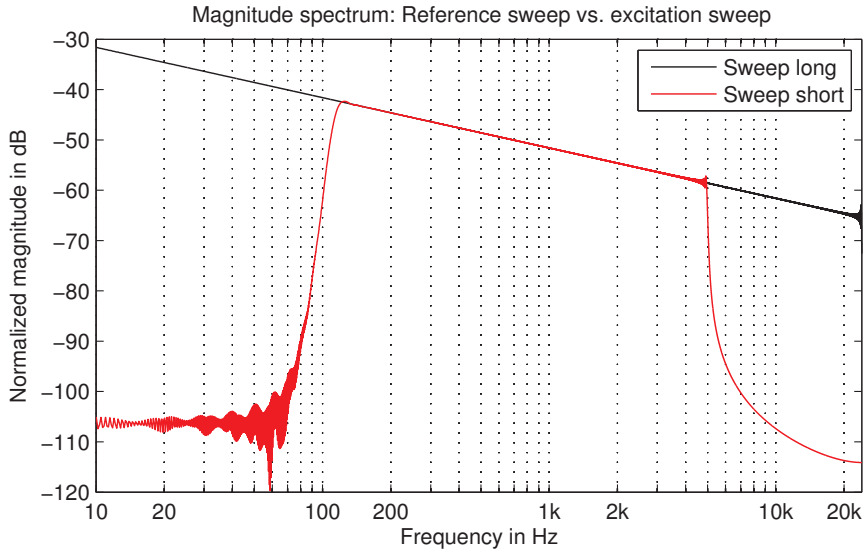

Figure 12.5.: Magnitude spectrum of the long reference sweep (black curve) in context with the short excitation sweep (red/gray curve).

describes the slope of the sweep frequency and can be calculated according to:

$$SR = \log_2\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)/D, \tag{12.11}$$

where the term $\log_2\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)$ yields the amount of certain octaves within the excitation frequency range and D is the duration. Hence the sweep rate (SR) describes the amount of octaves, which are passed through by the exponential sweep in one second.

The short excitation sweep fulfills the sweep rate, the specified frequency range and the duration.

The long reference sweep fulfills the sweep rate as well. Though another frequency range is defined, such that the long reference sweep has a frequency range beginning at the first frequency bin df and ending at the last frequency bin $f_s/2$. Therefore a new total duration D_{long} has to be found to satisfy the equation $SR_{\text{long}} = SR_{\text{short}}$ according to:

$$\log_2\left(\frac{f_s/2}{df}\right)/D_{\text{long}} = \log_2\left(\frac{f_{\text{stop}}}{f_{\text{start}}}\right)/D, \tag{12.12}$$

where $df = \frac{f_s}{N_{\text{FFT}}}$ is the resulting frequency resolution after FFT [ZAA+08, Mey09], which equals the first frequency bin in the frequency spectrum. The resulting FFT has a length of N_{FFT} samples an can be derived with $D_{\text{long}} = \frac{N_{\text{FFT}}}{fs}$.

Because no analytical solution could be found for equation (12.12), D_{long} is iteratively calculated, by a gradually increasing of the sweep length D_{long} until equation (12.12) is satisfied.

Subsequently the long reference sweep $x_{\text{long}}[n]$ can be generated with the length D_{long} and the resulting frequency $f_{\text{start}} = df = \frac{f_s}{N_{\text{FFT}}} = \frac{1}{D_{\text{long}}}$.

The excitation sweep can be generated with two methods:

- Filtering the long reference sweep $x_{\text{long}}[n]$ with a pass-band filter according to the desired frequency range.

- Zeroing or deleting the long reference sweep $x_{\text{long}}[n]$ at all frequencies below f_{start} and above f_{stop}. This is easy to perform as the current-time frequency of the reference sweep $x_{\text{long}}[n]$ can be derived. In addition a fading can be introduced using a suitable window function. This is the method which is preferred in this thesis.

12.2.3. Frequency-domain Approach

The frequency-domain approach was implemented according to Müller [MÖ1].

A performance evaluation is made in order to compare the accuracy of the time-domain and the frequency-domain approach. It has turned out that the frequency-domain approach does not yield more accurate measurement results compared to the time-domain approach.

The comparison of the measurements is illustrated in figure 12.6. Due to the small differ-

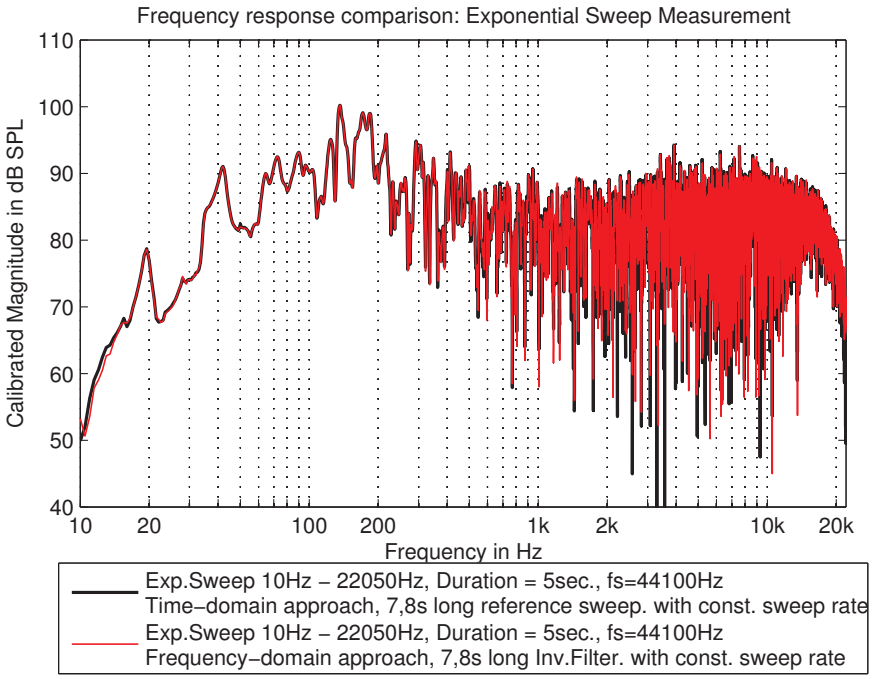

Figure 12.6.: Comparison of two FR measurements with exponential sweep excitation: Time-domain approach (black curve) vs. frequency-domain approach (red/gray curve).

ences it is obvious that the frequency domain-approach is not used in the developed software tools RoLoSpEQ and D-MLCNC.

12.3. Circular Convolution/Deconvolution via FFT

The discrete convolution is required in several task of digital signal processing (e.g. signal filtering or exponential sweep measurement). A fast convolution of two discrete sequences $x_1[n]$ and $x_2[n]$ is enabled by using the FFT [Mey09]. In the frequency domain the convolution is exchanged by a multiplication, which is performed element by element:

$$x_1[n] * x_2[n] \qquad \circ\!\!-\!\!\bullet \qquad X_1[m] \cdot X_2[m]\,. \tag{12.13}$$

Though attention has to be paid on the correct performance of this calculation in the frequency domain, as the discrete convolution performed by a multiplication in the frequency domain is always a *circular convolution* [Mey09].

Therefore the two sequences have to be zero padded in minimum to double length of the longest sequence to ensure equal results for a circular convolution in comparison to a linear (non-circular) convolution. This is due to discrete Fourier transform characteristics [Mey09].

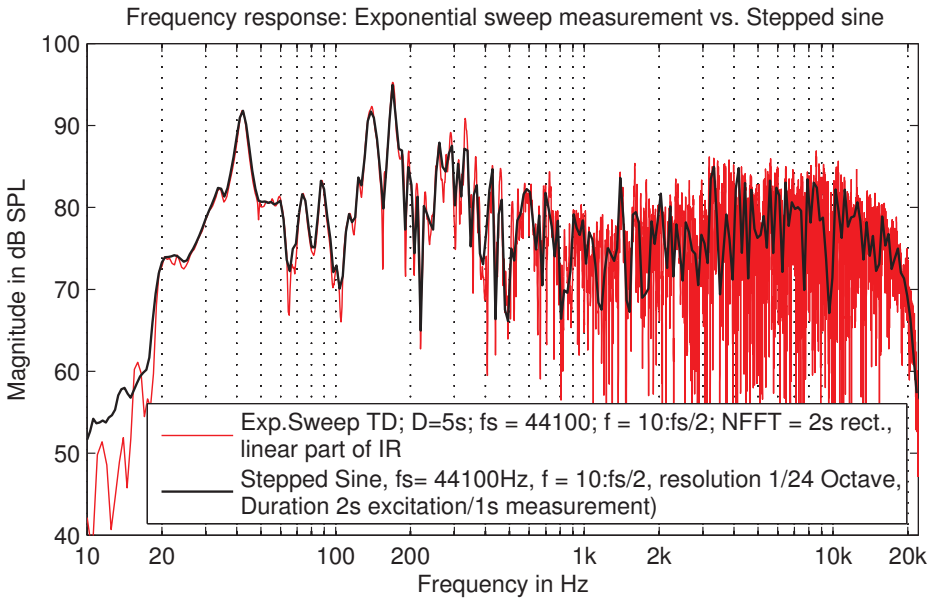

Figure 12.7.: Comparison of two frequency response measurements: Stepped sine method with 1/24 octave resolution (black curve) vs. exponential sweep method (red/gray curve – unsmoothed raw data).

12.4. An Objective Evaluation

In order to prove the accuracy of FR measurements performed with the exponential sweep measurement, the differences between an exponential sweep measurement and a stepped sine measurement is evaluated.

For this purpose a small living-room with a length of $522cm$, a width of $391cm$ and a height of $260cm$ is measured with the exponential sweep method and in addition with the stepped sine method (see chapter 11.1). A pillar loudspeaker (Visaton VOX200) is used for the excitation of the room and the response is recorded with an omnidirectional microphone (DPA 4006-TL) in an arbitrary position. The excitation level is empirical chosen to a common living-room level.

The exponential sweep is generated with the following specifications: $f_s = 44.1kHz$, $D = 5s$, $f_{\text{start}} = 10Hz$, $f_{\text{stop}} = f_s/2$. The deconvolution is performed with a long reference sweep, which is described in more detail in chapter 12.2.2. Certain distortions are deleted in the resulting IR, thus only the linear part of the IR is observed. The frequency response is derived using a $N = 2s \cdot f_s = 88200$ samples FFT.

The stepped sine is generated with a step size of 1/24 octave.

In figure 12.7 the response measurements of both methods are illustrated. A high correlation can be observed especially at lower frequencies. With increasing frequency the exponential sweep FR becomes more and more unsteady, which is certainly forced by the logarithmic frequency axis, as the *FFT bins are equidistant* [Mey09]. However, in average both measurements seem to be very similar. Thus a smoothing of the exponential sweep FR is assumed to increase the correlation.

In figure 12.8 the stepped sine measurement is kept untouched, whereas the exponential

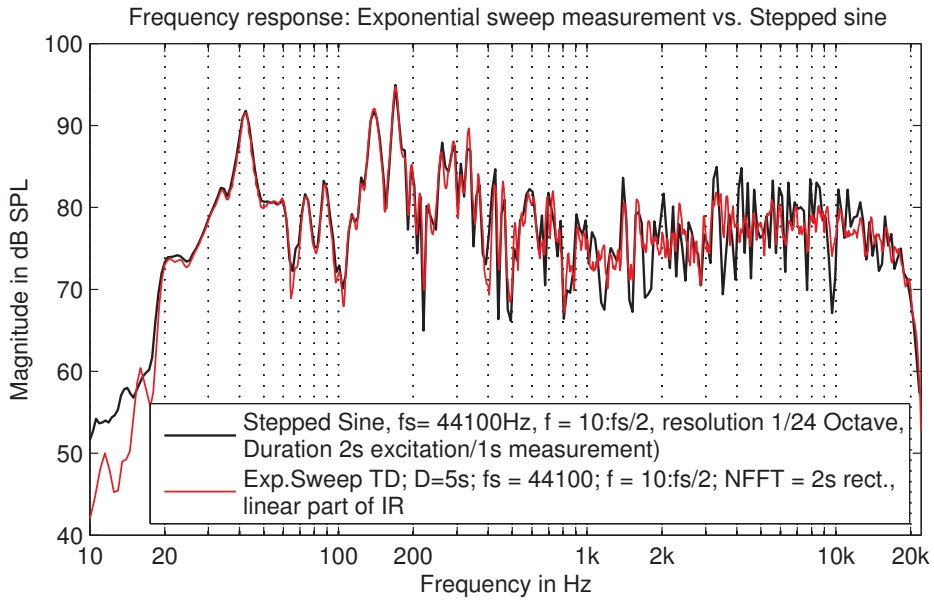

Figure 12.8.: Comparison of two frequency response measurements: Stepped sine method with 1/24 octave resolution (black curve) vs. exponential sweep method with a smoothing of 1/24 octave resolution (red/gray curve).

sweep FR is smoothed with an appropriate smoothing function. An increasing of the correlation can be observed. Overall the exponential sweep FR becomes more steady by the smoothing, because the frequency resolution is changed to a logarithmic resolution according to the stepped sine measurement characteristic. For the smoothing process a *moving average technique* is used with a frequency resolution of 1/24 octave (see chapter 13).

Conclusion:
In comparison with the stepped sine method, which is reported to be highly accurate [MÖ1], the exponential sweep seems to be also appropriate to get accurate measurements of frequency and impulse responses. Based on these facts and due to the very short measurement duration the exponential sweep measurement is used for several measurement tasks in the developed software tools RoLoSpEQ and D-MLCNC (see part V and VI).

13. Spectral Smoothing

On the one hand *spectral smoothing* enables a more clearly represented and apparently more accurate representation of FR measurements, which are performed with the exponential sweep method (see chapter 12).

On the other hand the spacing of the data points can be adapted to the logarithmic frequency spacing, which is mainly used for the FR representation.

Moreover the smoothing technique can be used to save memory, which can be important for practical implementations where memory might be limited.

13.1. The Moving Average Technique

The moving average technique is used for spectral smoothing applications of measured frequency responses in this thesis. The used approach enables a smoothing with *constant relative bandwidth*, according to the logarithmic frequency axis of frequency response (FR) measurements. The FR measurements originally offer a linear data spacing due to FFT characteristics [Mey09].

The moving average technique is a simple method for the smoothing of FFT based frequency response measurements, whereas simply the magnitude response $|H(j\omega)|$ is considered. However the performance of this technique is appropriate to smooth *room and loudspeaker responses* for the task of *room and loudspeaker equalization* (e.g. within the tool RoLoSpEQ – see part V). This arises from the fact, that a correction of phase deviations is assumed to be not relevant in practice (see chapter 4.1.3). Thus several filtering tasks are performed with minimum-phase filters during this thesis.

13.1.1. Derivation and Discrete Implementation

In the case of the moving average technique a data smoothing is performed by averaging a certain set of data points within a defined bandwidth. For the constant relative bandwidth smoothing the frequency range is defined according to a bandwidth of $1/n$ octaves. This ensures a constant smoothing bandwidth with respect to the logarithmic frequency axis of the magnitude spectrum.

The basic averaging procedure of one single averaging step is illustrated in figure 13.1.

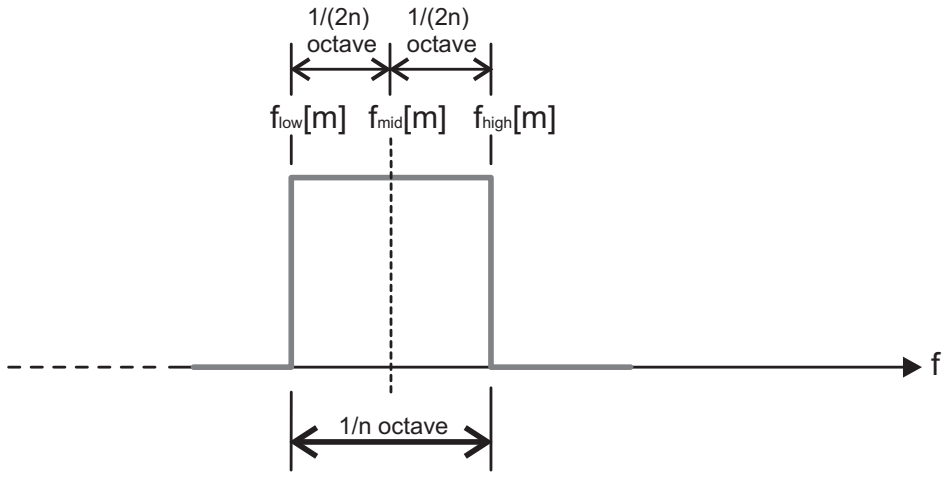

Figure 13.1.: Illustration of the average bandwidth in the case of constant relative bandwidth smoothing in the case of a logarithmic frequency axis.

The averaging bandwidth can be divided into two parts with a bandwidth of $1/(2n)$ octaves, one part below and one above the mid frequency. Thus two frequencies $f_{\text{low}} = f_{\text{mid}}/2^{1/(2n)}$ and $f_{\text{high}} = f_{\text{mid}} \cdot 2^{1/(2n)}$ can be calculated. The data points between f_{low} and f_{high} are averaged and the value is assigned to f_{mid}.

After this averaging process the averaging "window" is moved upwards side by side to the previous "window position" in order that:

$$f_{\text{low}}[m] = f_{\text{high}}[m-1], \tag{13.1}$$

where $m = 0 \ldots M-1$ denotes the m^{th} averaging step for in maximum M averaging steps. Therefore several frequencies can be calculated beginning at a specified frequency f_{start} according to:

$$\begin{aligned} f_{\text{low}}[m] &= f_{\text{start}} \cdot 2^{\frac{m}{n}} \\ f_{\text{mid}}[m] &= f_{\text{start}} \cdot 2^{\frac{1+2m}{2n}} \\ f_{\text{high}}[m] &= f_{\text{start}} \cdot 2^{\frac{1+m}{n}}, \end{aligned}$$

where n denotes the smoothing bandwidth in $1/n$ octaves.

The start frequency is set to the first frequency bin of the magnitude response $|H(j\omega)|$. The last possible frequency $f_{\text{end}}[M-1]$ within the magnitude response $|H(j\omega)|$ is $f_s/2 \geq f_{\text{end}}[M-1]$. Therefore the maximum amount of averaging steps can be calculated according to:

$$M \leq n \cdot \log_2\left(\frac{f_s}{2 \cdot f_{\text{start}}}\right). \tag{13.2}$$

Summarizing the moving average process is performed according to the following steps:

1. Calculation of the first frequency bin in the magnitude spectrum $f_{\text{start}} = df$, where df is the frequency resolution.

2. Calculation of the maximum amount of averaging steps M within the whole magnitude spectrum for the specific smoothing resolution of $1/n$ octaves. Round M downwards to the next integer value, if necessary.

3. Calculation of several frequencies: $f_{\text{low}}[m]$, $f_{\text{mid}}[m]$ and $f_{\text{high}}[m]$ for $m = 0 \ldots M-1$.

4. Calculation of several averages for the frequency ranges between $f_{\text{low}}[m]$ and $f_{\text{high}}[m]$. These averaged values are assigned to the mid frequency $f_{\text{mid}}[m]$.

5. Add the first and the last data point of the magnitude spectrum without averaging, as no bandwidth is left to perform further averaging.

Part IV.

Practical Preparation

In this part the practical preparations for the implementation of the software tools RoLoSpEQ and D-MLCNC in MATLAB (see part V and VI) are explained.

Initially chapter 14 deals with the practical application of discrete-time filters in MATLAB and the particularities, which have to be considered. In chapter 15 the focus is put on the practical filter implementation of digitally simulated analog filters for the use in D-MLCNC. Subsequently in chapter 16 basic FIR applications for the use in D-MLCNC are discussed. In addition the minimum-phase FIR filter generation procedure, which is used in RoLoSpEQ is described. Chapter 17 deals with the discrete interpolation, which is also important in RoLoSpEQ. Finally in chapter 18 the MATLAB feature expansion for multichannel audio support using the utility "Playrec" is discussed. "Playrec" provides the basis for the developed quasi real-time block processing approach, which is also described in this chapter.

14. Discrete-Time Filtering in MATLAB

In this chapter suitable practical filter strategies for the developed software tools RoLoSpEQ and D-MLCNC are discussed, which consider a high and accurate filter performance in order to enable the development of a quasi real-time signal processing approach.

14.1. Filtering with Discrete-Time Filter Objects

In general there are a couple of possibilities to perform a filtering of any audio signal in MATLAB with IIR or FIR filters. One possibility is the use of the native MATLAB function "filter". Nevertheless this function is merely used in the software tool D-MLCNC, because mainly IIR filters are applied, and the "filter" function was found to have the best speed-performance[1] for IIR filtering in MATLAB.

In the software tool D-MLCNC in each equalizer channel a serial connection of several different filter stages (e.g. high-pass filter and low-pass filter) has to be achieved (see figure 23.1). One possibility is to perform a sequential filtering of certain output signals of each filter stage, by filtering the output signal of the previous filter stage in the next filter stage and so on.

With regard to a high speed-performance this method is not suitable. For this purpose a special filter object for discrete-time filters is used, namely the "dfilt"-object, which is a part of the "Filter Design Toolbox" in MATLAB. Among others the "dfilt"-object enables the possibility of a simple serial connection of multiple filters, which is represented in one "dfilt"-object. Additionally the overall transfer function can be represented in a series of second order filters to minimize the risk of unstable filters (see chapter 14.2).

The resulting discrete-time filter object $H_{\mathrm{d},1\ldots8}$ can be directly processed with the MATLAB function "filter".

[1] The speed-performance of the filter process is extremely important for the realization of a quasi-real time implementation (see chapter 18.2). The MATLAB function "filter" is based on a direct-form II realization [Mat08] according to equation (7.1), which is a suitable realization of IIR filters (see chapter 8).

In the case of FIR filters in the software tool RoLoSpEQ, the MATLAB "dfilt"-objects are not used, because on the one hand a serial connection of filters is not necessary (there is only one filter per channel) and on the other hand series of second order filters are not required because FIR filters never get unstable [Mey09]. In addition the MATLAB "dfilt"-objects are not suitable for use in the MATLAB function "fftfilt", which is used in the software tool RoLoSpEQ instead of the MATLAB function "filter" (see chapter 14.3).

14.2. LTD System Representation – Practical Problems and Improvements

In the following paragraph two different representations for digital systems are introduced. Disadvantages in the practical implementation are discussed and improvements are explained.

14.2.1. Transfer Function Representation

The transfer function (TF) representation is the standard representation for LTD systems (see chapter 7.2). In general a LTD system is represented in the discrete z-domain as a rational function according to equation (7.3). For a full characterization of any LTI-system merely the filter coefficients $b_{1...N}$ of the numerator and $a_{1...M}$ of the denominator have to be known.

Certainly problems might appear using the TF representation for higher order LTD systems in MATLAB due to a limited numerical resolution of the filter coefficients[2]. The result is a shifting of poles and zeros in the complex z-plane, which yields uncertainties in the filter frequency response as well [Mey09]. If a pole is shifted to a position outside the unit circle in the complex z-plane the LTD system is unstable.

These negative effects are various for different filter structures. For the direct-form II implementation, which is discussed in more detail in chapter 8 the negative effects due to the shifting of poles and zeros are worst, because all coefficients b_i of the numerator have an impact on the position of each zero and respectively all coefficients a_i have an impact on the position of each pole [Mey09].

Therefore an alternative system structure is used in the software tool D-MLCNC, namely the cascade structure, which is described in the following paragraph.

14.2.2. Cascade Structure – Serial Connection of Biquad Filters

Another type of filter structure, which offers an improved behavior concerning the filter coefficient quantization is termed *cascade structure*. The benefit is, that the position of poles and zeros depends merely on a few filter coefficients [Mey09]. Therefore the cascade structure is suitable for the practical IIR implementation.

[2] The filter coefficients of LTD systems are quantized as well.

A short derivation:
The numerator and the denominator of the basic TF (equation (7.3)) are factorized, leading to [Mey09]:

$$H(z) = b_0 \cdot \frac{\prod_{i=1}^{N} (z - z_{\text{Zero},i})}{\prod_{i=1}^{M} (z - z_{\text{Pole},i})} \cdot z^{M-N}, \tag{14.1}$$

where $z_{\text{Zero},i}$ is the position of the i^{th} zero and $z_{\text{Pole},i}$ is the position of the i^{th} pole in the complex z-plane. Within all zeros and poles there can be real valued zeros and poles and complex conjugated zeros and poles.

Several complex conjugated zero couples are, with a multiplication combined to real valued second order subsystems. The same procedure is performed for all complex conjugated poles.

Several real valued zeros and poles represent single real valued first order systems.

In general the cascade structure is described according to [Mey09]:

$$H(z) = b_0 \cdot \frac{\prod_{i=1}^{N/2} b_{0i} + b_{1i} \cdot z^{-1} + b_{2i} \cdot z^{-2}}{\prod_{i=1}^{M/2} 1 + a_{1i} \cdot z^{-1} + a_{2i} \cdot z^{-2}}, \tag{14.2}$$

where i denotes the i^{th} second order subsystem, which are also termed biquad filters. In the case of a real valued first order subsystem i the corresponding i^{th} biquad degrades to a first order system as well [Mey09]. All in all a cascade structure is resulting.

Practical Application:
In MATLAB the cascade structure can be easily generated. For this purpose any desired filter is specified with its positions of poles and zeros at first[3]. Then the MATLAB function "zp2sos" is used to transform the zero, pole and gain representation of the filter in a series connection of biquad filters[4]. This function is part of the "Signal Processing Toolbox" in MATLAB. Finally the cascaded biquad filters are represented in direct-form II realization in a MATLAB "dfilt" object. This is achieved with the MATLAB function "dfilt.df2sos".

14.3. High-Performance FIR Filtering – Fast Convolution of Long Sequences

The speed performance of the filter process is extremely important in the case of a quasi real-time implementation (see chapter 18.2), which is desired in the software tools RoLoSpEQ and D-MLCNC. Generally any filter in MATLAB can be applied with the native MATLAB function "filter", which performs the filtering in a direct-form II implementation (see figure 8.1) according to equation (7.1).

Certainly in the case of FIR filters, which are mainly important in the software tool RoLoSpEQ (see part V) the filter coefficients also represent the FIR filter IR $h[n]$ (see chapter 9). Thus, this filter IR is directly applicable as FIR filter to any audio input signal $x[n]$

[3] This representation is also known as zero, pole and gain representation (zpk) in MATLAB.
[4] In MATLAB a series connection of biquad filters is termed *second order section* (sos).

in order to achieve the desired filtered output signal $y[n]$ according to:

$$y[n] = h[n] * x[n]. \tag{14.3}$$

which is a basic convolution operation [Mey09]. Alternatively the convolution could be performed via FFT, which is describe in more detail in chapter 12.3.

More suitable for the use in the tool RoLoSpEQ is a *FFT based convolution method for long sequences*, which makes use of an *overlap and add method* to achieve a strong acceleration of the convolution process for filter sequences with a length of $N > 30$ samples [Zö8]. More detailed explanations and derivations can be found in further literature [OSB99, Zö8].

E.g. in the software tool RoLoSpEQ the filter sequence lengths can be variated in a range between $N = 1024$ and $N = 65536$. Hence the filter process performed with a common convolution process requires quite a long time[5] and is therefore not suitable for a quasi real-time implementation (see chapter 18.2).

The desired FFT based convolution method for long sequences is natively provided in the "Signal Processing Toolbox" in MATLAB. The appropriate function is termed "fftfilt" and is used for the quasi-real time filter simulation processing in the software tool RoLoSpEQ.

[5] E.g. a convolution in MATLAB requires approximately $1.8s$ for the convolution of two 32768 samples signal frames with a standard PC (Intel Core2Duo $2.4GHz$). In comparison the FFT based convolution method merely requires approximately $0.015s$ for the same operation.

15. Implementation of IIR Filters

In this thesis *IIR filters* are mainly used in the software tool D-MLCNC (see part VI). D-MLCNC enables a quasi real-time simulation of a digital eight channel loudspeaker crossover network and is completely implemented in MATLAB [Mat08]. The main intention of D-MLCNC is a flexible simulation of commonly used filter types and filter approximation. As very much research on loudspeaker crossover networks is based on analog filters (e.g. passive and active crossover networks [Dic05]) a digital simulation of analog filters is preferred for the use in D-MLCNC. The digital simulation of analog filters is described in more detail in chapter 8.1.

Summarizing the practical implementation of digitally simulated analog filters is only possible with IIR filters. Several IIR filter types, which are used in the tool D-MLCNC and specialties concerning the practical implementation in MATLAB are discussed in the following paragraphs.

15.1. High-pass and Low-pass Filters

In the case of loudspeaker crossover networks merely low-pass and high-pass filters are required for the basic application of frequency splitting for different loudspeaker drivers in a multi-way loudspeaker (see figure 10.1). In general for mid-range loudspeaker drivers a band-pass filtering is required. Certainly the pass-band behavior can also be achieved by a serial connection of a low-pass filter and high-pass filter, which is the case for band-pass filtering in the software tool RoLoSpEQ.

15.1.1. Butterworth Filters

The Butterworth approximation is the only used filter approximation in the software tool D-MLCNC.

Filter approximation are necessary, because ideal filters cannot be realized in practice (see chapter 6.4). Thus filter approximations are used to approximate the behavior of their ideal counterparts. One possible filter approximation is the Butterworth approximation. Further details about filter approximations and a short derivation of the Butterworth approximation can be found in chapter 6.4. Furthermore the filter calculation for basic filter types is discussed in chapter 6.5. Subsequently the resulting analog Butterworth filter with filter order n has to be transposed from the analog s-domain in the digital z-domain. This is often performed with the bilinear z-transform. An example derivation of a digitally simulated second order Butterworth low-pass filter can be found in chapter 8.1.1.

For the practical realization several calculation steps are natively implemented in MATLAB [Mat08]. Nevertheless the theoretical background discussed in part II is necessary to ensure an accurate use of the provided filters. Furthermore several filters are derived with the *zero, pole and gain representation* and are transposed to a *serial connection of biquad filters*, as

uncertainties of filter frequency responses or even unstable higher order filters can be avoided (see chapter 14.2).

For the implementation of Butterworth high-pass and low-pass filters the MATLAB function "butter" is used. This function derives a digitally simulated Butterworth filter with a desired order and cut-off frequency. The calculation of several basic filter types (see chapter 5.1) is possible, however merely high-pass and low-pass filters are needed in the software tool D-MLCNC.

In the software tool D-MLCNC Butterworth filters up to order $N = 8$ are available.

15.1.2. Linkwitz-Riley Filters

Linkwitz-Riley filters (L-R filters) are another possible choice for high-pass and low-pass filtering in the software tool D-MLCNC. L-R filters are not a further filter approximation (see chapter 10.2.6). L-R filters are based on the Butterworth approximation [Lin78, Dic05, Boh05] and are realized by serial cascading of similar Butterworth filters. Hence L-R filters offer merely even filter orders of $N \geq 2$.

For the digital simulation of a L-R filter with even filter order N in MATLAB, a Butterworth with similar cut-off frequency and filter order $N/2$ is derived (see chapter 15.1). Subsequently two of the Butterworth filters are serial cascaded or the input signal $x[n]$ is filtered twice.

In the software tool D-MLCNC L-R filters up to order $N = 8$ are available.

15.2. Parametric Filter Structures

On the one hand loudspeaker crossover networks aim at a splitting of the audio signal into separate frequency bands, which are accurate for certain loudspeaker drivers in a multi-way loudspeaker [Dic05, GW06].

On the other hand loudspeaker crossover networks enable an active influence on the overall frequency response of a single- or a multi-way loudspeaker [Dic05, GW06] (see chapter 10.2.5). For this purpose in the software tool D-MLCNC two parametric filter structures are implemented namely the *peak filter* and the *shelving filter*. Peak filters and shelving filters are special weighting filters, which are used in practice in a series connection [ZÖ8] of a low-shelving filter, several peak filters and a high shelving filter (see figure 10.4).

Parametric filter structures enable direct access to the parameters of a filter transfer function [ZÖ8]. Hence the variation of associated parameters yields a variation of e.g. the cut-off frequency of a shelving filter. The benefit is that the variation of certain filter specifics (e.g. the cut-off frequency) does not induce a recalculation of several filter coefficients.

15.2.1. Shelving Filter

In the software tool D-MLCNC in maximum five digital first order parametric low-shelving and high-shelving filters are available in all filter channels (see part VI).

The low-shelving filter is implemented according to equation (10.4) and the high-shelving filter is implemented according to equation (10.7). For this purpose two MATLAB functions are created, which derive the desired transfer function coefficients according to the defined parameters f_c, V_0 and f_s. In addition the transfer function is transposed to zero, pole and

gain representation (see chapter 14.2) and is converted in a special MATLAB filter object (see chapter 14.1).

15.2.2. Peak Filter

In the software tool D-MLCNC in maximum five digital second order parametric peak filters are available in several filter channels.

These peak filters are implemented according to equation (10.11). For this purpose a MATLAB function is created, which derives the desired transfer function coefficients according to the defined parameters f_m, V_0, Q and f_s. In addition the transfer function is transposed to zero, pole and gain representation and is stored in a special MATLAB filter object.

16. Implementation of FIR Filters

16.1. Signal Delaying, Inversion and Gain Controlling

Signal delaying is used in the tool RoLoSpEQ in order to achieve a correction of displacement errors of loudspeakers in a common stereo loudspeaker system (see part V). Additionally signal delaying is used in the tool D-MLCNC to achieve a correction of a certain displacement of the acoustic center of different loudspeaker drivers in a multi-way loudspeaker (see part VI).

The *signal inversion* is an important task to avoid cancellation effects at the crossover frequencies of two separate loudspeaker drivers in a multi-way loudspeaker (see chapter 10). Furthermore the signal inversion could be interesting in the tool RoLoSpEQ for an automatic correction of reverse connected loudspeakers in a stereo loudspeaker setup. However this is not implemented yet.

A *controlling of the signal gain* is used in both tools as well, to achieve a gain adjustment of two separate loudspeakers in the tool RoLoSpEQ or of different loudspeaker drivers in the tool D-MLCNC.

The realization of a non-fractional signal delay is very simple in the case of FIR filters, which are used within both tools. Non-fractional signal delays apply a delay, which is an integral multiple of the basic sampling frequency f_s. In discrete-time sequences the sampling instants are uniformly spaced with $T_s = 1/f_s$ seconds distance [Zö8].

In chapter 9 the transfer function of a FIR filter is presented according to:

$$H(z) = \sum_{i=0}^{N} b_i \cdot z^{-i} = b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2} + ... + b_N \cdot z^{-N}. \tag{16.1}$$

As $H(z) = \frac{Y(z)}{X(z)}$ [OSB99, Mey09] the filter output $Y(z)$ can be expressed according to:

$$Y(z) = X(z) \cdot \sum_{i=0}^{N} b_i \cdot z^{-i}, \tag{16.2}$$

where $X(z)$ is the input of the filter.

The term z^{-i} in equation 16.2 represents the time-delay, which is applied on discrete filter coefficients $b_0 \dots b_N$, where i is an integral multiple of the basic sampling frequency f_s. Thus an additional fractional delay D can be applied according to:

$$Y(z) = X(z) \cdot \left[\sum_{i=0}^{N} b_i \cdot z^{-i} \right] \cdot z^{-D}, \tag{16.3}$$

where D is an integer value representing the delay in samples. If $b_0 = 1$ and several filter coefficients b_i are chosen to $b_i = 0$ for $i > 1$, a certain change of the complex spectrum of

$X(z)$ is avoided, as a convolution of any sequence with a unit pulse $x[n] * 1 = x[n]$ does not have an influence. Thus a delayed unit pulse does not change the input signal, but due to the convolution process, a delaying of the input sequence appears according to the certain fractional delay D. The FIR delay filter is derived according to:

$$Y_{\text{delayed}}(z) = b_0 \cdot X(z) \cdot z^{-D}. \tag{16.4}$$

The signal delaying can be seen more clearly if equation 16.4 is transformed in a differential equation [Mey09] according to:

$$y_{\text{delayed}}[n] = b_0 \cdot x[n - D]. \tag{16.5}$$

The delay time Δt in seconds can be recalculated according to:

$$\Delta t = D \cdot T_s = \frac{D}{f_s}. \tag{16.6}$$

For a given delay time this equation is transposed. The solution has to be rounded to an integer value.

The gain of $x[n]$ can be variated according to the free choice of b_0. E.g. a multiplication of $x[n]$ with $b_0 = 0.5$ yields an attenuation of $20 \cdot \log_{10}(0.5) \approx -6dB$.

Certainly in the case of a *signal inversion* b_0 has merely to be negated (e.g. $b_0 = -1$).

Summary:
The signal delay, the signal gain and a signal inversion are controlled together in one FIR filter according to equation 16.4 in the tool RoLoSpEQ and in the tool D-MLCNC as well.

16.2. Filter Generation via Hilbert Transform

Minimum phase FIR filters are used for the task of *room response equalization* inside the tool RoLoSpEQ (see part V), because merely corrections of the magnitude response are desired (see chapter 4.1.3). For this purpose an adequate filter design method was found by using the *Hilbert transform* and the *complex cepstrum*. Further details are discussed in chapter 9.1.1.

The desired filter responses, which are derived in the tool RoLoSpEQ contain no phase information at all. However in the case of minimum phase systems the phase response is directly related to the magnitude response with the Hilbert transform. This relation is discussed in more detail in chapter 9.1.1.

According to the calculation instruction given in chapter 9.1.1 the derivation of a minimum phase filter IR $h_{\text{mp}}[n]$ is performed using the desired filter magnitude response $\left|H_{\text{mp}}\left(e^{j\omega}\right)\right|$. The obtained filter IR $h_{\text{mp}}[n]$ is directly applicable as minimum phase FIR filter to any audio input signal $x[n]$ to achieve the desired filtered output signal $y[n]$, which is performed with a fast FFT based convolution method in MATLAB (see chapter 14.3).

16.2.1. Practical Derivation of the Filter Generation Process

The following paragraph aims at an illustration of the FIR filter generation procedure, which is realized in the tool RoLoSpEQ (see part V).

Based on a set of measurements, the developed algorithm of the tool RoLoSpEQ derives a desired equalizing filter magnitude response, which is illustrated in figure 16.1.

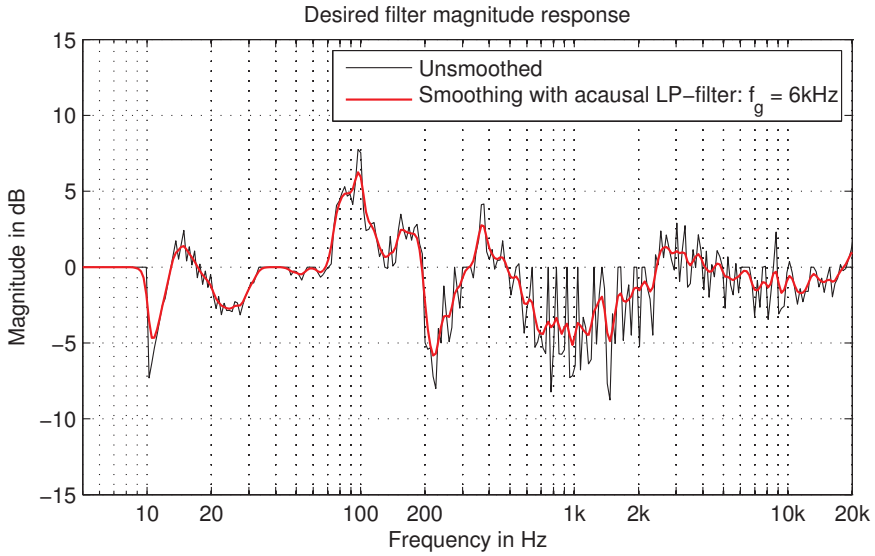

Figure 16.1.: Desired magnitude response (1/24 octave resolution due to the measurement smoothing, $f_s = 44.1kHz$) – calculated with the tool RoLoSpEQ. Comparison of the unsmoothed magnitude (black curve) and a smoothed magnitude response (red/gray curve). The smoothing is based on a simple filter processing.

This desired (unsmoothed) magnitude response (black curve) offers a very unsteady behavior. The unsteady behavior is further smoothed by applying a special filter technique, which is subsequently explained:

The illustrated unsmoothed magnitude response is considered to be a real valued time sequence $x[n]$ sampled at equidistant points. However, in reality the data points in this virtual sequence $x[n]$ offer no equidistant spacing. This is a result of the 1/24 octave smoothing of the basic measurement responses (see chapter 13). Thus the data points in $x[n]$ are spaced in logarithmic manner, according to the logarithmic frequency axis.

The virtual sequence $x[n]$ is filtered acausal[1] with a digital-simulated 1^{st} order Butterworth low-pass filter with a cut-off frequency $f_g = 6kHz$ at a fixed sampling frequency $f_s = 48kHz$. This filter process leads to a smoothed version of the virtual sequence $x[n]$, which henceforth is the desired equalizing filter magnitude response (red/gray curve).

The benefit of this approach is a *consistent smoothing* over the complete frequency range.

After this smoothing process the magnitude response is interpolated in $N/2 = 2^{14}$ equidistant points with a linear interpolation. This procedure is explained in more detail in chapter 17.

In figure 16.2 the calculation of a double sided spectrum is illustrated. For this purpose the smoothed and interpolated equalizing filter magnitude spectrum with equidistant frequency bins is mirrored and shifted according to Fourier transform characteristics[2].

[1] The acausal filtering is used to avoid a frequency dependent shifting caused by the non-linear behavior of the filter. With acausal filtering even a zero phase filtering is possible. Further details about acausal filtering with recursive filters can be found in literature [OSB99, Mey09].

By using the acausal filtering the smoothed magnitude response remains at the "old" position.

[2] The second half spectrum represents the negative frequencies of the Fourier spectrum.

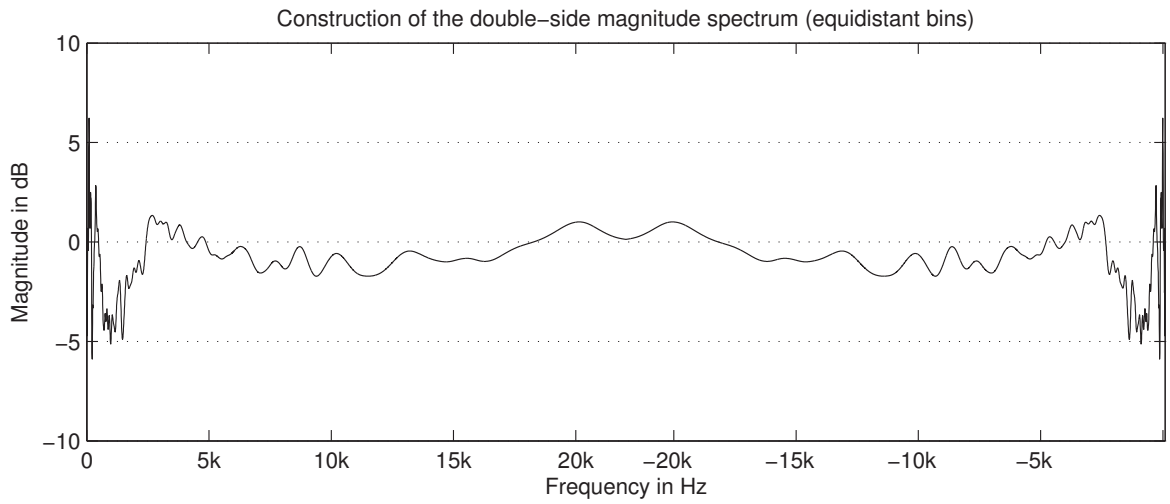

Figure 16.2.: Construction of a double sided magnitude spectrum for a frequency range of $df ... f_s$ according.

Essentially the magnitude spectrum would be periodically repeated, but this task is not necessary for the implementation in MATLAB. Hence the complete $N = 2^{15}$-point filter magnitude spectrum $\left|H\left(e^{j\omega}\right)\right|$ is defined.

Subsequently the FIR filter generation is performed according to the calculation instructions in chapter 9.1.1. The magnitude response $\left|H_{\mathrm{mp}}\left(e^{j\omega}\right)\right| = \left|\mathrm{FFT}\left\{h_{\mathrm{mp}}\left[n\right]\right\}\right|$ of the resulting FIR filter $h_{\mathrm{mp}}\left[n\right]$ is illustrated in figure 16.3 in comparison with the originally desired filter magnitude response.

Summarizing in figure 16.4 the complete filter generation process used in the software tool RoLoSpEQ is illustrated in a data flow graph.

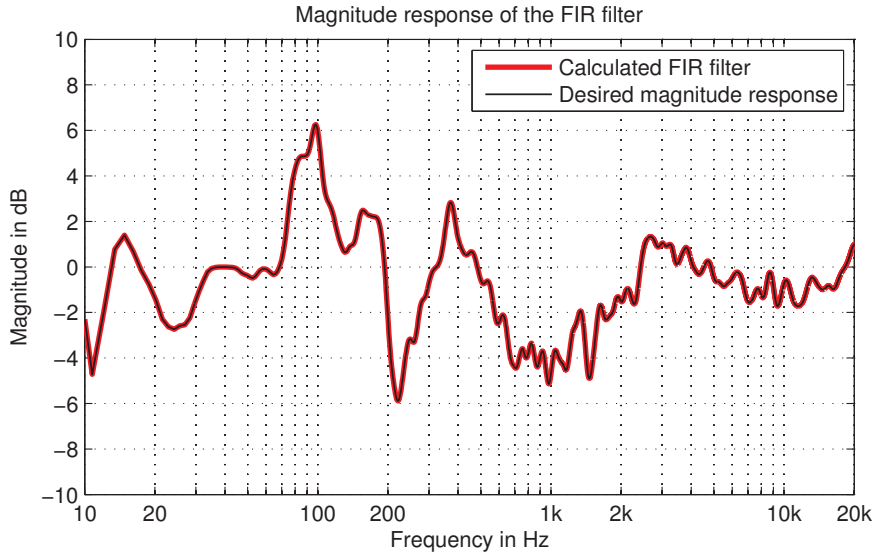

Figure 16.3.: Magnitude response $\left|H_{\mathrm{mp}}\left(e^{j\omega}\right)\right|$ of the calculated FIR filter IR $h_{\mathrm{mp}}[n]$ (red/gray curve) vs. desired filter magnitude response (black curve).

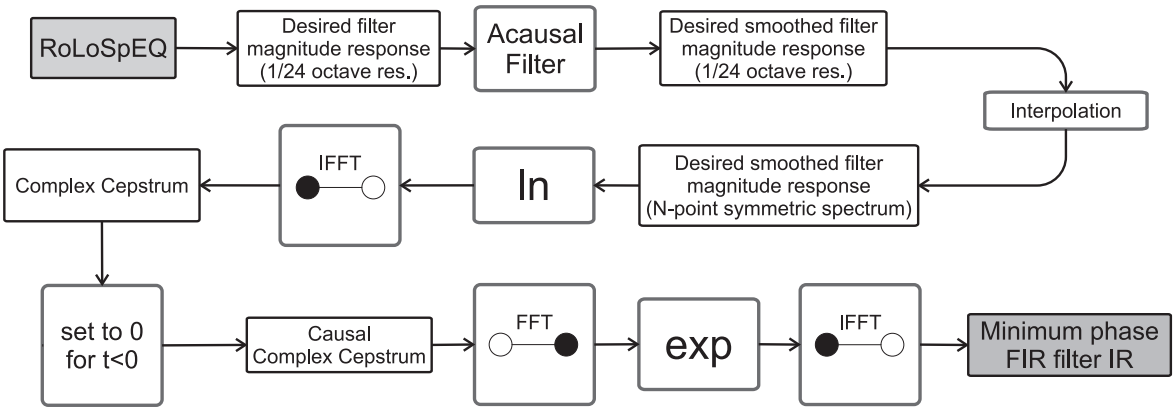

Figure 16.4.: Data flow graph of the FIR filter generation process in the software tool RoLoSpEQ.

17. Discrete Interpolation in MATLAB

Generally discrete interpolation is a special kind of approximation aiming at a minimization of the interpolation error at predefined interpolation points [Bar01]. The determination of intermediate values between two adjacent interpolation points is termed *interpolation*.

In this thesis discrete interpolation is used if a set of data points doesn't offer a desired data spacing. Therefore the interpolation is used in two cases:

1. Interpolation of microphone magnitude responses, which are delivered with calibration files of professional measurement microphones (e.g. DPA 4006-TL).

2. Interpolation of filter magnitude responses with arbitrary resolution (e.g. 1/24 octave resolution).

The intention of the data interpolation is to achieve a data spacing according to a required FFT resolution $df = \frac{f_s}{N_{\mathrm{FFT}}}$, which is dependent on the FFT length N_{FFT} and the sampling frequency f_s.

In this thesis merely the linear interpolation is used, which is explained in the following paragraph.

17.1. Linear Interpolation

The linear interpolation is a simple approach to derive the intermediate values between predefined data points. For this purpose the data points are virtually connected with a straight line. For two adjacent data points $(x_n; f(x_n))$ and $(x_{n+1}; f(x_{n+1}))$ several values on the straight line can be derived with a basic straight line equation [Bar01] according to:

$$f(x) = f(x_n) + \frac{f(x_{n+1}) - f(x_n)}{x_{n+1} - x_n}(x - x_n), \tag{17.1}$$

where x is a user-defined position between x_n and x_{n+1}. Thus an arbitrary spacing resolution can be chosen.

In figure 17.1 ten randomly chosen equidistant data points (red/gray points) and a linear interpolation according to equation (17.1) with five times more data points with a logarithmic spacing (black pluses) are illustrated. The straight line behavior of the linear interpolation can be easily observed. Furthermore an arbitrary data spacing can be used for the interpolation.

In addition the other way round can be used as well, that is a linear interpolation of arbitrary spaced original data in order to receive a linear data spacing. This case is illustrated in figure 17.2, where a microphone magnitude response, which can be used for a frequency response measurement calibration, can be seen.

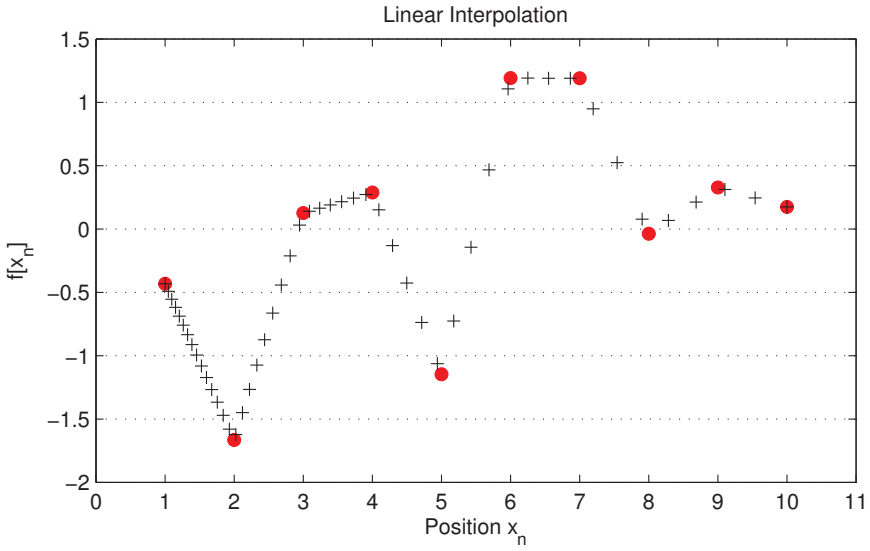

Figure 17.1.: Linear interpolation of a simple data set consisting of ten equidistant data points (red/gray points). The linear interpolation is derived with 50 positions of x_n, which are logarithmic spaced (black pluses).

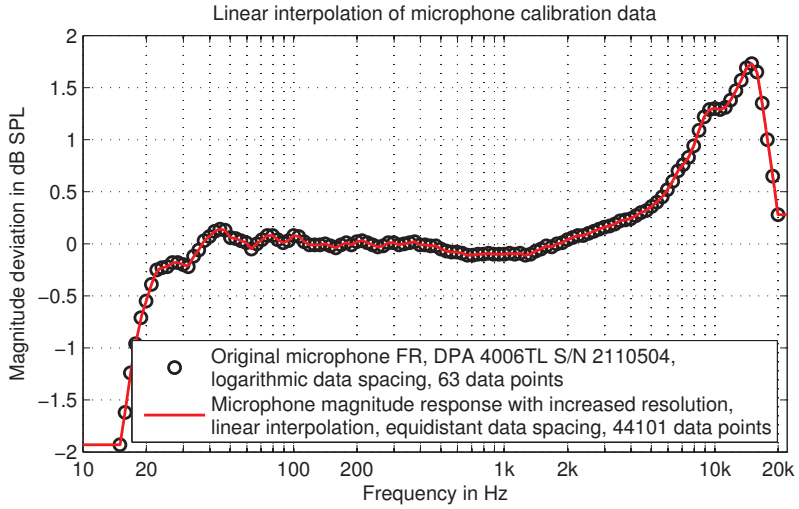

Figure 17.2.: Comparison of original microphone calibration data with 63 logarithmic spaced data points between $20Hz - 20kHz$ (black circles) and a linear interpolation of the same data set with 24001 data points between $0Hz - 24kHz$ (red/gray curve). The interpolation in the frequency range outside the data range of the original data is not used. The first and the last data point of the original data are sustained instead.

In order to enable the correction of the magnitude spectrum of any measurement spectra by a simple subtraction, the microphone correction data has to be equally spaced due to FFT characteristics. The FFT is used for measurement analysis. Furthermore the first and the

last data point of the original data has to be used for several values outside the original data space, because no accurate interpolation is possible.

The linear interpolation is used in the software tool RoLoSpEQ in order to:

1. Perform a microphone magnitude response correction of measurements.
2. To generate FIR filters for room equalization (e.g. see chapter 16.2.1).

For this purpose the native MATLAB function "interp1" is used to generate linear interpolated data sets. Non-defined data regions, which are not included in the original data set are chosen manually – either in the way as described above for microphone magnitude responses or by setting these non-defined values to a corresponding to $0dB$ gain for a filter magnitude response.

18. MATLAB Feature Expansion – Multichannel Audio Support with ASIO

18.1. "Playrec" for MATLAB

"Playrec" is a MATLAB utility ("mex" file) enabling an access to any soundcard using PortAudio. PortAudio is an open source cross-platform audio API (application programming interface), which enables a flexible multichannel support in MATLAB.

"Playrec" can be compiled for different platforms and different audio host APIs (e.g. ASIO[1], MME[2]).

During this thesis "Playrec 2.1.1" was compiled for both MATLAB 7.6 32bit on Windows XP Professional 32bit and MATLAB 7.6 64bit on Windows 7 Professional 64bit, according to the detailed compilation instructions on the "Playrec" website [Hum08]. ASIO and MME are chosen as audio host API for the use with "Playrec". On the one hand ASIO enables a multichannel support for several ASIO soundcards (in maximum eight synchronous output channels are required for the software tool D-MLCNC). On the other hand if no ASIO is supported by the soundcard the developed software tools can be used anyway, as MME is natively supported on the Windows platform.

"Playrec" was developed by Robert Humphrey [Hum06]. Further details about the development can be found in publications [Hum06, Hum08].

18.1.1. Benefits of "Playrec" for MATLAB

The following benefits are crucial for the use of "Playrec" in the software tools RoLoSpEQ and D-MLCNC:

- "Playrec" is compatible with MATLAB.
- "Playrec" is compatible with different platforms (Windows, Max and Unix) due to the use of "PortAudio". In addition a flexible compilation enables support for both, MATLAB 32bit and MATLAB 64bit.
- A multichannel support is achieved for several soundcards and drivers with arbitrary amount of input and output channels.
- All samples are buffered, in order that MATLAB can continue with other processing.
- All new buffered samples are automatically appended to older ones. Thus e.g. a block-processing is enabled, which is discussed in the following paragraphs.

[1] ASIO (Audio Stream Input/Output) is developed by Steinberg and is a multi-platform audio transfer protocol. ASIO enables a connection between software and professional multichannel audio hardware.

[2] MME (Windows MultiMedia Extension) is developed by Microsoft and is a standard audio API for Windows. In the case of MME merely two input and output channels are available at the same time.

18.2. Quasi Real-Time Multichannel Digital Audio Signal Processing in MATLAB

Among others the software tools RoLoSpEQ and D-MLCNC (see part V and VI) enable a simulation of audio filters. Actually the filtering of an audio signal can be achieved by filtering the complete input signal (e.g. a piece of music). Certainly if any changes are performed in the filter specifics whilst processing or playback the change to the signal output (the filtered signal) will be not audible until the whole input signal is again completely filtered with the new filter. In the case of long audio signals and the request of sensible perception of small differences[3] in the output signal this approach is not practical.

Therefore an alternative approach is developed based on an audio block-processing. On the one hand an audio block-processing enables a piecewise parallel filtering of multiple output signals. On the other hand a quasi real-time processing is enabled by an interruption of the block-processing which leads to a slow-down of the execution. Hence it is possible to change filter specifications during runtime and additionally to achieve a perception of the changes with a reasonable low latency – in this thesis this is termed quasi real-time processing.

In the following paragraph the basic approach for simultaneous signal playback with filtering is explained for the filtering of one channel. This is due to the complexity of the explanation for a multichannel audio signal processing. However in the case of a multichannel audio signal block-processing (e.g. in the software tools RoLoSpEQ and D-MLCNC) the approach can be easily extended by an one by one filtering for multiple channels at each processing step.

18.2.1. Signal Playback with Simultaneous Filtering

The development of a block-processing approach for a quasi real-time audio signal processing in MATLAB is one of the main acquisitions within this thesis and is based on the features of a special MATLAB "mex" file which is used for several implementations (see chapter 18.1). The approach enables a quasi real-time multichannel signal processing within the standard MATLAB environment.

In figure 18.1 the developed block-processing approach for signal playback and simultaneous filtering is illustrated. Generally a desired input signal $x\,[n]$ is filtered with a defined filter $h\,[n]$. In addition the filter $h\,[n]$ can be changed during runtime of the signal block-processing. Therefore an interruption of the block-processing is desirable.

Due to special features of the utility "playrec", which is part of the used MATLAB "mex" file this block-processing with interruption is enabled and subsequently explained.

First step:
At the first block-processing step $i = 1$ the first input signal frame is defined according to:

$$x_{i=1}\,[n] = \begin{cases} x\,[n] & \text{if } 1 \leq n \leq N_{\text{frames}}, \\ \emptyset, & \text{otherwise}, \end{cases} \tag{18.1}$$

where N_{frames} is the buffer size in samples. Then the signal $x_1\,[n]$ is filtered:

$$y_{q=1}\,[n] = x_1\,[n] * h\,[n]\,, \tag{18.2}$$

[3] Small differences in similar audio signals can merely be recognized if a fast switching between the different audio signals is possible.

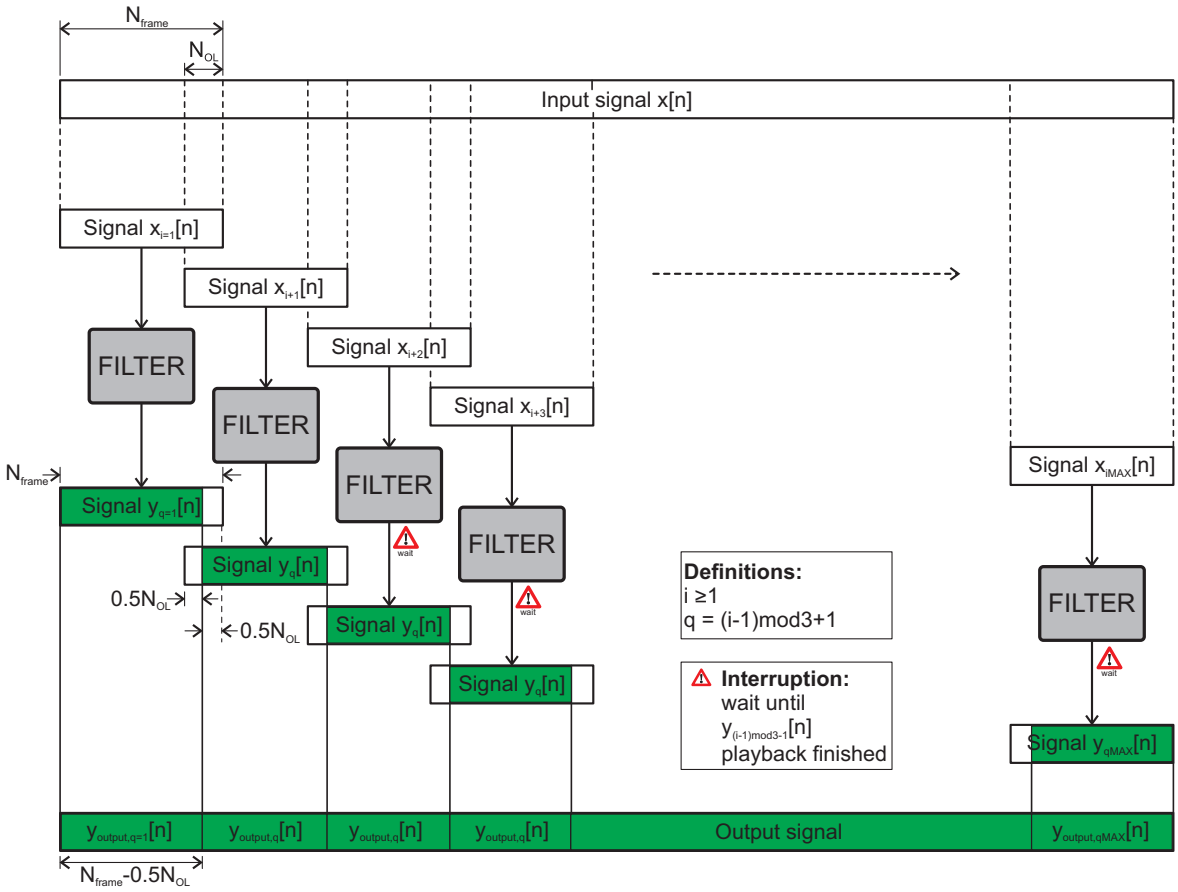

Figure 18.1.: Quasi real-time block-processing for signal playback and simultaneous filtering.

where $h\,[n]$ is the current filter and $q = (i-1) \bmod 3+1$ is a further variable beside i denoting the current position in the output buffer[4]. This buffering is introduced to save memory, as former played output signals are no longer needed in the case of a quasi-real time audio signal processing. Thus an output signal buffer of three output frames $y_1\,[n]$, $y_2\,[n]$ and $y_3\,[n]$ is created.

The output signal is derived according to:

$$y_{\text{output},q=1}\,[n] = \begin{cases} y_{q=1}\,[n] & \text{if } 1 \leq n \leq N_{\text{frames}} - 0.5 \cdot N_{\text{OL}}, \\ \emptyset, & \text{otherwise,} \end{cases} \tag{18.3}$$

where N_{OL} is the frame overlapping in samples, which is necessary for a distortion free playback due to a certain rise time of the used filter. Without using an appropriate frame overlapping an *impulsive click* can be perceived at the beginning of each filtered signal frame. This effect is additionally forced by the quasi rectangular windowing of the input signal[5].

Finally $y_{\text{output},1}\,[n]$ is directly sent to the function "playrec('play')", which performs a subsequential playback of available signal frames.

[4]For ascending integer values of i, $q\,[i] = (i-1) \bmod 3 + 1$ results in a sequence according to $q = [1 \quad 2 \quad 3 \quad 1 \quad 2 \quad 3 \quad \ldots]$.

[5]A steep slope at the beginning of one signal block can appear due to the blocking of the input signal.

Further steps:
Further block-processing steps for $i > 1$ can be derived more general. The input signal $x_i[n]$ is calculated according to:

$$x_i[n] = \begin{cases} x[n] & \text{if } i\,(N_{\text{frames}} - N_{\text{OL}}) + 1 \leq n \leq (i+1)\,N_{\text{frames}} - iN_{\text{OL}}, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{18.4}$$

The corresponding output signal $y_i[n]$ is achieved by a basic filtering with the current filter $h[n]$:

$$y_q[n] = x_i[n] * h[n]\,. \tag{18.5}$$

The output signal is derived according to:

$$y_{\text{output},q}[n] = \begin{cases} y_q[n] & \text{if } 0.5 \cdot N_{\text{OL}} + 1 \leq n \leq N_{\text{frames}} - 0.5 \cdot N_{\text{OL}}, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{18.6}$$

Finally $y_{\text{output},q}[n]$ is directly sent to the function "playrec('play')".

Last step:
The block-processing has to be interrupted for $i > i_{\max} \leq \frac{N_{\text{signal}}}{N_{\text{frame}} - N_{\text{OL}}}$, where N_{signal} is the length of the input signal in samples. $i = i_{\max} + 1$ is the last step which can be performed for the block-processing. The input signal is calculated according to:

$$x_{i=i_{\max}}[n] = \begin{cases} x[n] & \text{if } i\,(N_{\text{frames}} - N_{\text{OL}}) + 1 \leq n \leq N_{\text{signal}}, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{18.7}$$

The corresponding output signal $y_i[n]$ is achieved by a basic filtering with the current filter $h[n]$:

$$y_q[n] = x_i[n] * h[n]\,. \tag{18.8}$$

The output signal is derived according to:

$$y_{\text{output},q}[n] = \begin{cases} y_q[n] & \text{if } 0.5 \cdot N_{\text{OL}} + 1 \leq n \leq N_{\text{signal}} - i\,(N_{\text{frames}} - N_{\text{OL}})\,, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{18.9}$$

Finally the last output frame $y_{\text{output},q}[n]$ is directly sent to the function "playrec('play')".

Block-processing interruption:
To achieve a quasi real-time effect on the output signal due to a certain change of the filter an automatic interruption of the block-processing has to be achieved. For this purpose the function "playrec('isFinished')" is a suitable choice, as the playback status of a certain output frame can be requested. For several processing steps $i \geq 3$ this function is used to get a status request of the second to the last output frame $y_{\text{output},q[i]=(i-1) \bmod 3-1}[n]$. The current frame i will not be processed until the playback of the output frame $y_{\text{output},q[i]=(i-1) \bmod 3-1}[n]$ is finished. On the one hand this approach avoids a conflict with the next output frame $y_{\text{output},q[i]=(i-1) \bmod 3+2}$, because the identic memory space is used. On the other hand the approach leads to a quasi real-time effect on the output signal due to a change of the filter with perceivable latency but with non-stop playback.

Part V.

RoLoSpEQ – Room and LoudSpeaker EQualizer

RoLoSpEQ (**Ro**om plus **Lo**ud**Sp**eaker **EQ**ualizer) is a software based tool, which was developed during this thesis.

RoLoSpEQ aims at an analysis and illustration of room plus loudspeaker responses and an improvement of the sound reproduction behavior of loudspeakers in small rooms, which is also termed room equalizing.

RoLoSpEQ enables a quasi real-time simulation of several calculated room equalizing filters. The simulation is based on the developed block-processing approach, which is discussed in more detail in chapter 18.2. Furthermore a special technique for an automatic calculation of target frequency responses is developed. This approach is discussed in detail in chapter 21. RoLoSpEQ is completely implemented in MATLAB [Mat08] and is realized with a graphical user interface (GUI), which was created with the native MATLAB function "GUIDE" (Graphical User Interface Development Environment). The complete theory, which is necessary for the implementation of RoLoSpEQ is discussed in part I - IV.

In this part the software GUI of RoLoSpEQ is presented. The different program sections, which are important for practical use are explained in detail. Subsequently the developed target generation approach is discussed. At the end of this part a list of several features of RoLoSpEQ can be found.

19. Hardware and Software Connection

In figure 19.1 the basic wiring of the hardware is illustrated. A soundcard is connected

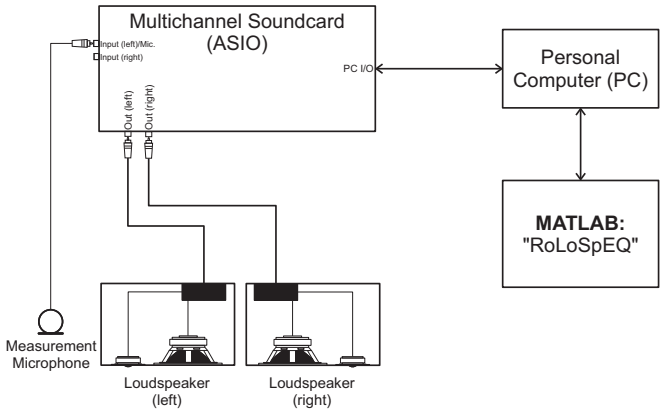

Figure 19.1.: RoLoSpEQ: basic connection of the hardware and the software. RoLoSpEQ is limited to a stereo loudspeaker system.

to the PC. The soundcard needs to have in minimum two input and output channels to use RoLoSpEQ. A measurement microphone has to be plugged in one input channel. The speakers are applied to the soundcard output channels.

The connection between software and hardware is enabled by RoLoSpEQ using "Playrec" (see chapter 18.1).

For the practical use of RoLoSpEQ a standard personal computer (PC) is sufficient.

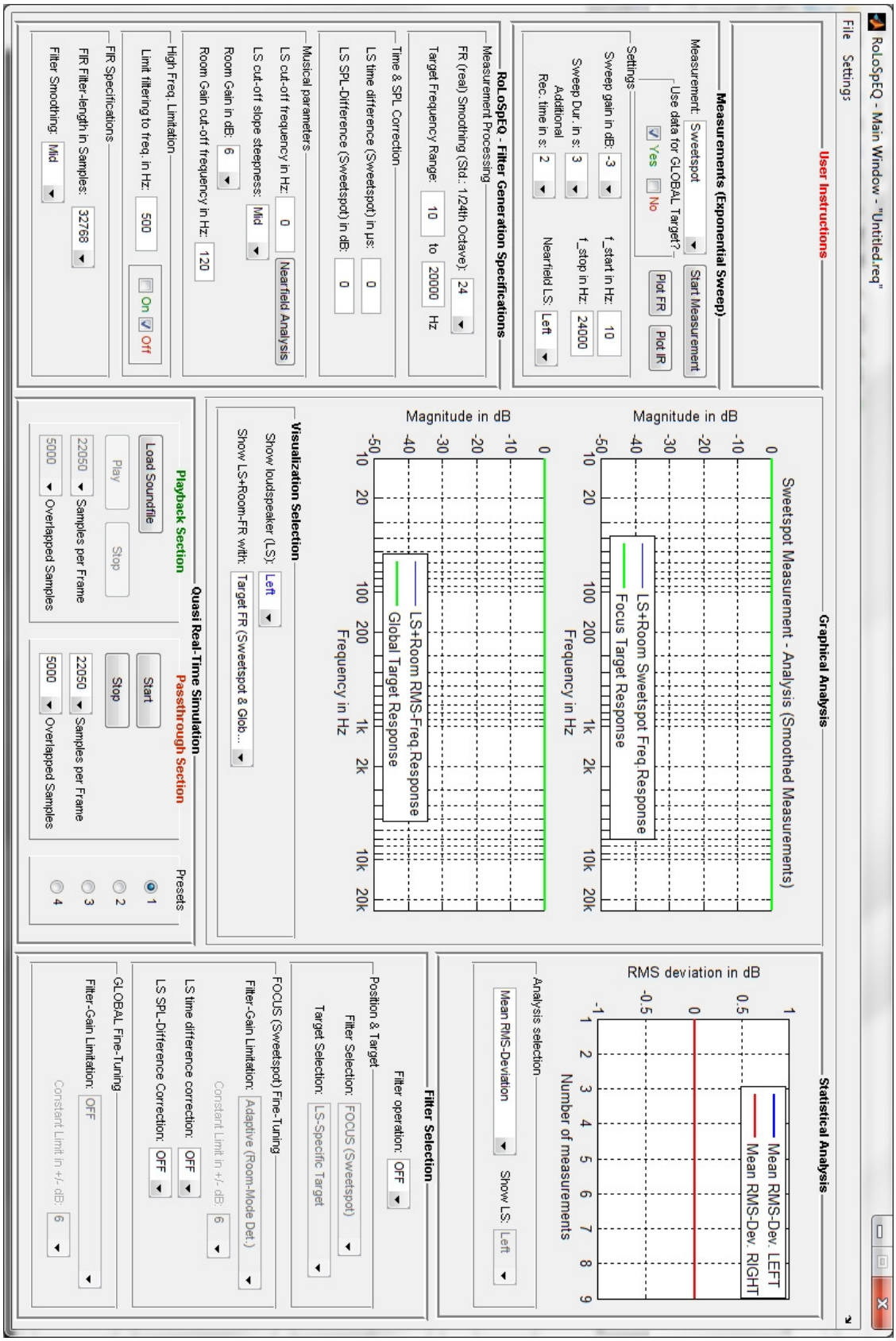

Figure 19.2.: RoLoSpEQ Main Window – GUI after starting the program.

20. Software GUI – Main Window

In figure 19.2 the main window of RoLoSpEQ is shown. Further program features are available in the main menu with the menu items *File* and *Settings*. In figure 20.1 the submenus of RoLoSpEQ are shown.

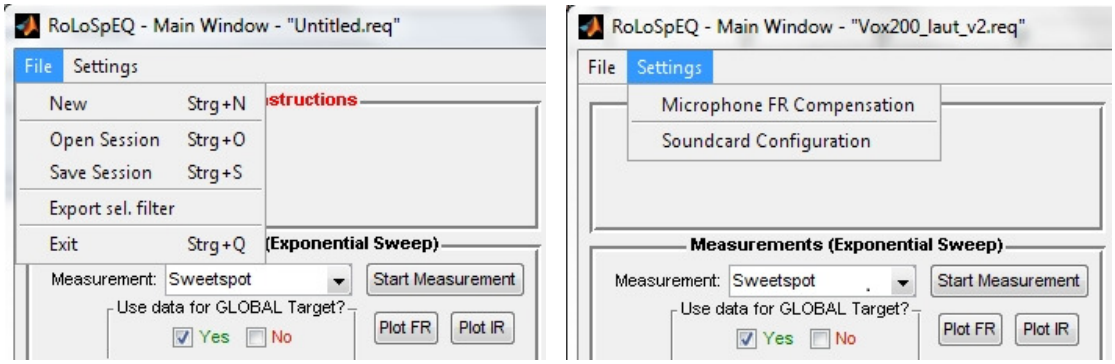

Figure 20.1.: RoLoSpEQ: Main menu selection opportunities. A complete section can be saved; Current selected filters can be exported as wave files for the use in other software solutions; Soundcard configurations can be changed (see chapter 20.1.1); Microphone frequency responses can be loaded (see chapter 20.1.2).

The main window is subdivided in program sections, where several equalizer specifications can be chosen. In figure 20.2 the internal data flow of RoLoSpEQ between different program sections is illustrated in correspondence with the main window structure.

In more detail the internal relationships can be explained with the following list:

- Measurements (Exponential Sweep).
  - Soundcard Configuration.
  - Microphone Frequency Response (FR) compensation.
- RoLoSpEQ – Filter Generation Specifications.
  - Near-Field Analysis.
- Graphical Analysis.
  - Visualization Selection.
- Filter Selection.
- Quasi Real-Time Simulation.
  - Filter Presets.
- Statistical Analysis.

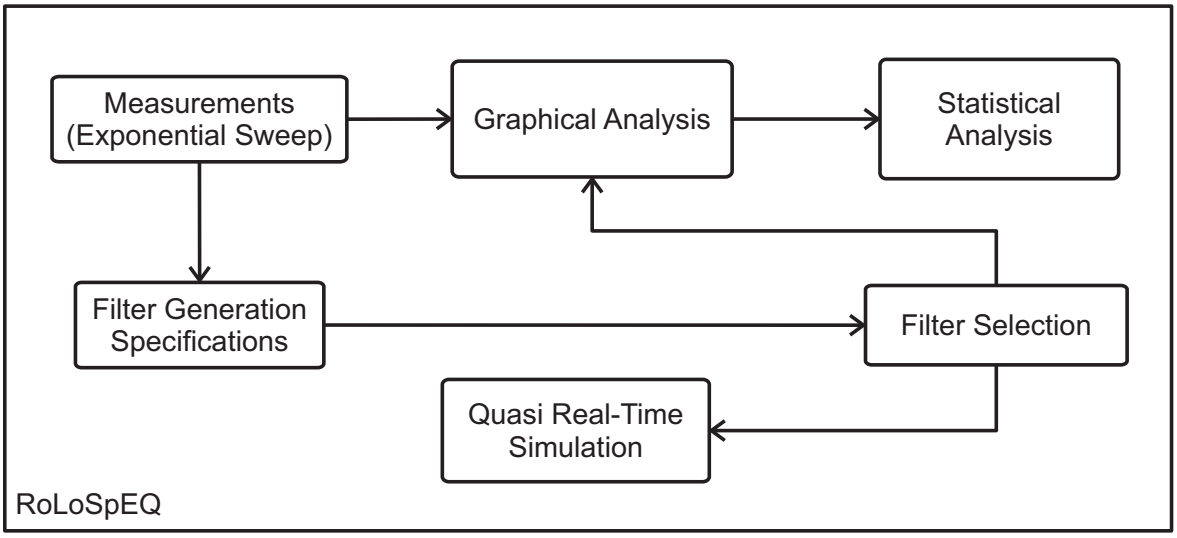

Figure 20.2.: RoLoSpEQ: Internal data flow graph of the main window between different program sections.

Particular program sections are discussed according to this sequence in the following paragraphs.

In general most specifications within the program sections are predefined to a standard value and do not have to be changed for a normal use of the software tool. Nevertheless several values can be adjusted according to personal objectives.

The program sections will be explained in the following paragraphs. For this purpose the main important specifications, which are useful for practical use of RoLoSpEQ are discussed.

20.1. Measurements (Exponential Sweep)

In figure 20.3 the measurement section is shown.

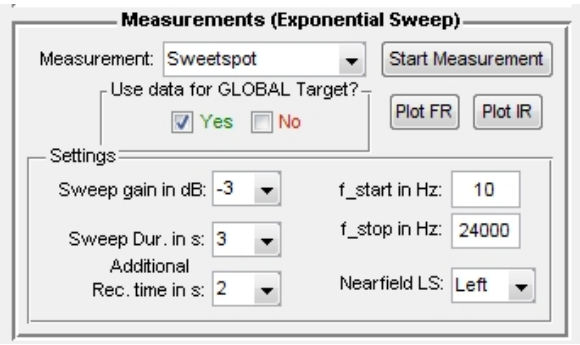

Figure 20.3.: RoLoSpEQ: Measurement (Exponential Sweep). In the drop-down list several measurement positions can be selected. Additionally certain measurements can be activated for global RMS averaging ("Use data for global target?"). The measurement can be started with the button *Start measurement*. After each measurement the measured data set (left and right loudspeaker) is automatically activated for global RMS averaging and the measurement selector is automatically increased to the next measurement position.

Several measurements, which are necessary for further calculation, can be performed in this section. Different measurement positions are available in the drop-down list. At the beginning of the measurement process the sweet spot (FOCUS) position is automatically selected. In total one sweet spot measurement and eight randomly chosen room positions are available. In order to ensure an accurate target response calculation (see chapter 20.2.1), additionally in maximum four near-field measurements, which can be used for further near-field analysis, can be performed .

The measurements can be started with the button *Start Measurement*. If the soundcard is not initialized, the soundcard configuration window will appear (see chapter 20.1.1). Otherwise the first measurement will be performed directly and stored for the selected microphone position (e.g. sweet spot). At first the left loudspeaker, then the right loudspeaker is used for the measurement. After the measurement process is finished, the drop-down menu will be automatically set to the next measurement position and the previous measurement data is enabled for use in the filter generation processing. Subsequently the measurement positions 1-8, which have to be randomly distributed within the room, have to be performed. It is recommended to perform all nine measurements to get an accurate filter generation with RoLoSpEQ.

If one measurement set is not required for further analysis, it can be disabled in the section *"Use data for global Target?"*. However the sweet spot measurement cannot be disabled, because in minimum one measurement is necessary for an accurate filter generation.

20.1.1. Soundcard Configuration

Figure 20.4 illustrates the soundcard configuration window, which is opened either automatically if the soundcard is required in RoLoSpEQ or with the menu item *Settings* in the main menu (see figure 20.1). At the top site the host API (see chapter 18.1) can be chosen. If an

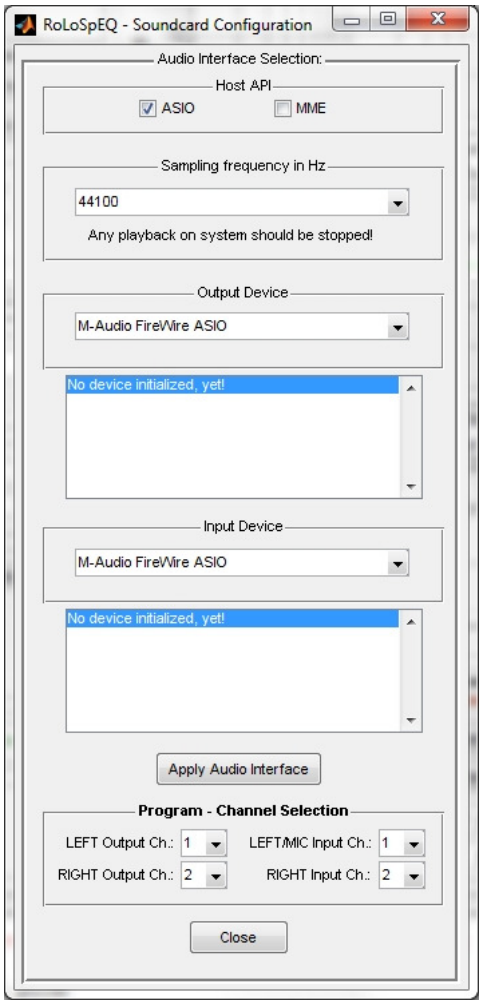

Figure 20.4.: RoLoSpEQ soundcard configuration window.

ASIO soundcard is found, the ASIO check-box is automatically enabled. If no ASIO device can be found merely MME can be selected.

Subsequently, if ASIO is the host API the sampling frequency f_s can be defined. Three different choices are available: $f_s = 44.1kHz$, $f_s = 48kHz$ and $f_s = 96kHz$. If MME is the host API the sampling frequency is fixed to $f_s = 44.1kHz$.

In the drop-down menus for the input device and the output device several identified audio devices can be selected, according to the enabled host API. If ASIO is enabled multichannel ASIO devices will be available there. After the desired audio devices are selected the choice can be applied with the button *Apply Audio Interface*. Further information about the selected devices are illustrated in the two lists below the device selectors in the soundcard

configuration GUI.

Finally the channel mapping can be adjusted to the personal audio environment. Therefore two output channels (left and right channel) and two input channels (left and right channel) can be assigned to arbitrary channels (stereo setup).

20.1.2. Microphone Frequency Response (FR) Compensation

In figure 20.5 the microphone frequency response (FR) compensation window, which can be opened with the menu item *Settings* in the main menu (see figure 20.1) is shown.

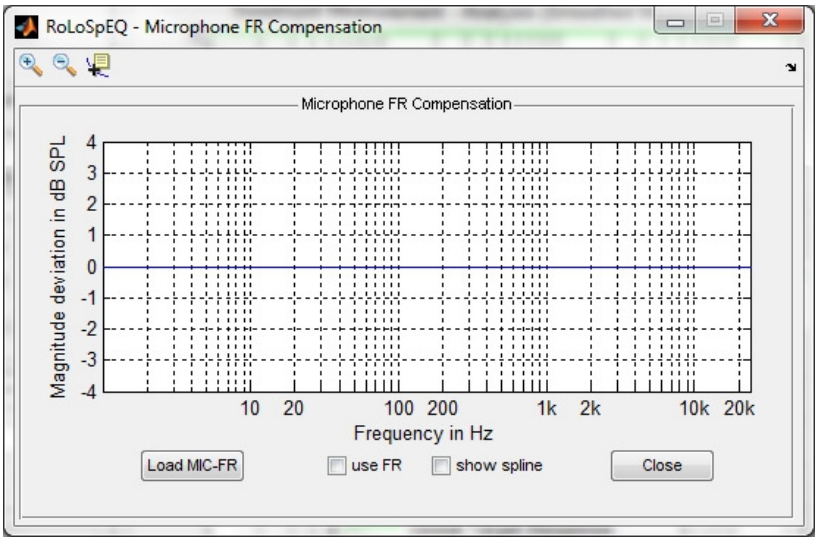

Figure 20.5.: RoLoSpEQ microphone frequency response compensation window.

In RoLoSpEQ the microphone FR compensation, which is derived in the measurement section is used in order to optimize the measurement data (see chapter 20.1).

With the button *Load MIC-FR* a standard microphone calibration file, which is for standard a two column ".txt" file can be chosen. The ".txt" ending should be exchanged by ".mic" before loading. After loading, the microphone frequency response is shown in the plot (see figure 20.5).

In addition the interpolation (see chapter 17) can be shown by activating the check-box.

If the microphone frequency response compensation shall be used for measurement analysis, the check-box *use FR* has to be active.

20.2. RoLoSpEQ – Filter Generation Specifications

In figure 20.6 the filter generation specification section is shown.

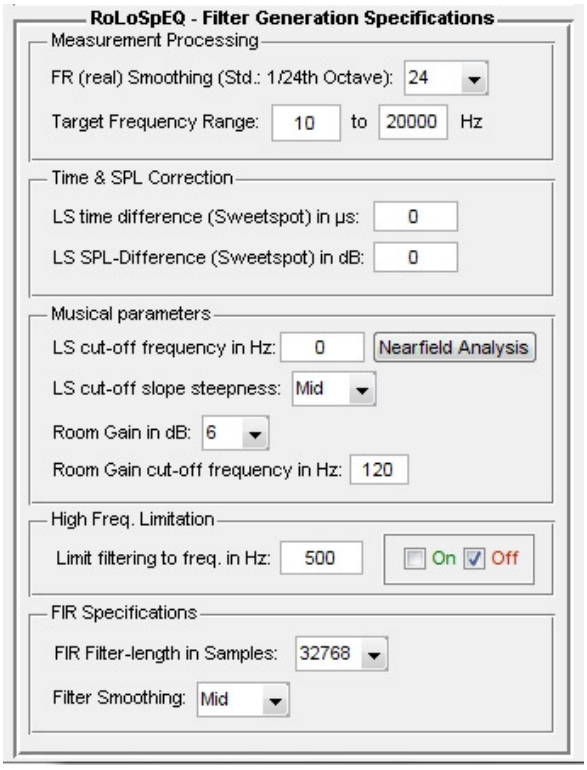

Figure 20.6.: RoLoSpEQ: Filter Generation Specifications.

A couple of important definitions have to be setup in this section to enable a calculation of accurate room equalizing filters with RoLoSpEQ:

- The **Target Frequency Range**, is the frequency range of the measurement data, which is used for the calculation of target frequency responses. Thus this frequency range should be defined to a realistic range according to the used loudspeakers. E.g. if large pillar loudspeakers are used the standard frequency range does not have to be changed.

- The **Time & SPL Correction** is automatically derived with the FOCUS (sweet spot) measurement data set. Thus these values do not have to be changed in practice. The time and SPL correction can be activated later in the *filter selection section* (see chapter 20.4).

- The **Musical Parameters** are very important for the consideration of specific human hearing habits (see chapter 4.2.4). Two different definitions have to be setup for an accurate target response generation:
  - The **Near-Field Analysis** is used to emulate the high-pass characteristic of loudspeakers in order to get accurate target responses. Therefore a suitable cut-off

frequency has to be selected either manually or with the near-field analysis tool (see chapter 20.2.1). Furthermore the slope steepness has to be manually chosen in the drop-down list according to the slope steepness, which can be observed in the *graphical analysis section* (see chapter 20.3).

– The **"Room Gain"** is used in RoLoSpEQ to preserve a natural timbre in small listening rooms. The user can adjust these values according to her/his own preferences, e.g. during listening to an audio file with the *Quasi Real-Time Simulation* feature in RoLoSpEQ (see chapter 20.5).

- The equalizer **High Frequency Limitation** aims at an limitation of the calculated room equalizing filters in order that high frequencies are not affected by the room equalizer. If the limitation is activated the effect can be observed in the *graphical analysis* section as well (see chapter 20.3). For this purpose the *filter visualization* has to be adjusted (see chapter 20.3.1).

- The **FIR Specifications** can be defined according to personal preferences as well. However the FIR filter length and the FIR filter smoothing can be adjusted, which can be observed in the graphical analysis section as well. Furthermore these specifications are also applied to the exported filters (see figure 20.1).

20.2.1. Near-field Analysis

Figure 20.7 illustrates the near-field analysis window, which can be opened with the button *Near-Field Analysis* in the filter generation specification section (see chapter 20.2). The

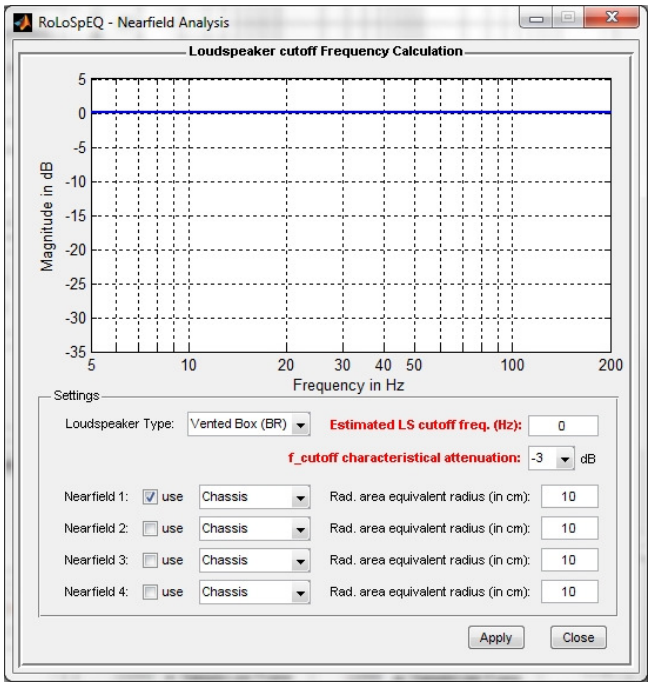

Figure 20.7.: RoLoSpEQ nearfield analysis window.

near-field analysis window enables analysis of near-field loudspeaker measurements, which are performed in the measurement section in the main window before.

The analysis aims at the calculation of a general loudspeaker cut-off frequency, which is necessary for an adequate loudspeaker target response calculation (see chapter 21).

Basically the loudspeaker type has to be defined in the drop-down menu. Available options are *Vented Box (BR)* or *Closed Box*.

In maximum four near-field measurements can be performed in the measurement section in the main window (see chapter 20.1). These measurements can be selected for near-field analysis by activating the check-box *use* in the near-field analysis window.

If a vented box is used, in addition the radiator type has to be chosen in a drop-down menu. Available are *Chassis* or *Bassreflex port*. A correct choice is necessary for an accurate near-field analysis.

Furthermore the radius of the equivalent radiation area of the radiator has to be defined. This is an important value for a correct scaling of the measurement data.

Finally the loudspeaker cut-off frequency can be individually chosen with a drop-down menu according to the selected attenuation between $-10dB$ and $-3dB$. This value is dependent on the used loudspeaker and should be optimized manually, in order to achieve a proper target response.

The calculated cut-off frequency can be confirmed with the button *Apply*. Otherwise the window can be closed without any change in the main window.

20.3. Graphical Analysis

In figure 20.8 the graphical analysis section is illustrated. In the standard configuration (see

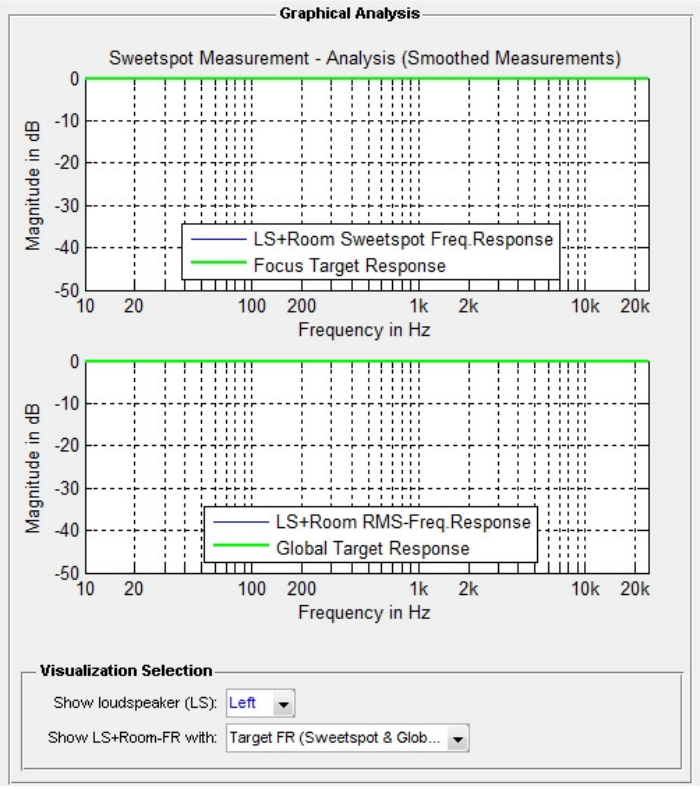

Figure 20.8.: RoLoSpEQ: Graphical Analysis section.

chapter 20.3.1) the top plot shows the FOCUS (sweet spot) measurement and the bottom plot shows the global average (multiple measurements should be available and activated for global averaging – see chapter 20.1). For this purpose measurement data has to be available. If new response measurements are performed in RoLoSpEQ (see chapter 20.1), the graphical analysis plots are automatically updated.

The contents of the plots in the graphical analysis section can be changed with the *visualization selectors* (see chapter 20.3.1).

20.3.1. Visualization Selection

In figure 20.9 the visualization selection, which is part of the graphical analysis section in RoLoSpEQ is illustrated.

The following visualization options are available:

- **Show Loudspeaker (LS)**: left or right. Here, the data set, which is used for the visualization, can be chosen. Thus either the measurements of the left loudspeaker or of the right speaker can be selected.
- **Show Loudspeaker (LS) plus Room Frequency Response (FR)**: On the one

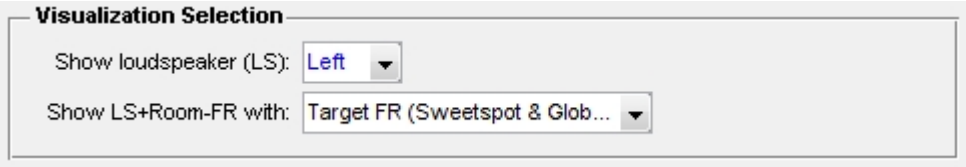

Figure 20.9.: RoLoSpEQ: Visualization Selection.

hand the visualization depends on the current selected filter (see chapter 20.4). On the other hand the visualization is dependent on the following selections:

- **Target FR (Sweet Spot & GLOBAL)**: This is the standard configuration. The top plot shows the FOCUS (sweet spot) measurement with calculated target response (green) and the bottom plot shows the global average with calculated target response (green) (multiple measurements should be available and activated for global averaging – see chapter 20.1).
- **Target & Room Equalizing Filter Frequency Response (FR)**: The top plot shows either the FOCUS (sweet spot) measurement or the GLOBAL RMS average with corresponding target responses. The bottom plot shows the selected filter, which is calculated according to the definition setup in the filter generation specification section (see chapter 20.2).

  However, if filtering is switched-off merely a flat filter will be observed in the bottom plot. Hence the filter visualization is dependent on the current filter selection (see chapter 20.4).
- **Room Equalizing Filter Simulation**: This mode can be used for the graphical simulation of the generated (see chapter 20.2) and selected (see chapter 20.4) room equalizing filter. The top plot shows either the FOCUS (sweet spot) measurement or the GLOBAL RMS average with corresponding target responses. The bottom plot shows the same data again, but the currently selected filter is used to simulate the magnitude response after room equalization.

  This simulation is merely based on magnitude responses. However an evaluation of the calculated filters is possible. Furthermore the filter specifications (see chapter 20.2) can be optimized according to the shown simulation, which is automatically updated if changes are performed.

20.4. Filter Selection

In figure 20.10 the filter selection section is illustrated.

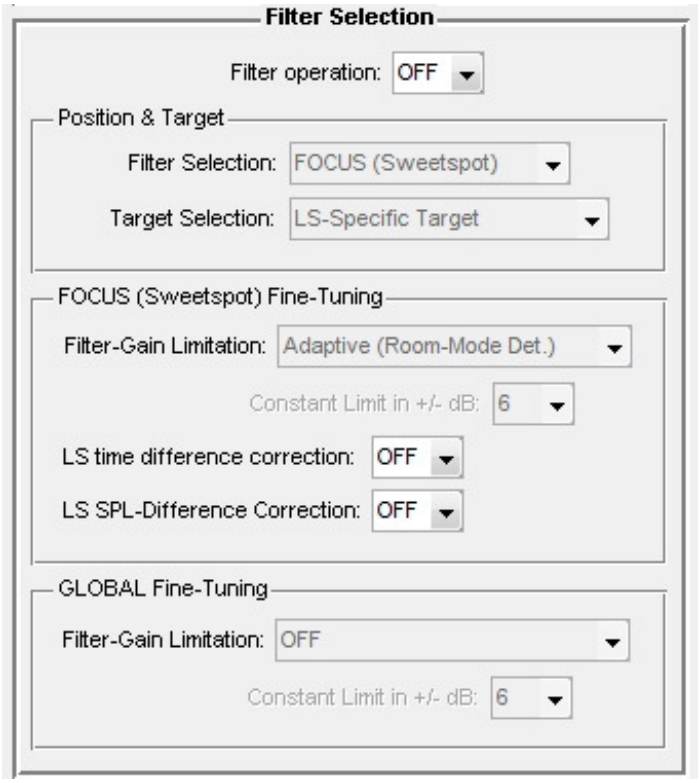

Figure 20.10.: RoLoSpEQ: Filter Selection. Different filters can be selected. Beyond others the adaptive room mode detection (ARD) can be used for FOCUS equalizing (see chapter 4.2.3).

The following filters selections are available:

- **Filter operation ON/OFF:** The room equalizing can be switched on or off here.
- **Position & Target:** The listening position and the used target response can be selected.
  - **Filter Selection:** The FOCUS filter is chosen for an equalization in the listening sweet spot. The GLOBAL filter is chosen for a global equalization.
  - **Target Selection:** The target response can be selected. Four different options are available:
    - **Loudspeaker specific target:** One specific target is calculated for both loudspeakers.
    - **Mean target:** An average of both specific targets is calculated, aiming at an timbral adjustment of the used loudspeakers.
    - **Use left LS target:** The specific target response calculated for the left loudspeaker is used for both room equalizers (left and right channel).

    - **Use right LS target:** The specific target response calculated for the right loudspeaker is used for both room equalizers (left and right channel).

- **FOCUS (sweet spot) fine tuning:** can be chosen if the FOCUS filter is selected.
  - **Filter gain limitation:** The sweet spot filter can be limited with different approaches to avoid an unsteady behavior of the sweet spot equalizer (see chapter 4.2.3).

    Additionally a time and SPL correction can be used to correct the image source position (see chapter 3.2.1).

    - **Off:** No filter limitation is used[1].
    - **Adaptive Room Mode Detection (ARD):** This approach is described in more detail in chapter 4.2.3.
    - **Constant** limitation: A constant filter limitation can be chosen within a range between $-12dB - 0dB$.

- **GLOBAL fine tuning:** can be chosen if the GLOBAL filter is selected.
  - **Filter gain limitation:** In the case of GLOBAL filters the limitation can be switched off or a constant limitation can be chosen within a range between $-12dB - 0dB$.

In general the filter limitation is forced by the FIR filter smoothing, which can be changed in the filter generation specification section (see chapter 20.2).

The filter selection has a direct impact on the graphical analysis section (see chapter 20.3) and the quasi real-time simulation section in RoLoSpEQ (see chapter 20.5).

20.5. Quasi Real-Time Simulation

In figure 20.11 the quasi real-time simulation section is illustrated. The quasi real-time simulation is based on a special block-processing approach, which was developed during this thesis. This approach is discussed in more detail in chapter 18.2.

The quasi real-time simulation section enables a simulation of selected room equalizing filters in RoLoSpEQ. The signal processing is performed in quasi real-time, that is playback and filtering are executed at the same time with reasonable time latency.

The benefit of this approach is the possibility of a direct filter change during audio signal playback or passthrough. Therefore it is possible to change several specifications in RoLoSpEQ whereas the simultaneous change of the output signal enables an immediate optimization.

In the quasi real-time simulation section two different simulation options can be chosen:

- **Signal Playback:** A stereo wave or mp3 file can be loaded with the button *Load Soundfile*. After loading is complete the frame size and the frame overlapping can be chosen (see chapter 18.2). These are important definitions, which are predefined to a suitable value for a quasi-real time simulation.

[1] However a hard limiting at $+6dB$ is always used in RoLoSpEQ in order to avoid a signal clipping. Hence a signal amplification is merely possible below $+6dB$.

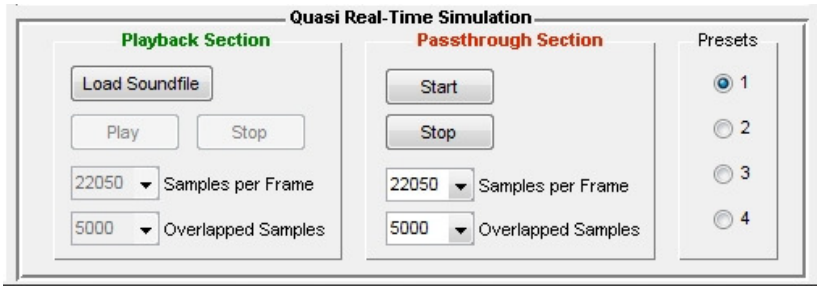

Figure 20.11.: RoLoSpEQ: Quasi Real-Time Simulation. The current selected filter (see chapter) can be simulated in quasi real-time. Either a playback of an audio signal or a passthrough of an external audio source, which is plugged in at the soundcard input channels can be used for the simulation. Additionally four global presets are selectable (see chapter 20.5.1).

The button *Play* starts the simulation of the selected room equalizing filters.

- **Signal Passthrough:** A stereo sound-source can be applied to the input channels of the soundcard (see figure 19.1). The button *start* enables the filter simulation of the input signal, which is recorded and filtered at the same time. The input and output channels are assigned in the soundcard configuration window (see chapter 20.1.1).

20.5.1. Filter Presets

In the quasi real-time simulation section another section, which is called **Presets** can be found (see figure 20.11).

Here, in total four different preset configurations are adjustable for simulation and the specification of several filters. Hence in total quasi four different main window surfaces are available in RoLoSpEQ, which can be switched with this preset selector.

Therefore multiple filter configurations can be defined in RoLoSpEQ and can be directly invoked with the preset selector.

This feature aims at a comfortable and effective filter switching whilst quasi real-time simulation of the room equalizing filters is performed. The presets enable a simple and fast comparison option for different room equalizing filter configurations. This is a helpful tool for the subjective evaluation of generated room equalizers.

20.6. Statistical Analysis

In figure 20.12 the statistical analysis section is illustrated. This section is an additional tool for statistical measurement data analysis. This analysis is used in order to find an automatic quality criterion for the validity of the measurement data in order to economize the measurement process and to save measurement repetitions.

The following options are available:

- **Mean RMS deviation:** The averaged RMS deviation in $[dB]$ is calculated according

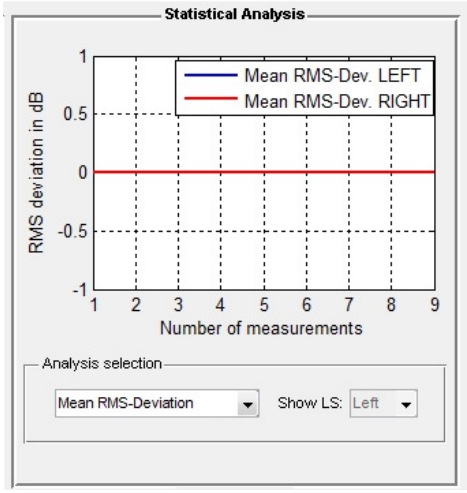

Figure 20.12.: RoLoSpEQ: Statistical Analysis.

to [Ped07]:

$$P_{\mathrm{rms,M}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|20\cdot\log_{10}\frac{P_{\mathrm{rms,M}}\left(f\left(i\right)\right)}{P_{\mathrm{rms,ref}}\left(f\left(i\right)\right)}\right|^{2}}, \tag{20.1}$$

where M is the number of used measurements, i is the i^{th} frequency bin of in total N frequency bins, f is the frequency in $[Hz]$ and P is the sound pressure in $[Pa]$. $P_{\mathrm{rms,M}}$ is calculated according to equation (3.1) for a desired number of measurements, whereas $P_{\mathrm{rms,ref}}$ is a reference calculation using all available measurements.

With equation (20.1) the descending of the averaged RMS deviation can be illustrated.

- **Mean Standard Error (SE):** The mean standard error is considered to yield a prediction of the measurement quality without a priori knowledge. According to the basic calculation (see [Bar01]) the calculation in RoLoSpEQ is performed according to:

$$\mathrm{SE}_{\overline{P}} = \frac{s}{\sqrt{M}} = \frac{1}{N}\sum_{i=1}^{N}\left\{\sqrt{\frac{M}{M-1}\sum_{n=1}^{M}\left(P_n\left(f\left(i\right)\right)-\overline{P}\left(f\left(i\right)\right)\right)^{2}}\right\}, \tag{20.2}$$

$$\overline{P}\left(f\right) = \frac{1}{M}\sum_{n=1}^{M}P_n\left(f\left(i\right)\right), \tag{20.3}$$

where s is the standard deviation, M is the number of considered measurements, i is the i^{th} frequency bin of in total N frequency bins, f is the frequency in $[Hz]$ and P is the sound pressure in $[Pa]$.

The approach for the statistical analysis was not further pursued during this thesis. However further investigations could lead to a reduction of required measurement repetitions, which is desirable if the average quality (e.g. the average RMS deviation) cannot be increased with further measurements. For this purpose a reliable quality criterion could be developed.

21. Automatic Generation of Target Responses

An automatic target generation approach was developed during this thesis and is described in the following paragraphs. The target generation is based on up to nine measurements (GLOBAL target) or on one sweet spot measurement (FOCUS target).

The approach aims at an optimal overall matching of the target response with the basic measurement data set. The objective is a minimization of influences on the native loudspeaker characteristics, because mainly room influences are corrected (e.g. adjacent boundary effects).

The derivation of the GLOBAL target response and the FOCUS target response are mainly similar.

21.1. The GLOBAL Target Response (Multi-Position)

The basic measurement data for the GLOBAL target response is the global RMS average, which is an accurate predictor of the loudspeaker reproduction behavior in the room. This averaging method is discussed in more detail in chapter 3.2.4.

First step:
At the first step of the target generation, two regression lines are used to emulate the high frequency behavior of the room plus loudspeaker response. One regression line is calculated for the frequency range between $500Hz - 10kHz$ and another one above $10kHz$, which considers the high frequency slope. The measurement data and the two regression lines are illustrated in figure 21.1.

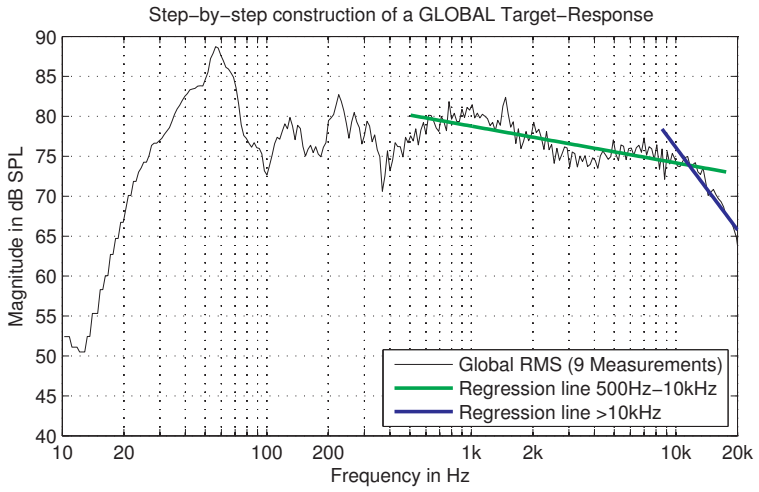

Figure 21.1.: RoLoSpEQ: GLOBAL target response calculation – Step 1.

Second step:
The mean SPL is calculated for a frequency between $400 - 600Hz$. At this frequency region interferences on the measurement data caused by isolated room modes or adjacent boundary effects can be ruled out (see figure 3.5). The new average value is illustrated in figure 21.2. This value characterizes the sensitivity of the used loudspeaker system.

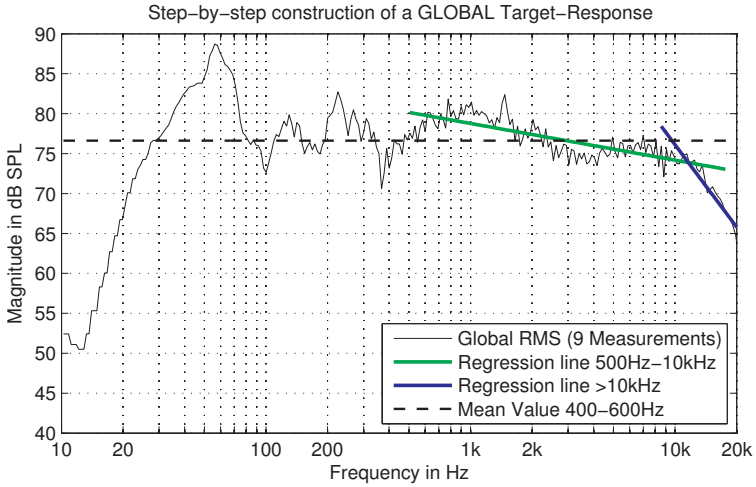

Figure 21.2.: RoLoSpEQ: GLOBAL target response calculation – Step 2.

Third step:
In the third calculation step the two high frequency regression lines and the average line are bond together in their intersection points. The resulting curve is illustrated in figure 21.3. The used intersection points for this measurement are approximately at $3kHz$ and $11kHz$.

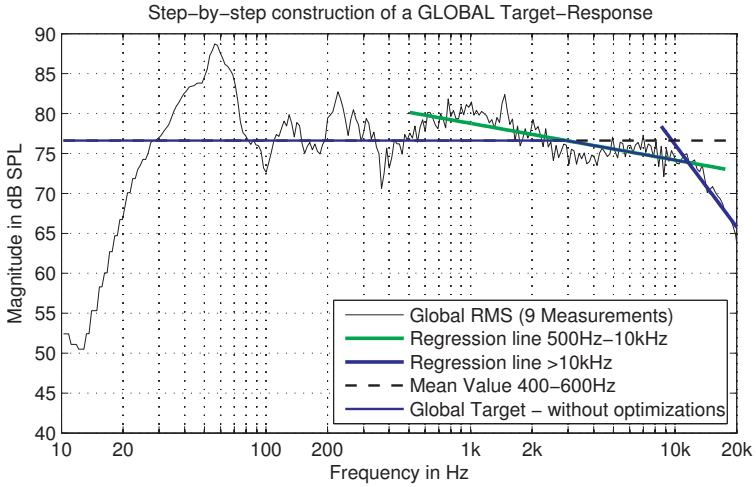

Figure 21.3.: RoLoSpEQ: GLOBAL target response calculation – Step 3.

Fourth step:
In the fourth calculation step a second order Butterworth high-pass filter behavior is applied to the target response to emulate the high-pass behavior of the loudspeaker. For this purpose a general high-pass magnitude response is multiplied with the target response. The high-pass optimized target response is illustrated in figure 21.4.

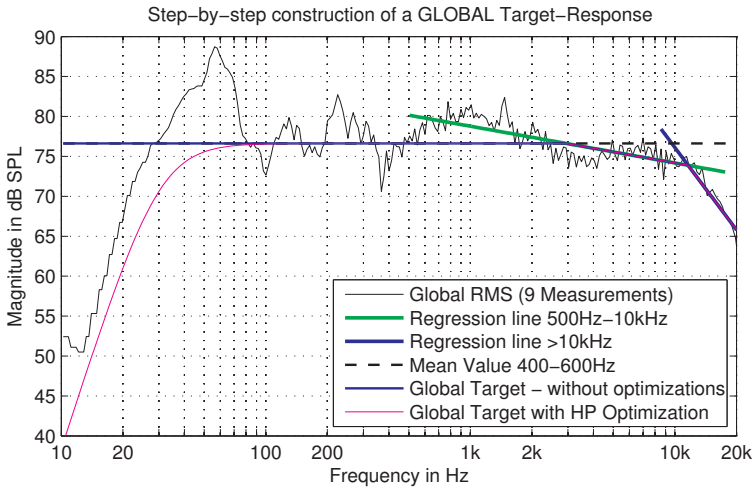

Figure 21.4.: RoLoSpEQ: GLOBAL target response calculation – Step 4.

Fifth step:
In the fifth calculation step the "Room Gain", which aims at a natural sound reproduction in the small listening room according to human hearing habits is applied. For this purpose a

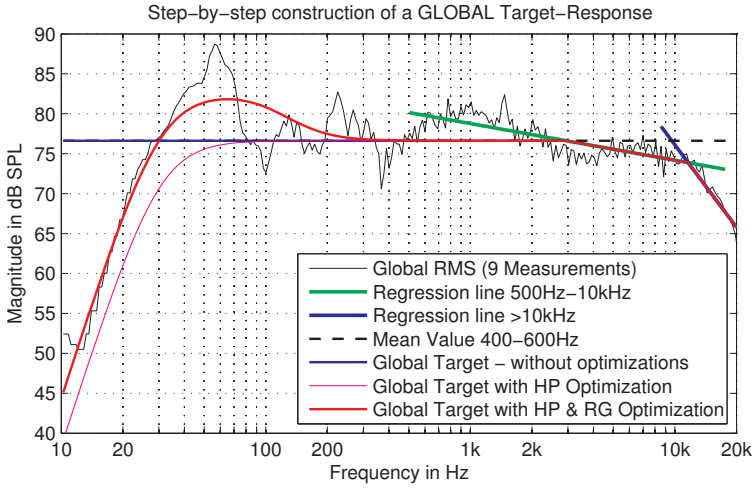

Figure 21.5.: RoLoSpEQ: GLOBAL target response calculation – Step 5.

low-shelving filter is applied to the target response. The resulting GLOBAL target response is illustrated in figure 21.5.

The low-shelving filter is a parametric filter, which is calculated according to equation (10.4). For this target response a $120Hz$ cut-off frequency and a $6dB$ gain is chosen.

The "Room Gain" is discussed in more detail in chapter 4.2.4. The optimal specifications for the room gain highly differ for different listeners, loudspeakers and rooms. The definition of the "Room gain" is one of the most important specification for a natural room equalizing with RoLoSpEQ. Thus the "Room Gain" is termed *musical parameter* in the main window and can be adjusted to the listener's preferences (see figure 19.2).

Summarizing:
The developed approach for an automatic calculation of a GLOBAL target response leads to suitable emulations of the shape of the basic measurement. The basic measurement is a global RMS average of nine different measurements, which yields an accurate prediction of the loudspeaker reproduction behavior in the room (see chapter 3.2.4). Certain timbral influences can be diminished using a GLOBAL equalizer, which is calculated according to equation (4.6).

21.2. The FOCUS Target Response

The basic measurement data for the FOCUS target response is one measurement in the listening sweet spot position. Therefore an unsteady frequency response behavior can be observed. Nevertheless the calculation process for the FOCUS target response is mainly similar to the calculation process of the GLOBAL target response, which is described in the previous paragraph.

The resulting filter, which is calculated with the generated FOCUS response has a very unsteady behavior, dominated by high peaks and dips. However after the basic filter calculation an optimization process is used to achieve a practical FOCUS filter, which is described in more detail in chapter 4.2.3.

The calculation process of the FOCUS target response differs mainly in the calculation of the average value between $400Hz - 600Hz$, which characterizes the sensitivity of the loudspeaker (see chapter 21.1 step 2). During the development of the approach it has turned out, that the usage of the SPL average of the GLOBAL target response yields more suitable results.

Hence in the second calculation step of the FOCUS target response, this GLOBAL SPL average is used instead of the FOCUS SPL average.

Fifth step:
In figure 21.6 the complete FOCUS target response generation process including several optimizations is illustrated.

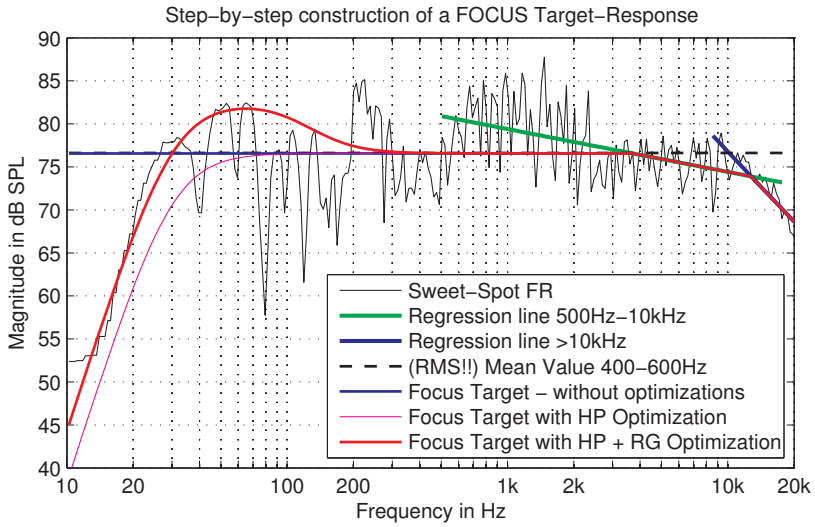

Figure 21.6.: RoLoSpEQ: FOCUS target response calculation – Step 5.

22. RoLoSpEQ Feature Summary

A list of the features of RoLoSpEQ:

- Graphical user interface implementation.
- ASIO and MME driver support. Flexible channel routing (stereo) for multichannel soundcards.
- Stereo loudspeaker (LS) system room equalizing.
- Room response measurements:
  - Automatic selection of measurement positions.
  - Smoothing of room response measurements.
  - Automatic averaging of room response measurements.
  - Near-field measurement and near-field analysis.
  - Room response analysis and visualization (RMS deviation and Standard Error (SE)).
- Four different presets can be defined in the *Real-Time Simulation* section.
- Automatic room equalizing filter generation:
  - Target response specifications (frequency range, LS cut-off calculation with near-field measurements, musical parameter ("Room Gain")).
  - Equalizing filter frequency limitation.
  - Definition of the FIR filter length.
  - FIR filter smoothing.
- Filter Simulation selector:
  - Filter operation: on/off.
  - Filter selector: FOCUS (sweet spot) or GLOBAL equalizing.
  - Target response selector:
    - LS-specific target calculation.
    - Mean Target target calculation (left plus right loudspeaker (LS).
    - Use left LS response for target generation.
    - Use right LS response for target generation.
  - FOCUS Equalizing:
    - Switch off (without optimization).

        * Adaptive Room Mode Detection.
        * Constant filter limitation.
        * LS time and SPL difference correction.
    - GLOBAL Equalizing:
        * Filter limitation on/off.

- Quasi real-time simulation of the room equalizer:
    - Signal playback.
    - Signal passthrough.

- Visualization:
    - Measurement visualization (sweet spot and global (RMS) average).
    - Room equalizing filter visualization according to selected filter.
    - Room equalizing filter simulation.

Part VI.

D-MLCNC – Digital Multichannel Loudspeaker Crossover Network

D-MLCNC (**D**igital – **M**ultichannel **L**oudspeaker **C**rossover **N**etwork **C**ontroller) is a software tool which was developed during this thesis.

D-MLCNC aims at a quasi real-time simulation of a digital eight channel crossover network. The software is developed for a stereo loudspeaker system, thus two input channels or a stereo sound file are supported for the simulation. D-MLCNC is completely implemented in MATLAB [Mat08] and is realized with a graphical user interface (GUI), which was created with the native MATLAB function "GUIDE" (Graphical User Interface Development Environment).

The basic theory which is necessary for the software implementation is discussed in part II and IV within this thesis. In this part the software GUI is presented and the basic features are explained.

23. Basic Functionalities

"D-MLCNC" supports stereo input and stereo output. Therefore one input channel (left or right) can be assigned to each crossover channel. In total eight output channels are supported by the software. Every crossover channel supports the following filter options:

- Gain Controlling – to adjust the signal gain of loudspeaker drivers.
- Signal Delaying.
- Signal Inversion.
- High-pass filter (Butterworth or Linkwitz-Riley Filter).
- Low-pass filter (Butterworth or Linkwitz-Riley Filter).
- Five parametric equalizers (Low-shelving, peak or high-shelving filters).

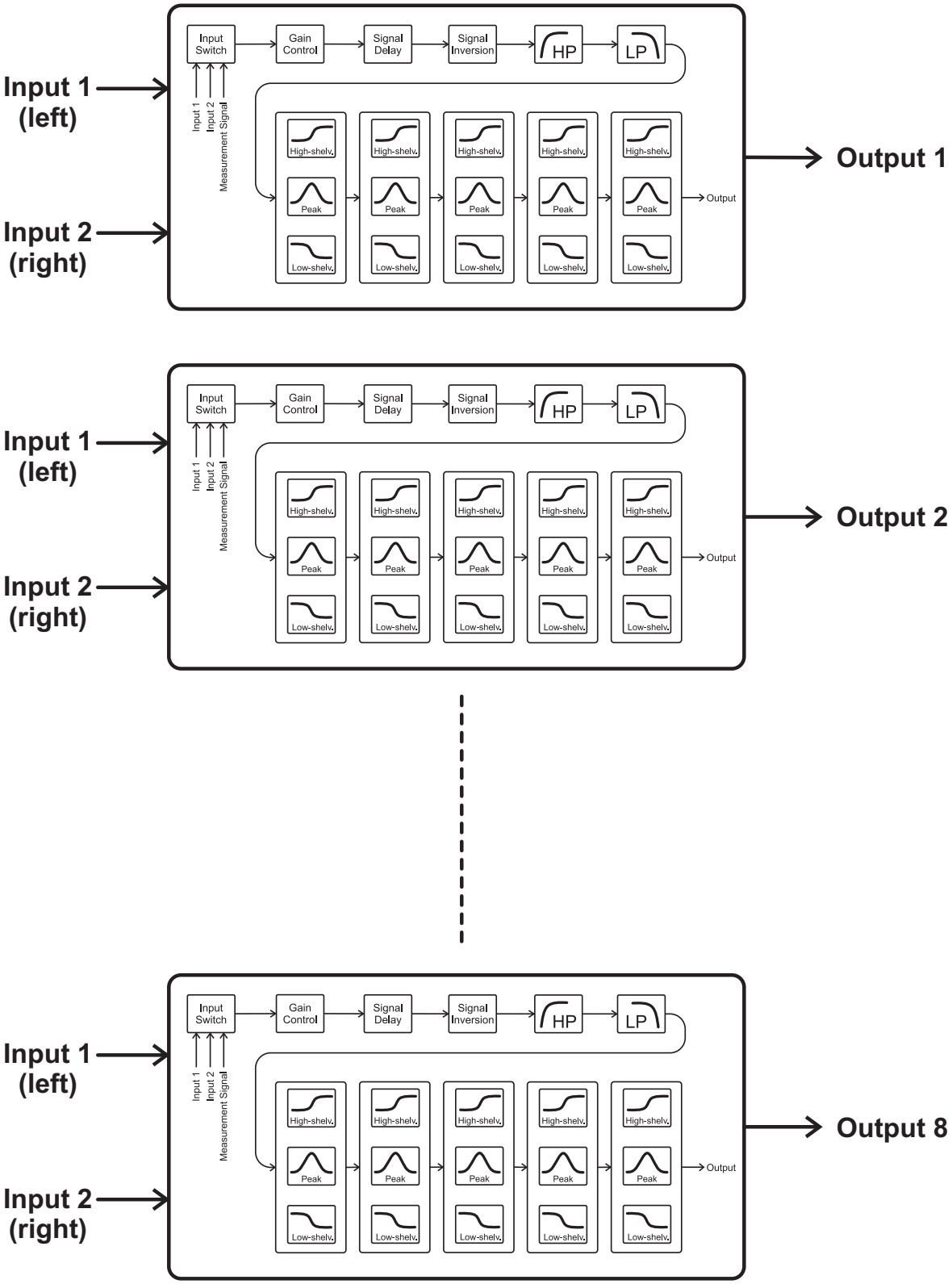

Figure 23.1.: D-MLCNC signal flow graph and equalizer stages.

The internal signal flow of D-MLCNC for eight channels with several filter configuration options is illustrated in figure 23.1.

Figure 23.2 illustrates one hardware wiring possibility, which can be chosen for the use of D-MLCNC with a two-way stereo loudspeaker system. In total eight crossover channels are available, thus e.g. two four-way loudspeakers crossover networks can be designed and simulated and measured with the applied loudspeakers. The soundcard is connected to the

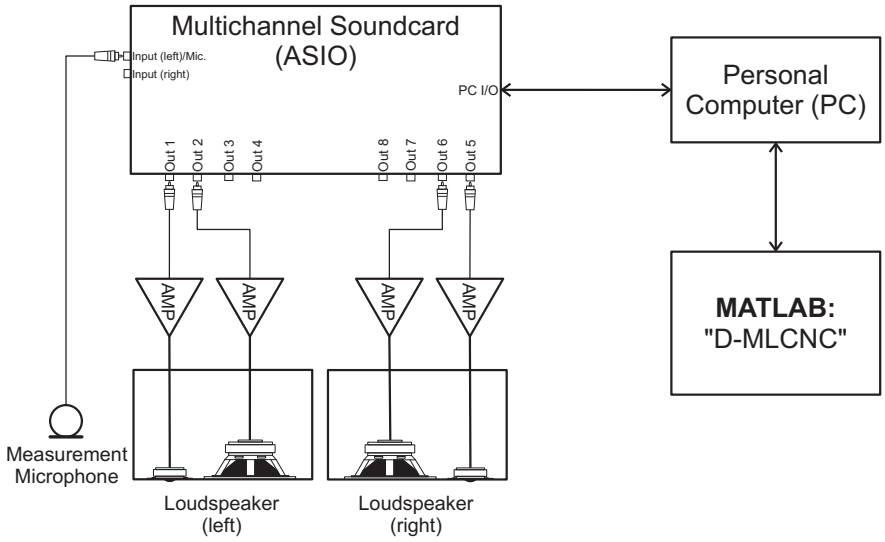

Figure 23.2.: D-MLCNC: Hardware and software connection.

PC. A measurement microphone can be optionally plugged in on one input channel in order to perform measurements (see chapter 24.4). The speakers are applied to the soundcard output channels.

The connection between software and hardware is enabled in D-MLCNC using the utility "Playrec" (see chapter 18.1). For practical use of D-MLCNC a standard personal computer (PC) is sufficient.

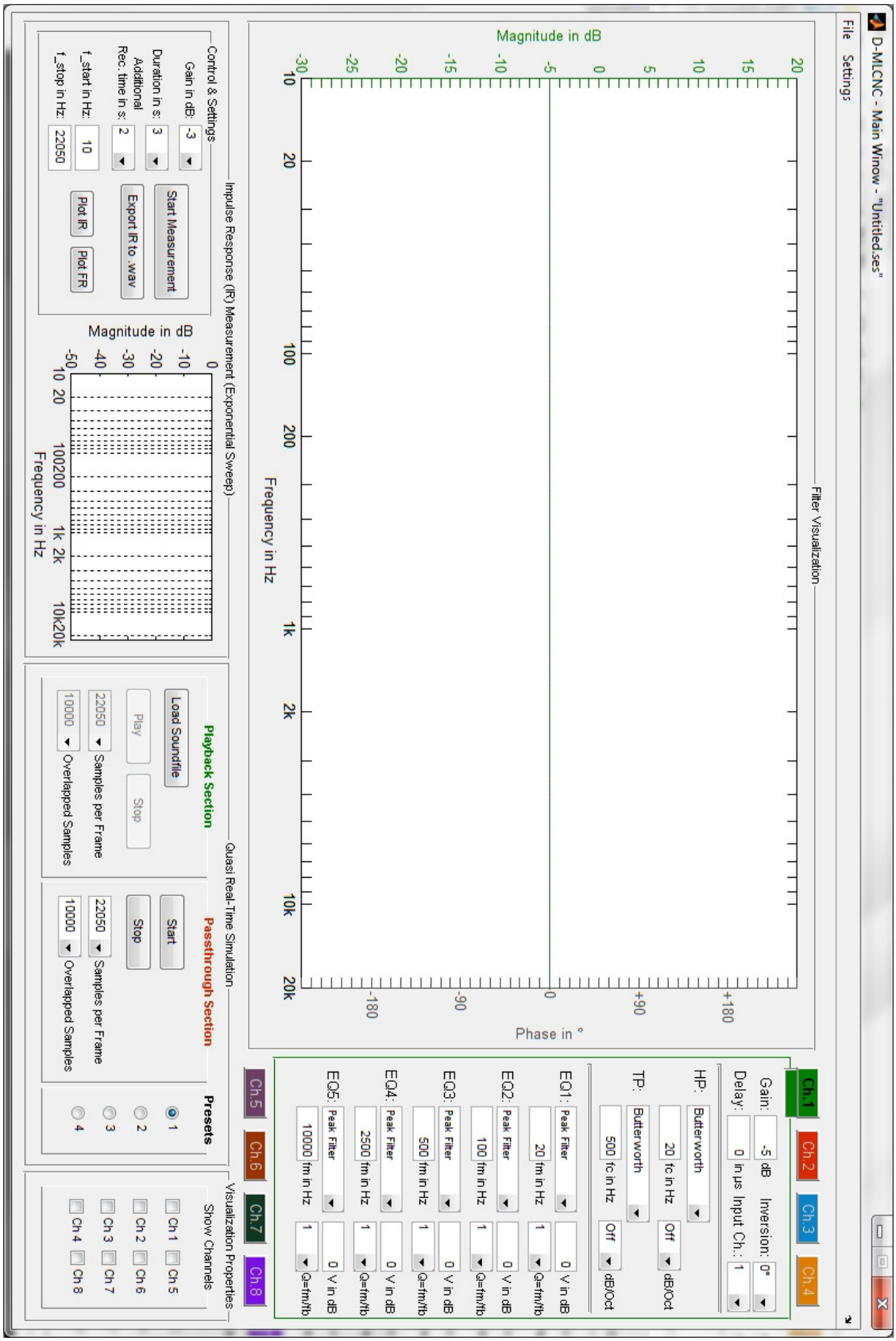

Figure 23.3.: D-MLCNC Main Window – GUI after starting the program.

24. Software GUI – Main Window

In figure 23.3 the main window of D-MLCNC is shown. Further program features are available in the main menu with the menu items *File* and *Settings*. In figure 24.1 the submenus of D-MLCNC are shown.

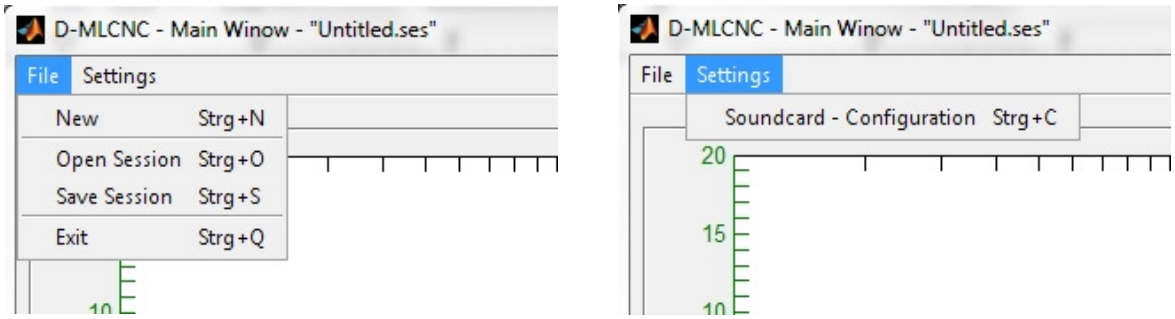

Figure 24.1.: D-MLCNC: Main menu selection opportunities. A complete section can be saved and opened; Soundcard configurations can be changed (see chapter 24.3.1).

The main window is subdivided in program sections with different functionalities. In figure 24.2 the internal data flow of D-MLCNC between different program sections is illustrated in correspondence with the main window structure (see figure 23.3).

Five different program sections can be observed:

- The **Filter Visualization** section on the top side.
  - The filter **Visualization Properties** section on the bottom right side, where additional channels can be enabled for additional visualization.
- The **Quasi Real-Time Simulation** section on the bottom side.
- The **Loudspeaker Crossover Filter Configuration** section on the right side. Different crossover channels can be selected with the buttons *Ch.1-8* (see figure 23.3). Several filter specifications can be separately chosen for every channel.
- An **Impulse Response (IR) Measurement (Exponential Sweep)** section on the bottom left side, where measurements can be performed and exported. For this purpose a measurement microphone has to be plugged in (see figure 23.2).

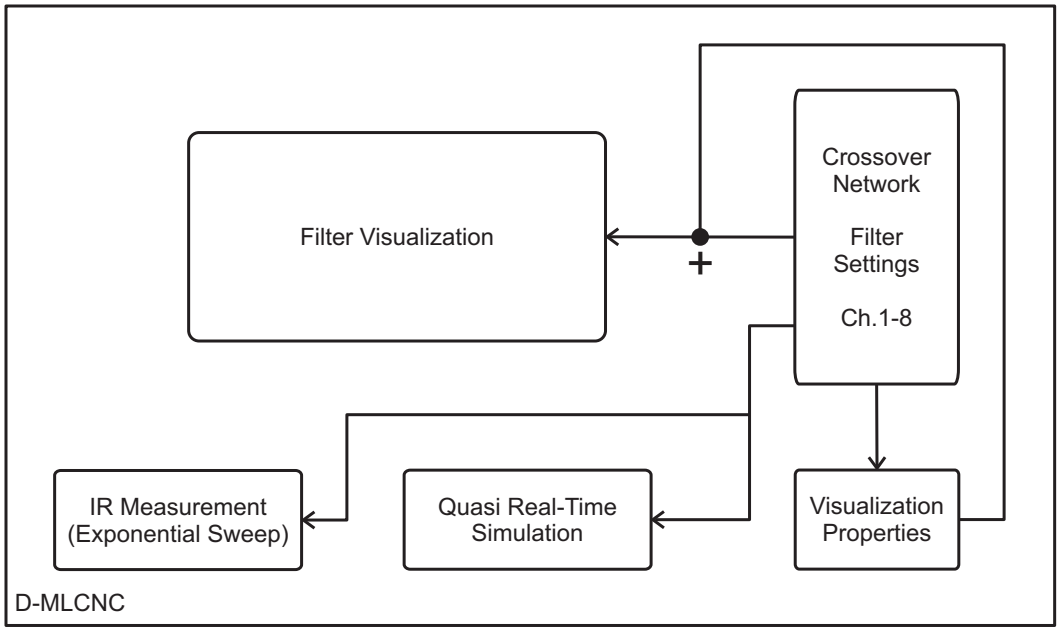

Figure 24.2.: D-MLCNC: Internal data flow graph of the main window between different program sections.

24.1. Filter Visualization

In the filter visualization section the frequency response (magnitude and phase) of the currently selected (see chapter 24.2) crossover channel is shown. Additional channels can be illustrated by activating desired channels in the visualization properties section.

24.1.1. Visualization Properties

In figure 24.3 the visualization properties section is illustrated. Additional crossover channel

Figure 24.3.: D-MLCNC: Visualization properties section. Additional crossover channels can be enabled for visualization.

filters can be enabled for the visualization in the filter visualization section.

The channel selection is limited by the maximum number of supported output channels of the used soundcard (after the soundcard initialization – see chapter 24.3.1).

24.2. Loudspeaker Crossover Filter Configuration

In figure 24.4 the loudspeaker crossover filter configuration section, which is positioned on the right side of the main window is shown. Certain crossover channels can be activated

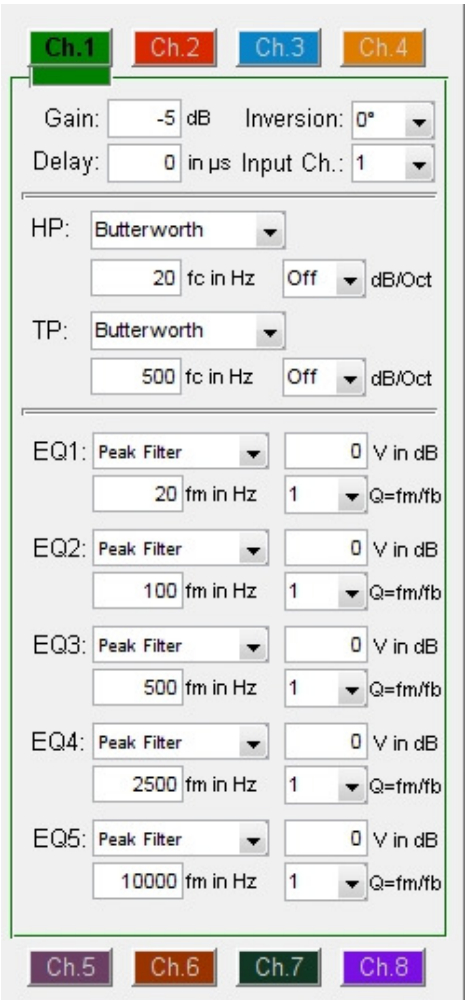

Figure 24.4.: D-MLCNC: Loudspeaker crossover filter configuration section.

for a filter configuration with the buttons *Ch.1-8* (tabbing structure). Several filters can be activated and configured according to the requirements of the user and the used loudspeakers.

The possibility to choose a certain channel with the buttons Ch.1-8 is limited by the soundcard and the maximum number of provided output channels respectively.

A certain filter change whilst quasi real-time simulation has an effect on the signal output as well (see chapter 24.3).

24.3. Quasi Real-Time Filter Simulation

In figure 24.5 the quasi real-time simulation section is illustrated.

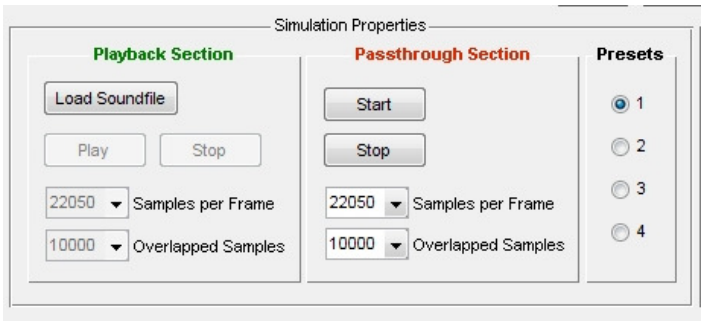

Figure 24.5.: D-MLCNC: Quasi real-time simulation section. On the one hand the quasi real-time simulation can be performed (more details in chapter 18.2). On the other hand a global preset switching is available, which enables a fast switching of different crossover filter presets whilst playback.

In the this section two different simulation options can be chosen:

1. **Signal Playback:** A stereo wave or mp3 file can be loaded with the button *Load Soundfile*. After loading is complete the frame size and the frame overlapping can be chosen (see chapter 18.2). These are important definitions for a quasi-real time simulation.

   The button *Play* starts the simulation of the defined filters. The input channel mapping has to be chosen in the crossover channel definition section on the right side of the main window. This input channel selector is part of the loudspeaker crossover filter configuration section (see chapter 24.2) and is illustrated in figure 24.6 as well.

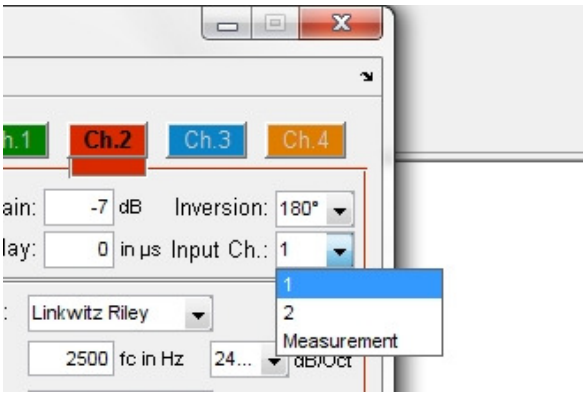

Figure 24.6.: D-MLCNC input channel selection (Channel 1, Channel 2 or Measurement).

2. **Signal Passthrough:** A stereo sound source can be applied on two input channels. The button *start* enables the filter simulation of the input signal, which is recorded and

filtered at the same time. The input channels are assigned with the channel selector as well.

If the soundcard is not initialized the soundcard configuration window will appear first (see chapter 24.3.1).

24.3.1. Soundcard Configuration

In figure 24.7 the soundcard configuration window is shown, which can be opened with the button *Settings* in the main menu (see figure 24.1). In D-MLCNC the configuration of the

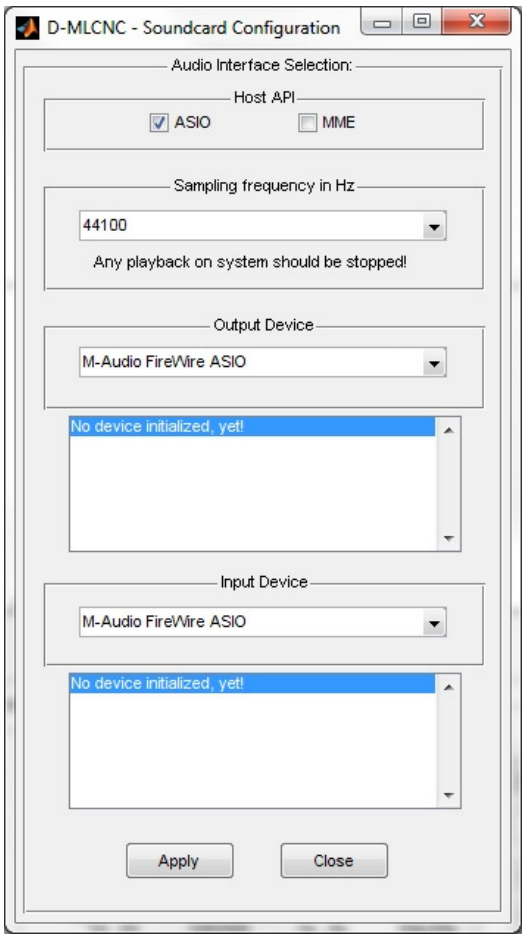

Figure 24.7.: D-MLCNC soundcard configuration window.

soundcard is mainly similar to RoLoSpEQ and is also discussed in chapter 20.1.1.

The only difference is that in D-MLCNC the channel mapping feature is not implemented, yet. Hence the channel mapping in D-MLCNC is assigned according to the soundcard driver default.

24.4. Impulse Response (IR) Measurement (Exponential Sweep)

In figure 24.8 the impulse response measurement section is illustrated. In this section the

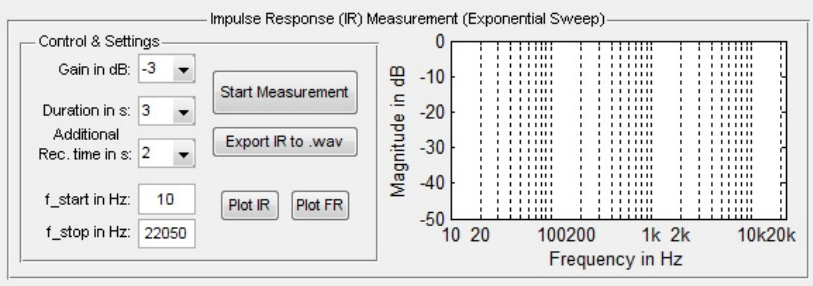

Figure 24.8.: D-MLCNC: Measurement section. Measurements can be performed according to the chosen crossover configuration. The measurement signal has to be the input channel of the desired measurement channels (see figure 24.6).

crossover network can be measured with the loudspeakers. The measured impulse responses can be exported as wave file and analyzed in further software tools.

In order to assign the measurement signal (exponential sweep – see chapter 12) the input channel selector (see figure 24.6) has to be set to *Measurement*. This adjustment has to be performed for every measurement channel.

Then the measurement can be started with the button *Start Measurement*. The measurement results can be checked with the buttons *Plot IR* and *Plot FR* – a separate window is opened containing either a plot of the measured impulse response or of the corresponding magnitude response.

Finally the impulse response measurement can be exported with the button *Export as .wav*.

25. An Objective Evaluation

In this paragraph, different measurements of a two-way loudspeaker are presented. For this purpose a two-way loudspeaker is optimized and measured with D-MLCNC. Special emphasis is put on a balanced overall transmission behavior and especially on a correct radiation behavior at the crossover frequency. This is achieved using mainly 4^{th} order Linkwitz-Riley filters, a gain controlling, a signal delay and a peak filter. The IR measurements are either evaluated with the magnitude responses or with vertical radiation patterns at the crossover frequency.

The used loudspeaker is a two-way bass-reflex system with a Visaton AL170 woofer and a Visaton G25FFL tweeter. The loudspeaker box is rectangular and the size is $420mm$ for the height, $220mm$ for the width and $320mm$ for the depth and is illustrated in figure 25.1.

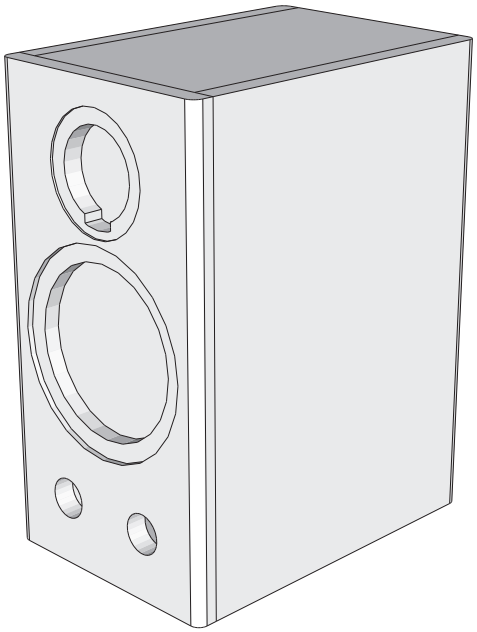

Figure 25.1.: Two-way loudspeaker box, which is used for the measurements.

The measurements of different crossover network configurations and the loudspeaker are performed directly in D-MLCNC (see chapter 24.4). For this purpose an omnidirectional DPA 4006-TL microphone is used. The impulse responses (IR) are further analyzed in MATLAB considering mainly the direct signal without reflections. Several frequency response measurements are mainly analyzed above $200Hz$, because on the one hand a low-frequency optimization of the two-way loudspeaker system is not desired an on the other hand the measurement resolution is limited due to the windowing of the IR.

In order to evaluate the basic transmission behavior of the both loudspeaker drivers in the loudspeaker box, measurements are performed without filters. The magnitude responses of the loudspeaker drivers are illustrated in figure 25.2. A higher efficiency can be observed

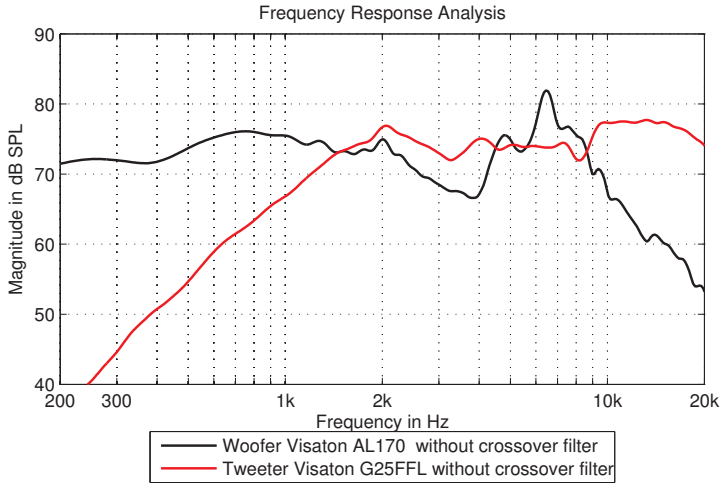

Figure 25.2.: On axis measurement of a two-way loudspeaker with Visaton woofer AL170 (black curve) and Visaton tweeter G25FFL (red/gray curve). No crossover filters applied at the drivers. Separate measurements.

for the tweeter. Therefore, a $+4dB$ gain is defined for the woofer in order to achieve a level adjustment.

Furthermore a crossover frequency of $f_c = 2.5kHz$ is generally defined to enable a balanced overall transmission behavior of the two-way loudspeaker. This crossover frequency is further used for the separation of the frequency ranges of both crossover paths at several crossover network configurations.

In figure 25.3 two magnitude responses of the loudspeaker drivers are again illustrated. For the woofer signal a 4^{th} order Linkwitz-Riley (L-R) low-pass filter is used with $f_c = 2.5kHz$. For the tweeter signal a 4^{th} order L-R high-pass filter is used with $f_c = 2.5kHz$. The drivers are measured separately.

In figure 25.4 the overall transmission behavior of the two-way loudspeaker is illustrated. The L-R filters are used and the tweeter signal is delayed in order to achieve an adjustment of the acoustic centers of both chassis at the crossover frequency (see chapter 10.2.4). The delay is defined according to $\Delta t_1 = 150\mu s$, which equals a displacement $d_1 \approx 5.1cm$ of the acoustic center.

However, the transmission behavior is not optimal, because a large dip appears at the crossover frequency due to an inverse phase behavior. Hence the tweeter signal has additionally to be inverted. The overall transmission of the same configuration with a signal inversion for the tweeter is illustrated in figure 25.5. A balanced on-axis transmission behavior of the two-way loudspeaker can be observed. Nevertheless, the focus should be also put on the vertical directivity behavior of the two-way loudspeaker at the crossover frequency in order to evaluate the accuracy of the filter configuration. For this purpose a set of 19 frequency response measurements are performed at different vertical angles between $-90°$ and $+90°$ in

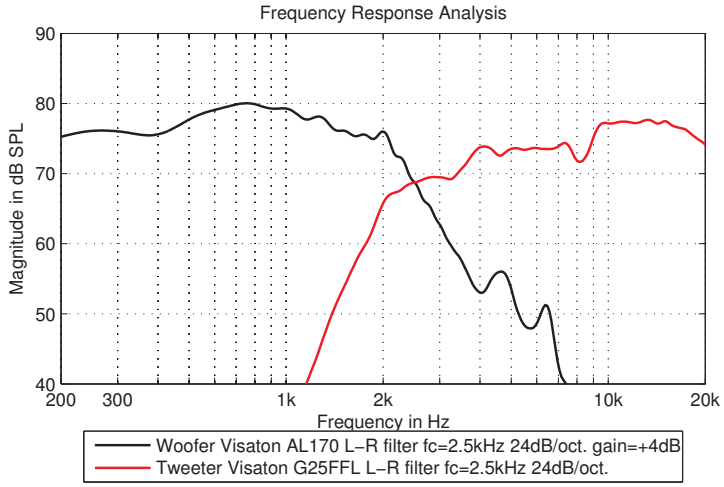

Figure 25.3.: On axis measurement of a two-way loudspeaker with Visaton woofer AL170 (black curve) and Visaton tweeter G25FFL (red/gray curve). Separately measured loudspeaker drivers with L-R crossover filter network. 4^{th} order L-R filters with a crossover frequency of $f_c = 2.5kHz$ are used.

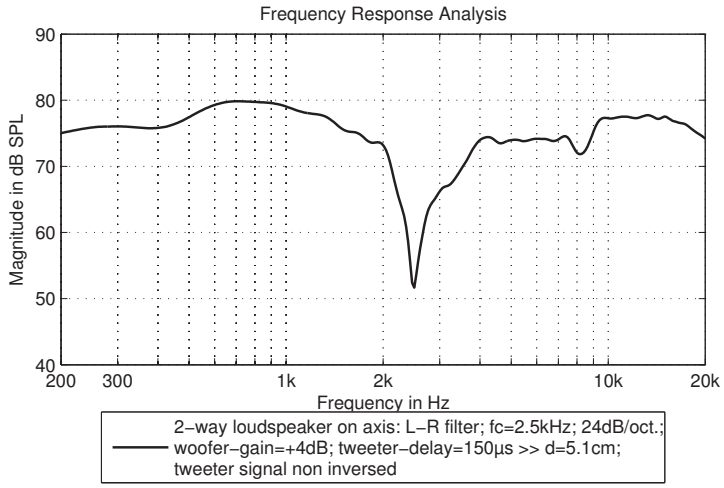

Figure 25.4.: Two-way loudspeaker with Visaton woofer AL170 (black curve) and Visaton tweeter G25FFL (red/gray curve). 4^{th} order L-R crossover filter network with a crossover frequency $f_c = 2.5kHz$. The tweeter signal is delayed in order to achieve an adjustment of the acoustic centers at the crossover frequency ($\Delta t_1 = 150\mu s$). A large dip appears at the crossover frequency due to an inverse phase behavior.

front of the loudspeaker baffle.

In figure 25.6 four measurements in different vertical directions are illustrated. An increased directivity behavior can be observed for higher vertical measurement angles.

Several measurements are used to derive a vertical radiation pattern of the two-way loud-

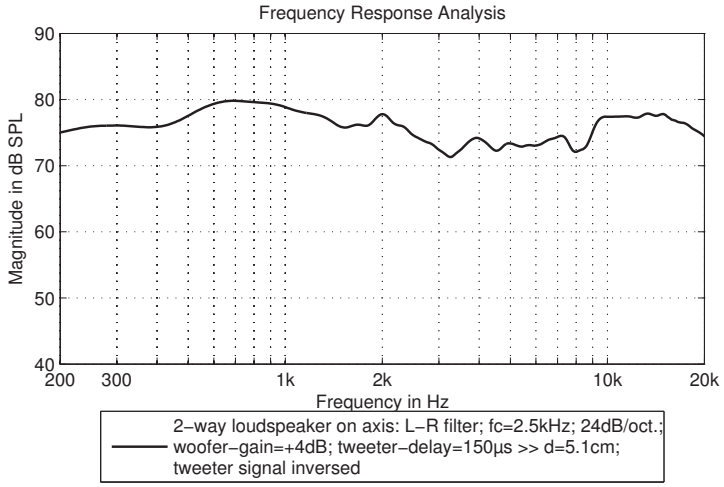

Figure 25.5.: On axis measurement of a two-way loudspeaker with Visaton woofer AL170 (black curve) and Visaton tweeter G25FFL (red/gray curve). 4^{th} order L-R crossover filter network with a crossover frequency $f_c = 2.5kHz$. The tweeter signal is delayed in order to achieve an adjustment of the acoustic centers at the crossover frequency ($\Delta t = 150\mu s$). The tweeter signal is inverted to avoid a dip at the crossover frequency.

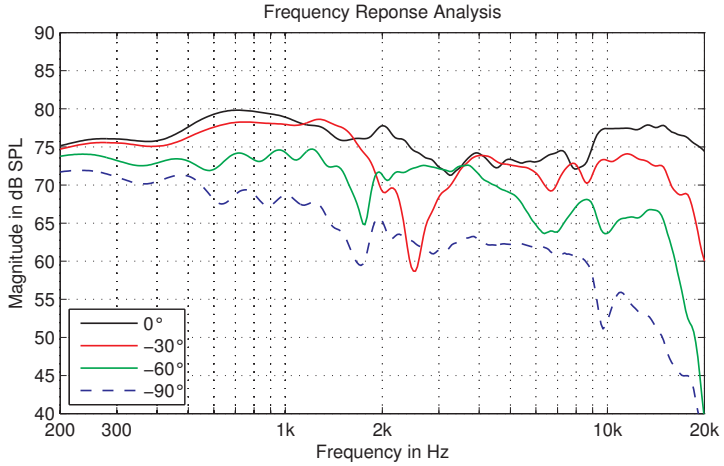

Figure 25.6.: Four measurements with different vertical angle in front of the baffle of a two-way loudspeaker. An increased directivity behavior can be observed for higher vertical measurement angles.

speaker at the crossover frequency $f_c = 2.5kHz$. The vertical radiation pattern is illustrated in figure 25.7. The typical lobing of a two-way loudspeaker at the crossover frequency can be observed in the radiation pattern (see also figure 10.5). The crossover configuration seems to be very good, as the main lobe points exactly on-axis to the baffle and no radiation tilt can be observed.

In order to illustrate the impact of the tweeter delay, a second delay is used. For this

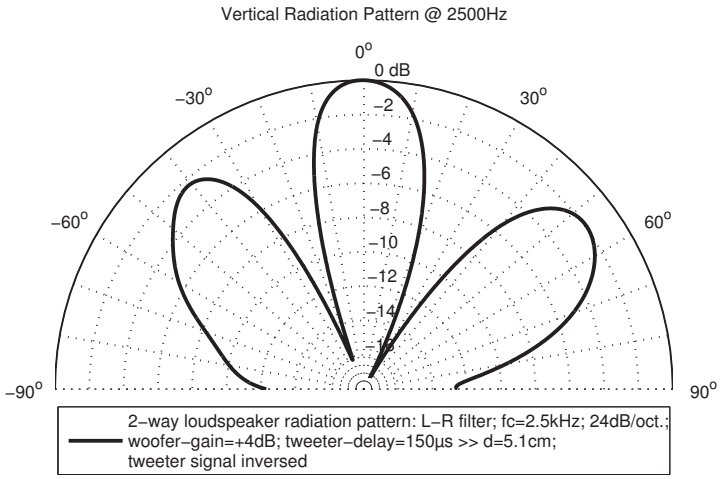

Figure 25.7.: Vertical radiation pattern of a two-way loudspeaker at the crossover frequency. On axis: 0°, loudspeaker bottom: +90°, loudspeaker top: −90∘. 4^{th} order L-R crossover filter network, $f_c = 2.5kHz$, woofer gain $V = +4dB$, tweeter delay $\Delta t = 150\mu s$, tweeter signal inverted.

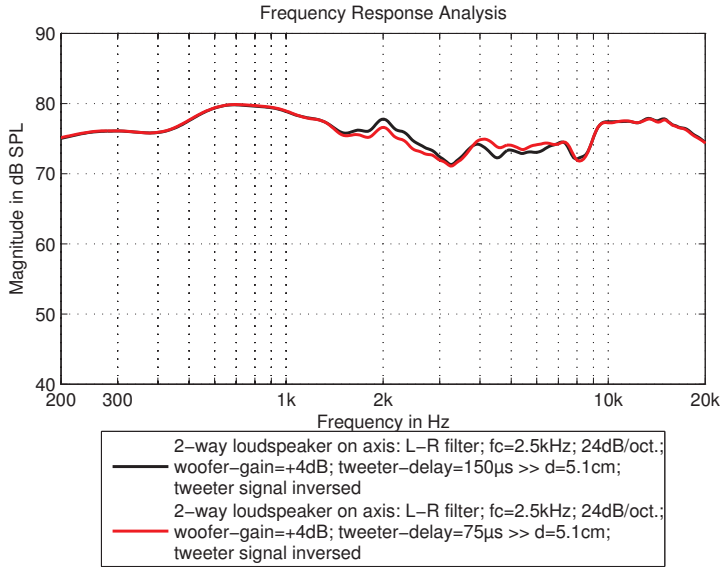

Figure 25.8.: On-axis measurement of a two-way loudspeaker. Two different tweeter signal delays are used: $\Delta t_1 = 150s$ (black curve) and $\Delta t_2 = 75s$ (red/gray curve).

purpose the tweeter signal delay is set to $\Delta t_2 = 75\mu s$, which equals a displacement of the acoustic center of approximately $2.5cm$.

In figure 25.8 the on-axis magnitude responses of the different delay configurations is illustrated. Both crossover configurations offer a balanced on-axis transmission behavior.

For a more detailed evaluation the vertical radiation behavior is again analyzed.

In figure 25.9 the radiation pattern of the two-way loudspeaker at the crossover frequency

with different tweeter signal delays is illustrated. For $\Delta t_2 = 75\mu s$ a radiation tilt to the

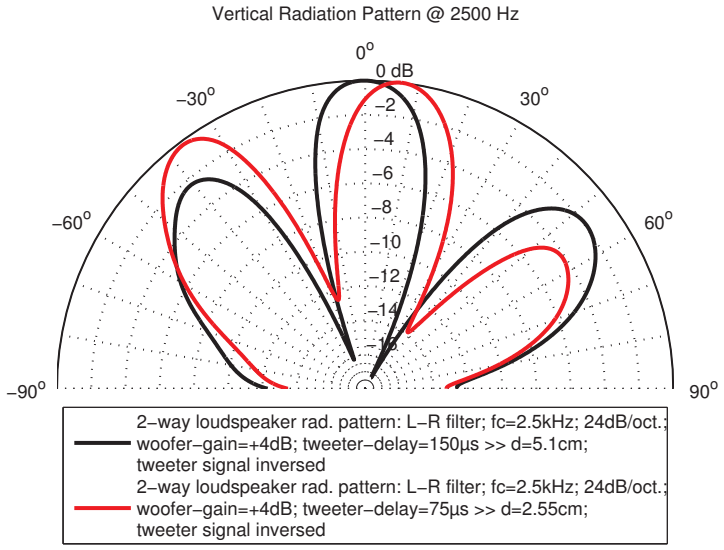

Figure 25.9.: Vertical radiation pattern of a two-way loudspeaker at the crossover frequency $f_c = 2.5kHz$. Two different tweeter signal delays are used: $\Delta t_1 = 150\mu s$ (black curve) and $\Delta t_2 = 75\mu s$ (red/gray curve).

loudspeaker bottom, which is a negative effect can be observed (see chapter 10.2.6).

Furthermore a higher tweeter signal delay $\Delta t_3 = 300\mu s$ is used. This delay equals an offset of the acoustic center of $d = 10.2cm$. In consideration of the crossover frequency $f_c = 2.5kHz$, this is almost a wavelength of $\lambda_c = \frac{c_0}{f_c} \approx 13.6cm$. In addition a peak filter is introduced in order to equalize the transmission behavior in the frequency range at $600Hz$.

In figure 25.10 two magnitude responses of the two-way loudspeaker, which are measured with the tweeter signal delay $\Delta t_3 = 300\mu s$ are illustrated. Due to the high offset of the acoustic center, the tweeter signal inversion leads to a dip in the on-axis magnitude response due to an inverse phase behavior (black curve). In order to achieve a balanced on-axis transmission behavior the tweeter signal has to be non-inverted (red/gray curve). In addition the impact of the peak filter can be observed.

Even more interesting is the comparison of the vertical radiation behavior of the described configuration at the crossover frequency $f_c = 2.5kHz$, which is illustrated in figure 25.11. As expected, the vertical radiation pattern of the configuration with inverted tweeter signal shows a sound cancellation in on-axis direction. Furthermore a strong radiation tilt to the loudspeaker top appears.

On the other hand, the vertical radiation pattern of the configuration with non-inverted tweeter signal results exactly in an inverse radiation behavior. However a small radiation tilt to the loudspeaker bottom can be observed.

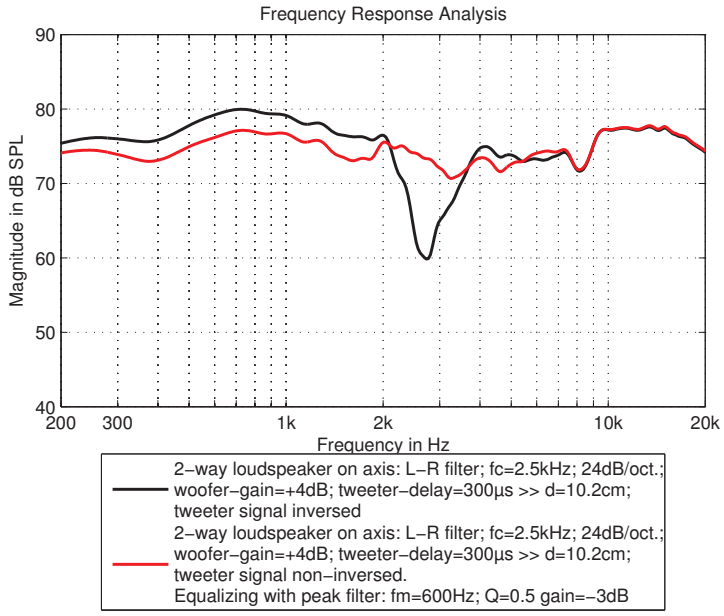

Figure 25.10.: On axis measurement of a two-way loudspeaker. A high tweeter signal delay $\Delta t_3 = 300\mu s$ is used. Two different cases are illustrated: tweeter signal inverted (black curve) and tweeter signal non-inverted (red/gray curve). In addition a peak filter is used in order to achieve an equalization at approx. $600Hz$ (red/gray curve).

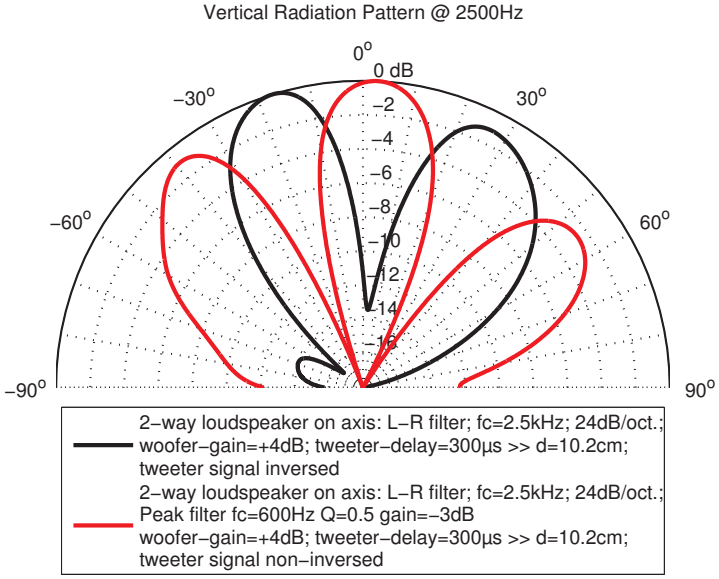

Figure 25.11.: Vertical radiation pattern of a two-way loudspeaker at the crossover frequency $f_c = 2.5kHz$. A high tweeter signal delay $\Delta t_3 = 300\mu s$ is used. Two different cases are illustrated: tweeter signal inverted (black curve) and tweeter signal non-inverted (red/gray curve). An opposite radiation behavior can be observed.

26. Conclusion

With RoLoSpEQ a ready-to-use software tool for room plus loudspeaker response equalization within small rooms is developed. RoLoSpEQ enables an automatic derivation and simulation of room plus loudspeaker response equalizers, which can also be exported for the use in other software tools. A measurement routine based on exponential sweeps is developed and basically used for the analysis of the loudspeaker sound reproduction characteristic in a specific room. With the consideration of the fundamental interaction between loudspeakers and small listening rooms an automatic target generation approach is developed, which still enables a manual adaption. That is, compared to other approaches RoLoSpEQ provides and automatic but adaptable target response generation, which allows the adaption to human hearing habits. The target responses are generated for two different listening positions, one for an equalization filter generation for a sound reproduction improvement in the entire room (GLOBAL) and another for an equalization filter generation for a sound reproduction improvement in the listening sweet spot (FOCUS). The derivation of equalization filters is performed completely automatically using these target responses. A further special contribution of RoLoSpEQ is the developed block-processing approach based on the MATLAB utility "Playrec", which enables a quasi real-time simulation and switching between generated equalization filters.

With D-MLCNC a ready-to-use software tool for quasi real-time simulation and measurement of multichannel loudspeaker crossover networks is developed. This multichannel capability is also enabled using the MATLAB utility "Playrec". D-MLCNC is convenient for a very flexible and fast loudspeaker crossover development, as different filter configurations can be easily switched and measured. For this purpose commonly used analog filters are digitally simulated in order to enable the optimization of the transmission behavior of arbitrary multi-way loudspeaker systems. This loudspeaker system can also be practically simulated using the developed block-processing approach. In addition the loudspeaker system can be measured for a precise loudspeaker analysis.

For further investigations a subjective evaluation of the sound reproduction quality improvement in different rooms with RoLoSpEQ is desirable, or whether sound colorations are audible. This subjective evaluation could also enable a more detailed investigation of the "Room Gain" effect, which seems to be a key factor for a natural room plus loudspeaker equalization. This evaluation could lead to a further automation of the equalization filter specifications. Nevertheless the manual adjustment of the equalization filters to human hearing habits should still be maintained. For further extensions of RoLoSpEQ a simplified and optimized software handling is desirable (e.g. definition of higher level parameters). In addition further features could be added to RoLoSpEQ, e.g. an automated signal inversion in the case of a reverse-poled loudspeaker connection, which can be easily detected in an impulse response measurement.

For a further practical optimization a combination of RoLoSpEQ and D-MLCNC in one single software tool is desirable. Furthermore the implementation could be extended for the use on

DSP hardware. For this purpose a transmission of final filter configurations after simulation, measurement and evaluation is conceivable. On the one such kind of software tool provides a simulation and measurement environment in order to find suitable filter configurations. On the other hand this software tool provides a user interface for a DSP hardware in order to transfer the desired filter configurations.

Bibliography

[Bar01] Hans-Jochen Bartsch. *Taschenbuch Mathematischer Formeln.* Fachbuchverlag Leipzig im Carl Hanser Verlag München Wien, 19th edition, 2001.

[Boh05] Dennis Bohn. Linkwitz-Riley Crossovers: A Primer. Puplished via website, 2005. Available online at `http://www.rane.com/note160.html`; visited on April 30^{th} 2010.

[Dic05] Vance Dickason. *Lautsprecherbau: Bewährte Rezepte für den perfekten Bau.* Elektor-Verlag Aachen, 3rd edition, 2005.

[EAP07] Fares El-Azm and Jan Abildgaard Pedersen. Natural Timbre in Room Correction Systems (Part II). *AES* 32^{nd} *International Convention: Hillerød, Denmark*, September 21-23 2007.

[EP09] F. Alton Everest and Ken Pohlmann. *Master Handbook of Acoustics.* McGraw-Hill/TAB Electronics, 5 edition, June 2009.

[GW05] Gerhard Graber and Werner Weselak. Skript zur Vorlesung Raumakustik. Technical report, Institut für Breitbandkommunikation, Technische Universität Graz, 2005.

[GW06] Gerhard Graber and Werner Weselak. Elektroakustik – Version 8.2. Technical report, Institut für Breitbandkommunikation, Technische Universität Graz, 2006.

[Hum06] Robert A. Humphrey. Automatic Loudspeaker Location Detection for use in Ambisonics Systems. 4th year project report for the degree of meng in electronic engineering, University of York, June 2006.

[Hum08] Robert Humphrey. Playrec Documentation. Puplished via website, 2008. Available online at `http://www.playrec.co.uk/index.php`; visited on May 20^{th} 2010.

[KP07] Matti Karjalainen and Tuomas Paatero. Equalization of loudspeaker and room responses using Kautz filters: direct least squares design. *EURASIP J. Appl. Signal Process.*, 2007(1):185–185, 2007.

[KS08] KRK-Systems. *ERGO - User Manual*, 2008. Available online at `http://www.krksys.com/manuals/ergo_manual.pdf`; visited on April 2^{nd} 2010.

[Kut09] Heinrich Kuttruff. *Room Acoustics.* Spon Press, 5th edition, June 2009.

[Lin78] Siegfried H. Linkwitz. Active Crossover Networks for Noncoincident Drivers. *J. Audio Eng. Soc.*, vol. 24, pp. 2-8, January/February 1978.

[MÖ1] Swen Müller. Transfer-Function Measurement with Sweeps. *J. Audio Eng. Soc., Vol. 49, No. 6*, June 2001.

[Mat08] The Math Works Inc. *MATLAB*, 2008.

[MBL07] Piotr Majdak, Peter Balazs, and Bernhard Laback. Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions. *Journal of the Audio Engineering Society, Vol. 55*, pages 623–637, 2007.

[Mey09] Martin Meyer. *Signalverarbeitung: Analoge und digitale Signale, Systeme und Filter*. Vieweg+Teubner, 5th edition, 2009.

[NA79] Stephen T. Neely and J. B. Allen. Invertibility of a room impulse response. *The Journal of the Acoustical Society of America*, 66(1):165–169, 1979.

[OSB99] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-time Signal Processing*. Prentice Hall, 2nd international edition, Jan. 1999.

[Ped03a] Jan Abildgaard Pedersen. ADAPTIVE BASS CONTROL – the ABC room adaption system. *AES* 23^{rd} *International Conference: Copenhagen, Denmark*, May 23-25 2003.

[Ped03b] Jan Abildgaard Pedersen. Adjusting a Louspeaker to its Acoustical Environment - the ABC system. 115^{th} *AES Convention: New York, NY, USA*, October 10-13 2003.

[Ped06] Jan Abildgaard Pedersen. Loudspeaker-Room Adaption for a Specific Listening Position using Information about the Complete Sound Field. 121^{st} *AES Convention: San Francisco, CA, USA*, October 5-8 2006.

[Ped07] Jan Abildgaard Pedersen. Sampling the Energy in a 3D Sound Field. 123^{rd} *AES Convention: New York, NY, USA*, October 5-8 2007.

[PHR94] Jan Abildgaard Pedersen, Kjeld Hermansen, and Per Rubak. The Distribution of the Low Frequency Sound Field and its Relation to Room Equalization. 96^{th} *AES Convention: Amsterdam, Netherlands*, February 1994.

[PM07] Jan Abildgaard Pedersen and Henrik Green Mortensen. Natural Timbre in Room Correction Systems. 122^{nd} *AES Convention: Vienna, Austria*, May 5-8 2007.

[PT07] Jan Abildgaard Pedersen and Kasper Thomsen. Fully Automatic Louspeaker-Room Adaption - the RoomPerfect System. *AES* 32^{nd} *International Convention: Hillerød, Denmark*, September 21-23 2007.

[Sch08] Hans W. Schüßler. *Digitale Signalverarbeitung 1: Analyse diskreter Signale und Systeme*. Springer, Berlin, 5th edition, March 2008.

[SPR99] J. Schoukens, R. Pintelon, and Y. Rolain. Broadband versus Stepped Sine FRF Measurements. In *Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE*, volume 2, 1999.

[SW82] Jorma Salmi and Anders Weckström. Listening Room Influence on Loudspeaker Sound Quality and Ways Of Minimizing It. 71^{st} *AES Convention: Montreux, Swiss*, March 2-5 1982.

[Too08] Floyd E. Toole. *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*. Focal Press, Jul. 2008.

[TS09] Ulrich Tietze and Christoph Schenk. *Halbleiter-Schaltungstechnik (German Edition)*. Springer, 13th edition, Oct. 2009.

[Wei08] Stefan Weinzierl, editor. *Handbuch der Audiotechnik*. Springer, Berlin, 1st edition, Dec. 2008.

[Wes09] Werner Weselak. Akustische Messtechnik 1. Technical report, Institut für Breitbandkommunikation, Technische Universität Graz, 2009.

[ZÖ8] Udo Zölzer. *Digital Audio Signal Processing*. Wiley, 2nd edition, Aug. 2008.

[ZAA+08] Udo Zölzer, Xavier Amatriain, Daniel Arfib, Jordi Bonada, Giovanni De Poli, Pierre Dutilleux, Gianpaolo Evangelista, Florian Keiler, Alex Loscos, Davide Rocchesso, Mark Sandler, Xavier Serra, and Todor Todoroff. *DAFX:Digital Audio Effects*. Wiley, 2nd edition, June 2008.