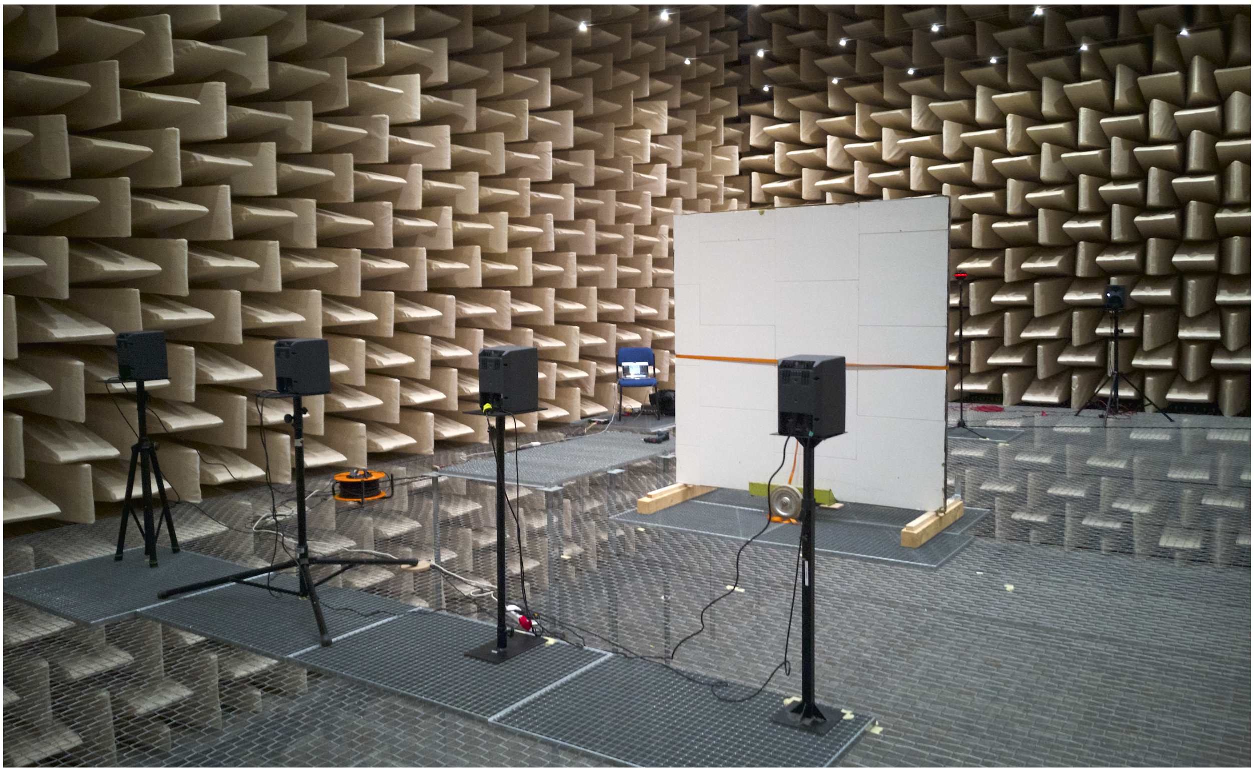

Directional Filtering of Recordings in Spherical Harmonic Domain for Auralization of Simulated Noise Barriers

Master's thesis in M.Sc Programme Sound and Vibration

BENEDIKT BECKENBAUER

Department of Architecture and Civil Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022
www.chalmers.se

MASTER'S THESIS 2022

Directional Filtering of Recordings in Spherical Harmonic Domain for Auralization of Simulated Noise Barriers

Supervisor and Examiner: Jens Ahrens, Applied Acoustics - Department of Architecture and Civil Engineering

Cover: Measurement setup for the measurement of spatial impulse responses of a noise barrier in the anechoic chamber of the TU Berlin.

Department of Architecture and Civil Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022

Directional Filtering of Recordings in Spherical Harmonic Domain for Auralization of Simulated Noise Barriers

BENEDIKT BECKENBAUER
Department of Architecture and Civil Engineering
Division of Applied Acoustics
Chalmers University of Technology

Abstract

The effects of traffic noise on humans can cause significant health problems and can impair the quality of everyday life. Current research and legal requirements mainly refer to physical sound pressure measurements of noise. However, the human perception of noise within its acoustic and visual context is not fully considered. With the development of growing population that is exposed to noise caused by roads, rails and planes, the need of an evaluation method that includes the perception of the noise within its environment gains in importance. At the same time a financially efficient tool is needed to reduce costs of construction projects for noise protection measures. Therefore, an evaluation tool should be evolved, that makes it possible to experience and rate noise protection measures based on a psychoacoustic model in virtual reality. Supported by findings from the virtual reality model, suitable noise protection measures can be built in real-life subsequently.

The aim of this thesis is to process spatial audio recordings in spherical harmonics domain in such a way that the resulting signal contains the spatial sound field of the recording including the effects of a synthesized noise barrier. Through this, the auralization of varying noise barriers can be realized for real life scenarios in a virtual environment. The processing is implemented by applying filter functions to directed sectors of a plane wave decomposed form of the signal. The filter functions approximate the transfer functions of simple single screen noise barriers, based on a secondary source diffraction calculation model. With this methodology it is possible to establish directional frequency filtering, which only affects components of the sound field that approach the receiver from certain directions. The model is used for filtering complex as well as synthetic generated sound fields and is subsequently analyzed thoroughly. Limitations of the processing are discussed and improvements are suggested.

Keywords: Spatial Audio, Ambisonics, Plane Wave Decomposition, Noise Barriers, Edge Diffraction, Spatial Impulse Response, Zero-phase Filtering, Spherical Harmonics

Acknowledgements

I would like to express my great appreciation to my primary supervisor from TU Berlin, Johannes Arend, who guided me throughout this project. I wish to acknowledge the TU Berlin and the staff of the audio communication group for their help, the constructive feedback and for making their measuring rooms available to me. Finally, I would like to offer my special thanks to Assoc. Prof. Jens Ahrens for his guidance.

Benedikt Beckenbauer, Berlin / Gothenburg, 09 2022

Contents

1

Introduction

The exposure to noise can have serious impact on the human health and well-being. Both short- and long-term consequences are caused by environmental noise like noise originating from roads, rails and aircrafts. The significance of those consequences is emphasized by the World Health Organisation, that estimates that the number of lost years of full health is above one million for countries of western Europe with a population of approximately 340 million [21]. Serious health issues caused by long-term exposure to traffic noise can be increasing risks of heart failure or hypertensive heart diseases [56] and metabolic issues like obesity and diabetes [40][14]. There are numerous investigations on the influence of noise on psychiatric illness [9], that can result in an increase of activation of stress hormones, the development of depression and anxiety disorders and conduct disorder of children [57]. Another major point is the noise-related disturbance of sleep [2], which itself can be responsible for further diseases and impairments of human daily life. A short term effect of exposure to noise is the reduction of cognitive abilities, which particularly has been explored for children [58]. The visual and auditory attention can be reduced [28], distraction occurs and the working memory can be degraded [62].

In the majority of recent studies, the specification of the measures of noise are given as simple (weighted) sound pressure levels, thereof sometimes as equivalent sound pressure level. Same applies for legal specifications related to noise protection measures, like construction projects for noise barriers and road surfaces to reduce traffic noise respectively tire rolling noise [12] [10]. Through the absence of extensive investigations, the subjective perception of noise and of noise in the context of its environment is not taken into account in those considerations and measures. 139 million people are exposed to significant traffic noise in the 33 EEA member countries (Turkey excluded) and will likely increase by 8% until 2030 [16]. With the development of population that is exposed to traffic noise, therefore the need of a new evaluation method based on perception of noise gains relevance.

The overarching goal of this project is to evaluate the potential of psychoacoustic parameters to describe traffic noise and common noise protection measures. Especially the assessment of noise protection measures including human perception of noise could be improved, corresponding project planning could benefit financially and noise related health care costs could be reduced. As part of this project the following master thesis focuses on the acoustic part of the virtual environment used for a participants study, to evaluate original scenarios and synthesized noise protection measures. Therefore, a digital processing methodology will be developed and analyzed to answer the question:

Can noise barriers be sufficiently well implemented synthetic into measurements of real acoustic environments evaluated according to subjective rating?

It is well known that the perception of sound is not only based on the sound pressure but also on several quantitative and qualitative parameters that describe temporal and spectral characteristics. Important basic research by Zwicker and Fastl has discovered measurable psychoacoustical quantities like loudness, sharpness and roughness. The focus has also already been placed on noise at that time [67]. Additionally situational and personal factors can influence the perception further [5] [1] [59]. Recently there were published several articles and studies on the research of noise annoyance related to traffic noise [22] [23] [46]. Here only a minority also focuses on other psychoacoustic quantities of noise, limited to only road traffic noise [17] [42] [38].
In the past few years, there have been some investigations on auralization of traffic noise and their synthesis [45], [44] [36]. Only a small number focused on noise reduction measures but on the sound sources. The "LISTEN"-project aims at a tool to auralize future traffic noise scenarios in order to make it possible to evaluate various noise mitigation solutions already at the planning stage. Within the project different auralization simulation approaches (time-domain and additive synthesis) where evaluated perceptually [20]. As part of the SONORUS-project on urban sound planning, an investigation on the overall acceptance of several planning options for the refurbishment of urban public spaces using Virtual Reality was performed [15]. In the project the auralization of an existing noisy environment was realized, including different synthesized noise barriers. To provide an acoustic environment, ambisonic recordings of the real environment were made and the effects of the synthesized sound barriers were calculated, using the detailed finite difference time domain method, which is capable of computing all physical phenomena involved like multiple diffraction and scattering [13]. Calculated frequency-dependent insertion losses were then used for filtering the aforementioned B-Format recordings. To provide an appropriate virtual acoustical environment, including synthesized noise protection measures, the processing must be observed in particular. The virtual experience, which was realized with the software "unity" and by using head-tracking and headphones, was rated as immersive and realistic. In a project of the university of Ghent, an auralization methodology is proposed that enables an auditive preview of noise abatement measures for road traffic noise [63]. In an investigation on the synthesis of wind turbine noise [25] sound attenuation by barriers (and air) is implemented as a linear phase FIR filter, which is applied to the source audio signal to simulate the sound pressure at the receiver position.

2

Theory

In this chapter the theoretical principles which are relevant for the investigations of this thesis are presented and explained in detail.

2.1 Fundamental Principles of Spatial Hearing

Initially the fundamentals of spatial hearing and sound perception are explained briefly, leading to the principles of spatial audio reproduction. The human auditory system consists of two sensory organs, the ears, which forward incoming air- and structure-borne sound to the sensory system. The ears are positioned, with a distance of about 20 cm, lateral on the sides of the head. The basic physical and psychological phenomena which enable the ability of spatial perception and the location of sound are well known and where studied for decades [4, 7, 52]. The physical phenomena depend on the properties of the sound waves arriving at the two ears. Differences between the two arriving sound waves at the ears are called interaural or binaural cues. which are mainly responsible for the differentiated lateral sound localization, i.e. left-right differentiation along the horizontal plane. Level and time related differences in the wave forms arriving at the two ears which are based on differences in the spatial sound propagation path between the ears and a lateral sound source are responsible for that. The maximum possible difference is reached, if the azimuth angle ϕ along the horizontal plane is $\phi = 90^{\circ}$ to the line of sight (cf. figure 2.1). The level differences are caused by the attenuation of the sound due to diffraction around the head for the ear located on the opposite side of the head. It was shown that the minimum audible angle along the horizontal plane is approximately 1° for frontal sources and increasing towards lateral positions of the source [43]. Additionally the sound waves arriving at the ears contain spectral cues, which cause a spectral coloration of the signal. The incident sound wave reaches the entrance of the ear canal not only on a direct path, but also after being reflected and diffracted at the head and torso. The multiple arriving waves interfere and therefore cause spectral patterns, which are depending on the location of the source and highly depending on the individual stature of the person. The spectral cues are mainly responsible for vertical localisation and front-back differentiation [64]. The minimum audible angle here is approximately two to four times higher then for the lateral source localisation [43]. The anatomy composition can be seen as a system redirecting the sound to the ear, containing azimuth and elevation angle-depending filter effects of the binaural and spectral cues. The corresponding head-related transfer function (HRTF) is topic of a variety of current scientific

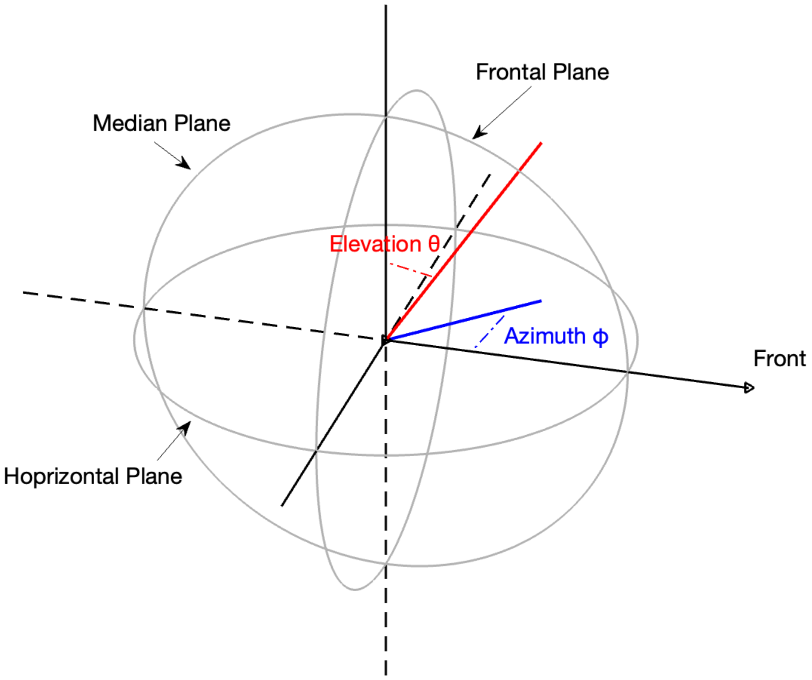

Figure 2.1: Spatial coordinate system. Azimuth angle along the horizontal plane ϕ, Elevation angle along the median plane θ

research and investigations.

2.2 Spatial Audio Reproduction

The reproduction of a sound field, that naturally has a spatial appearance, is a field of audio engineering that has its beginning in the work of Clement Ader in relation with the Paris Opera in 1881 [55]. The principle of modern stereophony was later introduced by Alan Blumlein with the XY recording (Blumlein pair) and playback technique [6]. The now well known principle of combining the signals of two microphones with different orientated directivities gave for the first time the possibility to reproduce the recorded sound field spatially. For the reproduction a pair of loudspeakers is necessary, which plays back the corresponding recorded signals and enables the positioning of sources along the horizontal plane between the speakers. Based on that the recording and playback of audio in three dimensions, the Ambisonics technology [11, 18] and the more sophisticated approach of the soundfield using spherical harmonic functions is described in the following chapter. Furthermore, the sampling of a soundfield on the sphere using microphone arrays is described.

2.2.1 Ambisonics and Spherical Harmonics

Ambisonics recording and playback was already introduced in the 1970's by Craven, Felgett and Gerzon, by describing the concept of utilizing continuous harmonic functions of scalable resolution. The three dimensional audio technique of first-order requires a microphone setup consisting of 4 microphones, which need to be setup in a defined geometric setup, depending on the characteristics of the microphones [66]. The resulting signals can then be decoded to a set of loudspeakers. Here the decoding computes the position depending gains for a source. The signals of the individual microphones can be combined to four channel signals, representing the soundfields omnidirectional, front-back, left-right and up-down components. Consequently the first channel (zeroth order) can be seen as reproducing the signal of a sensor with omnidirectional directivity and the three others reproducing those of sensors with having a figure-of-eight characteristic, each pointing along a different axis. The combination of directivities at the same time corresponds to the appearance of the first order spherical harmonics. Spherical harmonics are considered to be sets of three-dimensional orthogonal polar patterns. Orthogonal in this context means, that each spherical harmonic represents a unique directivity of the soundfield, independing of other spherical harmonics. The directional characteristic gets more complex and abstract towards higher spherical higher orders N_{sph} as seen in figure 2.2. As said before, Ambisonics, which is the representation of the soundfield as spherical harmonics or in spherical harmonics domain, of first order can be collected, using simple microphones and their directivity pattern [66]. Higher-order Ambisonics (HOA) are able to represent the soundfield with a finer resolution, but since the corresponding directional patterns can't be achieved using regular microphones, special spherical microphone arrays are needed.

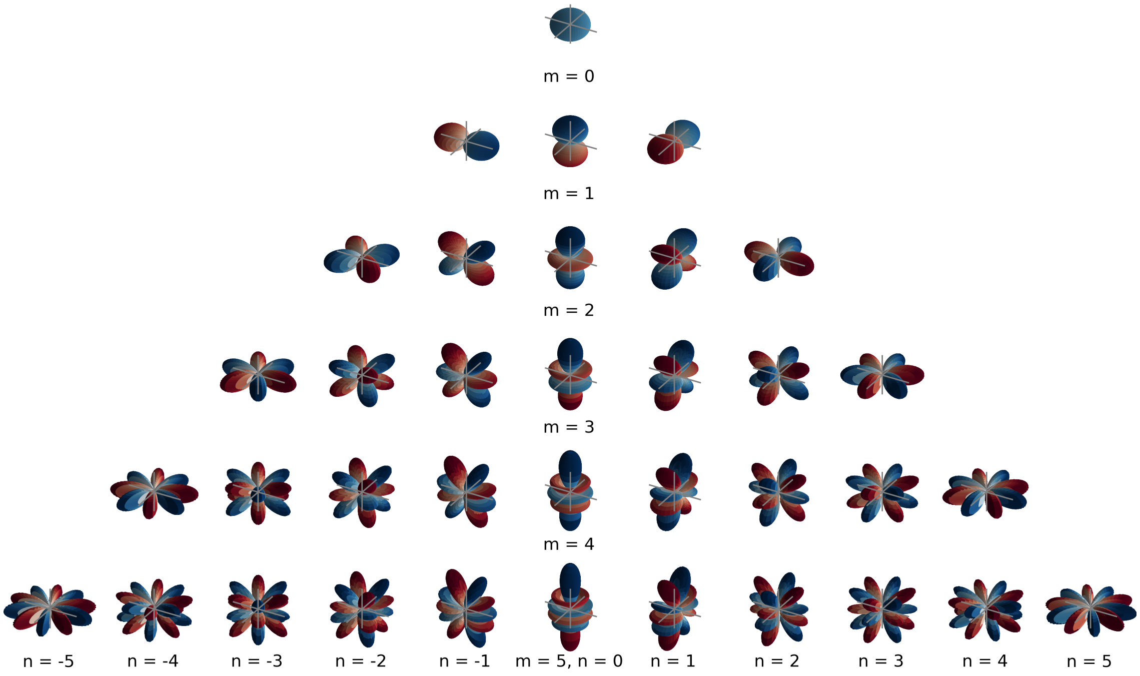

Figure 2.2: Spherical harmonics of order n and degree m

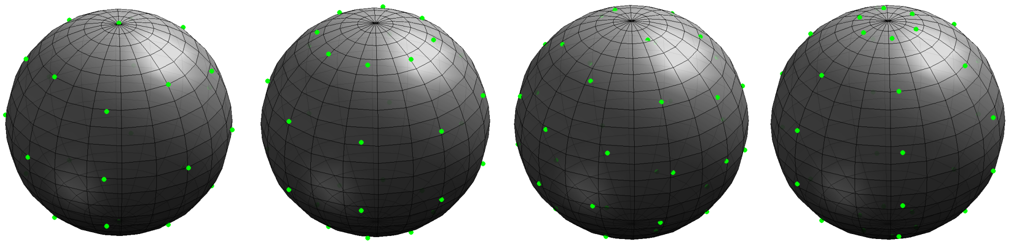

Figure 2.3: Sampling points on the unit sphere. From left to right: Lebedev ($Q = 26$), Gauss ($Q = 32$), T-Design ($Q = 36$), Equal-angular ($Q = 36$)

Spherical harmonics $Y_n^m(\theta, \phi)$ are basis functions on the sphere, which are defined as

$$Y_n^m(\theta, \phi) \equiv \sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^m(cos\theta)e^{im\phi}, \tag{2.1}$$

where m is the degree and n the order of the function. P_n^m are the associated Legendre functions [50, Chapter 1.2]. A function on the sphere $f(\theta, \phi)$ can now be represented by a weighted sum of spherical harmonics as

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\theta, \phi). \tag{2.2}$$

The weights f_{nm} can be derived from the spherical Fourier transform

$$f_{nm} = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) \left[Y_n^m(\theta, \phi)\right]^* sin\theta d\theta d\phi, \tag{2.3}$$

using the inverse of the spherical harmonics functions [50, Chapter 1.4].

2.2.2 Microphone Arrays and Spherical Sampling

For practical applications, a sphere needs to be represented by a limited number of spatial sampling points. This principle applies both to the digital computation of spherical inquiries, as well as the real application for spherical microphone arrays. Here the sampling points are represented by microphones along the surface of a sphere. There is a variety of different sampling methods, which are based on different principles and which have different advantages and disadvantages. A selection of some common sampling grids is shown in figure 2.3. A main requirement of spherical sampling for audio applications is to map and sample the soundfield and thereby estimate the sound pressure function on the sphere. Here, the number of microphones is tried to kept small to limit the complexity of the microphone array, which at the same time decreases the reconstruction accuracy of the original soundfield. A well established approach of spherical sampling is to solve the integral representing the surface of the sphere, using a quadrature formulation [50, Chapter 3]. Starting from a unit sphere function $g(\theta, \phi)$, the value of the definite integral

$$S = \int_0^{2\pi} \int_0^{\pi} g(\theta, \phi) \sin\theta d\theta d\phi \tag{2.4}$$

is numerically approximated by the quadrature method

$$S \approx \sum_{q=1}^{Q} \alpha_q g(\theta_q, \phi_q), \tag{2.5}$$

where Q is the number of sampling points and α_q are the sampling weights. Substituting $g(\theta_q, \phi_q)$ by the inverse Fourier transform $f(\theta, \phi) \left[Y_n^m(\theta, \phi)^*\right]$ from equation 2.3 gives the spherical harmonics representation by a discrete number of sampling points:

$$f_{nm} = \sum_{q=1}^{Q} \alpha_q f(\theta_q, \phi_q) \left[Y_n^m(\theta_q, \phi_q)^*\right]. \tag{2.6}$$

A first approach to gather the sampling points (θ_q, ϕ_q) is to arrange the points equidistantly along the points along the azimuth (ϕ) and the elevation (θ) plane, which is achieved by $2(N_{\text{sg}} + 1)$ sampling points along both planes, where N_{sg} is the sampling grid order. The sampling grid order sometimes is considered individually and different from the spherical harmonic order. For this investigation, the sampling grid order will always be considered to be the same as the spherical harmonic order. As visible in figure 2.3, this kind of sampling does not result in evenly distributed sampling points on the sphere. The density of sampling points increases towards the poles. Similar to that, the Gaussian sampling also uses an azimuthal equal-angle scheme, while the elevation is represented by only $(N + 1)$ samples. Obviously, less sampling points are needed, compared to the equal-angle sampling, but sampling errors may be introduced along the elevation.
More sophisticated approaches are to sample the sphere using vertices of convex Platonic solids as t-designs [24], deriving the grid in relation to a octahedral rotation (lebedev grid) [32] or other nearly-uniform sampling methods like the Fliege-Maier-nodes [19].

2.3 Spatial Audio Signal Processing

Audio signal processing is a fundamental and well studied part of acoustic signal analysis. In the following, the audio processing of spatial audio applications will be discussed in detail. Here, the focus is directed to spatial filtering.

2.3.1 Plane Wave Decomposition

As is widely known, a complex sound field can be described by superposition of single sound waves, if the related system is linear. This principle is a well-known tool for vividly describing the composition of complex wave forms, generated for example by an instrument [53]. Based on this principle, also a complex sound field can be described in terms of multiple plane waves. A plane wave is a special case

Figure 2.4: ZYLIA ZM-1 spherical microphone array [68]

of a travelling sound wave, where sound pressure and velocity are constant along the plane perpendicular to the direction of propagation along x. The plane wave S following can be described as

$$S\left(\vec{x},\omega\right) = \overline{S}(\omega)\ e^{\mathrm{i}\vec{k}(\omega)\vec{x}}, \tag{2.7}$$

where $\overline{S}(\omega)$ is the amplitude of the plane wave and k is the wave number. According to the superposition principle, a complex sound field can then be represented by a infinite number of plane waves, travelling towards the observation point. With equation 2.7 we obtain the plane wave representation of a complex soundfield:

$$S\left(\vec{x},\omega\right) = \oint_{\Omega\in S^2} \overline{S}(\theta,\phi,\omega)\ e^{\mathrm{i}\vec{k}(\theta,\phi,\omega)^T\vec{x}}\mathrm{d}\Omega, \tag{2.8}$$

where Ω is an infinitesimal element of the surface of the unit sphere S^2 and (θ,ϕ) indicate the direction of propagation of the plane wave. Converting equation 2.8 for practical applications, the number of sampling points on the sphere must be broken down from an infinite amount of unit sphere elements Ω to a discrete number of sampling points J, called ξ sector elements from here on. Which turns equation 2.8 into

$$S\left(\vec{x},\omega\right) = \sum_{\xi}^{J} \overline{S}(\theta,\phi,\omega)\ e^{\mathrm{i}\vec{k}(\theta,\phi,\omega)^T\vec{x}}. \tag{2.9}$$

The behaviour of the plane wave decomposition can be shown with a following simple example. The soundfield of single plane-wave is plane wave decomposed. By definition, the expected result would be a replication of the original plane wave as a delta function at the direction of arrival of the single plane wave. However, using a spherical microphone array, the resulting function is replaced by the spatial function

$$f(\theta,\phi) = \frac{N+1}{4\pi(cos\Theta-1)}\left[P_{N+1}(cos\ \Theta) - P_N(cos\ \Theta)\right], \tag{2.10}$$

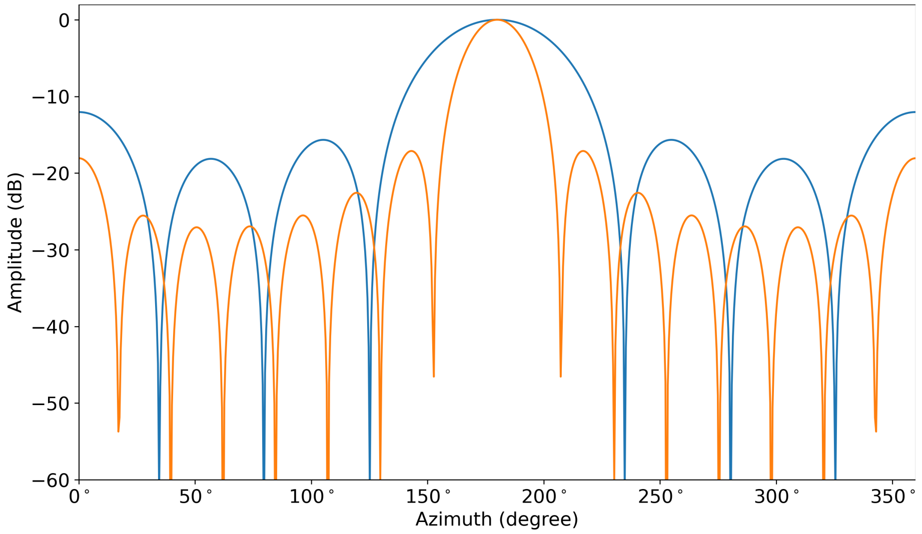

Figure 2.5: Plane wave representation of a spatially band-limited Dirac pulse at $\phi = 180^{\circ}$ of $N = 3$ (blue) and $N = 7$ (orange)

where Θ is the angle between θ and ϕ [50, Chapter 1.5]. The corresponding graph of the angle depending amplitude is shown in figure 2.5. The function consists of a order depending number of side-lobes and a main-lobe, which points in the direction of the plane wave, where the amplitude is the highest. Though, the function has a significant amplitude level at a wide range of directions besides the main direction. Thus, an important property of the processing chain is the width of the main-lobe, which also is order depending. The lobe width is directly related to the spatial resolution of the microphone array and is further discussed in section 2.3.3. The smaller the width of the main-lobe, the more accurate is the differentiation of two plane waves, incident from two angles close to each other.

2.3.2 Spatial Aliasing

A consequence of the discrete spatial sampling is spatial aliasing. In signal processing, the discrete collection and recording of a time signal is restricted by the sampling frequency. The sampling theorem describes the requirement to be met for representing a continuous sound wave by a series of discrete samples. The wave can only be reproduced accurately up to a frequency, which is half of the sampling frequency, the commonly known Nyquist frequency [39]. Signal components of higher frequencies will be projected as lower frequency aliases, since the repetition rate of the samples is to low to represent these higher frequencies properly. This principle describes sampling phenomena in time domain. Similar to that, during the spatial sampling of a soundfield spatial aliasing can occur in the same way. Here, however, the upper limit is not a frequency, but the maximum modal order that is the order of the sampling grid. Similar to the sampling process in the time domain, the limited number of spatial sampling points is not able to represent higher modes, which is why modal aliases are projected into lower modes. Spatial aliasing occurs over the whole frequency range for low as well as for high frequencies, but the level of artifacts increases for higher frequencies [3]. However, the spatial aliasing increases significantly for $kr > N$ [49, Chapter 4.4], with the radius of the sphere r. A spatial aliasing frequency can following be determined as [3, Chapter 3.8]

$$f_{\mathrm{sA}} = \frac{N_{\mathrm{sg}} c}{2\pi r}. \tag{2.11}$$

For practical applications spatial aliasing is negligible for frequencies below f_{sA}.

2.3.3 Modal Weighting

While creating the plane wave components for the plane wave decomposition, modal weighting is applied to the plane waves to achieve certain properties. Properties here mainly are the directivity of the plane wave, the lobe width and the control of unwanted side-lobes. The directivity is specified as directivity index (DI), which, similar to the directivity index of a microphone, is defined as the ratio between the response of the array to a plane wave in the direction of reference. Thereby, the modal weighting is order depending. Rafaely [51, Chapter 6.1] derives a maximum directivity beamformer, resulting in an axis-symmetric directivity function, which returns modal weighting as a constant factor of

$$c_n = \frac{4\pi}{(N+1)^2}. \tag{2.12}$$

The maximum directivity (max DI) beamformer as intended by its name, has the highest directivity index, compared to other weighting functions. Resulting from the directivity, the smallest possible angle between two arriving plane waves, that can be spatially separated by a sampling grid or respectively a microphone array, can be computed. Using max DI beamformer, this angle, or lobe width, can be approximated with

$$\Theta_0 \approx \frac{\pi}{N}, \text{ for } N > 4, \tag{2.13}$$

as proved by [48]. A corresponding minimum angle to separate two arriving plane waves, for a plane wave decomposition using max DI weighting and order $N = 5$ would be $\Theta_0 = 36°$. The high directivity though results in the disadvantage of side-lobes which contain a considerable part of the energy. Another common approach here is to use max R_e weighting. It is commonly used in audio applications, since the weighting optimizes the energy distribution towards the pointing direction and side-lobes are suppressed [65]. A comparison of both functions is displayed in figure 2.6 (right). The max r_E weightings are approximated using the Legendre Polynomials as

$$c_n = P_n\left(\cos\left(\frac{137.9°}{N+1.51}\right)\right). \tag{2.14}$$

2.4 Zero-phase Frequency Filtering

The design and application of filters that affect the spectrum of a signal is a well established tool in audio signal processing [41, i.a.]. A filter, either digital or analog, generally is an element that manipulates the amplitude of a signal depending on

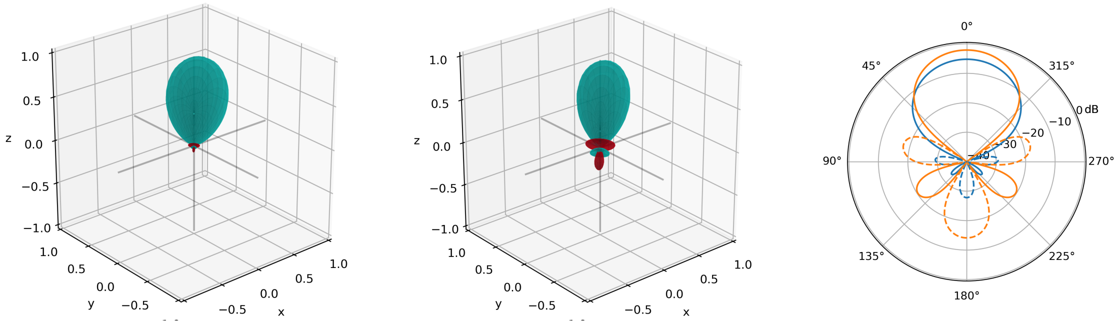

Figure 2.6: Left and Center: Weighting functions of plane wave sectors in 3D (Pointing direction: $\phi = 45°, \theta = 45°$): max r_E (left), hypercardioid (center), Right: Comparison of both functions in one 2D polar presentation. max r_E (blue), hypercardioid (orange)

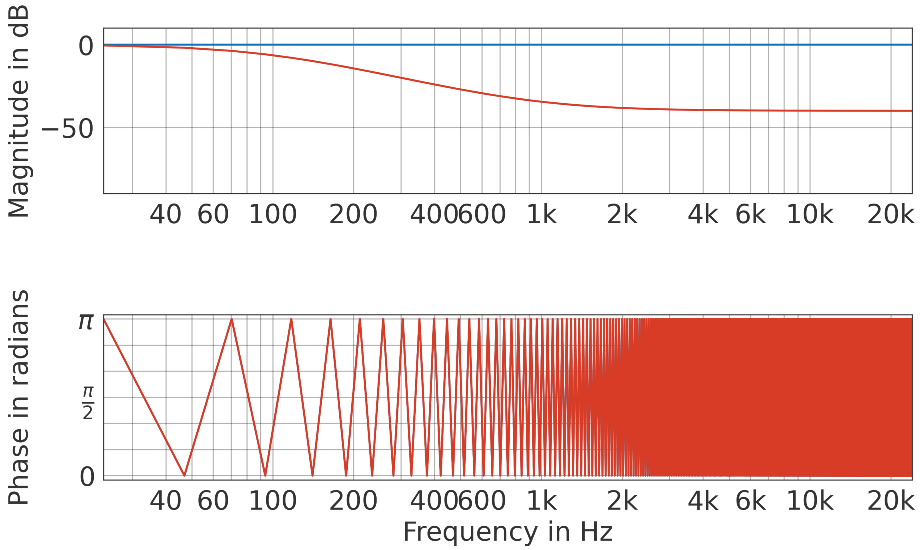

Figure 2.7: Frequency and phase response of a forward-backwards first order high shelf filter ($f_c = 945$ Hz, $G = -20$ dB)

the frequency. Here, the filter can occur with different features, whereby common examples are lowpass, highpass, bandpass or allpass filters. Unlike a lowpass filter, which increasingly attenuates a signals level in the stopband towards high frequencies, a shelving filter attenuates (or enhances) the signals level to a certain value. An example is shown in figure 2.7. A filter can be applied to the signal in frequency domain, by multiplication of the frequency response of the signal $X\left(e^{j\omega}\right)$ to the frequency response of the filter $H\left(e^{j\omega}\right)$ or in time domain by convolving the time signal with the filters impulse response h. The filtered result in frequency domain $Y\left(e^{j\omega}\right)$ is obtained by

$$Y\left(e^{j\omega}\right) = X\left(e^{j\omega}\right) \cdot H\left(e^{j\omega}\right). \tag{2.15}$$

With the application of the filter, a change of the signals phase response is introduced.

For some applications it is crucial, that the resulting signals phase is not delayed or distorted. For this purpose, non-zero filtering can be used. A common approach for that is to apply a filter twice. Here, the filter is first applied forwards and after that backwards. That's why it is also known as "forward-backward filtering".

$$Input \longmapsto Filter \longrightarrow TimeRev \longrightarrow Filter \longrightarrow TimeRev \longrightarrow Output$$

Time reversing here simply corresponds to the conjugate complex of the signal. Therefore equation 2.15turns into

$$Y_{rev}\left(e^{j\omega}\right) = conj\left(X\left(e^{j\omega}\right) \cdot H\left(e^{j\omega}\right)\right) = X^*\left(e^{j\omega}\right) \cdot H^*\left(e^{j\omega}\right) \tag{2.16}$$

after time reversing. Next, applying the filter forward and backwards results in

$$Y_{fb}\left(e^{j\omega}\right) = conj\left(X^*\left(e^{j\omega}\right) \cdot H^*\left(e^{j\omega}\right) \cdot H\left(e^{j\omega}\right)\right),$$

which simplifies to

$$Y_{fb}\left(e^{j\omega}\right) = X\left(e^{j\omega}\right) \cdot \left|H\left(e^{j\omega}\right)\right|^2. \tag{2.17}$$

The forward-backward filtering ends up in multiplying the spectrum of the signal with the squared magnitude response of the filter, which by definition only has real and positive values. Therefore the phase response is zero and no phase shift or similar is introduced. Since the magnitude is squared, the level changes induced by the filtering are twice as high as if the filter would be applied regularly.

An important property of the forward-backward filter is its non-causality. So its output is not only depending on current and past input values, but it is also depending on its future input values. Therefore the whole signal needs to be available for this kind of filtering, that's why it is not applicable for real-time filtering.

The implementation of such a filter will later be important for this investigation, when it comes to the processing of signals that are depending on each other.

2.5 Sound Propagation and Edge Diffraction

If the surface of an object vibrates and is inside a propagation medium, it forces the particles of the medium close to the object to vibrate. The particles get displaced from their equilibrium position, approach neighbouring particles resulting in a pressure increase and pass on the vibration. A pressure wave propagating away from the source is generated depending on the properties of the medium and the source. The pressure of the wave travelling away from the source can be defined as

$$p(r,t) = \frac{A_+}{r} e^{j(\omega t - kr)}, \tag{2.18}$$

derived from the wave equation of a harmonic spherical wave and depending on distance r and time t depending. A_+ is the Amplitude at 1m distance, ω the angular frequency and k the wavenumber [29].

An important acoustic property here is the acoustic impedance Z, which describes the resistance a system has when acting against an applied acoustic pressure

$$Z_{\text{ac}} = \frac{p}{v}. \tag{2.19}$$

If a travelling wave hits an object, the impact can be acoustically be described as an impedance change of the travelling medium. From the ratio of the impedance of the object and the travelling medium, the following behaviour of the sound wave can be derived. If the ratio is significant, most of the wave is reflected. If it is small and the impedances are similar, the wave is transmitted. If a sound wave hits the edge of a object the sound wave is diffracted. Sound diffraction describes the deflection of propagation direction of a wave from its initial path, when it hits the edge or corner of a rigid object. Edge diffraction of sound waves in particular, which follows similar phenomena of diffraction of electromagnetic light waves, was investigated over the last decades by different scientists and different models for the calculation and prediction of diffracted sound were proposed. The models hereby are based on different approaches. Published in 1968, Maekawa presented a empirical model, which describes the magnitude of the diffracted sound depending on the detour of the sound propagation path around the diffracting object [35]. The model represents a simple approach, which, especially for higher order diffraction, is not very accurate [54]. High order diffraction stands for the sound diffraction around multiple edges. A more sophisticated approach leads to Svenssons secondary source model of edge diffraction [61], which also enables higher order diffraction calculations. The model is based on the Biot-Tolstoy expressions and Medwin's extension for this. For the calculations of edge diffraction impulse responses a set of secondary sources along the diffraction edge is introduced. By means of analytical directivity functions, the impulse responses can be computed.

2.6 Noise barriers

A noise barrier is a physical object that is placed between a noise source and a receiver in order to significantly reduce noise immissions at the receiver position. Noise barriers are a common instrument to counteract environmental noise and are widely used against road traffic, railway and ground-operating aircraft noise. Thereby, depending on the use-case and the operation location, noise barriers appear in many different physical configurations. As shown in figure 2.8, the direct sound is interrupted to travel from the source to the receiver. As it hits the barrier, a part of the direct sound is reflected back from the barrier, some sound energy is absorbed and translated into vibration and thermal energy and a part is transmitted through the barrier. Additionally the sound that hits the top and side edges of the barrier is diffracted around the barrier, to the so called acoustic shadow zone, and therefore reaches the receiver. The amount of reflected, transmitted, absorbed or diffracted sound energy strongly depends on the geometrical and material-wise configuration of the noise barrier and depends on frequency. Generally noise barriers are made of rigid framework in combination with sound absorbing elements in between. When the mass per unit area is above 20 $\frac{\mathrm{kg}}{\mathrm{m}^2}$ sound transmission through the barrier can be controlled and be kept to a minimum [30]. Hereby the heavy mass of the noise barrier corresponds to a high acoustic impedance, especially towards high frequencies. A sudden change of acoustic impedance is introduced, which results in a high reflection coefficient for incident sound waves. Noise barriers have different common appearances depending on their operational location. Worth mentioning are

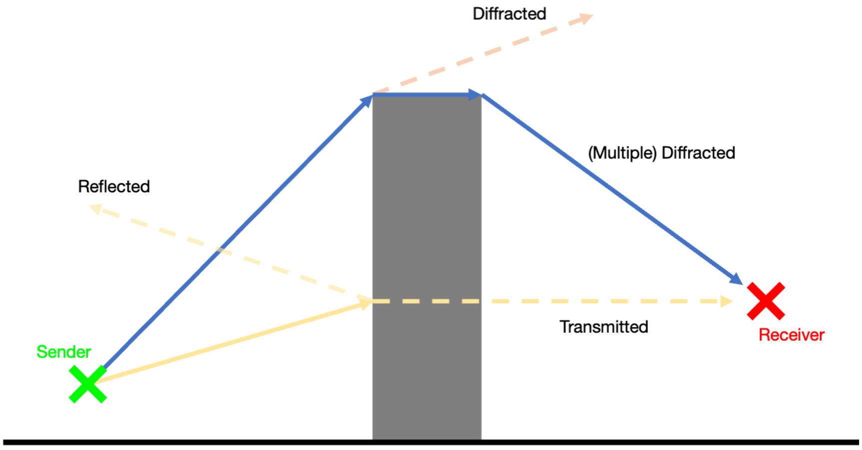

Figure 2.8: Schematic presentation of a noise barrier an the propagation paths between sender and receiver

natural barriers, which can cleverly take advantage of topographical configurations or be constructed from earth-berms. Here, vegetative growth can have additional beneficial absorbing properties [30]. Besides that there are tunnel configurations or commonly known noise screens along the noise source, which can appear in single or multiple top edge configurations, where multiple order diffraction becomes particularly important. There is a number of different shapes and materials that can be used for the construction of noise barriers, which have to be selected based on performance, stability and duration considerations. Studies have shown that the top edge has a significant influence on the performance of the noise barrier, whereby it was determined that T-shaped and scattering surface barriers with soft top covers are most effective [31].

3

Methods

Following the implementation of directional frequency filters and the corresponding signal processing, the calculation of impulse responses of sound barriers and their application as filters to recordings of real traffic scenarios in spherical harmonic domain is explained. Furthermore the measurement setup for the recording of spatial room impulse responses in a controlled environment are described. The measurements are the basis for the evaluation of the implementation of the directional filters using noise barrier transfer functions.

3.1 Sound Barrier Impulse Responses

Next the calculation of impulse responses are described in detail. The impulse responses later can be used as filter functions, thus representing the acoustic influence of the sound barriers. Their calculations are based on the secondary source model by Svensson, described in chapter 2.5.

The starting point of the calculation is an acoustic free field. Therefore, it consists of an infinite large ground plane and no additional reflecting surrounding surfaces. In this environment a point source and a receiver are introduced, both at fixed points. A cuboid of finite length, width and depth, which is connected to the ground on one side, is placed between the source and the receiver to function as a sound barrier. If the barrier is placed in direct line between source and receiver, no direct sound can subsequently reach the receiver. Due to the absence of sound reflecting obstacles, only sound waves diffracted around the barrier reach the receiver. Depending on the sound absorption characteristics of the ground, sound waves reflected from the ground could also reach the receiver. The resulting paths of the sound which reach the receiver are shown in figure 3.1. Here a simplified two-dimensional presentation of the model is displayed, where only the sound paths along the top edge of the barrier are shown. Following statements apply in a similar way to the paths around the side edges of the barrier.

There are four paths from source to receiver:

$P_{\text{d-d}}$: Source -> Edge -> Receiver
$P_{\text{d-r}}$: Source -> Edge -> Ground reflection -> Receiver
$P_{\text{r-d}}$: Source -> Ground Reflection -> Edge -> Receiver
$P_{\text{r-r}}$: Source -> Ground Reflection -> Edge -> Ground Reflection -> Receiver

To make the model feasible for calculations in MATLAB, it has to be adapted. Since an infinite large ground is impracticable, the model is translated to an image source model, which is displayed on the right in figure 3.1. Here the ground is removed, an

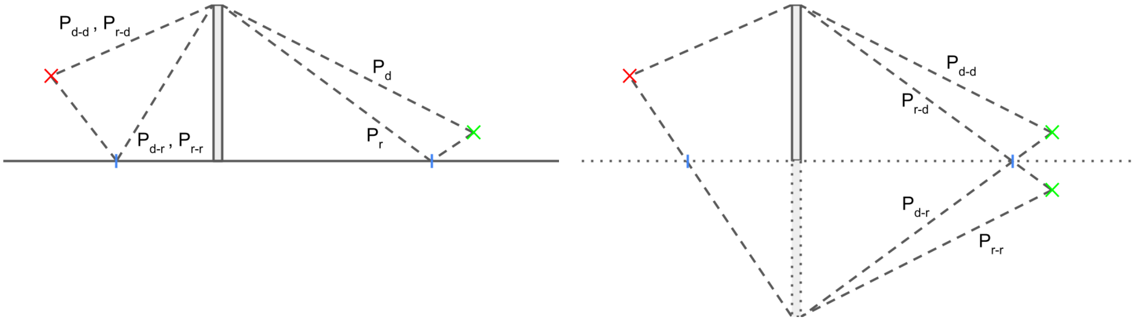

Figure 3.1: Schematic representation of the sound barrier model in 2D: Actual setup on the left and image source model on the right.

image source is added by mirroring the source along the original ground line. The cuboid barrier is doubled in height by mirroring it along the ground line as well. Consequently the paths listed above can perfectly be reproduced.

3.1.1 Implementation in MATLAB

The model is implemented in MATLAB and the use of the EDtoolbox [60]. An exemplary geometrical implementation of the image source model for the calculation of the impulse responses is shown in figure 3.2. The toolbox offers a variety of functions to calculate scattered sound pressures for polyhedral scattering objects with rigid surfaces. The sound pressure is computed at the receiver position considering direct sound, specular reflected sound and first and higher order diffracted sound, originating from a point source. Since the free-field radiating monopole source emits a normalized dirac impulse with an amplitude of 1, the resulting sound pressure can be viewed as a impulse response. Direct sound and specular reflected components are computed by means of visibility functions $V_{\mathrm{S|IS,\ R}}$. If the path from source S to receiver R is not obstructed it takes the value $V_{\mathrm{S,R}} = 1$. On the other hand, if it is obstructed $V_{\mathrm{S|IS,\ R}} = 0$. For the case of specular reflections the path is constructed from the image source IS to the receiver R, using the image source method. The impulse responses are then defined as free-field radiating monopole sources

$$\mathrm{IR}_{\mathrm{dir|spec}}(t) = \frac{\delta(t - \frac{r}{c})}{r}\, V_{x_{\mathrm{R}},\ x_{\mathrm{S|IS}}} \tag{3.1}$$

derived from equation 2.18. First and higher order diffraction components are then calculated, as described in section 2.5.

3.1.2 Assumptions and Limitations of the Model

The implementation of the secondary source model comes along with certain limitations, which have to be taken into account for further investigations.

The model computes the impulse responses in an acoustic free-field environment. This kind of environment can only be assumed for some specific scenarios in reality. Especially for traffic noise scenarios, which are considered relevant for noise protection measures, reflecting surrounding surfaces are present frequently, for example if buildings surround the scene. Furthermore, the image source model makes

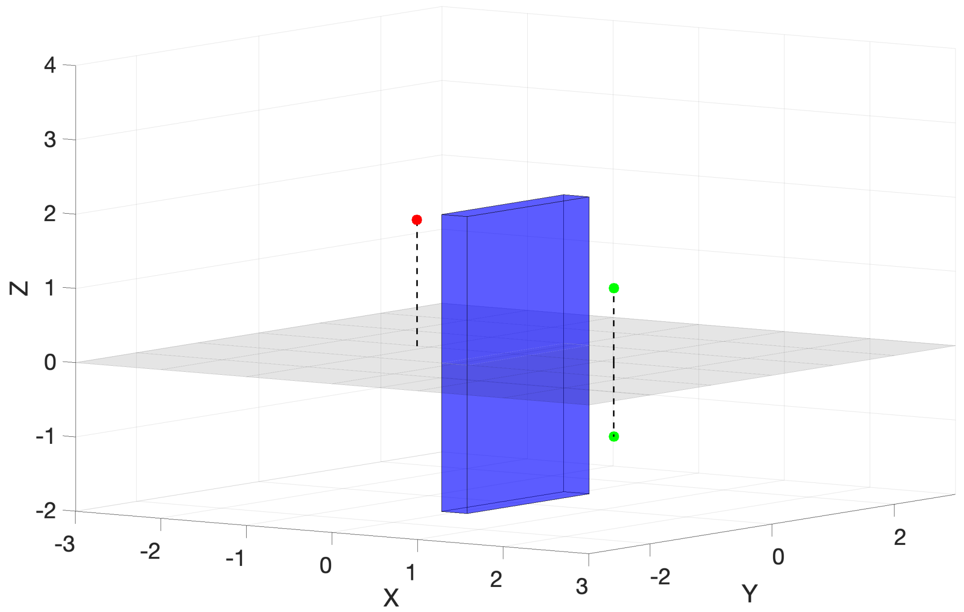

Figure 3.2: MATLAB implementation of the image source model for a calculation of a impulse response. sound barrier (blue): length = 2m, height = 2m (mirrored along the floor), thickness = 0.3m; Source and image source (green); receiver (red)

it possible to implement the model containing a ground, however, the ground can only modeled to be fully reflecting and perfectly flat. In practice, ground has a much more complex structure and reflecting or absorbing properties. Additionally to that, the wall is also implemented as perfectly reflecting object consisting of only perfectly straight edges. This means, no sound is transmitted through the barrier. As described in chapter 2.6, sound transmission through sound barriers usually can be neglected, but for thin barriers with insufficient mass per unit area, transmission especially at low frequencies has to be considered.
It must be mentioned that the relatively short length of the wall in figure 3.2 was set this way because it corresponds to the real wall measured later, as described in section 4.2. In doing so, the length had to be restricted according to the size of the measurement room. Sound barriers in real traffic noise situations are known to be significantly longer. However, this restriction does not affect the quality of the methodology.

3.2 Measurements

To create validation data for the processing of ambisonic recordings and the simulation of sound barriers, measurements are made in a controlled manner. The goal is to achieve comparable sets of measurements of similar geometrical setups with and without a physical wall. Subsequently, the measurements without a physical wall can be processed by applying the simulated impulse responses of a corresponding sound barrier, described in chapter 2.5. As a result, the measurement of a physical

wall can be directly compared to the processed ones with a simulated wall. Thus, a comparison between the real and the simulated impact of a sound barrier on the soundfield is possible objectively, through spectral and spatial analysis, as well as subjectively with a listening test.
Originally, one intention was to carry out the measurements in situ. To be able to realize measurements for this purpose in situ, certain conditions that must be met by the measurement location. The environment needs to be a quasi acoustic free-field environment so no strong reflecting surfaces other than the ground should be nearby. Furthermore, the ground should be relatively even and uncontrolled background noise sources must not be nearby. Another main requirement is that there are two similar locations like that, where one contains a sound barrier or a similar wall of reasonable size and one which contains no wall. Since no matching locations were available, the measurements were moved to an anechoic environment. Consequently, measurement inaccuracies caused by atmospheric and surrounding conditions could be eliminated, which would be difficult to control in an in situ environment.

3.2.1 Setup

The idea of the measurements is to record spatial impulse responses which later give the possibility to convolute them with any desired audio signal for further investigation. Here, a requirement is to have the possibility to simulate a moving source, that moves behind a wall and a source that simulates background noise. Therefore, the measurement setup consists of a receiver position, four source positions, which later can be used as moving source, and a fifth additional source position for background noise. For the measurements including a wall, a 2 by 2 meter plate is located in such a way, that it interrupts the line of sight of at least one source to receiver.
To obtain data of different geometrical configurations, four different geometrical setups were used for the measurements, for which the distance of the sources and the receiver differ, respectively. Origin for the following dimensions is the center point at the bottom within the wall. The coordinates are shown in table 3.1 and 3.2.1

Table 3.1: Coordinates of the different setups

Coordinates	X	Y	Z
Receiver	see Tab. 3.2.1	0	1.7
Source 1	see Tab. 3.2.1	-3	1.5
Source 2	see Tab. 3.2.1	-1.5	1.5
Source 3	see Tab. 3.2.1	0	1.5
Source 4	see Tab. 3.2.1	1.5	1.5
Source 5	-6	1	1.5

Table 3.2: Varying X-Coordinates for sources and receiver

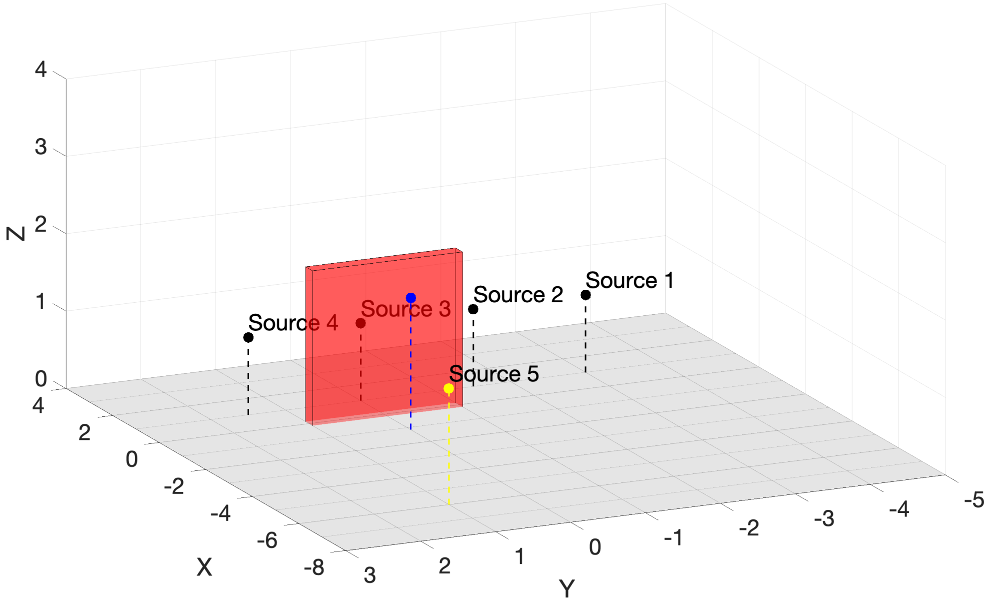

Figure 3.3: Sketch of the geometrical setup of the measurements.
Black: Position of loudspeakers in a line source configuration. Yellow: Position of a secondary source loudspeaker. Blue: Position of the receiver microphone array

Setup	Setup 1	Setup 2	Setup 3	Setup 4
X_{Receiver}	1	4	1	4
$X_{\text{Source 1-4}}$	1	1	3	3

Setup 2 is shown in figure 3.3 exemplary. The measurements took place in the anechoic chamber of the TU Berlin, with a volume of $V = 1070$ m^3 and a lower cut-off frequency of $f_{\text{low}} = 63$ Hz.

3.2.2 Equipment

The wall is made of medium density fibreboard with a thickness of 25 mm and a mass per unit area of 17.53 kg/m^2. An absorbing material (Rockfon SONAR-G) with a thickness of 20 mm is glued on top of the plate. As receiver, a Zylia ZM-1-3E microphone array is used. The array consists of 19 omnidirectional digital MEMS microphone capsules, which are distributed on a sphere of radius $r_{\text{array}} = 49$ mm and has a spatial aliasing frequency of $f_{\text{sA}} = 3.3$ kHz (cf. equation 2.11) and a speed of sound of $c = 345\ \frac{\text{m}}{\text{s}}$. It is capable of recording up to $N_{\text{sph}} = 3$. As sources five Neumann KH 120 loudspeakers, which are connected to a MacBook through a RME Fireface UC, have been used.
The recordings were made with a sampling frequency of $f_{\text{s}} = 48$ kHz. To measure the impulse responses a sine sweep was used as measurement with a sweep-length of 2^{19} samples.

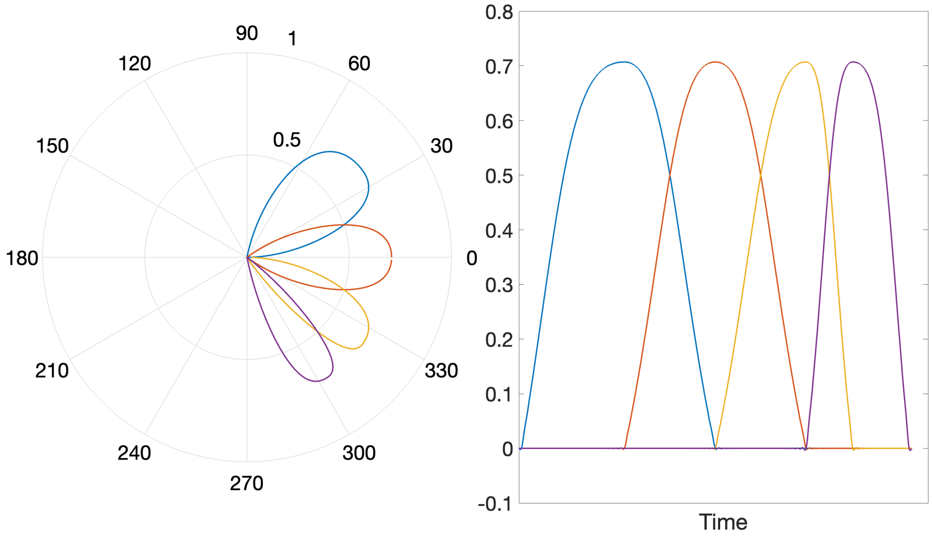

Figure 3.4: Channel gains using vector based amplitude panning for a moving source from 280° to 80° represented by the four source positions form setup 1. Channel 1 (Purple), Channel 2 (Yellow), Channel 3 (Orange), Channel 4 (Blue). Left: Polar representation; Right: Time domain representation

3.2.3 Simulation of Moving Sources

The simulation of a moving source can be implemented using simple vector based amplitude panning (VBAP). The gain of each source here is increased or reduced depending on the distance vector length between the position of the virtual source and the receiver. For such two dimensional cases, the position of the virtual source can be represented by two real sources [47]. In figure 3.4 the channel VBAP-gains for a moving source are shown exemplarily. Here, the moving source is simulated using the source positions from setup 1 to generate a source moving from left to right in 10 seconds.

3.3 Spatial Frequency Filtering using Plane Wave Decomposition

To access discrete directional components of the soundfield a spatial filterbank was used, which basically performs a plane wave decomposition of the soundfield [26]. Here the soundfield is decomposed in Q directional plane waves, called sectors from now on, which represent the soundfield. For this purpose a matrix with a set of spatial sector weightings is applied to the input signal in SH domain σ_{nm}^{in} (equation 3.2). The matrix of shape $[Q \times (N_{\text{sph}} + 1)^2]$ originates from the spherical harmonic basis functions $\mathrm{Y}_{Q,nm}$ evaluated using modal weighting c_n.

$$s_\xi = \sigma_{nm}^{in} \; \mathrm{Y}_{Q,nm} \; diag_{N_{sph}}\left(c_n\right) \tag{3.2}$$

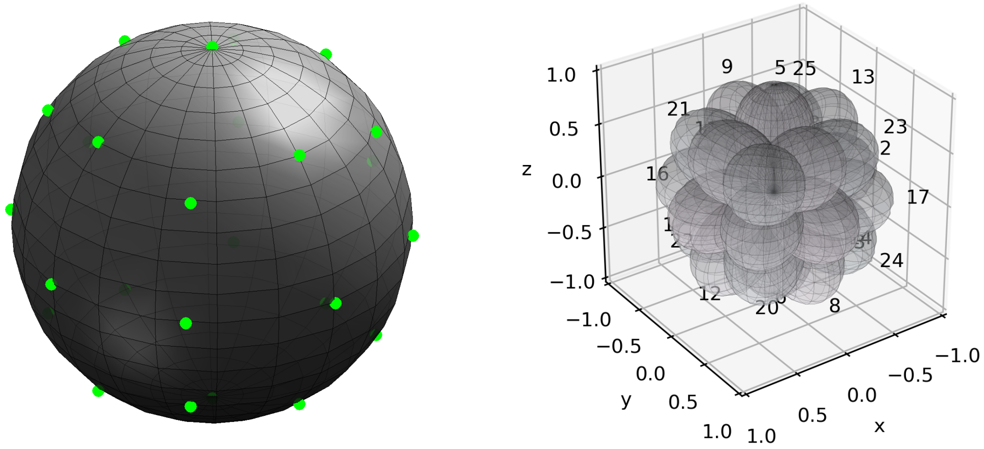

Figure 3.5: Sampling of the sphere (left) and visualized sectors resulting of the plane wave decomposition (right) using a $N_{\mathrm{sph}} = 3$ Lebedev sampling grid (Q = 26)

The operator $diag(\cdot)$ expands each vector entry m times to a diagonal matrix. This results in s_ξ, which contains $\xi = 0, ..., Q$ sector signals. max$_r E$ weights are chosen as modal weights for the plane wave decomposition, since their high front- to side-lobe ratio reduces the root-mean-square error for the plane wave decomposition and reconstruction as described in section 2.3.3. If combined again the plane waves result in a reconstruction of the initial soundfield. The reconstruction is carried out by applying the pseudo-inverse matrix to the sector signals.

The number of sectors is limited to the spherical harmonic order of the input signal and the type of the grid sampling methods, described in section 2.2.2. Since the Zylia microphone array for the recordings of real soundfields that is available for the investigations in this thesis, is capable of recordings up to a spherical harmonic order $N_{\mathrm{sph}} = 3$, the following evaluation mainly focuses on that specific order. An example of the spherical sampling and the sectors is shown in figure 3.5. Subsequently, potential performance improvements of the processing using higher spherical harmonic order $N_{\mathrm{sph}} = 7$ is analysed later. As it can be seen in the figure of the visualized plane wave sectors, single sectors overlap each other. To get a better insight into the overlapping, another figure displays the directivity of three plane waves of three neighbouring sectors ($\xi = [0,\ 2,\ 14]$) as two dimensional polar plots in figure 3.6. As mentioned earlier, the level of the side-lobes is very small compared to the main-lobe due to the max r_E weighting. It becomes clear that the processing with such a low order produces plane wave components, which overlap in a wide directional range. After decomposing the input signal into its plane wave components, those sectors can be analysed and manipulated individually. At this stage the frequency filtering takes place, where the filters can be applied on single sectors, pointing at certain directions. Particular caution is required while filtering. The sector signals represent the complex sound field which is why they depend on each other. Especially the phase relation between sectors that are filtered and that are not filtered should not be shifted, since relative phase changes will distort the time signals of the sectors between each other. The phase relation of unfiltered and filtered sectors is shown in section 4.3 more in detail. For this reason, the filter is applied, using forward-

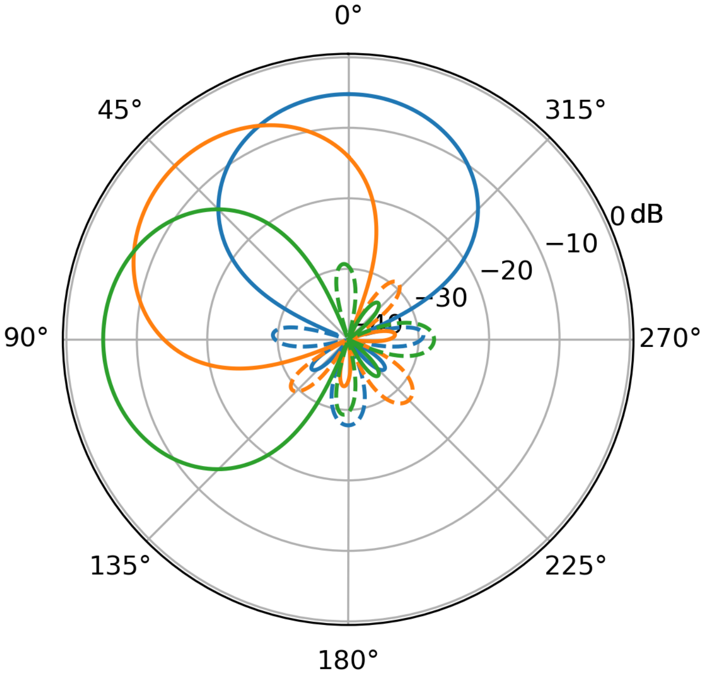

Figure 3.6: Polar representation of neighbouring sectors of a lebedev sampling grid ($N = 3$) using max r_E weighting

backward filtering, described in section 2.4, which has no effect on the phase of the signal. As noted previously the filtering should only affect sectors, that contain the plane wave components which represent the sound field, arriving from the direction in which the sound barrier should be placed. Consequently the sectors which point in the direction of the wall need to be detected. Therefore as a first step, the wall surface is projected on the unit sphere. Afterwards it is checked for each sector whether its direction vector intersects the projected surface. If so, the corresponding surface is filtered. An example is shown in figure 3.7. The projection of the wall for setup 1 & 3 is well mapped by the selected sectors, since the directions of the sectors fit well to the edges of the wall coincidentally in this case. Looking to the projected wall of setup 2 & 4 it can be seen though that the small spherical harmonic order and consequently a small number of sampling points leads to an inaccurate representation of the wall by the sectors. Applying the described procedure, the wall is here only represented by one sector. The direction of this sector doesn't point at the middle of the wall, so it is slightly pointing upwards, which can be corrected by rotating the sampling grid correspondingly. Also, as mentioned in section 2.3.3, the sectors overlap due to the applied $\max_{rE}$ weightings, which results in directions at the edges of the wall, which are represented both by filtered and unfiltered sectors. It becomes clear that the sectors can only represent the dimensions of the noise barrier with limited accuracy. This can be counteracted to a certain extent by a clever arrangement of the sampling grid.

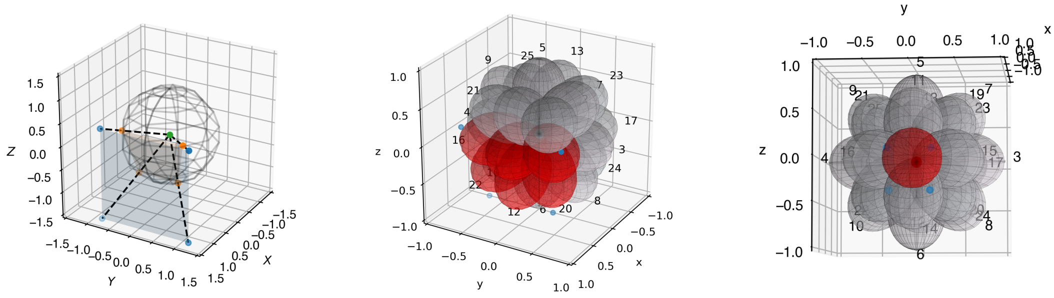

Figure 3.7: Left: Visualization of the projection of the wall (blue) onto the unit sphere (orange). Sector that intersects with the projection highlighted in red for setup 1 & 3 (center) and setup 2 & 4 (right)

3.4 Signal Chain

After gathering all needed impulse responses through the measurements described in section 3.2, any desired moving or static source can be auralized. The two signal chain logic for the auralization of a static source and a moving source are shown in figure 3.8 and figure 3.9 respectively. Starting point for both scenarios are the measured impulse responses, described in section 3.2. The measured impulse responses contain the response of the transfer path between the respective source position to the receiver position in the spherical harmonic domain. Convoluted with a mono input signal it generates the, at the receiver position arriving signal in spherical harmonics domain, played by the speaker at the particular source positions. In the case of a desired moving source, the input signal is convoluted with the impulse responses of all four source position. Subsequently, the corresponding VBAP-gain (section 3.2.3) of each source position is applied to all 16 channels and the four resulting source signals are summed channel-wise. If desired, the calculated impulse response of a sound barrier, described in section 3.1, can be applied to the signal. For this, the signal is plane wave decomposed, as described in section 3.3. Hereinafter, the desired sectors are frequency filtered by convolving the sector signal with the impulse response of the filter forwards and backwards, as explained in section 2.4. After the filtering, the sectors are reconstructed again, which returns the, at the receiver position arriving signal of a moving or static source, generated by the corresponding source setup, including a desired sound barrier. As a final step, the resulting spherical harmonics signal can be decoded for binaural playback, using individual head related impulse responses including head-tracked rotation of the acoustic scene [66]. Using this signal chain, three different scenarios can be generated, all using either a static single source position or a moving source, along the line of speakers. Those scenarios are:

- **"Empty" scene without a noise barrier**, using impulse response measurements without a physical noise barrier.
- **Scene with a real noise barrier**, using impulse response measurements in the presence of a physical noise barrier.
- **Scene with a simulated noise barrier**, using impulse response measure-

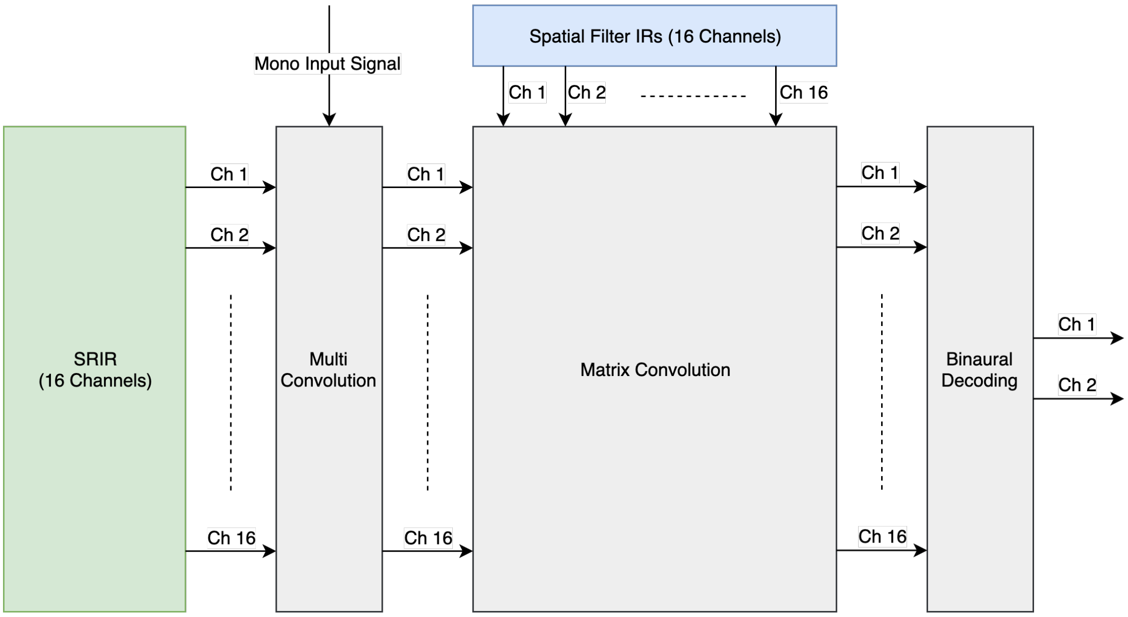

Figure 3.8: Signal block chain for the auralization of a static source. Measured impulse responses (green), Simulated sound barrier impulse responses (blue)

ments without a physical noise barrier and applying impulse responses of a simulated noise barrier within the plane wave decomposition.

On the basis of those scenes, a MUSHRA listening test [27] on the evaluation of the quality of the directional frequency filtering using the plane wave decomposition of a complex soundfield in the spherical harmonic domain, can be conducted. Quality parameters should be created from a selection of the spatial audio quality inventory [34].

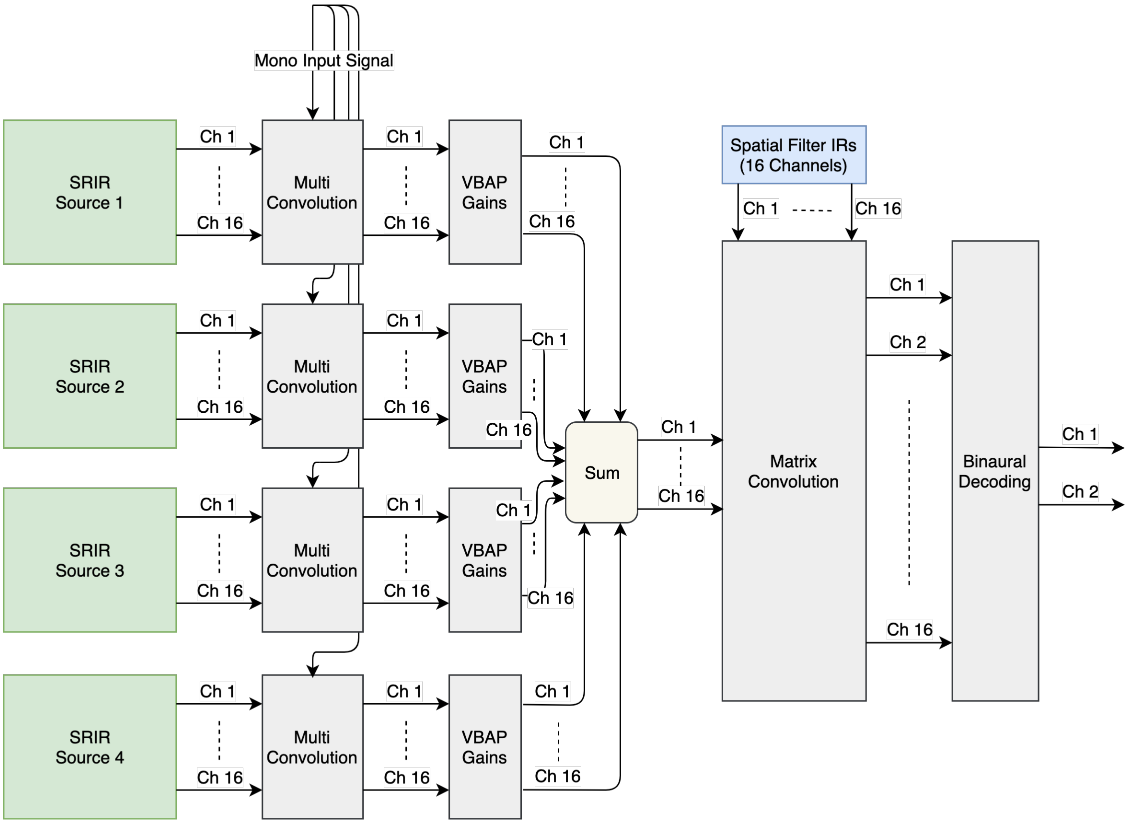

Figure 3.9: Signal block chain for the auralization of a mvoing source. Measured impulse responses (green), Simulated sound barrier impulse responses (blue)

4

Results

In this chapter results of the previously described methods for the calculation of noise barrier impulse responses, the measurements of spherical room impulse responses, the auralization of scenes using measured and calculated impulse responses and the evaluation of the directional frequency filtering are displayed and discussed more in detail.

4.1 Secondary Source Diffraction Model

As a starting point, the results of the simulation of noise barriers and the computation of corresponding impulse responses is elaborated here.

4.1.1 Noise Barrier Impulse Responses

The simulation of impulse responses of noise barriers described in section 3.1 is implemented. The resulting impulse responses are displayed in frequency domain as frequency responses. The calculations were carried out with the speed of sound in air $c_{air} = 344 \frac{\mathrm{m}}{\mathrm{s}}$ and the air density $\rho_{air} = 1.21 \frac{\mathrm{kg}}{\mathrm{m}^3}$. Here, only high order diffraction up to the second order is included. This limitation was made in order to save computation time. Minor differences in the impulse responses, visible in figure 4.1, due to higher order diffraction are not needed for the project. Second order diffraction is sufficient. In figure 4.2 the calculated frequency responses of all 4 setups from section 3.2 are shown. Since positions of source 2 and source 4 are in a geometrical symmetric configuration with respect to the wall and the receiver position, the calculations for those two positions result in two identical frequency responses. As described in section 3.1, no second mirror source or mirrored wall was used for the calculations here, because measurements were conducted in a fully anechoic chamber without sound reflecting floor. There are a several acoustic phenomena that are relevant for the appearance of the calculated frequency responses.

- Attenuation due to sound propagation
- Diffraction
- Superposition and interference of sound waves

The frequency responses of the sources whose lines of sight are not intercepted by the wall, are dominated by the direct sound component and are mainly flat over the frequency range. Here the attenuation due to basic rules of sound propagation is visible best. The direct sound is attenuated depending on the distance of the source to the receiver. For example the distance of source 1 in setup 2 is $d_{S1} \approx 5.8$ m leads

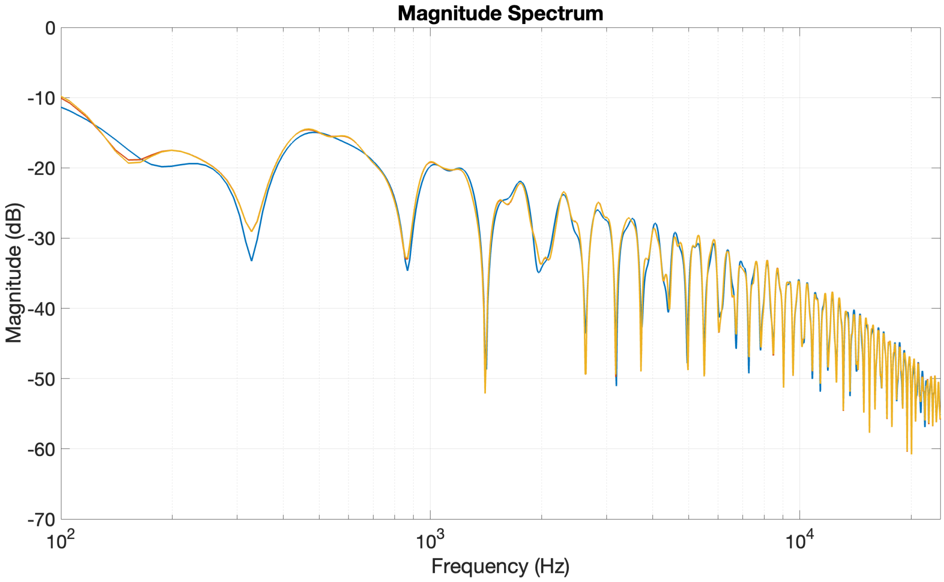

Figure 4.1: Simulated frequency responses (Setup 1, Source 3) using different amount of higher order diffraction. $N_{\text{diff}} = 2$ (blue), $N_{\text{diff}} = 3$ (orange), $N_{\text{diff}} = 4$ (yellow)

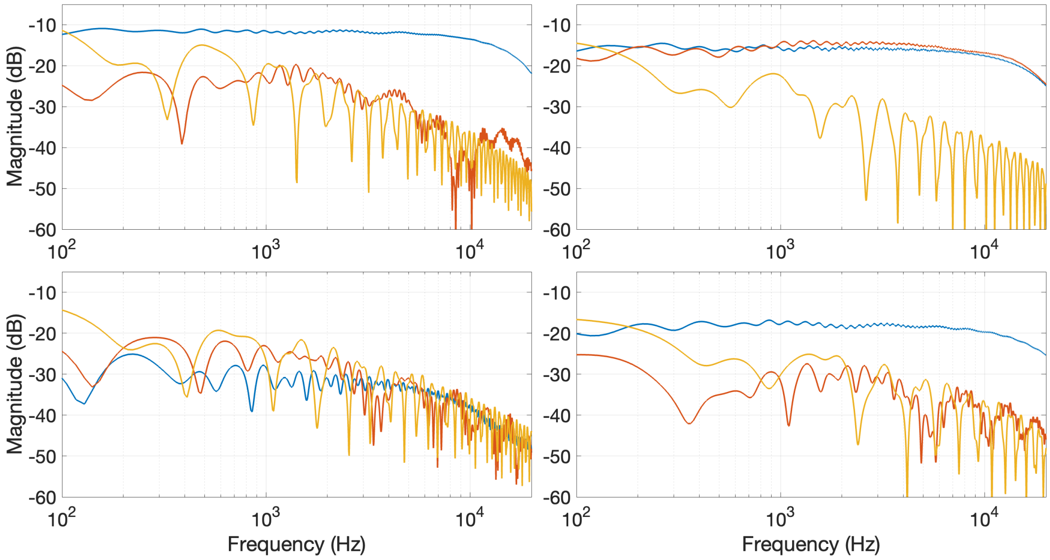

Figure 4.2: Simulated Impulse Responses of all measurement setups. Top left: Setup 1, Top right: Setup 2, Bottom left: Setup 3, Bottom right: Setup 4. Source 1 (blue), Source 2 and 4 (orange), Source 3 (yellow)

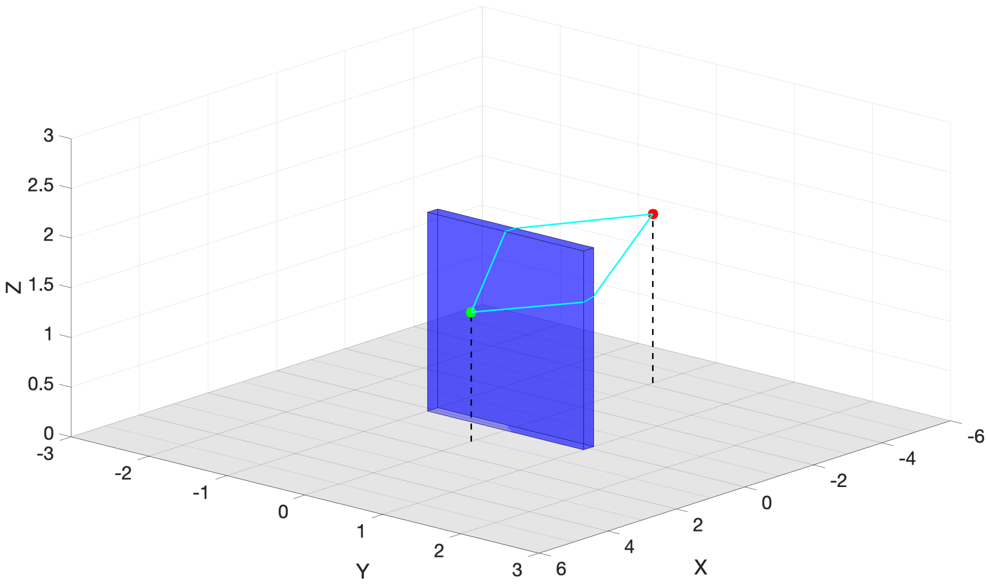

Figure 4.3: Exemplary illustration of different propagation paths that cause interference patterns in the frequency responses

to a level difference of $20 \log \left(\frac{1 \text{ m}}{d_{s2,S1}} \right) dB \approx -15.3$ dB. Only some small deviations are visible. The deviations compared to the frequency response of a perfect dirac impulse are caused by the diffraction and reflection components, which interfere with the direct sound and with each other. The effect of the superposition of sound waves is visible a lot more significantly in the frequency responses of the paths from source to receiver, which are blocked by the barrier. Here, no direct sound reaches the receiver but only diffracted sound. The frequency responses show repeating pattern of strong notches towards high frequencies - a comb filter like behaviour. The notches are caused by the previous stated superposition of sound waves. Here they are more prominent, since no predominant direct sound is present. The superposition occurs, because there are different paths of different length around the barriers edges. So the sound arrives in parts that are differently delayed to each other. If the delay between two waves corresponds to a phase shift of 180°, destructive interference occurs at corresponding frequencies. For instance, looking at Setup 2 and source 2 in figure 4.2, there are notches at 550 Hz, 1660 Hz, 2720 Hz, 3780 Hz, 4840 Hz and so forth. The notches are equidistant by about approximately 1060 Hz, which corresponds to a wavelength of 0.32 m. Looking at the path differences, it becomes clear that there is a matching difference of path lengths for the path along the top edge ($l_{top} = 5.5$ m) and the paths around the wall ($l_{left/right} = 5.85$ m) (cf. figure 4.3). The paths around the left and the right edge of the wall are same, since source and receiver are positioned central to the wall. The difference between the top path and the side paths is 0.319 m. For setups, where the source is not center in front of the wall, but is blocked by the wall concerning the direct sound path, a more complicated interference pattern arises.

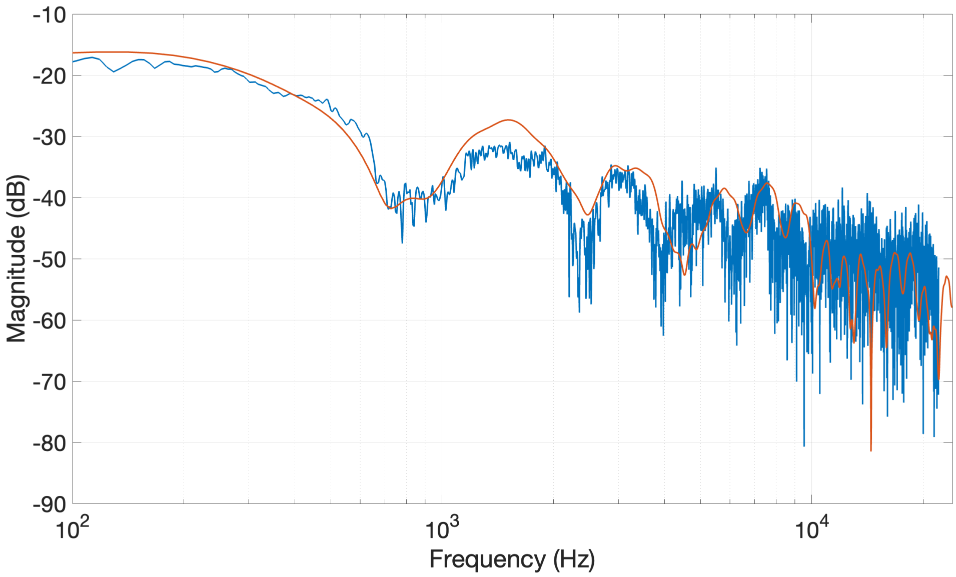

Figure 4.4: Measured frequency response (blue) and calculated frequency response (orange) of setup 1

4.1.2 Evaluation of the Noise Barrier Model

To test the performance of the secondary source diffraction model described in 2.5, the frequency response of a calculated impulse response from the model is compared with a measured frequency response. The evaluation is based on one scene taken from the database of the *Benchmark for Room Acoustics Simulation* (BRAS) project database [8]. More precisely, scene 2 with simple reflection and diffraction on, and around a finite square board is used and implemented in the model to calculate the impulse response for the path from source LS04 to receiver MP05. An detailed illustration of the scene is shown in figure A.1 the appendix as an excerpt from the BRAS project report. The corresponding frequency response is shown in figure 4.4 together with a calculated frequency response using the secondary source model, configured with the same geometrical arrangement. The calculation results in a overall well fitting frequency response, compared to the measured result. An average attenuation of 6 dB per octave can be seen for both graphs. Both follow the same resonant level peaks and drops with a slight shift in frequency. This is for example visible for the relative maximum at around 5000 Hz, at which the measured maximum is slightly lower in frequency than the calculated maximum. Deviations between the calculated and measured frequency responses are also affected by the non uniform frequency responses of the speaker and the microphone itself.

4.1.3 Simplification

As described before in section 4.1.1, the frequency responses of the noise barriers show a low-pass filter characteristic with some additional interference pattern. It already became clear that the interference pattern and the corresponding level notches, are strongly depending in the geometrical arrangement of source, receiver

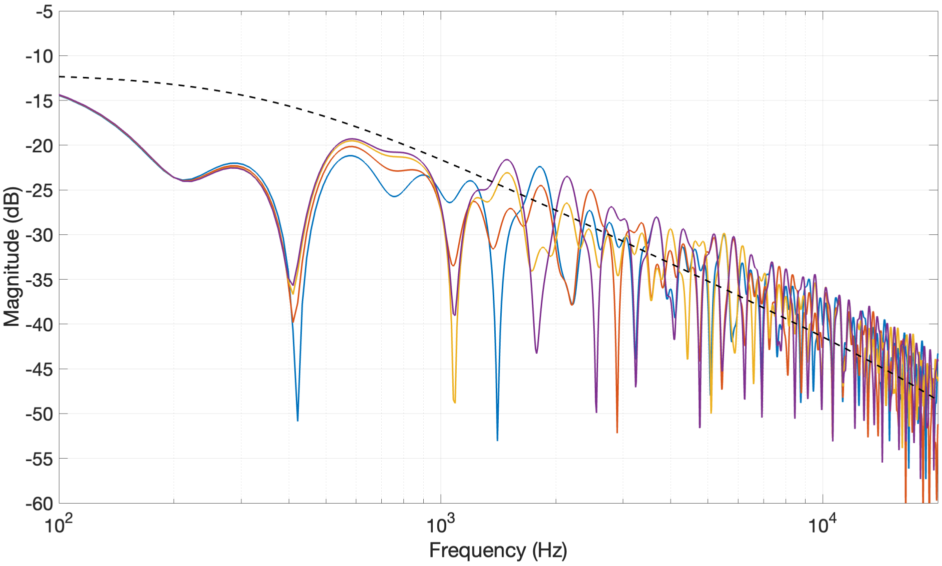

Figure 4.5: Calculated frequency responses: Noise barrier transfer paths with slight lateral displacement (0.1 m) of the source position (coloured). First order lowpass filter with $f_c = 350$ Hz and overall attenuation of -12 dB (dashed black)

and noise barrier. Picking another iteration, a close look is directed at the dependency of the interference on the geometrical arrangement in figure 4.5. Here the transfer functions are caluclated, using the coordinates of table 3.1 and 3.2.1 for Setup 3. Thereby, the Y-Coordinate was set to $\mathrm{Y} = [0,\ -0.1,\ -0.2,\ -0.3]$, to see how the interference pattern changes for small differences of the path lengths around the barrier. It can be seen that the frequency responses of the four setups differ a lot, when looking at the level notches due to interference in the range above 1 kHz. As already described in section 3.3, the directional frequency filtering is not very accurate regarding the directivity and extension of the sectors. Since one sector represents a certain angular range and only one frequency filter function can be applied to it's signal, it can't represent the positional differences. Thus the frequency response of the noise barriers is simplified by only considering their low-pass characteristic. As an approximation, a first order low-pass filter with a cut-off frequency of $f_c = 350$ Hz is shown in figure 4.5. Corresponding to the distance between source and receiver, for the calculations of the frequency responses and the following level attenuation, a overall level reduction of $20 \log\left(\frac{1}{4}\right) \approx -12$ dB is applied to the frequency response of the low-pass. The frequency response of the low-pass fits the shape of the calculated frequency responses very well and is consequently used as frequency filter function for the following investigations.

4.2 Measurements

The measurements are conducted as described in section 3.2. The atmospheric conditions in the anechoic chamber were stable at a temperature of 22.2 °C and a relative humidity of 45 %. The measurements of all four positions were made in the

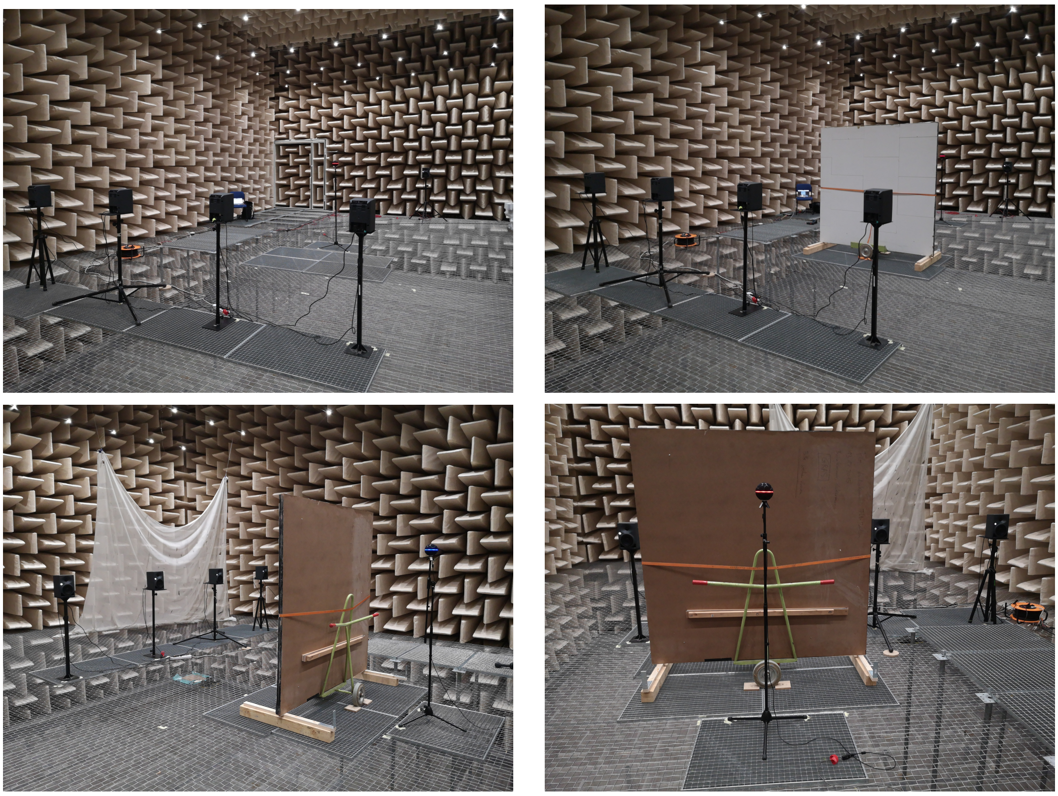

Figure 4.6: Pictures from the execution of measurements. Top: Setup 3 without (left) and with a wall (right). Bottom: Setup 3 receiver site (left) and Setup 1 receiver site (right)

course of one day, resulting in a total of 760 impulse responses:

	Microphone Capsules	Source Positions	Setups	Wall	Sum
Number of impulse responses	19	5	4	2	760

The positions of the sources and the receiver were rearranged between the measurements to achieve the four different geometrical setups. Here, in a first iteration all four setups were measured without wall and subsequently with the wall. To ensure identical source and receiver positions for both iterations, the distances were measured with a laser range finder. Additionally all positions were marked with tape. The reproducibility of the measurements is exemplarily shown in figure 4.7. Here the frequency response of source position 5 is measured in two iterations between which the source speaker and the receiver microphone array were removed and placed back at the same position. The shown frequency responses were measured with microphone capsule number 4 of the microphone array. The two frequency responses overall look similar. Only small deviation are visible which can be classified as measurement inaccuracies.
Some remarks have to be made concerning the measurement accuracy. As seen in the pictures of the execution of the measurements in figure 4.6, the wall was posi-

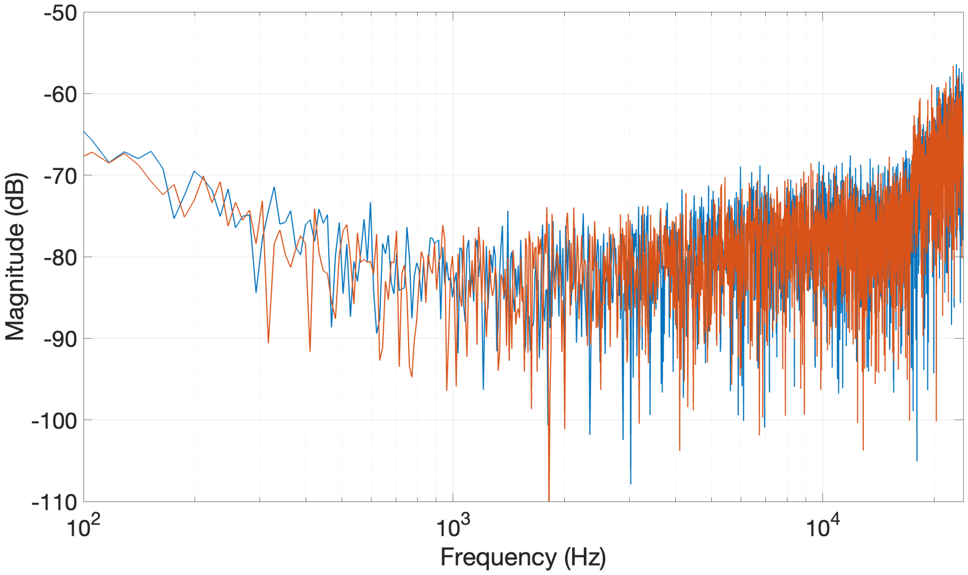

Figure 4.7: Measured frequency response of source 5 in two iterations

tioned in the room by means of a panel trolley and was stabilized by wooden beams, which can have a influence of the measurement accuracy. Their exact impact on the accuracy can not be determined and is neglected for the further investigation.
As a first step, the raw impulse response data is sorted and normalized relative to the highest absolute value of all measured impulse responses. Subsequently, it is truncated and windowed, using a Blackman-Harris window.

4.3 Auralization

In this chater, the results of the auralization, described in section 3.4, are shown using spectograms. Figure 4.8 and 4.9 display the spectograms of two moving sources are displayed. The signals were generated using a white noise source signal and a speech signal, respectively and the impulse responses measured for source, receiver and wall positions of setup 1. The binauralization was realized by using the sparta_ambiBIN binaural ambisonic decoder plugin [37]. The decoding settings were set to:

- Decoding Order: 3rd Order
- Method: Magnitude-LS
- Format: ACN, SN3D
- Apply maxRE weights: True
- Diffuse Conv. Constraint: False
- Apply truncation EQ: True

The plane wave decomposition was performed as explained in section 3.3, while using a lebedev grid to sample the sphere, which at $N_{sph} = 3$ returns $Q = 26$ sectors. After decomposing the signal, using the plane wave decomposition, sector signals $\xi = [0,\ 11,\ 14,\ 15,\ 19 \text{ and } 21]$ are filtered, as shown in the center of figure 3.7. Filtering took place using a first order high shelf filter with a cut-off frequency of

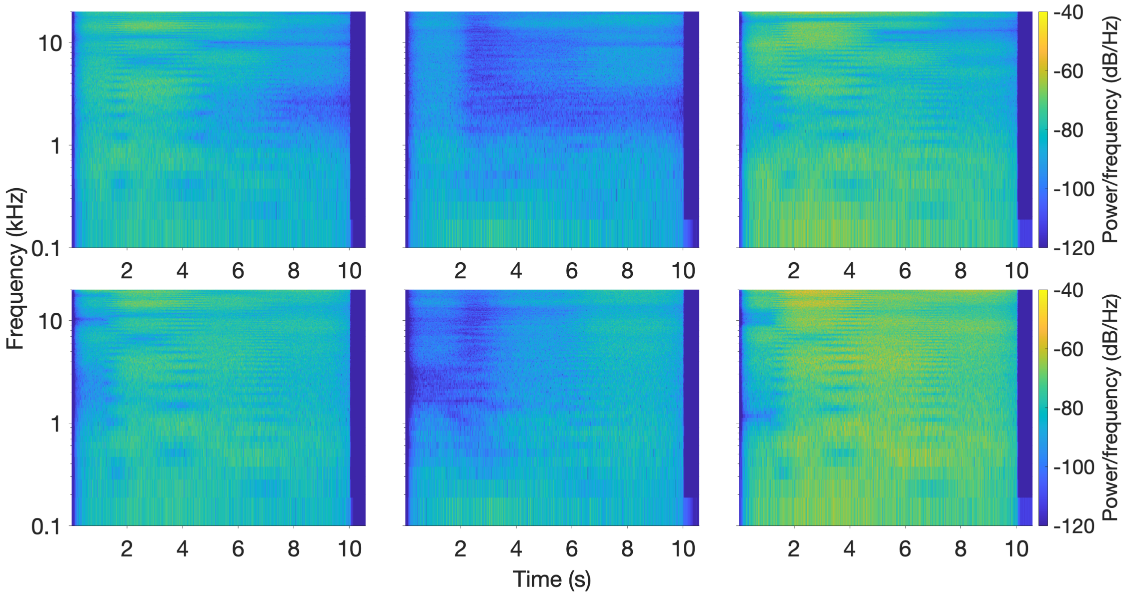

Figure 4.8: Spectogram of binaural scene auraliztaions (left channel: top row, right channel: bottom row) with a moving noise source: Empty scene (left), Real noise barrier (center), simulated noise barrier (right). Left channel (top) and Right channel (bottom)

$f_c = 950$ Hz. The filter is applied using the forward-backward filtering as described in section 3.3. In the two figures it can be seen, that the signal moves from left to right, since the spectograms show higher levels at the starting time on the left channel, whereby the level is lower on the right channel and vice versa at the end. Here the level differences due to the position of the source is particularly visible at mid frequencies. In the spectograms for the scene auralization of a real wall (center) it can also be seen, that the source is hidden behind the wall first at the beginning. Here the level above 500 Hz is attenuated significantly by the wall. The source then moves to the right, where it appears from behind the wall towards the end. Here the spectogram shows values, which are identical to the ones without a wall. Looking at the spectogram representing the auralization of the scene with a simulated wall and comparing it to the spectograms of the scene with a real wall, it can be seen that the directional frequency filtering does not return expected results. The overall level is increased. There is no filtering effect or level attenuation of high frequencies visible. The described visual results are also audible, whereby the auditory perception matches the described visible impression. Unfortunately, the signal processing and the described methodology didn't result in the expected objective and subjective outcome. The directional frequency filtering using the plane wave decomposition of a complex soundfield in the spherical harmonics domain could not sufficiently reproduce the soundfield of a real noise barrier in third spehrical harmonic order.

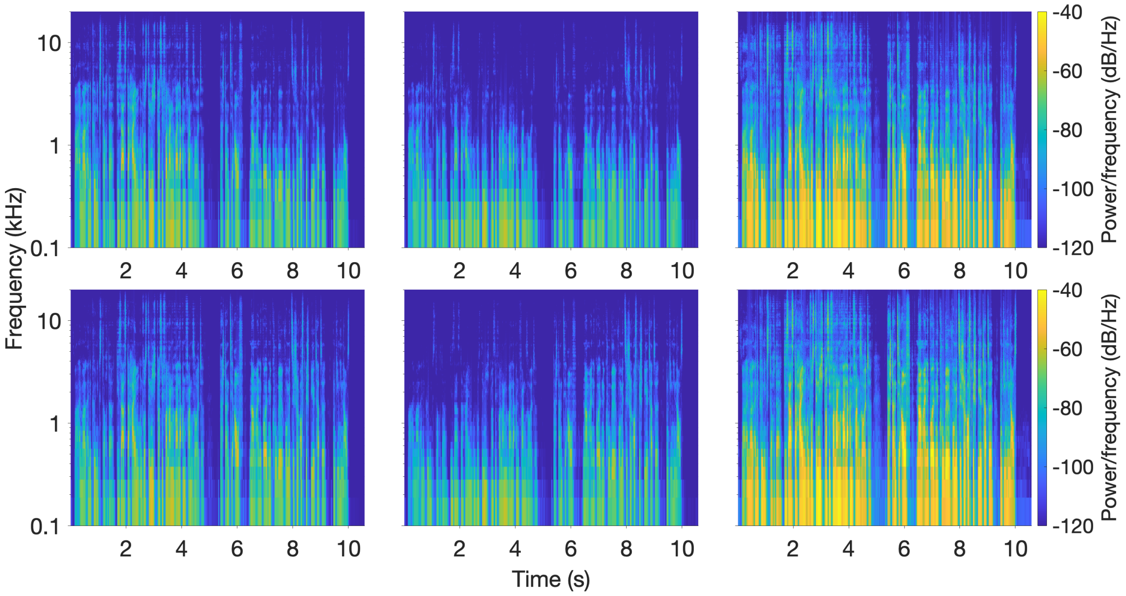

Figure 4.9: Spectogram of binaural scene auraliztaions with a moving speech source: Empty scene (left), Real noise barrier (center), simulated noise barrier (right)

4.4 Evaluation of Directional Filtering

Due to the unsatisfactory performance of the directional frequency filters and the auralization, listening tests were no longer considered for this investigation but the filtering within the plane wave decomposition is analysed more in detail in this section. To evaluate the quality of directional frequency filters, perfect synthetic signals are used before applying them to real recordings. Consequently a perfect dirac impulse incident as a single plane wave is used as initial signal in the first place. In the following the results and effects of the processing are shown. Figure 4.10 shows root-mean-square energy (RMS) distributions over the sphere for four different stages of processing from top left [a)] to bottom right [d)] in common English reading direction:

a) plane wave (before plane wave decomposition): perfect input signal
b) plane wave (after plane wave decomposition and re-composition): result of the perfect input signal decomposed in sectors and recomposed in the spherical harmonics domain without any further processing
c) plane wave filtered (before plane wave decomposition): frequency filtered perfect input signal
d) plane wave filtered (within plane wave decomposition): result of the perfect input signal decomposed in sectors, frequency filtering of desired sectors and recomposed to SH domain

As before, a first order high-shelving filter with $f_c = 950$ Hz and a gain of $G = -20$ dB is applied by forward-backward filter to the selected sectors. The signals' RMS plots are achieved by applying a plane wave decomposition to the signal. The plane wave decomposition is carried out as described in section 3.3. $\max_{rE}$ weights, displayed in figure 2.6, and a grid of 91 (azimuth ϕ) by 61 (elevation *theta*) of equally

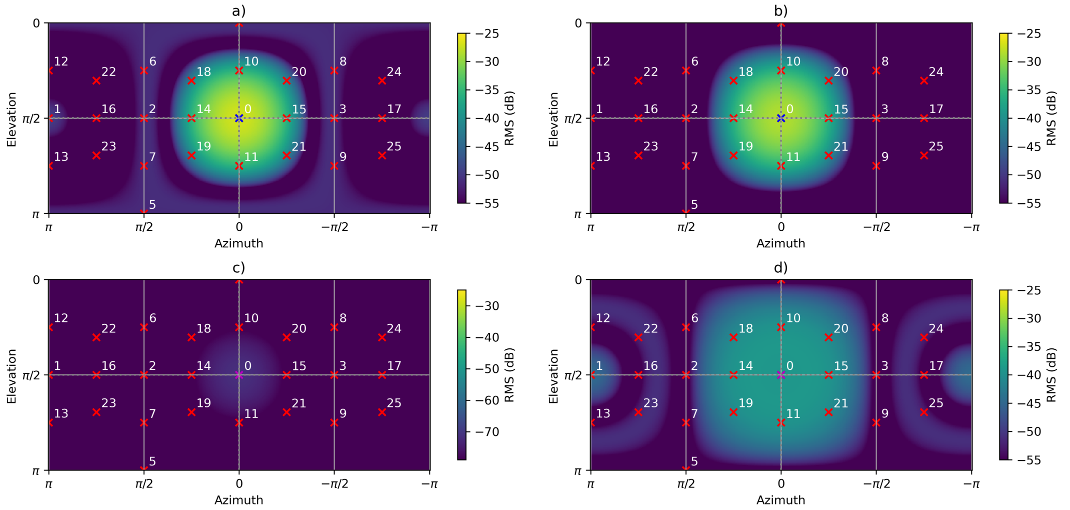

Figure 4.10: Signal RMS of different processing stages of a single plane wave. Sector directions of plane wave decomposition (red), incident direction of the plane wave (blue) and filtered sector (purple)

distributed sampling points are chosen for the decomposition. Subsequently, the RMS value of each resulting plane wave section is computed in dB and displayed as a function of both ϕ and θ. The RMS values correspond to the colorbar shown on the right of each diagram.

4.4.1 Single Direction Incident Plane Wave

First of all, to analyse the effects of the processing chain (plane wave decomposition, frequency filtering of a sector, re-composition to spherical harmonic domain) the most basic input signal is used. A single dirac impulse plane wave that arrives from the same direction as one of the sectors of the plane wave decomposition. Point 0 ($\phi = 0°, \theta = \frac{\pi}{2}$) is chosen arbitrarily.

In the figure of the initial plane wave [a)] in figure 4.10 it can be seen that the energy doesn't only distribute only at the incident direction (($\phi = 0°, \theta = \frac{\pi}{2}$)), but spreads out around neighbouring sectors because of the size of the main lobe of the plane wave beam. Additional to that, a symmetrical pattern of low energy distribution is visible in directions far from the incident direction. The occurrence of sound energy in those directions is caused by the side lobes of the plane wave beams, described in section 2.3.1. Looking at the de- and re-composed plane wave [b)] compared to the initial plane wave [a)], it can be seen that the energy distributions differ. The overall RMS is 2 dB lower for [b)] compared to [a)]. To emphasize the difference caused by the de- and re-composition, the dirac signals of the corresponding in and output signal in spherical harmonics domain is displayed in figure 4.11. First of all, it can be seen that the dirac is only represented by channels whose directivity is non-zero towards $\phi = 0°, \theta = \frac{\pi}{2}$ (cf. figure 2.2), i.e. channels $0, 3, 6, 8, 13, 15$. It can be seen that the amplitude of the signals of all channels, except for channel 0 which has an omnidirectional directivity, is reduced due to the plane wave decomposition. This is the reason for the slight difference in [a)] and [b)]. The omnidirectional

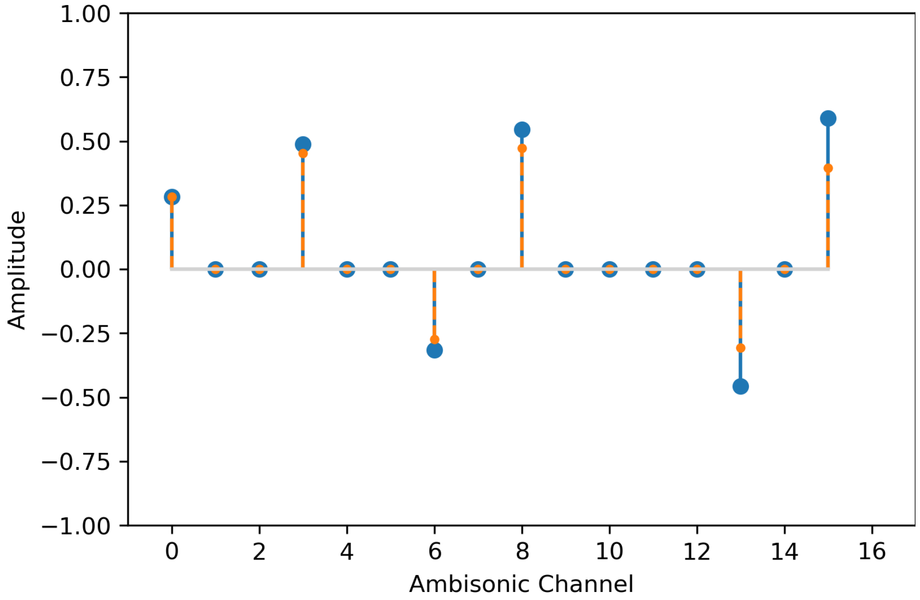

Figure 4.11: SH domain signal of the initial plane wave [a)] (blue) and the plane wave after de- and re-composition [b)] (orange)

energy is relatively increased for [b)], why the emphasis of the side-lobes and the corresponding distribution of energy around the direction of the initial plane wave is not as high as in [a)].
Comparing the frequency filtered plane wave (before plane wave decomposition) [c)] to the input plane wave [a)], it can be seen that the overall RMS is lower. Due to the filtering, high frequency components are attenuated and consequently the overall energy is reduced as expected by approximately 40 dB RMS. The filtered plane wave (before plane wave decomposition) [c)] could be seen a "perfect" filter result reference. However, looking at the filtered plane wave (within the plane wave decomposition) [d)] the resulting output appears to differ a lot from the desired output. Although the RMS at the direction of the filtered sector 0 is reduced, a lot of energy is distributed to areas where no signal energy was expected. The RMS of the filtered plane wave (within the plane wave decomposition) [d)] exceeds the perfectly filtered plane [c)] by more than 37 dB overall and 24 dB at the direction of sector 0.

4.4.2 Multiple Direction Incident Plane Wave

An initial signal composed of multiple plane waves incident from different directions is now used as a more complex signal. In figure 4.12 the signal contains three plane waves incident from $[\phi = 0, \theta = \frac{\pi}{2}]$, $[\phi = \frac{\pi}{2}, \theta = \frac{\pi}{4}]$ and $[\phi = -\frac{\pi}{2}, \theta = \frac{3\pi}{4}]$, which match sectors 0, 6 and 25, to achieve an asymmetric layout. Frequency filtering is applied to sectors 0 and 6.
As before, it can be seen that [a)] and [b)] show small deviations due to the de- and re-composition. The filtering of the input signal (before [c)] shows that the input signal is dominated by the unfiltered plane wave, arriving from $[\phi = -\frac{\pi}{2}, \theta = \frac{3\pi}{4}]$, as expected. The influence of the two filtered plane waves is clearly visible. The

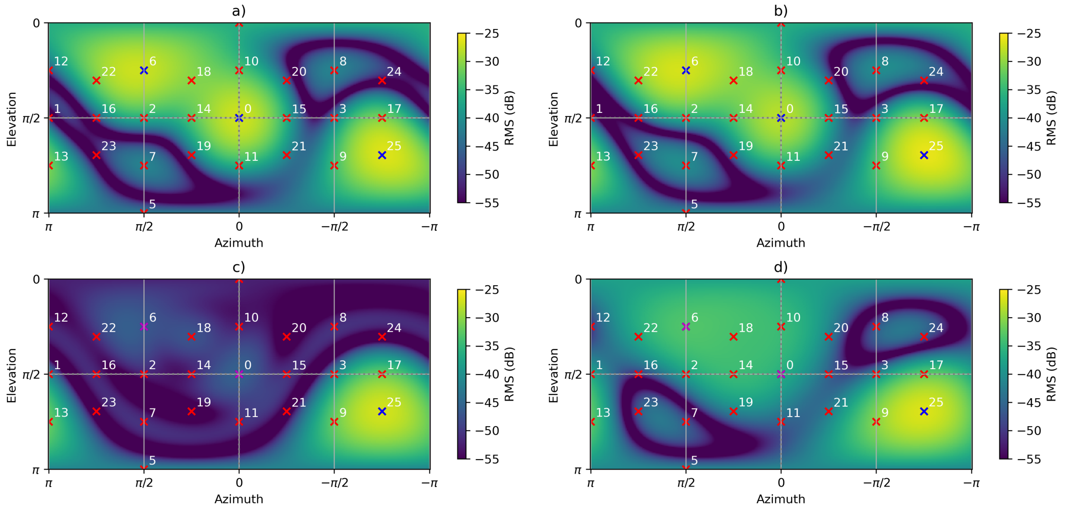

Figure 4.12: Signal RMS of different processing stages of a multiple plane wave. Sector directions of plane wave decomposition (red), incident direction of the plane wave (blue) and filtered sector (purple)

symmetric energy pattern is distorted due to the side lobes of the filtered plane waves around $[\phi = 0, \theta = \frac{\pi}{2}]$ and $[\phi = \frac{\pi}{2}, \theta = \frac{\pi}{4}]$, as describe before. Again, the result of the filtering within the plane wave decomposition [d)] differs from the "perfect" reference. The energy around the directions of the filtered beams is significantly higher. Compared to [b)] the RMS is only reduced by about 10 dB in those directions. Also the overall picture of the energy distribution is distorted, which is particularly evident in the areas around sector $\xi = [1, 8, 23, 24]$. Even though the direction of the incident plane wave is directed in such a way to match the direction of some of the sectors of the plane wave decomposition, also other sectors then the matching ones contain energy after the decomposition. The reason for that is the big main lobe width and the existing side-lobes of the beams of the plane wave decomposition. Even though it is not applicable for real scenarios, in the following all non matching sectors, which ideally should not contain any energy, are set to zero. Therefore it is possible to see a cleaner picture of the processing method if no distortions due to side-lobes would be existent. In figure 4.13 the initial filtered plane wave (before plane wave decomposition) and result of the filtering within the decomposition are displayed. It can be seen that the result looks a lot more like the "perfect" reference. The symmetric side-lobe energy distribution is more pronounced when filtering within the decomposition, which is why the filtered sectors can not stand out and are barely visible.

4.4.3 Non ideal incident plane wave signals

Next, the incident direction of the initial plane wave signals from figure 4.10 will be shifted. Thus, the directions of the incident plane waves will not be aligned with the directions of the plane wave sectors of the decomposition. First, the signal from section 4.4.1 is shifted by $\Delta\phi = \frac{\pi}{8}$ in positive azimuth direction and the corresponding figures are displayed in figure 4.14. The RMS energy distribution plots show the

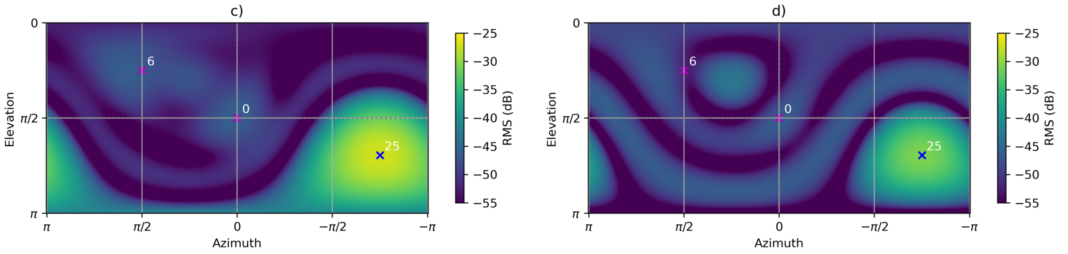

Figure 4.13: Signal RMS of different processing stages of multiple plane wave and non relevant sectors neglected: plane wave filtered before plane wave decomposition (left); plane wave filtered within plane wave decomposition (right). Sector directions of plane wave decomposition (red), incident direction of the plane wave (blue) and filtered sector (purple)

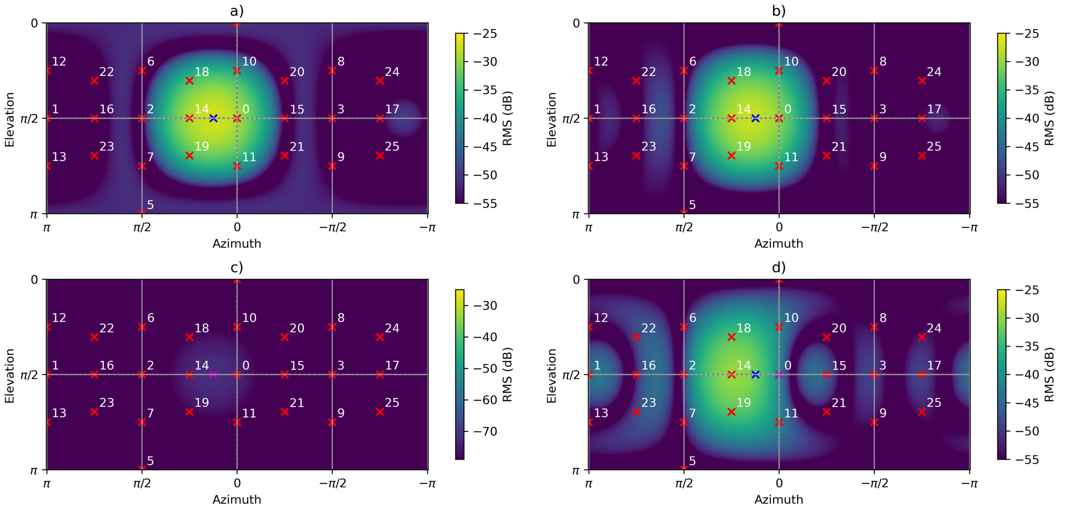

Figure 4.14: Signal RMS of different processing stages of a offset single plane wave

expected behaviour, especially for the signals before the plane wave decomposition. Here, the plot is just shifted as a whole compared to plot 4.10 by $\frac{\pi}{8}$ and is identical otherwise. Differences can be seen in the energy distribution of the de- and re-composed signals. While the energy distribution of [b)] only shows minor differences, the mismatching of the direction of incident plane wave and the direction of the plane wave decomposition sectors cause a distorted distribution of energy for the frequency filtering (within the plane wave decomposition) [d)]. Additionally to the distorted distribution, the side-lobe directions contain a significant amount of energy relative to the main lobe direction. Following the same procedure is applied to the input signal of section 4.4.2 and corresponding results are shown in figure 4.15. Similar as before, the initial signals [a) and c)] are identical as in figure 4.12 but shifted. The de and re-composed signal [b)] also matches the input signal with a slightly distorted energy distribution. Looking at the filtered signal (within the plane wave decomposition) [d)], it can be seen that the attenuation due to filtering of $\xi = [0, 6]$ is very low. Compared to [b)] the overall distribution is distorted even more.

Here it can be seen again that the proposed filter processing procedure can't produce

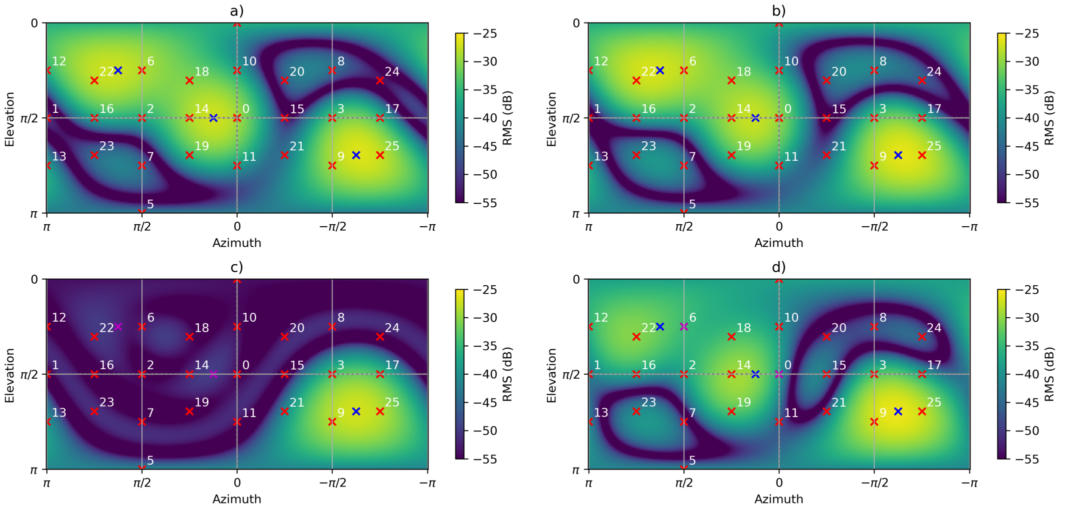

Figure 4.15: Signal RMS of different processing stages of a offset single plane wave

accurate results as expected. The big width of the main lobe and the side-lobes of the plane wave beams significantly contribute to a distortion of the energy distribution and inaccurate results. Furthermore, the discrete directions of the beams and the thereby low resolution of the sphere sampling amplify the named observations. Even though the signals that were used in this section, are perfectly generated artificial signals, the results could not satisfy the expectations. This especially is also the case for signals, which match the grid directions of the plane wave decomposition and are not as strongly affected by a insufficient grid resolution as non ideal and complex input sound fields.

4.4.4 Frequency Analysis of Sector Signals

To further analyse the properties of the processing and the changes introduced by filtering within the plane wave decomposition, frequency response plots are made at different stages of the processing. Here, the sector signals are of special interest, why they are studied after having the initial plane wave de-composed into sectors in a first step. In the following the filtered sectors within the plane wave are shown. After that the sectors are recomposed. To see how the sectors interact, they are decomposed in another iteration. In figure 4.16 the spectra of the sector signals of the single incident plane wave from section 4.4.1 are shown. Here not all sectors are shown but a selection, since some sector signals are identical due to the symmetry of the signal and the sampling grid. As expected the initial plane wave sectors have a flat frequency response which corresponds to the perfect dirac in time domain. Furthermore, sector 0, which points into the direction of the arriving plane wave, has the highest amplitude, followed by sector 10 and 18, which are the nearest neighbours to sector 0. The amplitudes match the recent observations and underline the influence of the inaccuracies introduced by the plane wave decomposition due to the width of the main and side-lobes of the beams. As intended only the signal of sector 0 is filtered and frequencies above 1 kHz are attenuated by 40 dB. The rest of the sectors are unaffected. The picture differs for the sector signals of the repeated

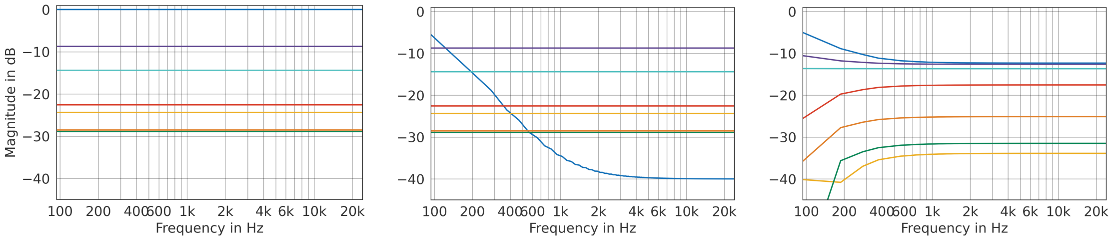

Figure 4.16: Frequency response plots of sector signals: Initial decomposed plane wave (left), Filtered decomposed plane wave (center), Filtered plane wave recomposed and decomposed again for analysis (right). Sectors: 0 (blue), 1 (orange), 2 (yellow), 10 (purple), 12 (green), 18 (turquoise), 22 (red)

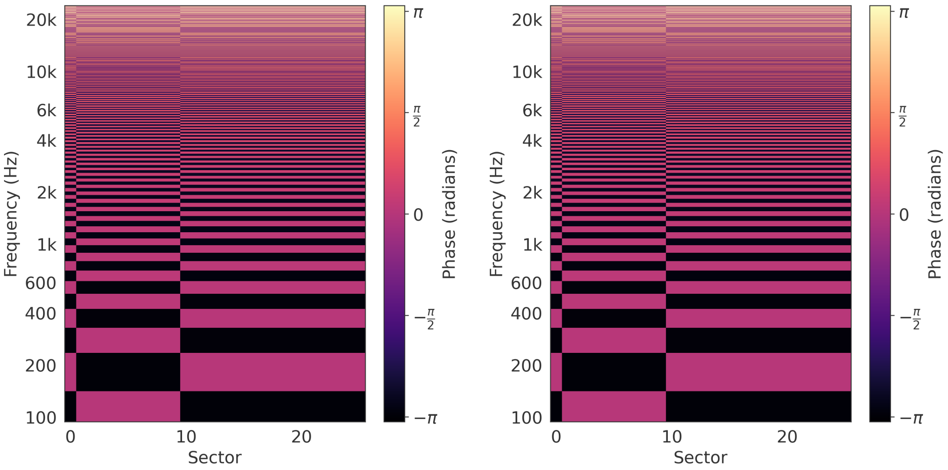

Figure 4.17: Phase responses of sector signals unfiltered (left) and filtered (right)

plane wave de-composition after the filter was applied and sectors were re-composed in a first iteration. Most importantly, it can be seen that the filtering of sector 0 cannot be reproduced as intended. The frequency response follows the shape of the actual desired attenuation for low frequencies, but ends up at a stop band with a level of -12 dB, which is much higher than expected. In addition to that, the overall levels of the other sectors are attenuated (sectors 2, 10 and 12) or amplified (sectors 1, 18 and 22) and are not constant along the frequency axis. A level attenuation can be observed for most of the sectors, whereas sector 10 also follows the filter response shape of sector 0 with increasing levels towards low frequencies.

The evaluation of the frequency responses has once again made clear that the overlapping of the sector signals results in signals influencing each other. The non-linearity of the system causes side effects concerning the filtering of the sector signals. Not only the frequency response of the selected sector is changed by the filter. Due to the composition of the sector signals also the frequency responses of the other sectors are affected by the filtering of one specific sector. Furthermore the phase responses of the sector signals of the filtered and unfiltered signal are shown in figure 4.17. Here it can be seen that the filtering does not have an influence on the phase of the signals.

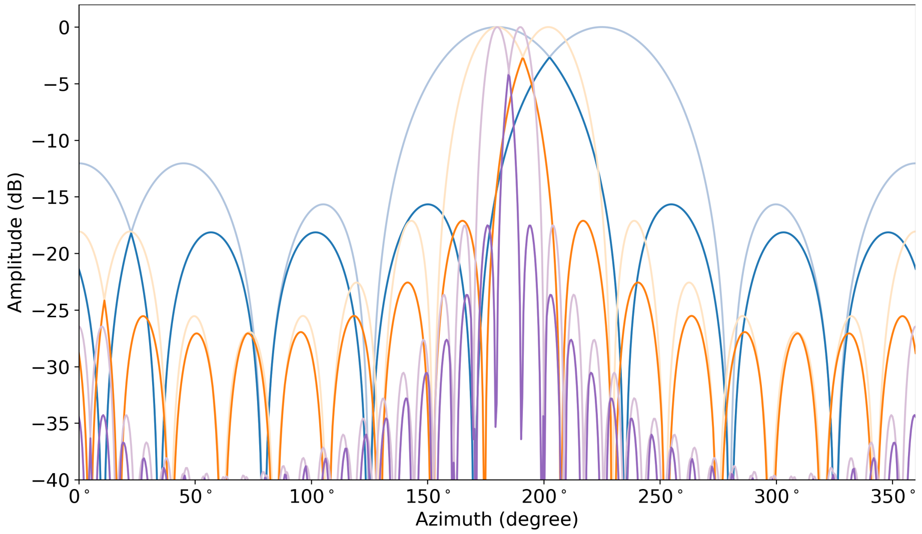

Figure 4.18: Directivity functions of neighbouring beams along the azimuth (light colored) and overlapping directivity parts (dark colored): $\max_{rE}$ weightings for $N_{\mathrm{sph}} = 3$ (blue, $\xi = [1, 17]$), $N_{\mathrm{sph}} = 7$ (orange, $\xi = [1, 69]$) and $N_{\mathrm{sph}} = 20$ (purple, $\xi = [1, 285]$

4.5 Improvements

Some limitations of the methodology and the technical equipment were described previously. In this section, those and additional limitations and their effects on the performance of the processing model and the auralization of noise barriers will be discussed more in detail. Possible improvements related to that are presented.

4.5.1 Spherical Harmonic Order

Since the microphone array available for this investigation is only able to record Ambisonics up to the third spherical harmonic order, the main part of this document also focused on this number of order. However, since the signal processing depends heavily on the overlap of the beams and the side lobes, as described in section 4.4, and these differ for higher orders, this section will discuss in more detail the possible improvements that could be achieved by using higher spherical harmonic orders. In figure 4.18, the directivity levels for neighbouring lobes and their overlap for different spherical harmonic orders are shown. Because the lobe width decreases and the level ratio between main- and side-lobes increases for higher orders, the angle ranges and the levels, where neighbouring sectors overlap decrease for higher orders. In combination with the increase of the grid sampling points, shown in figure 4.19, the sound field can be reproduced much more accurately without the influence of the overlapping main and side lobes. No measurements could be performed with such high orders due to the lack of corresponding technical measurement equipment. Therefore, improvements can only be shown based on synthetic soundfields. Some examples will be described using perfect dirac plane wave input signals, similar to section 4.4. Results are shown in figure 4.20, where the incident plane wave

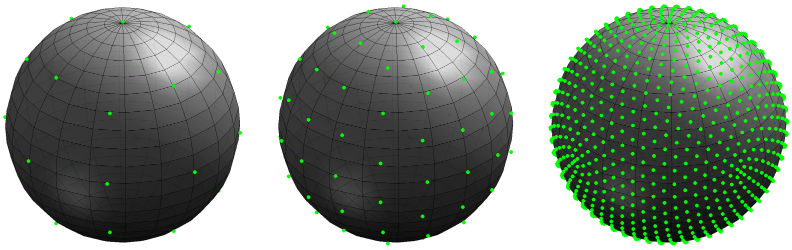

Figure 4.19: Sampling points of Lebedev sphere sampling: $N_{\text{sph}} = 3$; $Q = 26$ (left), $N_{\text{sph}} = 7$; $Q = 86$ (center) and $N_{\text{sph}} = 20$; $Q = 1202$ (right)

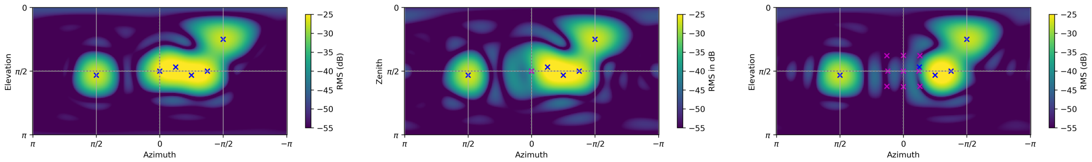

Figure 4.20: Signal RMS of different processing stages of a multiple plane wave with $N_{\text{sph}} = 7$: Unfiltered input signal after plane wave decomposition (left), filtered sector $\xi = 0$ (center), filtered sectors $\xi = 0$ and neighbouring (right); both within the plane wave decomposition

directions do not match the sector directions. Compared to the energy plots of the signals created with $N_{\text{sph}} = 3$ in section 4.4.2 it can be seen that a much more accurate picture of the initial plane waves can be reproduced, also when frequency filtering is applied and if the filtered sectors don't match the directions of the incident plane waves. The differences of the energy distribution due to distortion caused by the filtering are smaller compared to the results using a lower spherical harmonic order. Although it can be seen, that the level reduction at the direction of the filtered plane waves is not as high as expected, compared to "perfect" filtering when only one sector is filtered. Due to the higher resolution the sectors are closer to each other and the filtering can be improved by filtering adjacent sectors as well. Using an even higher spherical harmonic order, a line source of multiple incident plane waves is generated and shown in 4.21 exemplarily. Here the sectors that are filtered are chosen to intercept the line source, like a short sound barrier would do. It can be seen that the filtering works very accurate and sensitive, through which the line source can be interrupted very well. Here, the potential of the methodology using a higher order spherical harmonic processing chain becomes apparent.

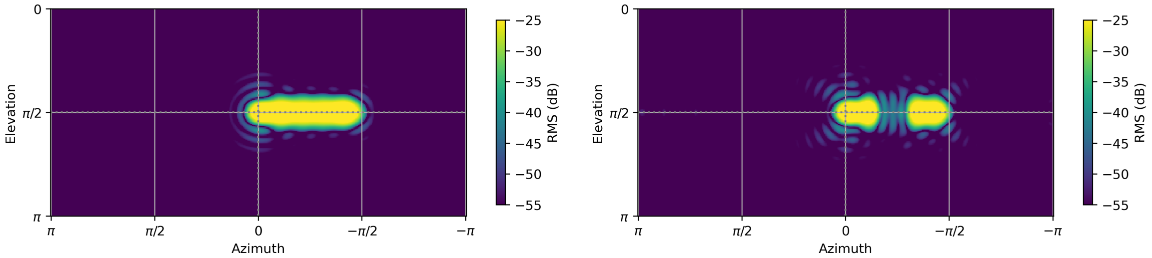

Figure 4.21: Signal RMS of multiple plane wave with $N_{\text{sph}} = 20$: Unfiltered input signal after plane wave decomposition (left) and filtered signal after plane wave decomposition (right)

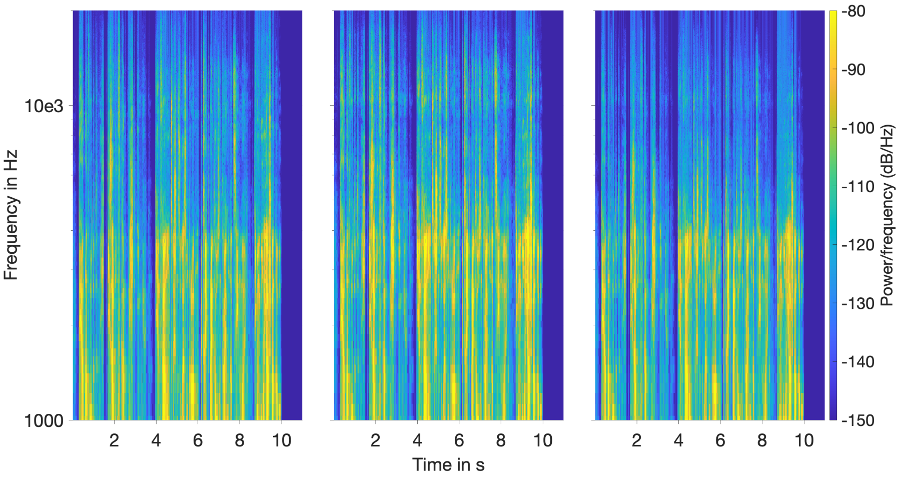

Figure 4.22: Spectograms of signals converted to mono: Input speech signal (left), processed signal $N_{\text{sph}} = 3$ without spectral compensation (center), with spectral compensation (right)

4.5.2 Spectral Compensation Filter

As described in 2.3.2 spatial aliasing artifacts arise, because the spatial sensitivity is low and the assignment of the directions of plane wave components is impossible due to the limited discrete sampling of the sphere, especially for high frequencies. Therefore, no reasonable plane wave decomposition is possible above f_{sA} [3]. Some approaches on how to reduce spatial aliasing artifacts. Looking at equation 2.11, the most obvious approach would be to either reduce the radius of the sampling sphere, thus the microphone array or to increase the spherical harmonic order and consequently the grid sampling points. In addition to the advantages of higher spherical harmonic order processing, mentioned in the previous section, this again emphasizes the improvements that can be achieved by increasing the spherical harmonic order. As said before, spatial aliasing cannot be avoided practically due to discrete sampling. Above all it is perceivable too [3, Sec. 5.6]. To compensate for the spectral aliasing artefacts, a spectral compensation filter in appearance of a first order low-pass filter with $f_{\text{c}} = f_{\text{sA}}$ should be used. The resulting spectograms of a quick application of such a spectral compensation filter can be seen in figure 4.22. It shows that the compensation filter works well at frequencies above f_{sA}. At around 4-5 kHz the level is increased for the processed but not compensated signal compared to the input signal. The compensation filter equalizes the level amplification properly. For higher frequencies a slight over compensation is visible. Filters should be discussed in further detail and should be incorporated into the signal processing chain.

4.5.3 Implementation of the Noise Barrier Geometry

For the investigations of this thesis, only approximated filters were used to represent the transfer functions of the sound barriers. Since the processing chain using $N_{\mathrm{sph}} = 3$ spherical harmonic order could not produce sufficiently accurate results, the quality of the approximation of the noise barriers, described in section 4.1.3, was not evaluated more in detail. The evaluation of the approximation could be part of listening tests, if higher order measurements are available. Part of this could also be the implementation of an algorithm using the ED-toolbox, that calculates multiple impulse responses of transfer paths, which represent the paths from discrete points along a line source to the receiver position. In contrast, in this thesis only the impulse responses of a single source position to the receiver were calculated.

4.6 Audio Files

Audio files of the different previously described auralization scenarios can be found at the following link:
https://zenodo.org/record/6980854#.YvPtF-xBx44
The audio files contain binaural for the direct playback via headphones. Furthermore the underlying ambisonic files are provided to enable playback considering head rotation with the use of corresponding plugins and further processing.
The filename consists of the spherical harmonic order, the setup index, the signal type, which type of noise barrier it auralizes (**no** barrier, measurement of a **real** barrier or a **simulation** of a barrier) and the format. Following the audio file of an auralization of a speech source moving along the line source of setup 1 in presence of a real wall in binaural format would be named as:
SH3_Setup1_Speech_real_binaural.wav
All scenes with all three types of barrier are auralized using a moving source once with a noise signal and one a speech signal for $N_{\mathrm{sph}} = 3$. Therefore the measured and calculated impulse responses are used as described in section 3.4. The files created with $N_{\mathrm{sph}} = 7$ (cf. section 4.5.1 are only available with a simulated barrier and without a barrier, because no real measurements could be made within this spherical harmonic order. A comparable moving source is simulated using the AmbiENC plugin [37]. The speech signal is taken from Dieter Leckschats audio file database of human voice stimuli for use in research and development in virtual acoustics [33].

5

Conclusions

The aim of this thesis was to implement a directional frequency filter processing chain for complex soundfields in spherical harmonics domain, to simulate the presence of a virtual noise barrier in the recording environment of the soundfields. Therefore, a model was created which is based on the decomposition of the soundfield into directed plane wave sectors, which represent sound components arriving from certain corresponding directions. Through this plane wave decomposition, access was granted to distinct directional parts of the soundfield, which subsequently can be processed individually. The processing here consists of frequency filtering selected sectors. The filtering should mimic the attenuating effects of a noise barrier, which would be positioned in the direction of the sector. Therefore the transfer function of noise barriers, which interrupt the direct path between a main source of interest and the receiving position, had to be calculated. A model that calculates impulse responses using direct sound, reflected sound, and second-order diffraction was used to calculate the transfer functions of the sound barriers.

To evaluate the quality of the application of the directional frequency filter processing, measurements were made. For this purpose, measurements of spatial room impulse responses were made in an anechoic chamber. This involved placing four loudspeakers in a line in front of and a fifth loudspeaker behind a Zylia ZM-1-3E microphone array, which is capable of third order spherical harmonics recordings. The loudspeakers serve as sources for the impulse response measurements, and they can later be used to simulate a moving source, using vector based amplitude panning or as a background noise source respectively. The measurements were made in two runs, each run consisting of measurements of the impulse responses in four different geometric arrangements of the source and receiver positions. For one of the runs, a sound barrier was placed between the line source and the microphone array. The processing model could be applied to the measurements without sound barrier and subsequently the results could be directly compared with measurements of exact the same acoustic scene with a real sound barrier in objective analysis and subjective listening tests.

The calculation results of the transfer function of the noise barriers showed to produce accurate results, compared to real measurements. To simplify the transfer functions, which are very sensitive to small position changes, that don't fit the spatial resolution of the spatial processing model, they're are approximated by low-pass filter functions. Furthermore it could be shown, that the processing chain using third order spherical harmonics is not suitable for the intended applications. The objective analysis has shown that signal processing per se works well. However, several factors directly related to the low spherical harmonic order cause the results to have little subjective agreement compared to real sound barriers. For this reason no subjective listening tests were conducted, as initially intended. Instead, the focus was put onto the analysis of the behaviour of the processing at different stages of the processing and the possibilities of higher spherical harmonic order processing. The following factors were found to be responsible for the unsatisfying results:

Spatial resolution
Due to the low spherical harmonic order, the sphere could only be sampled by few discrete points, which leads to a low spatial resolution. The sectors of the plane wave decomposition following need to cover a large directional area.

Lobe overlap
The overlap of the beam lobes of neighbouring sectors relative to their lobe width is constant over different spherical harmonic order. However, the absolute overlap increases with decreasing order. Therefore additional to the low spatial resolution, the directional area of neighbouring lobes interacting with each other increases, which especially leads to errors in the re-composed soundfield, if the neighbouring sectors are treated different frequency wise.

Side-lobe directivity
The side-lobe directivity levels increase for lower spherical harmonic order, which leads to higher errors within the plane wave decomposition.

Spatial Aliasing
For lower spherical harmonic order, the spatial aliasing frequency, above which a correct mapping of the sound field is not possible and spatial artifacts distort the results, is lower. Following the artifacts are distributed over a bigger frequency range with higher levels, compared to a signal processed with higher spherical harmonic order.

As an outlook, it is proposed to continue the methodology studied and discussed in this thesis and implement it with higher spherical harmonic order. In the process, the technical conditions could be improved and recordings could be made and processed that satisfy higher spherical harmonic order. This is a prerequisite if real acoustic noise scenes are to be evaluated in virtual environments in the future. Furthermore, acoustic scenarios with higher spherical harmonic order could also be synthesized and processed. Finally, if no recordings are available, the processing method could be performed in listening experiments without comparative reference. In particular, the following characteristics could be addressed: Quality and authenticity of the sound barrier simulation, application of spectral compensation filters, differences due to the approximation of the sound barrier transfer functions and differences due to the simplification of the geometric arrangement for the calculation of the transfer functions.

Bibliography

[1] Aletta, F., Van Renterghem, T., and Botteldooren, D. Influence of personal factors on sound perception and overall experience in urban green areas. a case study of a cycling path highly exposed to road traffic noise. *International Journal of Environmental Research and Public Health 15*, 6 (2018).

[2] Basner, M., and McGuire, S. WHO environmental noise guidelines for the European region: A systematic review on environmental noise and effects on sleep. *International Journal of Environmental Research and Public Health 15*, 3 (2018).

[3] Bernschütz, B. *Microphone arrays and sound field decomposition for dynamic binaural recording.* Doctoral thesis, Technische Universität Berlin, Berlin, 2016.

[4] Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization.* The MIT Press, 10 1996.

[5] Blauert, J., and Braasch, J. *The Technology of Binaural Understanding.* 01 2020.

[6] blumlein, a. d. British patent specification 394,325 (improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems). *Journal of the Audio Engineering Society 6*, 2 (april 1958), 91–98, 130.

[7] Boring, E. *Sensation and Perception in the History of Experimental Psychology.* Century psychology series. D. Appleton-Century Company, Incorporated, 1942.

[8] Brinkmann, F., Aspöck, L., Ackermann, D., Opdam, R., Vorländer, M., and Weinzierl, S. A benchmark for room acoustical simulation. concept and database. *Applied Acoustics 176* (2021), 107867.

[9] Clark, C., and Paunovic, K. WHO environmental noise guidelines for the european region: A systematic review on environmental noise and quality of life, wellbeing and mental health. *International Journal of Environmental Research and Public Health 15*, 11 (2018).

[10] Council of European Union. Council regulation (EU) no 49/2002, 2002. `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32002L0049`.

[11] Craven, P., and Gerzon, M. Coincident microphone simulation covering three dimensional space and yielding various directional outputs. *U.S. Patent, no. 4,042,779* (1977).

[12] DIN 18005-1:2002-07, Noise abatement in town planning - part 1: Fundamentals and directions for planning, July 2002.

[13] Ding, L., Renterghem, T., and Botteldooren, D. Estimating the effect of semi-transparent low-height road traffic noise barriers with ultra weak variational formulation. *Acta Acustica United with Acustica 97* (01 2011), 391–402.

[14] Dzhambov, A. M., and Dimitrova, D. D. Exposures to road traffic, noise, and air pollution as risk factors for type 2 diabetes: A feasibility study in bulgaria. *Noise & health 18*, 82 (May-Jun 2016), 133–142.

[15] Echevarria Sanchez, G. M., Van Renterghem, T., Sun, K., De Coensel, B., and Botteldooren, D. Using virtual reality for assessing the role of noise in the audio-visual design of an urban public space. *Landscape and Urban Planning 167* (2017), 98–107.

[16] European Environment Agency. Environmental noise in europe — 2020, 2020-03.

[17] Felcyn, J. The influence of a signal's time structure on the perceived noise annoyance of road traffic noise. *Journal of Environmental Health Science and Engineering 19* (04 2021).

[18] Fellgett, P. Ambisonic reproduction of directionality in surround-sound systems. *Nature 252* (1974), 534–538.

[19] Fliege, J., and Maier, U. The distribution of points on the sphere and corresponding cubature formulae. *IMA Journal of Numerical Analysis 19*, 2 (04 1999), 317–334.

[20] Forssén, J., Kaczmarek, T., Alvarsson, J., Lundén, P., and Nilsson, M. Auralization of traffic noise within the LISTEN project - preliminary results for passenger car pass-by.

[21] Fritschi, L., Brown, A. L., Kim, R., Schwela, D., and Kephalopoulos, S. *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*, first edition ed. WHO Regional Office for Europe, Copenhagen (Denmark), 2011.

[22] Gilani, T. A., and Mir, M. S. A study on the assessment of traffic noise induced annoyance and awareness levels about the potential health effects among residents living around a noise-sensitive area. *Environmental science and pollution research international 28*, 44 (11 2021), 63045–63064.

[23] Gille, L.-A., Marquis-Favre, C., and Morel, J. Testing of the european union exposure-response relationships and annoyance equivalents model for annoyance due to transportation noises: The need of revised exposure-response relationships and annoyance equivalents model. *Environment International 94* (2016), 83–94.

[24] Hardin, R. H., and Sloane, N. J. Mclaren's improved snub cube and other new spherical designs in three dimensions. *Discrete Comput. Geom. 15*, 4 (apr 1996), 429–441.

[25] Heutschi, K., Pieren, R., Müller, M., Manyoky, M., Wissen Hayek, U., and Eggenschwiler, K. Auralization of wind turbine noise: Propagation filtering and vegetation noise synthesis. *Acta Acustica united with Acustica 100* (01 2014), 13.

[26] Hold, C., Politis, A., McCormack, L., and Pulkki, V. Spatial filter bank design in the spherical harmonic domain.

[27] Method for the subjective assessment of intermediate quality level of audio systems. Standard, Radiocommunication Sector of International Telecommunication Union, Geneva, CH, Oct. 2010.

[28] Jafari, M., Khosrowabadi, R., Khodakarim, S., and Mohammadian, F. The effect of noise exposure on cognitive performance and brain activity patterns. *Open Access Macedonian Journal of Medical Sciences 7* (08 2019).

[29] Kleiner, M. *Acoustics and Audio Technology (3rd Edition).* J. Ross Publishing, Inc., 2012.

[30] Kurra, S. *Environmental Noise and Management: Overview from Past to Present.* Wiley Series in Acoustics Noise and Vibration. Wiley, 2020.

[31] Laxmi, V., Thakre, C., and Vijay, R. Evaluation of noise barriers based on geometries and materials: a review. *Environmental Science and Pollution Research 29* (2021), 1729 – 1745.

[32] Lebedev, V. I. Quadratures on a sphere. *USSR Computational Mathematics and Mathematical Physics 16*, 2 (1976), 10–24.

[33] Leckschat, D., and Epe, C. Aufnahmen von Sprecherinnen und Sprechern zur Verwendung in der Virtuellen Akustik, Feb. 2020.

[34] Lindau, A. Spatial audio quality inventory (saqi). test manual. Tech. rep., Technische Universität Berlin, 2014.

[35] Maekawa, Z. Noise reduction by screens. *Applied Acoustics 1*, 3 (1968), 157–173.

[36] Maillard, J., Kacem, A., Martin, N., and Faure, B. Physically-based auralization of railway rolling noise.

[37] McCormack, L., Hold, C., and Politis. Sparta suite. `https://github.com/leomccormack/SPARTA/releases/tag/v1.5.2`, Nov. 2021.

[38] Moehler, U., Huth, C., and Liepert, M. Case studies on the application of psychoacoustic methods for traffic noise. In *Forum Acusticum* (Lyon, France, Dec. 2020), pp. 2421–2424.

[39] Nyquist, H. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers 47*, 2 (1928), 617–644.

[40] Oftedal, B., Krog, N. H., Pyko, A., Eriksson, C., Graff-Iversen, S., Haugen, M., Schwarze, P., Pershagen, G., and Aasvang, G. M. Road traffic noise and markers of obesity – a population-based study. *Environmental Research 138* (2015), 144–153.

[41] Oppenheim, A., and Schafer, R. *Discrete-Time Signal Processing: Pearson New International Edition PDF eBook.* Pearson Education, 2013.

[42] Paszkowski, W., Kotus, J., Poremski, T., and Kostek, B. Evaluation of sound quality features on environmental noise effects – a case study applied to road traffic noise. *Metrology and Measurement Systems vol. 25*, No 3 (2018), 517–531.

[43] Perrott, D. R., and Saberi, K. Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America 87*, 4 (1990), 1728–1731.

[44] Pieren, R., Bertsch, L., Lauper, D., and Schäffer, B. Improving future low-noise aircraft technologies using experimental perception-based evaluation of synthetic flyovers. *Science of The Total Environment 692* (07 2019).

[45] Pieren, R., Heutschi, K., Wunderli, J. M., Snellen, M., and Simons, D. G. Auralization of railway noise: Emission synthesis of rolling and impact noise. *Applied Acoustics 127* (2017), 34–45.

[46] Preisendörfer, P., Liebe, U., Bruderer Enzler, H., and Diekmann, A. Annoyance due to residential road traffic and aircraft noise: Empirical evidence from two european cities. *Environmental Research* (2021).

[47] pulkki, v. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society 45*, 6 (june 1997), 456–466.

[48] Rafaely, B. Plane-wave decomposition of the sound field on a sphere by spherical convolution. *The Journal of the Acoustical Society of America 116*, 4 (2004), 2149–2157.

[49] Rafaely, B. Analysis and design of spherical microphone arrays. *IEEE Transactions on Speech and Audio Processing 13*, 1 (2005), 135–143.

[50] Rafaely, B. Fundamentals of spherical array processing. *Springer Topics in Signal Processing* (2019).

[51] Rafaely, B., Peled, Y., Agmon, M., Khaykin, D., and Fisher, E. *Spherical Microphone Array Beamforming.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[52] Rayleigh, L. On our perception of the direction of a source of sound. *Proceedings of the Musical Association 2* (1875), 75–84.

[53] Roederer, J. G. *Superposition and Successions of Complex Tones and the Perception of Music.* Springer US, New York, NY, 1975, pp. 143–170.

[54] Schröder, D., Pohl, A., Svensson, U., Stephenson, U., and Vorlaender, M. On the accuracy of edge diffraction simulation methods in geometrical acoustics. vol. 7.

[55] The telephone at the paris opera. *Scientific American 45* (1881), 422–423.

[56] Seidler, A., Wagner, M., Schubert, M., Dröge, P., Römer, K., Pons-Kühnemann, J., Swart, E., Zeeb, H., and Hegewald, J. Aircraft, road and railway traffic noise as risk factors for heart failure and hypertensive heart disease—a case-control study based on secondary data. *International Journal of Hygiene and Environmental Health 219*, 8 (2016), 749–758.

[57] Stansfeld, S., and Clark, C. *Mental Health Effects of Noise*, vol. 1. 12 2011, pp. 683–689.

[58] Stansfeld, S. A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., Haines, M. M., Head, J., Hygge, S., Van Kamp, I., and Berry, B. F. Aircraft and road traffic noise and children's cognition and health: A cross-national study. *Lancet 365*, 9475 (2005), 1942–1949.

[59] Sun, K., Echevarria Sanchez, G. M., De Coensel, B., Van Renterghem, T., Talsma, D., and Botteldooren, D. Personal audiovisual aptitude influences the interaction between landscape and soundscape appraisal. *Frontiers in Psychology 9* (2018), 780.

[60] Svensson, P. Edge diffraction matlab toolbox. `https://github.com/upsvensson/Edge-diffraction-Matlab-toolbox`, 2021.

[61] Svensson, U. P., Fred, R. I., and Vanderkooy, J. An analytic secondary source model of edge diffraction impulse responses. *The Journal of the Acoustical Society of America 106*, 5 (1999), 2331–2344.

[62] Szalma, J. L., and Hancock, P. A. Noise effects on human performance: a meta-analytic synthesis. *Psychological bulletin 137*, 4 (July 2011), 682—707.

[63] Thomas, P., Wei, W., Van Renterghem, T., and Botteldooren, D. Measurement-based auralization methodology for the assessment of noise mitigation measures. *Journal of Sound and Vibration 379* (2016), 232–244.

[64] Xie, B., Dr. Ning Xiang, R., and Blauert, J. *Head-Related Transfer Function and Virtual Auditory Display: Second Edition.* A Title in J. Ross Publishing's Acoustics: Information and Communication Series. J. Ross Publishing, 2013.

[65] zotter, f., and frank, m. all-round ambisonic panning and decoding. *Journal of the Audio Engineering Society 60*, 10 (october 2012), 807–820.

[66] Zotter, F., and Frank, M. *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality.* 01 2019.

[67] Zwicker, E., and Hugo, F. *Psychoacoustics: Facts and Models*, 3rd edition ed. Springer-Verlag Berlin Heidelberg 1990, Heidelberg, 1999.

[68] Zylia Sp. z o. o. Zylia zm-1 microphone. `https://www.zylia.co/zylia-zm-1-microphone.html`, 2018.

A

Appendix 1

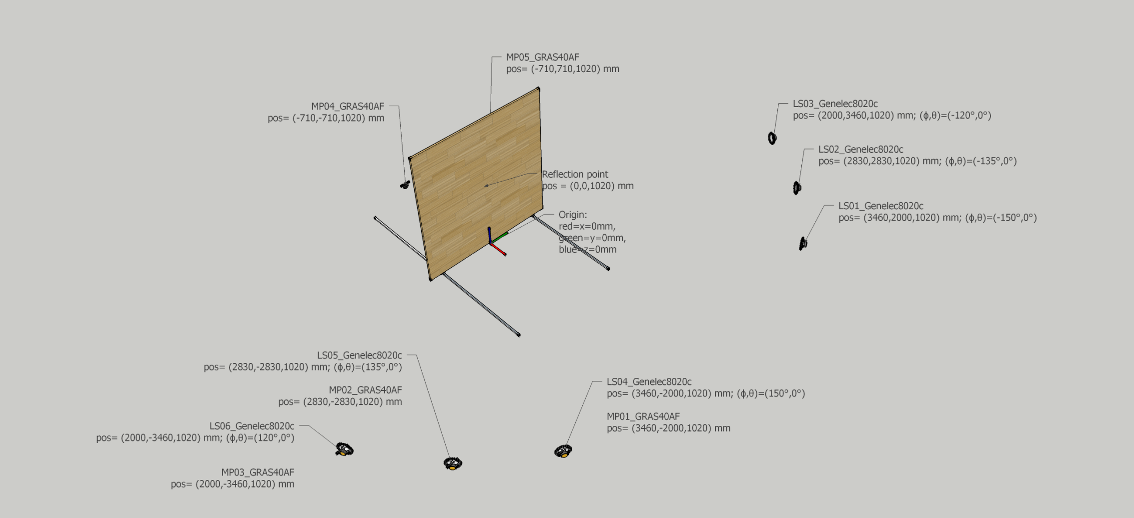

Figure A.1: Illustration of the measurement setup of Scene 2 of the BRAS model [8]

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY