A Soundgenerator for Wind- and Rolling Noises

Master Thesis by

Julian Koch, BSc

Matr.Nr. 01230639

Graz, September 21, 2021

Host Institution:
Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz

Graz University of Technology

In collaboration with Magna Steyr Fahrzeugtechnik

Assessor: Univ.Prof. Dipl.-Ing. Dr.techn. Alois Sontacchi
Supervisors: Univ.Prof. Dipl.-Ing. Dr.techn. Alois Sontacchi
Dipl.-Ing. Dr.Ing. Thorsten Bartosch

institut für elektronische musik und akustik

Statement of Originality

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

Julian Koch
September 21, 2021

Abstract

Nowadays, the majority of vehicle testing and development takes place virtually. Besides the visual and mechanic components, an authentic soundscape plays a vital role in appropriately assessing specific measures and situations. In this thesis, a real-time capable driving noise generator for wind- and rolling noise has been developed and merged with a previously developed engine noise generator to obtain a Simulink model that is able to synthesize the total driving noise inside a vehicle. The developed wind- and rolling noise generator can be parametrized by feeding the model with actual measurements. Spatial reproduction of the sounds is achieved by using Ambisonics, including the option to switch the playback system between a loudspeaker setup (as it can be found in driving simulators) and binaural headphones. The synthesized sounds have been validated by a listening test that concluded that there is no perceptible difference between the majority of the synthesized sounds and their real counterparts. Furthermore, the final driving noise generator has been assessed in the static driving simulator at the Virtual Vehicle Graz. A listening test was conducted to investigate the importance of different driving noise components (wind-, rolling- and engine noise) on the estimation of driving velocity. While this listening test did not prove the assumption that particular driving noise components improve the velocity estimation in different velocity regions, it still has shown that the velocity estimation itself is heavily influenced by the driving sounds and their respective loudness. Likewise, the test subjects were asked to subjectively evaluate the perceived realism of the driving sound module by filling a questionnaire after the listening test in the driving simulator. Most test subjects have positively evaluated the sound module, but the written feedback has shown that there is still room for improvement.

Kurzfassung

Der Großteil der Entwicklung und Testung von Fahrzeugen findet heutzutage virtuell statt. Neben der visuellen und mechanischen Komponente spielt ein authentisches Soundbild eine ausschlaggebende Rolle in der Beurteilung von speziellen Maßen und Situationen. In dieser Arbeit wurde ein echtzeit-fähiger Geräusch-Simulator für Wind- und Rollgeräusche entwickelt, welcher mit einem bereits vorhandenen Motorgeräusch-Simulator zu einem vollständigen Fahrgeräusch-Simulator kombiniert wurde. Der entwickelte Wind- und Rollgeräusch Simulator kann mittels realen Aufnahmen parametrisiert werden. Die räumliche Wiedergabe wird mittels Ambisonics erreicht, wobei die Möglichkeit besteht, zwischen einer Wiedergabe auf einer Lautsprecher-Anordnung (wie in Fahrzeug-Simulatoren üblich) und einer binauralen Wiedergabe auf Kopfhörern umzuschalten. Die synthetisierten Fahrgeräusche wurden mithilfe eines Hörversuchs validiert, welcher das Ergebnis lieferte, dass kein perzeptiver Unterschied zwischen den meisten synthetischen Fahrgeräuschen und den realen Gegenstücken besteht. Des Weiteren wurde der finale Fahrgeräusch-Simulator am statischen Fahrzeug-Simulator des Virtual Vehicle Graz beurteilt. Ein Hörversuch wurde durchgeführt, welcher die Wichtigkeit von verschiedensten Fahrgeräusch-Komponenten (Wind-, Roll- und Motorgeräusch) zur Einschätzung der gefahrenen Geschwindigkeit erforschen sollte. Obwohl der Hörversuch nicht statistisch nachweisen konnte, dass die Präsenz von bestimmten Fahrgeräusch-Komponenten die Geschwindigkeitswahrnehmung in verschiedenen Geschwindigkeitsbereichen verbessern kann, konnte beobachtet werden, dass die Geschwindigkeitswahrnehmung sehr wohl von den Fahrgeräusch-Komponenten und deren Lautheit beeinflusst wird. Abschließend wurden die Versuchspersonen gebeten, einen Fragebogen über die subjektiv wahrgenommene Realitätstreue der Fahrgeräusche auszufüllen. Die meisten Versuchspersonen haben den Fragebogen positiv evaluiert, jedoch zeigte das geschriebene Feedback, dass noch Raum zur Verbesserung besteht.

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Noise composition in a car 2

1.2.1 Wind Noise 3

1.2.2 Rolling Noise 5

1.3 Previous works 6

2 Wind Noise **8**

2.1 Analysis 8

2.1.1 Spectral Analysis 8

2.1.1.1 Velocity dependence 9

2.1.1.2 Different car models 13

2.1.1.3 Binaural differences 13

2.1.1.4 Yaw angle 16

2.1.2 Spatial Analysis 17

2.2 Synthesis 19

2.2.1 Spectral Shaping Methods 19

2.2.2 The Kautz Filter 20

2.2.2.1 Pole Detection - BU Method 22

2.2.2.2 Frequency Warping 23

2.2.3 Signal Generator 25

2.2.4 Turbulent flow conditions 26

2.3 Results 27

3 Rolling Noise **29**

3.1 Analysis 29

3.1.1 Spectrogram 30

3.1.2 Fan Chirp Transform 30

3.1.2.1 Acoustic Roller Dynamometer . . . 32
3.1.3 Vold-Kalman Filter . . . 33
3.1.4 Rolling Circumference . . . 36
3.1.5 Road Bumps . . . 37
3.2 Synthesis . . . 39
3.2.1 Noise Synthesis . . . 39
3.2.2 Tire Order Synthesis . . . 39
3.2.3 Road Bumps . . . 39

4 Playback - Ambisonics **41**
4.1 Encoding . . . 42
4.2 Decoding . . . 42
4.2.1 Sampling Decoder . . . 43
4.2.2 Mode Matching Decoder . . . 43

5 Implementation in Matlab/Simulink **44**
5.1 Matlab - Parametrization . . . 45
5.1.1 Parametrization of the wind noise . . . 45
5.1.2 Parametrization of the rolling noise . . . 49
5.2 Simulink Model . . . 50

6 Listening Test **55**
6.1 Listening Test 1 . . . 56
6.1.1 Calibration . . . 56
6.1.2 Phase 1 - Quality Evaluation . . . 57
6.1.2.1 Evaluation per Test Subject . . . 59
6.1.2.2 Evaluation per Sound . . . 62
6.1.2.3 Evaluation in total . . . 64
6.1.3 Phase 2 - Recognizing Velocity Differences . . . 64
6.1.3.1 Results . . . 67
6.2 Listening Test 2 . . . 68
6.2.1 Driving Simulator . . . 69
6.2.1.1 Connection . . . 69
6.2.1.2 Playback Setup . . . 69
6.2.1.3 Loudness calibration . . . 71
6.2.2 Velocity perception based on different stimuli . . . 71
6.2.2.1 Results . . . 72

6.2.3 Subjective Evaluation 78
6.2.4 Conclusion 79

7 Conclusion and Outlook **80**

Bibliography **82**

Chapter 1

Introduction

1.1 Motivation

Driving simulators begin to play an increasingly important role in the development cycle of a car. Instead of the original "trial and error" method, many development steps are nowadays executed virtually. This helps to save costs on the one hand and time on the other hand. Besides the visual and haptic components, auralization plays a crucial role when evaluating certain car measures, such as acceleration and speed. Furthermore, the driving sound also influences purely perceptive measures, such as driving comfort.

In the course of this master thesis, a sound generator for rolling- and wind-noises has been developed within Simulink/Matlab and integrated into an already existing Simulink model, containing an engine-sound simulation.

This driving-sound simulator will be utilized in car simulators that aim to test driving comfort. Therefore, while designing the sound generator, the focus was held on an authentic-sounding rolling-/wind-noise rather than on an exact reproduction based on car parameters and the physics of the sound sources.

This thesis contains seven chapters.
Chapter 1 gives a quick overview of this thesis and the general types of driving noise to be expected in a driving car, as well as an overview of the literature that has covered driving noise simulators and the commonly used methods to synthesize driving sounds.
Chapter 2 and 3 deal with the analysis/synthesis of wind noise and rolling noise, respectively.

Chapter 4 contains a short introduction into Ambisonics, which has been used as playback format during the listening test in the driving simulator.
Chapter 5 explains the usage and implementation of the previously presented algorithms in Matlab/Simulink and can be seen as a "manual" on how to use, parametrize and modify the driving noise simulator.
Chapter 6 describes the execution and the results of a listening test that was conducted at the Virtual Vehicle Graz to get an impression of the driving sound simulator and to validate the importance of driving sound for the general velocity impression.
Chapter 7 sums up the results of this thesis and gives an outlook for further improvements of the developed driving noise simulator.

As the sound generator will be worked on further by other parties in the future, the focus of this thesis also lies in a detailed explanation of the used analysis/synthesis techniques to lay the foundation for an easier improvement/development of the sound generator.

1.2 Noise composition in a car

This section contains a general overview of the driving noise components inside a moving car and their physical background.

The driving noise originates from a multitude of different moving parts of a vehicle and the interaction of the vehicle with its environment. Though many sources and effects contribute to the noise composition created by a driving car, it is basically dominated by three main components: the engine noise, the rolling noise, and the wind noise. While all those three components show a velocity dependence, their relative increase in volume is different for every noise component due to their different physical principle of origin. Therefore, certain components dominate the total noise composition in different velocity regions.

The velocity dependent SPL[1] of the different noise components inside a limousine can be described by (see [Zel12, p. 202]):

Engine noise SPL (with k and n_0 depending on the load):

$$L_{Eng,appr.} \approx L(n_0) + 10lg\left(\frac{n}{n_0}\right)^k \tag{1.1}$$

$$k \approx 3 \text{ for full engine load}$$

Rolling noise SPL:

$$L_{R,appr.} \approx L(v_0) + 10lg\left(\frac{v}{v_0}\right)^k \tag{1.2}$$

$$\text{with } k \approx 2$$

Wind noise SPL:

$$L_{W,appr.} \approx L(v_0) + 10lg\left(\frac{v}{v_0}\right)^k \tag{1.3}$$

$$\text{with } k \approx 6$$

n and v are describing the engine revolution and driving velocity, respectively, with n_0 and v_0 being reference values. While the SPL of the driving noise components at an arbitrary reference driving velocity/engine revolution $L(v_0)/L(n_0)$ is depending heavily on the car model, its construction, and its acoustic design, the SPL increase can generally be approximated quite accurately by the previous equations.

1.2.1 Wind Noise

The wind noise physically results from fluctuations of the airflow in the proximity of the vehicle, caused by edges and gaps of the vehicle shape and surface. Wind noise sources mainly have dipole character, which leads to $k \approx 6$ in equation 1.3 and subsequently to the level dominance in higher velocity regions.
Generally, wind noise is perceived as broadband-colored noise. Exceptions of this model arise under particular car shapes and building conditions. For example, an unfavorable car shape design can lead to howling and whistling sounds: cavities in the car might get excited by the airflow. A poorly designed side-mirror shape might

[1]Sound Pressure Level

also lead to sharp resonances and deterministic components. To obtain a perceptually pleasant wind noise, deterministic components should be avoided. Furthermore, a concrete localization of the wind noise can be identified as an error source and therefore is perceived as disturbing [Zel12, p. 205].

Factors that influence the sound and level of the wind noise include

- The car shape
- The car surface
- The sealing of doors, windows, etc.
- The presence of potentially resonating cavities

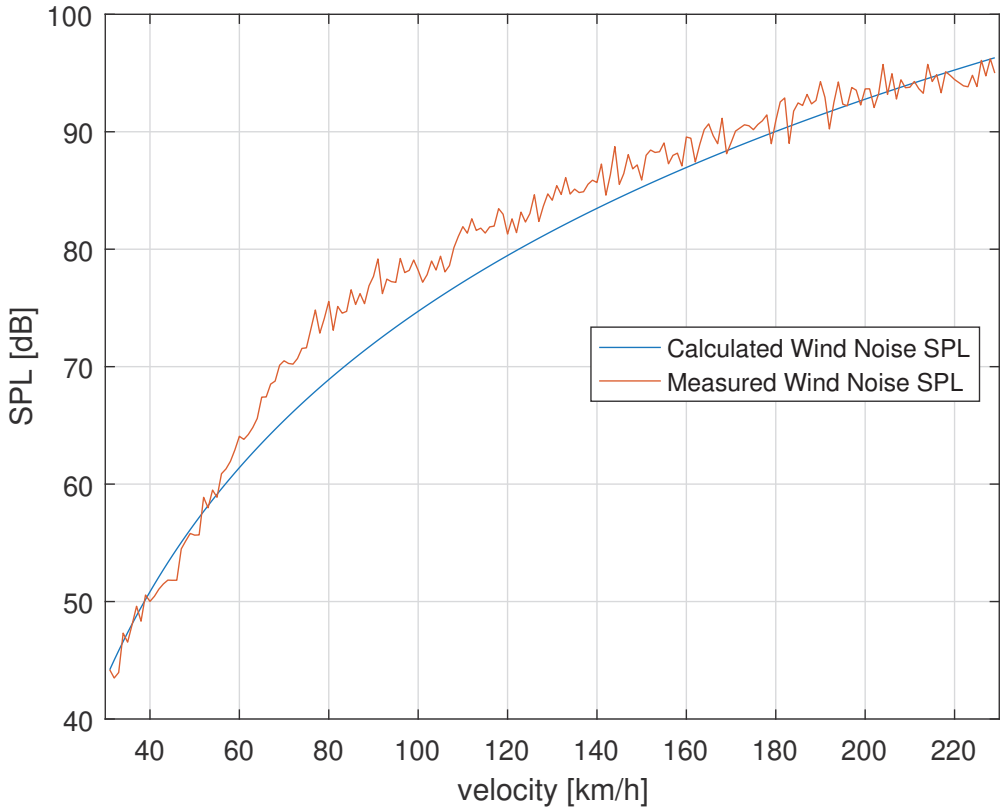


Figure 1.1: Comparison between the sound pressure level of a wind channel measurement and the theoretical calculation using equation 1.3

1.2.2 Rolling Noise

The emergence of the rolling noise is generally caused by two different phenomena [BKHK08]:

Time-varying contact forces on the tire lead to vibrations of the tire surface which subsequently get radiated both into the air as air-borne noise, and are transmitted via the front- and rear vehicle axle as structure-borne noise [Zel12, p. 203]. While the road surface with its random nature leads to a stochastic excitation, the rotation of the tire and the repetition of the tread pitches cause a periodic excitation of the tire, resulting in deterministic components. To avoid the radiation of single dominant frequencies, the size of the tread pitches is varied, and the pitches have an irregular shape/form. The variation in size of those pitches basically acts like a frequency modulation that spreads the radiated energy over a broader frequency range [IY96].

The second formation mechanism of rolling noise is called "air pumping". It is caused by air that is displaced near the contact region of the tire. Contrary to the tire vibrations, the rolling noise created by air pumping is solely air-borne and reaches the driver cabin mainly via the window- and door sealings.

Furthermore, the rolling noise gets amplified by different side effects:
The horn effect occurs due to the geometric shape that is formed by the tire and the road at their contact point. It supports very efficient sound radiation in a broad frequency range. Measurements with a loudspeaker buried in the road surface side had shown an increase of over 10 dB SPL for frequencies above 300 Hz when the tire was present, compared to measurements without the tire [IY96].
Several resonance phenomena such as the tire cavity modes amplify the rolling noise in specific frequency bands. Cavity modes typically lie in the frequency region between 180 Hz and 270 Hz. The exact frequency depends on several factors such as driving velocity and temperature of the tire [BS09]. Another resonance phenomenon may occur due to the tire groves. Compressed air in the tire groves that starts to oscillate basically behaves like an organ pipe, amplifying frequencies around 1 kHz [Zel12, p. 286].

1.3 Previous works

While the synthesis of car engine noise is a well-researched field (different methods to synthesize engine noise have been outlined e.g. in [Sta18]), works regarding the synthesis of car internal wind noise and rolling noise sounds are sparse.

There are two main reasons for this:

- The interest in engine sound synthesis is higher, due to its important role, especially in the field of engine noise reconstruction in electric cars
- Most of the exact used methods for driving sound synthesis for driving simulators are used commercially and therefore are kept confidential

Adams and van de Ponseele proposed a model for the synthesis of engine noise and structure-borne rolling noise [AVdP01] using an SQE[2] model. The structure-borne rolling noise has been modeled as a colored noise process. A third-octave filterbank has been used to filter white noise, and the resulting noise bands have been weighted and summed up accordingly. The velocity-dependent weights have been obtained for several discrete velocities by using a coherence based approach:
First, the wheel inputs that have been measured using tri-axial vibration sensors on the wheel hubs have been orthogonalized, using virtual coherence analysis (as the wheel inputs are always partially correlated). In the next step, the coherence between these orthogonal components and an acoustic receiver[3] has been calculated, and only the summed coherent auto-power was used for the spectral synthesis. The coherent autopower is defined as the auto-power of the acoustic receiver that is coherent with one of the orthogonalized wheel inputs. The summed coherent auto-power can be obtained by summing up the coherent auto-power regarding each orthogonalized wheel input. This way, only the structure-borne rolling noise is included in the synthesis.

Adams and van de Ponseele have also investigated the modeling of time structures in rolling noise synthesis in [AVdP01], as, due to the irregularities of road surfaces, the rolling noise often additionally shows an impulsive character. They came to the conclusion that their original idea of modeling those time structures as level variations could not provide a sufficiently realistic sound impression. A wavelet analysis of the

[2]Sound Quality Equivalent
[3]Adams and van de Ponseele used a binaural head, placed at the passenger seat

rolling noise has shown that every road impact appeared twice in the spectrum, caused by the excitation of the front wheels, followed by the excitation of the rear wheels. The time difference between those two impacts depends on the car velocity and the distance between the tires. Also, the spectrum of the impulse caused by the front wheels is different from the one caused by the rear wheels, resulting from different transfer paths from the wheel to the acoustic receiver for front- and rear wheels. Proposals on how to solve this problem have been mentioned but have not been explicitly solved in [AVdP01].

Sarrizin et al. proposed two models using higher resolution solutions than the standard SQE 1/3-octave filterbank to model the noise components of the driving sounds: using a 1/12-octave filterbank approach and the so-called narrow-band approach [SCJvdA13].

The first of the two is basically a modification of the filterbank method, used for SQE modeling (like in [AVdP01]). Instead of an 1/3-octave filterbank, a 1/12-octave filterbank is used to analyze and synthesize the driving noise.

The latter one works on a different principle. It relies on the fact that the phase information of rolling noise can be considered random [SCJvdA13]. Contrary to the SQE models, where the amplitude of each band is calculated, the narrow-band approach uses the complete amplitude information of the signal. During the synthesis stage, for every new input velocity, a so called "seed" signal is generated by applying the IDFT[4] to the amplitude information corresponding to the input velocity and randomly generated phase information. The seed signals are consecutively windowed with a half-sine-window, overlapped by 50 %, and summed up to obtain the output signal. If the velocity does not change for consecutive seed signals, it is sufficient to repeat the same seed signal with a random phase shift for each instant instead of repeatedly calculating new seed signals.

[4]Inverse Discrete Fourier Transform

Chapter 2

Wind Noise

2.1 Analysis

Multiple datasets of wind-channel recordings have been analyzed to find appropriate methods for the synthesis. Those datasets have not been recorded during this Master Thesis but have been provided by Magna Steyr.

This section presents the results of different analysis methods to classify wind noise.

2.1.1 Spectral Analysis

The spectral analysis is conducted by calculating the PSD[1]. The PSD $S_{xx}(\omega)$ is describing the power of a signal over its frequency. The PSD is defined as the Fourier transform of the autocorrelation function of a signal. In the digital domain, it can be described as the DFT of the autocorrelation sequence.

$$S_{xx}(\omega) = \sum_{m=-\infty}^{\infty} r_{xx}[m] e^{-j\omega m} \tag{2.1}$$

For a stationary and finite signal $x[n]$, the autocorrelation function r_{xx} is defined as

$$r_{xx}[m] = \sum_{n=-\infty}^{\infty} x[n]x[n+m] \tag{2.2}$$

[1]Power Spectral Density

The cross-power spectral density is defined as

$$S_{xy}(\omega) = \sum_{m=-\infty}^{\infty} r_{xy}[m] e^{-j\omega m} \tag{2.3}$$

where r_{xy} is the cross-correlation for finite and stationary signals

$$r_{xy}[m] = \sum_{n=-\infty}^{\infty} x[n] y[n+m] \tag{2.4}$$

The following sections show the PSD analyses of different wind noise-related data to investigate the impact of different conditions such as velocity, car shape, angular flow conditions on the wind noise spectrum. Also, the spectral differences between both ears for binaural recordings have been examined.

2.1.1.1 Velocity dependence

The data used to investigate the effect of different driving velocities on the wind noise spectrum originates from different wind channel run-ups. The dataset contains the measured SPL at different positions in the car, with the wind velocity being increased linearly from 40 km/h up to 230 km/h (some measurements cover the velocity range from 30-230 km/h). Those recordings have been cut into several time frames that span over a velocity range of 10 km/h to be able to obtain velocity-dependent spectral information from the dataset.
Analyzing these time frames results in a spectral average over the selected velocity region. After calculating the power spectral density, the resulting spectrum approximately represents the frequency information of the mean velocity in each frame.
From now on, every time the spectrum at a certain velocity is mentioned in this context, it is referring to the spectrum of the cut time frames as described above, with the mentioned velocity being the mean velocity of that frame.

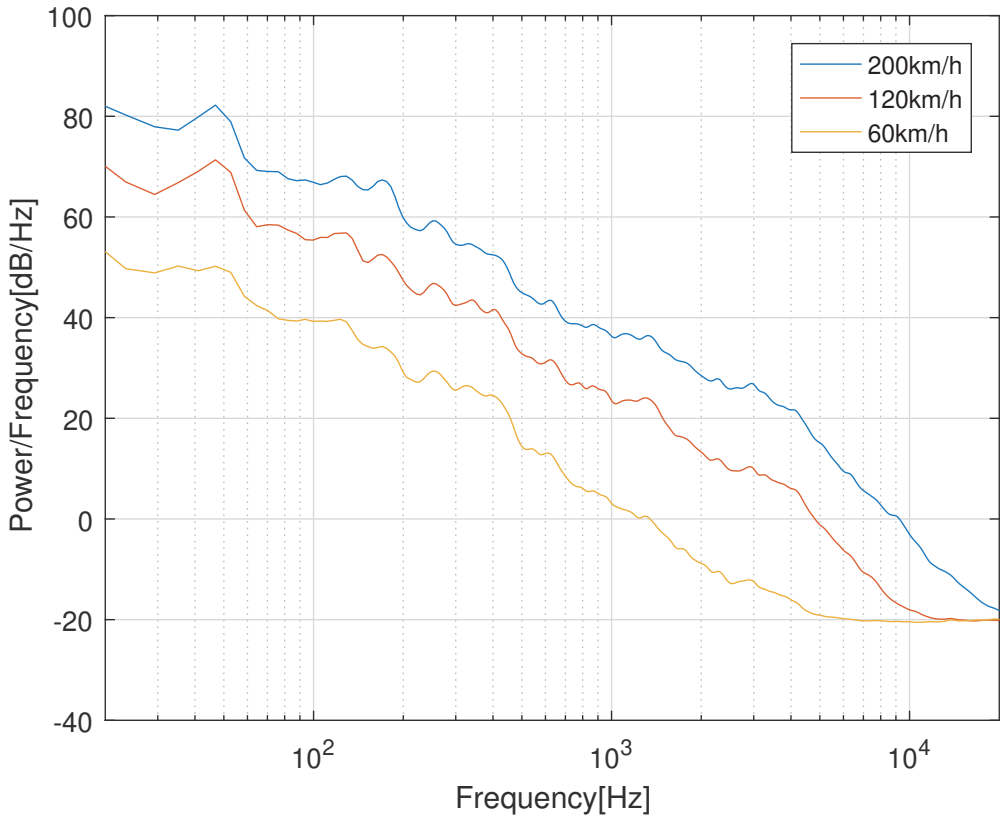


Figure 2.1: PSD of the wind noise for different velocity regions, cut out of a wind channel run-up, smoothed with a 1/12-octave filter

Figure 2.1 is showing the power spectral density of one car model for different velocities.

It can be observed that the overall level of the wind noise depends on the velocity. Additionally, the spectral coloration of the wind noise is changing for different velocities. Higher velocities emphasize the upper-frequency region. This is even easier visible in figure 2.2 and 2.3. Figure 2.2 is showing the PSD of different velocity regions referenced to the wind noise PSD at 150 km/h.

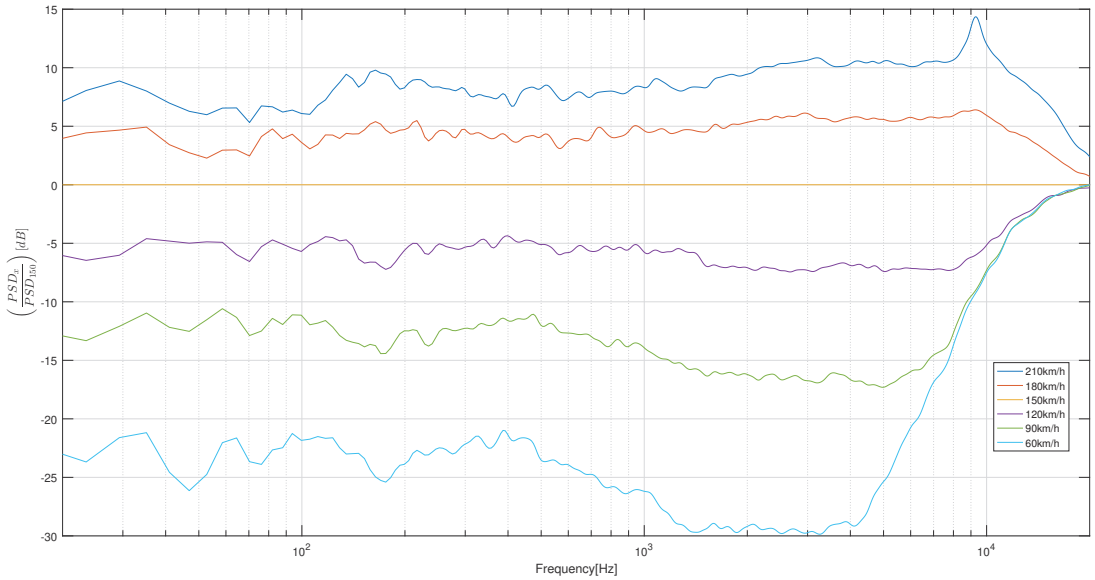


Figure 2.2: PSD of the wind noise for different velocity regions, related to a reference PSD

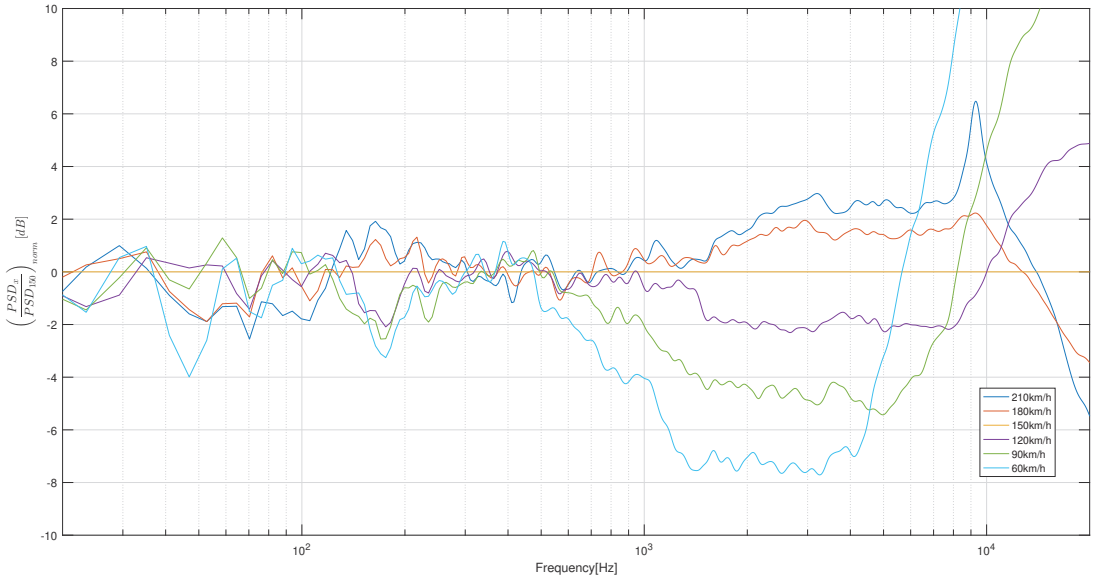


Figure 2.3: PSD of the wind noise for different velocity regions, related to a reference PSD and normalized to the power in the frequency region from 0-500 Hz

To make the spectral colorations even more recognizable, figure 2.3 additionally normalizes all the PSDs that are shown in figure 2.2 by their power in the frequency region from approximately 0 Hz to 500 Hz. This frequency region for normalization was explicitly chosen to obtain the best visibility of spectral changes.

$$H_{v,norm}(k) = \frac{H_v(k)}{\sum_{k=0}^{k\equiv 500\,Hz} H_v(k)} \tag{2.5}$$

Below 500 Hz, the spectral differences between the velocity regions stay within a region of -4 dB to +2 dB, whereas the most prominent differences occur around 40-50 Hz and 180 Hz. Above 500 Hz, a significant spectral coloration with a high-shelve-like characteristic can be observed.

Another phenomenon that can be seen in figure 2.3 is the sharp resonance around 9 kHz in the velocity region of 210 km/h. Such high frequency wind noises could origin from several sources such as the side windows, windscreen wipers, radio antennas or generally cavities and gaps on the car surface [Zel12, p. 205]. Also, a resonance caused by the wind channel for high wind velocities could be an explanation for this peak. Further investigation of the origin of this deterministic wind noise component was not possible with the provided dataset, but was also not crucial for the development of the wind noise simulator part.

The steep ascent of high frequencies for low velocities that is visible in figure 2.3 originates from the noise floor induced by the measurement and is physically not present. As the wind channel run-up measurements contain velocities ranging from about 40 km/h up to 230 km/h in a single measurement, the dynamic range that has to be recorded by the microphones is very big. Furthermore, the spectrum of the wind noise is spectrally tilted towards the low frequencies, so higher spectral components at low velocities may fall below the measurement noise floor.
Looking at figure 2.1, it is evident that the noise floor for this measurement lies at approximately -20 dB.

Another important observation in figure 2.1 is that the resonance peaks in the lower frequency region (47 Hz, 123 Hz, 170 Hz, 252 Hz) only change their amplitudes, but not their center frequencies. This fact allows for simpler filter structures to model the wind noise properly.

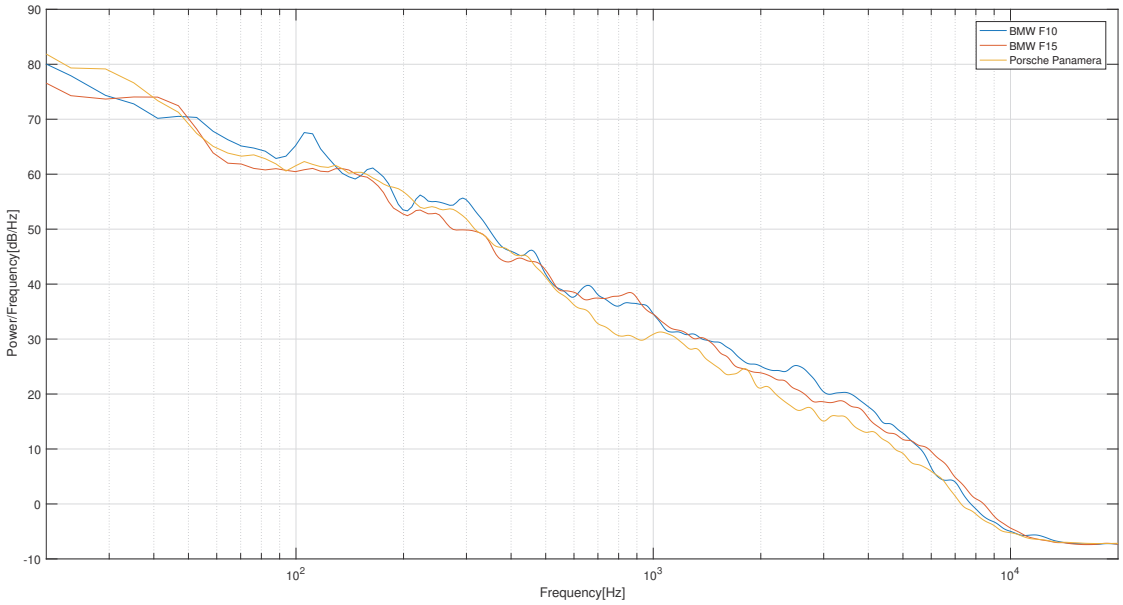
2.1.1.2 Different car models

Different car models have different wind noise characteristics that are originating from their respective car shapes. Figure 2.4b shows the power spectral density of different cars at the same velocity. All those recordings originate from a single omnidirectional microphone placed at the driver's ear position and have been recorded in the same wind channel.

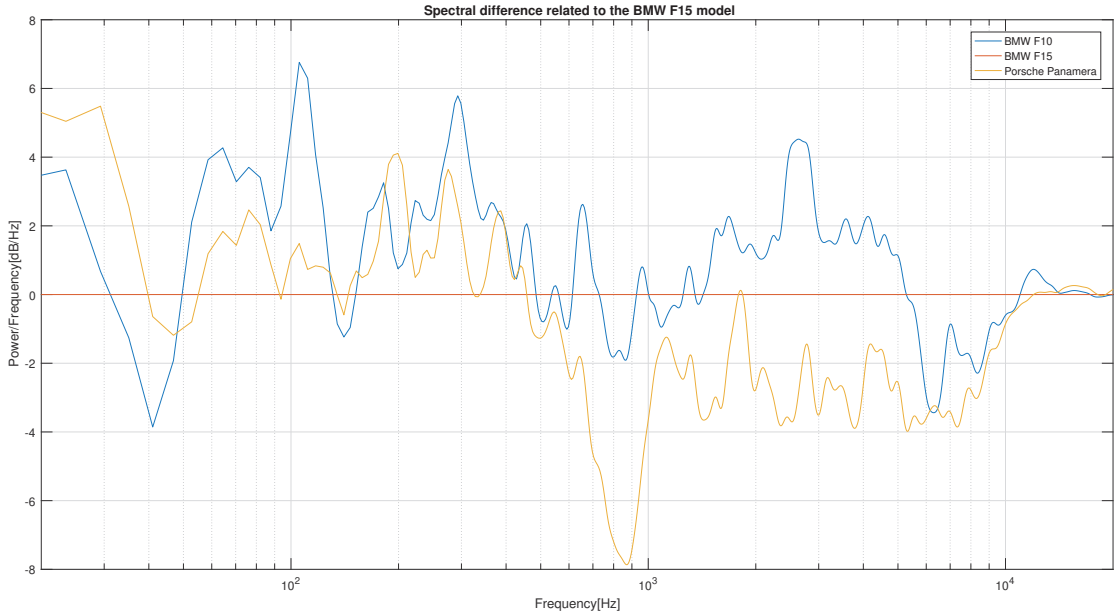
All three wind noise spectra have a similar general spectral shape but differ significantly in their characteristic resonance frequencies, as the shape and the acoustic measures differ significantly from car model to car model. For example, the resonance around 105 Hz is only observable for the BMW F10 model, not for the other two car models.

2.1.1.3 Binaural differences

As the driver's head position is not located symmetrically related to the car shape, spectral differences and level differences between both ears of the driver will occur. Several recordings in a wind channel, using an artificial binaural head, have been analyzed to investigate those differences. Figure 2.5 is showing the PSD of each ear microphone as well as the difference between the two ears for a wind channel measurement at 120 km/h. When looking at the difference between the spectra, it can be observed that the overall level for the left ear is higher than for the right ear, especially in the higher frequency range. This can be explained by the fact that on the driver side, the left ear is closer to the side window, where most of the wind noise is leaked into the car [Zel12, p. 206].



(a) PSD of the different car models



(b) Relative level differences compared to the BMW F15 model

Figure 2.4: PSD of the wind noise for three different car models at a velocity of 120 km/h, smoothed with a 1/12-octave filter

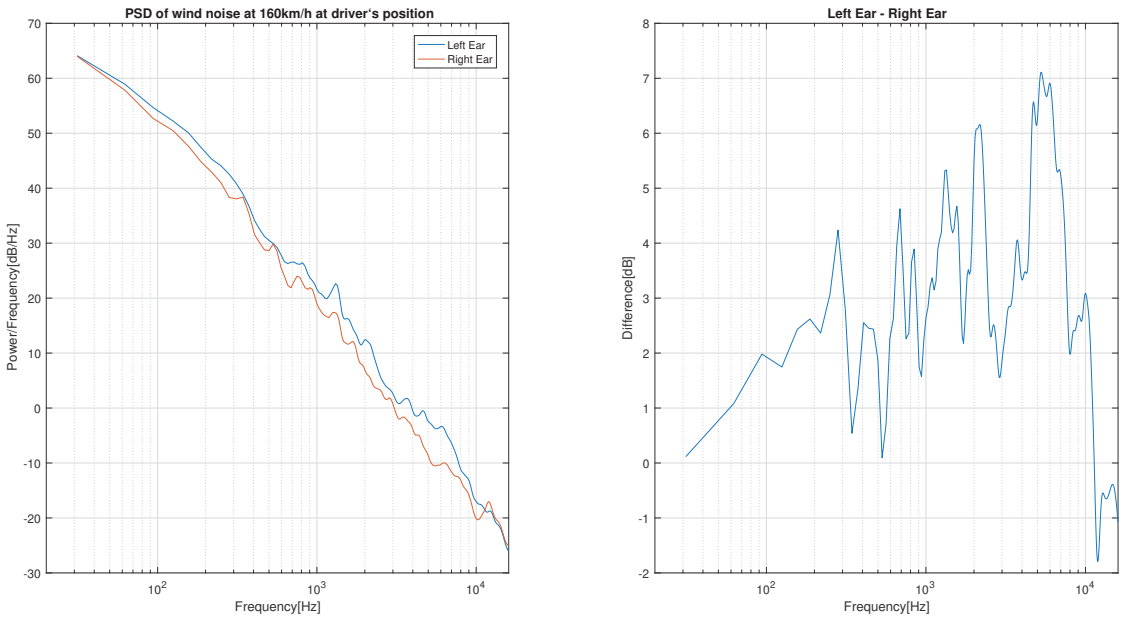


Figure 2.5: PSD of artificial head recordings of a Mercedes S-Class model at the driver's position and the difference between both ears, smoothed with a 1/12-octave filter

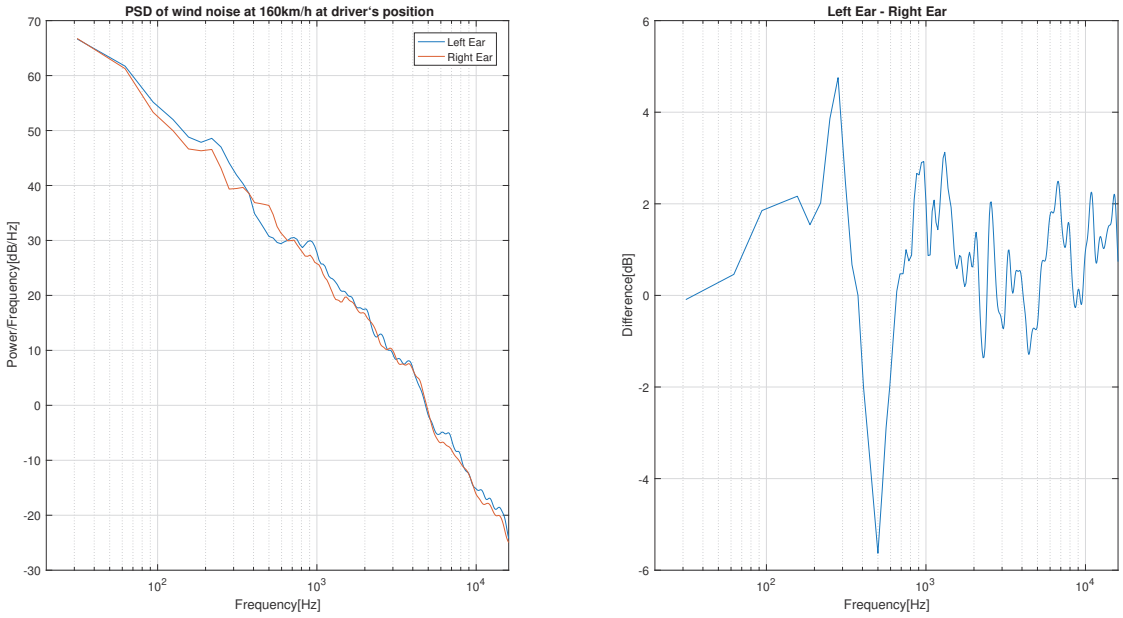


Figure 2.6: PSD of artificial head recordings of a Tesla Model S at the driver's position and the difference between both ears, smoothed with a 1/12-octave filter

2.1.1.4 Yaw angle

Besides all the previously mentioned influences, the angle of air inflow also alters the spectrum and level of the wind noise [RWH06]. A non-frontal incident flow onto the car chassis can occur due to atmospheric conditions or during steering processes. While the car is being steered left or right, the car's nose does not line up with the driving direction anymore. The so-called yaw angle describes the difference between the direction of travel and the orientation of the car. If the flow in coming straight at the vehicle, a positive yaw can be interpreted as rotating the vehicle clockwise while keeping the flow direction constant.

Figure 2.7 is showing the effect of different yaw angles on the spectral coloration and overall level of the wind noise. The analyzed dataset has been measured in a wind channel at a constant wind velocity of 150 km/h using a microphone at the driver's head position. The vehicle itself has been placed at different angles related to the inflow direction to obtain the data for yaw angles ±10°.

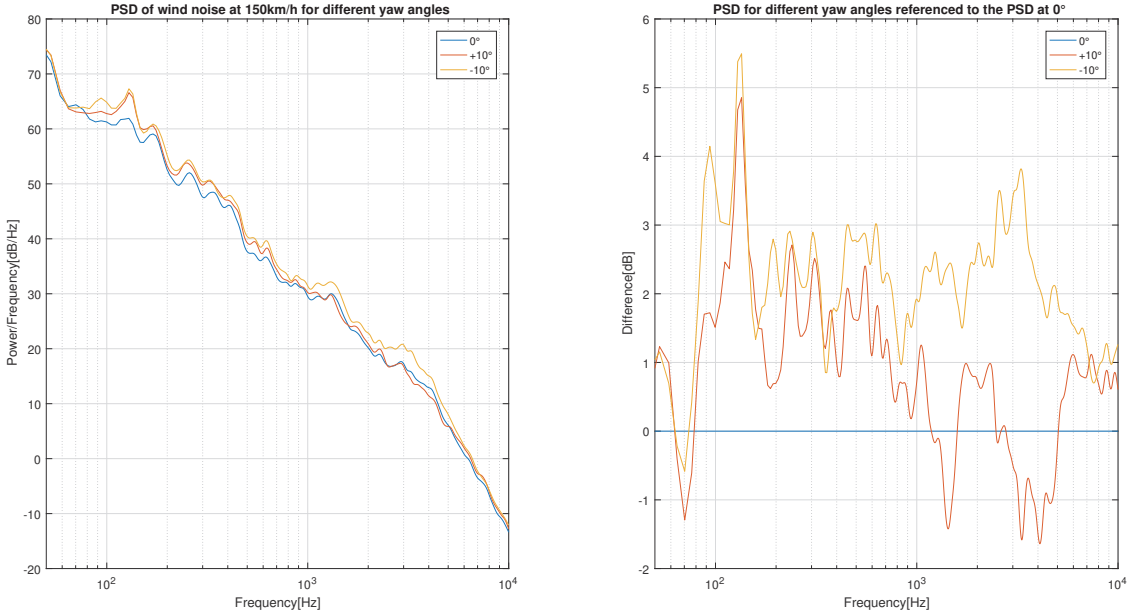


Figure 2.7: PSD of the wind noise of the same car at a velocity of 150 km/h for different yaw angles and the spectral differences when referenced to the PSD at an yaw angle of 0°, smoothed with a 1/12-octave filter

It can be observed that below 1 kHz, both PSDs at an inflow angle of +10° and −10° show similar spectral differences when referenced to the one at 0°. For a yaw

angle of $-10°$, the left front edge of the chassis is acting like an obstacle for the flow, creating the most turbulences out of the depicted scenarios. This circumstance also explains why the wind noise is the loudest for a yaw angle of $-10°$.

2.1.2 Spatial Analysis

As the model for the driving simulator will use spatial sounds, the auralization of the synthesized sounds plays an important role.

Besides the knowledge about the spectral differences between both ears, which was already investigated in section 2.1.1.3, it is also essential to identify whether distinct directions of the driving sounds can be perceived.

A valuable measure for the spatial analysis is the so-called coherence. It describes whether there is a linear dependence between two signals and is often used in dual-microphone speech enhancement algorithms. The magnitude coherence function can be described as ([YKL10])

$$|\Gamma_{Y1,Y2}(f)| = \frac{|P_{Y1,Y2}(f)|}{\sqrt{P_{Y1}(f)}\sqrt{P_{Y2}(f)}} \tag{2.6}$$

where $P_{Y1}(f)$ and $P_{Y2}(f)$ are the power spectral densities of the signals $y_1[n]$ and $y_2[n]$, respectively, and $P_{Y1,Y2}(f)$ describes the cross-power spectral density between $y_1[n]$ and $y_2[n]$ (see section 2.1.1).

The magnitude coherence can hold values between 0 and 1. If the coherence becomes 1, there is a purely linear dependence between $y_1[n]$ and $y_2[n]$. A coherence < 1 suggests that there are uncorrelated components in the microphone signals.

In dual-microphone speech enhancement applications, the magnitude coherence can indicate whether the speech signal is present in certain frequency bands, provided that the speech signal is correlated and the noise signal is uncorrelated. Analogue to that, the magnitude coherence can also be used to investigate whether the wind noise recordings contain correlated components that can be assigned to a specific direction or mainly correspond to a diffuse sound field without any perceived sound source direction.

Relating the coherence measure to the sound field attributes, a coherence of 1 occurs when a pressure field is present, whereas the coherence converges to 0 for a diffuse sound field.

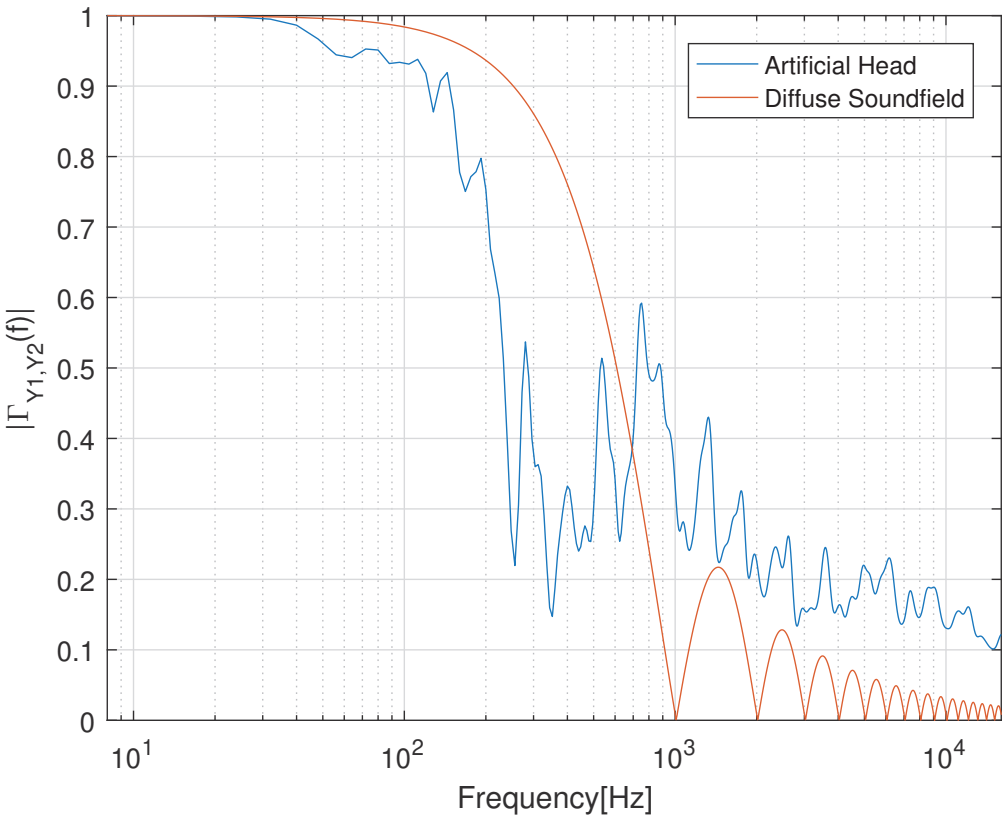


Figure 2.8: Coherence between the two microphones of an artificial head in a Mercedes S-Class at 120 km/h, smoothed with a 1/12-octave filter, compared to the coherence of an ideal diffuse sound field

It is also important to mention that a coherence value of 0 cannot be achieved over the whole frequency range for small microphone distances, even for entirely diffuse sound fields. This circumstance results from the fact that the phase relation between the two microphones is very similar for long wavelengths compared to the distance between the microphones.

The coherence of a perfectly diffuse sound field can be described by an unnormalized sinc function as described in [Pie78]

$$\gamma^2_{diff}(f) = \left(\frac{sin\left(2\pi f \frac{d}{c}\right)}{2\pi f \frac{d}{c}} \right)^2 = sinc\left(f\frac{d}{c}\right)^2 \tag{2.7}$$

with d being the distance between the two microphones and c being the speed of sound ($\approx 343\,\mathrm{m/s}$)

Figure 2.8 is showing the magnitude coherence between the two microphones of an artificial head recording at 120 km/h. It was calculated with the Matlab command `mscohere`. The figure also contains the calculated coherence of a diffuse sound field for two points spaced d = 0.2 m apart (the approximate ear distance). It can be observed that the coherence of the artificial head is mostly corresponding to the one of a diffuse sound field, except for a peak around 800 Hz, where a directional component seems to occur. As this directional component only appears in a very narrow frequency range, it does not play a significant role in the general spatial perception of the wind noise. Therefore, the spatial properties of the wind noise can be approximated by a diffuse sound field.

2.2 Synthesis

As already stated and as it can be seen in figure 2.1, wind noise is basically a broadband noise process with several resonance peaks. The basic idea for wind noise synthesis is to spectrally shape a noise process according to the driving velocity.

2.2.1 Spectral Shaping Methods

Most previous works (section 1.3) have used a filterbank approach to achieve this spectral shaping. A noise process is used as input to the filterbank, which consists of several bandpass filters that run in parallel. The output of each stage is weighted with a gain factor, and the outputs are summed up again to obtain the final output signal. The design of the filterbank can be arbitrary regarding the filter order and center frequencies of the individual bandpass filters. The only restriction for this use case is that the filterbank should be nearly perfectly reconstructible.

Common filterbank designs include filterbanks with linearly or logarithmically spaced center frequencies. Considering the resolution of the human auditory system, filterbanks with logarithmically spaced center frequencies are the preferred choice for this application. Usually, filterbanks with a 1/3-octave resolution are used for the synthesis of noise-like driving sounds.

The higher the resolution (i.e., the number of filters) of the filterbank, the narrower the bandpass filters can be. Followingly, the modeling of the velocity-dependent spectral changes also becomes more accurate. The downside of a high filterbank resolution is the quickly increasing computational effort. Doubling the filterbank resolution would result in a duplication of required filters and a massive increase in

required computational resources. Therefore, it is not feasible to use the filterbank approach if the resulting noise model should also be able to model spectral details.

As an alternative to the inflexible filterbank approach, this thesis investigates Kautz filters to spectrally shape the noise.
Kautz filters have been used in different fields, including loudspeaker equalization [KP06], identification of mechanical systems [DS11] and automotive applications, such as modeling exterior vehicle sound from internal recordings [FHRB16]. One of the major advantages over the filterbank method is the flexible filter structure of the Kautz filter. An arbitrary number of poles can be placed freely along the frequency range. Therefore, the trade-off between computational effort and precision of the synthesis can be adjusted almost continuously. It is also possible to lay the focus on perceptually more relevant frequency ranges. The placement of the poles can be done algorithmically or manually, which adds further flexibility compared to the filterbank approach. Table 2.1 compares the advantages/disadvantages of the Kautz filter over the filterbank approach.

Filterbank	Kautz-Filter
+ Easy implementation	- Not too easy to understand without background knowledge
+ Direct adjustment of the quality by number of bands	- More poles do not necessarily improve the quality
- Specific frequencies cannot be targeted well	+ Can model specific spectra more exact
- Inflexible structure	+ Automatic pole detection/ manual pole tuning possible

Table 2.1: Comparison between Filterbank and Kautz-Filter structure

2.2.2 The Kautz Filter

The most basic Kautz filter consists of a filter structure shown in figure 2.9.

It consists of several output stages that are each separated by an all-pass filter that achieves orthogonality between every output stage.

The transfer function of the whole filter structure is given as ([KP06])

$$H(z) = \sum_{i=0}^{N} w_i G_i(z) = \sum_{i=0}^{N} w_i \left(\frac{\sqrt{1 - z_i z_i^*}}{z^{-1} - z_i^*} \prod_{j=0}^{i} \frac{z^{-1} - z_j^*}{1 - z_j z^{-1}} \right) \tag{2.8}$$

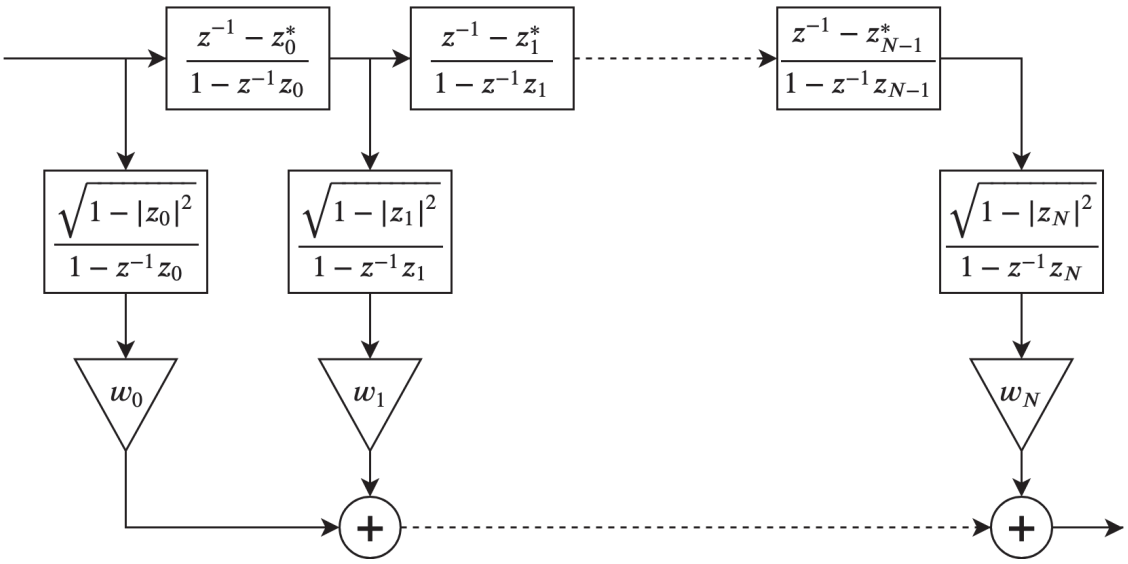


Figure 2.9: Basic Kautz filter structure with N+1 tap outputs

or in time-domain as the impulse response of the filter structure

$$h[n] = \sum_{i=0}^{N} w_i g_i[n] \tag{2.9}$$

with $g_i[n]$ being the impulse responses of the single tap outputs, i.e. the inverse z-transformed transfer functions of each tap output stage $\mathscr{Z}^{-1}\{G_i(z)\}$.

The orthonormal behavior of the output stages leads to the following attributes:

$$\langle g_i, g_j \rangle = 0 \text{ for } i \neq j \tag{2.10}$$

$$\langle g_i, g_j \rangle = 1 \text{ for } i = j \tag{2.11}$$

In order to describe a Kautz filter completely, only two different parameter sets have to be determined: the poles z_i and the respective tap output weights w_i.

While the determination of the pole-set is a non-trivial task, the calculation of the filter weights w_i is a straightforward process. As the tap output signals are orthogonal, the optimal weights can be calculated by solving a simple least-squares problem. The gain w_i for each output stage can be calculated by

$$\begin{aligned} x_i[n] &= h[-n] * g_i[n] \\ w_i &= x_i[0] \end{aligned} \tag{2.12}$$

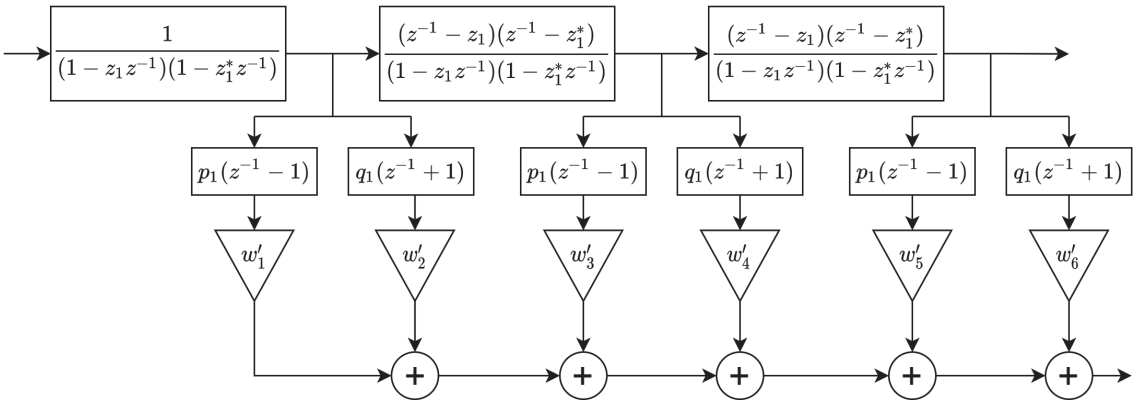


Figure 2.10: Efficient implementation of a Kautz filter structure for conjugate complex pole pairs

An efficient realization for conjugate complex pole pairs that lead to real tap-output responses can be seen in figure 2.10. The coefficients p_i and q_i are calculated as shown in equation 2.13.

$$\begin{aligned} \rho_i &= |z_i|^2 \\ \gamma_i &= -2\Re\{z_i\} \\ p_i &= \sqrt{(1-\rho_i)(1+\rho_i-\gamma_i)/2} \\ q_i &= \sqrt{(1-\rho_i)(1+\rho_i+\gamma_i)/2} \end{aligned} \tag{2.13}$$

2.2.2.1 Pole Detection - BU Method

As the determination of the best fitting pole positions is a non-trivial problem found in many application fields, the amount of methods that can be applied is pretty significant.
Karjalainen and Paatero suggested the usage of an iterative method called the "BU-method" [KP06]. The name origins from Brandenstein and Unbehauen, who proposed this method in 1998 [BU98].

The application of the "BU-method" regarding the determination of ideal poles for Kautz filters is described compactly in [PK03]:

Basically, the algorithm is based on the approximation of the all-pass term:

$$\hat{A}^{(k)}(z) = \frac{z^{-N}D^{(k)}(z^{-1})}{D^{(k-1)}(z)} \tag{2.14}$$

$D^{(k)}(z)$ are polynomials that are generated iteratively (k = 1,2...):

$$D^{(k)}(z) = 1 + \sum_{i=1}^{N} d_i^k z^{-i} = 1 + D_1^{(k)}(z) \tag{2.15}$$

$\hat{A}^{(k)}(z)$ converges to an all-pass when $||D^{(k)}(z) - D^{(k-1)}(z)|| \to 0$

The output $U^{(k)}(z) = \hat{A}^{(k)}\,[X(z)]$ should be minimized, with $X(z) = z^{-L}H(z^{-1})$, which represents the z-transform of the time-inverted impulse response $\mathscr{Z}\{h[-n]\}$.

Defining $Y^{(k)}(z) = X(z)/D^{(k-1)}(z)$ and substituting results in

$$Y^{(k)}(z)z^{-(N-1)}D_1^{(k)}(z^{-1}) = U^{(k)}(z) - z^{-N}Y^{(k)}(z) \tag{2.16}$$

which can be written as a matrix equation

$$\begin{aligned} \mathbf{A}^{(k)}\mathbf{d}^{(k)} &= \mathbf{u}^{(k)} + \mathbf{b}^{(k)} \\ \mathbf{u}^{(k)} &= \mathbf{A}^{(k)}\mathbf{d}^{(k)} - \mathbf{b}^{(k)} \end{aligned} \tag{2.17}$$

The square norm of $\mathbf{u}^{(k)}$ gets minimized for

$$\mathbf{d}^{(k)} = \mathbf{A}^{(k)}/\mathbf{b}^{(k)} \tag{2.18}$$

This calculation is repeated iteratively, calculating the next values for $Y^{(k)}(z)$ by filtering $h[-n]$ with $1/D^{(k-1)}(z)$ and solving the matrix equation again.

After several iterations, $D^{(k)}(z)$ that produces the smallest error in a least-squares sense is selected as the best fitting polynomial. To obtain the poles, the roots of the polynomial have to be calculated.

2.2.2.2 Frequency Warping

To additionally optimize the pole placement to the human auditory system, it is beneficial to place more poles in the lower frequency region than in the higher frequency region. One possibility to achieve this without fundamentally changing the pole detection algorithm is to frequency warp the FFT spectrum before executing the algorithm. Basically, the spectrum is getting transformed so that low frequencies get "stretched" over a more extensive frequency range, while high frequencies get compressed to a smaller frequency region.

One suitable frequency representation, that was also used in audio application case 2 of [PK03], is the Bark scale [Zwi61]. Instead of evenly spaced frequency bins, it splits the frequency range into 24 bands that correspond to the first 24 critical bands of the human auditory system. Table 2.2 is showing all 24 Bark bands with their respective center frequency f_c, lower band edge f_1 and higher band edge f_2.

Bark	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
f_c	50	150	250	350	450	570	700	840
f_1	0	100	200	300	400	510	630	770
f_2	100	200	300	400	510	630	770	920
Bark	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
f_c	1000	1170	1370	1600	1850	2150	2500	2900
f_1	920	1080	1270	1480	1720	2000	2320	2700
f_2	1080	1270	1480	1720	2000	2320	2700	3150
Bark	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
f_c	3400	4000	4800	5800	7000	8500	10500	13500
f_1	3150	3700	4400	5300	6400	7700	9500	12000
f_2	3700	4400	5300	6400	7700	9500	12000	15500

Table 2.2: Center frequencies f_c and band edges f_1, f_2 of the 24 Bark bands

An approximation of the Bark frequency distribution from a common spectrum can be achieved by using a first-order bilinear transformation with the unit circle mapped to itself (DC component mapped to DC component and $fs/2$ mapped to $fs/2$)[SA95]

$$z = A_\rho(\zeta) = \frac{\zeta + \rho}{1 + \zeta\rho} \qquad\qquad \rho = \frac{\zeta_3 - z_3}{1 - z_3\zeta_3} \tag{2.19}$$

where ρ is an all-pass coefficient and is the only degree of freedom that needs to be tuned to fit the Bark scale as good as possible

$$\rho^* = \arg\min_\rho \{||a(\omega) - b(\omega)||\} \tag{2.20}$$

where $||a(\omega) - b(\omega)||$ describes the L2-norm of the difference between the approximated bark scale and the actual bark scale. Unfortunately, an easy minimization is not possible as the frequency error is nonlinear in ρ. Therefore,

$$\arg\min_\rho \left\{||e^{ja(\omega)} - e^{jb(\omega)}||\right\} \tag{2.21}$$

is used instead, which yields similar results.

After further derivations cf. [SA95], we get

$$\rho^* = \frac{\sum_{k=1}^{K} \{cos\,[b(\omega_k)] - cos(\omega_k)\}\, v(\omega_k)}{\sum_{k=1}^{K} \{cos\,[b(\omega_k) + \omega_k] - 1\}\, v(\omega_k)} \qquad v(\omega_k) = \frac{1}{1 + \rho^2 - 2\rho cos(\omega_k)} \tag{2.22}$$

As $v(\omega_k)$ is depending on ρ, and therefore is unknown, the unweighted solution for ρ^* is calculated first by setting $v(\omega_k) = 0$. Subsequently, $v(\omega_k)$ is calculated with ρ^* instead of ρ and the weighted solution is recalculated by using the calculated values for $v(\omega_k)$.

2.2.3 Signal Generator

As wind noise can be described as a colored broadband noise process, white noise generators are a commonly used tool that is also available on most signal processing devices. Still, the constant power spectral density over the whole frequency region for white noise does not fit the power spectral density of the wind noise (shown in figure 2.1) well.

A better suitable frequency distribution can be achieved by using pink noise. Pink noise also contains all frequencies, but contrary to white noise, the power spectral density of pink noise is distributed according to $\frac{1}{f}$. This results in constant energy in each single octave band. Looking at figure 2.1, it can be concluded that pink noise would fit the spectral trend of wind noise better than white noise.

To obtain an even better starting point for the spectral shaping, the pink noise generator can be replaced by a looped wavetable, containing pre-colored noise that represents the wind noise spectrum at a common velocity. This way, vehicle-specific resonance characteristics can already be found in the noise process and the required filter complexity to achieve a satisfying result can be lowered.

The recreation of the noise will be most accurate around the reference velocity chosen for the creation of the wavetable. As wind noise becomes the most prominent part of the driving sound composition for velocities above 100 km/h, the reference velocity in this model has been set at 120 km/h.

To remove non-stationarities, eventually present non-linear effects, and other disturbances (that might be repeated constantly, as the wavetable is played back periodically) from the reference noise, the basic idea is to spectrally analyze the wind

channel recording at the reference velocity and recreate the spectrum by filtering a random noise process. Directly using the wind channel run-up recording for this purpose would not only bring the previously mentioned problems but would also result in a constant rise of volume within the wavetable sample due to the constantly increasing wind velocity of the run-up recording.

A 1/12-octave filter bank has been used to analyze and recreate the spectral attributes of the wind channel recording. Basically, the power in each filter bank band of the reference wind channel recording has to be calculated. The random noise process then gets filtered with the same filter bank, and every band gets multiplied with the corresponding calculated power of the reference wind channel recording. As additional spectral smoothing of the wind channel recording is desired before the analysis, the complete analysis and calculation of the power in each band is executed in the frequency domain. The detailed calculation is outlined in section 5.1.

2.2.4 Turbulent flow conditions

The standard wind channel for aeroacoustic vehicle measurements is producing laminar flow conditions. Contrary to this, the airflow on the street is never perfectly laminar, but comes from different angles and fluctuates in flow velocity. Weather conditions and other vehicles on the street are disturbing the airflow even before the air hits the vehicle of interest. This phenomenon results in different acoustically relevant effects such as level fluctuations and changes in the spectral coloration of the wind noise.

Several papers proposed methods to simulate on-the-road wind noise in wind channels. These methods include modifications of the measurement setup that directly influences the wind channel measurements [RWH06] and manipulation of wind noise data recorded "conventionally" in a wind channel under laminar flow conditions ([CD19],[RWH06],[OSWD14],[KRW09]).

The majority of those papers investigate this issue due to sound design reasons. Stationary wind noise recordings originating from a wind channel that sound acceptable for listeners may not be acceptable for them anymore at all with the modulations that occur on the street, as the human ear is sensitive to volume fluctuations. By simulating this data from wind channel recordings, the evaluation of optimizations of the car shape and consecutively their influences on the acceptability of the wind

noise can become a lot faster and straightforward. Furthermore, the findings of those papers can also be used to improve realism in the synthesis of wind noise for driving simulators.

Commonly, the simulation of turbulent wind noise is achieved by piecing together various stationary wind channel measurements at different flow velocities and yaw angles according to measured or simulated flow conditions [KRW09]. The disadvantage of the presented methods is the lack of real-time-capability necessary for a driving noise simulator.

The developed synthesis framework here makes the a real-time implementation of a similar algorithm possible. Analogue to the proposed algorithms, a wind noise sample recorded under stationary conditions gets modified according to the flow velocity and the yaw angle. As the coloration of the wind noise sample is executed by the Kautz filter structure, all that has to be done is modulating the input parameters (velocity and yaw angle) accordingly.

Unfortunately, the provided dataset did not contain any appropriate measurements to evaluate the airflow behavior regarding vehicles on the street. Therefore, instead of measured data, a high-pass filtered white noise process has been used to modulate the input parameters. As this was done without any backing of real data, this part of the synthesis was tuned to sound perceptually well and should be revisited later, when appropriate data is available.

2.3 Results

Figure 2.11 is showing the PSD of real and synthesized wind noise samples for both 100 km/h and 200 km/h. The PSDs have intentionally not been smoothed to emphasize the spectral similarity between the real and synthesized sounds. The visible ripple in the PSD of the synthesized sounds for high frequencies is caused by the filterbank used to create the basic wavetable. While the ripple is clearly visible in the figure, it does not create any audible difference as the frequencies are too high and the level difference is too low. The perceptual evaluation of the synthesized wind noise is part of the listening test in section 6.1.2.

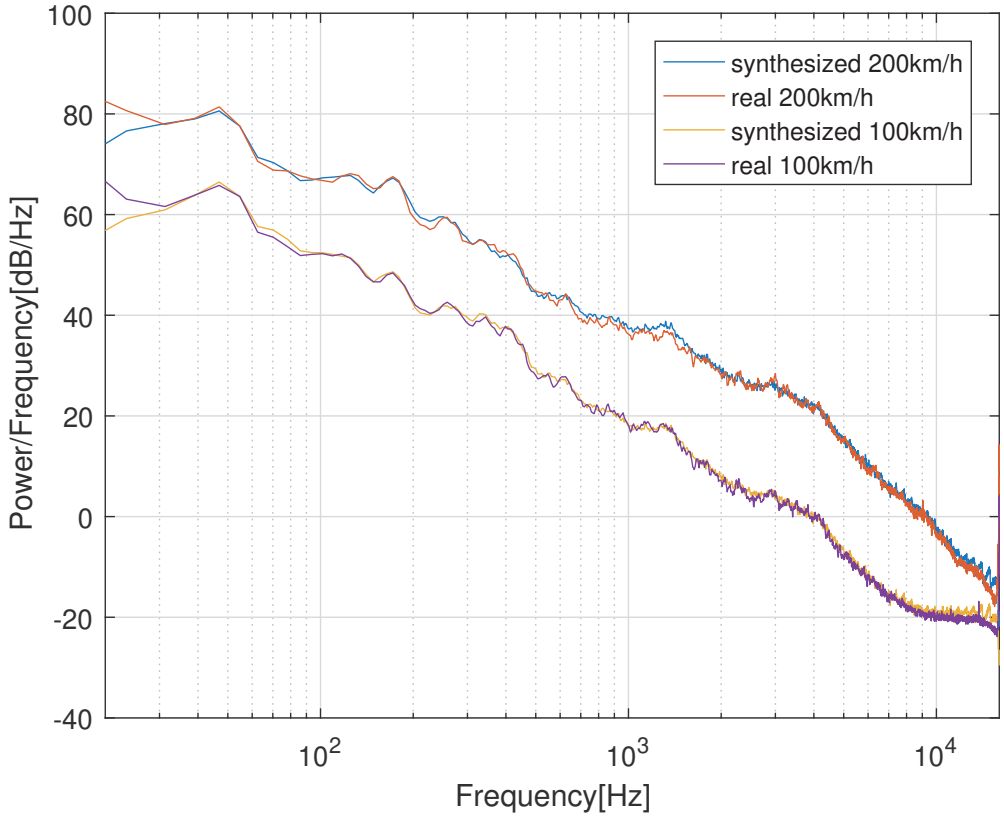


Figure 2.11: Comparison between the PSD of real and synthesized wind noise

Chapter 3

Rolling Noise

3.1 Analysis

Similar to the wind noise, real road measurements always contain other noise sources and therefore are suitable only to a limited extent to analyze the rolling noise over a sufficiently large velocity range. The increasing dominance of wind noise with increasing velocity limits the usability of road measurements for high velocities. The influence of the engine noise can be minimized by measuring the rolling noise after accelerating the car and disabling the engine. The velocity range of measurements using this method is often limited by the length of the test track, as the velocity is only decreasing slowly when the vehicle is rolling out without the engine brake slowing down the vehicle or the brake being engaged actively.

Another specific method to measure the rolling noise only is using an acoustic dynamometer roll.

The rolling noise data provided for this thesis contains real road measurements on different road surfaces in a limited velocity range as well as measurements on the acoustic roller dynamometer that cover a wider velocity range.

Contrary to wind noise, rolling noise also contains harmonic components (their physical origin has already been described in section 1.2.2). Due to the changing frequency components, it makes sense to first look at the rolling noise data in the frequency domain using a time-frequency representation.

3.1.1 Spectrogram

Basically, a spectrogram is a visual representation of the result of a STFT[1]. The idea of the STFT is to extend simple frequency analysis (like usage of a DFT[2]) by a time component. This becomes necessary for the analysis of non-stationary signals. Mathematically, the STFT can be described by the DFT of the input signal multiplied with a shifted window function [OS09].

$$X[n, \lambda) = \sum_{m=-\infty}^{\infty} x[m]w[m-n]e^{-j\lambda m} \tag{3.1}$$

The spectrogram then visually represents the STFT with the abscissa and ordinate describing time and frequency, whereas the absolute value of each STFT bin is represented by a color, according to its magnitude.

The selection of the window length plays a crucial role in the application of the STFT. Short windows lead to good time resolution but bad frequency resolution, whereas the opposite is the case for long windows. Therefore, it is always important to adjust the window length to the properties of the signal to be analyzed.

A rolling noise measurement on an accelerating dynamometer contains tire harmonics that constantly change their frequencies according to the velocity. The basis function of the DFT (also used in the STFT) is a complex exponential function that describes the oscillation for a particular frequency that is constant within each DFT frame. As this is not the case for the tire harmonics of a dynamometer measurement, a complex exponential sweep would be a better fit. This leads to the fan chirp transform.

3.1.2 Fan Chirp Transform

The fan chirp transform in its general form is described by ([WK07])

$$X(f, \alpha) \widehat{=} \int_{-\infty}^{\infty} x(t)\sqrt{|\phi'_\alpha(t)|}e^{-j2\pi f \phi_\alpha(t)}dt \tag{3.2}$$

where ϕ_α is a second-order polynomial with α being the chirp rate

$$\phi_\alpha(t) \widehat{=} (1 + 0.5\alpha t)t \tag{3.3}$$

[1] Short-Time Fourier Transform
[2] Discrete Fourier Transform

The instantaneous frequency corresponds to the derivative of the phase and therefore is given by

$$v(t) = f\frac{d\phi_\alpha(t)}{dt} = (1+\alpha t)f \tag{3.4}$$

which basically represents linear sweeps, whose slope is dependent on the frequency f and the chirp rate α.

As the calculation of the fan chirp transform is computationally very expensive when being done this way, an alternative approach is commonly used. Instead of using equation 3.4, $x(t)$ is first time-warped and consecutively a simple FFT is conducted. This method takes advantage of the higher efficiency of the FFT algorithm.

$$X(f,\alpha) \widehat{=} \int_{-\infty}^{\infty} x(\phi_\alpha^{-1}(\tau))e^{-j2\pi f\tau}d\tau \tag{3.5}$$

Similar to the STFT (equation 3.1) to obtain a frequency as well as time-dependent analysis, the transformation is applied to windowed time segments of the whole signal.

Using the methods described above, it is now possible to analyze the rolling noise measurements in more detail.

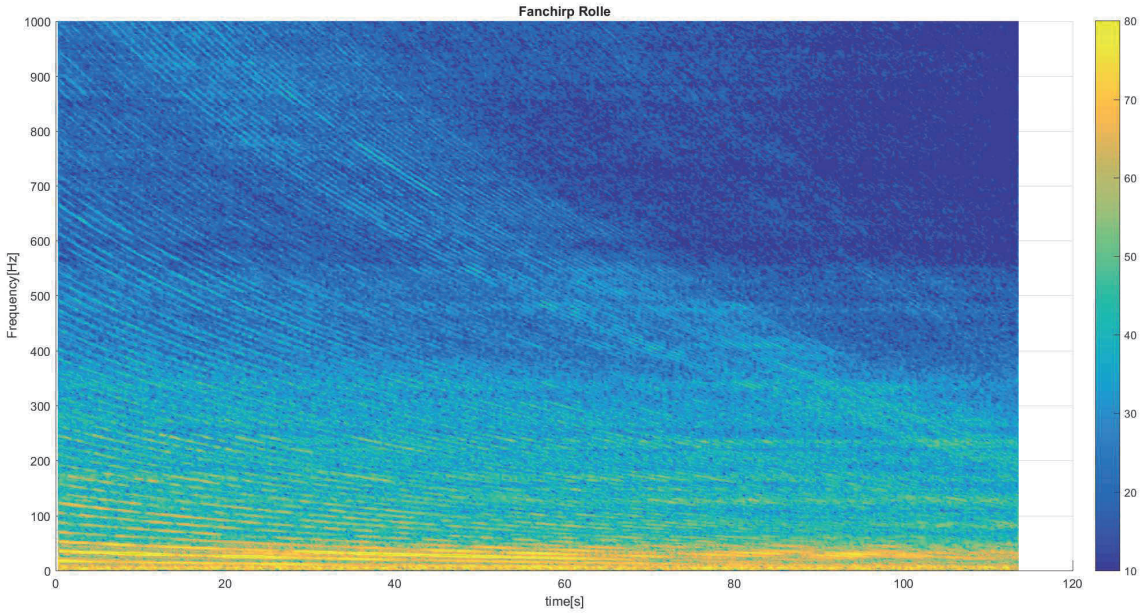


Figure 3.1: Time-frequency representation of the dynamometer measurement using the Fan Chirp Transform

Figure 3.1 is showing the result of the fan chirp transform being applied to the measurement of the acoustic roller dynamometer. The dynamic range of the picture starts at 10 dB (dark blue) and goes up to 80 dB (yellow). The velocity starts at 120 km/h (for t = 0 s) and goes down to 20 km/h (for t = 113 s) at a non-linear slope.

Besides the noise-like spectrum, several tire orders can be observed. The horizontal regions with a higher magnitude are resonances (e.g., caused by the tire cavity modes, structural conditions, etc. - see section 1.2.2). The noise component caused by the tire profile pitches can be observed as a broad region with higher noise magnitude, reaching from around 1 kHz at t = 0 s down to approximately 300 Hz at t = 113 s.

3.1.2.1 Acoustic Roller Dynamometer

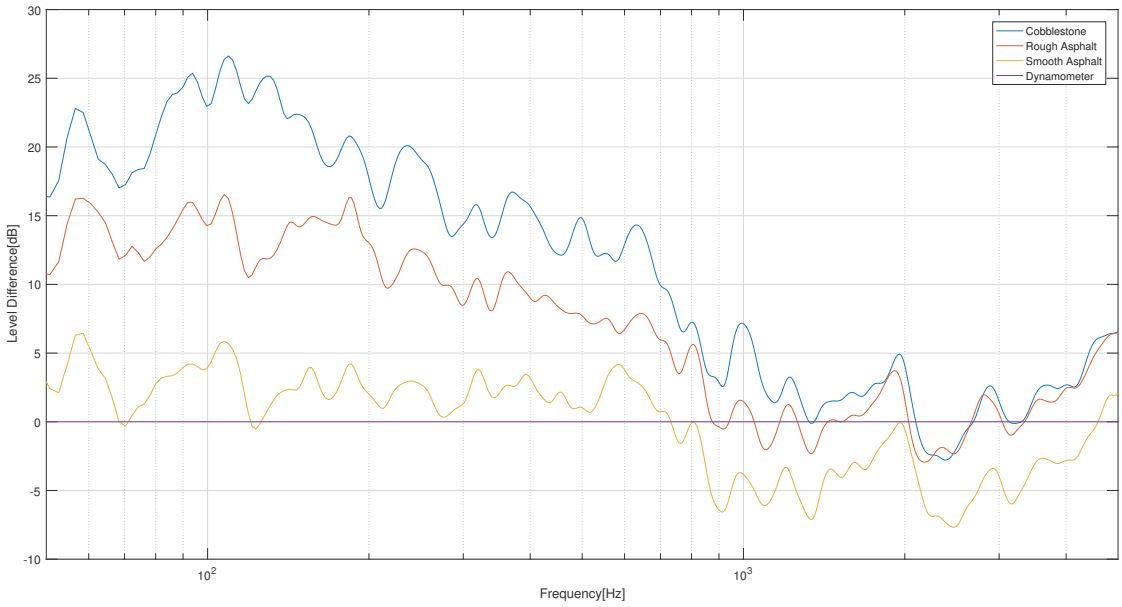
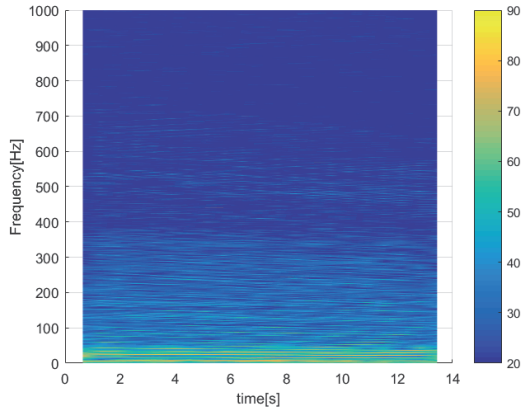


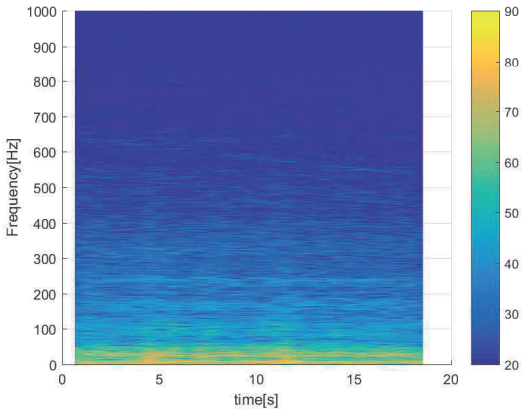
Figure 3.2: Level difference between different street measurements and an acoustic dynamometer roll measurement at an average velocity of 55 km/h, smoothed with a 1/12-octave filter

Figure 3.2 is showing the relative level difference of different road surfaces compared to the dynamometer measurement. The overall level of the rolling noise, especially in the frequency region below 700 Hz, is becoming louder as the surface becomes rougher.

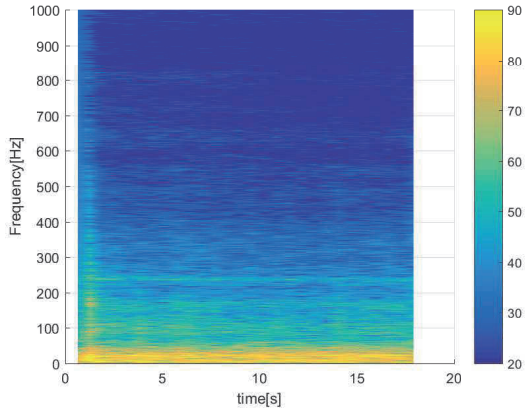
Comparing the fan chirp transform of road measurements (figure 3.3b, 3.3d and 3.3c) to the measurement on the acoustic roller dynamometer (figure 3.3a), an interesting behavior can be observed. The tire order structure that is clearly visible for the dynamometer measurements is barely visible anymore under real road conditions. It gets smeared over a wider frequency range.



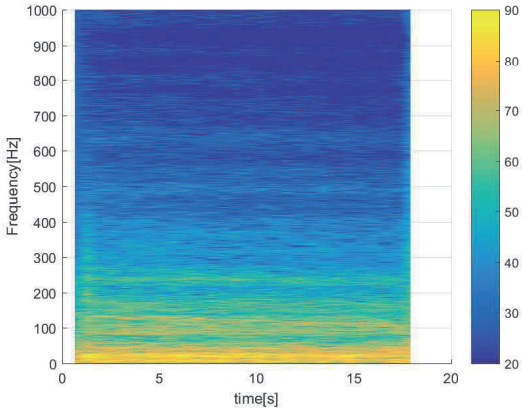
(a) Fan Chirp Dynamometer 60-50 km/h



(b) Fan Chirp Smooth Asphalt 60-50 km/h



(c) Fan Chirp Rough Asphalt 60-50 km/h



(d) Fan Chirp Cobblestone 60-50 km/h

Figure 3.3: Time-frequency representation of multiple coast downs from 60 km/h to 50 km/h on different surfaces

The harmonics that appear in the dynamometer measurement have to be filtered out first to obtain the noise characteristics of the rolling noise.

3.1.3 Vold-Kalman Filter

The Vold-Kalman filter was first proposed in 1993 by Håvard Vold and Jan Leuridan [VL93]. A second generation of the filter was proposed in 1997. The second-generation

Vold-Kalman filter basically is an order tracking filter with many advantages, such as a flat passband, easy specification of the filter bandwidth, decoupling of close or crossing orders, and no phase error [VMB97]. Since the phase of the extracted orders is not manipulated, those orders can theoretically also be used for re-synthesis.

The principle of this second generation Vold-Kalman filter ([VMB97]) is briefly summed up here:

The orders to be tracked are described as complex phasors, weighted with a time-varying, order-dependent complex envelope term $A_k(t)$, containing the magnitude and phase information. The whole order structure can then be described as a sum over the phasors, weighted with their respective complex envelope term:

$$X(t) = \sum_k A_k(t)\Theta_k(t) \tag{3.6}$$

$$\Theta_k(t) = e^{j2\pi k \int_0^t \omega(u)du} \tag{3.7}$$

with k describing the order of the phasor.

The integral in the exponent describes the cumulated phase term and is necessary due to the time-variant instant frequency of the oscillations. As the resulting signal is real-valued, the orders have to occur in complex conjugates, i.e., for every positive k, there has to be its negative counterpart.

The idea of the Vold-Kalman is based on two local constraints:

- the complex envelope terms $A_k(t)$ are smooth (structural equation)
- the sum over the orders should approximate the total measured signal, except for the noise component (data equation)

Those constraints can be described mathematically by:

Structural Equation:

$$\frac{d^\xi A_j(t)}{dt^\xi} = \epsilon(t) \tag{3.8}$$

or for sampled data

$$\nabla^\xi A_j[n] = \tilde{\epsilon}[n] \tag{3.9}$$

where $\epsilon(t)$ or $\tilde{\epsilon}[n]$ contains higher orders that cannot be described with a low-order polynomial approximation.

Data Equation:

$$X[n] - \sum_{j \in J} A_j[n]\Theta_j[n] = \tilde{\delta}[n] \tag{3.10}$$

where $\tilde{\delta}[n]$ contains the noise component and any modelling error.

The phasors can be calculated from corresponding RPM[3] measurements. The structural equation and the data equation then are used to determine the unknown complex envelope terms $A_k(t)$. The bandwidth of the Vold-Kalman filter can be influenced by the relative weighting between structural- and data equation.

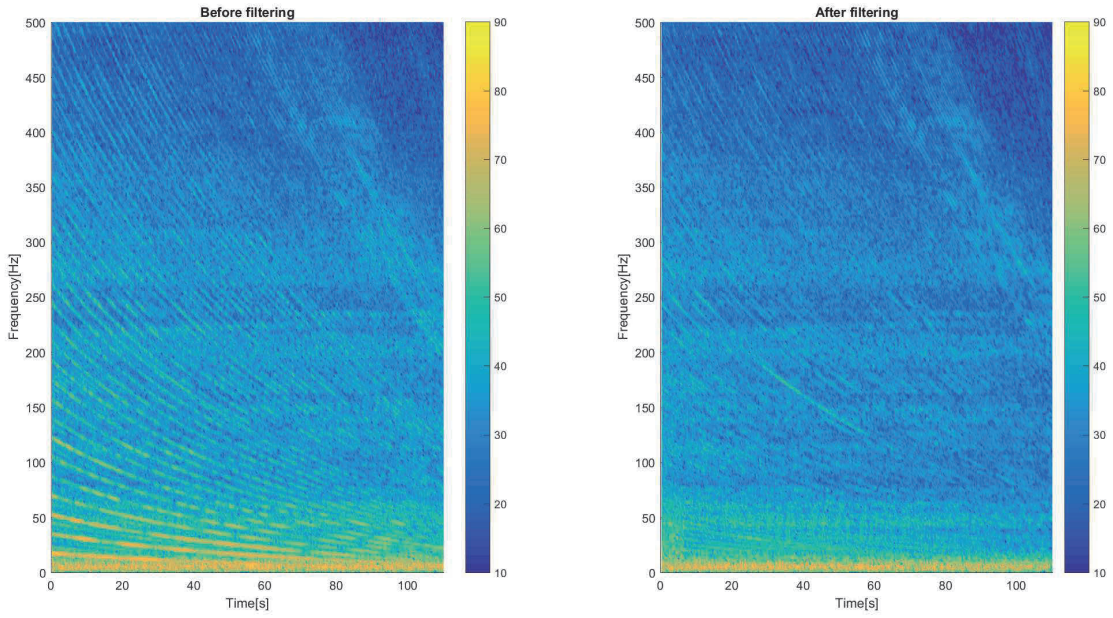


Figure 3.4: Spectrogram of an acoustic roller dynamometer measurement before and after removing the first 100 tire orders using a Vold-Kalman filter

After removing the tire orders from the recorded data, a predominantly noise-like signal remains. This signal can be treated similarly as it was already done previously in section 2.1.1.

[3]Revolutions per Minute

3.1.4 Rolling Circumference

The previous sections required knowledge about the tire rotation frequency. As this measure was not present in the dataset, the rotation frequency of the tires had to be estimated from the vehicle velocity. Those two measures are associated by the constant C - the rolling circumference

$$f_{0,r} = \frac{v}{C} \tag{3.11}$$

where $f_{0,r}$ is the rotation frequency of the tire and v is the velocity of the vehicle.

For the first rough estimation, $f_{0,r}$ has been read out of a spectrogram, and C has been calculated by using equation 3.11. This result is only a first approximation, as the read-out value for $f_{0,r}$ is discretized and therefore not necessarily the actual value. Consecutively the correlation between the signal and a complex phasor is calculated (similar method to the Vold-Kalman filter in section 3.1.3).

$$A(C) = \sum_{k=0}^{L} \sum_{n=0}^{N} x[n, C] e^{-j\Theta[n]k} \tag{3.12}$$

$$\Theta[n, C] = \frac{2\pi}{fs} \sum_{n=0}^{N} \frac{v[n]}{C} \tag{3.13}$$

where L is the number of orders to be summed up. This calculation has been repeated with varying C until the correlation $A(C)$ is maximized.

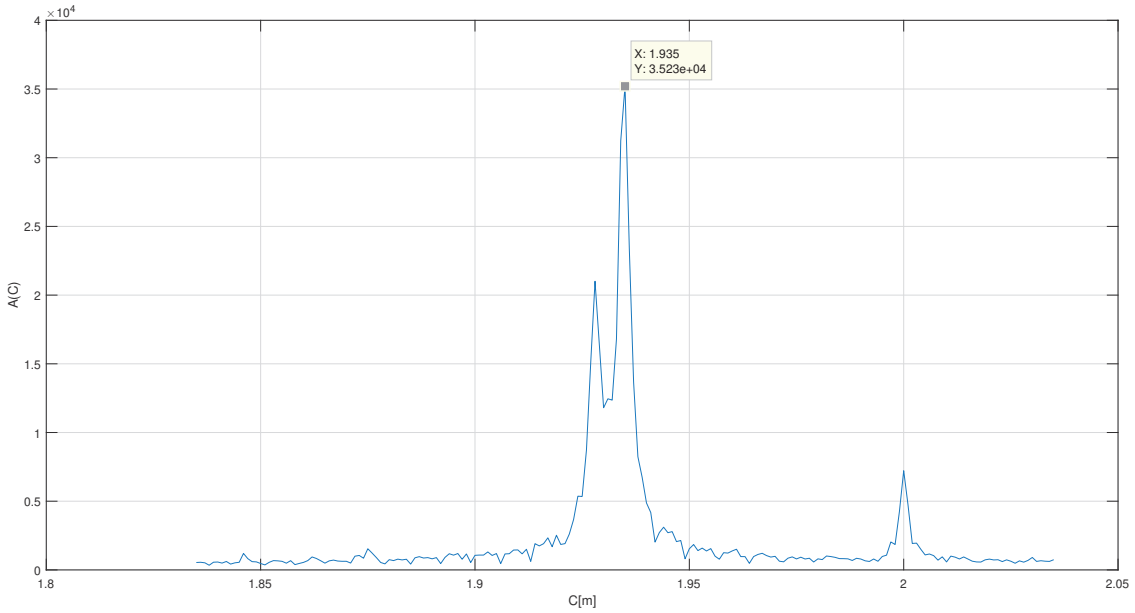
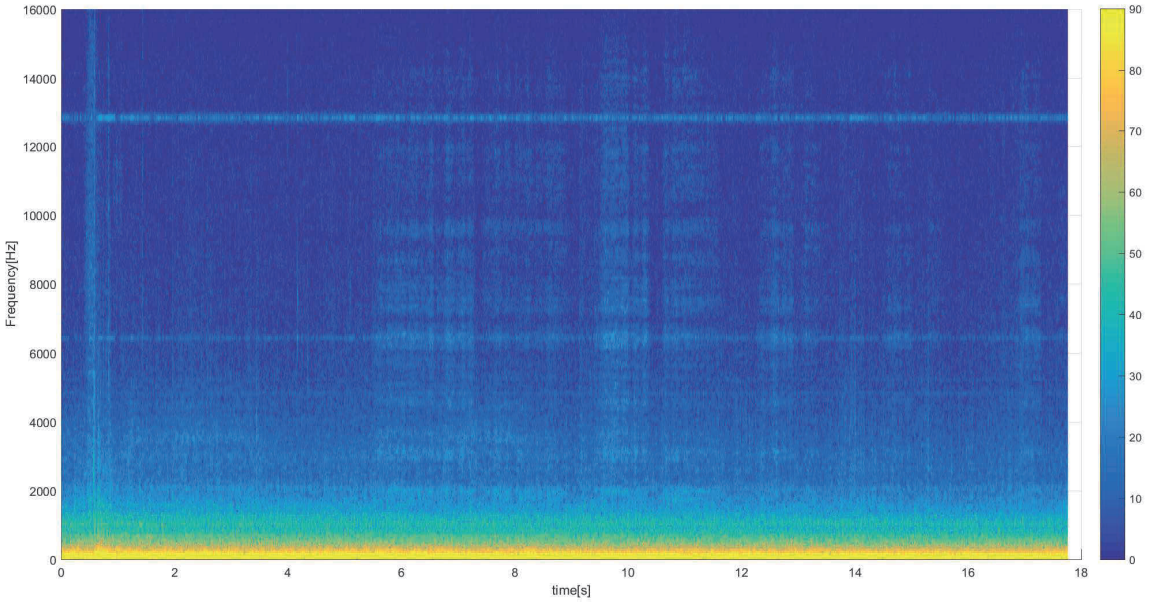


Figure 3.5: Finding the value for C so that the correlation is maximized

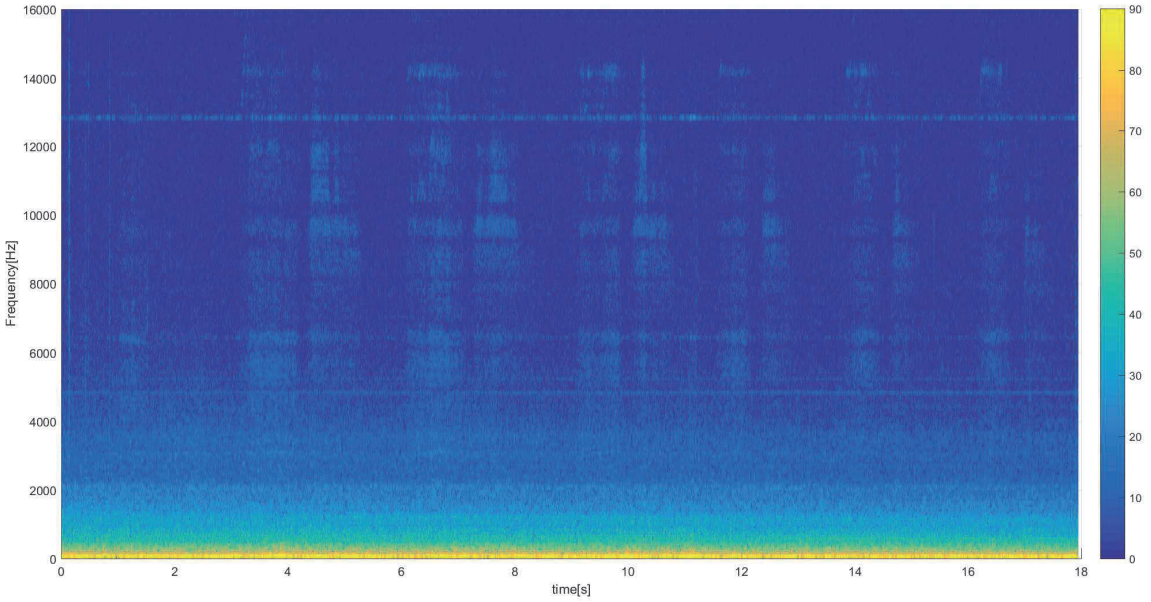
3.1.5 Road Bumps

As real roads do not have a homogeneous surface, the excitation of the tire also varies. The rougher a road surface, the more impulsive elements are added to the rolling noise. So the cobblestone measurements, for example, contain a lot of audible impulsive events.

Figure 3.6a shows the spectrogram of a cobblestone measurement. The FFT length was shorter ($N = 512$) to better recognize the impulsive events. Those impulsive events can be identified as the thin vertical blue lines. Figure 3.6b shows the spectrogram with the same settings for a smooth asphalt measurement for comparison.



(a) Spectrogram of a rolling noise measurement on cobblestone



(b) Spectrogram of a rolling noise measurement on smooth asphalt

Figure 3.6: Comparison between two spectrograms of road measurements at approximately 50 km/h

3.2 Synthesis

3.2.1 Noise Synthesis

To synthesize the noise portion of the rolling noise, the same techniques as used for the wind noise (section 2.2) can be applied to the rolling noise data after the deterministic components have been removed with a Vold-Kalman filter (section 3.1.3).

Contrary to the wind noise synthesis, for the creation of the wavetable and the calculation of the Kautz coefficients, two different measurements have been used. The road measurements are used to generate the wavetable, as they represent a more natural sound. The Kautz coefficients have been calculated using the dynamometer measurements, as they span over a more extensive velocity range. Other than that, the synthesis principle works analog to wind noise synthesis.

3.2.2 Tire Order Synthesis

While the harmonic content of the engine noise is a substantial part of the driving noise, the synthesis of the tire harmonics is rather irrelevant. Especially when looking at the real road measurements in figure 3.3, the sharp orders that have been visible at the dynamometer measurements are smearing over a broader frequency range.

Therefore, a synthesis of the tire orders with similar methods used to synthesize the engine orders seems unreasonable, and a method to recreate the tire orders computationally cheap by using a frequency modulated sine wave has been tested.
As the addition of the modulated sine wave did not bring any perceptive advantages over the synthesized variant without it, the decision has been made that the explicit synthesis of the tire orders is not necessary, and the noise synthesis using Kautz filters is sufficient for realistic synthesis of the rolling noise.

3.2.3 Road Bumps

As the road bumps play an essential role in the sound of rough road surfaces, an experimental way to create randomness has been tested. Kronecker-Delta pulses have been created at random points in time. Every created impulse was followed by a second one: the first impulse representing the front wheel hitting a road bump, the second impulse representing the back wheel hitting the same road bump.

The time difference between those impulses is calculated as

$$\Delta t = \frac{d}{v} \tag{3.14}$$

where d is the distance between the front and back wheel and v is the current velocity of the vehicle.

The impulses then have been filtered with impulse responses that have been created from TPA[4] measurements, representing the transfer function from the front and back wheel axle to the driver's ear, respectively.

Unfortunately, the results of applying this technique have not been satisfactory. Compared to real road noise samples with road bumps, the synthetically added bumps sound too monotonous and worsen the quality of the synthesized rolling noises. A randomization of the bump amplitude and impulse response phase could improve the realism of the road bumps. As the realism of the total road noise was not suffering significantly from the absence of the road bumps, the synthesis of the road bumps has been disabled in this model and remains a topic for further investigations.

An alternative method to include the road bumps in the synthesized sounds is described in the following paragraph. This method was not implemented or tested but can be seen as an inspiration for future improvements:

The current implementation creates a new wavetable by recreating the spectrum of the real road measurements using a noise process and a 1/12-octave filerbank. This way, all non-linearities (including the road bumps) have been removed. However, using the road measurements directly as a looped basis sound would result in prominent areas (such as loud bumps) being repeated over and over, which could become recognizable to the listener's ear after some time. Therefore, a granular synthesis approach could be appropriate, where several short sound snippets are cut out of the real rolling noise sample and are randomly played back. This approach has already been used to simulate traffic noise (see [MJ13]) and is generally a commonly used method for sound texture synthesis (see [FH09]).

[4]Transfer Path Analysis

Chapter 4

Playback - Ambisonics

This chapter gives a small, compact overview of a spatial sound format called Ambisonics, which will be used as the playback method in the static driving simulator. In general, Ambisonics is splitting the sound-field into its spherical harmonics. This brings several advantages over traditional playback methods, one of them being the independence on the speaker setup - the exact loudspeaker configuration is only necessary for the decoder. The encoded signal is independent of it. This fact makes it the perfect format to work with varying loudspeaker arrangements, as found in static driving simulators.

The Ambisonics order N determines the order of the spherical harmonics used to represent the signal and therefore influences the spatial resolution. A higher Ambisonics order also widens the sweet spot that normally lies right in the middle of the speaker setup [ZF19]. The number of required channels K for a certain order N is $K = (N + 1)^2$ for 3D Ambisonics and $2N + 1$ for 2D Ambisonics.

Another advantage of Ambisonics lies in the easy manipulation of the sound field (panning, rotating, etc.), using simple mathematical operations. Many VST plugins executing those operations have already been programmed and can effortlessly be integrated into the Simulink model for driving sound synthesis. The IEM-Plugin Suite[1] and the ambiX - Ambisonics plugin suite by Matthias Kronlachner[2] for example are freely available to download.

[1] https://plugins.iem.at/

[2] http://www.matthiaskronlachner.com/?p=2015

Using Ambisonics as playback format also opens up the possibility to render the spatial signals to headphones, using the "Binaural Decoder" plugin from the IEM-Plugin Suite. This plugin is using pre-processed HRTFs[3] to directly convert the Ambisonics signals to binaural signals [SZH18].

4.1 Encoding

First-order Ambisonics can be recorded directly using different microphone array setups, using a least three microphones for 3D and two microphones for 2D. While the recording of higher-order Ambisonics signals can be a rather complicated process, the encoding of mono sources into the Ambisonics domain of arbitrary order is a straightforward process. Single sound sources can be encoded into specific directions by using a simple vector multiplication of the signal with the respective spherical/circular harmonics, evaluated at the respective panning angle [ZF19]. The following equation describes the encoding process for 2D Ambisonics

$$\chi_{\mathbf{N}} = \mathbf{y}_{\mathbf{N}}(\varphi_s)s \tag{4.1}$$

where $\chi_{\mathbf{N}}$ represents the encoded result, N represents the circular harmonic index, $\mathbf{y}_{\mathbf{N}}(\varphi_s) = [\Phi_{-N}(\varphi_s), ..., \Phi_N(\varphi_s)]^T$ represents the circular harmonics, evaluated at the panning angle φ_s and s represents the signal.

The circular harmonics can be calculated by

$$\Phi_m = \sqrt{\frac{1}{2\pi}} \begin{cases} \sqrt{2}sin(|m|\varphi) & m < 0 \\ 1 & m = 0 \\ \sqrt{2}cos(|m|\varphi) & m > 0 \end{cases} \tag{4.2}$$

4.2 Decoding

To obtain the speaker signals from the encoded Ambisonics signal, an Ambisonics decoder is necessary. The two most simple Ambisonics decoding strategies are the sampling decoder and the mode-matching decoder.

[3]Head Related Transfer Functions

4.2.1 Sampling Decoder

Using the sampling decoder approach for 2D Ambisonics, the loudspeaker signals $\mathbf{x}$ are calculated by [ZF19]

$$\mathbf{x} = \boldsymbol{D}\text{diag}\{\mathbf{a}_N\}\chi_{\mathbf{N}} \tag{4.3}$$

where $\boldsymbol{D}$ is a matrix, containing the circular harmonics, calculated at each loudspeaker angle φ, normalized by the factor $\sqrt{\frac{2\pi}{L}}$ with L being the number of loudspeakers:

$$\boldsymbol{D} = \sqrt{\frac{2\pi}{L}}\mathbf{Y}_N^T = \sqrt{\frac{2\pi}{L}}\left[\mathbf{y}_N(\varphi_1), ..., \mathbf{y}_N(\varphi_L)\right]^T \tag{4.4}$$

a_m is a side-lobe suppressing weighting factor

$$a_m = cos\left(\frac{\pi m}{2(N+1)}\right) \tag{4.5}$$

4.2.2 Mode Matching Decoder

The mode matching decoder uses a different principle compared to the sampling decoder, which requires the calculation of the pseudoinverse of $\mathbf{Y}_N$.

$$\mathbf{Y}_N = \left[\mathbf{y}_N(\varphi_1), ..., \mathbf{y}_N(\varphi_L)\right] \tag{4.6}$$

$$\boldsymbol{D} = \sqrt{\frac{L}{2\pi}}\mathbf{Y}_N^T(\mathbf{Y}_N\mathbf{Y}_N^T)^{-1} \tag{4.7}$$

One of the requirements for the mode matching decoder is the existence of the term $(\mathbf{Y}_N\mathbf{Y}_N^T)^{-1}$, which can be achieved by having at least as many loudspeakers as harmonics($L > 2N + 1$ for 2D-Ambisonics). Still, for irregular loudspeaker designs, the inversion of the matrix boosts the loudness for positions sparsely equipped with loudspeakers which could lead to a numerically ill-conditioned result [ZF19].

It is worth noting that both sampling decoder and mode matching decoder are not ideal decoder strategies for irregular loudspeaker arrangements [Fra14]. Other decoder strategies such as AllRAD decoding deliver better results for those cases [ZFP13]. Still, the simpler decoding techniques have been used for this thesis, mainly due to time reasons.

Chapter 5

Implementation in Matlab/Simulink

The implementation of the previously described algorithms has been conducted in Matlab and Simulink (Version R2018b). Matlab is used to calculate the required model parameters by analyzing real recordings and measurements.

Two Matlab scripts are responsible for this calculation: `init_windnoise` is responsible for the calculation of the required wind noise parameters, whereas the function `init_rollingnoise` is responsible for the rolling noise parameters. This way, it is possible to combine different wind noise and rolling noise characteristics of different car models without redundant recalculations. The parametrization is an off-line process and the resulting parameters are stored in a .mat-file called `windnoise_data.mat` and `rollingnoise_data.mat`, respectively. The exact functionality of the parametrization is documented using comments directly in the code. An overview of the approach is given in section 5.1.

The real-time synthesis of the driving noise is executed in Simulink. There are three different main blocks responsible for the engine noise, the wind noise, and the rolling noise, respectively, followed by an output block, responsible for the Ambisonics encoding of the sources and decoding to the loudspeakers. The model sections and blocks are outlined more thoroughly in section 5.2.

5.1 Matlab - Parametrization

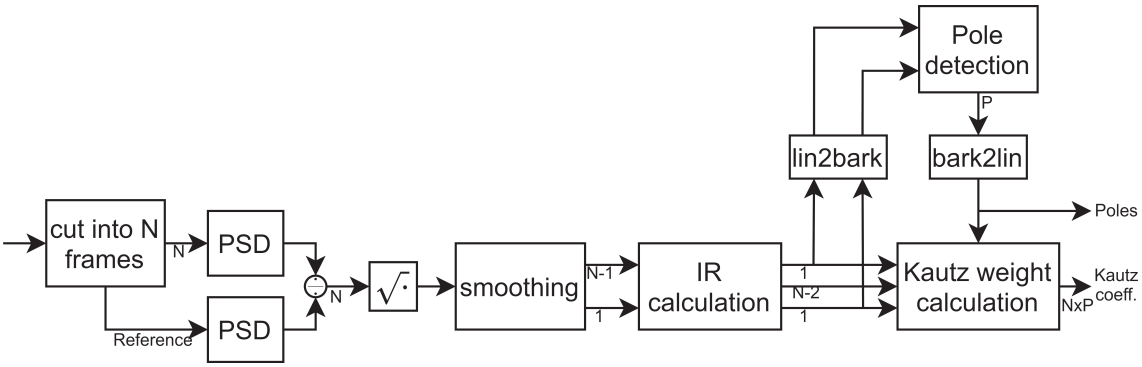


Figure 5.1: Simplified schematic of the pole and Kautz coefficient calculation

5.1.1 Parametrization of the wind noise

The wind noise parameters can be calculated by running the Matlab script `init_wind_noise.mat`

The following parameters at the top of the script have to be adjusted accordingly:

- x_out: The wind noise run-up data for the driver's left ear
- x_in: The wind noise run-up data for the driver's right ear
- vmin: The starting velocity of the provided run-up [km/h]
- vmax: The end velocity of the provided run-up [km/h]
- fs_ file: The sampling rate of the provided run-up [Hz]
- v_ ref: The reference velocity that is used for the creation of the wavetable [km/h]
- x_0: Wind noise data at an inflow angle of 0 deg
- x_p10: Wind noise data at an inflow angle of +10 deg
- x_m10: Wind noise data at an inflow angle of −10 deg
- x_sw: The previous data concatenated into one matrix

- angle_sw: The corresponding angles to the previous data
- fs: The sampling rate of the model
- fc: Everything lower than fc in the wind noise will be played back in mono
- S: Before creating the wave-table and calculating the Kautz coefficients, the PSDs get smoothed by a 1/S-octave-filter. $S = 0$ disables the smoothing
- len: the length of the created wave-table, which gets looped

The script then resamples the data to the sampling rate of the model, partitions the wind channel run-up into 10 km/h snippets, and selects the reference data for the creation of the wavetable. Subsequently, the function `create_noise_smooth` is called.

`create_noise_smooth`

`[n_l,n_r] = create_noise_smooth(x_l,x_r,fs,len,fc,S)`
creates two independent wind noise wavetables by "printing" the frequency response of the reference data `x_ref_l` and `x_ref_r` onto a random pink noise process.

Output arguments:

- n_l: wavetable that mimics the spectrum of x_l
- n_r: wavetable that mimics the spectrum of x_r

Input arguments:

- x_l: signal, whose spectrum should be mimicked by n_l
- x_r: signal, whose spectrum should be mimicked by n_r
- fs: sampling frequency
- len: length of the wavetables
- S: determines the width of the spectral smoothing filter

First, a partial-octave filterbank is designed using the `octaveFilter` function from the audio systems toolbox. Subsequently, all the filters from the filterbank are transformed into the frequency domain. Then the PSD of the signal is calculated by using Matlab's `pwelch` function and subsequently smoothed in the logarithmic domain (to avoid sharp frequency spikes). The corresponding filterbank weights are then calculated by multiplying the PSD of the signal with each of the magnitude squared transfer functions of the filterbank filters.

The wavetables are then calculated by creating pink noise at a pre-determined length, filtering it with the same filterbank, multiplying the bands with the previously calculated weights, and summing them up. To obtain a coherent sound for low frequencies, the low-frequency bands of wavetable 2 get replaced by the time signal found in wavetable 1. The threshold for this procedure is determined by the variable `fc`. The frequency band closest to `fc` is searched, and for all bands equal or below the determined band, the signal is modified to be mono.

As the wavetables are repeated periodically, the transition between the repetitions should not be noticeable. Audible clicks typically occur due to jumps in the time signal. Therefore, the created wavetables have been cut so that the last sample differs from the first sample only minorly. Personal evaluation has shown that the transitions between the repetitions are not audible.

After the creation of the wave-tables, the poles and coefficients for the Kautz filter are calculated by the function

`create_kautz_filter`

This function can be called with 5 arguments as
`[p,k_coeff] = create_kautz_filter(x_1,fs,x_2,S,plot_flag)`
to calculate the parameters for a one-dimensional interpolation table or with 7 arguments as
`[p,k_coeff] = create_kautz_filter(x_1,fs,x_2,S,plot_flag,x_sw,angle_sw)`
to calculate the parameters for a two-dimensional interpolation table. Usually, the function call with 5 input arguments is used for parametrizing the rolling noise (Kautz coefficients depending on the velocity). In contrast, the function call with 7 input arguments is used for calculating the wind noise parameters (Kautz coefficients depending on the velocity and the inflow angle).

Figure 5.1 shows a schematic of the approach.

Output arguments:

- p: vector with the poles of the Kautz filter
- k_coeff: matrix of the Kautz coefficients for different velocities

Input arguments:

- x_1: signal, whose spectrum is the starting point for the Kautz filter
- x_2: signal/s, whose spectrum should be modeled by the Kautz filter
- fs: sampling frequency
- S: determines the width of the spectral smoothing filter
- plot_flag: if plot_flag = 1, the spectrum of the Kautz approximation and the original spectrum is plotted for different velocity regions
- x_sw: wind noise data at a fixed velocity for different inflow angles as a matrix (columns contain data at a fixed inflow angle)
- angle_sw: the corresponding angles to the data in x_sw

First, the transfer functions to be modeled are calculated by

$$|H_v(e^{j\theta})| = \sqrt{\frac{S_v(\omega)}{S_{v,ref(\omega)}}} \tag{5.1}$$

where $S_{v,ref(\omega)}$ is the reference PSD and $S_v(\omega)$ is the target PSD. Both are obtained by using the Matlab function `pwelch` with the corresponding audio snippets.

The obtained transfer functions $|H_v(e^{j\theta})|$ then get smoothed by a 1/N-octave smoothing filter and get frequency warped according to the Bark scale (section 2.2.2.2). This is done by using the Matlab functions written by Jonathan S. Abel[1].

Subsequently, an impulse response corresponding to the smoothed transfer function is calculated. This is done by calculating the minimum-phase reconstruction of the signal. This step is repeated for all transfer functions to be modeled by the Kautz filter.

[1]Downloadable at https://ccrma.stanford.edu/ jos/bbt/ (last accessed on 1st December 2020)

In the next step, the optimal pole positions have to be determined. This task is done by using the BU-Method described in 2.2.2.1. To achieve this, the Matlab script `braun1`[2] by Tuomas Paatero is used. To compensate for eventual deviations in resonance frequencies of the wind noise for different velocity region, the algorithm is run twice: once for a transfer function $|H_v(e^{j\theta})|$ where $v << v, ref$ is small and once for a transfer function where $v >> v, ref$. Those two pole sets are subsequently merged into one pole set.

The previous frequency warping allows for a pole placement focused on low frequencies and therefore corresponds better to the human auditory system. Another given option is the usage of the variable `downsampling_factor` which downsamples the impulse response from f_s to $f_{s,ds}$, so frequencies higher than $\frac{f_{s,ds}}{2}$ are not taken into consideration for the pole positioning anymore. This feature can be useful when used on only one of the two transfer functions fed into the BU-method algorithm, as redundant poles in the high-frequency region are removed. Furthermore, the high-frequency information for low velocities and therefore also the transfer function for $v << v, ref$ is not useful due to the noise floor caused by the measurement of low velocities.

If the approximation of the transfer function is not satisfactory, additional poles can be added by manually entering the corresponding frequencies and pole radii into the variables fp and r, respectively.

5.1.2 Parametrization of the rolling noise

The rolling noise is parametrized similarly to the wind noise. The only difference here is that the wavetable is generated from a different measurement (real street measurement) than the Kautz filter parameters (acoustic dynamometer measurement).

As tire orders may be pretty dominant in dynamometer measurements, the script gives to option to remove them by using a Vold-Kalman filter (see section 3.1.3). By setting the flag `ro_flag` to 1, this calculation will be performed, while by setting `ro_flag` to 0, a previously filtered file can be loaded (to save computation time).

[2]Downloadable at http://legacy.spa.aalto.fi/software/kautz/kautz.htm(last accessed on 12th June 2021)

If `ro_flag` = 1, the following parameters have to be adapted:

- x_roll: the time signal of the dynamometer measurement
- f_0: the fundamental frequency of the orders that should be filtered out
- p: the order and therefore steepness of the Vold-Kalman filter
- bw: the bandwidth of the Vold-Kalman filter
- save_path: the path where the calculated data is saved so that it can be loaded again at a later time

5.2 Simulink Model

This section further describes the blocks and sections of the model.

Initialization

The initialization parameters for the Simulink model are set in the file `init_ess`. It contains model parameters, such as the sampling frequency `fs`, the buffer size `buf_sz`, the loudspeaker angles `ls_pos` and calculates the corresponding Ambisonics decoder matrix. Furthermore, it loads the previously calculated Kautz coefficients, wavetables and headphone impulse responses into the workspace.

1) Input Stage

The input stage is the source of the variable model input parameters. It is possible to switch the input between real-time data coming from the UDP interface (switch up) and sample data that has been recorded from a test drive in the driving simulator (switch down). The input data contains

- engine rpm [1/min]
- throttle [%]
- velocity [m/s]
- yaw angle [rad]
- surface friction - not used in this model
- clutch [boolean]

The Saturation-Block in the velocity path is used to limit the velocity to only positive values. In certain extreme driving situations, the driving simulator led to a negative velocity that crashed the model.
The Rate Transition-Blocks adapt the sampling rate of the input data (0.01 s) to the sampling rate of the model.
The Subtraction-Block in the yaw angle path was necessary to correct an offset value of the transmitted yaw angle data.

2) Engine Noise

The engine noise is generated in the PSOLA[3] block. PSOLA is a synthesis technique that was originally developed to alter the pitch and length of speech signals [CS86]. The engine noise synthesis for this driving noise simulator was already developed prior to this thesis in [Sta18].

Right below the engine noise block, a speed-dependent gain can be found. It is used to adjust down the engine noise volume at higher velocities. This adjustment was conducted totally subjective and is not based on any measurements as no proper data to analyze the volume behavior of the engine noise at different velocities has been available.
Similarly, the Sporty-Switch can be utilized to increase the proportion of the engine noise in the total noise composition.

3) Rolling Noise

Figure 5.2 is showing the substructure of the Rolling Noise-Block. The synthesis of the rolling noise consists of three paths.

The upper path contains an experimental method to create random road impacts that should make the rolling noise sound more natural. Unfortunately, the result did not sound natural and did not increase the perceived realism of the rolling noise. Therefore, this path has been commented out, but not deleted to allow for further development of those road impacts (see section 3.2.3).

[3]Pitch Synchronous Overlap and Add

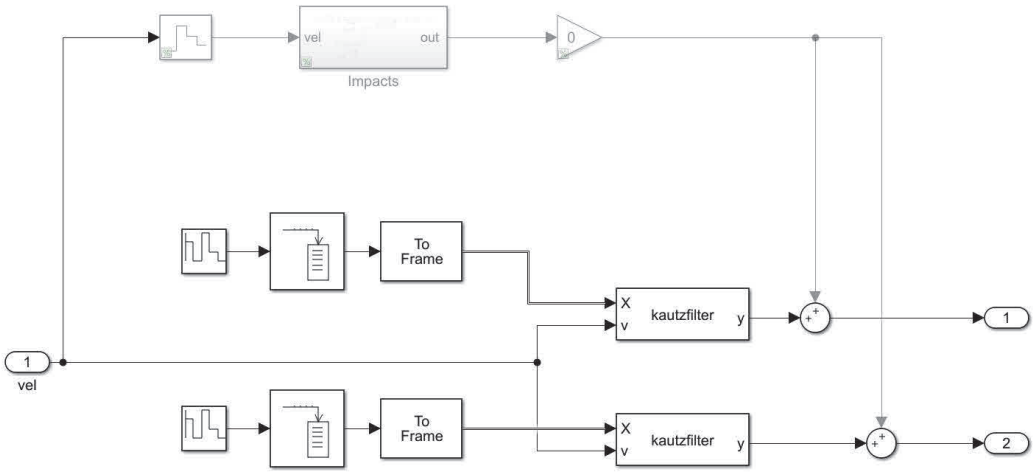


Figure 5.2: Rolling Noise Block of the Simulink Model

The two lower paths are responsible for the noise characteristics of the rolling noise. Both Repeating Stair Sequence-Blocks contain colored noise samples (section 5.1.2) that are uncorrelated to each other. This way, it is possible to place two independent rolling noise sources in the sound field to achieve a wider sound perception. The Kautz Filter blocks are responsible for the velocity-dependent coloration of the noise. They contain the Kautz filter structure (as described in 2.2.2), implemented as a Matlab System Block. This noise component makes up the most prominent and most important part of the total rolling noise.

4) Wind Noise

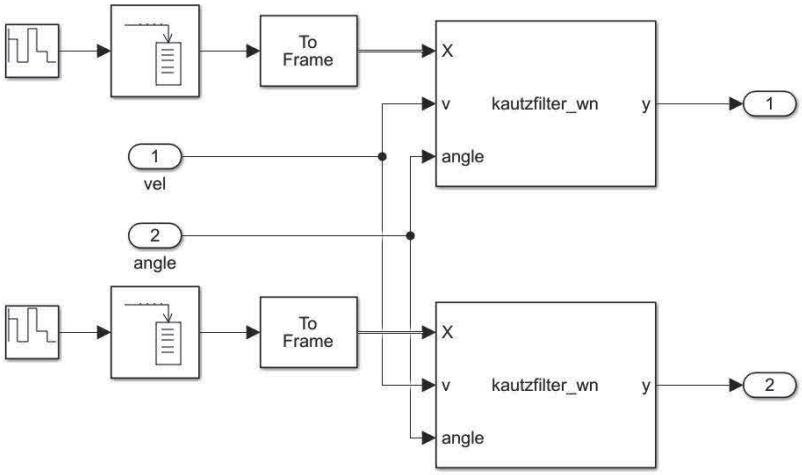


Figure 5.3: Wind Noise Block of the Simulink Model

Similar to the rolling noise synthesis, the wind noise block (figure 5.3) contains two separate synthesis paths to be able to spread out the wind noise in the sound field. Again, the velocity-dependent spectral shaping and level adjustment is conducted by two Kautz filter blocks. Contrary to the method used with the rolling noise, the interpolation table for the Kautz coefficients is two-dimensional. It also contains the yaw angle as a second parameter besides the velocity.
A possibility to add realism to the model was implemented in the `Random deviation` block. It adds a noise filtered noise signal to the input velocity and yaw angle for the wind noise synthesis, which creates an impression of increased fluctuation of the wind noise (see section 2.2.4).

5) Ambisonics Decoding

The Ambisonics decoders are implemented as Matlab System Blocks. As only 2D-Ambisonics is required here, the decoders have been programmed explicitly instead of using existing 3D solutions to decrease the computational effort.
The panning angles of the sound sources are supplied to the decoders as block inputs. This way, those panning angles could be altered during the runtime of the simulation (for example, to slightly alter the position of the wind noise sources over time, to achieve even a higher realism[4]).
The size of the input angles is [2x1] for both wind noise and rolling noise (two separate sources to place) and [1x1] for the engine noise.

6) Widening

As in a real-world situation, the driving sounds do not originate from discrete directions but rather surround the driver, an approach to widen the point-like audio sources has been integrated into the Simulink system.
This was done by using an Ambisonics widening VST-Plugin, which is part of an Ambisonics plugin suite developed by Matthias Kronlachner[5].
This plugin is using a frequency-dependent panning angle to achieve the perception of increased width ([ZFKC14]). The intensity of the widening effect can be adjusted in the Matlab Function code itself by tweaking `obj.sp.ModDepth` in the `stepImpl` function.

[4]This method was not tested and is only based on an assumption
[5]Free download at http://www.matthiaskronlachner.com/ (Accessed on Nov, 5th 2020)

In the default Simulink model, the Widening-Block is commented out, as, in the driving simulator, there was subjectively no need for a widening of the sound sources. Nevertheless, widening might be helpful in other situations.

7) Sound Selector

For the sake of testing, a switch has been implemented to be able to quickly change single driving noise components or the whole sound composition.

8) Output Stage

For the output stage, there is the possibility to switch between headphones- or loudspeaker output. It is recommended to comment out the method that is currently not used to save computational power.

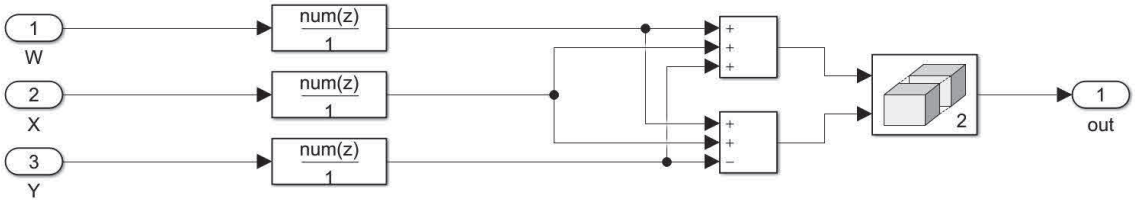


Figure 5.4: Headphone Decoder Block of the Simulink Model

When selecting the headphone decoder path, the Ambisonics signals get converted to binaural headphone signals by using a set of impulse responses that was taken from the BinauralDecoder of the IEM Plugin Suite[6]. The binaural decoder converts the Ambisonics signals directly to the binaural headphone by using pre-processed HRTFs[7] as in [SZH18].

The loudspeaker decoder path contains the Ambisonics decoder and an equalizer to linearise the speaker responses. A simple mode matching decoder (like described in section 4.2.2) is used here. The equalizer is currently adjusted for the driving simulator at the Virtual Vehicle Graz Inffeldgasse.

The output signal is then going through a highpass filter to roll off inaudible frequencies below 40 Hz and subsequently is passed to the soundcard outputs via the `Audio Device Writer` block.

[6] plugins.iem.at
[7] Head-Related Transfer Function

Chapter 6

Listening Test

Several papers have shown the impact of the auditory component on velocity estimation [HP08] [Eva70] [MJ11] [DRKM+12]. One of the most crucial factors influenced by the auditory component is the perceived velocity of the vehicle. The previously mentioned papers have proven that without the auditory component or with an attenuation of the driving sound volume the vehicle velocity is generally underestimated. In real-world-situations this could lead to a tendency to drive faster and therefore to a more dangerous driving behavior (see [AVS06]).

While most studies focus on the impact of the driving sounds in total, the influence of different driving sound components on the velocity estimation has not been investigated yet.

Two listening tests have been independently conducted in the course of this thesis. Both listening tests took place on different dates, in different places, with (partially) different test subjects, and had different goals. The following points should be evaluated with the corresponding listening tests:

- Listening Test 1 at Magna Steyr
  - Phase 1) Quality evaluation of the synthesized sounds
  - Phase 2) Ability to discern between different velocities based on wind noise only
- Listening Test 2 at VIF[1] Graz
  - Importance of different driving noise components for the velocity perception
  - Subjective evaluation of the driving sound generator in a driving simulator by the test subjects (questionnaire about their acoustic impressions of the driving simulator)

[1] Virtuelles Fahrzeug/Virtual Vehicle

The execution of Listening Test 1 was necessary before Listening Test 2, as the quality of the synthesized sounds had to be validated prior to their usage in Listening Test 2. Furthermore, the ability to discern between different velocities based on wind noise only has been investigated in Phase 2 of Listening Test 1 to ensure that such a differentiation between driving velocities based on a single driving noise component will also be possible in Listening Test 2.

Parts of the listening tests have already been published in [KBSR20].

6.1 Listening Test 1

The first listening test has been executed at a quiet meeting room at the engineering center of Magna Steyr AG & Co KG in Graz. Eight male employees of Magna Steyr participated in the listening test. The stimuli were presented to the test subjects via AKG K-271 MKII closed-back headphones (to block out background noise) and an M-Audio MobilePre audio interface.

6.1.1 Calibration

The data set used for the listening test contains the sound pressure levels of the respective wind- and rolling noise samples. Therefore, the unit of the sample data is given in pascal. To be able to write this data to an audio file (in this case, a wave file), it has to be ensured that the maximum and minimum values of the data must not exceed 1 and -1, respectively. This is achieved by scaling the data by an arbitrary factor to ensure this condition is valid.

To calibrate the system, a 1 kHz test tone has been generated. The auditory threshold of humans lies around 20 μPa at 1 kHz, so the amplitude of the test tone has to be set at

$$A = p_0 \cdot \sqrt{2} \cdot a \tag{6.1}$$

where $p_0 = 20 \cdot 10^{-6}$ and a is the scaling factor that has been previously applied to the data to ensure the values lie between -1 and 1. The factor $\sqrt{2}$ is used for the conversion from the RMS[2] value to the peak value.

[2]Root Mean Square

Consecutively, the system has been calibrated to the hearing threshold. This was achieved by adjusting the output volume of the headphone amplifier so that the 1 kHz test tone can be just heard. This calibration is necessary for the second phase of the listening test and has been conducted only once previous to the beginning of the listening test by the investigator himself.

The calibration is only relevant for phase 2 of the listening test, as the sound samples for phase 1 have been volume normalized according to their Zwicker loudness.

6.1.2 Phase 1 - Quality Evaluation

Recorded driving sounds have been compared to the synthesized sounds both for rolling noise and wind noise to examine the quality of the synthesized driving sounds.

This is achieved by playing back different monaural sound excerpts including real wind noise/rolling noise recordings and their synthesized counterparts to the test subjects. The real wind noise excerpts were taken from wind-channel recordings of different car models with the microphone placed at the driver's ear position. For the rolling noise, excerpts from real street measurements of one particular car model that have been recorded on different street surfaces (smooth asphalt, rough asphalt, and cobblestone) have been used. The rolling noise excerpts have been cut out of coast-downs from 65 km/h down to 55 km/h. In this velocity region, the wind noise proportion is negligible compared to the rolling noise proportion. To effectively compare wind- to rolling noise, the wind noise excerpts that have been used also represent the velocity region around 60 km/h. The synthesized excerpts have been created using the Simulink model. For the creation of the synthesized wind noise excerpts, the artificial turbulence generation has been switched off. All the sounds have been volume-adjusted according to their Zwicker loudness ([Zwi61],[ZFW$^+$91]). Without this adjustment, wind noise and rolling noise would be easily discernible by their volume differences alone.

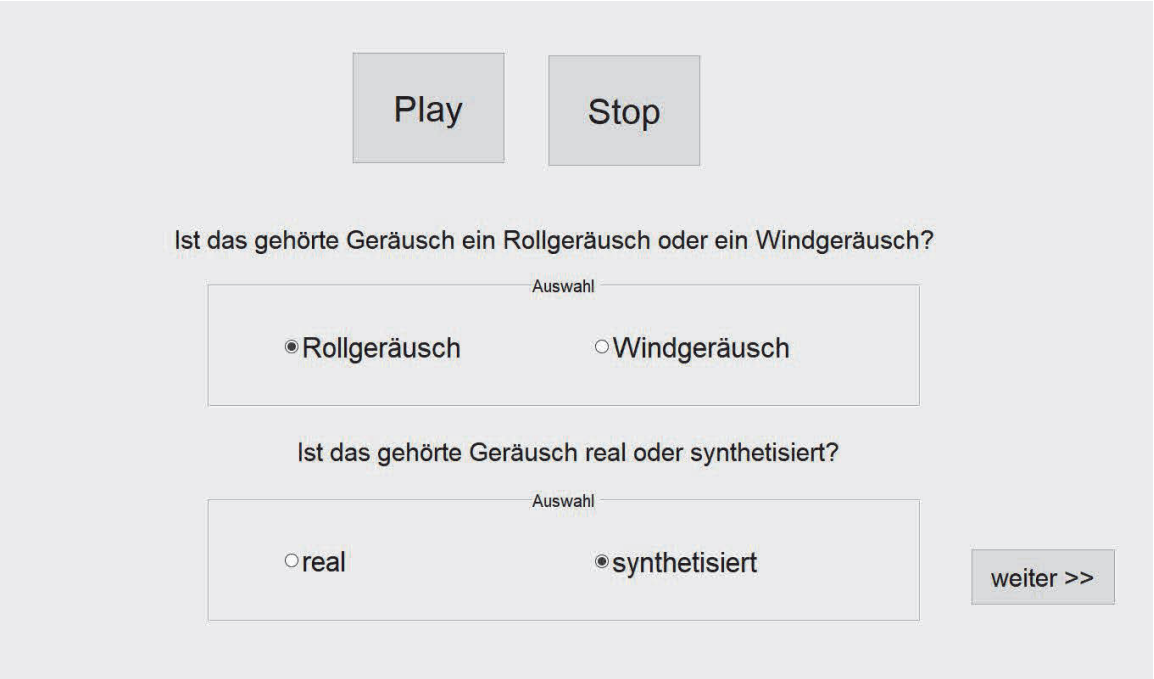


Figure 6.1: User interface for the first phase of the first listening test

The listening test started with a learning phase, where the test subjects listened to loudness-adjusted sound excerpts of real wind/rolling noise. Those sound excerpts have been chosen to be similar to the excerpts presented later on but are not identical to preclude a direct recognition of the sounds in the following phase.
After the learning phase, several stimuli without any further information have been presented to the test subjects. For each of those stimuli, the test subjects had to decide whether the sound stimulus is perceived as wind noise or rolling noise and whether they think it is an actual recording or a synthesized sound. The presented stimuli contained three actual street measurements of the same car on different surfaces for the rolling noise, three wind channel recordings of different car models, and their synthesized counterparts. Every stimulus has been used twice to facilitate statistical analysis (i.e., reliability of subjects and refined quantification). That resulted in a total of 24 stimuli that needed to be classified by each test subject. The sequence of the stimuli has been randomized for each test subject, but with the restriction that the same excerpt will not appear twice in a row.

6.1.2.1 Evaluation per Test Subject

The first thing to evaluate is whether the test subjects have generally been able to discern between wind- and rolling noise. This can be investigated by looking at the number of correct answers. If we assume that the test subjects can distinguish wind- and rolling noises, the statistical framework for this scenario is as follows:

- Null hypothesis H_0: "test subjects cannot discern between wind and rolling noise - they are guessing" ($p = 0.5$ for giving a correct answer)

- Alternative hypothesis H_1: "test subjects can discern between wind and rolling noise" ($p \neq 0.5$)

Assuming the null hypothesis to be true, the listening test results can be described statistically by a binomial distribution. The corresponding PDF[3] and CDF[4] over the number of correct answers are shown in figure 6.2 and 6.3 exemplary for a maximum of 24 correct answers.

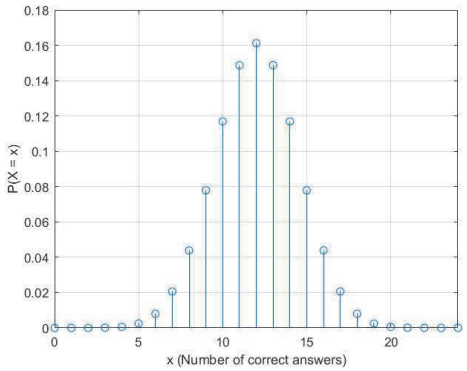


Figure 6.2: Binomial PDF for $p = 0.5$ and $N = 24$

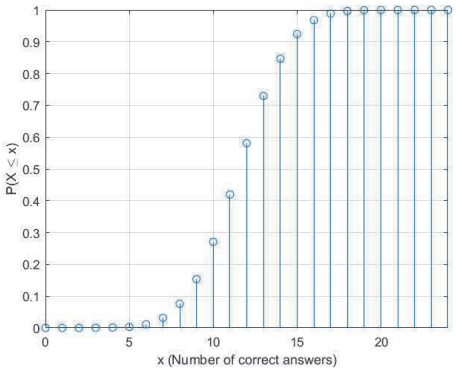


Figure 6.3: Binomial CDF for $p = 0.5$ and $N = 24$

The corresponding values of the CDF with the random variable X describing the number of correct answers are shown in table 6.1.

In this case, using a 5 % significance level and a two-tailed test, the null hypothesis can be rejected if the number of correct answers is smaller than seven or bigger than seventeen ($X < 7 \vee X > 17$). When evaluating wind noise and rolling noise samples

[3]Probability Density Function
[4]Cumulative Distribution Function

x	0	1	2	3	4
$P(X \leq x)$	0.00 %	0.00 %	0.00 %	0.01 %	0.08 %
x	5	6	7	8	9
$P(X \leq x)$	0.33 %	1.13 %	3.20 %	7.58 %	15.37 %
x	10	11	12	13	14
$P(X \leq x)$	27.06 %	41.94 %	58.06 %	72.94 %	84.63 %
x	15	16	17	18	19
$P(X \leq x)$	92.42 %	96.80 %	98.87 %	99.67 %	99.92 %
x	20	21	22	23	24
$P(X \leq x)$	99.99 %	100.00 %	100.00 %	100.00 %	100.00 %

Table 6.1: Binomial CDF values for $p = 0.5$ and $N = 24$

separately, the maximum number of correct answers goes down to 12, which brings the number of correct answers to reject the null hypothesis to $X < 3 \vee X > 9$.

The results of the evaluation per test subject regarding the differentiation between wind- and rolling noises are shown in figure 6.4. The red dashed line represents the confidence interval, whereas, for better readability, it has been set between the integer numbers (e.g., for $X < 7 \vee X > 17$ the red dashed line is drawn at $X = 6.5$ and $X = 17.5$). It can be observed that all test subjects except test subject 8 can discern wind- and rolling noises or are at least not guessing blindly.

As each sound was presented twice to the test subjects, it also makes sense to look at the consistency of the answers for the same sound. The number of consistent answers is shown in figure 6.5. A consistency of 24 means that the duplicate sound samples have been assigned to the same attributes all the time. A consistency of 0 means that the duplicate sound samples have always been assigned to different attributes. The consistency factor can provide information about the reliability of each test subject. Furthermore, a consistency factor around 50 % could hint that the test subjects are guessing the answers. It can be observed that the consistency is generally higher for the differentiation between wind- and rolling noise than for the differentiation between real and synthesized sounds. This supports the assumption that the test subjects can discern between wind- and rolling noises but are mostly guessing regarding the decision whether the heard sound is a real or a synthesized recording. The overall consistency for test subjects 6-8 is quite low, especially when looking at the differentiation between wind- and rolling noise.

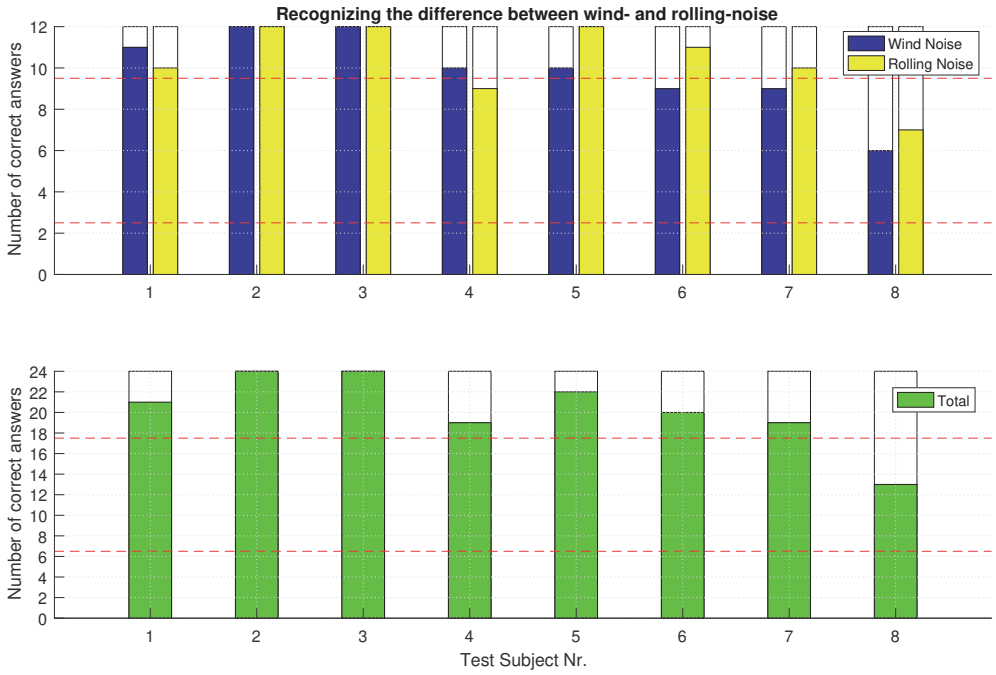


Figure 6.4: Number of correct answers regarding the differentiation between wind- and rolling noise per test subject. The red line represents the confidence interval.

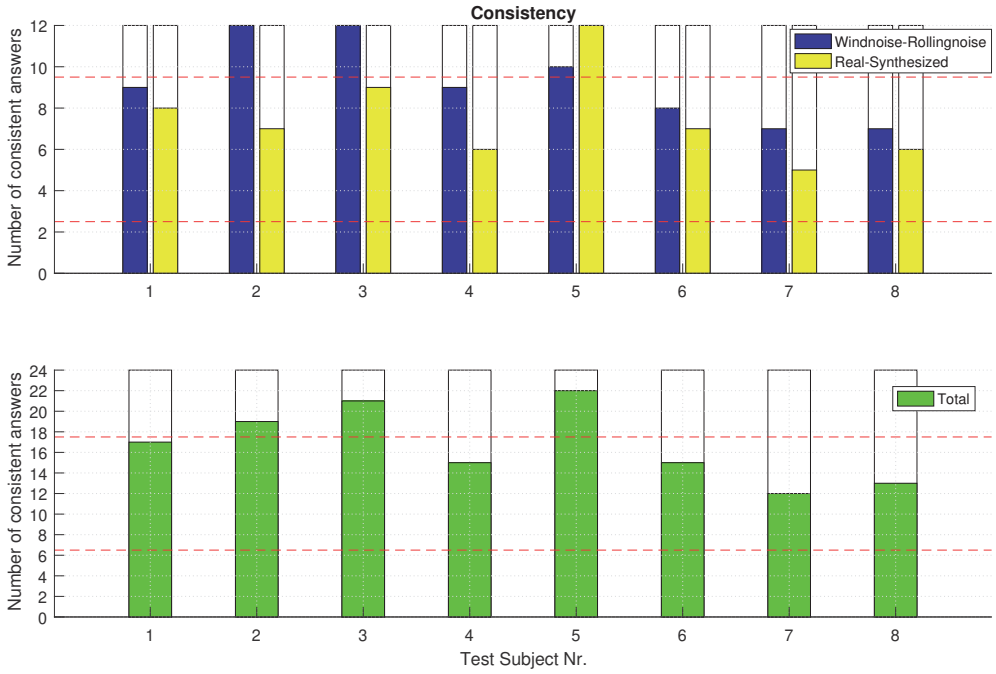


Figure 6.5: Consistency factor for the different test subjects

After evaluating the results and consistency per test subject, test subject 8 has been excluded from further evaluations. Besides not being able to discern between wind- and rolling noise and having a small consistency factor, test subject 8 claimed to wear a hearing aid due to hearing loss. Therefore, the results from subject 8 have not been representative.

6.1.2.2 Evaluation per Sound

After the exclusion of test subject 8, we can look at the number of correct answers per sound.

By evaluating the results in the decision between rolling-/wind noise for real sound excerpts, it is possible to detect whether the test subjects are able to discern the different noise sources and compare that result to the synthesized sounds. Ideally, there should be no difference in the results between real and synthesized noises.

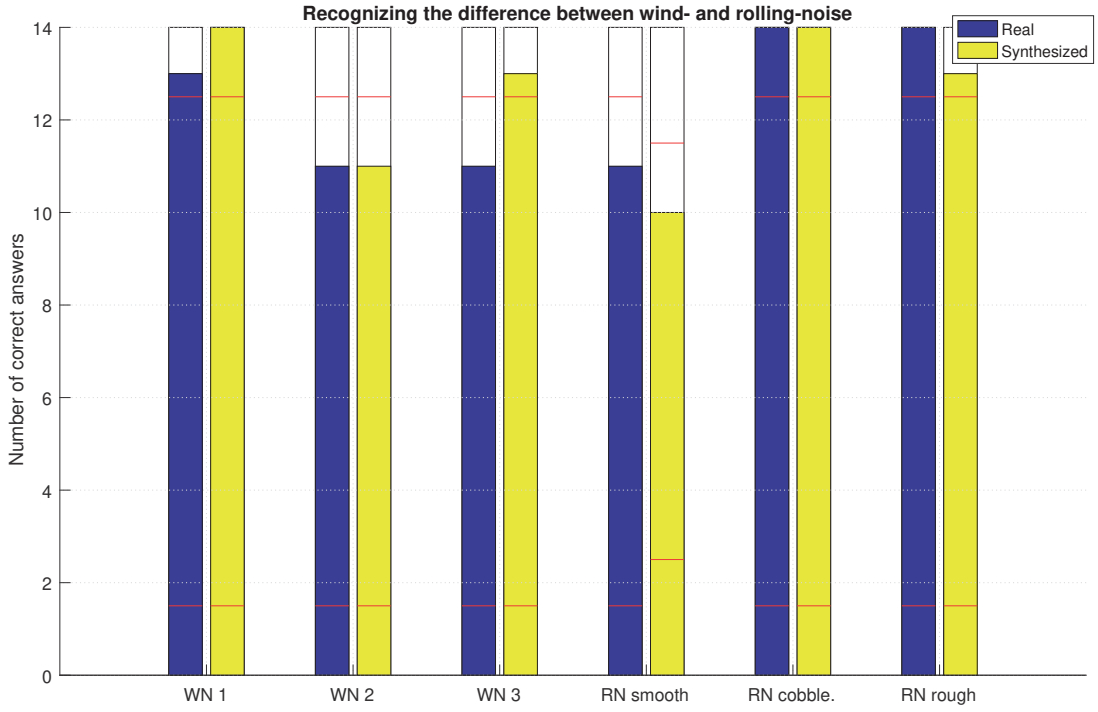


Figure 6.6: Number of correct answers regarding the differentiation between wind- and rolling noise for each sound. The red lines are showing the confidence intervals after applying the Bonferroni-Holm method.

Figure 6.6 is showing the number of correct answers for the decision between wind- and rolling noise per sound. Here, the Bonferroni-Holm method has been applied to adjust the significance levels, as multiple subsets of data have been compared directly [Hol79]. Without any correction, this would lead to an alpha error accumulation. For most sounds, the total number of correct answers is high enough to reject the null hypothesis at a 95 % confidence level and claim that the differentiation between wind- and rolling noise is possible for each sound.

However, for a few specific sounds, it has been harder to discern between wind- and rolling noise. "RN smooth", for example, is the rolling noise that sounds closest to the wind noise samples, as due to the smooth road, the characteristic road impacts are negligible.

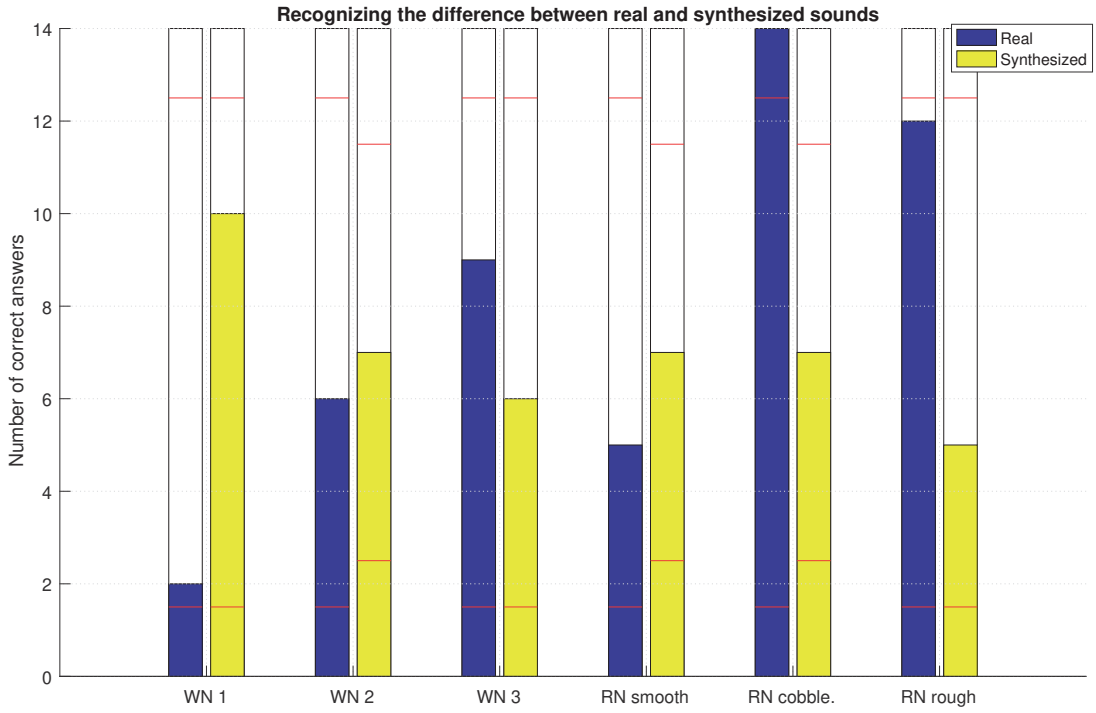


Figure 6.7: Number of correct answers regarding the differentiation between real and synthesized samples for each sound. The red lines are showing the confidence intervals after applying the Bonferroni-Holm method

Figure 6.7 is showing the results per sound for the decision between "real" and "synthesized", again corrected with the Bonferroni-Holm method. If those decisions show that the test subjects are guessing (for $\alpha = 5\,\%$, both-sided), this will guarantee that the quality of the synthesized noise is sufficiently given.

It can be seen that for the sound excerpts “RN cobble. - Real” and “RN rough - Real”, it was easier for the test subjects to correctly assign them as a “real” sound (all test subjects assigned the cobblestone sample correctly, for the rough asphalt sample, 12 out of 14 answers have been correct). This can be explained by the fact that especially these particular sound excerpts had audible random impulsive events originating from the road structure. As this phenomenon was not modeled for the synthesized sounds, a correct assignment of the synthesized sounds was more challenging. For the real smooth asphalt sound excerpt, the test subjects could not assign it correctly due to its rather noisy nature without any significant impulsive events. For nearly all wind noise excerpts, the null hypothesis “the test subjects are guessing” cannot be rejected. Therefore, it can be assumed that they have not been able to discern real and synthesized sounds. WN1 is an exception here - most test subjects have classified the real sound excerpt as synthesized.

6.1.2.3 Evaluation in total

Figure 6.8 is showing the results for the decision between “real” and “synthesized” of all test subjects and sound samples put together, but only taking those answers into consideration that have assigned the sound correctly to the “wind noise or rolling noise” attribute. It has already been proven in section 6.1.2.1 that the test subjects are generally able to discern between wind- and rolling noise. Therefore, it makes sense to set this as a requirement.

Again, 95 % confidence intervals for H_0 = “test subjects are guessing” have been calculated. Those are represented in figure 6.8 as red lines. Only for the real rolling noise samples, the null hypothesis that the test subjects are guessing can be rejected. For all other categories, H_0 cannot be rejected. Furthermore, for most sounds, the rate of correct answers is close to 50 %, which means that the test subjects are guessing. Therefore, the quality of the synthesized sounds is sufficiently given.

6.1.3 Phase 2 - Recognizing Velocity Differences

The goal in the second phase of the first listening test was to evaluate the ability of the test subjects to differentiate between different driving velocities by only listening to different wind noise samples. It has already been shown that the engine noise, with its prominent velocity-dependent tonal components, makes it easier for people to keep a steady speed [DRKM+12]. Also it has been shown that at the absence of driving sounds, the test subjects in [DRKM+12] underestimated their driving velocity to a greater extent than with the driving noise present.

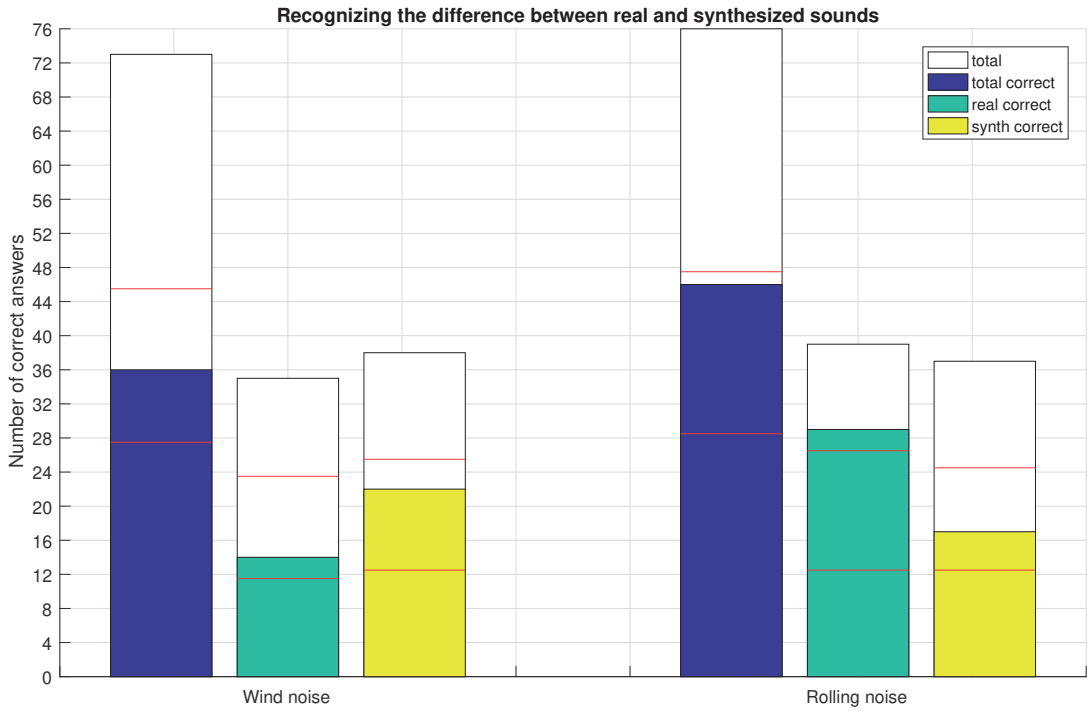


Figure 6.8: Number of correct answers regarding the differentiation between real and synthesized samples for all samples that have been correctly assigned to wind-/rolling noise. The red lines are showing the 95 % confidence intervals for H_0 = "test subjects are guessing".

The first question to be investigated is whether the presence of wind noise alone also alters the velocity perception. First, it has to be evaluated whether the test subjects are able to recognize different velocities by only listening to the wind noise. The test subjects had to order five different wind noise samples according to the perceived velocity to validate this assumption. As the authenticity of the synthesized wind noises has already been proven in the previous phase of the listening test, synthesized wind noise sounds have been used in this phase. This also brings the advantage that constant velocities can be used, which could not be obtained by cutting out the excerpts from the wind noise run-up data directly.

During Task 1, the test subjects had to sort five different wind noise snippets that significantly differ in velocity (50 km/h, 80 km/h, 100 km/h, 130 km/h, 160 km/h and 200 km/h). The snippets have been approximately 3 seconds long and were looped. The user interface for this phase of the listening test is shown in figure 6.9.

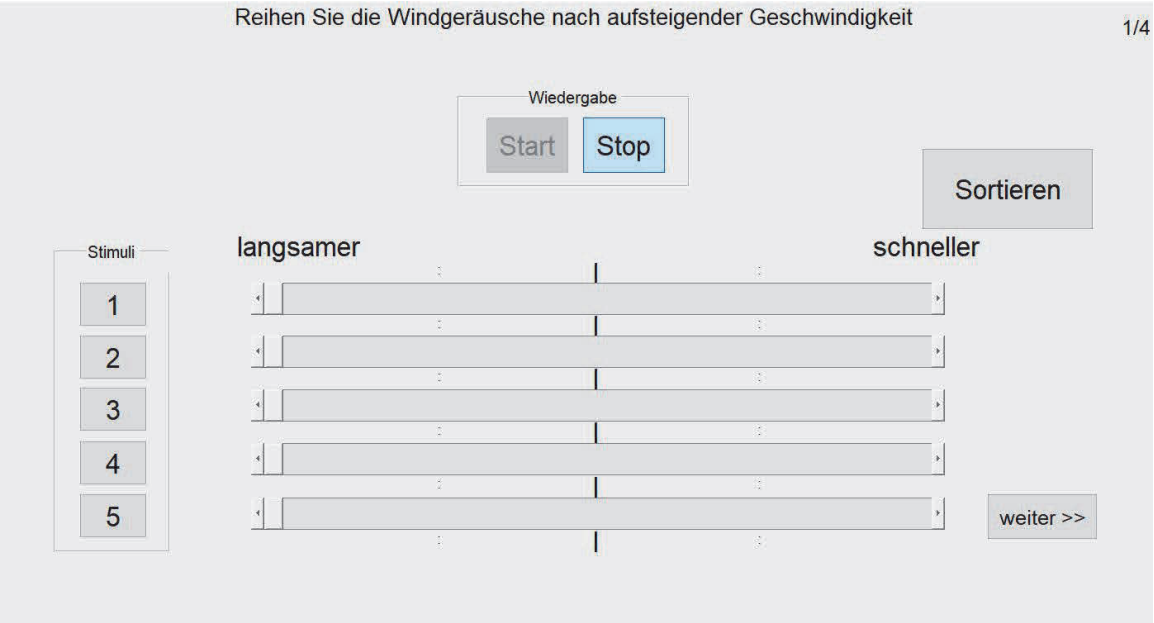


Figure 6.9: User interface for the second phase of the first listening test

The test subjects could listen to different sound snippets by clicking on the corresponding buttons in the "Stimuli" section. They then had to use the sliders to order them according to the perceived velocity. By clicking the "Sortieren" button, the test subjects could rearrange the stimuli by the corresponding slider position.

In the consecutive tasks, the ability to differentiate between minor velocity differences only by listening to the wind noise is investigated. The velocity differences between those stimuli are chosen so that the least difference is pretty close to imperceptible, and the largest difference should be easily discernible. Used differences are -5 km/h, -2 km/h, 0 km/h, +3 km/h and +10 km/h around a certain working point. This is done with the same velocity differences for 3 different working points: 50 km/h, 100 km/h and 160 km/h, respectively. That way, it can be examined whether the sensitivity for the acoustically perceived velocity difference stays the same over the whole velocity range or changes at the extrema. The arrangement of Task 2, Task 3, and Task 4, as well as the arrangement of the wind noise samples within each task was randomized for every test subject. Task 1 was always conducted first.

[km/h]	1	2	3	4	5
Task 1	50	80	100	130	180
Task 2	45	48	50	53	60
Task 3	95	98	100	103	110
Task 4	155	158	160	163	170

Table 6.2: Velocities of the synthesized wind noise sounds for phase 2 of the listening test 1

6.1.3.1 Results

The results of the second phase of the first listening test are shown in table 6.3 - 6.6.

		assigned position				
		1	2	3	4	5
correct position	1	7	0	0	0	0
	2	0	7	0	0	0
	3	0	0	7	0	0
	4	0	0	0	7	0
	5	0	0	0	0	7

Table 6.3: Results for Task 1

		assigned position				
		1	2	3	4	5
correct position	1	5	0	0	0	0
	2	0	5	0	0	0
	3	0	0	5	0	0
	4	0	0	0	5	0
	5	0	0	0	0	5

Table 6.4: Results for Task 2

		assigned position				
		1	2	3	4	5
correct position	1	6	0	0	0	0
	2	0	5	1	0	0
	3	0	1	5	0	0
	4	0	0	0	6	0
	5	0	0	0	0	6

Table 6.5: Results for Task 3

		assigned position				
		1	2	3	4	5
correct position	1	6	1	0	0	0
	2	1	5	1	0	0
	3	0	1	6	0	0
	4	0	0	0	7	0
	5	0	0	0	0	7

Table 6.6: Results for Task 4

Unfortunately, some of the user inputs have been faulty (several sound samples have not been assigned before confirming). Therefore, the total number of answers to evaluate differs from task to task. In total, there have been 7 test subjects. All of

them gave a valid answer for task 1 and task 4, while there have been only 5 valid answers for task 2 and 6 valid answers for task 3.

In the case of substantial velocity differences (Task 1), all subjects were able to rank the five stimuli in the correct order. The small sample group exhibited accurate ranking in the velocity range of 50 km/h (Task 2) for large (7 km/h) and small (2 and 3 km/h) speed differences. Within the velocity range of 100 km/h (Task 3), an error of roughly 10 % occurred for small speed differences. Moreover, around 160 km/h (Task 4), this error increases by a factor of two. Nevertheless, it was observed that even minor velocity differences can be distinguished.

6.2 Listening Test 2

The second part of the listening test took place at the Virtual Vehicle Graz and used the static driving simulator situated there. This phase of the listening test had two goals:

- The influence of different driving sound components on driving velocity perception
- The validation of the driving noise in a driving simulator environment

While the latter is evaluated by taking a survey with open questions at the end of the listening test, the first goal can be evaluated statistically.

Figure 6.10: The static driving simulator located at the Virtual Vehicle Graz/ *Image Source: https://www.v2c2.at/neue-trends-bei-fahrsimulatoren-das-out-of-the-loop-erlebnis/*

6.2.1 Driving Simulator

6.2.1.1 Connection

The computer that is responsible for the synthesis of the driving sounds is receiving the driving data from the driving simulator via CAN[5]. The Simulink model then synthesizes the driving sounds according to the received driving data and plays them back via an RME Fireface UC to the sound system of the driving simulator.

6.2.1.2 Playback Setup

For the playback of the driving sounds, the driving simulator uses the Logitec Surround Sound speaker Z906 System. It contains five quasi-full-range loudspeakers (L - left, R - right, C - center, BL - back left, BR - back right) and one subwoofer, which is positioned outside the car. Both left and right loudspeakers are installed into the bottom of the car doors while center- and both back loudspeakers are positioned at shoulder height. This makes the loudspeaker arrangement non-regular. Furthermore, the frequency response of the sound system has to be linearized at the driver's head position to obtain an authentic playback situation.

The angles and distances of the corresponding loudspeakers referenced to the driver's head position are shown in table 6.7.

Channel	Position	Angle[deg]
1	Front Left	-31.1
2	Front Right	56
3	Center	24.7
4	Back Left	-114.2
5	Back Right	123.5

Table 6.7: Angles and distances of the speakers to the driver's head

An ideal set-up for two-dimensional Ambisonics would contain a regular speaker layout on a circle with linear speakers, that are positioned at ear-height, and a placement of the listener right in the center of the loudspeaker arrangement. Unfortunately, none of those conditions is fulfilled for the given sound system of the driving simulator.

Nevertheless, the loudspeakers can be equalized to obtain a more accurate frequency response.

[5]Controller Area Network

To achieve this, the frequency response of each loudspeaker has been measured with a measurement microphone at the driver's head position. Subsequently, the resulting frequency response has been roughly linearized using an octave-band graphic equalizer structure with center frequencies ranging from 31.5 Hz to 16 kHz implemented into the Simulink model. The following adjustments have been applied:

Frequency[Hz]	31.5	63	125	250	500	1000	2000	4000	8000	16000
Front Left	0	0	0	+5	+5	0	-3	-3	-3	0
Front Right	0	0	0	+5	+5	0	-3	-3	-3	0
Center	0	0	0	0	0	-10	-10	-10	-10	0
Back Left	0	0	0	0	0	0	0	0	0	0
Back Right	0	0	0	0	0	0	0	0	0	0

Table 6.8: Gain values in dB, applied to the octave-band equalizer to linearize the sound system of the driving simulator

Further documentation and results of the measurement can unfortunately not be provided in this thesis, as the measurements have been conducted by a different party.

In the parametrization procedure for the listening test, wind channel recordings of a BMW X1 model have been used to obtain the wind noise parameters and recordings of a Honda Civic (both road and dynamometer roll measurements) have been used to parametrize the rolling noise.

The different driving sound sources have been placed intuitively in the sound field as described in table 6.9. The reason for the selection of these angles was the theory that the majority of the wind noise is entering the cabin via the side windows. The structure-borne rolling noise mainly finds its way into the car via the front and back axle.

Component	Angle[deg]
Engine noise	0
Rolling noise 1	0
Rolling noise 2	180
Wind noise 1	-60
Wind noise 2	90

Table 6.9: Positions of the driving sound sources

The utilized Ambisonics method was a 1st order Ambisonics mode matching decoder (described in section 4.2.2). Although this type of decoder technically is not the ideal choice for this particular loudspeaker setup, it has proven itself as being sufficient for this purpose.

6.2.1.3 Loudness calibration

The loudness calibration for the listening test has been done by playing back a real wind noise file that has been recorded at 100 km/h. Consecutively, the output volume was adjusted so that an A-weighted SPL meter placed at the driver's head position showed the physically correct level. As all the driving noises generated by the simulation software represent the physical sound pressure level, it is sufficient to calibrate the levels only for one sound with a known sound pressure level.

6.2.2 Velocity perception based on different stimuli

The listening test itself consists of two sub-phases. The goal of the first phase was to familiarize the test subjects with the virtual car simulator and gain some further knowledge about its driving behavior. The test subjects were allowed to drive freely on a realistic track (a freeway with other cars on the road) for 5 minutes. Here, all stimuli (auditory, visual, and sensory) have been switched on. The test subjects have been instructed to drive in different velocity regions and also observe the tachometer so that they can develop a feeling for the according velocities.

In the second phase of the listening test, the tachometer got hidden. The test subjects have then been driving on a test track with three long straights. Each straight had a corresponding length, so that it is easily achievable to accelerate up to 200 km/h. The test track has been designed so that before each of the three straights, the test subjects had to drive through a serpentine to slow down and start the following acceleration to the target velocity at a low velocity. The test subjects got acoustic instructions to accelerate up to a certain velocity right at the end of each serpentine and had to confirm when they thought they had reached the target velocity on the subsequent straight. The confirmation was given by pressing the clutch down. The clutch had no other functionality in this scenario, as the driving simulator was simulating a car with an automatic gearbox. There were three given velocities (50 km/h, 100 km/h, and 160 km/h) that had to be reached, each one for one of the three straights. The order of those three velocities was randomized for every single test subject.

This was repeated several times for different given velocities and different active sound scenarios:

- Engine noise only
- Rolling noise only
- Wind noise only
- Engine-, rolling- and wind noise

The sequence of stimuli combinations also got randomized to prevent a learning effect.

It is expected that the addition of rolling noise to the engine noise will improve the velocity estimation at lower velocities and the addition of the wind noise will improve the velocity estimation at higher velocities, as those are the velocity regions where that driving noise components respectively are the most prominent. The addition of rolling- and wind noise to the engine noise, therefore, should improve the velocity estimation over the whole velocity range.

Unfortunately, some test subjects developed motion sickness during the listening test. Some of them were not even able to finish the test and therefore have been excluded from the results. After excluding those test subjects, 18 test subjects (15 male, 3 female) remained. 17 test subjects have been employees of Magna Steyr, 1 test subject has been an employee of the Virtual Vehicle Graz.

6.2.2.1 Results

The difference between the indicated velocity v_{in} and the target velocity v_t has been calculated for every single test subject, for every target velocity, and for every sound scenario.

$$\Delta v_{v_t} = v_{in} - v_t \tag{6.2}$$

Considering that the deviation from the target velocity is the subject to be evaluated and there are three different target velocities in different velocity regions, it is questionable whether the absolute velocity difference is an appropriate measure for this task. However, from an intuitive perspective, a 10 km/h deviation seems to be way more noticeable at 50 km/h than at 160 km/h. Therefore, the relative deviation is used for further evaluations.

$$\Delta v_{v_t,rel} = \frac{v_{in} - v_t}{v_t} \tag{6.3}$$

For $\Delta v_{v_t,rel} < 0$ the test subject has driven at a lower velocity than the target velocity and therefore has overestimated the current velocity when operating the indicator. On the other hand, for $\Delta v_{v_t,rel} > 0$, the test subject has underestimated the current velocity.

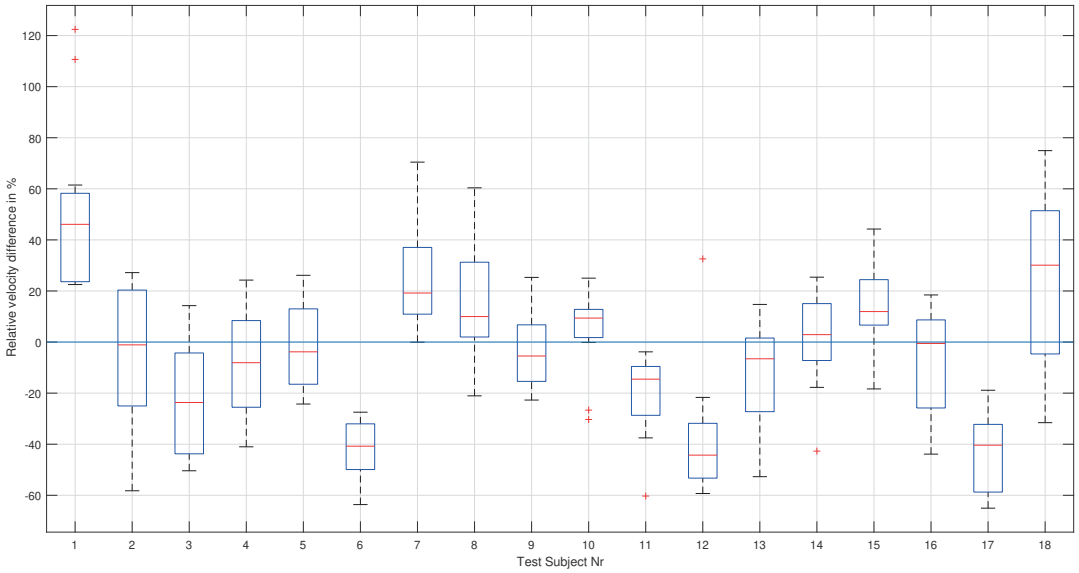


Figure 6.11: Boxplot of the relative velocity deviations of all scenarios and target velocities per test subject

Figure 6.11 is breaking down the relative velocity deviations of all sound scenarios and velocities for each test subject in a box plot. It can be seen that the statistical scattering of the answers is quite high, both for the answers that each test subject gave and between different test subjects. Paired t-tests, which are subsequently used for the statistical evaluation, only consider the differences between various driving scenarios per test subject. Therefore, a z-score normalization for each test subject to adjust the mean and standard deviation has been omitted on purpose.

Three different target velocities (50 km/h, 100 km/h and 160 km/h) had been given which resulted in three different datasets (each containing all test subjects and all driving scenarios). A Lilliefors test ([Lil67]) has been conducted using the pairwise difference between the target velocity datasets ($\Delta\mathbf{v_{150,rel}} - \Delta\mathbf{v_{100,rel}}$, $\Delta\mathbf{v_{150,rel}} - \Delta\mathbf{v_{50,rel}}$, $\Delta\mathbf{v_{100,rel}} - \Delta\mathbf{v_{50,rel}}$) to check for normal distribution. The results and the corresponding p-values are shown in table 6.10.

Data	p-value	Reject H_0?
$\mathbf{\Delta v_{150,rel} - \Delta v_{100,rel}}$	> 0.5	NO
$\mathbf{\Delta v_{150,rel} - \Delta v_{50,rel}}$	0.2212	NO
$\mathbf{\Delta v_{100,rel} - \Delta v_{50,rel}}$	0.4761	NO

Table 6.10: Results of the Lilliefors Test for the difference between the datasets for different target velocities

For each investigated difference between the datasets, the null hypothesis of an underlying normal distribution cannot be rejected. Therefore, we assume a normal distribution is present.

Consecutively, to check whether those three difference datasets origin from a normal distribution with 0-mean value, a t-test is conducted. Basically, this procedure corresponds to a paired t-test between all combinations of two different target velocities. The results of the t-tests are displayed in table 6.11

Data	p-value	Reject H_0?
$\mathbf{\Delta v_{150,rel} - \Delta v_{100,rel}}$	0.2807	NO
$\mathbf{\Delta v_{150,rel} - \Delta v_{50,rel}}$	0.0058	YES
$\mathbf{\Delta v_{100,rel} - \Delta v_{50,rel}}$	0.0001	YES

Table 6.11: Results of the t-test for the difference between the datasets for different target velocities

It shows that at a 95 % confidence level, $\Delta\mathbf{v_{150,rel}} - \Delta\mathbf{v_{50,rel}}$ and $\Delta\mathbf{v_{100,rel}} - \Delta\mathbf{v_{50,rel}}$ do not origin from a 0-mean distribution. Although this null-hypothesis cannot be rejected for $\Delta\mathbf{v_{150,rel}} - \Delta\mathbf{v_{100,rel}}$, the datasets for different target velocities have to be analyzed separately.

Figure 6.12 is showing the statistical measures for each velocity and driving sound scenario as a boxplot. Certain trends can be observed: for higher velocities, the spread of the relative velocity deviation becomes smaller.

Table 6.12: Relative mean velocity deviation and standard deviation for each sound scenario and target velocity in %

	Engine Noise		Rolling Noise		Wind Noise		All Driving Noises	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
50 km/h	−11.97	35.76	−13.63	40.49	3.38	44.56	−26.83	29.29
100 km/h	8.66	34.27	1.97	35.83	6.66	31.54	−9.36	29.46
160 km/h	3.85	21.32	0.46	21.15	−2.26	21.29	−6.39	17.63

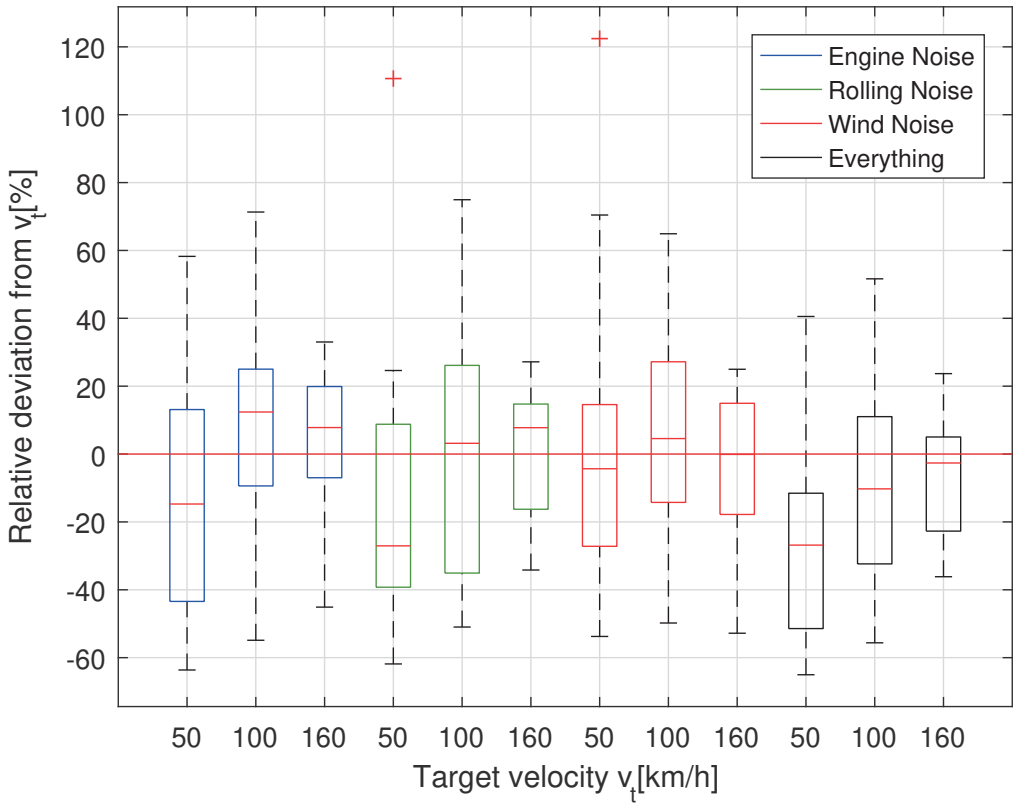


Figure 6.12: Boxplot of the relative velocity deviation for each sound scenario and target velocity over all test subjects

Engine Noise

The test subjects overestimated their actual velocity for a target velocity of 50 km/h. For $v_t = 100$ km/h the velocity has been underestimated, similarly as for $v_t = 160$ km/h. As the engine sound is most prominent in lower velocity regions, it makes sense that the perceived velocity is estimated to be higher there.

Rolling Noise

The test subjects overestimated their actual velocity for a target velocity of 50 km/h. However, for the target velocities of 100 km/h and 160 km/h, the mean relative velocity deviation is close to 0, so the test subjects have been able to determine the current velocity more exact on average. The trend of the mean relative velocity deviation increasing for higher velocities also corresponds to the volume trend of the rolling noise in general, which has a significant contribution to the total driving sound in the lower velocity regions. Also for a target velocity of 50 km/h one significant outlier can be observed.

Wind Noise

The opposite is observable for the wind noise. For a target velocity of 160 km/h, the relative mean velocity difference is lower than for a target velocity of 50 km/h. This also makes sense as the wind noise becomes most prominent in the higher velocity regions. Similar to the rolling noise scenario, at 50 km/h another significant outlier (caused by the same test subject) can be observed.

All Driving Noises

The presence of all driving noises should theoretically yield the most natural experience in the driving simulator, and therefore also the lowest deviations were expected. Obviously, the exact opposite is the case here. For every target velocity, the actual velocity has been overestimated. This could be caused by several reasons such as a subjectively perceived mismatch between the acoustic and optical stimulus that was reported by several test subjects (see section 6.2.3).

When looking at the difference between two calculated relative deviations (equation 6.3) for different sound scenarios, the target velocity v_t in the nominator cancels out. Therefore, we can only see the trend in the velocity estimation, regardless of whether the estimation moves closer to the target velocity or further away. Equation 6.4 exemplary shows this for a target velocity of 50 km/h and when comparing the "wind noise only"-scenario to the "rolling noise only"-scenario.

$$\mathbf{\Delta v_{50,wn-rn,rel}} = \mathbf{\Delta v_{50,wn,rel}} - \mathbf{\Delta v_{50,rn,rel}} = \mathbf{v_{50,wn,rel}} - \mathbf{v_{50,rn,rel}} \tag{6.4}$$

	WN - RN	EN - RN	EN - WN	ALL - WN	ALL - EN	ALL - RN
50 km/h	17.01 %	1.66 %	-15.35 %	-30.21 %	-14.86 %	-13.20 %
100 km/h	-4.69 %	6.69 %	2.01 %	-16.02 %	-18.02 %	-11.33 %
160 km/h	-2.72 %	3.39 %	6.10 %	-4.13 %	-10.23 %	-6.85 %

Table 6.13: Percentage-wise mean deviation regarding the velocity estimation between two different sound scenarios. Green cells show statistical significance using a paired t-test with $\alpha = 0.05$, Bonferroni-Holm corrected

In table 6.13 it can be observed again that at a target velocity of 50 km/h and when only hearing the rolling noise, the test subjects have estimated their actual velocity on average 17.01 % higher than when driving with the wind noise only. Conducting a paired t-test using two different sound scenarios provides information on whether the differences between two sound scenarios are statistically significant. Scenarios showing statistical significance at a confidence level of 95 % are highlighted as a green cell in the table. Also, the Bonferroni-Holm method was applied separately for each target velocity to prevent an alpha error accumulation. Here, only several comparisons between "all driving sounds" and individual driving sounds show significant differences. In those cases, "all driving sounds", the current velocity is estimated to be significantly higher.

To find out whether certain sound scenarios support the velocity estimation in particular velocity regions, the difference between the absolute velocity deviations was investigated (again displayed in the following example for wind/rolling noise at 50 km/h):

$$\mathbf{\Delta v_{50,|wn|-|rn|,rel}} = |\mathbf{\Delta v_{50,wn,rel}}| - |\mathbf{\Delta v_{50,rn,rel}}| \tag{6.5}$$

Table 6.14 is showing the mean results over all test subjects for the difference between the absolute velocity deviations in different sound scenarios and at different target velocities, such as described exemplary in equation 6.5.

	WN - RN	EN - RN	EN - WN	ALL - WN	ALL - EN	ALL - RN
50 km/h	-1.21 %	-2.17 %	-0.96 %	1.93 %	2.89 %	0.72 %
100 km/h	-3.77 %	-0.91 %	2.85 %	-0.48 %	-3.34 %	-4.25 %
160 km/h	-1.57 %	-0.63 %	0.94 %	-2.22 %	-3.15 %	-3.79 %

Table 6.14: Percentage-wise mean improvement/impair of estimation of the target velocity when comparing two different sound scenarios - no statistically significant results

Adhering to the wn - rn example, positive values for $\mathbf{\Delta v_{50,|wn|-|rn|,rel}}$ would imply that the test subjects were able to estimate the velocity more precisely when only hearing the rolling noise than when only hearing the wind noise. This case would be expected as the rolling noise contributes a bigger part to the total driving sound composition at 50 km/h than the wind noise. Surprisingly, the mean value for $\mathbf{\Delta v_{50,|wn|-|rn|,rel}}$ over all test subjects is -1.21 %, which is the opposite of what was to be expected. The test subjects gave better velocity estimations for a target velocity of 50 km/h when only hearing the wind noise. The values for target velocities of 100 km/h (-3.77 %) and 160 km/h (-1.57 %) also stayed negative. A trend for better velocity estimations at higher target velocities when only hearing the wind noise cannot be observed clearly. The same can be said about the difference between all other scenarios. Using a paired t-test using two different sound scenarios here leads to no significant results, regardless of which sound scenarios are compared.

6.2.3 Subjective Evaluation

After the listening test has been concluded, a short questionnaire was handed to the test subjects. They had the possibility to rate the driving noise in three different categories and state their opinion on the sound and their experience in the driving simulator in general.

Three questions have been asked to be rated from 1 (very unrealistic) to 10 (very realistic):

1. How realistic would you rate the sound of the driving sounds in the driving simulator?
2. How realistic would you rate the spatial impression of the driving sounds in the driving simulator?
3. How realistic would you rate the balance between the driving sound components in the driving simulator?

Question Nr	Mean	Std.
1	6.58	1.80
2	6.79	1.40
3	6.53	1.68

Table 6.15: Mean and standard deviation for the subjective assessment of the driving sounds

The written feedback repeatedly included statements like

- low-frequency rumble (rolling noise) is too loud
- background noises made by the driving simulator is disturbing
- visual representation is unrealistic (looking slower than the actual velocity)
- engine noise is too low in volume
- wind noise is sounding good but too low in volume

6.2.4 Conclusion

Unfortunately, the second listening test did not statistically prove the assumption that certain driving sounds improve the velocity perception in different velocity regions. Still, the tendency of estimating the velocity higher when the driving sounds are louder is clearly visible in the data. Especially for all driving sounds together (the loudest sound scenario), the test subjects estimated their velocity significantly higher compared to other scenarios when they only heard individual driving sounds.

Several occurrences before and during the listening test potentially have influenced the outcome. Originally, a fifth sound scenario was planned where no driving sound should have been present at all. However, the shakers of the driving simulator have been producing audible deterministic order sounds that have correlated with the engine noise, and it was not easily possible to shut them off. Due to time issues, the decision to skip this sound scenario has been made, as, with the shakers engaged, it would not have brought any usable data. The lack of the "no sound at all"-scenario aggravated the evaluation of the measured data.

What still can be concluded from this test is that the subjective impression of the synthesized driving sound was evaluated positively, and the personal feedback has brought valuable insights on how the driving sounds can be improved further.

Chapter 7

Conclusion and Outlook

A sound generator for wind- and rolling noises has been developed in Simulink/Matlab. The quality of the synthesized sounds has been validated in a listening test, and the driving noise generator has been field-tested in a static driving simulator at the Virtual Vehicle Graz.

The quality of the synthesized sounds has overall been proven as being sufficiently given in a listening test conducted on headphones (section 6.1.2). However, there is still room to improve the synthesized sounds further, but that would have gone beyond the scope of this thesis.
The synthesized rolling noises could be improved further by adding more realism to them. Two potential methods to achieve this by considering road bumps have been proposed in section 3.2.3. One of them has been discarded due to unsatisfactory results (further work would be needed), the other one has not been implemented and tested due to time reasons.
The synthesized wind noise samples have been indiscernible from the real wind noise samples recorded in a wind channel. Still, as the synthesized sounds have a static volume structure, the realism can be improved by applying a modulation scheme to the synthesized sounds to simulate "on-the-road"-conditions. The basic principle of this modulation has been implemented in the Simulink model. However, the parameterization and tuning of the method and, therefore, its usage have not been possible due to the lack of appropriate data. Here, further work to improve the realism of the driving noise simulator would be desirable.

While the quality of the synthesized sounds has been validated in a headphone scenario, the listening test in the static driving simulator has shown that the portation from headphones to the driving simulator, especially the spatial playback of the

sounds, is not a trivial task. Unfortunately, the listening test under the given circumstances could not prove the assumption that specific sound scenarios are improving the test subjects' ability to identify their current velocity. Still, it could be seen that louder sound scenarios increased the perceived velocity of the test subjects.
The subjective evaluation of the driving sounds in the static driving simulator (section 6.2.3) has shown that the driving sounds have been perceived as rather realistic than unrealistic, but there is still room for improvements. To improve the perception in the driving simulator, changes in the set-up would have to be deployed and tested directly at the driving simulator, which would be a rather costly and time-consuming task and therefore not possible within the scope of this thesis. Especially the spatialization of the driving sounds in driving simulators, including the choice of the Ambisonics decoder strategy for this purpose, is a topic that still requires more research. In conclusion, while the basic integration of the driving noise module into different driving simulators is easy and should yield sufficiently good results, optimal results can only be achieved when spending enough time to adapt the driving sound module to the particular driving simulator.

Bibliography

[AVdP01] M. Adams and P. Van de Ponseele, "Sound quality equivalent modeling for virtual car sound synthesis," *SAE transactions*, pp. 1906–1912, 2001.

[AVS06] L. Aarts and I. Van Schagen, "Driving speed and the risk of road crashes: A review," *Accident Analysis & Prevention*, vol. 38, no. 2, pp. 215–224, 2006.

[BKHK08] T. Beckenbauer, P. Klein, J.-F. Hamet, and W. Kropp, "Tyre-road noise prediction: A comparison between the speron and hyrone models-part 1," *Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3388–3388, 2008.

[BS09] C. Bederna and E. Saemann, "Contributions to a better understanding of tire cavity noise," *NAG/DAGA Rotterdam*, 2009.

[BU98] H. Brandenstein and R. Unbehauen, "Least-squares approximation of fir by iir digital filters," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 21–30, 1998.

[CD19] D. J. Carr and P. Davies, "A method to simulate wind noise in cars under time-varying wind conditions," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 259, no. 4. Institute of Noise Control Engineering, 2019, pp. 5763–5771.

[CS86] F. Charpentier and M. Stella, "Diphone synthesis using an overlap-add technique for speech waveforms concatenation," in *ICASSP'86. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 11. IEEE, 1986, pp. 2015–2018.

[DRKM+12] S. Denjean, V. Roussarie, R. Kronland-Martinet, S. Ystad, and J.-L. Velay, "How does interior car noise alter driver's perception of motion? multisensory integration in speed perception," in *Acoustics 2012*, 2012.

[DS11] S. Da Silva, "Non-parametric identification of mechanical systems by kautz filter with multiple poles," *Mechanical Systems and Signal Processing*, vol. 25, no. 4, pp. 1103–1111, 2011.

[Eva70] L. Evans, "Speed estimation from a moving automobile," *Ergonomics*, vol. 13, no. 2, pp. 219–230, 1970.

[FH09] M. Fröjd and A. Horner, "Sound texture synthesis using an overlap–add/granular synthesis approach," *Journal of the Audio Engineering Society*, vol. 57, no. 1/2, pp. 29–37, 2009.

[FHRB16] M. Frank, R. Höldrich, M. Resch, and S. Brandl, "Generating consistent exterior vehicle sound from interior sound," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 253, no. 4. Institute of Noise Control Engineering, 2016, pp. 3858–3865.

[Fra14] M. Frank, "How to make ambisonics sound good," in *Forum Acusticum,(Krakow)*, 2014.

[Hol79] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian journal of statistics*, pp. 65–70, 1979.

[HP08] M. S. Horswill and A. M. Plooy, "Auditory feedback influences perceived driving speeds," *Perception*, vol. 37, no. 7, pp. 1037–1043, 2008.

[IY96] K. Iwao and I. Yamazaki, "A study on the mechanism of tire/road noise," *JSAE review*, vol. 17, no. 2, pp. 139–144, 1996.

[KBSR20] J. Koch, T. Bartosch, A. Sontacchi, and W. Reinalter, "Real-time capable wind and rolling noise synthesis for a more realistic vehicle simulator experience," SAE Technical Paper, Tech. Rep., 2020.

[KP06] M. Karjalainen and T. Paatero, "Equalization of loudspeaker and room responses using kautz filters: Direct least squares design," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–13, 2006.

[KRW09] S. Krampol, M. Riegel, and J. Wiedemann, "A procedure to simulate the turbulent noise interior of cars," in *NAG/DAGA congress. Rotterdam*, 2009.

[Lil67] H. W. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *Journal of the American statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.

[MJ11] N. Merat and H. Jamson, "A driving simulator study to examine the role of vehicle acoustics on drivers' speed perception," 2011.

[MJ13] J. Maillard and J. Jagla, "Real time auralization of non-stationary traffic noise-quantitative and perceptual validation in an urban street," *Proc. of AIA-DAGA*, 2013.

[OS09] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. USA: Prentice Hall Press, 2009.

[OSWD14] N. Oettle, D. Sims-Williams, and R. Dominy, "Assessing the aeroacoustic response of a vehicle to transient flow conditions from the perspective of a vehicle occupant." *SAE International journal of passenger cars. Mechanical systems.*, vol. 7, no. 2, pp. 550–558, 2014.

[Pie78] A. Piersol, "Use of coherence and phase data between two receivers in evaluation of noise environments," *Journal of Sound and Vibration*, vol. 56, no. 2, pp. 215–228, 1978.

[PK03] T. Paatero and M. Karjalainen, "Kautz filters and generalized frequency resolution: Theory and audio applications," *Journal of the Audio Engineering Society*, vol. 51, no. 1/2, pp. 27–44, 2003.

[RWH06] M. Riegel, J. Wiedemann, and M. Helfer, "The effect of turbulence on in-cabin wind noise–a comparison of road and wind tunnel results," in *Tagungsband der „6th MIRA International Vehicle Aerodynamics Conference ", Gaydon, Großbritannien, Heritage Motor Centre*, 2006.

[SA95] J. O. Smith and J. S. Abel, "The bark bilinear transform," in *Proceedings of 1995 Workshop on Applications of Signal Processing to Audio and Accoustics*, 1995, pp. 202–205.

[SCJvdA13] M. Sarrizin, C. Colangeli, K. Janssens, and H. van der Auweraer, "Synthesis techniques for wind and tire-road noise," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 247, no. 6. Institute of Noise Control Engineering, 2013, pp. 2303–2312.

[Sta18] B. Stahl, "Improved engine noise synthesis algorithm," Master's thesis, TU Graz, 2018.

[SZH18] C. Schörkhuber, M. Zaunschirm, and R. Höldrich, "Binaural rendering of ambisonic signals via magnitude least squares," in *Proceedings of the DAGA*, vol. 44, 2018, pp. 339–342.

[VL93] H. Vold and J. Leuridan, "High resolution order tracking at extreme slew rates, using kalman tracking filters," SAE Technical Paper, Tech. Rep., 1993.

[VMB97] H. Vold, M. Mains, and J. Blough, "Theoretical foundations for high performance order tracking with the vold-kalman tracking filter," *SAE transactions*, pp. 3046–3050, 1997.

[WK07] L. Weruaga and M. Képesi, "The fan-chirp transform for non-stationary harmonic signals," *Signal Processing*, vol. 87, no. 6, pp. 1504–1522, 2007.

[YKL10] N. Yousefian, K. Kokkinakis, and P. C. Loizou, "A coherence-based algorithm for noise reduction in dual-microphone applications," in *2010 18th European Signal Processing Conference*. IEEE, 2010, pp. 1904–1908.

[Zel12] P. Zeller, *Handbuch Fahrzeugakustik - Grundlagen, Auslegung, Berechnung, Versuch - 2. Auflage*, 200th ed. Berlin Heidelberg New York: Springer-Verlag, 2012.

[ZF19] F. Zotter and M. Frank, *Ambisonics*. Springer, 2019.

[ZFKC14] F. Zotter, M. Frank, M. Kronlachner, and J.-W. Choi, "Efficient phantom source widening and diffuseness in ambisonics," *10.14279/depositonce-4103*, 2014.

[ZFP13] F. Zotter, M. Frank, and H. Pomberger, "Comparison of energy-preserving and all-round ambisonic decoders," *Fortschritte der Akustik, AIADAGA,(Meran)*, 2013.

[ZFW+91] E. Zwicker, H. Fastl, U. Widmann, K. Kurakata, S. Kuwano, and S. Namba, "Program for calculating loudness according to din 45631 (iso 532b)," *Journal of the Acoustical Society of Japan (E)*, vol. 12, no. 1, pp. 39–42, 1991.

[Zwi61] E. Zwicker, "Subdivision of the audible frequency range into critical bands (frequenzgruppen)," *The Journal of the Acoustical Society of America*, vol. 33, no. 2, pp. 248–248, 1961.