Aachener Beiträge zur Hörtechnik und Akustik

Christian G. Dreier

Virtual Acoustics: Synthesis of Traffic Sound Sources

IHTA Institute for Hearing Technology and Acoustics
RWTH AACHEN UNIVERSITY

Virtual acoustics:

Synthesis of traffic sound sources

Von der Fakultät für Elektrotechnik und Informationstechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN

genehmigte Dissertation

vorgelegt von

M.Sc.
Christian G. Dreier

aus Kamp-Lintfort, Deutschland

Berichter:

Univ.-Professor Dr. rer. nat. Michael Vorländer
apl. Professor Dr.-Ing. Peter Urban

Tag der mündlichen Prüfung: 28. Juli 2025

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

One man's noise is another man's signal.

EDWARD NG, 1990

Abstract

Being a preconscious environmental factor in modern societies, the acoustics of urban environments has an unavoidable impact on our living conditions and public health. The growing understanding of sound effects on humans was accompanied by the development of methods for sound field simulation. Introduced in 2002, the European Commission's Environmental Noise Directive (END) obliges the neighboring states to prepare and continuously update noise maps that are representing the noise originating from road, rail and air traffic as well as industrial sound sources in form of equivalent noise level calculations. Auralizations aim for producing audible sound based on these simulations, for example of auditoriums being planned, vehicle interiors or the built urban environment. The auralization technique should enable subjective evaluations of acoustics when a purely numerical and level-based assessment using sound pressure levels or spectra is not meaningful enough.

Models for sound sources provide the necessary input information for the sound wave propagation simulation between the location of sound generation and the position of the listener. The objective of this thesis is to characterize, model and implement road-traffic and air-traffic sound sources, thereby achieving a holistic auditory fit of urban auralizations with the real-world experience. However, the modeling effort varies depending on the specific application and affects the resulting configurability, plausibility and applicability for real-time auralizations. The modeling of moving sound sources is particularly challenging, as distortion-free acquisition of the source information must be carried out exclusively at non-moving measurement points, and thus from a considerable spatial distance. In addition, larger source expansions can make it difficult to measure from all spatial directions.

This work systematically introduces the measurement, modeling and synthesis techniques of moving traffic sound sources for auralization. The aim of the presented content is to reproduce properties of real-world sound sources in acoustic virtual reality. For this purpose, methods are presented for systematically obtaining spatial and the spectro-temporal properties of moving sound sources, its algorithmic extraction and further integration into auralization frameworks.

Kurzfassung

Die Akustik urbaner Umgebungen stellt einen vorbewussten Umwelteinfluss auf die Lebensbedingungen moderner Gesellschaften und die öffentliche Gesundheit dar. Mit wachsendem Verständnis der Auswirkungen von Schall auf die Menschen ging die Entwicklung von Methoden zur Schallfeldsimulation einher. Die 2002 eingeführte Environmental Noise Directive (END) der Europäischen Kommission verpflichtet die Mitgliedstaaten, Lärmkarten zu erstellen, die den vom Straßen-, Schienen- und Flugverkehr sowie von industriellen Schallquellen ausgehenden Lärm in Form äquivalenter Lärmpegelberechnungen darstellen. Auralisationen zielen auf die Hörbarmachung dieser Berechnungen beispielsweise von in Planung befindlichen Auditorien, Fahrzeug-Innenräumen oder urbaner Bebauung ab. Die Auralisationstechnik soll es ermöglichen, subjektive Bewertungen der Akustik vorzunehmen, wenn eine rein zahlen- und pegelbasierte Betrachtung mittels Schalldruckpegeln oder Spektren zu wenig aussagekräftig ist.
Für die Simulation der Schallwellenausbreitung bis zur Hörposition stellen Modelle der Schallquellen die nötigen Eingangsinformationen bereit. Das Ziel dieser Arbeit ist es, Schallquellen von Straßenverkehrs- und Fluglärm zu charakterisieren, modellieren und implementieren, um so eine ganzheitliche klangliche Anpassung der städtischen Auralisationen an die reale Erfahrung zu erreichen. Der Modellierungsaufwand schwankt jedoch abhängig vom konkreten Anwendungszweck und wirkt sich auf die resultierende Parametrierbarkeit, Plausibilität und die Möglichkeit zum Einsatz für Echtzeit-Auralisationen aus. Besonders herausfordernd ist die Modellierung bewegter Schallquellen, da eine verzerrungsfreie Erfassung der Quellinformationen an nicht-bewegten Messpunkten, und somit aus erheblicher räumlicher Distanz, erfolgen muss. Zudem können größere Quellausdehnungen selbst die Messung aus allen Raumrichtungen erschweren.
Diese Arbeit befasst sich mit der Messung, Modellierung und Synthese bewegter Verkehrsschallquellen für Auralisationen. Ziel der vorgestellten Inhalte ist es, die akustischen Eigenschaften von Schallquellen aus der realen Welt in der virtuellen Realität zu reproduzieren. Dazu werden Methoden vorgestellt, mit denen sowohl die räumlichen als auch die spektro-temporalen Eigenschaften von bewegten Schallquellen systematisch erfasst, signaltheoretisch ausgewertet und in Auralisationssysteme integriert werden können.

Contents

Glossary xiii

1 Introduction 1
1.1 Motivation . . . 1
1.2 From noise mapping to traffic auralization: The past 25 years . . 4
1.3 Contribution of this thesis . . . 9

2 Fundamentals of sound source characterization 11
2.1 Traffic sound auralization chain . . . 11
2.2 Sound source model calibration . . . 13
2.3 Modeling strategies . . . 15
2.3.1 Black and white-box models . . . 15
2.3.2 Bottom-up and top-down approaches . . . 15
2.3.3 Models for wave-based vs. ray-based auralization frameworks 16
2.4 The 4 dimensions of source characterization . . . 17
2.4.1 Spectral properties . . . 18
2.4.2 Temporal properties . . . 19
2.4.3 Spatial properties . . . 20
2.4.4 Operating state properties . . . 22
2.5 Postprocessing methods . . . 23
2.5.1 Time-variant wave backpropagation . . . 23
2.5.2 Signal separation . . . 25
2.6 Sound synthesis techniques . . . 26
2.6.1 Signal model synthesis . . . 27
2.6.2 Physical model synthesis . . . 28

3 Passenger car auralization model 31
3.1 Introduction . . . 32
3.1.1 State of the art . . . 34
3.1.2 Contribution of this work . . . 37
3.2 Experimental setup for source characterization of road-traffic sound emission . . . 38
3.2.1 Array design for pass-through measurements . . . 39
3.2.2 Microphone calibration . . . 39

3.3 Postprocessing: Time-variant wave backpropagation 41
3.3.1 Source extension-dependent spreading loss $\mathbf{A_{div}}$ 42
3.3.2 Ground effect $\mathbf{A_{gr}}$ 43
3.3.3 Uncertainty considerations 44
3.4 Measurement 49
3.4.1 Vehicle classes 49
3.4.2 *In situ* array installation on test track 50
3.4.3 Measurement procedure 50
3.5 Results 53
3.5.1 Sound power of road-traffic sound sources 53
3.5.2 Directivity patterns of road-traffic sound sources 54
3.6 Powertrain sound synthesis 60
3.6.1 Combustion engines 60
3.6.2 Electric engines 62
3.7 Tire/road noise synthesis 67
3.7.1 Wet and dry asphalt filter curve 69
3.8 Real-time implementation as VST plugin 70
3.9 Summary and discussion 71

4 Aircraft auralization model 73
4.1 Introduction 73
4.1.1 State of the art 74
4.1.2 Contribution of this work 76
4.2 Source characterization of simulated subsonic jet noise 81
4.2.1 Spatial sound field decomposition 83
4.2.2 Jet noise source localization 89
4.2.3 Equivalent source model of jet noise 92
4.3 Far-field jet noise radiation 94
4.4 Aircraft sound synthesis 97
4.4.1 Extraction of fan noise modulation parameters from flyover recordings 98
4.5 Summary and discussion 103

5 Drone auralization model 105
5.1 Introduction 105
5.1.1 State of the art 105
5.1.2 Contribution of this work 106
5.2 Source characterization of drone sound emission 107
5.2.1 Modeling concept 107
5.2.2 Measurement setup 109

5.2.3 Measurement equipment 109
5.3 Analysis and synthesis of drone sound emission 110
5.3.1 Broadband spectrum 110
5.3.2 Analysis of tonal amplitude and frequency modulations . 111
5.3.3 Synthesis of stochastic drone sound component 114
5.4 Directivity model 115
5.4.1 Radiation in the median and frontal plane 116
5.5 Synthesis results 117
5.6 Summary and discussion 119

6 Case study: Auralization and visualization of complex urban environments 123
6.1 Virtual urban environment IHTApark 124
6.2 VR-based listening experiment 128
6.2.1 Urban scenarios description 128
6.2.2 Experimental design and procedure 128
6.2.3 Results 129
6.3 Summary and discussion 129

7 Conclusion and Outlook 133

Appendices 137

A Complex boundary conditions at the ground 139

B Wet asphalt auralization filter data 143

C Vehicle noise sources 145

D Semi-empirical aircraft noise model input parameters 147

Bibliography 163

Curriculum Vitae 165

Acknowledgements 167

Glossary

Acronyms

ADC	Analog-to-Digital Conversion
AM	Amplitude Modulation
AR	Augmented Reality
AVAS	Acoustic Vehicle Alert System
BEM	Boundary Element Method
CPX	Close-Proximity Method
DAW	Digital Audio Workstation
DSP	Digital Signal Processing
DUT	Device Under Test
FDTD	Finite-Difference Time-Domain Method
FEM	Finite Element Method
FIR	Finite Impulse Response
FM	Frequency Modulation
GA	Geometrical Acoustics
GUI	Graphical User Interface
HMD	Head-Mounted Display
HOA	Higher-Order Ambisonics
HpTF	Headphone-to-Ear Transfer Function
HRTF	Head-related Transfer Function
ICE	Internal Combustion Engine
IIR	Infinite Impulse Response
IR	Impulse Response
ISO	International Standard Organization
LTI	Linear Time-invariant System
NMF	Non-negative matrix factorization
PCA	Principal Component Analysis
PSD	Power Spectral Density
PSTD	Pseudospectral Time-Domain Method
RT	Ray Tracing

SOS	Second-Order Section
SPL	Sound Pressure Level
STFT	Short-Time Fourier Transform
VA	Virtual Acoustics
VR	Virtual Reality
VST	Virtual Studio Technology

1

Introduction

1.1 Motivation

Living quality perception in urban environments is closely linked to acoustical factors. Traffic is the most dominating source of environmental noise. Current research associates exposure to traffic noise with a variety of effects on human health and well-being. Resulting from a comprehensive meta-analysis, the "Guidelines for community noise" were authored by an expert task force from the World Health Organization (WHO) in 1999 in order to consolidate scientific knowledge on the health impacts of noise and support environmental health authorities to protect the population from harmful noise effects [18]. This report was followed by regulatory frameworks and noise regulations, such as the notable Environmental Noise Directive (END) by the European Union [30]. The approaches being discussed can be classified into three subcategories: Noise emission reduction (at the sound sources), noise mitigation (immission reduction) [90] [126] and soundscape creation (psychoacoustical treatment by masking) [91] [157]. Noise not only has an impact on health, but recently was enhanced by an economic dimension too. By heading at 65% reduction of perceived aircraft noise emissions within the next decades, the strategic paper Flightpath 2050 by the European Commission [50] resulted in a regulatory market mechanism as the noise-related taxation of air traffic was introduced. The International Civil Aviation Organisation (ICAO) provides guidance on noise-related charges [32].

These regulatory and economic frameworks have led to comprehensive research on noise mapping tools and noise-preventing countermeasures. For the assessment of noise, measurements and simulations in the form of A-weighted equivalent continuous sound levels $L_{\mathrm{A,eq}}$ are the worldwide standard. They are useful for simple comparisons and necessary as a basis for decision makers. Striving for a harmonization of the multiple national parallel developments regarding noise mapping in Europe, the CNOSSOS-EU framework (2010-2015, Common NOise

aSSessment methOdS) consistently implements numerical source and directivity models for road-, rail-, and aircraft-traffic sound sources [92]. It provides noise predictions in octave bands in the frequency range from 125 Hz to 4 kHz for road-traffic noise and from 50 Hz to 10 kHz for aircraft noise, respectively. However, for the aspects of social acceptance and participation, the consideration of psychoacoustical measures – beyond the limited scope of level metering using $L_{\mathrm{A,eq}}$ values – becomes more and more relevant.

Focusing on bridging this gap, auralization technique strives for producing audible sound from a listener's perspective (Fig. 1.1) based on sound field simulation using physics-based computer models [177]. From the listener's perspective, an acoustically complex urban scene is composed by superposition of a large number of direct and reflected sound paths (black lines). Each (direct or reflected) path orinates from a sound source, such as cars, airplanes or birds (black objects). Light grey objects indicate visible objects that are not necessarily audible.

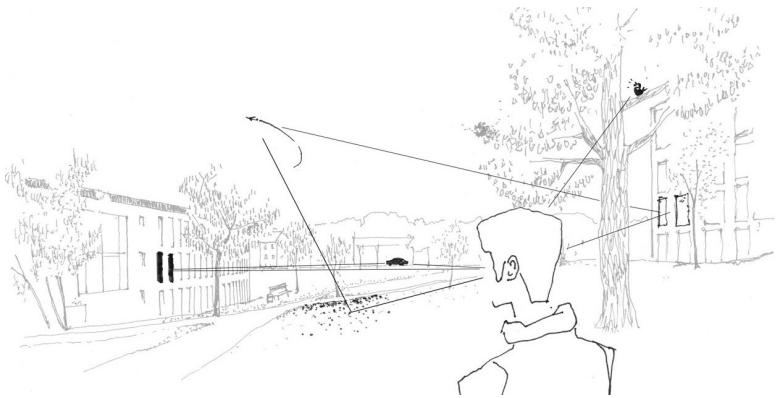


Figure 1.1: Graphical representation of the auralization approach. Drawing by J. Llorca-Bofí from [111].

Shown at a higher level of abstraction, Figure 1.2 shows the paradigm for auralization. The signal processing chain consists of three distinguished and concatenated elements. First, being highlighted as grey-colored box, a (recorded or synthesized) sound source signal is fed into the sound propagation rendering module in order to account for the acoustical medium and surface material properties along each geometrically valid (direct, reflected and diffracted) path in the space between the source position and the receiver position. The sound propagation model uses data from the scene description (blue) and the physical acoustic laws (violet). Third, a reproduction module combines this information (which is coded for each path by

the parameters time delay, amplitude and incidence angle) with a receiver model in order to render spatial audio. Finally, for playback this spatial audio stream is decoded to either headphone-based or loudspeaker-based playback systems.

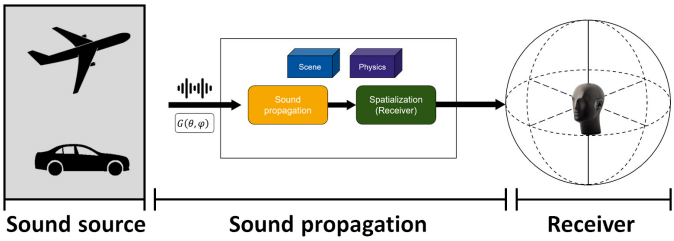


Figure 1.2: Typical auralization chain along the sound source emission and propagation to the receiver.

The subjective plausibility of auralizations, i.e. their agreement with the listener's expectation towards a corresponding real acoustical event [108], depends on the individual quality of many factors along the auralization chain (Fig. 1.2), such as the accuracies of sound source characterization (including models of their spatial and spectro-temporal characteristics), algorithms for computing the sound propagation effects [163], models for surface material properties, and the receiver model. Moreover, interaction- and hardware-related aspects are relevant, especially regarding an aurally correct equalization of the playback system response, computation time (offline vs. real-time rendering), number of degrees of freedom, avoidance of audible artifacts, and additional visualization of the noise sources using a head-mounted display (HMD).

This thesis focuses on methods for sound source characterization, subsequent modeling of moving traffic sound sources and synthesis procedures for their application in auralizations of complex urban environments. In particular, the specific acoustical features of road-traffic (passenger cars and motorcycles) as well as air-traffic sound sources (conventional aircraft and drones) will be discussed. It is structured as follows: The remaining part of this chapter sketches the general path from noise mapping to traffic auralization including publications from the past 25 years. More specific, source type-related literature is referenced in the corresponding chapters. Chapter 2 summarizes technical details of the traffic sound auralization chain, the dimensions of source characterization, source modeling and calibration procedures. The source type-related chapters focus on specific modeling techniques reaching beyond the according state of the art. In particular capturing the hemi-spherical directivity patterns of moving road-traffic sound

sources (chapter 3), reducing the complexity of aeroacoustic sources (chapter 4), and reproducing the statistical emission signal properties of a drone for enhanced plausibility (chapter 5). Finally, some applications of the models are presented in chapter 6.

1.2 From noise mapping to traffic auralization: The past 25 years

During the course of this chapter the evolution of the main branches contributing to state-of-the-art traffic sound auralization is summarized by highlighting notable scientific contributions. In the bigger picture, a parallel evolvement of politically-motivated noise regulations and according development of simulation tools for standardized noise prediction can be observed. Furthermore, the implementation of the noise regulations adopted at European level in the form of national directives has also led to a lively interplay between various research institutions and software implementations. Notable frameworks focusing on standardized noise mapping (without audible sound) are presented in the following.

Noise mapping tools

Systematic noise research has been carried out since the 1950s. As a prominent example, Leo Beranek describes in his autobiography the derivation of a perception-based assessment metric tailored to aircraft noise in 1957 [16]. Karl Kryter's book titled "The Effects of Noise on Man" from 1970 is a milestone in this field since it compiles the essential knowledge of that time with regard to damage risk from noise exposure and broad knowledge on environmental noise evaluation [97].

Finished in 2000, the Nord2000 project presented a comprehensive prediction scheme comprising all important noise types, including road, rail and air traffic as well as industrial plants and wind turbines [1]. It considers sound propagation effects due to geometrical divergence, atmospheric absorption, ground reflections, diffraction and meteorological refraction. Nord2000 predicts noise levels in 27 one-third octave bands in the frequency range 25 Hz to 10 kHz, assuming wave propagation simulations based on geometrical acoustics. In comparison to ISO 9613-2 [82], Nord2000 introduces a segmentation of the area between a sound source and a receiver position into approximately 10-15 elements with acoustical material parameter definitions for impedance and roughness. The reflections and attenuations caused by the ground are based on geometric ray theory.

The EU-funded, large-scale HARMONOISE project (2001-2004) specifically focused on close-to-ground sources and propagation distances up to 1 km with moderate atmospheric refraction conditions [87]. In course of the project, new predictions schemes for noise emission from road and railway vehicles were obtained. Moreover, two propagation models were developed – a reference model based on numerical simulations and an engineering model based on analytical and heuristic solutions. The Harmonoise framework computes source-to-receiver excess attenuations ΔL in the frequency domain.

The IMAGINE project (2003-2006, Improved Methods for the Assessment of the Generic Impact of Noise in the Environment), aimed to extend the Harmonoise road and rail source databases and to develop prediction methods for aircraft and industrial noise sources. Equal to Nord2000, noise levels in 27 one-third octave bands in the frequency range 25 Hz to 10 kHz are predicted.

As mentioned before, the CNOSSOS-EU framework (2010-2015) revises and integrates the methods developed as part of the Harmonoise and Imagine projects into a unified method by consistently implementing numerical noise predictions in octave bands in the frequency range from 125 Hz to 4 kHz for road-traffic noise, from 50 Hz to 10 kHz for aircraft noise and in one-third octave bands between 50 Hz and 10 kHz for rail-traffic noise, respectively [92]. It became a European Commission directive in 2015 and since January 1, 2019, strategic noise maps in Europe must be produced according to this method.

Road-traffic auralization frameworks

Generally, the auralization of outdoor sceneries is highly complex due to great variability in simulation parameters, time-variant processes of the sound generation mechanisms at the sources and during sound progation, as well as the (theoretically) arbitrarily high number and motion of natural and technical sound sources. Several virtual road-traffic scenarios have already been implemented in European research projects.

Between 2000 and 2003, exterior vehicle noise was measured and pass-by simulations were realized in the European research project SVEN (Sound quality of exterior vehicle noise). A major research outcome indicated the importance of psychoacoustical assessments when considering the physiological reactions of residents acutely triggered by road traffic noise. Compared to the SVEN project, the European projects Quiet City Transport (2005-2009) and CityHush - Acoustically green road vehicles and city areas (2010-2012) increased the variablility of motor vehicles with combustion engines, selected electric vehicles were mea-

sured on test tracks, and various two-wheeled vehicles were considered, as well as exemplary scenarios on the psychoacoustic effects of specific road traffic sounds were auralized [47]. During course of this project, the connection to the PTV VISSIM traffic flow simulation software was implemented to generate road traffic noise for realistic traffic scenarios considering the interaction between vehicles. In addition, exemplary initial studies have been initiated on the potential impact of Acoustic Vehicle Alert Systems (AVAS) on the resulting overall traffic noise and associated noise annoyance [53].

Comprehensive research activity on auralization of passenger car pass-by noise can be traced back to work at Chalmers University, among others. The LISTEN project (2008-2011) presented an acoustic demonstrator of three urban scenarios [127]. All scenarios demonstrated the noise-reducing effect of noise barriers close to the source based on precalculation of the diffracted spectra. The plausibility was enhanced by playback of traffic-free background sound recordings. However, no visual informations of the scenes were included to the demonstrations.

A group from the Centre Scientifique et Technique du Bâtiment (CSTB) integrates different traffic sound auralization models to their software MithraSOUND. Maillard and Jagla, as part of the European HOSANNA project (2009-2013, Holistic and sustainable abatement of noise by optimized combinations of natural and artificial means), have implemented auralization of moving and accelerating cars in urban environments, with real-time user interaction [83]. Based on tire/road noise and engine noise recordings in the near field, emission signals were synthesized with granular synthesis enhanced by a pitch-synchronous overlap-and-add method [84]. The model has been updated to include equalization filters for synthesizing the effect of engine load on the emitted sound, induced by current vehicle operating parameters such as speed, acceleration, gear and road slope [114]. In later work, Maillard used this technique to assess noise exposure along pedestrian and bike paths in urban areas. Listening tests rated the perceived realism of the auralizations on a scale from 0 to 10, with an average score of 6.9 [115].

Combining input spectra generated by using the SPERoN model, which predicts third-octave band sound pressure levels of tire/road noise in the range from 315 Hz to 2000 Hz [15], with the aforementioned LISTEN demonstrator, Forssén produced pass-by auralizations by enhancing this limited spectral range by using backpropagation of broadband pass-by recordings in a horizontal plane [51]. Hoffmann psychoacoustically compared the similarity of this hybrid model and pass-by recordings in a listening experiment regarding loudness, sharpness, roughness and pleasantness [71].

In the foreground of organizing comprehensive knowledge exchange between urban noise researchers and urban planners, a European training network was funded by the European Commission in the SONORUS project (2012-2016). A subproject implemented a pre-compiled, web-based listening test using combined auralization and visualization at a virtual model of the Piazza Vittoria in Naples, Italy [85]. However, limitations resulted from sound source and propagation simplifications, e.g. by treating the traffic and natural sound sources as point sources without directivities or variations and by neglection of air absorption and diffraction.

Hornikx addresses questions regarding the relevance to include various propagation effects to wave-based urban sound simulations and their computational methods [72]. Related to this work using wave-based urban auralization approaches, Georgiou [56] investigated in his PhD thesis an incorporation of directivity patterns to a pseudo-spectral time domain method (PSTD) simulation framework. For passenger car pass-by auralization, generically synthesized free-field signals of a passenger car combined with directivity patterns from the Harmonoise model are convolved with PSTD-simulated impulse responses (IR) from a street canyon at pre-calculated emission and immission points [54] – and in another paper measured IR, respectively [55].

Pieren focused his PhD thesis on the auralization of environmental acoustical sceneries [132]. Among models for railway and wind turbine noise, spectral modeling synthesis is described to generate emission signals for auralization of a combustion engine and characteristics of accelerated passenger cars [134]. Fu presented in his PhD thesis a framework combining integrated sound design features of a game engine with additional visualization [52]. In this approach, different tools – Unity3D, Pure Data, Wwise, and Matlab – are concatenated with the underlying vehicle pass-by model being based on the horizontal directivity patterns from the Harmonoise model [87] and the propagation effect simulation assuming a first order image source of a sound hard ground. In his PhD thesis, Alkmim [6] used anechoic laboratory recordings of static cars and drones to compare their auralization in combination with different sound propagation rendering techniques.

Air-traffic auralization frameworks

Several research groups develop tools dealing with the auralization of aircraft noise. The most notable differences between these frameworks are underlying

directivity models, propagation algorithms, sound reproduction (and optional visualization) possibilites and real-time capabilities.

The Netherlands Aerospace Centre (NLR) develops the Virtual Community Noise Simulator (VCNS) [3]. The system is conceptualized to inform communities on the effects of aircraft noise for their environment including special setups such as drones. In order to multimodally extend the experiments, it enables three-dimensional modeling of urban scenarios based on open data for visualization. The NASA Auralization Framework (NAF) is developed by a research group at NASA Langley Research Center [9] and follows a modular apporach that is extensible through a plugin architecture. The swiss Federal Laboratories for Materials Science and Technology (Empa) is working on AURAFONE [133]. The Carmen software module is developed by the Office national d'études et de recherches aérospatiales (ONERA) and integrated into the Infrastructure d'Evaluation de Systèmes de Transport Aérien (IESTA) framwork [149]. In a round robin test, auralizations from the latter three frameworks were compared for a common set of predefined scenarios [144].

A real-time capable rendering module was developed at the Institute for Hearing Technology and Acoustics (IHTA), RWTH Aachen University [1]. It is based on the open-source auralization framework Virtual Acoustics (VA) [74]. Its algorithms are written in C++ and enable reproduction on listener-tracked loudspeaker-based and headphone-based systems. It includes a ray tracing algorithm for inhomogeneous, moving atmospheres [152] and produces binaural or higher-order Ambisonics (HOA) signals. Also based on VA, Sahai et al. [147] implemented an interactive simulation of aircraft noise for virtual environments. In their system the user is able to listen to and see the aircraft while moving freely through a Cave Automatic Virtual Environment (CAVE).

Rail-traffic auralization frameworks

Rail-traffic noise is dominated by rolling noise which is generated by wheel and/or rail roughnesses at their contact patch. The most elaborated models for the generation of wheel/rail rolling noise were implemented in the TWINS package (Track–Wheel Interaction Noise Software) [171]. It predicts emission spectra of wheel, rail and sleepers in one-third octave bands between 100 Hz and 5 kHz. Additionally, in course of the Harmonoise/Imagine projects a rail a parametrical model for railway traction noise was developed, being dependent on operational

[1] Auralization examples available at www.akustik.rwth-aachen.de/go/id/dzhe

parameters such as driveshaft speed and power settings, and duty cycles of compressors, fans and valves [31]. Klemenz implemented a synthesizer for traction noise and discussed its psychoacoustic optimization potential for practical application [93]. It was found that subtractive synthesis using one-third octave band filtering of broadband noise is not sufficient to plausibly auralize rail-traffic noise in all its nuances, as it is not able to recreate the resonances that characterize the timbre of the wheel/rail interaction noise [135]. The main goal of Kacem's dissertation is to implement an auralization model adapted to the physics of noise to improve sound realism [88].

The auralization frameworks discussed cover various types of traffic noise. They employ standardized and largely validated model catalogs to predict the emitted sound source levels and spectra, sometimes considering their behavior at different operating points. However, to make these sound predictions audible and furthermore considering their spatial properties, additional dimensions are required. There is a lack of standardization in this area, leading to the use of different analysis and synthesis techniques in auralization practice for each individual case. Generally, traffic sound source syntheses in existing literature are tailored for specific use cases and only consider some of the required dimensions. Thus, chapter 2 will systematically introduce which dimensions should ideally be considered, explore existing modeling concepts, and coherently discuss techniques for measurement, analysis, and synthesis to make traffic noise sources audible.

1.3 Contribution of this thesis

This dissertation focuses on the synthesis of traffic sound sources and its integration to dynamic, real-time auralization frameworks. For their systematic characterization, four source dimensions are identified and introduced in section 2.4. Today, although there are numerous source models for different traffic sound sources available, they still lack completeness regarding these dimensions.

a) Temporal characteristics

The core problem of three-dimensional acoustical source characterization is the decomposition of sound measurements or numerical simulations into source and propagation components. On the source side, elementary oscillators, noise generators, filters, temporal interaction effects have to be identified and implemented in real-time capable synthesizers, and corresponding directivity patterns have to be calculated. Existing air-traffic source models particularly lack modulation

parameters in the temporal dimension, such as fan noise and drone emissions. In contrast, in the domain of road-traffic source models, although effects from the temporal dimension are modeled, the parameterization of driving speed in the spatial dimension is often neglected. In the context of this dissertation, missing parameters in source models are identified and enhanced by according implementations.

b) Directional characteristics

Another fundamental contribution of this thesis, is a remote sensing algorithm using an inverse measurement method for calculating three-dimensional directivity patterns of moving sound sources (see section 2.5). The subsequent decomposition into the partial sound sources was demonstrated by adaption to different road-traffic and air-traffic sound sources (in chapters 3 and 5). This approach in characterizing moving sound sources is novel because they cannot be measured in a fully surrounding microphone array, as is the state of the art for acoustic speech, instrument, or household appliance sources, and their radiation through aeroacoustic turbulent layers cannot be adequately simulated to date. Using this technique, speed-dependent directivity patterns of road-traffic vehicles are measured, modeled and published in an open database. Finally, it discusses the source model implementation to an urban sound auralization framework to simulate the sound transmission in a built urban surrounding including complex propagation effects and investigates their acoustical convincingness in virtual reality by evaluating immersion and presence in chapter 6.

c) Complexity reduction

This work also demonstrates complexity reduction of simulated subsonic jet noise using spherical harmonics decomposition and provides a proof of concept for far-field simulation of an aeroacoustic sound source by reducing the complexity of high-resolution near-field data (in chapter 4). Within this work, the often-neglected temporal patterns deviating from the steady-state of a sound in sound source auralization models are addressed, resulting in the extraction of modulation parameters and the implementation of exemplary auralization models.

2

Fundamentals of sound source characterization

The general goal of source characterization as basis for use in auralization is the description of a sound source operating under free-field conditions, i.e. excluding any influences from room reflections or non-absorbing boundaries. This chapter introduces the strategies, dimensions and techniques for characterizing sound sources used throughout this thesis. After briefly introducing the auralization concept, the modeling approaches for sound sources are systematically presented, followed by information on the integration and calibration of the models in auralization tools.

2.1 Traffic sound auralization chain

Auralization is the method of creating audible sound from computational acoustic data. Auralizations usually follow a three-step paradigm of discrete models for sources, wave propagation and receivers [177]. The block diagram in Figure 2.1 shows the signal chain for traffic sound auralization using the open-source auralization framework Virtual Acoustics (VA) [74]. In the scope of the VA framework, this thesis focuses on sound sources, represented by the green-colored input elements of the block diagram. When creating an auralization scenario, sound sources are defined by assigning source properties, directivities and time-variant spatial positions on a trajectory. Subsequently, the sound propagation characteristics between source and receiver in the form of a set of impulse responses are taken into account. The depiction emphasizes the necessity of a source model with quasi-freefield properties since all surface reflections are separately simulated by an outdoor propagation module (calculating the effects according to the ISO 9613-2 standard [82]), and an urban sound propagation module [151]. Finally, aurally correct equalization for the reproduction system can be included, such

as head-related transfer functions (HRTF) and equalization of the reproduction system such as headphone-to-ear transfer function (HpTF).

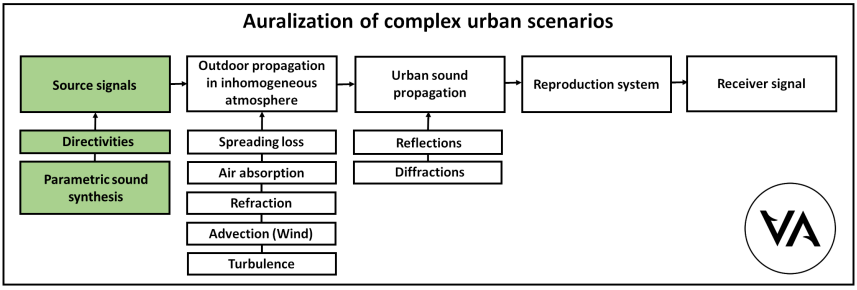


Figure 2.1: Signal chain for environmental noise auralization in VA.

For more extensive explanations regarding the theory and models of outdoor sound propagation, the reader is referred to the corresponding textbooks and ISO standards with respect to outdoor sound propagation [98], [8], [81], [82]. Focusing on urban sound propagation including the calculation of directivity-dependent reflections and diffractions of building surfaces the reader is referred to [41] and [151]. The framework enables rendering to multiple spatial headphone- or loudspeaker-based reproduction systems. Due to VA's server-based nature, the framework can be connected to common gaming visualization engines. Referring to the highlighted blocks in Figure 2.1, the concept for a sound source auralization model is emphasized in Figure 2.2.

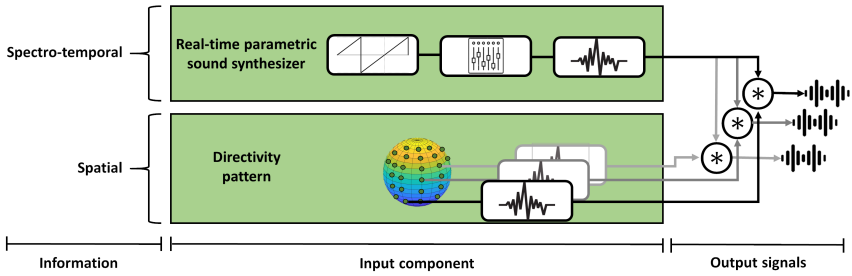


Figure 2.2: Concept for sound source auralization models.

The spectro-temporal information is encoded into the omnidirectional input component and convolved with spatial information from the directivity. It is imple-

mented as procedural or wavetable synthesizer and produces a single-channel audio stream, therefore producing the spectro-temporal patterns which are observable from all radiation directions. This audio stream is convolved with higher-order spatial components that are stored in a spatially discretized form of magnitude spectra or impulse responses on a spatially discretized grid around the center of the sound source. The multifold of output signals resulting from convolution describes the sound field on a virtual three-dimensional hull surface around the center position of the sound source.
With regard to ideal boundary conditions, the overall goal of an sound source model for use in virtual acoustics is the description of the sound field under free-field conditions. In practice, this condition can rarely be fulfilled for (moving) traffic sound sources during operation. Therefore, recordings obtained with non-ideal boundary conditions must be compensated for wave propagation effects. In course of this thesis, in-situ source recordings will therefore be postprocessed by using time-variant wave backpropagation (cf. examples for road-traffic sound sources in section 3.3 or drones in section 5.4).

2.2 Sound source model calibration

In most auralization frameworks, a sound source's playback level is calibrated at the very end of the auralization chain by reproducing a target sound pressure p_{cal} at a virtual receiver in the free field at a defined distance from the source, cf. ② in Figure 2.3. In opposite, the auralization models presented throughout this thesis comprise an iterative calibration option directly at the source based on a target sound power P_{cal}, cf. ①. From a technical perspective, as shown in Figure 2.3, the calibration procedure aims to reproduce a source's sound power P_{cal} or target sound pressure at a receiver p_{cal} by computing a single calibration gain factor G_{cal} for application to the emission signal path.
The advantages of using sound power calibration is based on its frequency- and direction-independent properties, therefore being the most convenient single-valued measure for comparing acoustical source emissions. Moreover, its calculation is standardized for different sound field conditions, such as the free field [78], over a reflection plane [79] or in a reverberant environment [80]. In the first case, radiated sound power level L_{w} is calculated based on the expression for a measurement on a surface with segments having unequal areas:

$$L_{\mathrm{w}} = 10 \cdot \log_{10} \left[\frac{1}{\mathrm{S}} \sum_{\mathrm{i}=1}^{\mathrm{M_N}} \mathrm{S_i} \times 10^{0.1 L_{\mathrm{p,i}}} \right] \mathrm{dB} + 10 \cdot \log_{10} \left(\frac{\mathrm{S}}{1\,\mathrm{m}^2} \right) \tag{2.1}$$

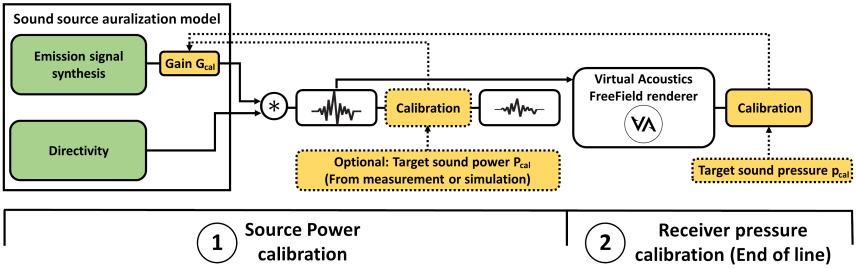


Figure 2.3: Source calibration approaches: Computation of a calibration gain factor G_{cal} from 1) a target sound source sound power P_{cal} or 2) a target sound pressure p_{cal}.

with $L_{\mathrm{p,i}}$ being the sound pressure levels at the ith measurement position on the sphere, S_i being the according area segment on the spherical surface and S being the total surface area.
Resulting in a single-valued measure, sound power is suitable as a starting point for both, the level calibration of sound sources as well as the dynamic object parameterization of a scene in the auralization practice. The calibration procedure is dependent on three prerequisites:

1. The digital full-scale value of the source signal synthesis must be referenced to a defined sound pressure (in the unit of Pascals).
2. The according directivity data are stored in form of linear-scaled amplitude values.
3. The area represented by each directivity data point on the unit sphere is known.

The procedure is as follows: An overall calibration gain factor G_{cal} (small yellow box) is defined as the ratio of the target sound source power P_{cal} and uncalibrated sound source power P_{uncal}

$$G_{\mathrm{cal}} = \frac{P_{\mathrm{cal}}}{P_{\mathrm{uncal}}} \tag{2.2}$$

with P_{uncal} being calculated by

$$P_{\mathrm{uncal}} = \sum_{n}^{N} I_n \cdot w_n \tag{2.3}$$

with the area weight w_n being the proportional fraction of each represented directivity data point on a unit sphere. The sound intensity I_n is computed for

each directivity surface element n as

$$I_n = \sum_{f}^{F} \frac{d_{n,f} \cdot s_{n,f}^2}{Z_0} \tag{2.4}$$

with d_f being the directivity's linearly scaled amplitude factors of each frequency bin, s_f being a frequency bin's amplitude of the uncalibrated emission signal and $Z_0 = \rho_0 \cdot c$ being the characteristic wave impedance.

2.3 Modeling strategies

2.3.1 Black and white-box models

Auralization modeling of real-world systems involves building computer models from measurements or physical-based equations. Analogous to a system-theoretical point of view, auralization source modeling can generally be classified into two different approaches:

1. White-box modeling: involves developing an auralization model based on the physical interactions of the object's sound radiating elements. These equations are derived from the physical principles governing the dynamics of these systems.
2. Black-box modeling: involves taking calibrated measurements of the sound emission under laboratory conditions, and creating a digital system that replicates perceptually relevant aspects of the sound source's behaviour.

2.3.2 Bottom-up and top-down approaches

The selection of a specific modeling approach depends on various factors, such as characterization dimension (introduced in section 2.4), the complexity of its system dynamics and the time budget, respectively. Therefore it must be individually chosen for each partial sound source. As methodologically shown in Figure 2.4, the associated modeling processes can be classified into top down and bottom up. A technical dimension of each modeling approach is the computational cost of the digital signal processing (DSP). The cost of each DSP algorithm is determined by two factors: First, the complexity of the process being modeled and second, the ease with which the resulting model can be implemented in the discrete-time domain (see according synthesis techniques in section 2.6). Furthermore, however

nowadays a minor issue, the algorithmic efficiency matters due to its memory requirements.

2.3.3 Models for wave-based vs. ray-based auralization frameworks

Generally, two sound field simulation concepts – wave-based and geometrical approaches – can be distinguished for auralization, which in turn determines the mathematical form of sound sources as input data.

Geometrical Acoustics (GA) assumes sound wave propagation in a fluid as a ray or particle, analogous to geometrical optics, where light is regarded as a ray which travels in straight lines in a homogeneous medium and changes its direction at the boundaries of different materials [150]. One of the main advantages of GA is based on its wavelength-independent nature, with sound path tracing algorithms producing abstract intermediate results, which can be computationally efficiently postprocessed across the entire frequency range of interest. However, GA methods largely ignore wave effects, leading to inaccuracies when these effects become significant, especially at low frequencies. Compared to wave-based methods, GA significantly reduces computational cost since only the contribution of each sound source to a specific receiver location is computed. Therefore, GA is ideally suited for virtual acoustic environments that (unlike noise maps) need calculations at selected locations, considering a limited number of sources. The effort to implement sound source properties in GA is rather low, as these merely need to be stored in the form of frequency domain coefficients with a defined spectral and spatial resolution, which for reproduction can be quickly retrieved from simple data tables.

In contrast to GA, wave-based methods describe the behavior of sound by means of wave mechanics. Consequently, sound propagation is computed by solving the wave equation, which contains partial differential terms and necessitates idealized environmental conditions. Being computationally heavy especially for three-dimensional problems, this approach is being limited in the context of urban traffic auralization to well-aligned problems including strict patterns in the built environment, for example, a raster of buildings or a repetition of street canyons. For improved practicability, numerical computations discretize dimensions like space and frequency or time in order to linearize sections of the propagation problem. Urban sound propagation has been simulated using Finite-Difference Time Domain (FDTD) [174], Pseudo-Spectral Time Domain (PSTD) [175] and Fast-Multipole Boundary Element (FM-BEM) methods [89]. Compared to GA, wave-based methods exhibit different advantages and disadvantages regarding

traffic sound auralization. Being able to accurately simulate low-frequency sound waves, aligns well with incorporating diffraction in urban environments. However, the high memory demand and computational complexity limits the upper frequency, making it even difficult to reach the mid-to-high frequencies of human hearing. The effort to accurately implement sound source properties for wave-based simulations is rather high, as source data from multiple discrete grid points on an extended surface must be transformed at first into an equivalent source model by means of spherical harmonics decomposition, monopole or multipole decomposition. Furthermore, uncertainties in acoustic boundary conditions significanty affect the precision of wave-based calculations. After all, for purely theoretical reference cases these methods are essential but in practical cases they suffer from uncertainties of modeling and boundary conditions.

2.4 The 4 dimensions of source characterization

Acoustical source characterization comprises the description of a sound source concerning four dimensions: Spectral, temporal and spatial emission properties as well as its parameterization, i.e. the dynamic change of these properties depending on the source's state of operation (see Figure 2.4). It should be noted that traffic sound emission is a mixture of multiple partial sound sources. Each partial sound source therefore has to be individually modeled. Since the information on partial sources practical measurement setups during operation cannot be directly obtained, source signal separation is an important step in the analysis.

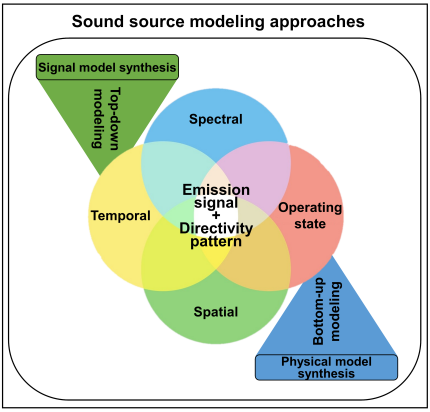


Figure 2.4: Modeling process overview and the four dimensions of source characterization.

2.4.1 Spectral properties

A crucial aspect of characterizing a sound source in frequency domain involves measuring the energy it emits across various frequency ranges. This information is essential for syntheses to precisely replicate the original sound's timbre. The relevant spectral information to be analyzed depends on the regarded signal type. For tonal components, the amplitude, frequency, and phase of each harmonic oscillator are required for additive synthesis (cf. section 2.6.1). For stochastic signals, the amplitude in each frequency bin is used for subtractive synthesis. In practice, due to the limited frequency resolution of human hearing, it is often adequate to store information at a coarser resolution, such as in one-third octave bands. For transient signals, which inherently have an infinitely broad spectrum, additional data on the decay over time at each frequency bin is needed.

Power spectral density

Noise signals are characterized as a series of random variables, making it a stochastic process. When estimating the frequency-dependent power of stochastic signals using multiple periodograms, each repetition exhibits significant variance. Therefore, power spectral density (PSD) is usually estimated using Welch's method [180]. By averaging of K segment estimations this method reduces the variance by a factor of 1/K. Based on the normalized periodogram of the m-th block of a time-discrete sound pressure series $p[n]$

$$P_{\mathrm{m}}(f) = \frac{1}{M} \left| \sum_{n=0}^{N-1} p_{\mathrm{m}}[n] e^{-\mathrm{j}2\pi kn/N} \right|^2 \tag{2.5}$$

the Welch PSD estimate is obtained by averaging of K blocks

$$\mathrm{PSD}(f) = \frac{1}{K} \sum_{m=0}^{K-1} P_{\mathrm{m}}(f). \tag{2.6}$$

Throughout this thesis, these values are expressed as PSD in decibels

$$\mathrm{PSD}_{\mathrm{dB}}(f) = 10 \cdot \log_{10} \left(\frac{\mathrm{PSD}(f)}{P_{\mathrm{ref}}^2} \right) \tag{2.7}$$

with $P_{\mathrm{ref}} = 20\mu\mathrm{Pa}$.

2.4.2 Temporal properties

Besides adequate parameterization, the production of a plausible, perceptually convincing sound is a major challenge of sound source characterization. Temporal patterns deviating from the steady-state of a sound, alas, are often neglected in sound source auralization models, resulting in sterile and implausible syntheses. Two families of temporal phenomena are relevant, modulations and transients.

To achieve perceptually accurate sound syntheses, it is essential to account for modulations in the sound field, whether they originate at the sound source or occur during sound propagation. These phenomena have been demonstrated in both scenarios: for drones in the first case [38], and for aircraft [107] and wind turbine noise in the latter [24].

Any abrupt changes in a signal, such as discontinuities in the waveform, its slope or curvature, are considered as transients. Mathematically, a transient is defined as a signal whose Fourier expansion requires an infinite number of sinusoids. Conversely, a signal that can be represented by a finite number of sinusoids constitutes a steady-state signal. Analyzing and synthesizing transients is challenging. Technically, transient events can practically be defined as sudden broadband events in an otherwise stationary signal. For instance, in engine noise, backfire sound is transient, while revolving piston sound is a steady-state signal.

The process of analyzing modulated sounds is briefly described in the following and will be used in chapter 5.

Extraction of oscillator modulation parameters for sound synthesis

In the case of traffic sound sources, fluctuations are audible by means of frequency modulations (FM) and amplitude modulations (AM) in tonal sound sources, such as fan noise and electric motor emissions. The time-variant FM and AM parameter information can be directly extracted from the bandpass-filtered tones by computation of the discrete-time analytic signal, that in turn, is the basis for the computation of an instantaneous magnitude and phase [21]. An analytic signal $\tilde{s}(t)$ is given by the complex sum of the original signal and an imaginary part equal to its Hilbert transform by

$$\tilde{s}(t) = s(t) + is_H(t) \tag{2.8}$$

with the Hilbert transform being defined by

$$s_{H(t)} = \mathscr{H}[s(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} dt' \, \frac{s(t')}{t - t'} \tag{2.9}$$

After bandpass filtering around each oscillator frequencies, the time-varying amplitudes for synthesis are obtained from the envelope e_n for all N tonal oscillators by calculation of the instantaneous magnitude of the analytic signal by

$$e_n(t) = |\tilde{s}_n(t)| \tag{2.10}$$

The time-variant description for instantaneous frequency of each harmonic oscillator $\phi_N(t)$ is obtained by differentiation of the analytic signal phase

$$\phi_n(t) = \frac{1}{2\pi} \frac{d}{dt} [arg \, \tilde{s}(t)] \tag{2.11}$$

Assuming normal distribution, each individual tone's statistical information about amplitude and frequency modulation rate and depth is stored in form of variance values σ^2.

2.4.3 Spatial properties

In general, sound sources emit sound in all spatial directions. If sound propagation is not obstructed by boundaries, such as tire/road noise at the asphalt contact patch, a source characterization must account for variations across the multidimensional spatial angles by means of a so-called directivity pattern. The effect of directivity patterns is particularly noticeable with highly directed sound sources, such as the frontal sound of a trumpet compared to the rearward sound. An extensive database of 41 musical instruments was created by Shabtai et al. [158]. Additionally, spatial sound source characterization should consider the extent of the sound source and the distance of observation.

The behavior of acoustic energy from a sound-emitting object varies significantly with the observer's distance. Because wavelength depends on frequency, the far field onset also varies with frequency. The far field is defined as the region where sound pressure and acoustic particle velocity are in phase, and where sound pressure level decreases by 6 dB for each doubling of distance from the source. In the far field, the source appears as a point with no discernible size, and the spherical sound waves expand enough to approximate the wavefront as a plane wave. Here, sound intensity follows the inverse square law, resulting in reliable and predictable single microphone recordings. Assuming no reflections,

the sound pressure level decreases by 6 dB with each doubling of distance from the source. In contrast, in the near field, sound waves behave more complexly since there is no fixed relationship between sound pressure and distance, making single microphone measurements difficult and inconsistent. Very close to the emission area of complexly vibrating surfaces, sound energy circulates with the radiating surface, forming so-called evanescent waves that do not propagate away. A little farther out, a fraction of the sound field still circulates while another fraction propagates. Accurate near-field measurements typically require multiple microphones to capture the energy of both circulating and propagating waves.

Another critical aspect of characterizing a sound source's spatial behavior is due to its spatial extension. Usually, the source region of technical sound sources, such as the four wheels of a car or the turbulent sound field of a jet engine, is distributed. This has a significant influence on the waveform's shape, determining whether it is spherical, cylindrical, plane, or a mixture.

To further reduce the complexity of a sound field, spatial decomposition methods, such as spherical harmonics transform or mono-/multipoles [128], can be applied. The first is briefly described in the following and will be used in chapter 4.

Sound field decomposition using spherical harmonic transform

Given are the time-domain sound pressure signals p_i of a sound field from a (virtual) spherical microphone array. Since the sound pressures p_i are on a spherical surface, a mathematical expansion of the simulated discrete sound field points on the sphere into spherical harmonics exists. The real-valued spherical harmonics $Y_\mathrm{n}^\mathrm{m}(\Theta, \phi)$ of (positive) order n and degree m are defined as [181]

$$Y_\mathrm{n}^\mathrm{m}(\theta, \phi) = \sqrt{2}\sqrt{\frac{2\,(2\mathrm{n}+1)}{4\pi}\frac{(\mathrm{n}-\mathrm{m})!}{(\mathrm{n}+\mathrm{m})!}}\,P_\mathrm{n}^\mathrm{m}(\cos\theta)\cos(\mathrm{m}\phi) \tag{2.12}$$

with P_n^m being the associated Legendre polynomial. The real-valued spherical harmonic basis functions up to second order are shown in Figure 2.5. The spherical sound pressure data $p(r, \phi, \theta)$ at discrete points, defined by their azimuth ϕ and elevation θ angles at radius r, can be expressed for each wavenumber k by the sum of the SH basis functions using the discrete inverse spherical harmonic transform (ISHT)

$$p(r, \phi, \theta, k) = \sum_{\mathrm{n}=0}^{\infty}\sum_{\mathrm{m}=-\mathrm{n}}^{\mathrm{n}} Y_\mathrm{n}^\mathrm{m}(\theta, \phi)\,\mathrm{f}_\mathrm{nm}(r, k) \tag{2.13}$$

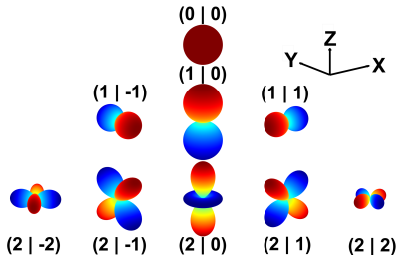


Figure 2.5: Real-valued spherical harmonic basis functions up to order and degree (n|m).

with $\mathrm{f_{nm}}$ being according discrete spherical wave spectrum coefficients that represent the source strength of each SH basis function. Since the sound field is simultaneously simulated at multiple points, a matrix system (with matrix $\mathbf{Y}$ and vectors $\mathbf{p}$ and $\mathbf{f}$) must be solved for computing the coefficients $\mathbf{f}_{\mathrm{nm}}$ by using the inverse $\mathbf{Y}^{-1}$ of the discrete spherical harmonic basis matrix

$$\mathbf{f}_{\mathrm{nm}} = \mathbf{Y}^{-1} \cdot \mathbf{p}. \tag{2.14}$$

The spherical wave spectrum $\mathrm{f_{nm}}$ must be order-limited for an exact ISHT, i.e. $\mathrm{f_{nm}} = 0$ for $n > N$. Hence, for the discrete spherical harmonic transform of a specific sound pressure $p(\phi_l, \theta_l)$

$$\mathrm{f_{nm}} = \sum_{\mathrm{l}=1}^{L} p(\phi_l, \theta_l)\, Y_{\mathrm{n}}^{\mathrm{m}}(\theta_l, \phi_l) \tag{2.15}$$

the Shannon-Nyquist criterion must be fulfilled, i.e. lobes in the directivity that are beyond the spatial resolution of the array must be avoided.

2.4.4 Operating state properties

The spectro-temporal characteristics as well as the directivity patterns of traffic sound sources are influenced by their operating condition, which are recorded as metadata for the source model parameterization. The complexity of the sound generation mechanism correlates with the complexity of the parameterization: Obvious parameters of the operating state include, for example, vehicle driving speed and an engine's rotational speed. Not so obvious in contrast, in case of aeroacoustic sound sources, both nozzle and propeller geometries, as well as fluid dynamic properties like the Strouhal number and medium viscosity, might be important.

2.5 Postprocessing methods

2.5.1 Time-variant wave backpropagation

As traffic sound sources are usually moving, the extraction of their directivity patterns is challenging. In order to obtain directivity patterns with quasi-freefield properties, a recorded source signal of a pass-by (fly-over) or pass-through (fly-through) array must be post-processed for inversely compensating several sound wave propagation effects (blue block in Fig. 2.6).

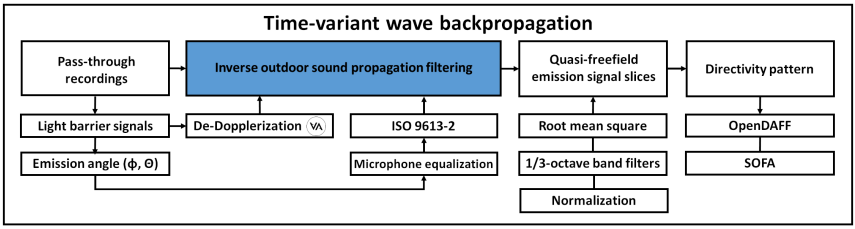


Figure 2.6: Block diagram of the signal postprocessing chain.

This process compensates for wave propagation effects between the receiver point and a virtual (hemi-)spherical surface around the moving source. Finally, it reconstructs the spectro-temporal properties of the recorded sound source in a virtualized close distance. In practical application, the wave backpropagation can be interpreted as projection of signals recorded at a 2-dimensional microphone array onto a 3-dimensional reconstruction grid around a vehicle's center point (see practical array realizations for passenger cars in Fig. 3.4 and drones in Fig. 5.3).

The algorithmic idea is exemplary illustrated in Figure 2.7. The spectrogram of a recorded pass-by consists of four sinusoidal tones (left), plotted as frequency spectrum (y-axis) over source-receiver distance d (x-axis). It shows distance-dependent attenuation of the high frequency range (due to air absorption) A_{atm}, a Doppler shift f_{Doppler} and the geometrical distance law-dependent loss A_{div}. Not explicitly shown in the figure, a reflection from the ground A_{gr} and an angle-dependent measurement microphone response A_{mic} must be compensated too, if relevant. After inversely compensating these effects, the resulting spectrogram (right) is free from Doppler shift, air attenuation and geometrical spreading loss.

Referring to Figure 2.6, the Doppler effect in the recorded data is compensated for each individual channel by applying dynamic resampling based on the speed

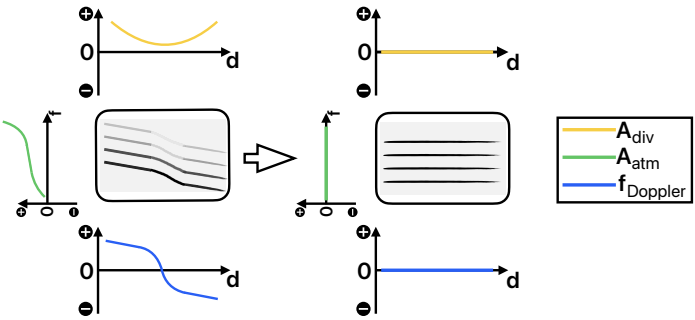


Figure 2.7: Schematic illustration of the time-variant wave backpropagation algorithm on a pass-by/fly-over recording spectrogram: State of parameters (green, blue and yellow curves) in recorded spectrogram (left) and compensated spectrogram (right).

information deduced from light barrier signals. For that purpose, the FreeField-Renderer module by the publicly available auralization software Virtual Acoustics [74] is used throughout this thesis. The directivity patterns are exported to the widely used data formats OpenDAFF [129] and SOFA [116], usually in one-third octave band magnitude spectra. The exact attenuations are computed according to ISO 9613-2 in order to calculate the total attenuation A_{tot} in decibels by adding partial contributions [82]:

$$A_{\text{tot}} = A_{\text{div}} + A_{\text{atm}} + A_{\text{gr}} + A_{\text{bar}} + A_{\text{misc}} \tag{2.16}$$

where A_{div} = attenuation due to geometrical divergence
A_{atm} = att. due to atmospheric absorption
A_{gr} = att. due to the ground effect
A_{bar} = att. due to diffraction
A_{misc} = att. due to miscellaneous other effects (e.g. turbulence).

When a receiver can be assumed to be in the acoustic far field, the source can be approximated as a point source. In this case, due to the inverse-square law, the field-field sound pressure is frequency-independent and just depends on the distance between sound source and receiver which is represented as A_{div} in eq. (2.16). All other terms are functions of frequency. The atmospheric attenuation A_{atm} depends on several physical parameters:

$$A_{\text{atm}} = \text{f}(f, RH, T, p_{stat}) \tag{2.17}$$

where f is frequency, RH is relative humidity in %, T is temperature and p_{stat} is static pressure.

The ground material determines amplitude and phase of the (complex) ground reflection coefficient. The superposition of direct and reflected wave at the receiver is represented by A_{gr}. Attenuations A_{misc} account for miscellaneous effects to the overall sound pressure at the receiver, such as fluctuations due to coherence variations of the incoming sound waves.

2.5.2 Signal separation

Signal separation using non-negative matrix factorization

Suppose, the spectrogram $\boldsymbol{V}$ (see for example Fig. 4.25) is a superposition of different source contributions from stochastic and tonal components. Then, technical sounds such as vehicle pass-bys can be decomposed into tire, exhaust and engine noise. In order to obtain the emission signals and/or directivity patterns for each source, the compensated spectrogram is separated into tonal and noisy components using non-negative matrix factorization (NMF) [100]. The NMF algorithm is applied to the transposed Short Time Fourier Transform (STFT) magnitude spectrogram $\boldsymbol{V} = |X(m,k)|^{\mathrm{T}}$ of a time signal $x(t)$, where k is the number of spectral bins at m time frames. Application of NMF means to factorize $\boldsymbol{V}$ by

$$\boldsymbol{V} \approx \tilde{\boldsymbol{V}} \stackrel{\mathrm{def}}{=} \boldsymbol{W}\boldsymbol{H} \tag{2.18}$$

with columns of $\boldsymbol{W}$ being spectral basis function for each source (templates), and the rows of $\boldsymbol{H}$ containing time-varying gains in each time frame (activations). All matrices represent energies and therefore are nonnegative.
From mathematical point of view, the factorization of equation (2.18) can be described as a constrained minimization problem, minimizing a measure of fit between $\boldsymbol{V}$ and $\boldsymbol{W}\boldsymbol{H}$ by

$$\min_{\boldsymbol{W},\boldsymbol{H}\geq 0} D(\boldsymbol{V}|\boldsymbol{W}\boldsymbol{H}) = \sum_{fn} d([\boldsymbol{V}]_{fn}[\boldsymbol{W}\boldsymbol{H}]_{fn}) \tag{2.19}$$

Like Principal Component Analysis (PCA) – that minimizes a quadratic cost function between $\boldsymbol{V}$ and its approximate $\boldsymbol{W}\boldsymbol{H}$ where all matrices are treated real-valued – NMF minimizes a distance measure d but is subject to nonnegativity of the values of $\boldsymbol{W}$ and $\boldsymbol{H}$. The distance measure used in this thesis, from the family of β-divergences – whose advantages in spectral audio signal processing

are more comprehensively described in [46] – is the Itakura-Saito cost function

$$d_{\beta,IS}(x|y) = \frac{x}{y} - \log\frac{x}{y} - 1 \tag{2.20}$$

When the spectro-temporal characteristics of different noise sources can be clearly distinguished (as fan noise compared to jet noise in the aircraft noise emission), spectral templates can be prescribed for initialization of matrix $\boldsymbol{W}$. The minimization problem is reduced to the update of matrix $\boldsymbol{H}$ during every iteration step

$$d_{\beta,IS}(\boldsymbol{H}) = \sum_{fn}\left(\frac{v_{fn}}{[\boldsymbol{WH}]_{fn}} - \log\frac{v_{fn}}{[\boldsymbol{WH}]_{fn}} - 1\right) \tag{2.21}$$

where v_{fn} denotes the power spectrogram defined by $v_{fn} = |x_{fn}|^2$.
The matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ are postprocessed after NMF application with alpha-related soft masks [109] to avoid audible artifacts due to binary (hard) masks. Finally, the backtransformation of the separated STFT magnitude spectra to time-domain signals is calculated for also considering phase reconstruction using the Grffin-Lim algorithm [59]. When characteristic spectral properties of the partial sound sources in the mixture are known, such as jet and fan noise in a fly-over measurement, their models can be translated to spectral templates in advance so that the NMF can be applied in supervised form, as mathematically described in equation (2.21). Therefore, templates for matrix $\boldsymbol{W}$ in equation (2.21) with characteristic spectra of sound sources are added to the decomposition algorithm as additional information.

2.6 Sound synthesis techniques

The source signal (cf. upper block in Fig. 2.2) is produced using a sound synthesis technique specific to the source emission. Numerous methods have been proposed, which can be classified into two main families. The distinction is made between physical models and signal models. Synthesis using physical models involves generating sound by modeling the physical phenomena that produce it, making in-depth knowledge of the sound-generation mechanisms indispensable. In contrast, signal models are concerned with sound from a signal point of view. Here, parameters of the sound signal (i.e. spectrum, phase, amplitude, etc.) are used to reproduce the sound. Generally, these parameters are determined during an analysis phase based on recordings. For plausible sound rendering, the recordings must be of such quality as not to distort the nature of the reference sound. For more extensive explanations of sound synthesis techniques, the reader

is referred to the textbooks by Raffaseder [140] and Zölzer [188]. The following section briefly describes the main types in each family.

The choice of a particular synthesis technique is significantly influenced by the noise characteristics as well as possible requirements for real-time constraints. For example, fan noise and jet/airframe noise might rather be synthesized with additive and subtractive synthesis respectively, whereas an engine might rather be physically-modeled using waveguide synthesis. For basic principles and examples for dynamically synthesizing a variety of technical sound sources, the interested reader is referenced to the textbook by Farnell [44].

2.6.1 Signal model synthesis

As mentioned above, signal model synthesis considers the properties of the sound signal radiated by the system under study. This family comprises several techniques, most of which are based on a phase known as sound analysis. This phase aims to characterize the sound under study, which will facilitate the choice of the appropriate synthesis method. In this section, the most popular signal model synthesis methods are presented.

Additive synthesis

Additive synthesis is motivated by the Fourier series principle, stating that any complex periodic waveform can be decomposed into a sum of sinusoids, whose frequencies are integer multiples (so-called harmonics) of the fundamental frequency. Assuming quasi-stationary, such as the emitted spectrum of an electric engine at a fixed rotational speed, the amplitude, frequency and phase parameters are estimated for each sinusoid of a recording in an analysis phase. The additive synthesizer implements a set of oscillators, each being associated to a specific harmonic, and being defined by its amplitude, frequency and phase. The sum of the oscillators' contributions re-synthesizes the original signal.

Subtractive synthesis

A subtractive synthesizer reproduces the timbre of a sound using a source-filter approach. Therefore, a frequency-dependent attenuation filter subtracts energy from a frequency-rich source signal, such as white noise. Its main advantage is the ability to mimic complex timbres with high computational efficiency by using digital filter structures based on cascaded biquads.

Sample-based synthesis

Two similar sample-based methods, wavetable and granular synthesis, can be distinguished. A wavetable synthesizer utilizes a predetermined table of sampled waveforms, which limits the degree of control over the fundamental components of sound synthesis. In contrast, a granular synthesizer produces a continous audio stream from sequentially concatenating sound grains. Typically, a sound grain has a duration between 10 and 100 ms and is produced from a recording of a sound source during steady-state operation.

Frequency-modulation (FM) synthesis

This method modulates the frequency of a waveform (carrier) with another waveform (modulator), creating a third, complex waveform as an output. This method produces complex modulation products, depending on involved frequencies, modulation depth settings, number and inner connection of carriers and modulators.

2.6.2 Physical model synthesis

Synthesis by physical modeling consists in generating sound from the resolution of the set of equations describing the physical behavior of the system under study. The model's input parameters are the physical properties of the materials involved in sound generation. The main approaches to physical model synthesis are described below.

Digital waveguide synthesis

The basic idea behind digital waveguide synthesis is to model the medium in which the acoustical wave will evolve. An excitation signal is send into a delay line with the delay corresponding to the time taken for the wave packet to propagate from one end of the waveguide to the other. Frequency-dependent losses are introduced by digital filters. Despite computational cost-effectiveness, the sonic plausibility of the sound rendering is promising for synthesizing technical and musical sounds. The waveguide synthesis technique has been widely exploited to reproduce the sounds of various musical instruments [160] or engines [12].

Modal synthesis

The modal synthesis technique is based on mimicing the vibrational behavior of an object in terms of its resonance modes, which are assumed to be independent of each other and oscillating at a specific frequency. When excited, the response of a vibrating body is expressed as the sum of the contributions of these modes. For synthesis, each mode is defined by a set of modal parameters, such as resonant frequency, damping and amplitude. These modal data can be obtained either by solving equations describing the object's motion, or from experimental characterization measurements. This approach has been widely used to synthesize noises linked to interactions between objects, such as friction, rolling or shocks.

Numerical simulations

Numerical simulations characterize the behavior of a vibrating object by numerically solving the physical equations that describe it. First, the according system of governing partial differential equations, which combine spatial and temporal derivatives, are established. For solving the equation system in a discrete space, the original space is spatially or temporally discretized. For example, the Finite Element Method (FEM) and the Boundary Element Method (BEM) are often used to characterize the vibro-acoustical behavior of structures.

3

Passenger car auralization model

Parts of this chapter have been published in:
Dreier, C., Vorländer, M. (2025). Speed-dependent directivity patterns of road-traffic vehicles. The Journal of the Acoustical Society of America, ***157****(4), 2735–2749 [39], and*
Dreier, C., Hahn, J., Heck, J., Llorca-Bofí, J., Vorländer, M. (2022). Real-time vehicle pass-by noise synthesis for urban sound auralization. In: Proc. Int. Cong. Ac., Gyeongju, South Korea, 1–8 [33], and
Dreier, C., Vorländer, M. (2023). Road-traffic auralization: Modeling and synthesis of electric drives. In: Proc. Tecniacústica. Cuenca, Spain, 187–190 [37].

This chapter presents the design and implementation of a perceptionally plausible road-traffic auralization model for real-time capable and dynamic traffic auralization. According to the concept presented in Figure 2.2, the source model consists of two separate components, a emission signal synthesizer (mimicing the spectro-temporal properties) and speed-dependent directivity patterns (mimicing the spatial properties). The emission signal is produced from distinguished engine sound (section 3.6) and tire/road noise synthesis (section 3.7) models. Referring to the modeling strategies introduced in section 2.3, the directivity patterns and the tire/road noise emission models are obtained from a black-box and top-down approach, whereas the engine sound is obtained from a white-box and bottom-up approach. With regard to real-time constraints and parametrical controllability, the proposed synthesizer is implemented in C++ and relies on procedural audio synthesis using digital waveguide networks and spectral shaping. Striving for the auralization of urban scenarios in VR using multiple object instances, the model integrates well being computationally efficient.

3.1 Introduction

Acoustic emissions in urban environments are dominated by road traffic noise, such as emissions from passenger cars, motorcycles, buses and trucks. Vehicle exterior noise consists of speed-dependent contributions from tire/road, powertrain and aeroacoustic noise, each of those contributing to overall sound emission by characteristic directivity patterns. However, directivity pattern extraction of moving sound sources is a complex measurement problem, as discussed in the following. Therefore, noise mapping is based on simple analytical models, neglecting speed effects or further segment-specific emission characteristics of road-traffic sound sources. For illustrating the variation of pass-by frequency spectra of a passenger car at different driving conditions, Figure 3.1 shows three horizontal pass-by spectra of a single C-segment plug-in hybrid electric vehicle (PHEV). It compares the resulting emission at different non-accelerated driving speeds by keeping remaining vehicle operating parameters and trajectories constant. Furthermore, variations due to different drivetrain configurations are compared by using an electric driving mode with activated Acoustic Vehicle Alerting System (AVAS) as well as an internal combustion engine (ICE).

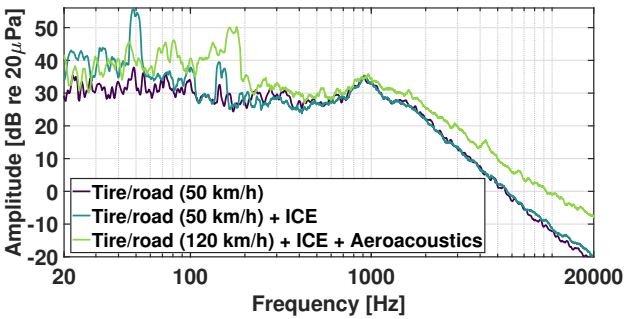


Figure 3.1: Typical pass-by frequency spectra of a C-segment plug-in hybrid electric vehicle (PHEV) at different non-accelerated driving speeds and powertrain configurations.

However, since the pass-by measurement result projects the information from a broad horizontal emission angle range to a single curve, directional and frequency-specific information cannot be derived. From a metrological point of view, simultaneous sound measurement in all radiation directions of a moving sound source is challenging and only possible by conventional methods if inherent measurement errors are acceptable:

1. The speed-dependent sound radiation of the tire/road noise and load-dependent engine sound components can be computed from separate measurements using a roller-type chassis dynamometer. The obtained results neglect advective wave propagation effects due to aerodynamic flow around the vehicle.

2. A separate measurement of the stationary passenger car in a wind tunnel is used to record the speed-dependent aeroacoustic radiation pattern. A disadvantage of this method arises from the positioning of the microphone capsules outside the flow, so that the received signals are distorted by measuring through the moving medium. In order to avoid distortions in the capsule, microphone placement is limited to positions outside the flow in wind tunnel measurements. Furthermore, the shape of measured directivity itself is distorted due to advection [68].

Striving to eliminate these metrological limitations, this work presents a method to compute speed-dependent directivity patterns of road-traffic vehicles from stationary hemi-circular microphone array signals. Subsequently, the recorded time signals are post-processed by applying time-variant wave backpropagation in order to reconstruct the emission signals on a virtual spherical surface around the moving sound sources. A multitude of approaches with individual foci on road-traffic auralizations is reviewed in the following. Beyond the focus of the reviewed works, the approach pursued in this thesis focuses on the real-time implementation of the auralization models as well as their integration capability and controllability in 3D visualization software.

This chapter is structured as follows: Section 3.2 introduces the methodology for *in situ* directivity patterns measurements using a pass-through microphone array. The subsequent signal postprocessing chain using time-variant wave backpropagation and according uncertainty considerations is presented in section 3.3. In section 3.4, the considered vehicle classes as well as the actual measurement procedure are defined. The results by means of speed-dependent directivity patterns and according sound power levels are presented and compared to results from the analytical Harmonoise model in section 3.5. The synthesis models for engine sound and tire/road noise are presented in sections 3.6 and 3.7, respectively. Their real-time implementation as VST plugin is presented in section 3.8. Finally, the results are discussed in section 3.9.

Table 3.1: Sound power dependency of tire parameters, according to [173].

Parameter	Proportionality	Exponent value n
Rolling speed v	$P \propto v^n$	$3 \sim 4$
Static load p_L	$P \propto p_L^n$	$1/3$
Tire inflation pressure p_s	$P \propto p_s^n$	$1/2$
Tire width b	$P \propto b^n$	$3/2$
Tire radius a	$P \propto a^n$	-3

3.1.1 State of the art

The following considerations are important in order to develop the measurement setup and its conditions in such a way that the measurement procedure is representative. The most recent knowledge on vehicle exterior noise and sound generation mechanisms is extensively compiled in chapters 86, 87 and 120 of Crocker's Handbook of Noise and Vibration Control [29], of which the most important briefly summarized in Table C.1. Regarding tire/road noise generation mechanisms, the vibration-related mechanisms mainly influence the frequency range below 1 kHz [14].

Tire/road noise models

Focusing on the tire/road noise component as shown in Table 3.1, Ullrich derived proportional relationships for dominant physical tire parameters with regard to their individual influences on resulting sound power P [173].

From the data, it can be approximated that tire/road noise varies in a level range of about 40 dB depending on a rolling speed between 10 and 150 km/h. Tire/road noise depends on the combination of both surface textures. Theoretically, a variation potential in the sound pressure level of 7 dB can be associated with the vehicle and 9 dB associated with the pavement [103]. Therefore, depending on the specific tire/road combination the dynamic range can vary by $\approx$ 15 dB [14], thus increasing the overall dynamic range of tire/road noise to $\approx$ 55 dB. Focusing on a tire/road noise emission relationship with vehicle speed v, a proportional increase in about 30 to 35 times $\log(v)$ can be experimentally observed, although a very wide range in the multiplier has been reported. In addition to overall levels, the effects of case-specific tire/road interactions on directivity patterns are not discussed in the literature. For comparison to tire/road noise, aeroacoustic noise depends on vehicle speed v by about 60 times $\log(v)$ [29].

The effect due to applied torque during acceleration causes variations up to 8 dB, depending on the tire design and on the applied torque [29]. Focusing on spectral effects, the vibration-related sound generation mechanisms of tire/road noise, such as radial/tangential vibrations as well as stick-snap and stick-slip effects, mainly influence the frequency range below 1 kHz [14]. Meng analyzed the location and emission spectra of partial sound source at moving passenger cars based on array measurements [119]. In a comprehensive review on general influencing factors of tire/road noise emissions, Li defined five subcategories: Tire-, driver-, pavement-, tread pattern- and environmental-related parameters [103]. Moreover, Li *et al.* extensively reviewed 74 tire/road noise models in [104] and categorized them into 37 deterministic, 18 statistical and 19 hybrid models. Noteworthy from those, Heckl considered the mechanical equations of motion to derive theoretical vibration patterns of tires in dependency of their material parameters [66]. A comprehensive tire/road noise model resulted from work made inside the so-called Sperenberg project, known as the SPERoN model (Statistical Physical Explanation of Rolling Noise) [15]. It predicts third-octave band sound pressure levels of tire/road noise in the range from 315 Hz to 2000 Hz.

Road-traffic directivity models

As stated in the final report of the Harmonoise project [2], information on directivity patterns and sound power of road-traffic sound sources is rare in the literature. Therefore, it includes generic formulae for horizontal and vertical directivities. Assuming omnidirectionality in the lower frequency range $f \leq 1250$ Hz, for higher frequencies $f \geq 1600$ the horizontal directivity corresponds to a symmetric cardioid pattern that is constricted perpendicular to the direction of travel level and shows a level reduction of 5 dB compared to the maximum radiation in the direction of travel. Similar to the vertical directivity, the minimum of a symmetric cardioid pattern is vertically above the sound source. Depending on frequency and point source height, its level of attenuation varies between 2 and 8 dB [87]. For a critical evaluation, the following section compares the computational directivity pattern results to the analytical patterns of category 1 (cars) from the Harmonoise catalogue as a reference. The Harmonoise model, which is used in common noise mapping procedures, describes cars by two vertically distributed source points with source heights 0.01 m and 0.3 m, respectively. Each of those is attributed to a specific pattern in eight octave bands between 50 Hz and 10 kHz. Both, horizontal and vertical directivity patterns are plotted in Figure 3.2. Important to note that the directivity patterns do not depend on the vehicle speed except from an overall level scaling factor. Emphasized as balloon plots

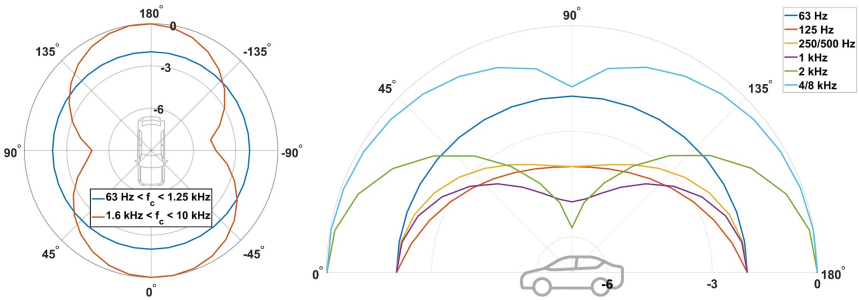
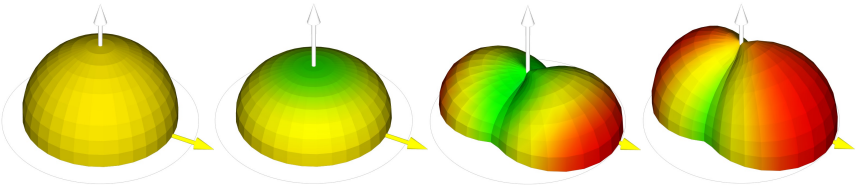


Figure 3.2: Reference: Horizontal (left) and vertical (right) directivity patterns from the Harmonoise car model.

in Figure 3.3, the formulae for the lower source describe four different spatial shapes in four frequency ranges:

1. Omnidirectionality is assumed for very low frequencies up to 80 Hz;
2. In the range of low-mid frequencies between 100 and 1250 Hz, the omnidirectional characteristic is increasingly attenuated above the vehicle for higher frequencies in order to simulate the shadowing of the tire/road noise by the vehicle body;
3. In the high-mid frequency range above 1600 Hz, an azimuthal component is added, which represents the amplified sound radiation through the horn effect;
4. In the range of high frequencies above 3150 Hz, only the azimuthal component remains, which describes a dipole that is reduced by 5 dB orthogonally to the direction of travel compared to the level in the direction of travel. The vertical level deflection is reduced (compared to lower frequency ranges), leading to an increase of about 4 dB.

However, as a data resource, the document refers to a conference paper abstract lacking any data, so that these models are not independently reproducible. The most elaborated approach for modeling road-traffic vehicle directivities was published by Tsukui *et al.* by approximating the directivities of three types of road vehicles at a fixed speed of 40 km/h in three octave bands around 500, 1000 and 2000 Hz [172] in the horizontal and vertical planes. For this specific vehicle speed, similar directivities for all three vehicles in shapes of the vertically compressed hemisphere are reported with the A-weighted sound pressure levels in the upward direction being 9–10 dB(A) lower than lateral ones. Moreover, an asymmetri-



(a) 50 Hz **(b)** 1000 Hz **(c)** 2000 Hz **(d)** 5000 Hz

Figure 3.3: 3-D car directivity patterns of the Harmonoise model in different octave bands. The yellow arrow indicates the driving direction, the white arrow points vertically upwards.

cal drift of the horizontal directivity according to an elevation angle of $\theta \approx 9°$ towards the backward direction can be observed for higher frequencies. At the current state of research, no speed-dependent directivities have been reported. However, details on the measured vehicles as well as a description of the precise post-processing steps are not well documented in any of the reviewed models. Furthermore, the equalization of the receiver characteristics has been neglected too.

Apart from directivities, two approaches for the determination of the sound power of a single vehicle have been described in the literature. On the one hand, Tachibana established a measurement procedure in a reverberant tunnel in order to practically apply the diffuse field method [165]. On the other hand, assuming a line source approximation, the sound power of vehicle flow can be estimated as a function of average speed and traffic flow rate [142].

In summary, the analytical road-traffic emission models developed in the Harmonoise project are referenced in this work by means of their wide integration with noise prediction simulation tools. However, this model is limited in terms of lacking drivetrain variability and neglecting the influence of driving speed on resulting directivity patterns.

3.1.2 Contribution of this work

This work focuses on speed-dependent directivity patterns of multiple vehicle classes with conventional internal combustion engine (ICE), hybrid, and electric drivetrains. Necessary data are obtained using the inversion method introduced in section 2.5.1. The presented method accounts for including the complex wave propagation effects into a far-field directivity pattern that neither can be simulated – even with elaborated numerical methods – nor can be directly obtained

from wind tunnel or roller-type chassis dynamometer measurements. Microphone arrays integrated in wind tunnels are located outside the flow field around the non-moving measurement object and are therefore limited to measuring laterally radiated sound. The following effects are part of the resulting directivity patterns: Advection due to the object's speed and the laminar flow around the moving object, refraction due to turbulent flow at body cavities, and diffraction due to object geometry. The generated data can be used for calibrated auralizations according to the procedure described in [38]. Beyond the scope of previously published research, this work presents the following novelty aspects:

- A technique to extract free-field radiation characteristics of moving sound sources.
- A supplementary directivity pattern database using the OpenDAFF [129] and SOFA (Spatially Oriented Format for Acoustics) [116] format conventions [1].
- Sound power computations based on single vehicle measurements (in opposite to established diffuse field or traffic flow-based line source measurement approaches).

3.2 Experimental setup for source characterization of road-traffic sound emission

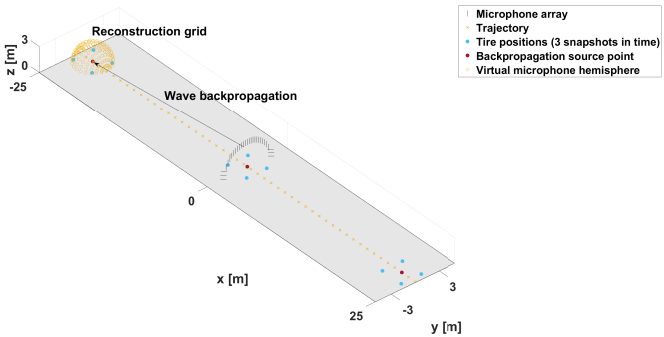


Figure 3.4: Projection of signals obtained with the pass-through microphone arc to a virtual surrounding microphone hemisphere. Note that the vehicle is moving through the array. The figure just shows a snapshot in time.

[1] Link to complete directivity pattern database (Last viewed 12/16/2024)

The general experimental approach is time-variant wave backpropagation, as introduced in section 2.5.1. However, specific adaptions for road-traffic sound sources are presented in section 3.3. In order to obtain a directivity pattern with quasi-freefield properties, the original source signals are recorded using a pass-through array. As shown in Figure 3.4, wave backpropagation can be interpreted as a projection of signals recorded at a two-dimensional microphone array onto a three-dimensional reconstruction grid around a moving vehicle's center point.

3.2.1 Array design for pass-through measurements

The microphone array is conceptualized for *in situ* measurements during a pass-through procedure. The microphone array consists of 30 microphones that are placed on a hemi-circular arc with a radius of $r = 3$ m (Figure 3.5).

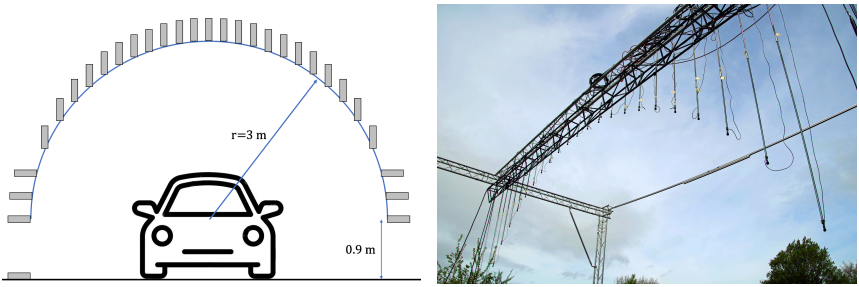


Figure 3.5: Construction scheme and realization of pass-through array for directivity measurements of moving traffic sound sources.

A NTI Audio M2230 measurement microphone (Class 1 acc. to IEC 61672 and ANSI S1.4) was installed flat on the ground, while omnidirectional Sennheiser KE4 microphones were installed on the remaining 29 positions on the hemi-circular arc. All microphones were connected to an RME Octamic XTC audio interface, set to $f_s = 96$ kHz sampling rate and $N_{\mathrm{ADC}} = 24$ bit depth. The microphones were equipped with windshields to avoid aeroacoustic distortion.

3.2.2 Microphone calibration

Each array microphone including its mounting using a threaded rod was individually equalized for the elevation range $0° < \Theta < 90°$ with a resolution of $10°$. The frequency responses were measured in the hemi-anechoic chamber of the

Institute for Hearing Technology and Acoustics using the substitution method according to [73] by placing the receivers in the center of a hemi-circular measurement arc (Fig. 3.6, left). Both, individual directivity patterns in the frontal hemi-circle with a resolution of $\delta_\theta = 10°$ as well as transducer sensitivities in V/Pa were measured with sweep excitation and subsequent post-processing for reflection removal. The signal-to-noise ratio (SNR) of the measurement chain, i.e., including preamplifiers, cables and analog-to-digital converters, were ≈ 85 dB for all channels.

The ideal sensor for the purpose of time-variant backpropagation would have spatially omnidirectional and spectrally flat properties. The microphone capsules are packaged in a custom-made cylinder (Fig. 3.6, right), which contains the electronics and XLR cable connector. The combination of the packaging with the threaded rod mounting causes a local diffaction of the sound field and therefore influences the overall frequency response. Therefore, the directivities of all microphones used in the measurement array were individually measured by measuring the transfer functions on an equiangular grid with a 2.5° x 2.5° resolution.

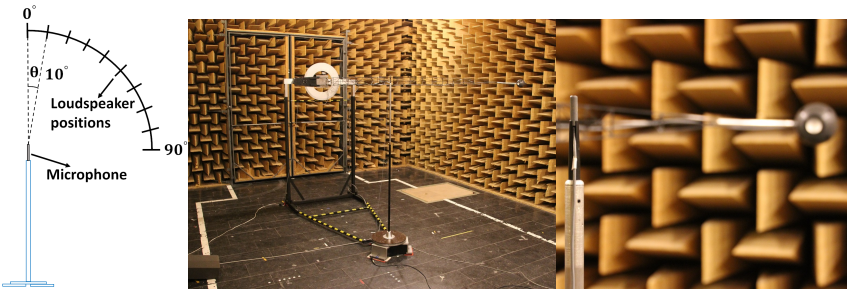


Figure 3.6: Directivity measurement setup in the hemi-anechoic chamber. Left: Schematic. Middle: Installation. Right: KE 4 capsule close up with measurement loudspeaker in the background.

For example, the measurement result (Fig. 3.7) shows the transfer functions of a single Sennheiser KE4 array microphone including threaded rod mounting over different incidence angles θ in the frequency range 2 kHz $< f <$ 20 kHz. In this plot, a perfectly omnidirectional receiver would appear as a spectrally flat transfer function at 0 dB. However, the result shows for off-axis positions the directivity pattern is approximately omnidirectional only in the frequency range $f < 2$ kHz. For higher frequencies $f > 2$, kHz the result indicates higher sensitivity and directionality. A maximum gain of 4.5 dB can be observed at 20 kHz for incident waves approaching from the off-axis direction. The same trend applies

to all microphones in the measurement setup and is individually equalized in the measured data. Finally, striving for an angle-dependent equalization of the recorded data, the resulting equalization filters are obtained by spectral inversion.

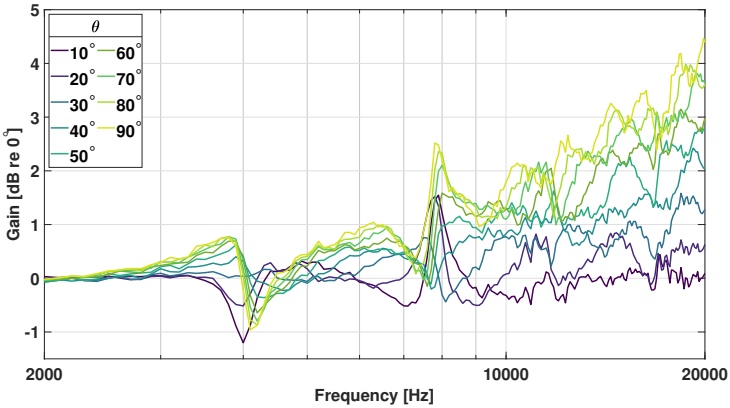


Figure 3.7: Off-axis frequency response (elevation range $10° < \theta < 90°$) of the Sennheiser KE4 capsule with custom-made cylindrical housing and threaded rod mounting. Free-field sensitivity measurement acc. to [73].

3.3 Postprocessing: Time-variant wave backpropagation

The recorded pass-through signals are postprocessed by a time-to-space conversion in order to obtain the directivity patterns on a specific sampling grid. During pass-through, the emitted sound wave approaches the array from a continuously changing angle. The backpropagated time signals are discretized by slicing and subsequent mapping to the grid points according to their respective azimuth and elevation angles. Overall, the directivity is reconstructed at 397 spatially discrete points on an equiangular grid with an angular resolution of 6° x 6° (Fig. 3.8). The according frequency information is computed from time signal slices with a spectral resolution of third-octave bands. The signal postprocessing is realized using a time-variant filter chain. The time-variant filters, which are equidistantly updated 500 times for each recorded pass-through, being equivalent to a distance of 0.1 m on the trajectory. All contributions from equation (2.16) are added to calculate the total attenuation in decibels according to the standard as:

$$A_{\mathrm{tot}} = A_{\mathrm{div}} + A_{\mathrm{atm}} + A_{\mathrm{gr}} \tag{3.1}$$

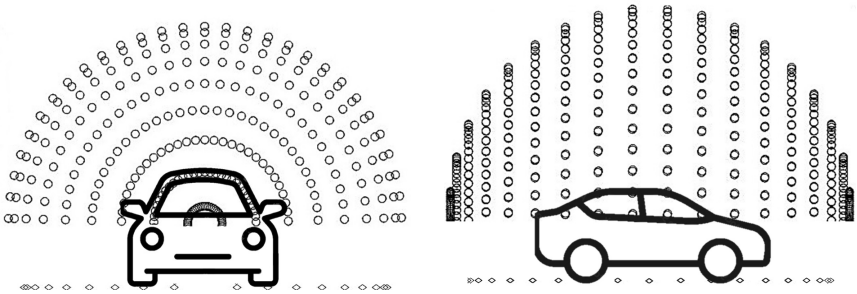


Figure 3.8: Sampling grid for the directivity reconstruction of moving traffic sound sources.

with A_{div} being the attenuation due to geometrical spreading (details in section 3.3.1), A_{atm} due to air absorption and A_{gr} due to the ground effect (details in section 3.3.2). Since the measurement microphones are placed at different heights over the asphalt surface, A_{gr} reflects the effect of two superimposed waves at the receiver in form of a comb filter in the frequency domain. The air absorption A_{atm} depends on frequency and is calculated by integration over the absorption coefficient α [13] in the limits of the source-to-receiver distance.

3.3.1 Source extension-dependent spreading loss $\mathbf{A_{div}}$

In the case of cars, the distribution of partial sources can be approximated by a finite line of uncorrelated broadband sources. This implies the sound pressure p to decay dependent on lateral distance d_{lat}, i.e., horizontally perpendicular to the driving direction, in a range between $p \propto \frac{1}{d_{\mathrm{lat}}^2}$ and $p \propto \frac{1}{d_{\mathrm{lat}}}$ (see Fig. 3.9).

Therefore, the tire/road noise component of the passenger car auralization model must consist of at least two point sources with a distance of the desired wheelbase. This work assumes a geometrical divergence of a finite line source by

$$A_{\mathrm{div}} = -10 \log d_{\mathrm{lat}} - 11 + 10 \log \left[2 \tan^{-1} \left(\frac{d_{\mathrm{wheel}}}{2\, d_{\mathrm{lat}}} \right) \right] \tag{3.2}$$

with d_{wheel} being the front-rear axle wheelbase of the road vehicle.

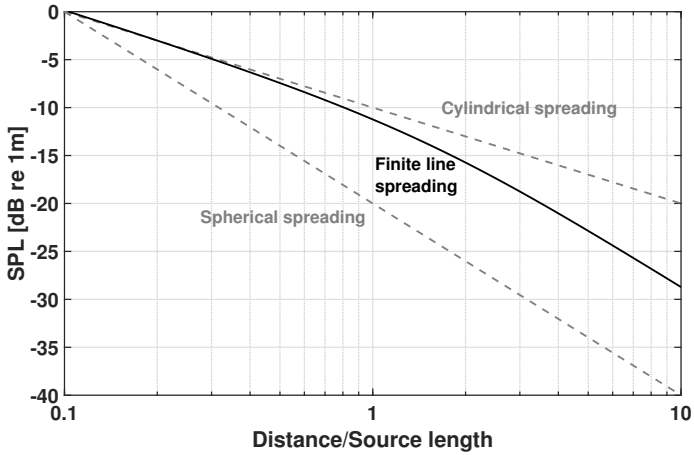


Figure 3.9: Source extension-dependent spreading loss of spherical, cylindrical and finite line (exemplarily shown with a length of 4m) sound sources [8].

3.3.2 Ground effect $\mathbf{A_{gr}}$

From a theoretical point of view, this work treats the ground effect $\mathrm{A_{gr}}$ as the problem of spherical sound wave propagation along a finite impedance plane as comprehensively discussed in [27] and in appendix A. Determining the grazing angle Θ in Figure A.1, this work assumes virtual source heights according to the Harmonoise model, i.e., 0.01 m for passenger cars and 0.3 m for motorcycle. Therefore, frequency-dependent absorption coefficient of the asphalt material α is derived from its spherical-wave reflection coefficient Q by

$$\alpha = 1 - |Q|^2 \tag{3.3}$$

with

$$Q = R_p + (1 - R_p)F(w) \tag{3.4}$$

being composed of the plane-wave reflection coefficient

$$R_p = \frac{Z_s \sin\Theta - \chi}{Z_s \sin\Theta + \chi} \tag{3.5}$$

consisting of the material-dependent complex surface impedance $Z_s = Z_c \cdot \coth(-ikd)$ assuming a rigid-backed layer thickness d, the grazing angle Θ and

$$\chi = [1 - (k_0/k)^2 \cos^2\Theta]^{1/2} \tag{3.6}$$

being computed from complex characteristic ground impedance Z_c, complex wave number k, $k_0 = 2\pi f/c$ (c being the homogeneous sound speed of air). The term $(1 - R_p)F(w)$ in (3.4) accounts for a surface wave component, where

$$F(w) = 1 + i\sqrt{\pi}w \exp(-w^2) \operatorname{erfc}(-iw) \tag{3.7}$$

is a virtually extended reaction surface being described by

$$w^2 = 2ik_0 r_2 \, \chi^2 \left[Z_{\text{norm}}(1 - R_p)\right]^{-2} \tag{3.8}$$

using the normalized surface impedance $Z_{\text{norm}} = Z_s/Z_0$ and r_2 being the path length of the ground-reflected wave in Fig. A.1. In this work Z_c and k are resulting from the Miki model equations [120], depending on the flow resistivity parameter R_s. The asphalt surface impedance Z_s is calculated assuming a rigid-backed layer with asphalt thickness $d = 0.04$m. In the apparent case of non-porous asphalt, a flow resistivity value R_s for stone mastic asphalt (SMA) was chosen to ($R_s = 6 \cdot 10^7 \frac{\text{Ns}}{\text{m}^4}$). Being rather reflective, a maximum absorption coefficient of 0.105 is reached at 20 kHz and shows good agreement with empirical results based on impedance tube measurements from literature (cf. Fig. 5 from [176]).

3.3.3 Uncertainty considerations

Uncertainties in the backpropagation calculations

By using the backpropagation method, an inherent error Δ_{BP} occurs in the A_{div} compensation term in equation (3.1). This error is due to compensating spreading loss by an analytical decay, whereas the real-world sound source is distributed and neither being a point source nor a finite line source. However, the choice of a spreading loss type affects the overall sound pressure level reconstruction. To estimate the resulting error regarding the directivity pattern and sound power, both decays are compared to a simulated baseline emission of a simplified car. It consists of four incoherent point sources representing the tire/road noise emission on an acoustically hard surface. A wheelbase of 2.8 m with a track width of 2.6 m is assumed. First, the directivity pattern of the simplified model is numerically calculated and referenced to a baseline of 0 dB in Figure 3.10. Second, pass-through signals at a virtual receiver array are simulated using the auralization software Virtual Acoustics [74]. Third, the simulated data are backpropagated to a single coordinate in the center of the four point sources by using the same algorithm described in section 2.5.1. The backpropagation is performed twice, on the one hand assuming spreading loss of a point source (PS) and on the other

hand assuming spreading loss of a finite line source (FL). Finally, both resulting backpropagation levels are referenced to the baseline level. The horizontal error depends on the azimuth angle and can be estimated to a remaining error range between $-1.3 < \Delta_{\mathrm{BP,FL}} < +1.8$ dB and $+3.0 < \Delta_{\mathrm{BP,PS}} < +5.0$ dB, respectively. By means of sound power, FL compensation overestimates the baseline by about 1.4 $\mathrm{dB_{SWL}}$, whereas PS compensation is 5.7 $\mathrm{dB_{SWL}}$ higher.
The smaller error of $\Delta_{\mathrm{BP,FL}}$ motivates the selection of the finite line source compensation for A_{div} in case of cars. Nevertheless, without further knowledge of the exact locations of all contributing sound sources of a car, the remaining error cannot be further reduced.

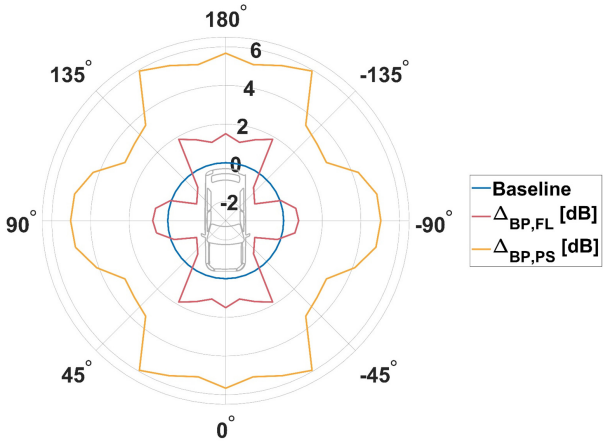


Figure 3.10: Horizontal reconstruction errors (in decibels) for finite line (FL) and point source (PS) spreading loss compensation.

Using a similar measurement apparatus, Witew estimated the worst-case influence of the aluminum truss on the sound field in Appendix B of his thesis [182]. The maximum disturbance of A_{div} (see equation (3.1)) in case of an orthogonal sound incidence is smaller than 0.5 dB at a minimum distance of 1 m between the microphone and the aluminium truss.

The compensation algorithm for air absorption A_{atm} in equation (3.1) is performed based on the analytical absorption model from [13] which depends, among others, on humidity and temperature. The accuracy of this postprocessing step, however, is limited by the thermo-hygrometer's accuracy used in the measurement setup. For resulting spectral wave backpropagation uncertainty is estimated by means of input parameter variation. The results (Fig. 3.11) are simulated as

worst-case estimate for a fixed distance of 25 m, corresponding to the largest distance of the measurement setup.

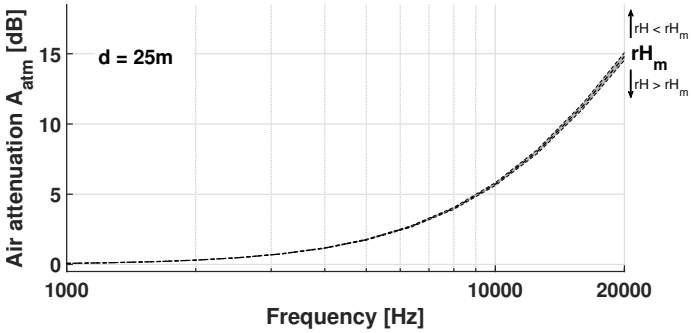


Figure 3.11: Backpropagation uncertainty due to the limited thermo-hygrometer's temperature accuracy.

The thermo-hygrometer's temperature accuracy of ±1°C corresponds to a worst-case uncertainty of ±0.28 dB at 20 kHz, and without plot, an uncertainty due to a humidity accuracy of ±3% corresponds to a worst-case uncertainty of ±0.02 dB at 20 kHz.

Uncertainties due to scattering on asphalt surface

The wave backpropagation algorithm is limited to correcting for the surface impedance and neglects the surface roughness. Therefore, the frequency-dependent

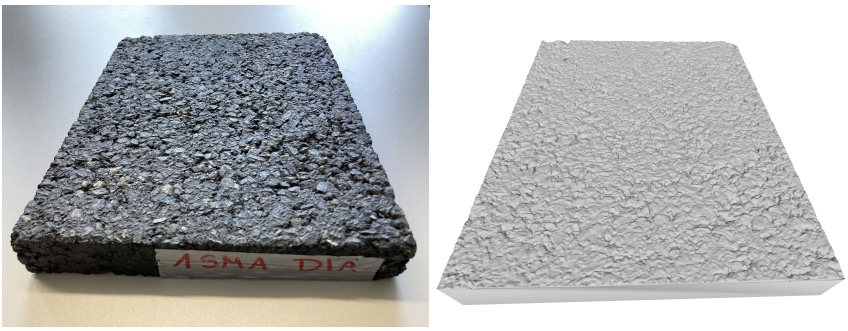


Figure 3.12: Original (left) and 3D-scanned split mastix asphalt road surface model (right).

error due to neglecting sound wave scattering at the asphalt surface is estimated in the following. The simulation model of the asphalt surface was developed by using an Artec Space Spider 3D scanner with an optical resolution of 0.1 mm (Figure 3.12). The resulting surface mesh is shown in Figure 3.13 (left). The actual scanning data were taken *in situ* at the same test track where the vehicle noise emission measurement campaign took place, cf. section 3.2. The asphalt layer is used for acoustical measurements of vehicle noise and complies with the ISO 10844:2021-12 standard [75]. Using the resulting mesh, the sound field was numerically simulated using boundary element method (BEM) for impinging sound from multiple zenith angles θ (Fig. 3.13, right).

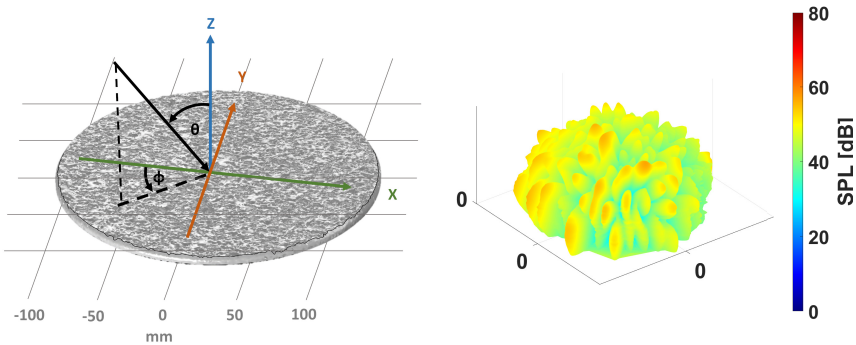


Figure 3.13: Left: Surface mesh of the scanned ISO asphalt surface with coordinate system for the numerical scattering simulation. Right: Scattering pattern at 20 kHz from the ISO asphalt model for grazing incidence ($\theta = 70°$, $\phi = 0°$).

From the BEM-simulated results, the angle-specific scattering coefficients δ_{θ_i,ϕ_j} of each angle of incidence (θ_i, ϕ_j) are calculated using the equation (3.9) by Mommertz [122]

$$\delta_{\theta_i,\phi_j} = 1 - \frac{|\sum_{k=1}^{n} \mathbf{p}_{\theta_i,\phi_j}(\mathbf{X}_k) \cdot \mathbf{p}^*_{0,\theta_i,\phi_j}(\mathbf{X}_k) \cdot w(\mathbf{X}_k)|^2}{\sum_{k=1}^{n} |\mathbf{p}_{\theta_i,\phi_j}(\mathbf{X}_k)^2 \cdot w(\mathbf{X}_k)| \cdot \sum_{k=1}^{n} |\mathbf{p}_{0,\theta_i,\phi_j}(\mathbf{X}_k)^2 \cdot w(\mathbf{X}_k)|} \tag{3.9}$$

with the individual surface weights of each simulated microphone position $\mathbf{X}$ on the hemisphere being taken into account. $\mathbf{p}_{\theta_i,\phi_j}(\mathbf{X})$ correspond to the simulated sound pressure results from the asphalt model, whereas $\mathbf{p}_{0,\theta_i,\phi_j}(\mathbf{X})$ correspond to results from a flat reference model. The number of simulated receiver points is 32401, resulting from an equiangular sampling grid with a resolution of $\theta = 1°$

and $\phi = 1°$. As shown in Figure 3.14, two kind of results are derived from the angle-specific scattering coefficients δ_{θ_i,ϕ_j}:

1. The elevation angle-dependent scattering coefficients δ_θ are derived by averaging over all corresponding azimuth angles, and
2. The random-incidence scattering coefficient δ_{ri} is calculated using a discretized modification of Paris' formula by averaging over n angle-specific scattering coefficients [178]

$$\delta_{\mathrm{ri}} = \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{\theta_i\phi_j} \cdot \frac{1}{n}\frac{\sin(2\theta_i)}{\sum_{i=1}^{n}\sin(2\theta_i)} \tag{3.10}$$

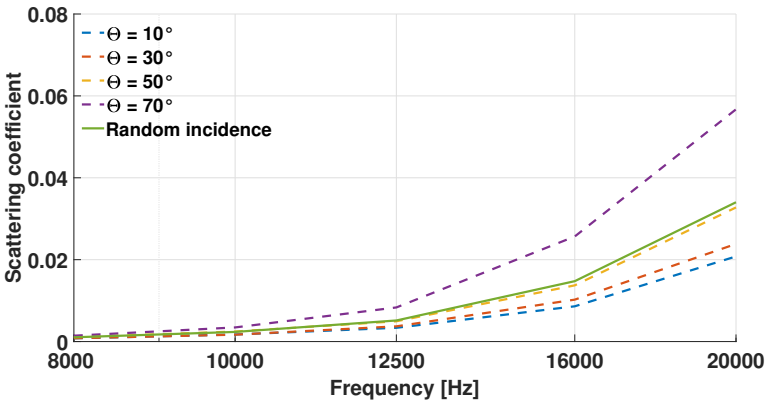


Figure 3.14: Resulting frequency-dependent random-incidence (solid line) and different unidirectional (dotted lines) scattering coefficients of ISO asphalt surface.

Based on the results shown in Fig. 3.14, it can be estimated that in the worst-case scenario, such as grazing incidence observed in vehicle pass-by noise, the scattered energy is below 10%. Considering the predominant frequency range of road-traffic sound emissions (below 5 kHz at a driving speed of 50 km/h or below 10 kHz at 120 km/h, as derived from a threshold of 30 dB attenuation in Fig. 3.1), this implies that more than 90% of the nominal energy is accounted for in the relevant frequency range for directivity pattern reconstructions. Therefore in worst case, this results in an underprediction of $\Delta\mathrm{SPL} = 20\log_{10}(\sqrt{0.9}) \approx 0.46\,\mathrm{dB}$ in the reconstructed sound pressure level.

3.4 Measurement

3.4.1 Vehicle classes

The range of characteristics of noise emissions is as broad as the multitude of variables in the vehicle configuration. For this reason, the following chapter compares noise emissions using reference cases of the most relevant road vehicle classes and also examines the influence of driving speed. Specifically, this experiment investigates six passenger cars from four different classes and a motorcycle (Fig. 3.17), according to the ISO 3833:1977 classification scheme. Their individual specification regarding tire and fuel type, engine and power is compiled in table 3.2.

Vehicle	Classific.	Tire specification	Fuel	Engine	P
BMW i4	D-segm. CE	Goodyear Vector 4S 225/55R17 101W	E	–	210
Ducati Monster 797	2-wheel	Pirelli Diablo Rosso 180/55 ZR17	G	Desmodue 803	54
Fiat Scudo	B-segm. large MP	Hankook Winter I*c 215/60R16 99H	D	165 Multijet	120
Opel Meriva B	B-segm. MP	Dunlop Winter Sport 205/55R16 94H	D	1.6 CDTI	81
Peugeot Boxer XL	M-segm. LC	Continental VanCont 225/75 R16C 121R	D	140 3 Multijet	103
VW Golf GTE	C-segm. CC	Dunlop Sport Maxx 225/45R17 91W	H/G	1.4 TSI	150
VW Golf Sportsvan	B-segm. MP	Uniroyal RainSport 225/45R17 91Y	D	2.0 TDI	110

Table 3.2: Technical specification of the measured vehicles (shown in Fig. 3.17) according to the ISO 3833:1977 classification scheme. Abbreviations: Power in kilowatts (P). Fuel types: Electric (E), diesel (D), gasoline (G) and hybrid (H). Classification: Compact car (CC), compact executive (CE), multi purpose (MP) and light commercial (LC).

3.4.2 *In situ* array installation on test track

The *in situ* measurements took place at the test track of the Institute for Automotive Engineering (IKA) of RWTH Aachen University [2]. The test track consists of a straight asphalt road with a length of 450 m and a turning circle with a diameter of 95 m, in which the vehicles accelerate. The test track is surrounded on one side by agricultural fields and warehouses on the other side, from which no noticeable background noise distorts the acoustical measurements.

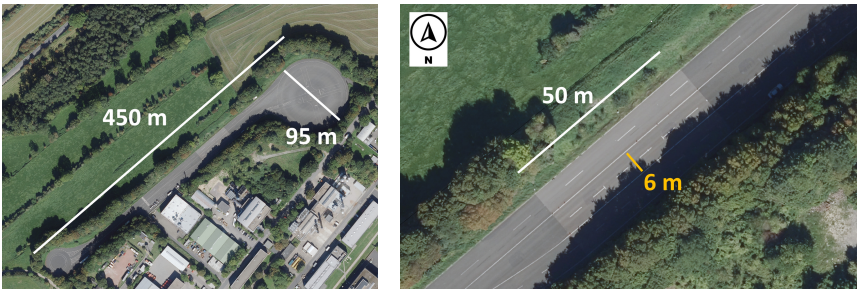


Figure 3.15: Left: Aerial view of the measurement site, right: Close up of array size, position and orientation (orange line).

Clearly visible as light gray-shaded area in Figure 3.15 (right), the test track features a certified asphalt layer with a length of 50 m, which is used for acoustical measurements of vehicle noise. The asphalt layer complies with the ISO 10844:2021-12 standard [75]. The standard describes requirements for asphalt roads for the purpose of noise emission measurements of road vehicles and their tires. According to the standard, the asphalt of the road section is subject to a maximum sound absorption of 8% in each one-third octave band between 315 Hz and 1600 Hz, a maximum grain size of 8 mm and an average tread depth of 0.5 mm with 0.2 mm tolerance. As indicated by the orange line in Figure 3.15 (right), the pass-through microphone array is placed in the center of the ISO-compliant asphalt area.

3.4.3 Measurement procedure

After accelerating the vehicle in the turning circle and turning onto the straight track, a constant vehicle speed was controlled by the driver with the assistance

[2] Link to OpenStreetMap: Geographical location of the IKA test track

Figure 3.16: *In situ* installation of the pass-through measurement array for measuring directivities of moving road-traffic sound sources.

of a GPS-based feedback or the vehicle's integrated speed control. In the case of vehicles with ICE, the engagement of the speed-typical gear was controlled by automatic transmission. Aiming to center the vehicle through the measurement array, a non-transparent adhesive strip was placed on the windshield in such a way that the driver's subjective perception corresponded with the center strip of the test track. For each vehicle, measurements with three repetitions were taken at discrete vehicle speeds between 10 and 120 km/h with 10 km/h steps. The atmospheric conditions (temperature, relative humidity and atmospheric pressure) were measured individually for each vehicle by using a digital thermo-hygrometer, type TFA Klima Bee 30.5036. By passing through the array, the sound emissions of the sources, i.e., cars or motorcycles, are recorded from multiple angles. The vehicle speed was externally validated by the placement of two light barriers at both ends of a trajectory with length $L = 50$ m. The passing trajectory for the backpropagation is $\pm$ 25 m to both sides of the array, thus limiting the elevation angle to a minimum of $\theta_{\min} = 6.84°$.

(a) BMW i4

(b) Peugeot Boxer

(c) VW Golf Sportsvan

(d) VW Golf GTE

(e) Opel Meriva B

(f) Fiat Scudo

(g) Ducati Monster

Figure 3.17: Road-traffic vehicles measured in the experiment.

3.5 Results

The results of the time-variant backpropagation are then reconstructed in the form of one-third octave band spectra at discrete grid points (Fig. 3.8) for each vehicle and speed, and averaged for three measurement repetitions (as introduced in section 3.4.3). In the following, notable results are discussed by using two-dimensional (2D) projections and three-dimensional (3D) directivity patterns. The complete data are available in the accompanying database [3] using two directivity exchange formats, OpenDAFF [129] and SOFA [116] by means of both, normalized magnitude spectra (for use in auralizations) and third-octave sound pressure levels (for noise mapping applications). We start, however, with the results for the radiated sound power.

3.5.1 Sound power of road-traffic sound sources

Based on the actual backpropagated measurement data, the speed-dependent sound power level L_W of each vehicle is computed using the procedure described in [78] by using the angular segments from the backpropagated data. According to the standard, the sound power is derived by summing the area-weighted sound intensities from each grid point. The latter in turn are obtained as the ratio of the squared effective sound pressures and the acoustical free-field impedance. The result is shown in Figure 3.18 in comparison to additional sound power levels from the Harmonoise model [87].

The obtained values calculated in this work agree with the analytical model in terms of magnitude. Individual dips in the curves can be explained by the gear engaged: In the range of dominating powertrain noise contributions (at low speeds), a decrease in sound level can be observed for increasing driving speeds when the engine's contribution to the sound diminishes from high levels (due to high engine speed in low gear) to lower levels at higher driving speeds (due to lower engine speed in a higher gear). As expected, the noise emissions of electrified vehicles (VW Golf GTE and BMW i4) are lower than those with a combustion engine, with reductions up tp 10 $\mathrm{dB_{SWL}}$ at low driving speeds. The obtained sound power level difference between motorcycles and passenger cars aligns with the predictions made by HARMONOISE. For an average passenger car (category I) compared to a motorcycle (category IV.a), propulsion noise emissions of motorcycles are approximately 5-20 $\mathrm{dB_{SWL}}$ higher.

[3] Link to complete directivity pattern database (Last viewed 12/16/2024)

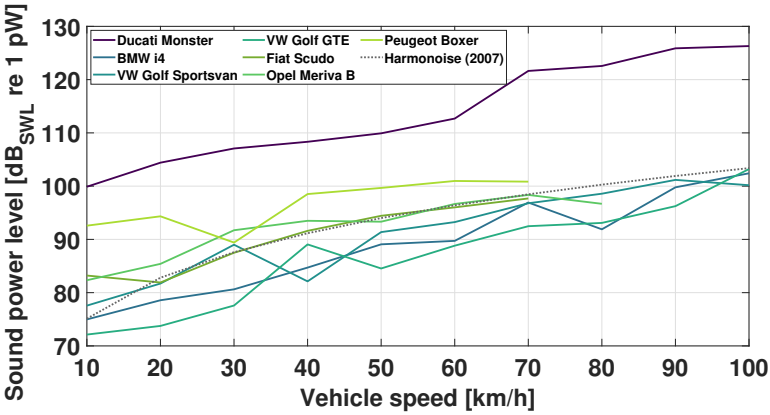


Figure 3.18: ISO 3745-compliant speed-dependent sound power levels of different road vehicle classes in comparison to Harmonoise model values [87].

3.5.2 Directivity patterns of road-traffic sound sources

As mentioned in section 3.3, the backpropagated time signals are discretized by slicing and subsequent mapping to the sampling grid points. In the following, exemplary analyses by means of 3D balloon plots as well as 2D cross-sectional plots are presented. Shown using a normalized logarithmic scale (i.e., in decibels, referenced to the maximum emission value), each balloon plot shows the spatial emission of a single third-octave band (cf. Fig. 3.19), cross-sectional plot shows multiple third-octave bands in a single plane (cf. Fig. 3.20).

Passenger cars

Without exception, the main sound radiation for all measured passenger cars can be observed along the vehicle axis. This finding supports the assumptions of the Harmonoise model in Figure 3.3. However, changes in speed have a complex influence on the directivity patterns. In general, the individual emissions of all four wheels are visible in the directivity patterns by local main lobes in the horizontal patterns, such as in Figure 3.20.

Vehicles with hybrid powertrains, such as the VW Golf GTE in this study, are advantageous for an isolated analysis of partial sound contributions to the overall emission and therefore exemplarily presented in the following. The aeroacoustical noise significantly contributes to the overall emission for higher speeds $v > 80$

km/h. The VW Golf GTE was measured in two different powertrain modi, with ICE, and in electric mode (with activated AVAS). For higher speeds $v > 30$ km/h, the AVAS component is deactivated in electric drivetrain mode, with remaining road/tire and aeroacoustical noise sources. Accordingly, the difference between the results of the two drivetrain modi can be attributed to the engine. Each modus is represented in the following by balloon plots in four different third-octave bands. The yellow arrow indicates the driving direction, the white arrow points vertically upwards.

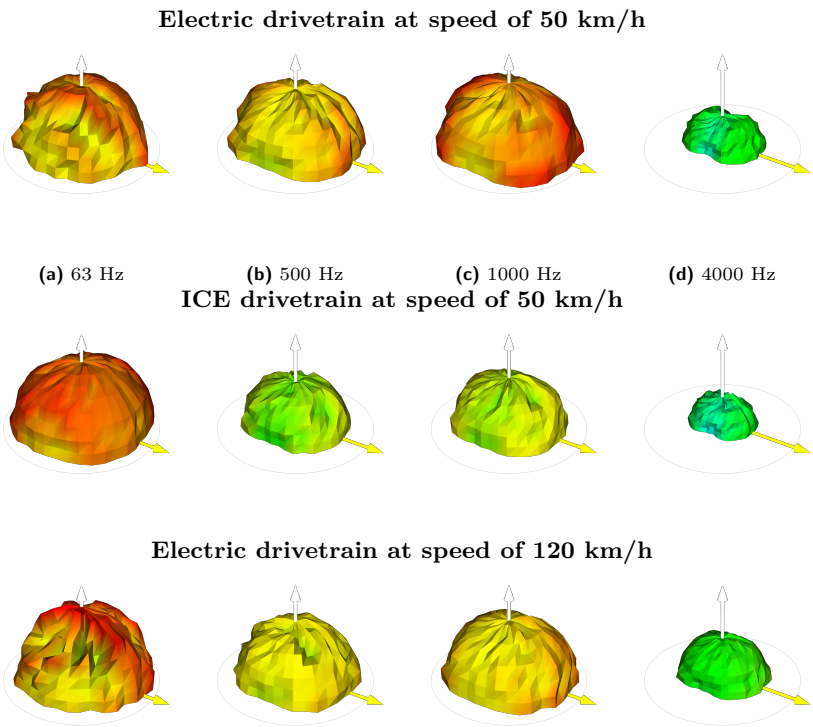
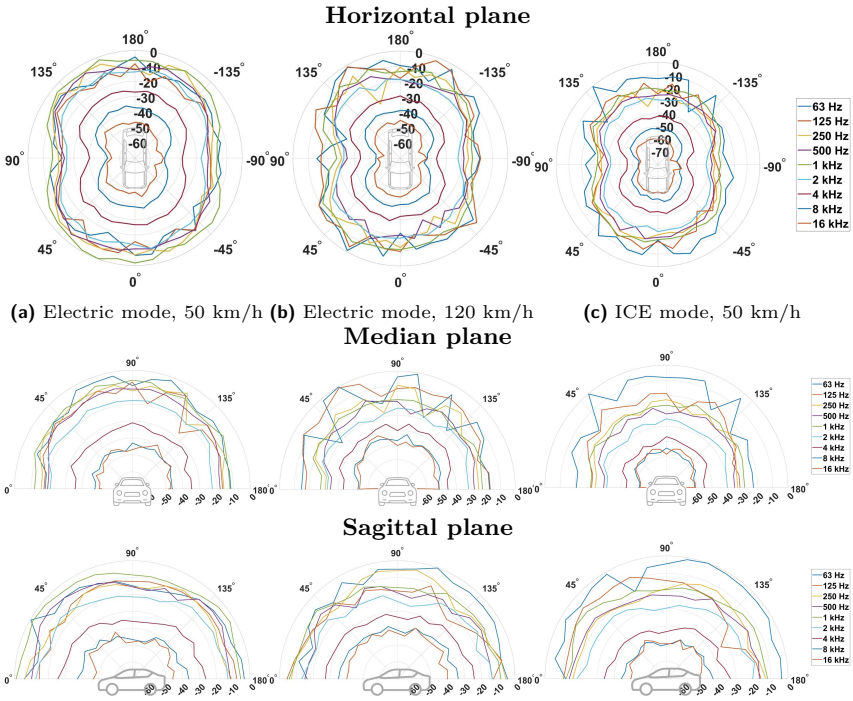


Figure 3.19: VW Golf GTE: 3-D third-octave band directivity patterns for different drivetrain modi and speeds.

Clearly visible at high speeds in electric powertrain mode, the median and sagittal cross-sectional plots (Fig. 3.20 (b)) indicate the aeroacoustical noise contribution – occurring with high sound levels, being dominant by > 10 dB in the lowest octave band – which is emitted obliquely upwards to the rear.



(a) Electric mode, 50 km/h **(b)** Electric mode, 120 km/h **(c)** ICE mode, 50 km/h

Figure 3.20: VW Golf GTE: 2-D directivity patterns for different drivetrain modi and speeds (cf. Fig. 3.19).

Focusing on inter-object differences between cars at a constant driving speed, Figure 3.21 compares multiple third-octave band directivity patterns in the horizontal, median, and sagittal plane. Moreover, no clear dominancy of a certain vehicle throughout all frequency bands can be observed. Rather than levels, broadness of the main radiation lobes correlates to a specific car. Furthermore, the median plane patterns show symmetry to a large extent, whereas the sagittal plane patterns are more directional. In contrast, Figure 3.22 focuses on intra-object differences by comparing the patterns for a single car at multiple driving speeds. A comparison therefore refers to the speed dependency. Considering that the plot covers a wide level range of 70 dB, it is clear that the directivity patterns are speed-dependent.

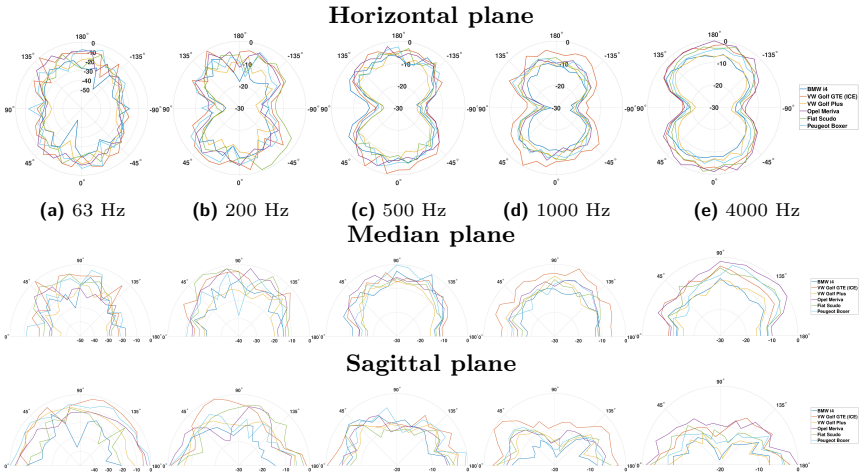


Figure 3.21: Comparison of multiple third-octave band directivity patterns of all cars at a driving speed of 50 km/h.

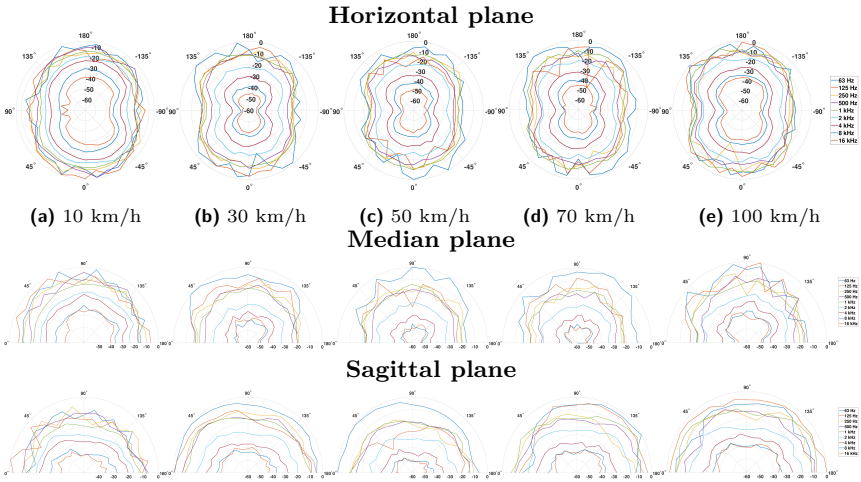


Figure 3.22: VW Golf Sportsvan: Comparison of normalized third-octave band directivity patterns at multiple driving speeds.

Two-wheeled motorcycle: Ducati Monster

In contrast to those of cars, the directivity pattern of the motorcycle is rather directed to the sides and depends less on speed. In particular at low speeds, the directivity patterns between 10 and 50 km/h are very similar, which could be explained due to the highly dominant sound emission by the engine. The spectrally narrow-band and highly directed sound emission by the exhaust (which is physically mounted upwards with an elevation angle of $\theta_{\mathrm{ex}} = 50°$) is clearly visible in the 3D balloon plot in Figure 3.23.

Figure 3.23: Detailed view of the Ducati Monster motorcycle (right) and 200 Hz directivity pattern of the Ducati Monster at a speed of 50 km/h (left).

Figure 3.24 shows the comparison of constant driving conditions with variation of the engaged gear. In the case of a motorcycle, the resulting patterns for gears

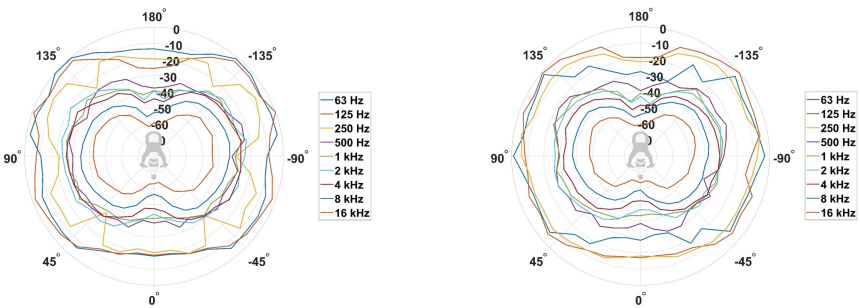


Figure 3.24: Motorcycle directivity pattern comparison of gear 2 (left) vs. gear 3 (right) at the same speed of 70 km/h.

2 and 3 at a constant speed of 70 km/h differ significantly in low frequency bands. Furthermore, the sound power level emitted with gear 2 engaged is about

2 dB_{SWL} higher than with gear 3 engaged. Some general trends can be observed from the motorcycle directivity patterns. In the high frequency range $f > 4$ kHz, slightly more energy is radiated to the driver's left side cf. horizontal and median plane plots in Figure 3.25, due to chain noise. The sagittal directivity pattern shows that, as speed increases, high-frequency amplitudes are gaining more rapidly than those at lower frequencies.

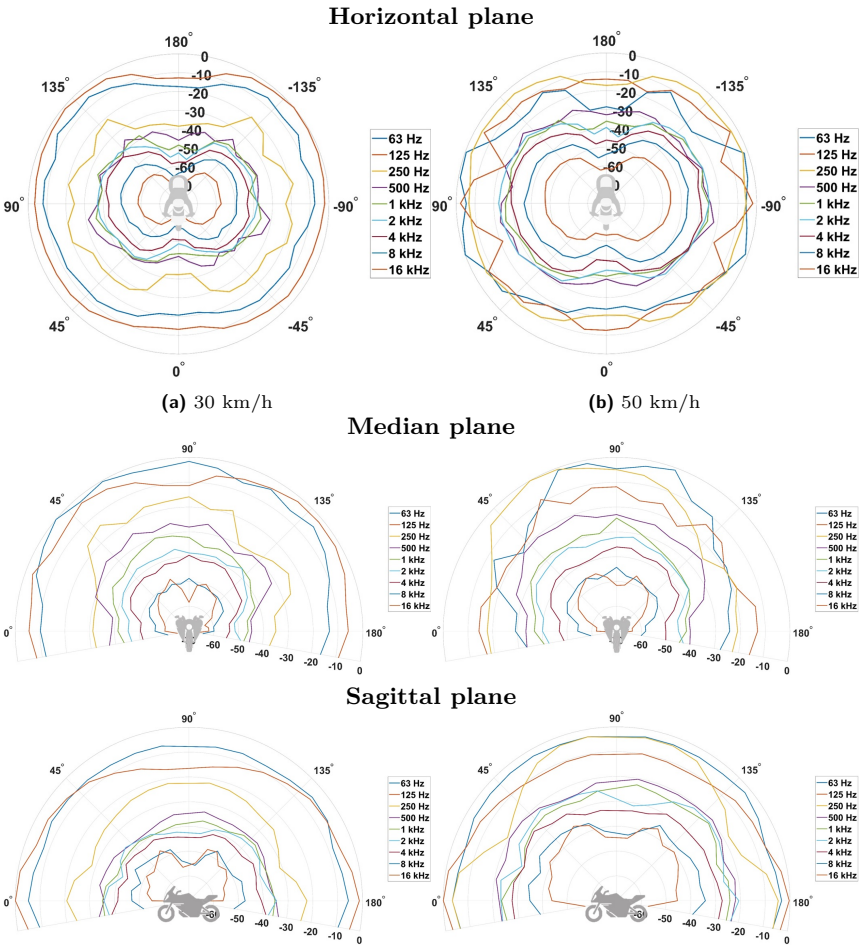


Figure 3.25: Ducati Monster motorcycle: 2-D directivity patterns at different speeds.

These findings document the great variety of directivity patterns between different vehicle class segments and speeds. First and foremost, from the obtained data it can be concluded that the directivity patterns of road vehicles depend on driving speed. Most influential on a resulting pattern is the driving condition of the vehicle, in particular the engine speed (as a result of the gear engaged). In general, no clear dominancy of a certain vehicle throughout all frequency bands can be observed - meaning that the mixture of contributing engine, tire/road and aerodynamic sound sources is highly dependent on the particular vehicle. Rather than levels, broadnesses of the main radiation lobes correlate to a specific car. Furthermore, the median plane patterns show symmetry to a large extent, whereas the sagittal plane patterns are more directional.

3.6 Powertrain sound synthesis

The following sections will present real-time synthesis implementations for two types of powertrains: internal combustion engines (section 3.6.1) and electric drives (section 3.6.2). These will be integrated with an additional tire/road synthesis (section 3.7) into a unified VST plugin (section 3.8).

3.6.1 Combustion engines

A real-time capable procedural synthesis model for generating emission signals of intake and exhaust noise contributions was proposed by Baldan et al. [12], which elaborates an initial approach by Farnell [44]. Corresponding to the duration of a full engine cycle (induction, compression, combustion, exhaust) a sawtooth-like trigger phasor linearly rises from 0 to 1. Four processes (piston motion, fuel ignition, intake valve, exhaust valve) are derived from this trigger signal and described by trigonometric functions:

- Piston moves up and down two times during one cycle → Cosine wave.
- Intake valve and exhaust valve opening → Positive half cycle of a sine wave.
- Variable pressure increase due to fuel ignition → Parametrically controllable width of the fuel ignition pulse.

The engine model fundamentally assumes the intake, cylinder and exhaust system to be a tube network with each tube being realized by a digital waveguide (introduced in section 2.6.2, [159]). The sound field inside a waveguide is composed of two time-discrete input signals $x_0[k]$ and $x_1[k]$ and two output signals $y_0[k]$ and $y_1[k]$ (Fig. 3.26).

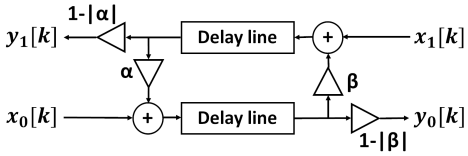


Figure 3.26: Schematic of the digital waveguide.

On each delay line the waves propagate in both directions: Due to its variable reflection coefficients α and β a certain part of an incoming wave is reflected at the end of a delay line. The total propagating wave results from the constructive or destructive interference of the outgoing and the return wave: For multiple integers of half wave lengths being equal to the waveguide length results in a standing wave of high amplitude. Except from this case the interference is destructive.
A digital waveguide consists of two delay lines that support very fast modulation of the implemented delay length without any audible artifacts or clicking noise, which is essential for a plausible result. Each delay line stores values of an audio stream. A simple storage in terms of a first-in-first-out memory (FIFO) would not be sufficient, since the variable cylinder tube length due to the piston motion must be mimiced. Therefore, the delay line is implemented as variable delay line (VDL, Fig. 3.27) [145]. Each VDL consists of one write position and two read positions where new values are written into, or read from the VDL.

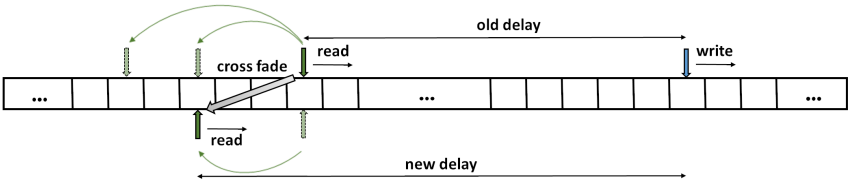


Figure 3.27: Schematic of the variable delay line (VDL) implementation.

The sound synthesis originates from a cylinder and is spectrally filtered by the intake and outlet elements. The latter are modeled as waveguides of fixed length, with modulated reflection at the valve end and a fixed reflection coefficient at the free end. Additionally, the intake path is fed with lowpass-filtered white noise to simulate turbulences in the air and the fuel mixture. The exhaust path consists of a centrally positioned muffler with inlet and outlet tubes. The spectrally complex modeling of the muffler attenuation is based on four digital waveguides with length-tuned delay lines resulting in destructive interferences.

With regard to emission levels, engine noise mainly depends on engine speed (revolutions per minute, rpm) and load with levels varying in a range of ≈ 20 dB [14] in traffic situations. A typical proportional increase of ≈ 30 times log(rpm) and increases by about 2 to 3 dB for partial versus full throttle can be observed [29]. Generally, power train noise does not show a constant relationship to vehicle speed, rather being a function of the gear selection.

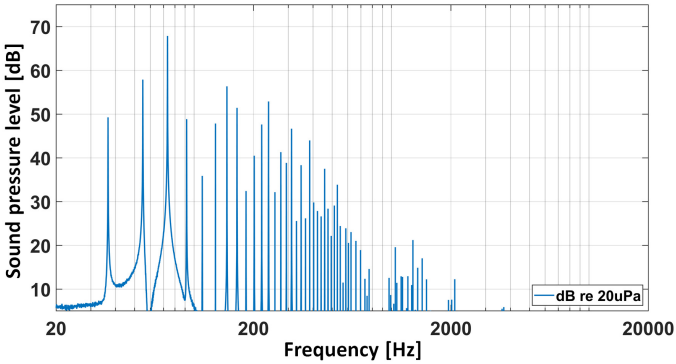


Figure 3.28: Exemplary spectrum of a synthesized four-stroke combustion engine with four cylinders at 2200 rpm.

Shown in Figure 3.28 at the example of a four-stroke combustion engine with four cylinders and a rotational speed of 2200 rpm, the synthesized spectrum contains clearly distinguishable engine orders with dominating contribution from harmonics of the explosion frequency

$$f_{\mathrm{ex}} = \frac{N_{\mathrm{c}}}{2} \frac{\mathrm{rpm}}{60} \tag{3.11}$$

with N_{c} being the number of cylinders. The additional harmonics stem from both the engine cycle, which encompasses two revolutions, and the revolution cycle. The synthesis model produces an artifact-free and plausible auralization. Since its spectrum aligns with objective expectations too, it was not further subjectively evaluated in listening tests.

3.6.2 Electric engines

Compared to the sound emission from vehicles equipped with conventional combustion engines, electrical powertrains will seriously change the soundscape of urban environments since their acoustical characteristics significantly differ (cf.

table 3.3). Most important, the drivetrain emission reduces to a purely harmonic source spectrum and due to increased torque of electric drives, the load dependency decreases. Furthermore, the number of partial sound contributors and their spatial distribution is reduced for electrical drivetrains.

Table 3.3: Sound source properties of combustion engine and electrical powertrains.

Powertrain type	**Combustion**	**Electric**
Sound spectrum	Harmonics + Stochastic noise	Harmonics
Parameters	Rotational speed + Load (strong)	Rotational speed + Load (weak)
Source distribution	Multiple, spatially distributed source points (Intake, exhaust, engine housing, gearbox)	Dense source location (Engine housing and gearbox)

Focusing on the sound quality assessment of trains with electrical drivetrains, Klemenz implemented a source signal synthesizer for tonal components [93]. The tool has a parametrical design and incorporates multiple variables, such as pole-pair number, wheel perimeter, acceleration and gear dimensions. Bösing et al. presented a comprehensive methodology for the acoustic modeling of electrical drives [23]. It classifies the noise sources in electrical drives into two categories: Airborne sound emissions are caused by aerodynamic noise from rotor and fan of the cooling, and structure-borne sound emissions are caused by housing and stator vibrations that are resulting from mechanical components (bearing and friction), electromagnetic force (radial force ripple, torque ripple and converter switching noise) or geometrical eccentricity. Their combined approach computes the machine-inherent structural vibration modes with subsequent prediction of its responses. Being applicable to all machine types and designs - such as slanted, axial flux, transverse flux or external rotor motors - the offline simulation provides input data for auralizing the electromagnetic noise excitation. It comprises the electrical, mechanical and structural-dynamic properties of the permanent magnet synchronous machine (PMSM) over the sound-radiating surface vibrations of the motor's housing to a binaural auralization for subjective evaluation considering the indoor (rooms) or outdoor sound propagation [57].
A PMSM uses permanent magnets in the rotor and a variable frequency current traveling through the stator to generate torque. As the rotor turns, the stator uses electric currents to generate a magnetic field that follows the speed of rotation of the rotor to generate a consistent torque. The electromagnetic forces generated

during its operation generate vibrations not only at the frequency of excitation but also at higher frequencies. These variations, called higher-order harmonics, appear at multiples of the first harmonic. Since the electric engine is a periodically rotating machine its spectrogram is composed of multiple sinusoidal harmonics (Figure 3.29, right).

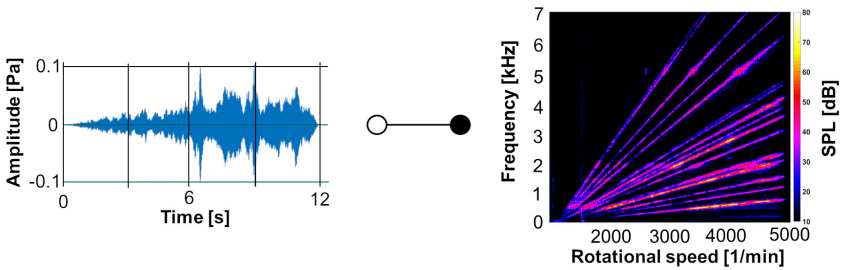


Figure 3.29: Time signal (left) and spectro-temporal transformation (right) of the reference electric engine sound emission during a run up.

Therefore, sound synthesis is achieved using the additive synthesis technique (introduced in section 2.6.1), featuring a low computational load, flexible rotational speed parameterization, and a modular architecture that easily adapts to data from any specific electric motor. The model input is an anechoic recording signal s(n) (Fig. 3.29, left) of an iolated electrical drive's run up on a test bench (Fig. 3.30).

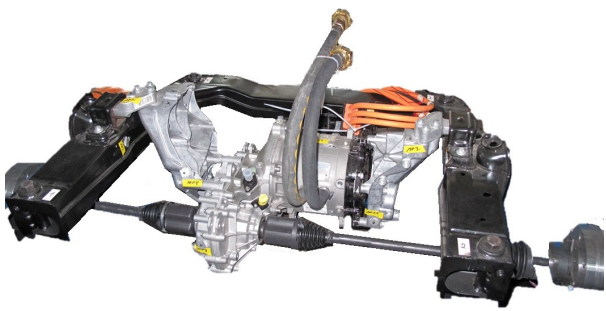


Figure 3.30: Reference electric engine on a test bench.

Generally, the input signal of the presented method can be a microphone recording or numerically simulated. As shown in the block diagram (Fig. 3.31), the signal processing chain consists of an analysis step that transforms the electric engine's

sound emission signal s(n) into a set of coefficients by means of discrete, rotational speed-dependent amplitude values $\hat{A}_p$. Based on these coefficients, the synthesis step produces audible sound by using additive synthesis.

Figure 3.31: Block diagram of the analysis-and-synthesis procedure based on the sinusoidal model.

Two different approaches can be chosen for the synthesis step: As in this work, it can be based on synchronized information on the rotational speed (e.g. from CAN bus data) or information on the PMSM geometry by means of its number of pole pairs. The Campbell plot of the anechoic run-up recording (Fig. 3.29, right) reveals the main oscillators contributing to the acoustic response of the PMSM at various speeds of rotation (rpm). Using the Fast Fourier Transform (FFT), peak amplitude values $\hat{A}_p$ are computed at discrete frequencies $\hat{\omega}_p$. For the given run-up signal, the plot reveals the signal to be composed by 15 dominant harmonics. It can be observed that each individual oscillator has a unique amplitude envelope (Fig. 3.32).

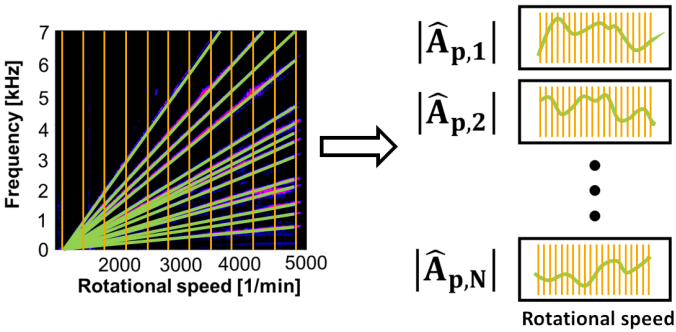


Figure 3.32: Rotational speed-discretized peak detection (orange slices) for each harmonic oscillator in the spectro-temporal domain (left, green) and extraction of discrete amplitude envelopes (right, green).

From a perception-related view, this specific pattern is responsible for the characteristic sound of an electric engine. From the time-frequency representation, only the amplitude magnitude $|\hat{A}_{p,N}|$ from each of the N harmonic oscillators is used for the synthesizer programming. For implementation of the electric engine sound

synthesizer, harmonic oscillator's amplitude values at 120 equidistant rotational speeds are stored in a look-up table (LUT) by using a matrix form. For auralization, the actual source signal of the electric engine must be a 1-channel audio stream. With regard to the decomposition ending up in multi-channel output information, additive synthesis is used in this work (Fig. 3.33). Each discretized amplitude magnitude $|\hat{A}_{p,N}|$ – that is stored in the LUT – individually scales the amplitude of the according harmonic oscillator ω_N where N denotes an integer multiple of the elementary sinusoidal oscillator's frequency ω_1.

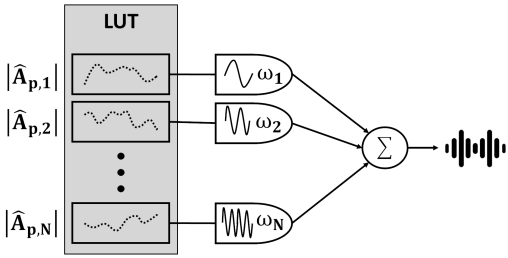


Figure 3.33: Block diagram of the synthesis step by superposition of N discretized amplitude magnitudes (left) and N sinusoidal oscillators (right).

Figure 3.34 compares the spectrograms of the reference run-up recording (left) and the synthesized result (right). The synthesis model produces an artifact-

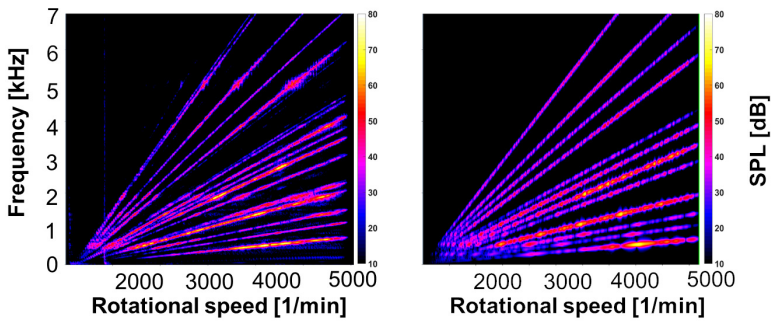


Figure 3.34: Spectrogram comparison of reference engine recording (left) and auralization synthesis result (right).

free auralization which is indistinguishable from the recorded reference machine, with exception from an unstable behavior audible in the reference for very low rotational speed.

3.7 Tire/road noise synthesis

This section proposes a procedure to model and synthesize the tire/road noise component – including all its complex aerodynamical and mechanical sound generation mechanisms (cf. appendix C). From its conceptualization, it can be based on input data obtained from established tire/road noise measurement techniques, such as the close-proximity method (CPX) [148]. In contrast to the tonal motor component being synthesized using additive synthesis (cf. section 3.6.2), tire/road noise can be implemented computationally efficient by using the subtractive synthesis technique, as introduced in section 2.6.1. In particular, this work proposes the according digital filter coefficients to be obtained by using an autoregressive (AR) modeling approach. It assumes purely stochastic input signals, i.e., free of tones (emitted by the powertrain). In order to find a compact model description of this noise, a two-step analysis-and-synthesis procedure is discussed in the following and depicted in an according flowchart (Fig. 3.35):

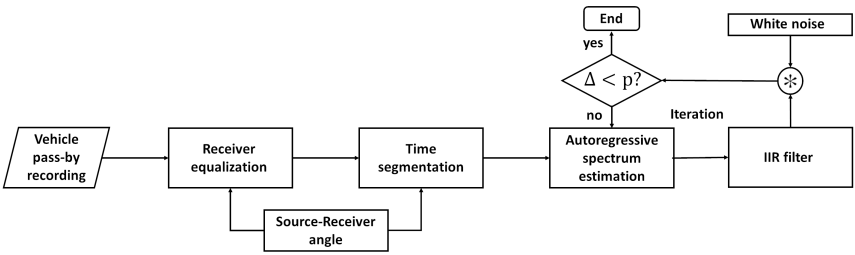


Figure 3.35: Flowchart of the iterative filter optimization approach for finding an IIR filter representation of vehicle pass-by noise.

In a first step, the statistical properties of a backpropagated and calibrated vehicle pass-by recording (see section 3.3) are analyzed. The goal of AR modeling is to determine numerical parameters of a p-th order all-pole infinite impulse response (IIR) filter that, by convolution with white noise at its input, produces a time signal with the same random process statistics as the backpropagated signal. Since the phase response of the digital filter has no audible impact on the resulting convolved signal with the white noise excitation – and the filter's group delay can be neglected for stationary excitation signals – an IIR filter structure is adequate for low-order, feature preservation and minimum storage requirements. The algorithmic procedure for finding the filter parameters is briefly described in the following.

The power spectral density (PSD) of a p-th-order autoregressive process is [64]

$$\mathrm{PSD}_{\mathrm{AR}}(e^{j\omega}) = \frac{|b(0)|^2}{\left|1 + \sum_{k=1}^{p} a(k)e^{-jk\omega}\right|^2} \tag{3.12}$$

with a and b being the IIR filter coefficients resulting from solving the Yule-Walker equations by means of the Levinson-Durbin recursion. Another advantage is that this method inherently produces a stable model. The PSD at the output of the IIR filter is given by the magnitude-squared of its frequency response multiplied by the variance of the white Gaussian noise input.

For real-time synthesis, the resulting coefficient vectors from the analysis step must be converted in a format conventional C++ libraries can handle. A common approach is the use of cascaded biquads: Therefore, the AR model transfer function is expanded into a series of second-order sections (SOS) form (Fig. 3.36).

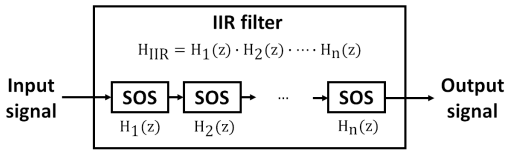


Figure 3.36: IIR filter design consisting of n cascaded second order sections (SOS).

As result, the overall IIR filter block consists of a cascaded chain of SOS, each of those being a limited number of coefficients from the solution of equation (3.12). Details on the algorithmic procedure are described in [121]. Note, depending on the auralization framework, the conversion to an alternative form by means of a parametric equalizer with high-pass, low-pass, parametric or shelving filter elements defined by its parameters frequency f, gain G and quality factor Q might be preferred.

The number of SOS elements depends on the coefficients to be handled. The required number is assessed by comparison of the vehicle's emission spectrum with synthesized noise spectra (Fig. 3.37). The vehicle emission stems from a backpropagated recording (introduced in section 3.3). At order 30, the coloration of the tire/road sound synthesis matches those of the recorded reference.

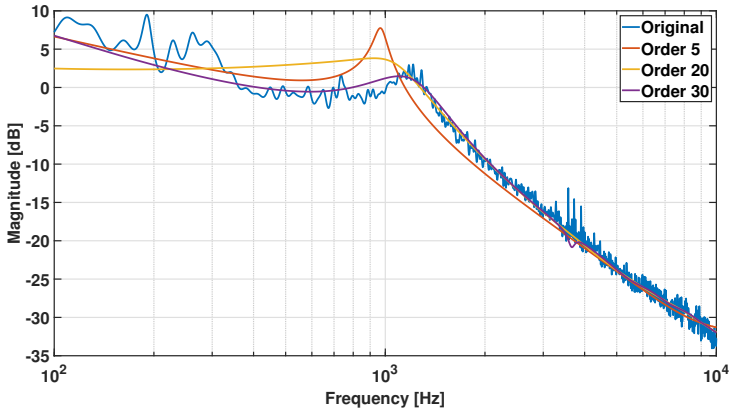


Figure 3.37: Comparison of averaged near-field radiation spectra with synthesized noise spectra of different filter orders.

3.7.1 Wet and dry asphalt filter curve

The audible impact of the surface wetness on vehicle exterior noise is significant. Although excited at the tire/road contact patch, the sound generation is not related to the mechanisms reported in table C.1. However, the surface wetness correction can be conveniently implemented to the tire/road auralization model as an optional audio filter. Shown in Figure 3.38, the filter curve is derived from two long-term road traffic sound recordings with dry and wet surface.

Based on the broadband immission spectrum for wet asphalt condition (light grey, dotted) its one-third octave band spectrum is derived (smooth grey curve), and compared to the asphalt condition (smooth black curve). The range of interest (ROI) is determined by a significantly higher energy of the wet condition in the frequency range $f > 1680$ Hz. The final audio filter curve (blue) is resulting from an optimization fitting of the difference spectrum in the frequency range $1680 < f < 20000$ Hz using a 12-band parametric equalizer. The result independently confirms the spectral difference curve from the Harmonoise project, as shown in [14]. For implementation, a table with digital filter parameters is given as parametric equalizer coefficients in appendix B.

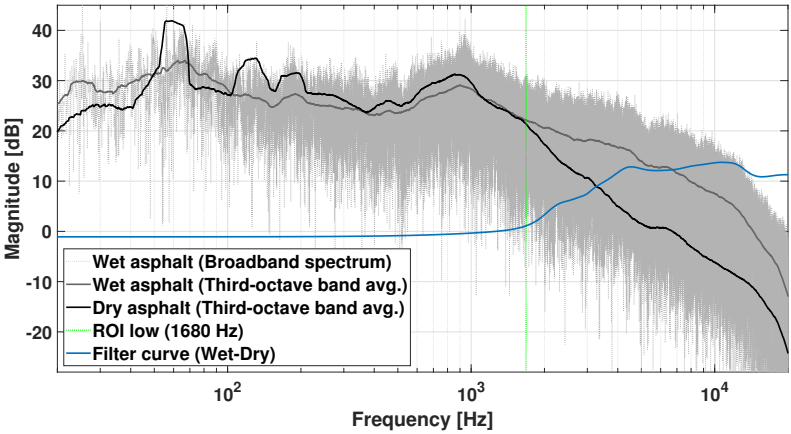


Figure 3.38: Spectral difference curve (blue) for wet (grey) and dry asphalt condition (black).

3.8 Real-time implementation as VST plugin

The proposed synthesis model for the superposition of all partial sources follows the overall goal to be a computationally efficient, considering acoustically free field conditions, being spectrally accurate and customizable model representation of an individual car's pass-by noise emissions. With regard to maximum flexibility towards automation and multiple host software compatibility, the synthesizer module is implemented in C++ as a Virtual Studio Technology (VST) plugin. This format enables plug-and-play capable integration to digital audio workstations. Furthermore, every parameter can be automated, for dynamically synthesizing auralization scenarios. As shown in Figure 3.39, the graphical user interface (GUI) includes 14 parameters, setting the operating state of the engine by the number of cylinders, cylinder size and dimensions of the exhaust system, as well as the revolutions per minute and the throttle. Furthermore, two-stroke and four-stroke engines can be selected. The level of different sound components like intake, vibrations, exhaust and tire noise are indicated by level sliders and can be automated in the hosting digital audio workstation (DAW).

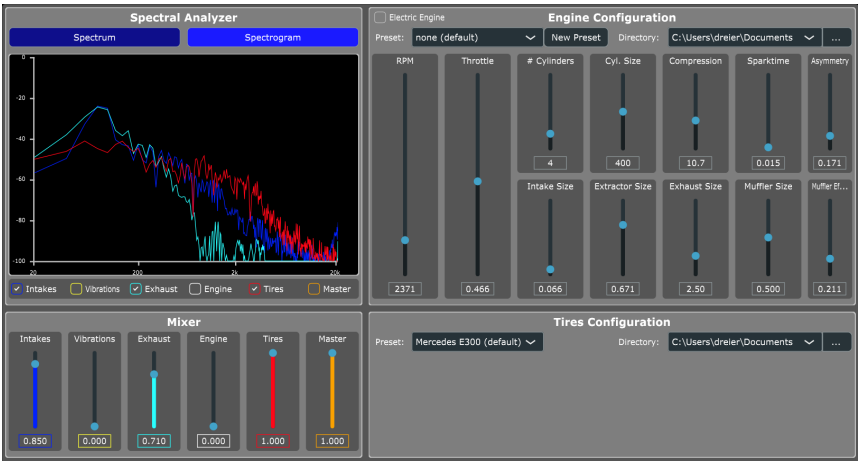


Figure 3.39: Graphical user interface of the VST plugin for road-traffic sound synthesis.

3.9 Summary and discussion

A comprehensive synthesis model was developed containing directivity patterns and sound power of seven road vehicles at ten speeds [3]. With particular novelty, an open-access database documents the great variety of directivity patterns between different vehicle class segments and speeds. First and foremost, from the obtained data it can be concluded that the directivity patterns of road vehicles depend on driving speed. This sheds new light on the numerical pattern descriptions of commonly used sound source models, such as the Harmonoise model, which merely scale the overall level according to vehicle speed. Most influential on a resulting pattern is the driving condition of the vehicle, in particular the engine speed (as a result of the gear engaged). Moreover, the higher the frequency band, the smoother the directivity pattern. In general, no clear dominancy of a certain vehicle throughout all frequency bands can be observed - meaning that the mixture of contributing engine, tire/road and aerodynamic sound sources is highly dependent on the particular vehicle. Rather than levels, the broadness of the main radiation lobes correlates to a specific car. Furthermore, the median plane patterns show symmetry to a large extent, whereas the sagittal plane patterns are more directional. It is clear from the obtained results that future road-traffic models should distinguish between the powertrain type, such as electric cars or those with ICE. Not only differences in the directivity patterns themselves, but also sound power level reductions up to 10 dB_{SWL} should be taken into

account. It is known from the literature that the level effect due to applied torque during acceleration causes variations up to 8 dB (depending on the tire design and on the applied torque), as well as increases by about 2 to 3 dB for partial versus full throttle [29]. Therefore, future work should evaluate if a variation in directivity patterns due to acceleration and load can be observed too. Furthermore, it should be evaluated if the tire-pavement combination is influencing the directivity patterns. Generally, power train noise does not show a constant relationship to vehicle speed, rather being a function of the selected gear. During the experiment presented, only the driver sat in the vehicle, therefore future work should evaluate if the emitted sound power levels significantly change in cases of more passengers and/or full trunk.

For further discovering perceptual aspects, future work should evaluate the audibility of speed-dependent directivity patterns using auralizations for listening tests. The results can be applied to urban noise auralizations, in which urban planning measures are evaluated by using audio examples.

Regarding sound signal synthesis, this chapter introduced model structures for all significant partial sound sources in road-traffic vehicles, along with their algorithmic implementations and integration into a unified tool for real-time emission signal rendering. The synthesis results, evaluated based on objective criteria such as spectral, temporal, and operating state representations, meet the sound characterization dimensions introduced in section 2.4. In future research, these results can serve as a foundation for integrating transient processes, such as the unstable load behavior of electric motors at low rotational speeds or sound generation on asphalt surfaces containing small particles like asphalt grains.

4

Aircraft auralization model

Parts of this chapter have been published in:
Dreier, C., Vogt, X., Schröder, W., Vorländer, M. (2023). Acoustic source characterization of simulated subsonic jet noise using spherical harmonics. The Journal of the Acoustical Society of America, ***154****(1), 167–178 [36], and*
Dreier, C., Vorländer, M. (2021). Sound source modelling by nonnegative matrix factorization for virtual reality applications. In: Proc. Inter-Noise. Washington DC, USA, 1053–1061 [34].

4.1 Introduction

In contrast to road-traffic noise auralization (cf. chapter 3), air-traffic noise auralization differs in two distinct ways: Firstly, the computation of sound propagation between source and receiver needs elaborated algorithms, since very large distances (several kilometers) are considered. Secondly, the overall emission mixture involves a greater number of sound sources, which are distributed and spatially separated around the fuselage and wings. The latter is relevant for the synthesis of the air-traffic sound emissions, as the shielding results in a frequency-dependent difference of their near and far field directivity patterns (see Fig. 4.1). However, these differences are not captured by semi-empirical source models, which are commonly implemented in modern air-traffic auralization frameworks. As a reference, this chapter introduces semi-empirical models in the next section 4.1.1. Moreover, this chapter will exemplarily discuss far-field directivity pattern predictions using high-fidelity flow simulations of jet noise.

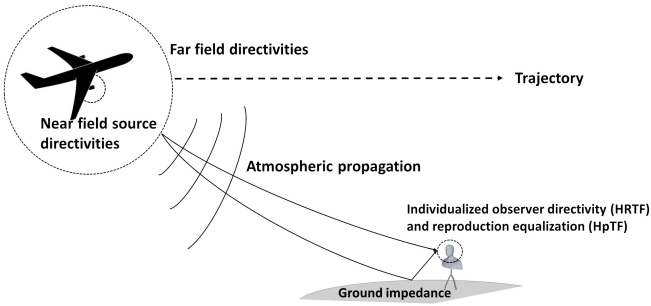


Figure 4.1: Schematic overview of air-traffic noise auralization.

4.1.1 State of the art

Over the last decades, many different prediction tools for aircraft noise have been developed. Referring to top-down and bottom-up approaches (cf. chapter 2.3.2), these tools can be separated into three groups: Empirical models (top down) rely on measurements, whereas physics-based or numerical models (bottom up) are obtained from mathematical equations with multiple parameters. By using models of the latter group, also non-existing designs and configurations can be predicted, allowing for the assessment of new aircraft configurations and identifying promising new designs. This methodology assumes that total aircraft noise can be calculated by modeling and summing all the different noise-emitting components of an aircraft. Furthermore, semi-empirical models provide a good compromise between fidelity and flexibility. Due to the increased scientific complexity, these models are more commonly found in research applications.

Semi-empirical prediction models

The most prominent prediction tool is the Aircraft Noise Prediction Program (ANOPP) developed by NASA beginning in the 1970s and 1980s. The main noise components predicted in this method are fan noise, jet noise, and airframe noise. The according source components are shown in Figure 4.2. ANOPP was updated continuously allowing for more noise sources to be considered, such as turbine noise or core noise. NASA introduced ANOPP2 in 2011 in order to account for noise prediction of non-conventional aircraft, too [112]. Other notable theoretical noise prediction tools are CARMEN by the Office national d'études et de recherches aérospatiales (ONERA) [149], PANAM by the Deutsches Zentrum

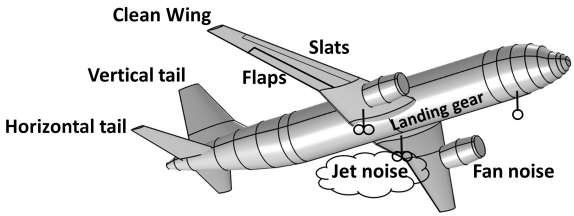


Figure 4.2: Semi-empirical prediction model components.

für Luft- und Raumfahrt (DLR) [19] and SAFT by the Chalmers University of Technology and the KTH Royal Institute of Technology [169].

In the following, established semi-empirical prediction models for fan noise by Heidmann [67], airframe noise by Fink [48] and jet noise by Stone [164] are used to calculate the directivity patterns and power spectral densities of the various partial noise sources of an aircraft. Several assumptions are made in semi-empirical models: atmospheric attenuation and ground reflections are neglected. Moreover, the receiver is assumed to be far away, thus the noise of all sources is assumed to radiate from a single point.

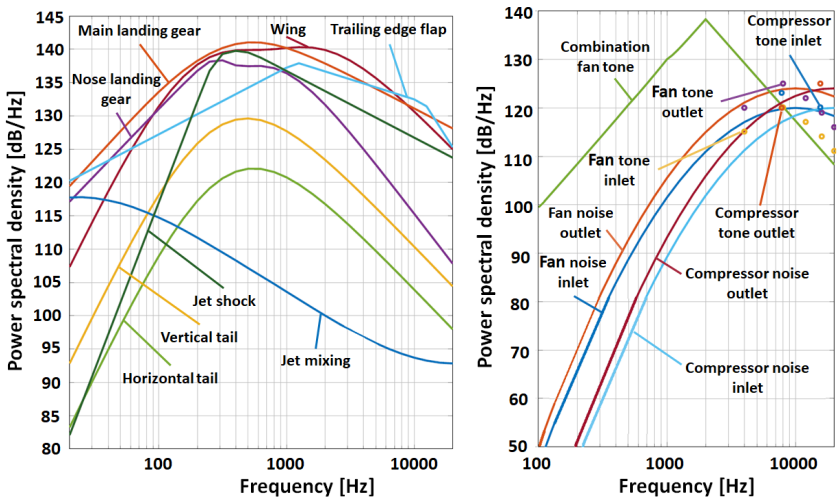
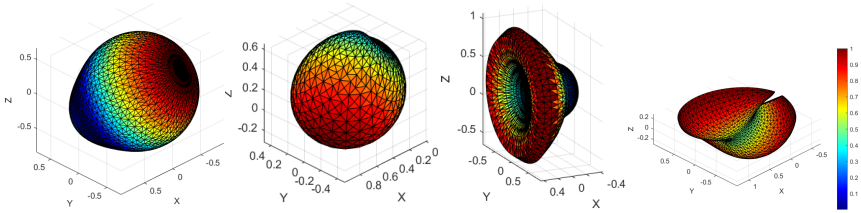


Figure 4.3: Power spectral densities of the partial sound source emissions for jet/airframe noise (left) and fan/compressor noise (right), respectively.

Referring to the source characterization dimensions being introduced in section 2.4, the objective of all semi-empirical models is to characterize the spatio-spectral emission in one-third octave bands and dependent on specific operating states. Note, the temporal properties are not covered in these models, and therefore have to be additionally synthesized (cf. section 4.4) for auralization. All prediction models require a broad range of input parameters, for example, mass flow rate of air passing through fan or compressor, span and area of wings, and jet speed and temperature. According to the aforementioned models, Figure 4.3 visually compares the power spectral densities (PSD) of all model contributors for jet and airframe noise (left) as well as fan noise emission (right). The PSDs were exemplary computed according to the input parameter set summarized in appendix D.

Moreover, the predicted directivity patterns of jet, trailing edge flap, combustion fan tone and nose landing gear are exemplarily shown in Figure 4.4 at their dominant emission frequency ranges.



(a) Jet mixing 50 Hz **(b)** Trailing edge flap 1000 Hz **(c)** Combination fan tone, 8000 Hz **(d)** Nose landing gear 250 Hz

Figure 4.4: Exemplary semi-empirical aircraft directivity patterns shown in different normalized one-third octave bands. The positive x-axis shows into flight direction, the positive z-axis vertically upwards.

Semi-empirical models are an elaborated tool for creating noise forecasts of individual flight events and standardized noise maps. As already mentioned, they are not suitable for direct auralization due to their lack of temporal dimension. In addition, the empirically observed sound spectra are based on a fixed set of design configurations that are more or less applicable in individual cases.

4.1.2 Contribution of this work

As subsonic jets remain one of the major contributions to aircraft noise emissions, near-field flow simulations should be included in aircraft design at an early stage

using quantitatively predicted sound pressure levels and the time-domain signal properties of the noise data. In this regard, the interface from the near-field data to the far-field radiation – under consideration of acoustical reflections from objects such as fuselage and wings – remains as bottleneck. This study presents the computation of a spherical equivalent source model of jet noise with minimal complexity by means of spherical harmonic coefficients. Using spherical Hankel extrapolation of sound pressure data from virtual, concentrical microphone arrays, the results of the determination for the radius, in which all acoustical sources of a flow field are confined, indicate the source radius around the end of the potential core to be equivalent five times the nozzle diameter. The result of the spherical harmonic transform shows the dominant energy contribution to be related to nine elementary sources. The resulting equivalent source model of jet noise provides a convenient format for further use in large-scale computational fluid dynamics simulations.

Regulatory frameworks to determine and reduce environmental noise exposure, such as the Environmental Noise Directive by the European Union [30], have led to comprehensive research on possible countermeasures, such as high-bypass ratio engines, acoustic liners or engine-over-wing concepts. In the field of aircraft noise emissions the focus is in particular on jet noise contributions. As well known for jet mixing noise, the proportional relationship between a jet flow velocity v and its sound intensity I is $I \sim v^8$ [105]. A decreasing, but asymptotically limited trend can be observed in the reduction of effective perceived noise levels over the last few decades, i.e. further noise reductions at the sources are harder to achieve [58]. Beyond reducing objective, level-based noise measures, perception-based approaches emerged. Therefore, psychoacoustical parameters, e.g. loudness and sharpness – which are affected by the spectral weighting between low and high frequencies – are considered in an early stage of aircraft design and used as an optimization criterion [146]. The impact of psychoacoustical optimization is commonly assessed in listening experiments based on auralized signals, i.e. audible sound at a virtual receiver from numerically simulated, measured or synthesized data [177]. In the context of aircraft noise emissions, auralizations can reproduce the spectro-temporal characteristics under consideration of weather-dependent atmospheric propagation [35] and urban surroundings including buildings [151].

Isolated measurements or simulations of near-field jet noise can quantify sound pressure levels and qualitatively describe their spatio-temporal emission characteristics. However, the computationally efficient prediction from these near-field data to the far-field characteristics – i.e. including the effect of reflections on fuselage and wing surfaces – remains a complex task and therefore has to be numerically approximated [162][137]. This chapter, therefore focuses on the possibility to

extrapolate the data to other radii by expressing the source characterization of simulated jet noise in a computationally efficient form based on spherical harmonics (SH). In order to find a minimal set of equivalent sources the radius in which all acoustical sources are confined is derived using the Hankel extrapolation technique with subsequent computation of order-dependent spatial errors. The resulting SH coefficients describe the strength of equivalent elementary sources (monopoles, dipoles, etc.), an according conversion procedure is described in [190]. Typically, these multipoles can be directly applied to many different simulation tools, such as wave-based aeroacoustical methods (e.g. FEM or discontinuous Galerkin) [28] or in directivity encoders for virtual reality applications, which is demonstrated as a proof of concept in section 4.3. This chapter is related to three research fields: flow noise generation mechanisms based on numerically simulated data, its measurement with spatial microphone arrays and its modeling using equivalent source descriptions.

Since Lighthill's paper on aeroacoustic analogies in which the general solution to the wave equation for a medium including turbulence and steady scattering objects is derived [105] numerous theoretical publications on jet noise sources have been published. The most intriguing research articles in this field have been reviewed by Tam [166]. Based on Lighthill's work, Proudman showed that the acoustical power of jet noise can be parameterized by the integral scale of turbulence and the turbulent Mach number under the assumption of isotropic turbulence [138]. To overcome the drawback of Lighthill's equation, i.e. interpreting purely kinematic processes such as refraction at shear or boundary layers as acoustical sources, Lilley discussed the influence of nonlinear compared to linear terms on acoustic convection and refraction [106]. Based on Lilley's analogy equation, Tester established a parametrical model to predict jet noise dependent on flow velocity and medium density [170].

Furthermore, research publications are focusing on the measurement-based studies of sound generation mechanisms: Laufer identified two main noise-producing regions for supersonic jets, an area downstream of the potential cone (comparable to subsonic jets) and an upstream region exhibiting Mach wave radiation [99]. Panda investigated the correlation between medium density fluctuations and far-field sound pressure fluctuations for jets with varying Mach number [130]. From the radiated sound of two coaxial jets (a normal-velocity profile representing a turbofan engine and the inverted-velocity profile configuration), Tanna concluded that the velocity ratio between primary and secondary flow is more dominating than the temperature ratio [168]. In a joint publication, some of the aforementioned authors point out the sound spectrum of jet noise to be domi-

nated in the low-frequency range by the large-scale turbulent structures and in the high-frequency range by small-scale turbulent structures inside the jet [167].

In addition to measurement-based research, numerical simulations revealed further insights: Based on large-eddy simulations (LES), Bogey studied the effect of the turbulence intensity at the inlet boundary condition on the shear layer development and the radiated sound field [22]. Using a two-step procedure – an LES coupled with a solution of the Ffowcs Williams-Hawkings (FWH) equation – Spalart investigated the spatial source locations for different fixed frequencies from sound pressure data and noise transmission through turbulent region [161]. Koh computed two coaxial jet configurations, with secondary streams identical to that of a single jet configuration, in order to show the impact of the additional shear layer on the sound field and revealed the Lamb vector fluctuations to represent the major acoustical sources of the isothermal jet and evaluating the additional noise source due to temperature gradient for the hot coaxial jet [95]. Koh employs a two-step method by combination of LES and acoustic perturbation equations (APE, [42]) to simulate the sound field, the same simulation approach as used in this chapter.

With particular focus on three-dimensional source characterization for further use in wave-based acoustic simulations and acoustic virtual reality applications, spherical harmonic (SH) transform is used in this chapter. The ideas of spatial sound field decomposition recently faced a renaissance driven by research on real-time spatial audio formats. Motivated by computationally efficient jet noise predictions, Mungur at first documented in 1974 the characterization of jet noise by means of spherical harmonics as solution of a modified wave equation including flow and flow gradient terms [125]. Two common ways to analyze jet noise source locations from the far field are the acoustic telescope [20] and the polar correlation technique [49]. Both techniques extract from 1-D microphone array measurements a sound field representation in terms of an equivalent line source density along the downstream axis of the jet. Using a cylindrical microphone array around the jet flow, Lee and Bolton [101] experimentally investigated the far-field radiation pattern of a subsonic jet with Mach number 0.26 based on cylindrical near-field acoustical holography (NAH) and subsequently expanded the measured sound field into its helical wave spectrum. This research revealed the analyzed jet to be composed of eleven uncorrelated dipole-, quadrupole-, and octupole-like elementary sources. Mathematically similar to cylindrical NAH, further studies on jet noise source distributions were based on Helmholtz Equation Least Squares (HELS) method proposed by Wang and Wu [179]. Both procedures make use of orthogonal expansions fulfilling the wave equation. Whereas in the NAH case the transform is in terms of helical wave functions, spherical wave functions are

used in HELS. Since the spherical harmonic coefficients are found by using a least-squares solution – instead of exploiting the orthogonality of the spherical harmonics over a spherical surface – the surface not necessarily has to be spherical and thus can accommodate arbitrarily shaped sources. Using HELS, Mazza [118] presented an experimental cylindrical array measurement around an installed engine's subsonic jet flow under consideration of an optimized array for high frequencies. The spherical holography procedure suggested by Maynard et al. [117] exploits the orthogonality of the spherical harmonics over a spherical surface. This forms the approach used in this chapter by taking the sound pressures on 8 concentric spherical surfaces with a fixed center coordinate including an optimized measurement grid for enabling the direct inversion of the equation system instead of being dependent on taking the pseudoinverse and least-squares error optimization.

In foreground of the chosen simulation-based approach, practical limitations of measurement-based methods should be mentioned. Inaccuracies in microphone positioning and signal-to-noise ratio limitations due to sensor noise and the acoustical properties of the measurement room not only affect the analysis resolution but result in error amplification due to matrix inversion in the SH transform. Furthermore, measured data for source localization using beamforming approaches include the effects of jet shear layer refraction and advection making the sources appear to emanate from further downstream [5]. Therefore, the fully simulated approach is used in this study. Beyond the scope of the presented research literature, this chapter covers the following novelty aspects:

- By using a simulation-based approach full-spherical sound pressure data from the complete three-dimensional flow field are considered.
- Analysis of the influence of the microphone array design on spatial and spectral resolution.
- Derivation of a spherical equivalent source model of jet noise from computation of spherical harmonic coefficients and according order-dependent errors.
- Derivation of the acoustic source dimension using multiple concentric spherical microphone arrays and Hankel extrapolation technique.
- The simulation-based approach avoids errors due to sensor-to-field interactions as well as sensor positioning inaccuracies, sensor mismatching and limited signal-to-noise ratio of the measurement equipment, hence minimizing numerical error amplification in the matrix inversion during SH transform.

This chapter is structured as follows: Section 4.2 summarizes the numerical flow field simulation and subsequent sound field computation. In section 4.2.1, the theoretical background of the sound field decomposition is introduced and a combined measure for spatial and power spectrum similarities between the original and the order-reduced flow fields is defined. Furthermore, the virtual microphone array design is presented and analyzed concerning its properties on spectral and spatial resolutions. The determination of the acoustic source radius based on Hankel extrapolation is described in section 4.2.2. In section 4.2.3, the jet noise source is acoustically characterized based on spherical harmonics and an order-reduced equivalent source model is computed. Finally, the results are discussed in section 4.5.

4.2 Source characterization of simulated subsonic jet noise

Numerical flow and acoustic field simulation

In this study, the flow field of the baseline geometry SMC000 of the Small Metal Chevrons [25] is investigated (Fig. 4.5). This specific convergent, axisymmetric nozzle is defined by its conical contraction of 5°. The computational domain for

Figure 4.5: Geometry of Small Metal Chevron nozzle SMC000.

the flow simulation has an extent of 48 D in the streamwise and 36 D in the crossflow directions (Fig. 4.6), with D being the nozzle diameter. The computational mesh contains several refinement patches with the highest resolution in the vicinity of the nozzle surface and in the near field of the nozzle outlet to enable the capturing of the turbulent small scale structures in this region. The minimum cell length of the grid is $\Delta c = 1.56 \cdot 10^{-2}$ D resulting in 64 cells along the nozzle diameter. In total the mesh for the turbulent jet flow simulation contains 168 million cells.

The acoustical sources are determined inside a domain extending over 10 D in the axial and 5.5 D in the crossflow directions. The finest grid resolution matches the resolution of the grid used for the LES. The overall acoustic domain extends over 30 D in the streamwise and 36 D in the normal and spanwise direction

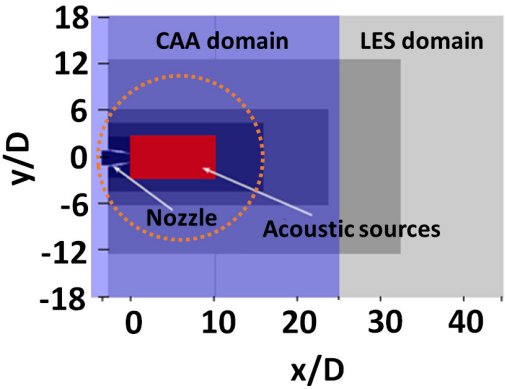


Figure 4.6: Two-dimensional cross section of the LES and CAA simulation areas. The dotted orange circle corresponds to the outermost sampling sphere in Figure 4.9.

and contains 5 refinement levels resulting in a computational grid containing 327 million grid points. To obtain the underlying simulation data used throughout this study, a hybrid computational fluid dynamics (CFD)/computational aeroacoustics (CAA) approach is used. In a first step, the unsteady compressible flow field is computed on an unstructured Cartesian grid using a highly resolved LES. Afterwards the acoustical field is computed by solving the APE [42], whose source terms are determined by the LES. Further information about the numerical framework and flow conditions can be found in the according journal paper [36].

Simulation result

The LES is run for approximately 560,000 timesteps, i.e., a simulation time of $930\,\mathrm{D}/u_j$, ensuring a statistically converged flow field. Subsequently, the simulation is continued for another 140,000 time steps with samples of the simulation data written out every 7th timestep, leading to 20,000 samples used for the calculation of the acoustical sources. The result of the sound source for the SMC000 is shown in Figure 4.7.

The noise production inside the shear layer can be clearly identified, whereas the sound source inside the potential core is negligibly small. Although the location and magnitude of the noise source is determined, further analysis is essential to calculate the propagation of the pressure waves to the far field and to identify the directivity and frequency spectrum of the noise. The calculated noise sources

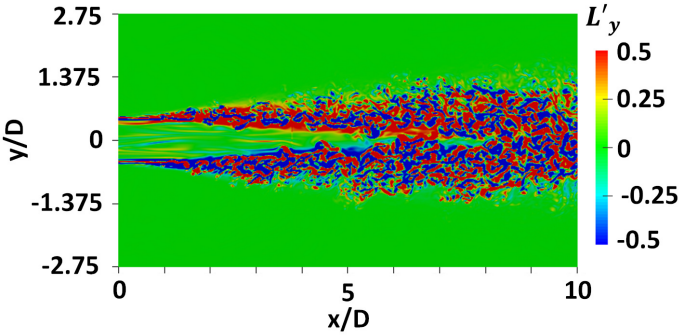


Figure 4.7: Visualization of the Lamb vector (Two-dimensional cross section).

are used as input data for the next step of the two-step LES-CAA method, i.e., the determination of the noise field. The acoustics computation is carried out over 175,000 time steps with a sampling interval of 50 time steps. The pressure waves travelling from the nozzle exit towards the far-field are shown exemplarily in Figure 4.8.

The CFD simulation flow solver used in this study has been extensively validated, e.g. by Hartmann et al. [62] and Schneiders et al. [155]. The acoustics solver has been validated by Ewert and Schröder, e.g. in [42] and [43]. Besides, the general capability of the hybrid CFD/CAA method for noise prediction, the applicability of the flow and acoustics solver to specific flow problems has been proven in several studies, e.g. by Gröschel et al. [60] and Koh et al.[94]. For details concerning the validation of the simulated chevron nozzle with references found in the literature the reader is referred to Pauz et al. [131].

4.2.1 Spatial sound field decomposition

The simulated flow field sound pressure data are exported on concentrical spheres around the end of the flow field's potential core (Fig. 4.9). The time-domain sound pressure signals p_i are taken from virtual spherical microphone arrays on each sphere. To further extend the source characterization with freely selectable distances, a SH transformation approach is applied. This approach also allows a detailed analysis of an equivalent source arrangement, which can serve as a basis for the calculation of the nozzle noise scattered at the aircraft fuselage and wings.

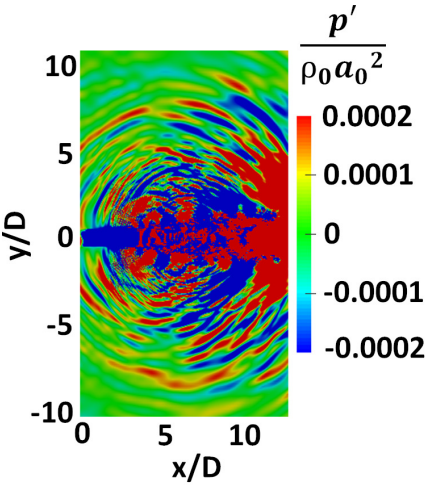


Figure 4.8: Visualization of the normalized sound pressure (Two-dimensional cross section).

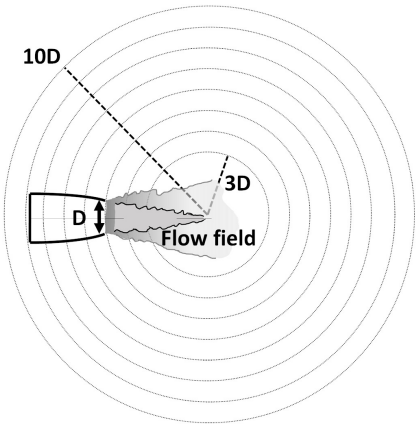


Figure 4.9: Simulation setup: The sound pressure data p_i are taken on concentrical spherical surfaces around the end of the flow field's potential core.

Spherical harmonic transform

Since the sound pressures p_i are obtained on a spherical surface, a mathematical expansion of the simulated discrete sound field points on the sphere into spherical harmonics exists. The CFD simulated data are given in form of real-valued sound pressure data so that the real spherical harmonic transform coefficients can be directly obtained [136]. The real-valued spherical harmonics $Y_\mathrm{n}^\mathrm{m}(\theta, \phi)$ of (positive) order n and degree m are defined as introduced and plotted in section 2.4.3.

Design of the virtual microphone array

The positions at which the sound pressure field on a spherical surface is evaluated are subject to different spatial sampling schemes. Each sampling scheme has trade-offs between aliasing error, sampling efficiency and condition number [189]. Instead of approximating the integration over the sphere, a hyperinterpolation sampling was derived in [183] for exact polynomial interpolation on the 2-sphere, exemplarily shown for order $\mathrm{N} = 30$ in Figure 4.10. In consequence, the corresponding spherical harmonic transform is exact for order-limited functions on the sphere. Since the transformation matrix $\mathbf{Y}$ is square for this sampling, the SH transform can be calculated according to eq. (2.14) using $\mathbf{Y}^{-1}$, hence giving an exact transform. The spherical basis functions are not orthonormal, i.e ${\mathbf{Y}_\mathrm{N}}^T \neq {\mathbf{Y}_\mathrm{N}}^{-1}$. It has to be noted that $\mathbf{Y}$ needs to be well conditioned as otherwise the solution might not be exact anymore, as will be discussed below.

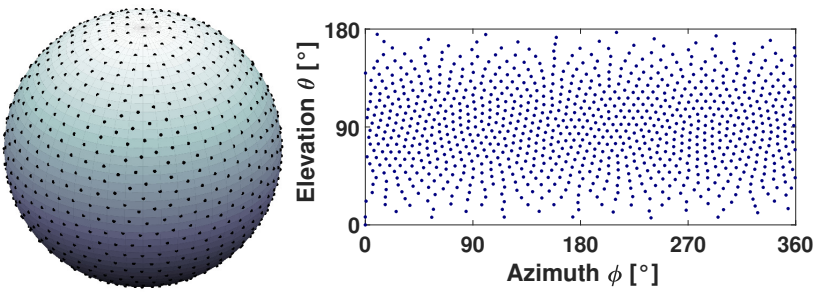


Figure 4.10: Discretization of each spherical array in Figure 4.9 using hyperinterpolation samplings. The shown example is a sampling grid of spherical harmonics order $\mathrm{N} = 30$.

Sampling-induced numerical error

The condition number $\kappa(\mathbf{A})$ is used as a measure to evaluate the feasibilty of a matrix $\mathbf{A}$. Generally, for a given system of linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{4.1}$$

where $\mathbf{A}$ and $\mathbf{b}$ are known, the error in the solution of $\mathbf{x}$ can be estimated as

$$\frac{||\Delta\mathbf{x}||}{||\mathbf{x}||} \leqslant \kappa \frac{||\Delta\mathbf{b}||}{||\mathbf{b}||} \tag{4.2}$$

for a given error in $\mathbf{b}$ and the condition number

$$\kappa(\mathbf{A}) = ||\mathbf{A}|| \cdot ||\mathbf{A}^{-1}||. \tag{4.3}$$

With regard to the matrix inversion of the spherical harmonic transform the condition number $\kappa(\mathbf{Y})$ can be applied to assess the feasibility of the SH base matrix $\mathbf{Y}$. The condition number of the spherical base matrix for the hyperinterpolation sampling with order 80 is $\kappa_{\text{hyper},80} = 12.76$. The simulated sound pressure data are virtually noise-free so that the numerical resolution is only limited due to rounding errors of the used data format. Given the computational accuracy of the double-precision floating-point format, i.e. having 16 decimal digits of accuracy, a condition number value of 12.76 can be considered as a negligible factor. Related to the machine precision $\epsilon = 2.22e - 16$, critical condition number values would exceed $\kappa_{\text{crit}} > 1/\sqrt{\epsilon} \approx 6.7e7$. For the given condition number according to the rule of thumb from [26]

$$\frac{\ln 12.76}{\ln 10} \approx 1.1 \tag{4.4}$$

approximately a single number of precision is lost in the spherical harmonic transform.

Spatial aliasing

Generally, the performance of a microphone array is limited depending on its sampling scheme in frequency resolution and spatial accuracy. Depending on the sampling scheme, artifacts due to the spatial discretization, as typically known from time domain sampling, are more or less prominent [139]. Concerning these aspects, in this study a hyperinterpolation sampling of order $\mathrm{N} = 80$ with $(\mathrm{N}+1)^2 = 6561$ sampling points is used. The specific features of the hyperinterpolation sampling comprises a sampling efficiency of 100%, i.e. the number of

spherical wave spectrum coefficients equals the number of sampling points as well as a square spherical base matrix $\mathbf{Y}$ that can be directly inverted. The sampling scheme allows a theoretically ideal reconstruction of a spherical expansion at every node, if the sampling nodes provide a well-conditioned matrix $\mathbf{Y}_{\mathrm{N}}$ [189]. Mathematically, an ideal and aliasing-free sampling with a limited maximum order N would result in the orthonormality matrix $\mathbf{O} = \mathrm{Y}^{-1}\mathrm{Y}$ being the identity matrix. An aliasing matrix is used to analyze the magnitude of the relation between each coefficient f_{nm} and $f_{\mathrm{n'm'}}$ and in consequence how the orders n' > N are reflected into the coefficients n.

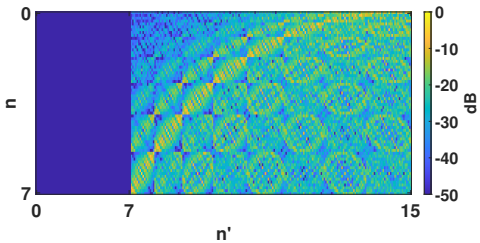


Figure 4.11: Spatial aliasing matrix for a hyperinterpolation sampling of order n = 7 and n' = 15.

For better visual clarity, the according aliasing plot in Figure 4.11 shows the aliasing effect at the order-reduced example by using a 7th order hyperinterpolation sampling (compared to the 80th order sampling used in the actual study). However, the general effect applies for hyperinterpolation samplings of any order. It can be observed that the sampling scheme is aliasing-free for SH orders n' < N, but shows strong spatial aliasing for SH orders n' > N.

Spectral aliasing

The magnitude and phases responses of the spherical array with a radius of the outermost sphere, i.e. r= 10 D, were simulated by an incident plane wave field expansion. The results in Figure 4.12 show the array to be spectrally flat below the 9 kHz cut-off frequency and having a $\pm 8°$ phase accuracy. Therefore, the array itself exhibits a negligibly small influence on the simulated data in the frequency range of $f < 2$ kHz.

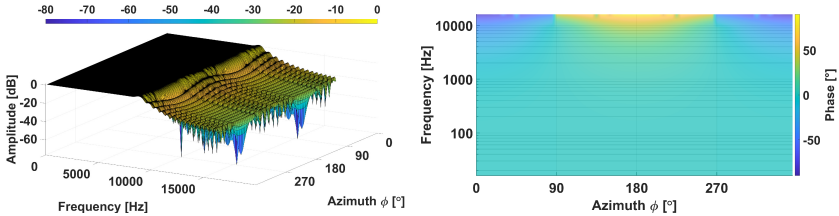


Figure 4.12: Magnitude (a) and Phase (b) response of the open spherical microphone array with an 80th order hyperinterpolation sampling with radius of the outermost sphere (r=10 D).

Spatial error measure

The order of truncation error of the spherical harmonic transform restricts the spatial resolution, but also affects its spectrum [11]. For quantification of the order-dependent error a measure is introduced that calculates the area-weighted spatial standard deviation between the original and the order-reduced flow field. Therefore, a modified formula was derived from Leishman's approach [102]. For each order of truncation error the area-weighted spatial standard deviation $\sigma(f)$ is calculated by

$$\sigma(f) = \sqrt{\frac{\sum_{\mathrm{g}=0}^{\mathrm{G}-1}\sum_{\mathrm{t}=0}^{\mathrm{T}-1} S_{\mathrm{g,t}}[L_{\mathrm{g,t}}(f)_{\mathrm{max}} - L_{\mathrm{g,t}}(f)_{\mathrm{min}}]^2}{\sum_{\mathrm{g}=0}^{\mathrm{G}-1}\sum_{\mathrm{t}=0}^{\mathrm{T}-1} S_{\mathrm{g,t}}}} \tag{4.5}$$

with $S_{\mathrm{g,t}}$ being the area weighting factor of each sampling point, i.e. the effective area per microphone on each sphere (resulting from Delaunay triangulation of the spatial sampling), L_{max} and L_{min} denote the smaller and greater value for each compared pair, derived from

$$L_{\mathrm{g,t}}(f) = 20\log_{10}\left[\ |H_{\mathrm{g,t}}(f)|\ \right] \tag{4.6}$$

where $|H_{\mathrm{g,t}}(f)|$ is the modulus of the complex frequency response function at the point g,t on the spherical array.

4.2.2 Jet noise source localization

As reference for the spatial error comparisons, the power spectral density (PSD) on the largest simulated sphere ($r = 10\,\mathrm{D}$) is shown in Figure 4.13 with a spatial segmentation of the measurement grid into concentric circles of 10° around the flow field axis. For PSD estimation, Welch's method [180] is used as introduced in section 2.4.1.

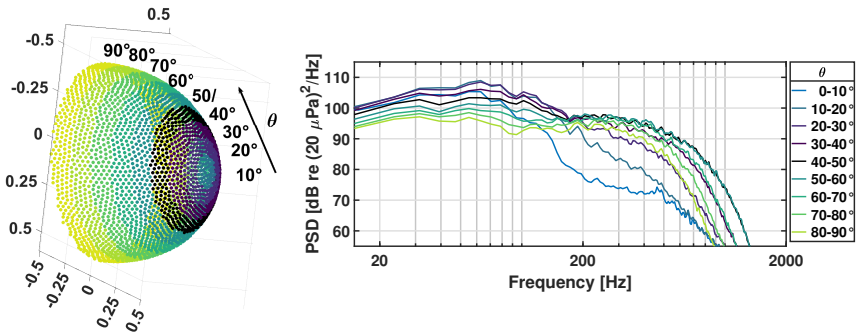


Figure 4.13: Downstream directivity of the simulated flow noise in radial segments with 10 degrees.

The downstream direction is defined as $\theta = 0°$, whereas the orthogonal elevation angle is defined as $\theta = 90°$. The following trend can be derived. For very low frequencies below 100 Hz the major sound power is radiated downstream ($\theta = 0°$) and in crossflow direction ($\theta = 90°$). In a transition frequency range between 100 and 200 Hz, the higher the frequency the more these two lobes approach so that the most sound power is radiated into directions of a narrow cone and a wide cone. This trend converges for frequencies higher than 200 Hz. In this spectral range it can be observed that the highest sound power is radiated to elevation angles between $40° < \theta < 60°$, with emphasis towards higher elevation angles. Specifically in this angular range, the emitted noise power is nearly constant up to an upper cut-off frequency (defined at the 10 dB roll-off) at about 700 Hz. In the range $f > 500$ Hz, with increasing frequencies the side lobes are vanishing towards a rather omnidirectional directivity with decreasing overall power.

The radiated sound power level of the simulated jet noise is calculated according to eq. (2.1). The simulated jet's actual sound power evaluation yields $L_\mathrm{W} = 132.07\,\mathrm{dB}$ or $P_\mathrm{ak} = 16.13\,\mathrm{W}$.

Hankel range extrapolation

In order to find the radius where a sphere fully confines the turbulent regions of the flow field - and thus the acoustic source terms - an analytical far-field extrapolation by means of Hankel functions is used in the domain of spherical wave functions (cf. 2.4.3). For this purpose, the spherical harmonic coefficients are calculated using eq. (2.14) and analytically extrapolated using eq. (4.7). The Hankel extrapolation technique is used to calculate back and forth between a sound field's near and far field in spatial domain. In case of an exterior problem, where all acoustical sources are confined within a sphere of radius r_0 and additional free field conditions in the vicinity of this sphere are assumed, the acoustical near field at r_0 can theoretically be extrapolated to any far-field point with greater radius $r_1 > r_0$ in terms of its spherical wave spectrum by

$$f_{\mathrm{nm}}(r_1, k) = f_{\mathrm{nm}}(r_0, k) \cdot \frac{h_{\mathrm{n}}^{(2)}(kr_1)}{h_{\mathrm{n}}^{(2)}(kr_0)}. \tag{4.7}$$

using the result from from eq. (2.15) and the second-order spherical Hankel function $h_{\mathrm{n}}^{(2)}$. The Hankel function is the only distance-dependent part so that the ratio of both Hankel functions at radial distances r_0 and r_1 can be multiplied with the SH coefficients obtained at the source radius $f_{\mathrm{nm}}(r_0, k)$. The behavior of this ratio can be considered as an order-dependent as well as frequency-dependent decay that is examplarily shown up to order $\mathrm{n} = 5$ in Figure 4.14 for the case of a radial doubling $r_1 = 2 \cdot r_0$, hence corresponding to a far-field decay of -6dB. The asymptotic behavior of the magnitudes of the spherical Hankel functions is clearly visible. Their near- and far-field regions can be distinguished. In the near field of the sound sources higher-order components decay more rapidly with increasing radius whereas in the far field all orders show very similar radial decay. It can be concluded that sources of order n must be $(6 \cdot \mathrm{n})$ dB times louder than a monopole in proximity of the source to generate the same far-field sound pressure.

In order to find a minimal equivalent source model for jet noise by means of spherical harmonic coefficients, the basic idea can be summarized as follows: Since smaller source radii generally correspond to smaller numbers of SH coefficients, the minimal source radius, in which all acoustical sources are confined, must be derived. Since the actual values on the next larger sphere are known from the simulation (or), they can be used as a reference and compared – by using the spatial error measure introduced in eq. (4.5) – with the extrapolated values (ex) of the smaller sphere. If all sound sources are confined, both sound fields should be identical in case of a non-moving medium.

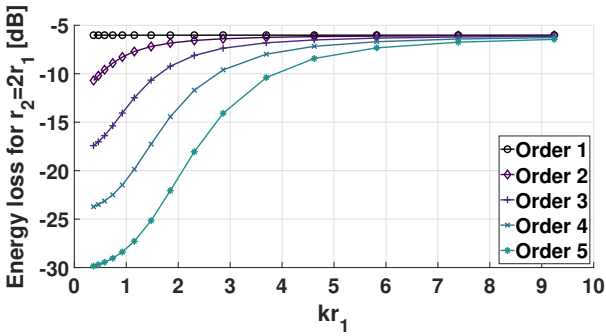


Figure 4.14: Decay of the spherical Hankel functions up to fifth order as a function of wavenumber and radius.

The results are shown in Figure 4.15. Each curve represents the error between an analytical freefield spreading loss prediction propagating from a source region with a non-freefield, turbulent and advected propagation simulation result. The latter stochastic processes remain in the input data for the comparison of the sound field data by means of eq. (4.5) and result in the non-vanishing spatial error. However, the trend indicates the acoustical sources to be confined within a radius of 5D around the end of the potential core.

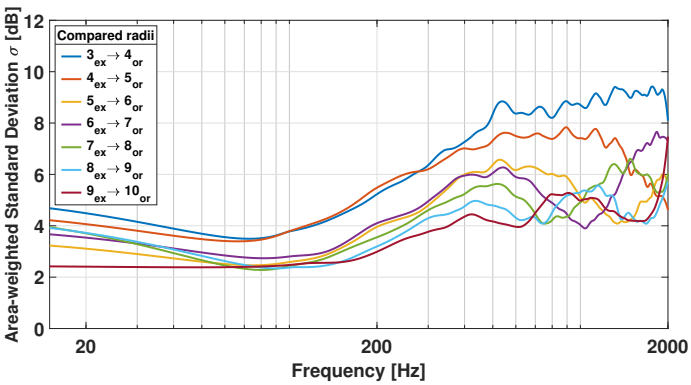
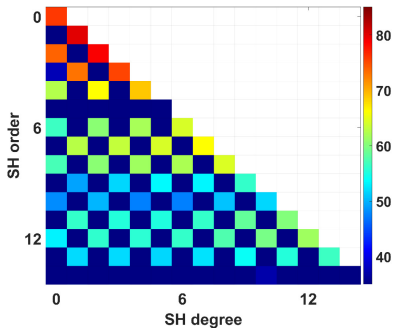


Figure 4.15: Area-weighted standard deviation between extrapolated sound fields (ex) and original sound fields (or).

4.2.3 Equivalent source model of jet noise

The relationship for the minimum number of SH orders of given directivity can be estimated by a rule-of-thumb formula with $N_{\text{trunc}} = \lfloor kr_{\text{min}} \rfloor + 2$ [124]. Assuming a minimum source radius of $r_{\text{source, min}} = 5.2 \cdot 0.0508\,\text{m} \approx 0.264\,\text{m}$ based on the results of 5, a maximum frequency of 1500 Hz (Fig. 4.14) and sound speed $c = 340\,\frac{\text{m}}{\text{s}}$ yields $N_{\text{trunc}} \approx 9$. Figure 4.16 shows the result of the spherical harmonic transform by means of the individual strengths of the spherical harmonic coefficients. The energy for orders $n > 13$ can be assumed as negligible so that the dominant excerpt of the full matrix is plotted. A trend of decreasing energy towards higher orders can be observed. As can generally be expected from the directivity, the elementary radiator of each SH order that is focused on in the downstream direction contains the most energy. To be more precise, the first order SH coefficient (dipole) in the downstream direction is the strongest source within the full matrix. Interesting to note, the fifth order SH only consists of weak elementary sources. As basis for the order-reduced reconstruction model



75.86				
	80.05			
73.86		78.76		
	73.08		75.27	
62.43		66.10		68.87

Figure 4.16: Logarithmically-scaled spherical harmonic coefficients of jet noise (left). In this detailed view the coefficients up to order 14 (of 80) and positive SH degrees are shown. The table shows the most dominant nine (of 6561) numerical coefficients for the order-reduced reconstruction model (right).

from this matrix and as shown in the according table in Figure 4.16, the nine most dominant elementary radiators were selected.

The results in Figure 4.17 show that jet noise cannot be fully reproduced by an array limited to low elementary radiator orders; rather, the accepted reproduction error is crucial. In the case that the sound field originally described with 6561 (order 80) elementary radiators is reproduced by $40^2 - 1 = 1599$ sources (order 40), the sound field can be reconstructed in the complete frequency range up to the cut-

off frequency of 600 Hz with a negligible error of < 1 dB. By dominance analysis of the original 6561 sound sources, a convenient equivalent source description based on nine elementary radiators was derived.

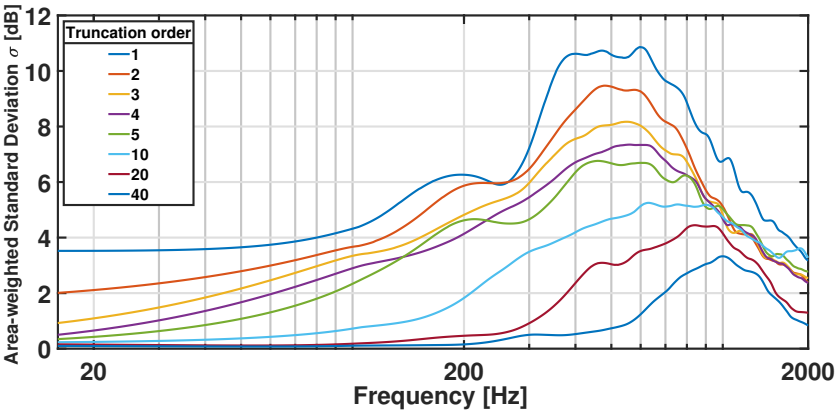


Figure 4.17: Order-dependent area-weighted standard deviation between the original and the order-reduced flow field model.

From the decomposed sound field an order-reduced jet noise model with less coefficients is derived. The resulting directivity of the order-reduced jet is shown in Figure 4.18. In comparison to the original directivity (Fig. 4.14), two frequency regions can be distinguished: Whereas the directivity is well reconstructed up to 200 Hz it smears out the conical shape of the original for higher frequencies. This specific equivalent source description approximates the original directivity well over the entire spectrum, but accepts errors in the radiated elevation angle between $40 < \theta < 60°$. The reconstructed source power level according to eq. (2.1) is $L_{W,rec} = 128.37\,\text{dB}$, compared to $L_W = 132.07\,\text{dB}$ of the original, i.e. an error of ~2.8%.

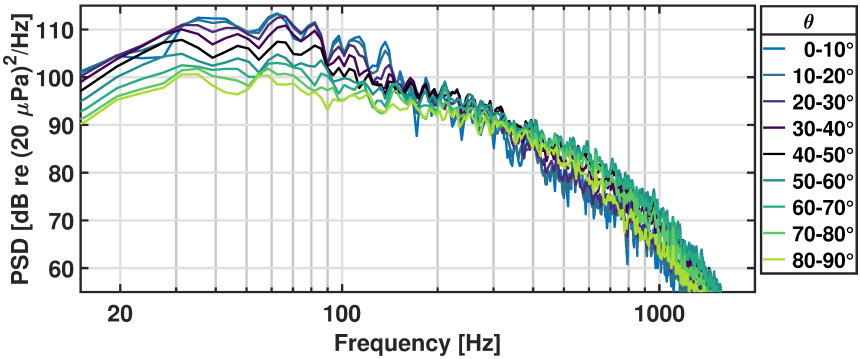


Figure 4.18: Downstream directivity of the reconstructed jet based on 9 selected (of 6561) elementary radiators.

4.3 Far-field jet noise radiation

As a proof of concept, this section describes the integration of the equivalent jet noise source model to a finite element method simulation (FEM) by using COMSOL Multiphysics® [28] to obtain the acoutic far field including wave reflection and diffraction effects at the fuselage and wing surfaces of a reference aircraft geometry. The aircraft geometry CSR-01 from the CeRAS database is chosen as a reference as shown in Figure (4.19, left) [143]. The aircraft geometry dimensions are (W x L x H) ≈ (34 x 37 x 7) m, with W being the wingspan, L being the fuselage length and H being the aircraft's height from the lower engine frame boundary to the top end of the vertical stabilizer.

Simulation setup

The cylindrical computational domain for the flow and acoustics simulation has an extent of 63 m in the streamwise direction and a diameter of 52.5 m in the crossflow direction, with the outermost 6 m being a perfectly matched layer. The source domain at the CSR-01 aircraft geometry is centered around the end of the potential core (cf. Fig. 4.9) as shown in Figure (4.19, right).
The simulation consists of three steps: Based on a first isolated laminar background flow simulation step and an intermediate background fluid flow coupling the acoustic field is finally simulated by solving the governing linearized Euler equations in time domain. The simulation includes a no slip boundary condition at the geometry surface. The domain meshes are automatically generated based

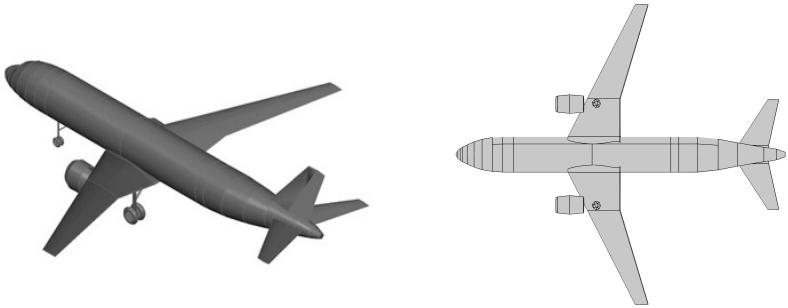


Figure 4.19: Left: CSR-01 aircraft geometry from CeRAS database [143]. Right: Source domain for the equivalent jet noise model.

on COMSOL's integrated physics-controlled meshing option. Motivated by the atmosphere's effect on sound waves to absorb high frequencies, the simulation is limited to an upper frequency of $f_{\max} = 500$ Hz. Depending on the actual atmospheric properties, this frequency limit is suitable to cover the relevant use cases for flight altitudes higher than 1000 m (cf. Fig. 8 in [35]). The resulting mesh for the flow simulation consists of 670000 volume elements and 23000 bundary elements, whereas the acoustics mesh consists of 4560000 volume elements and 248000 boundary elements. The simulation was performed using the high-performance computer (HPC) at the RWTH Aachen University data center, requiring 153.94 GB of physical memory and 5 hours, 52 minutes computation time, running in distributed mode using 1 nodes on 2 Intel Xeon Platinum 8160 CPUs at 2.10 GHz on 4 sockets with 8 cores.

Simulation results

Two simulation results are exemplarily depicted in the following. At first, Figure 4.20 shows the sound intensity resulting from a single-sided excitation with a sinusoidal signal with 400 Hz. The shielding effects of fuselage and wings as well as the no slip boundary condition are clearly visible in these plots. Moreover, the absorption efficiency of the perfectly matched layer can be observed at the lower frame since no relevant energy is reflected back into the simulation domain.

In comparison, the resulting sound intensity from a double-sided broadband excitation is shown in Figure 4.21.

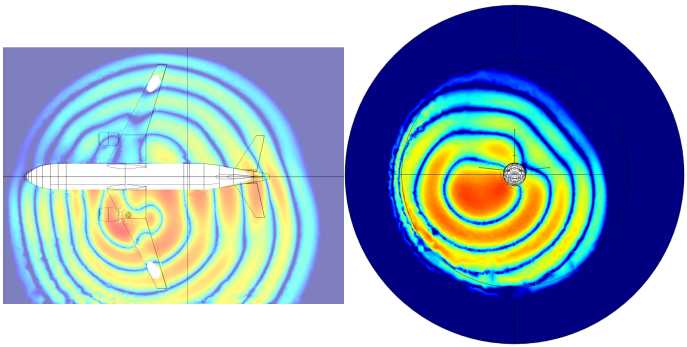


Figure 4.20: Cross-sectional sound intensity resulting at 400 Hz from the numerical far-field simulation.

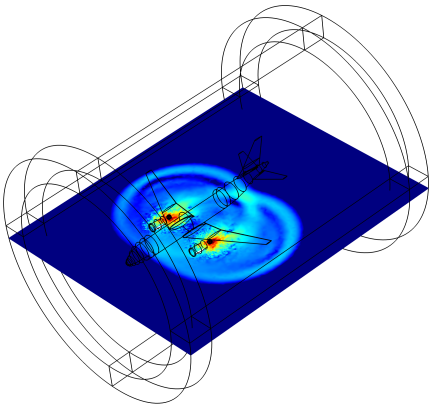


Figure 4.21: Sound intensity resulting from broadband excitation.

4.4 Aircraft sound synthesis

As discussed in the section 4.1.1, semi-empirical aircraft noise models characterize the spatio-spectral source properties. However, since temporal properties are not covered in these models, they have to be additionally synthesized for auralization.

Fan noise consists of multiple contributors, such as tones from fan and compressor and stochastic noise emanating from both (see Figure 4.3, right). The structures of a fan and a compressor are similar, both consisting of rotating blades (fan) in front of static blades (stator vanes). The Heidmann model [67] focuses on predicting a single fan or compressor stage and does not differentiate between both types. Therefore, every fan, high-pressure compressor, and low-pressure compressor stage has to be calculated individually.
According to the model, three components can be distinguished when looking at fan noise: Discrete tone and broadband noise. A third component is combination tone noise (due to its sound characteristics also called buzz-saw noise), which significantly occurs at high thrust settings with the rotor tip relative inlet Mach number being supersonic, which in turn leads to shock wave generation [63]. The combination tone frequencies are calculated from multiples $\mathrm{N_{rpm}}$ of the shaft rotational speed rpm by

$$f_{\mathrm{Fan,rpm,supersonic}} = \mathrm{N_{rpm}} \cdot \mathrm{rpm} \tag{4.8}$$

According amplitude values are assigned from the green curve in Figure 4.3 (right). In case of subsonic fan tip velocities, fan and compressor tones interaction can be observed for $\mathrm{N_{BPF}}$ fan blade passage frequency harmonics $\mathrm{BPF_{Fan}}$ and the high pressure compressor blade passage frequency $\mathrm{BPF_{HPC}}$:

$$f_{\mathrm{Fan,BPF,subsonic}} = \mathrm{N_{BPF}} \cdot \mathrm{BPF_{Fan}} \pm \mathrm{BPF_{HPC}} \tag{4.9}$$

where each $\mathrm{BPF_{Fan}}$ is the multiplied value of the number of fan blades of the emitting engine and the shaft rotational speed.

Due to its sinusoidal nature, additive synthesis (as introduced in section 2.6.1) is ideally suited for the sound synthesis of fan and compressor tones. However, as motivated in section 2.4.2, a purely sinusoidal synthesis of fan and compressor noise sounds implausible and sterile due to missing frequency and amplitude modulations that are audible in recordings of these sound sources. In the following, an algorithmic procedure to extract the according modulation parameters is presented.

4.4.1 Extraction of fan noise modulation parameters from flyover recordings

In order to obtain the fan noise modulation parameters from in-situ aircraft flyover recordings, they are compensated by three postprocessing steps (shown in Figure 4.22): After time-variant inverse filtering for the outdoor sound propagation effects (being introduced in section 2.5.1) and separation into tonal and stochastic signal components by non-negative matrix factorization (being introduced in section 2.5.2) the modulation parameters of individual fan tone harmonics are extracted by using Hilbert transform-based envelope extraction (being introduced in section 2.4.2). However, specific adaptions for aircraft flyovers are presented in the following.

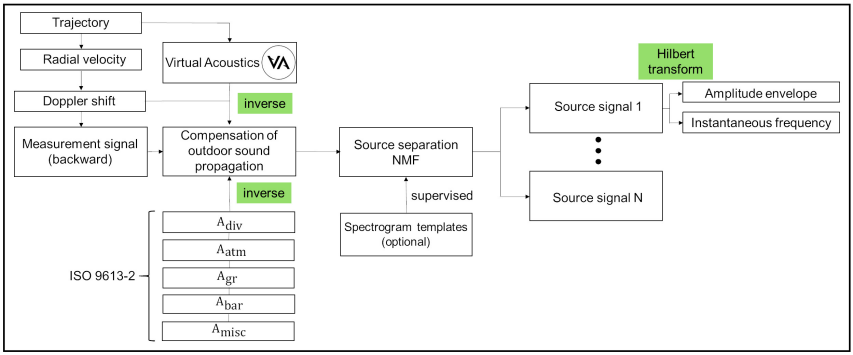


Figure 4.22: Block diagram for flyover postprocessing using time-variant wave backpropagation and source separation.

The in-situ flyover measurements for obtaining a reference input signal took place in a distance of about 4.5 km behind the point of departure and in line of the Düsseldorf Airport runway, Düsseldorf, Germany [1]. The measurement position is surrounded by meadow, from which no noticeable background noise distorted the acoustical measurements. As can be observed in the spectrogram (Figure 4.23), the recorded signal of an aircraft flyover contains time- and frequency-dependent effects from outdoor sound propagation, whose spectra are Doppler-shifted and attenuated by the atmosphere.

[1] Link to OpenStreetMap: Geographical location of the flyover measurement position

The Doppler effect can be observed in the spectrogram by the gradually decreasing frequency at about 2.2kHz (at t=0s) that originates from fan noise harmonics that itself operates at constant rpm.

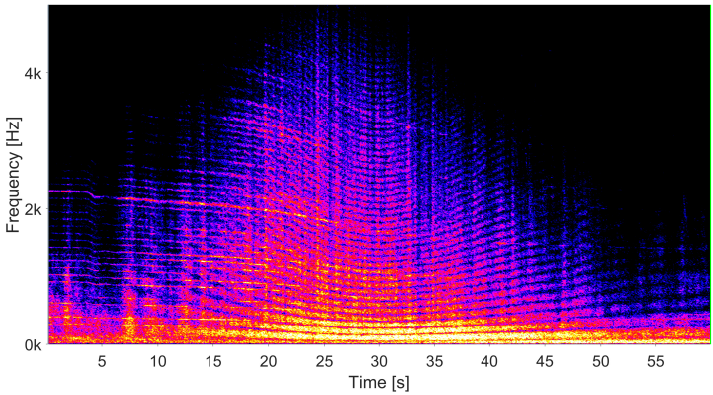


Figure 4.23: Spectrogram of in-situ measurement signal.

As basic requirement for removing the Doppler effect from the recording, the trajectory of the moving sound source must be tracked and calculated relatively to the microphone position. In the present case, the trajectory shown in Figure 4.24 was reconstructed using official flight tracking data supplied by Düsseldorf Airport using their track visualization tool TraVis [2]. The cut-back procedure, which is visible in the spectrogram as well as in the trajectory, was used for align the recording time stamp to the trajectory.

The actual de-Dopplerization is performed by using VA's binaural freefield renderer [74], which simulates a Doppler effect by an algorithm using variable delay lines (VDL) [145]. Therefore, a virtual source is assigned with the time-reversed flyover recording as input signal. The assigned source trajectory inversely follows the reconstructed trajectory, therefore inversely compensating the Doppler shift. As mentioned earlier, further wave propagation effects were compensated using time-variant wave backpropagation (as being introduced in section 2.5.1), considering the effects described in ISO 9613-2 [82]. To estimate the source-to-receiver sound propagation path, an atmospheric ray tracer was used [152], which was initialized with the local temperature and humidity parameters at the ground as

[2] Link to Düsseldorf Airport track visualization tool (Last viewed 12/16/2024)

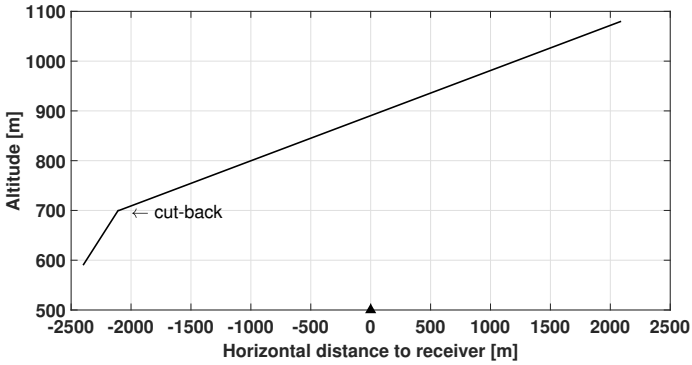


Figure 4.24: Tracked altitude of the aircraft according to the recorded flyover shown in Figure 4.23. Flight direction is from left to right. The indicated cut-back operation leads to an abrupt frequency shift. The measurement position at ground is indicated by the black triangle.

well as wind direction and speed. Altitude-dependent predictions of temperature and pressure variables until the flyover altitude assume International Standard Atmosphere (ISA) conditions [76], whereas wind properties and humidity assume constant conditions over altitude. The compensation filters are updated every 100ms and applied to the de-Dopplerized signal. As visible in Figure 4.25, the Doppler- and atmosphere compensated spectrogram reflects all steady-state harmonic sources as horizontal lines.

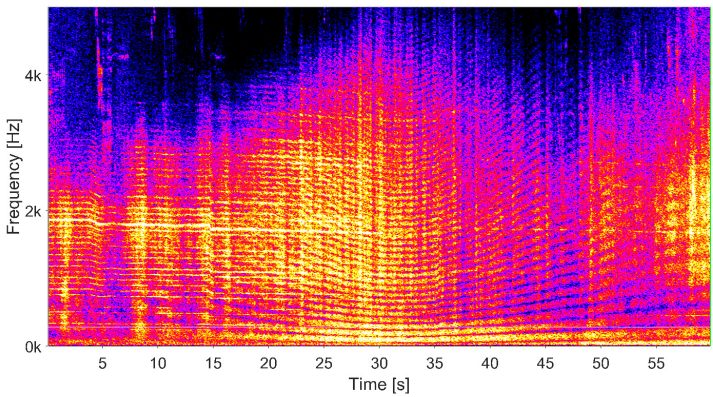


Figure 4.25: Spectrogram of de-Dopplerized and absorption-compensated flyover signal.

After backpropagating the flyover signal, the overall signal is separated into its stochastic and tonal signal components by informed non-negative matrix factorization (NMF).

As introduced in section 2.5.2, informed NMF depends on templates representing spectral characteristics of the components to be separated. For fan noise, a-priori knowledge was based on the datasheet of the recorded IAE V2500 engine. As presented in equation (4.9), each BPF_{Fan} is the multiplied value of the number of fan blades of the emitting engine and the shaft rotational speed rpm, as can be observed at 1791 Hz in Figure 4.26. No higher BPF harmonics frequencies can be observed in the recorded data since they are absorbed by the atmosphere.

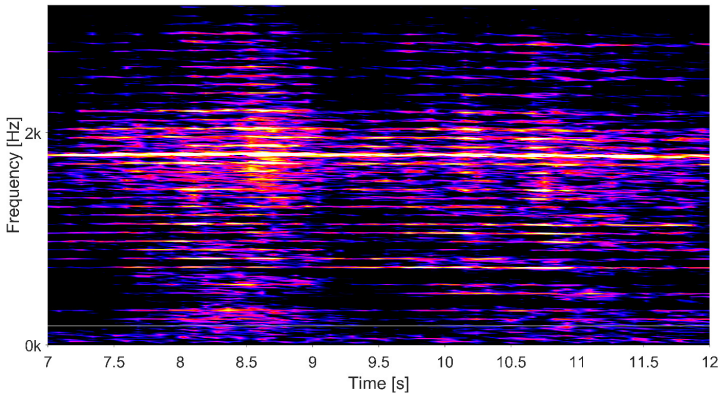


Figure 4.26: Close-up of compensated flyover spectrogram with $\Delta f_{FFT} \approx 5.8$Hz frequency resolution and amplitude range of 20 dB. The BPF at 1791 Hz is related to the shaft rotational speed of 81.4 Hz multiplied by 22 fan blades. Other harmonics are multiples of the shaft speed.

In the presented case a number of 22 fan blades was multiplied for calculation of the harmonics frequencies with the rotational speed at a stationary-working point. This rotational speed was found by division of the peak BPF at 1791 Hz in the fan emission spectrogram (Fig. 4.26) by the number of fan blades.

From the extracted fan component, the amplitude and frequency modulation parameters are extracted by using the method introduced in section 2.4.2. Shown at the example of the BPF in Figure 4.27, it can be observed that the instantaneous amplitude variation exhibit the range between 0.01 and 0.2 Pascal, whereas in the instantaneous frequency variations between 1790 and 1805 Hz can be observed. The information about amplitude and frequency fluctuations can be coded as

variance values for application to basic oscillators during sound synthesis, as further described in the following chapter 5.3.2 for drone sound emission.

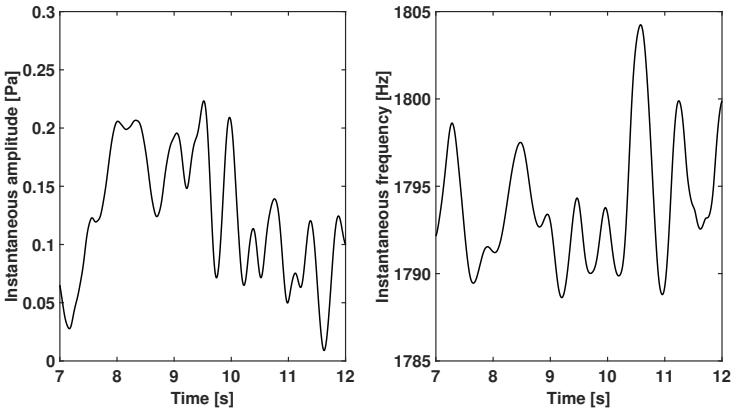


Figure 4.27: Example of instantaneous amplitude (left) and frequency (right) of the BPF (22nd harmonic according to Fig. 4.26) that are obtained from the analytic signal $\tilde{s}_{22}(t)$.

Compensation filter design issues

The practical implementation of atmospheric absorption compensation as described above has aspects that should be discussed in the foreground of electronic hardware limitations and its settings. Due to the lowpass characteristics of atmospheric attenuation A_{atm}, high frequencies are damped before reaching the measurement microphone. For the example of a typical flyover in a distance of 1000 m, the sound pressure level in the 4 kHz frequency range is reduced by about 30-50 dB compared to very low frequencies at 20 Hz [35]. This in turn leads due to the limited signal-to-noise ratio (SNR) of the recording equipment to a reduced SNR of the flyover recording at high frequencies. However, during the spectral compensation process it is necessary to apply high gains especially at high frequencies, alas, ending up in increased harshness. Therefore, the proposed procedure benefits if maximum bit depth setting of the measurement equipment is used.

4.5 Summary and discussion

Semi-empirical models for aircraft sound sources (such as fan, jet and airframe noise) are commonly implemented in modern auralization frameworks. Their objective is to characterize the spatio-spectral emission in one-third octave bands and their dependency on specific operating states. However, these semi-empirical models have two major limitations for auralizations:

Firstly, referring to the source characterization dimensions being introduced in section 2.4, the temporal characteristics are not covered in these models. Since semi-empirical models predict emission amplitudes at discrete frequencies, these information is suitable for additive and subtractive synthesis techniques (see section 2.6), which indeed is sufficient for stochastic components (such as jet and airframe noise). However, without further information, fan noise auralization sounds sterile and implausible, so that the temporal dimension has to be additionally modeled for their auralization, so that section 4.4 proposed a synthesis enhancement by extraction of fan noise modulation parameters based on flyover measurements.

The second limitation of semi-empirical models is their directivity pattern description of acoustical near field patterns by neglecting the shielding effect by fuselage and wings on acoustical far field patterns. In case of jet noise, this has been investigated in section 4.3 by a proof of concept based on high-fidelity flow simulations with complexity reduction using spherical harmonics. The presented study sheds new light on the results presented by Lee and Bolton [101]. First of all, their findings of a jet noise expansion to be dominated by dipole and quadrupole were confirmed independently. By choosing a simulation-based approach new insights beyond measurement-based restrictions were revealed by detecting other strong contributions from octupoles, the monopole and also hexadecapoles. The ability for detecting these additional radiators, that are not negligible in terms of their individual contributions to the reconstructed directivity and the overall sound power, are most likely due to the possibility of sensor placement within the flow. This raises the question of their physical origins. Since it appears obvious from other studies that the quadrupole-like radiation is related to mixing noise [105], the identification and matching of other noise generation mechanisms with according SH orders may be subject of future research. Therefore, the presented work could be extended to more nozzles and flow conditions. Compared to previous studies on jet noise source characterization, the chosen approach also proved to be helpful due to a much higher signal-to-noise ratio and according precision of the data processing at the numerical resolution limit. This, in turn, allowed

for high spatial resolution while keeping spatial and spectral aliasing errors at a negligibly low level. From the spherical harmonic coefficients were calculated according numerical source strengths for each elementary radiator. The results indicate that jet noise cannot be fully resolved in terms of its acoustical sources by a microphone array limited to a low number of microphones. Hence it cannot be fully reconstructed by a low number of elementary radiator orders. Rather, the accepted reproduction error is crucial. In this study by dominance analysis of the original 6561 sound sources, a convenient equivalent source description based on 9 elementary radiators was derived, which approximates the original acoustic field well over the entire spectrum, but accepts errors to the radiated elevation angle between $40 < \theta < 60°$. These coefficients provide a convenient format for an equivalent sound source description of jet noise for further use in numerical flow simulations. Since the spherical harmonic transform assumes a common source point, the question is raised to what extent the addition of spatial degrees of freedom can enable further optimization. A suitable approach could be based on the sound field decomposition using multipole or monopole decomposition. The reduced-order jet noise model can be used in broader simulation tasks. As demonstrated for jet noise by the proof of concept in section 4.3, future air-traffic auralization frameworks should integrate acoustical far-field directivity patterns of aircraft sound sources by additionally considering wave reflections at fuselage and wing surfaces and advection during sound propagation through the ambient medium.

5

Drone auralization model

Parts of this chapter have been published in: Dreier, C., Vorländer, M. (2024). Drone auralization model with statistical synthesis of amplitude and frequency modulations. Acta Acustica, ***35****, 1–13* [38].

5.1 Introduction

Model-based syntheses are especially challenging to produce when perceptual plausibility is a quality criterion. For perceptually plausible auralizations, the consideration of fluctuations in the sound field is crucial, regardless of whether they are generated at the source itself or during sound propagation. In the latter case, this has been shown, for example, in the case of aircraft noise [107] or wind turbine noise [24]. In the case of drones, plausible synthesis is especially critical with regard to time-variant amplitude and frequency modulations of the tonal component. These emanate, on the one hand, from the speed control of the electric motors and, on the other hand, from advection and refraction effects due to aerodynamic turbulence in the rotor surroundings. This work presents a drone auralization model accounting for the statistical amplitude and frequency modulation distributions of the tonal component. For model completeness, the drone directivity is computed by using a time-variant wave backpropagation method that is compensating for wave propagation effects as well as non-ideal, directional properties of the measurement microphones.

5.1.1 State of the art

General drone noise emission characteristics have been widely investigated in a large number of publications, cf. [153], [187]. In most of the publications, micro-

phone array signals are processed to compute one-third octave band spectra and the emitted sound power. An important finding is that the directivity pattern of drones can be assumed as independent of the rotational speed of the rotors and of the flight procedure [184]. By using a comprehensive circular microphone array, Herold [69] measured the in-flight sound power and directivity of a drone in an anechoic chamber and compared his results to analytical monopoles and dipoles. Furthermore, he reconstructed the flight path using a multi-dimensional beamforming analysis. Alkmim *et al.* [7] presented a spherical harmonics decomposition of drone noise emission based on static hemi-spherical array data in order to predict far-field sound pressure levels. In both studies, no compensation of sound propagation effects due to atmospheric absorption or of the microphone directivity were reported. Heutschi *et al.* [70] presented a drone auralization model with focus on a parameterization of the rotational speed dependency and derived a generic directivity. For auralization, these model parameters can be used in combination with an anechoic source signal recording from a hovering maneuver. Focusing on psychoacoustic noise assessments in the field of advanced air mobility (AAM), Lotinga *et al.* [113] found in a review paper sharpness and tonality to more accurately describe drone sound emission compared to level-based noise metrics regarding its comparability in terms of annoyance ratings.

5.1.2 Contribution of this work

Beyond previous publications, this work presents a drone auralization model focusing on perceptual plausibility by statistically modeling the amplitude and frequency modulation distributions in the drone's tonal sound emission. Both elements are calibrated by using a power matching approach so that the spectro-temporal source characteristics of the reference drone is fully replicated and optionally level-adapted to measurement- or simulation-based data of other drones (Fig. 2.3). A common approach of source signal synthesis for auralizations is to discretize the overall signal mixture into a procedural audio synthesis of tonal components and subtractive synthesis of white noise for representing stochastic components, cf. [132]. The basic idea for statistically modeling the modulations in the drone's tonal sound emission is sketched in Figure 5.1. A simple solution for auralizing the tonal component would be based on amplitude estimates A_N at N discrete frequencies from a spectrum of a drone recording. With the drawback of sounding unnatural and artificial, the synthesis would reproduce the amplitudes of purely mono-frequent sinusoidal oscillators by using additive synthesis (vertical black lines, see section 2.6.1). For enhanced plausibility, the presented work focuses on the statistical modeling of the instantaneous amplitudes and frequency

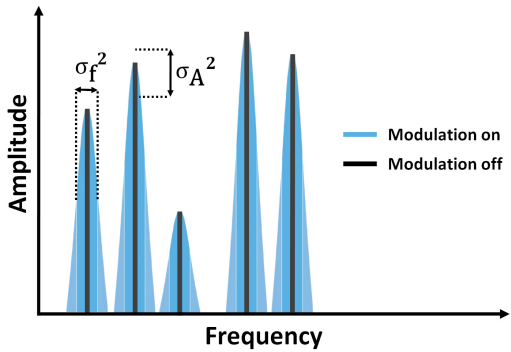


Figure 5.1: Schematic spectral effect in the tonal component synthesis with (blue) and without (black) modulations.

variations (blue Gaussian distributions, modulation variance σ^2) to include the characteristic "bee" sound of a drone in auralization. In conclusion, the core of this work comprises a method for estimating amplitude and frequency variations of drone noise tonal components from measurement in a controlled environment, a method for estimating drone noise emission directivity by time-variant wave backpropagation, and a validation of the procedural audio synthesis method including tonal and broadband components by comparing psychoacoustic metrics obtained on the recorded and auralized samples for the hovering drone.

5.2 Source characterization of drone sound emission

The acoustic source characterization of a reference drone (DJI Mavic Pro, shown in Figure 5.2) is based on measurements in the hemi-anechoic chamber of the Institute for Hearing Technology and Acoustics.

5.2.1 Modeling concept

The general goal of source characterization as basis for use in auralization is the description of a sound source operating under free-field conditions, i.e. excluding any influences from room reflections or non-absorbing boundaries. Note, although a drone radiates considerable amount of sound energy in the ultrasonic frequency range (Fig. 5.4), these frequencies are not considered in this auralization model since they are inaudible. The auralization model aims to reproduce the spectro-

Figure 5.2: Reference drone type DJI Mavic Pro.

temporal and spatial characteristics of recorded drone sound emission data. In this work, the model structure consists of two blocks, an emission signal in form of a synthesized audio stream and a directivity in form of angle-dependent magnitude spectra (Fig. 2.2). As described in more detail in section 2.2, both components are calibrated with regard to a correct sound pressure reconstruction on a virtual spherical surface with a radius of 1 m around the drone's center. For example, this model structure is compatible with the open-source auralization framework Virtual Acoustics (VA) [74] with additional directivity coding in the OpenDAFF format [129]. For emission signal modeling (cf. section 3), at first the tonal and stochastic contributions to the overall sound are obtained from recorded data as two separate components of a steady-state hovering maneuver. The stochastic contribution was modeled by using an auto-regressive approach to extract coefficients for a digital infinite impulse response (IIR) filter. For modeling the tonal contribution, after individual extraction of emitted tones, their statistical variation regarding frequency and amplitude modulations were analyzed. For directivity modeling, the radiated emission to the lower hemisphere in an angular range $0^\circ < \phi < 180^\circ$ and $20^\circ < \theta < 90^\circ$ is reconstructed on a virtual spherical surface with a radius of 1 m around the drone's center. In the case of drones with contra-rotating propellers, a symmetry of sound emission along the vertical median plane can be assumed, thus, reducing the number of microphones in the array (cf. Fig. 5.3). The reconstruction is computed from flyover array measurements using time-variant wave backpropagation, as generally introduced in section 2.5.1.

5.2.2 Measurement setup

The frequency range of the measurement is valid down to the hemi-anechoic room's cut-off frequency at about $f_{\text{cut-off}} \approx 125$ Hz. The sound emission was recorded in two settings, during a hovering maneuver and a flyover. The first measurement is the basis for the extraction of the drone's sound emission signal (Sec. 5.3) whereas the second measurement is post-processed to extract the emission directivity (Sec. 5.4). For the fly-over measurement shown in Figure 5.3, the drone repeatedly passes a non-uniformly spaced linear microphone array with a constant speed of $v = 0.75\,\frac{\text{m}}{\text{s}}$ in a fly-over altitude of $h = 2$ m. The drone speed and altitude were controlled by using its optical sensor-assisted automation [1]. The speed was externally validated by placement of two timer microphones on the ground at both ends of a trajectory with length $L = 10$ m. The microphone array is orthogonally placed to one side of the flight direction and symmetrically centered half-way on the trajectory. In the case of drones with contra-rotating propellers, the measurement data can be mirrored at the vertical median plane due to geometric symmetry of the drone's fuselage. The microphones are non-uniformly spaced on the floor in geometrical continuation to their position with an angular spacing of $\Theta = 10°$ on a hemi-spherical surface with radius $r = 1$ m around the drone's center. The emission angles for elevation Θ and azimuth ϕ, as shown in Figure 5.3 are calculated by

$$\Theta = \arctan\left(\frac{h}{\sqrt{(\frac{L}{2} - x)^2 + d^2}}\right) \tag{5.1}$$

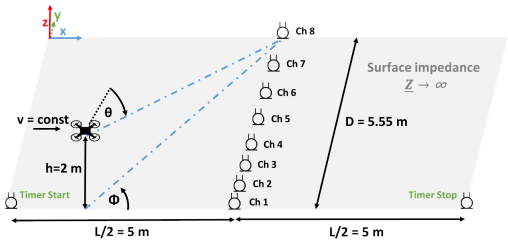
$$\phi = \arctan\left(\frac{d}{\frac{L}{2} - x}\right) \tag{5.2}$$

with x being the distance travelled and d being the lateral microphone distance.

5.2.3 Measurement equipment

The measurement was performed by using a set of eight calibrated measurement microphones placed on the rigid floor (Fig. 5.3). On flyover median plane (channel 1), an NTI Audio M2230 (Class 1 acc. to IEC 61672 and ANSI S1.4) was installed, whereas omnidirectional Sennheiser KE4 microphones were installed flat on the ground on the off-axis positions (channels 2-8). All microphones were connected

[1] A video example of a fly-over measurement is provided in the supplemental files of the journal publication at https://doi.org/10.1051/aacus/2024026 (Last viewed 12/16/2024)



Ch [#]	d [m]	$\theta_{\phi=90}$ [°]
1	0	90
2	0.35	80
3	0.72	70
4	1.15	60
5	1.68	50
6	2.38	40
7	3.46	30
8	5.55	20

Figure 5.3: Schematic of the drone measurement setup in the hemi-anechoic chamber. The table indicates the lateral microphone distances to the fly-over trajectory and the according emission elevation angle θ.

to an RME Octamic XTC audio interface, set to $f_s = 96$ kHz sampling rate and $N_{\mathrm{ADC}} = 24$ bit depth. The microphones were equipped with windshields to avoid aeroacoustic distortion on the capsules. The atmospheric conditions were 20°C temperature and 40% humidity.

5.3 Analysis and synthesis of drone sound emission

5.3.1 Broadband spectrum

The drone's sound emission is recorded at a 2 m distance during a hovering manoeuvre in the hemi-anechoic chamber. The sound pressure level at a 1 m radial distance from the source can be directly calculated from the calibrated audio recording due to the doubled sound pressure on the rigid floor (compared to free-field conditions). In the broadband spectrum, shown in Figure 5.4, tonal and broadband noise components can be distinguished by local level peaks and a shaped noise floor. Three different characteristic spectral ranges can be distinguished in the broadband spectrum (Fig. 5.4):

1. The frequency range f < 2.5 kHz is dominated by tonal components.
2. The range f > 2.5 kHz contains amplitude-modulated, shaped noise, with a non-perceptual tonal component.
3. A broad peak emission can be observed in the ultrasonic frequency range 35 kHz < f < 45 kHz.

Due to their clearly distinguishable characteristics, the first two components are separately synthesized in the presented drone auralization model. The presented

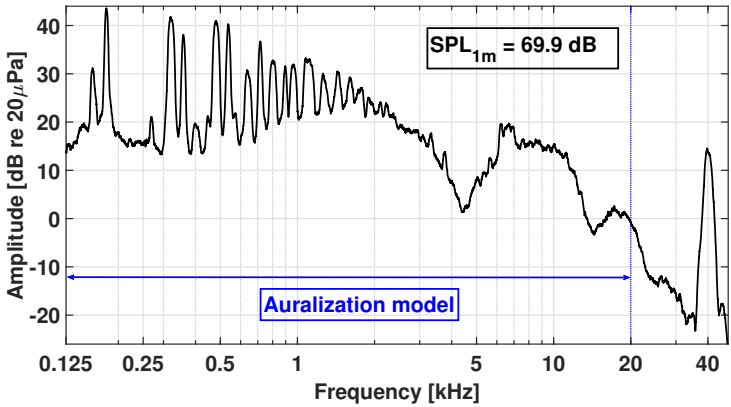


Figure 5.4: Reference drone broadband emission spectrum (125 Hz < f < 48 kHz) and sound pressure level (SPL), recorded during hovering manoeuvre at a 1m distance at emission angle $\theta = 45°$ and $\phi = 0°$ under free-field conditions. A 1/24-octave band smoothing is applied for a better visual clarity.

drone auralization model assumes the broadband noise to be inaudible in the low frequency range f < 2.5 kHz and the tonal components to be inaudible in the range f > 2.5 kHz due to spectral masking considerations [45]. Note, since the ultrasonic component is inaudible for human hearing it is not considered in the presented model.

5.3.2 Analysis of tonal amplitude and frequency modulations

Tone extraction

The broadband spectrum (Fig. 5.4) is a two-dimensional representation of the sources spectral characteristics, omitting details on temporal variations (cf. section 2.4). At first, in order to analyze the amplitude modulations (AM) and frequency modulations (FM) individually by using the analytic signal, as introduced in section 2.4.2, each tone needs to be extracted from the broadband spectrum using steep bandpass filters, symmetrically centered around the observed peak frequency. The recorded data of the hovering maneuver is sampled at f_s =96 kHz, which is necessary to realize a filter bank of steep bandpass filters. As a rule of thumb, a filter bandwidth of $B_n = \pm 0.1 f_{c,n}$ around the center frequency f_c of tone n (with n $\in [1, 22]$) was found necessary for a precise tone extraction. The passband ripple was chosen to 0.1 dB with a stopband attenuation of 60 dB.

The filters are realized as linear-phase finite impulse response (FIR) filters based on the Kaiser window method. The resulting spectra of the first 22 individual tones are plotted with colored lines in Figure 5.5 and compared to the broadband spectrum.

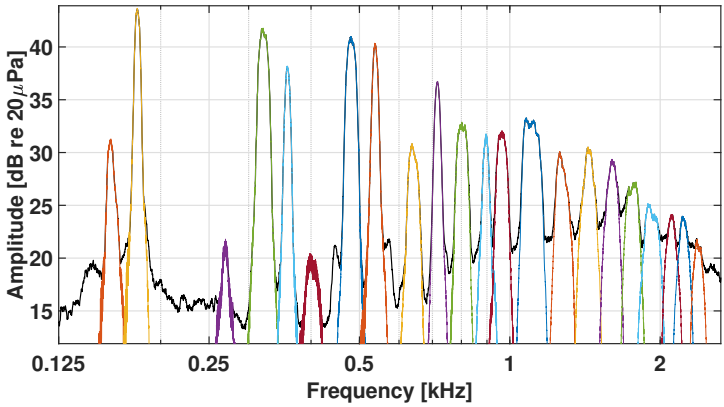


Figure 5.5: Spectra of the first 22 extracted tones (colored), compared to the originally recorded broadband spectrum. A 1/24-octave band smoothing is applied for a better visual clarity.

Instantaneous amplitude and frequency computation

For plausible drone sound synthesis, each of the drone's tonal oscillators can be described by its instantaneous amplitude and frequency. The time-variant information of both parameters can be computed from the signal processing procedure described in section 2.4.2. As an example, it can be observed in Figure 5.6 that the instantaneous amplitude variation of the fifth extracted tone exhibits the pressure range $\approx 0.005\,\mathrm{Pa} < p_5 < 0.033\,\mathrm{Pa}$, whereas instantaneous frequency variation is in the range $352\,\mathrm{Hz} < f_5 < 364\,\mathrm{Hz}$.

Assuming normal distribution, each individual tone's statistical information about amplitude and frequency modulation rate and depth is stored in form of variance values σ^2 [2]. These values are directly applicable to the oscillator properties for procedural sound synthesis. The time-variant plot of the extracted harmonics (Fig.

[2] All numerical values are provided in the supplemental files of the journal publication at https://doi.org/10.1051/aacus/2024026 (Last viewed 12/16/2024)

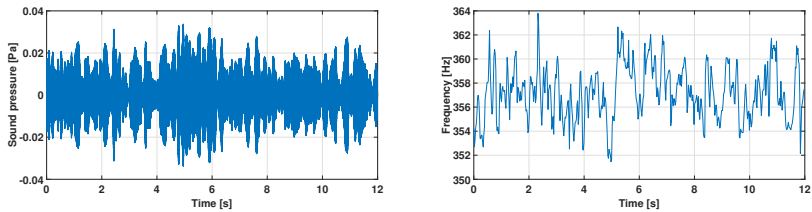


Figure 5.6: Sound pressure (left) and instantaneous frequency (right) of the fifth extracted tone with mean frequency $f_{5,\mathrm{mean}} = 358.75$ Hz.

5.7) shows the frequency modulations (FM) as well as amplitude modulations (AM). For excluding distortion in the statistics due to any possible drifting-based variation of the geometrical spreading loss between the measurement microphone and the noise-emitting oscillators, the zero-mean modulations are computed first.

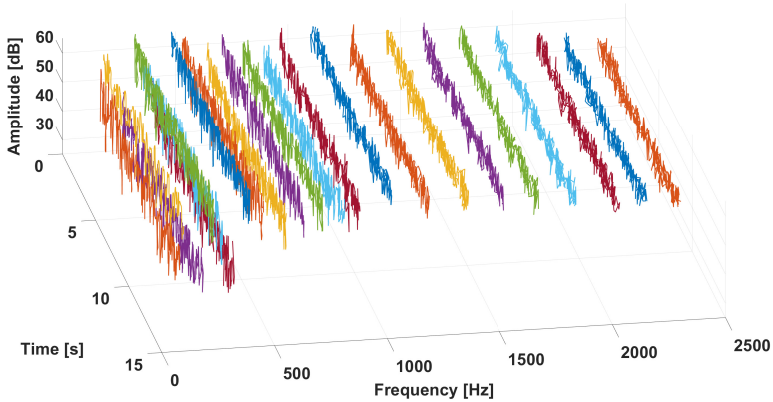


Figure 5.7: Instantaneous tonal amplitude and frequency fluctuation in the far-field sound measurement. Coloration of each tone according to Figure 5.5.

Since the sound is measured after travelling through the turbulent layer, each oscillator frequency is individually modulated and therefore varies independently of time. The statistical mean and variance values of the instantaneous amplitude and instantaneous frequency are shown in Figure 5.8 and according numerical values are provided as a separate table [2].

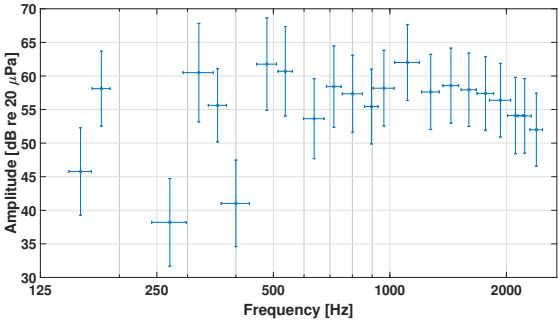


Figure 5.8: Instantaneous amplitude and frequency variance in the tonal component. The frequency variance is exaggerated in the plot by a factor of 10 for a better visual clarity.

5.3.3 Synthesis of stochastic drone sound component

The spectral characteristics of the noise component are modeled in this work using the same AR modeling process as proposed for tire/road noise synthesis (see section 3.7 for algorithmic details). The required number of SOS filter coefficients is assessed by comparison of the drone's emission spectrum with the synthesized noise spectrum (Fig. 5.9). At order 40, a sufficient approximation – that is indistinguishable in its audibility – of the original spectrum is achieved [3].

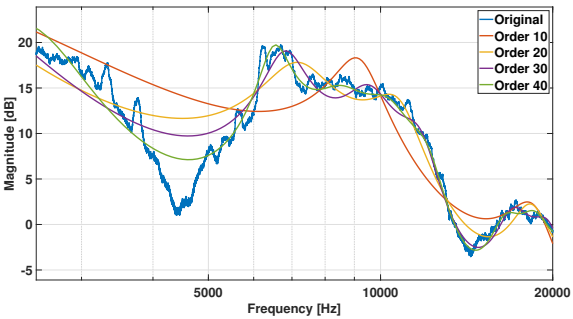


Figure 5.9: Comparison of the emitted high-frequency noise and the cascaded SOS filter responses of different order in the frequency range 2.5 kHz $< f <$ 20 kHz.

[3] The numerical discrete-time filter coefficients b and a for the IIR filter with cascaded second-order sections form are provided in the supplemental files of the journal publication at https://doi.org/10.1051/aacus/2024026 (Last viewed 12/16/2024)

5.4 Directivity model

The directivity is computed from flyover data by using the time-variant wave backpropagation method, as introduced in section 2.5.1. All receivers were individually calibrted according to the procedure described in section 3.2.2. Since most trajectories of drone auralizations refer to flight altitudes at ear level or above head height, the directivity is measured only in the lower hemisphere. A homogeneous atmosphere is assumed so that the path length of the direct wave can be determined along straight lines. Furthermore, the method assumes a rigid floor.

Measurement uncertainty from positioning inaccuracies

As stated earlier, the flight speed and altitude were controlled by using the drone's optical sensor-assisted automation. It is assumed that it can still lead to deviating positions due to a non-ideal control system. The averages of $n = 10$ measurement repetitions are evaluated. The averaging result for the recorded spectrum at microphone channel 1 is shown in Figure 5.10. It indicates the standard deviation to be within ± 2 dB for the lower frequency range $125\,\text{Hz} < f < 250\,\text{Hz}$ and within ± 1 dB for the frequency range above.

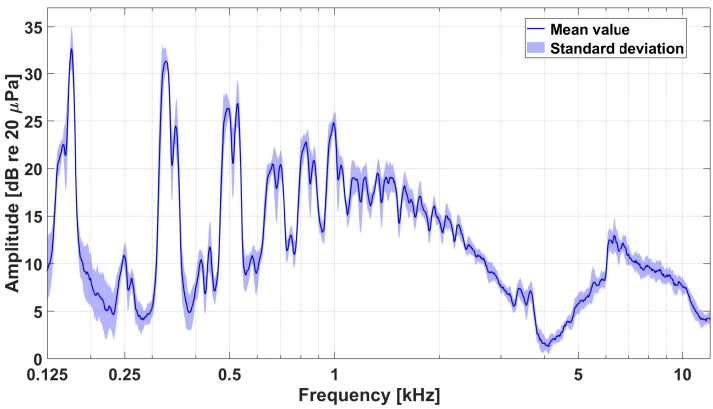


Figure 5.10: Spectral mean and standard deviation at microphone channel 1 for $n = 10$ flyover measurement repetitions.

5.4.1 Radiation in the median and frontal plane

The median plane (x-z plane in Fig. 5.3) directivity spectra (Fig. 5.11) show the tonal component to be more prominent for smaller angles, near to the x-y plane. As observed in other publications (e.g. [70], [69]), a trend towards a vertical dipole radiation characteristics can be observed in a broad frequency range. However, this dipole behavior cannot be generalized to every frequency and may cause an audible difference in auralization. Moreover, the dipole shows a rather broad lobe to the lower hemisphere. An audible frequency shift of the spectral notch in the frequency range 2.5 kHz$< f <$ 5 kHz can be observed towards lower frequencies for larger elevation angles. This notch in turn leads to a quadrupole directivity. The frontal plane (y-z plane in Fig. 5.3) directivity (Fig. 5.12) shows the resulting

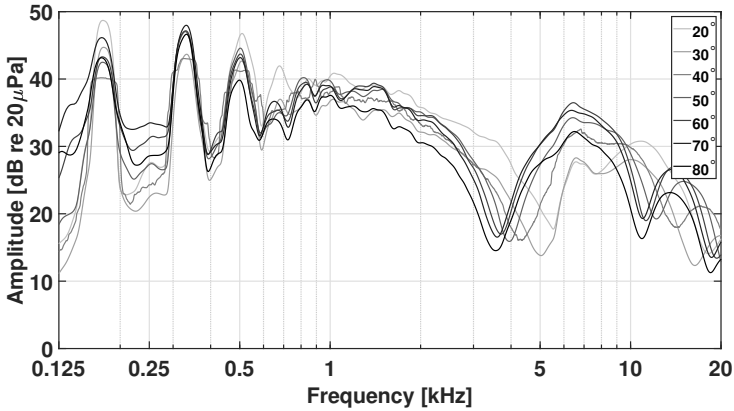


Figure 5.11: Median plane directivity spectra (x-z plane) for azimuthal angle $\phi = 0°$ and elevation range $15° < \theta < 85°$ (sliced into $\pm 5°$ segments around the indicated middle value).

spectra from each of the microphones. In this data, the spectral notch is only visible in the median plane emission angle $\phi = 0°$. The spectral notch is likely originating from a fixed phase relation between the front and rear axis propellers, since no ground or Doppler effect are apparent in the post-processed data. Again, a broad lobe of the vertical dipole is visible.

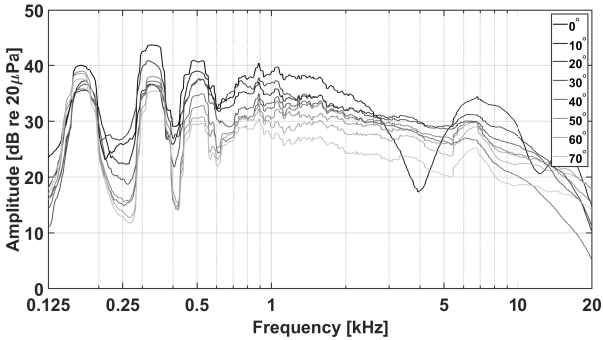


Figure 5.12: Frontal plane directivity spectra (y-z plane) in the elevation range $20° < \theta < 90°$.

5.5 Synthesis results

The spectrogram in Figure 5.13 visually compares the tonal component syntheses of the drone auralization model without modulations (in the time range 0 to 6 seconds) and with modulations (in the time range 6 to 12 seconds). Each oscillator's amplitudes are matched in both cases. In the following, all spectrogram are computed with an FFT block length of 8192 samples, 50 % overlap, Hanning window, and 96 kHz sampling frequency [4].

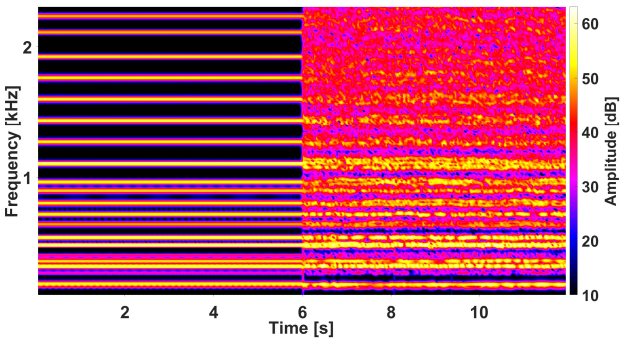


Figure 5.13: Synthesis result: Spectrogram of tonal component synthesis without (left) and with (right) modulations.

[4] For audible comparison, the auralization result is provided in the supplemental materials of the journal publication at https://doi.org/10.1051/aacus/2024026 (Last viewed 12/16/2024)

Focusing on the omni-directional emission signal, i.e. without the directivity pattern, the spectrograms of the reference recording and the synthesized emission from the presented drone auralization model are compared in Figure 5.14. The perceptual similarity is objectively evaluated by means of the psychoacoustic measures sharpness and tonality in table 5.1. Their computation is based on the reference recording and the synthesis signals for the hovering maneuver. The algorithms are considering the signal's spectro-temporal properties and their effects under consideration of a non-linear hearing model. In this sense, loudness can be understood as the aurally correct counterpart to level.

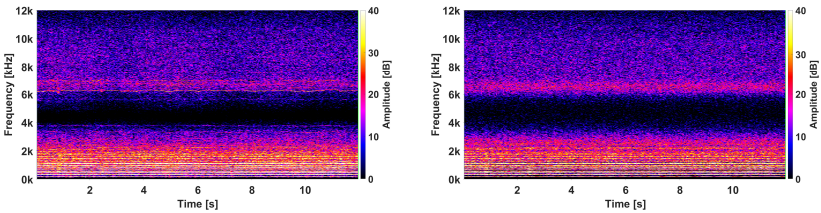


Figure 5.14: Spectrogram of the reference recording (left) and of the synthesized emission (right) in the frequency range $f < 12$ kHz.

Since a drone fly-over is a rather slowly changing event regarding its spectral content, the 5% percentile loudness values (N5) in this study were calculated according to the ANSI S3.4-2007 standard. This algorithm is based on a model by Glasberg and Moore [123] aiming at the loudness determination of stationary signals. Unlike the level description in decibels, loudness values on the sone-scale can intuitively be compared so that the ratio of two sone-values correspond to their subjectively perceived ratio. Since the tonal component is a main contribution to drone noise, the psychoacoustic parameter sharpness is used for evaluation. It is closely inverse related to sensory pleasantness and reflects the ratio of high frequency and low frequency contributions in a sound. In order to account for the relationship between the sensation of sharpness and loudness for technical sounds, this study is based on the free-field calculation method according to Aures [10]. The tonality is separately calculated according to ECMA-418-2 [40].
The auralization model synthesis shows its largest percentual deviation regarding its tonality reproduction, being about 1% more tonal. Slightly overstimating the reference, the model produces a 0.25% stronger loudness. Showing a very similar sharpness, the synthesis is about 0.5% sharper. Finally, these differences are not audible since they are below the just noticeable differences for loudness (7%), sharpness (≈ 0.05 acum), and tonality (≈ 0.05 $\mathrm{tu_{HMS}}$).

	Reference recording	Auralization model synthesis	Δ [%]
Loudness [sone,N5]	39.1	39.2	+0.25
Sharpness [acum]	3.85	3.87	+0.52
Tonality [tu_{HMS}]	0.300	0.306	+1.02

Table 5.1: Comparison of drone reference recording and the auralization model synthesis by means of calculated psychoacoustic measures.

5.6 Summary and discussion

This chapter expands the view of previous publications on drone auralization by focusing beyond the analysis of sound emissions to present a complete synthesis model including the oscillator and digital filter coefficients for procedural audio synthesis. The auralization model aims to reproduce the spectro-temporal and spatial characteristics of recorded drone sound emission data during flight in cruise mode with constant altitude. The general model structure consists of two blocks, an emission signal in form of a synthesized audio stream and a directivity as a set of angle-dependent magnitude spectra. Both components consider a combined level calibration with regard to the reconstructed sound pressure of the surface data. For directivity modeling, the radiated emission to the lower hemisphere in an angular range $0° < \phi < 180°$ and $20° < \theta < 90°$ is reconstructed on a virtual spherical surface with a radius of 1 m around the drone's center. The reconstruction is computed from flyover array measurements using a time-variant wave backpropagation. In the case of drones with contra-rotating propellers, a symmetry of sound emission along the vertical median plane can be assumed, thus, reducing the number of microphones in the array (cf. Fig. 5.3).
The synthesis model depends on an initial measurement phase. For this measurement phase, the conclusion can be drawn that a free-field recording of a drone flyover is required with the following conditions: For ensuring useful instantaneous amplitude and frequency computation from the analytic signal (cf. 5.3.2), the recording should be free from audible rotor speed variation caused by interruptions from the automatic positioning control. Furthermore, use of the equipment's maximum sampling rate and bit depth is recommended for an improved frequency domain resolution during postprocessing. In general, the measurement conditions play an important role when precise modeling of the tonal composition is required. Our measurements show for drone measurements

in a hemi-anechoic chamber that signal-to-noise ratios of $\mathrm{SNR_{HAC}} > 60$ dB can be achieved. In cases where the tonal details of the source are not so important or a large anechoic chamber is not available, it can be concluded from experience – without presenting details in this chapter – that measurements can also be made in the field, with signal-to-noise ratios between $20\,\mathrm{dB} < \mathrm{SNR_{outdoor}} < 30$ dB. In the second case, it is possible to determine the directivity with good accuracy. For tonal component feature extraction, however, care must be taken to ensure that prominent ambient or natural noises are not recorded.
Two effects can be observed from the directivity. First, the radiation of tonal components is more prominent for smaller angles towards the x-y plane. Second, a trend towards a vertical dipole radiation with a broad lobe can be confirmed independently, as observed in other publications (e.g. [69], [70]). However, compared to previous publications this dipole behavior cannot be generalized to every frequency, which may cause an audible difference in auralization. An audible frequency shift of a spectral notch in the frequency range 2.5 kHz$< f <$ 5 kHz can be observed towards lower frequencies for larger elevation angles and only in the x-z symmetry plane. This notch in turn leads to a quadrupole directivity. Since no ground or Doppler effect are apparent in the post-processed data, it can be concluded that the spectral notch is originating from a fixed phase relation between the emissions from the front and rear axis rotors.
For emission signal modeling, at first the tonal and stochastic contributions to the overall sound were separated from recorded data of a steady-state hovering maneuver. The stochastic contribution was modeled by using an auto-regressive approach to extract coefficients for a digital IIR filter chain. For modeling the tonal contribution, after individual extraction of emitted tones, their statistical variation regarding frequency and amplitude modulations were analyzed. The comparison of synthesized data with recordings by means of psychoacoustic measures shows good agreement. The presented model is limited regarding the auralization of dynamic flight procedures, such as climbing or sinking maneuvers. Furthermore, the extracted oscillator and filter coefficients describe a specific drone type (DJI Mavic Pro) and may not be valid for larger drones. However, the documentation of the modeling steps are applicable for further drone types and flight maneuvers. The tonal synthesis of the drone presented in the chapter is necessary for a very small vehicle, where tonal noise is clearly dominant. In heavier drones, the noise signature might be dominated by broadband noise. The drone auralization model can be applied in future research on urban soundscapes where UAVs may appear in a great variety of use cases. Since the model describes the source properties free from outdoor sound propagation effects, it can also be used as an input for auralization frameworks that specifically study the influence of different wind and weather conditions on drone noise. Furthermore, it can

deliver input data for simulation tools where the spatial radiation characteristics of a drone should be included, such as the development of array-based drone detection.

6

Case study: Auralization and visualization of complex urban environments

Parts of this chapter have been published in: Rehman, R., Dreier, C., Heck, J., Llorca-Bofí, J., Vorländer, M. (2025). Comparison of Virtual Reality and Web-Based Listening Experiments on the Perception in Complex Auralized Environments. The Journal of the Acoustical Society of America, ***157****(3), 2001–2017* [141].

This chapter describes the integration of the traffic sound source models presented in the chapters 3 and 5 to a real-time implementation of complex urban scenarios using combined auralization and visualization. Not explicitly added to the following case study, the air-traffic auralization model presented in chapter 4 has been used in a dedicated journal paper to assess weather-dependent effects on the psychoacoustical perception of aircraft noise [35]. A part of the results has been demonstrated in an according 360° video [1]. From an objective point of view, the traffic sound auralization models that were presented in the previous chapters should reproduce the spatial and spectro-temporal characteristics of its real-world counterparts. At the example of four different scenarios, the virtual urban environment is subjectively evaluated by means of a listening experiment in VR with regard to their acoustical convincingness and immersion.

Nowadays, the creation of immersive virtual reality scenarios – by means of real-time performance to freely interact with the virtual environment (meaning within 50 milliseconds system response time) and multimodality in 3D – is particularly advanced in game industries. The degree of immersion of a VR scene can be

[1] Link to 360° urban VR demo (Last viewed 12/16/2024)

experimentally inverstigated by different criteria, such as the plausibility of the environment (by means of a spatial and source perception), freedom to interact with the environment in the sense of being part of it, and causality, should ideally be fulfilled, which is commonly summarized under experience of presence [77]. However, in contrast to limiting the computational complexity through non-physical modeling as achieved in computer games, physics-based auralization of complex urban scenarios in real-time is a promising, but technically challenging approach: From a theoretical point of view, urban noise auralization deals with inherent non-LTI (i.e. linear time-variant) properties of the propagation medium (air) that occurs due to turbulence-induced advection as well as refraction due to temperature, humidity and pressure inhomogeneities. In the following section 6.1, technical details regarding an exemplary virtual urban environment and the auralization chain are presented.

6.1 Virtual urban environment IHTApark

Striving for establishing a tool to integrate latest research results and to investigate modifications of urban environments in VR, the IHTApark framework (see Fig. 6.1) has been developed at the Institute for Hearing Technology and Acoustics [110].

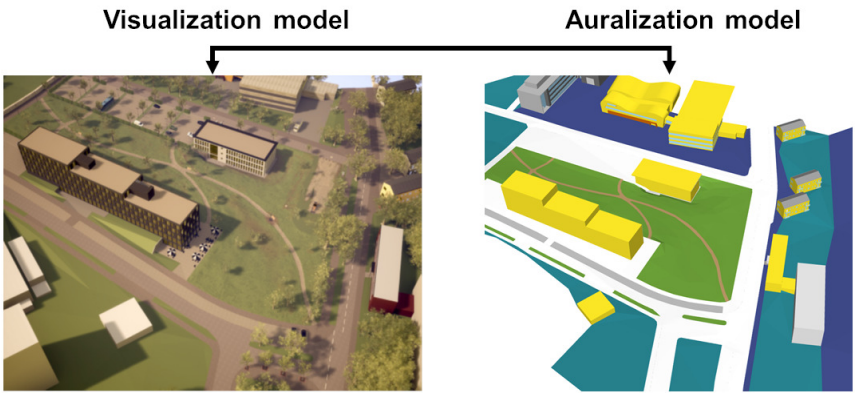


Figure 6.1: Visualization (left) and auralization model (right) of the virtual IHTApark framework[110].

Auralization chain

The sound is rendered with the open-source auralization framework Virtual Acoustics (VA) [74], based on the auralization model (Fig. 6.1, right), which contains geometrical information about the orientations of building surfaces and their acoustical material properties. VA is based on real-time capable algorithms written in C++ and generally enables sound reproduction using loudspeaker-based and headphone-based systems with motion tracking. It includes three features that prerequisitely enable the experiment presented in the following:

- An image source algorithm from the RAVEN software [156], and
- binaural signal rendering with head-tracking, and
- a plug-in interface to the 3D computer graphics software Unreal Engine (UE), which reproduces the visualization model (Fig. 6.1, left) using a head-mounted display (HMD). Furthermore, UE allows the user for free translational movement within the model by using the HMD motion controller.

The auralization is following the signal chain, as introduced in Figure 2.1, consisting of separated simulation modules for sound sources, propagation rendering and the reproduction module. The ambient background sound is a calibrated background recording using a spatial first-order Ambisonics format from the publicly available database published in [111]. The Ambisonics signals were postprocessed considering the equalization curves proposed by Ahrens [4].

The concept of vehicle pass-by sound auralization used in the experiment is shown in Figure 6.2. Since this thesis presents traffic sound source models coding spatial and spectro-temporal descriptions with free-field properties, the simulation engine for sound propagation has to add further information along all source-to-receiver paths.

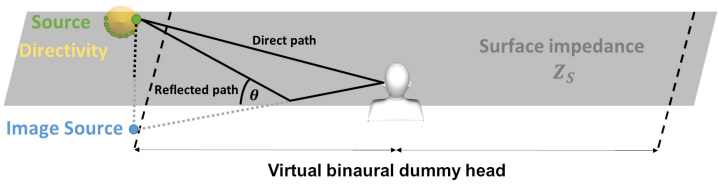


Figure 6.2: Digital twin model of vehicle pass-by noise auralization.

The propagation is simulated along multiple propagation paths so that the source model should have spherical properties (yellow sphere). Practically, this is imple-

mented using the OpenDAFF format with spatio-temporal discretized information (green circles). Along each path, acoustical properties can be calculated based on the propagation effects described in ISO 9613-2 [82]. The acoustical properties of surface materials, such as their absorption coefficients or surface impedances Z_s are provided to the propagation algorithm by means of parameter-based material models, such as shown for non-porous and porous asphalt in appendix A.

Higher-order reflections and diffractions in urban environments

As in room acoustics simulation, the sound propagation in urban environments can be based on measured or simulated methods. The biggest challenge for real-time outdoor sound propagation simulations are the big number of reflected and diffracted paths that propagate over very large distances – up to several kilometers at the example of aircraft noise. Aiming at real-time auralization, the calculation of the numerical solution of the wave equation (so-called wave-based acoustics, as introduced in section 2.3.3) is too computationally expensive for the audible frequency range, so that in this study, the auralization is based on the assumptions of geometrical acoustics (GA).

From the source's point of view, these outgoing waves emanate in different horizontal and vertical directions. The auralization renderer convolves every path's impulse (defined by its delay τ and gain g) with the according spatial radiation signal (see Fig. 6.2). One algorithm of GA is the image source method. Here, sound paths between sources and a receiver are found assuming wave propagation along straight lines and following the angular rules of incidence and reflection at model surfaces.

Since diffraction cannot directly be simulated by GA methods, modified diffraction models based on asymptotic formulations – where infinite edge's contribution (the diffraction wave emanating from the edge) are described by an explicit expression – can be used for efficient computation, e.g., the Uniform Theory of Diffraction (UTD) [96]. An open-source tool for finding sound paths in an urban environment including higher-order reflections and diffractions from a building's surfaces and edges by computation of a deterministic solution to this problem based on the image edge model was developed by Erraji [41].
As exemplarily shown in Figure 6.3, the second-order reflection and diffraction paths are simulated inside the IHTApark auralization model between positions of a fixed vehicle on the street (S) and a virtual receiver in the park (R), which have a distance of 58.4 m [110].

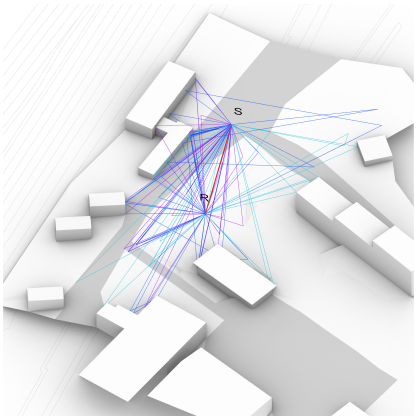


Figure 6.3: Simulated second-order reflection and diffraction paths inside the IHTApark auralization model between positions of a fixed vehicle on the street (S) and a virtual receiver in the park (R).

The second-order reflection and diffraction paths in the presented IHTApark scenario are added at the receiver and contribute with individual delays and amplitudes to the resulting impulse response shown in Figure 6.4 (left) and consists of 49 impulses (Fig. 6.4, left). The according transfer function (Fig. 6.4, right) reflects the spectral influences of the outdoor sound propagation effects as defined by [82], i.e. geometrical spreading loss, atmospheric attenuation, ground effect and diffraction. A theoretical influence by turbulences is neglected, assuming that they are not audible in urban environments. This transfer function (assuming second-order reflection and diffraction) can be compared to the excess attenuation in Figure A.4, resulting from first-order simulations and comparing non-porous and porous asphalt. Differences between GA-simulated and measured impulse responses in the IHTApark model has been critically discussed in [65].

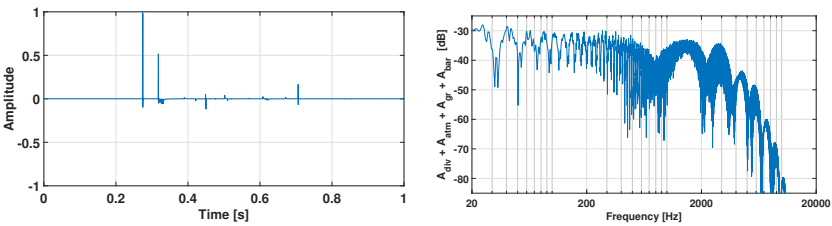


Figure 6.4: Impulse response (left, normalized) and according transfer function (right) between a vehicle on the street and a receiver in the IHTA park (Fig. 6.3).

6.2 VR-based listening experiment

6.2.1 Urban scenarios description

Using the presented real-time auralization and visualization framework, four different experimental scenarios were implemented (being part of a more comprehensive study on VR and web-based listening experiments, published in [141]). Presented in Figure 6.5, these four scenarios should represent four clearly distinguishable noise conditions regarding their objective noise measures by means of A-weighted sound pressure levels. As baseline of all four scenarios, the auralization consisted of a continuous car traffic at 50 km/h on the street next to the park, and additionally of a social picnic scene with four students talking on a blanket. The "Calm" reference condition had no additional noise sources. The other three conditions – "Bus", "Motorcycle" and "Drone" – included peak events from a passing bus and motorcycle, or from a drone continuosly looping on a track above the park.

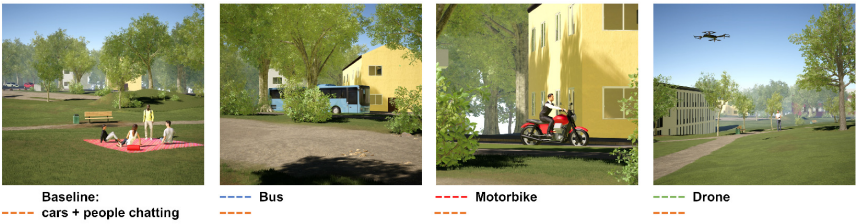


Figure 6.5: First-person representation of the four noise conditions: Baseline, bus, motorcycle and drone.

6.2.2 Experimental design and procedure

The experiment was structured as a within-subject study so that each participant experienced each noise condition in VR. For each condition, the participants moved freely in the IHTApark environment for a period of two minutes, while actively listening to the soundscape, similar to a real-world soundwalk procedure. The order of environments was randomized across the entire sample group. Following the two-minute exploration phase for each noise condition, the participants completed a questionnaire regarding their acoustical convincingness and immersion ratings. Answers were given on a five-point Likert scale (0-4, lowest-highest) The calibrated audio playback chain consisted of Sennheiser HD650 headphones,

connected to a Yamaha Steinberg UR44C audio interface. Finally, in order to compensate for the headphone characteristics on a GRAS 45CA headphone test fixture, the headphone-to-ear transfer function (HpTF) was compensated according to the method proposed by Schärer [154]. The visual playback system used a head-tracked HTC Vive Pro 2 HMD. For free movement in six degrees of freedom (6DoF), participants controlled translations with the HMD controller and managed rotations with head movements.

Sample group

The sample group consisted of 22 participants (15 male, 7 female), with an average age of 27.8 years. 13 participants were classified as locals, having previously visited IHTApark in real life.

6.2.3 Results

The ratings of the presented urban VR scenarios were assessed in terms of immersion and audio-visual convincingness. Immersion was evaluated using two items from the Igroup presence questionnaire [77]. One item focused on the sense of "being there" within the virtual world, while the other assessed the level of attention still given to the real world during immersion. The wording of the items was slightly adapted to suit the experiment.

The overall feeling of "being there" in VR was rated by means of mean μ and standard deviation σ – with respect to the answers being given on a five-point Likert scale between 0 (lowest) and 4 (highest) – with $\mu = 3.273$, $\sigma = 0.827$. The amount of attention the subjects paid was rated $\mu = 3$, $\sigma = 0.873$. From these results can be concluded that subjects were able to focus on the virtual worlds and black out the real world around them well.

Investigating visual and acoustical convincingness separately revealed the auralization to be perceived as more convincing ($\mu_{\mathrm{VR}} = 3.318$, $\sigma_{\mathrm{VR}} = 0.716$) than the visual rendering ($\mu_{\mathrm{VR}} = 2.591$, $\sigma_{\mathrm{VR}} = 1.054$).

6.3 Summary and discussion

All ratings evaluated in section 6.2.3 are high enough to label the presented urban VR environment and – as particularly in focus of this thesis – the real-time traffic sound auralization as convincing. These results pave the way for further explo-

ration of the approach's potential. Besides research, possible applications of such a physics-based auralization and visualization approach are general behavioral studies, urban planning and the comparison of the acoustical signatures of virtual prototypes under natural conditions.
Despite its promising potential, some limitations should be considered: Firstly, as the auralization was based using GA, it accounted for specular reflections and scattering, and did not include effects of an inhomogeneous medium and edge diffraction. Additionally, the study was conducted in a laboratory with high-quality calibrated equipment. As such, potential effects of audio-visual playback latencies or low-quality and uncalibrated user equipment were not investigated. Both factors could significantly degrade audio-visual convincingness and should be addressed separately before claiming it on large-scale demonstrators.
In the course of current research, the ecological validity of VR methods for urban sound assessments was discussed, that means the ability of laboratory experiment results to predict behavioral findings in real-world situations. Xu et al. conducted a comprehensive literature review highlighting the ecological validity of immersive virtual reality techniques for perceiving urban sound environments. They concluded VR techniques, particularly with the use of head tracking, to be an ecologically valid tool for soundscape and noise assessments [185]. Yang et al. studied the ecological validity of different audio-visual reproduction technologies with different degrees of authenticity, ranging from simple to high, in soundscape research. These technologies included field recordings, augmented reality (AR), virtual reality (VR), and audio-only presentations [186]. Their findings emphasize the necessity for selecting a reproduction system with a high authenticity rating, that means a setup consisting of a head-tracked audio-visual rendering by using HMDs in combination with headphones.

There is significant potential for integrating auralizations with 3D city models in future research, opening new paths for exploration of noise and soundscape perception. Significant scientific contributions to the understanding of noise perception are based on psychological and social-psychological research. Many international studies have shown that objective noise measures – such as the commonly used A-weighted equivalent continuous sound pressure level $L_{\mathrm{A,eq}}$ – merely explains a fraction of the statistical dispersion of the found subjectively assessed noise effects (mostly annoyance). In a landmark paper, Job calculated an average correlation coefficient of $r = 0.42$ in a meta-analysis of 39 studies [86]. This means that the variance of the subjective reaction to noise can be explained only up to about $r^2 \cdot 100 \approx 20\%$ by the objective noise measure itself. By optimising the objective noise measure – i.e. by inclusion of peak ($L_{\max}$) and percentile values (e.g. L_5), difference in peak exposure (L_5 compared to L_{95}), inclusion of the number of

noise events and inclusion of quiet periods – the extent of the explained variance could be increased by only about 10%. In the further course of research, the explained variance could be increased by a further third by inclusion of non-acoustic factors due to so-called moderators, , i.e. personal and social aspects of the residents [61]. At the current state of research, the last missing third of the explained variance has not been elucidated. To this day, the influencing human factors on noise assessments (in terms of annoyance ratings) are not yet fully discovered by research.

This in turn motivates the methodological advance of using virtual acoustics for noise research as being pursued with this thesis: A reproducible simulation framework to combine spatial auralizations and 3D visualization to enable reliable and controlled laboratory environments for noise experiments.

7

Conclusion and Outlook

Early auralization techniques supported three degrees of freedom (3DoF), enabling real-time rendering of head rotations with according yaw, pitch and roll informations. However, modern auralization strives for six degrees of freedom (6DoF) with additional translations, allowing the user to freely move around in a virtual scene. Given the high computational cost associated with the large number of moving sound sources, existing outdoor auralization frameworks often compromise by being either incapable of real-time operation or reliant on pre-computed intermediate results.

In this thesis, modeling and real-time capable synthesis methods for traffic sound sources (road and air) are proposed and implemented (chapters 3,4 and 5). Further, these are integrated to a real-time framework for dynamic auralization and visualization of complex urban scenarios (chapter 6). The perceptual ratings from a listening experiment (evaluated in section 6.2.3) are high enough to label the presented urban virtual reality environment and the real-time traffic sound auralization as convincing. The specific implementations stem from decisions shaped by a meticulous evaluation of established methodologies and implementation of promising concepts, which are critically concluded herein. Furthermore, the outlook explores the implications for future projects and addresses potential practical challenges.

In order to achieve a plausibe reproduction of a traffic sound source in all its acoustical details, it is necessary to focus beyond adequately reproducing spectral and spatial source characteristics at different operating states – which already has sufficiently been covered by a broad range of detailed semi-empirical source models – on the source's temporal phenomena (cf. section 2.4). In this regard, this thesis specifically focuses on modulation phenomena, which can be observed in drone sound (section 5.3) as well as fan noise emission (section 4.4). However, with regard to sound sources, further research is needed. Nevertheless, the evolution

of temporal phenomena integration can already be observed in the area of sound wave propagation simulation.

Semi-empirical models describe near-field directivity patterns, by neglecting the shielding effect of fuselage and wings on acoustical far field patterns. In case of jet noise, this has been investigated in section 4.3 by a proof of concept based on high-fidelity flow simulations with complexity reduction using spherical harmonics. This approach offers a computationally efficient solution for integrating complex sound fields of acoustical sources with a minimal number of parameters to large-scale simulations.

The following Figure 7.1 classifies the sound source-specific novelty aspects presented in this thesis, colored according to the dimensions of source characterization introduced in Figure 2.4.

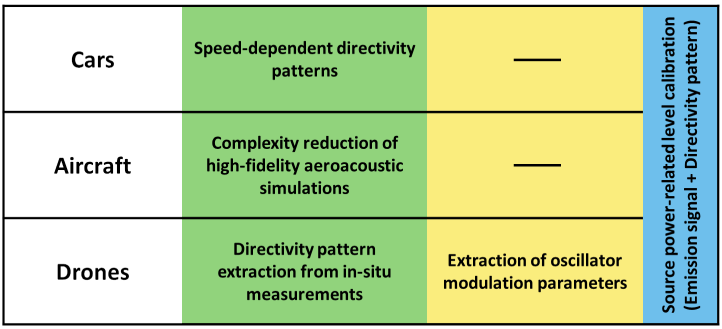


Figure 7.1: Overview of source-specific novelty aspects related to the dimensions of source characterization (cf. Fig. 2.4) regarding spatial (green), temporal (yellow) and spectral aspects (blue).

At the current state, less elaborate than aircraft noise auralization models are those for road-traffic sound sources. As reported in literature, directivity patterns have merely been considered based on generic patterns from noise mapping. However, these emission models neglect differences between vehicle classes, driving speed and its operating state. Based on the results presented in section 3.5, such dependencies on the directivity patterns can be clearly observed and have therefore been published in an open database. This database serves as a resource for researchers and practitioners in the acoustics field, enabling to develop more accurate and case-specific auralization models for road-traffic sound sources.

As presented in chapter 6, the developed road-traffic and drone models were integrated to a real-time framework for complex urban scenarios using combined

auralization and visualization. Using this framework, four different experimental scenarios were evaluated in order to assess their immersion and audio-visual convincingness. As mentioned above, all ratings evaluated in section 6.2.3 are high enough to label the urban VR environment and – as particularly in focus of this thesis – the real-time traffic sound auralization as convincing. These results motivate the methodological advance of using virtual acoustics for noise research as being pursued with this thesis: A reproducible simulation framework to combine spatial auralizations and 3D visualization to enable reliable and controlled laboratory environments for noise experiments.

Further research is needed for all traffic sound sources discussed within this thesis, particularly in two aspects: Firstly, except for aircraft, sound sources in urban environments are often close enough to the receiver so that they are audible as distributed sound sources (cf. section 2.4.3). Given their impact on computational cost, it is crucial to perceptually evaluate at what source-to-receiver distances and relative speeds these sound sources should be better modeled as distributed sources or single compact sources. This aspect is accompanied by the problem of installation effects, which involves a change in sound radiation between the installed and non-installed state of sound sources. In this thesis, primarily complete systems were examined and, if necessary, spectro-temporally extracted into subsystems. However, the opposite scenario is also of practical relevance: Auralization of systems with variable configuration, based on partial sound sources separately measured in the laboratory. Secondly, transient processes in sound syntheses, such as the backfire sound in engine emissions, have been reported to very small extent. The impact of these processes on the tonal quality of sound syntheses remains unexplored, however, being promising for further plausibility enhancement. As an example, in the context of urban auralization, this could include the sound of loose stones on gravel squares when cars drive over. These transient sounds play a significant role in creating a more believable and immersive auditory experience, as they capture the unique and often unpredictable acoustic events that occur in real-world environments.

To conclude, this thesis contributes to the field of traffic sound auralization. It presents advanced modeling techniques and real-time synthesis methods for sound sources that produce plausible auralization results, which are highly relevant for future research on noise in urban environments. By addressing the key challenges and identifying areas for further research, this work paves the way for the continued evolution of accurate, plausible and immersive urban auralizations. Future research will focus on enhancing the computational efficiency of traffic sound auralizations across various applications, including virtual reality environments and noise studies.

Appendices

A

Complex boundary conditions at the ground

The sound field of road-traffic vehicles in an urban environment can be assumed as a superposition of a direct wave, a ground reflected wave and reflections from building façades, with the ground reflection being most influential. Therefore, it is necessary to consider the acoustical material properties at the ground. Depending on the ground material, the delay and amplitude of the reflected wave is affected due to the frequency-dependent absorption coefficients. In order to compare the influence of a (noise-reducing) porous with a non-porous asphalt surface the sound field is simulated in the following using different ground material properties.

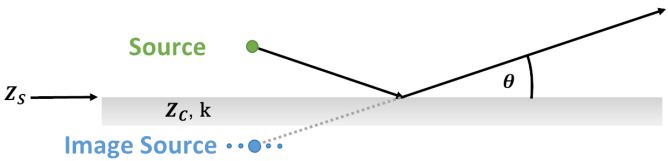


Figure A.1: Fuzzy image source at a non-locally reacting ground.

In the case of a non-locally reacting ground such as porous asphalt, the frequency-dependent absorption coefficient of a material α is derived from its plane-wave reflection coefficient R_p based on the geometry-dependent complex surface impedance Z_s, the complex characteristic ground impedance Z_c, complex wave number k, $k_0 = 2\pi f/c$ (c being the homogeneous sound speed of air), its layer thickness d and the grazing angle Θ by

$$\alpha = 1 - |R_p|^2 \tag{A.1}$$

with

$$R_p = \frac{Z_s \sin\Theta - \chi}{Z_s \sin\Theta + \chi} \tag{A.2}$$

with

$$\chi = \left[1 - (k_0/k)^2 \cos^2\Theta\right]^{1/2} \tag{A.3}$$

and (assuming a rigid-backed layer)

$$Z_s = Z_c \cdot \coth(-ikd). \tag{A.4}$$

Z_c and k are resulting from the above-mentioned Hamet [17] and Miki [120] model equations for equivalent fluids. Whereas the phenomenological Hamet model depends on three parameters (flow resistivity R_s, porosity Ω and tortuosity q^2, the Miki model merely depends on R_s. Compared to the more popular Delany-Bazley model equations, the Miki expressions perform better in particular for the very low frequency range $f/R_s < 0.01$ for double layer impedances. The Hamet model equations are:

$$Z_c = \rho_0 c_0 \frac{1}{\Omega}\sqrt{\frac{K}{\gamma}} \frac{\sqrt{1 - i\frac{f_\mu}{f}}}{\sqrt{1 - (1 - \frac{1}{\gamma})\frac{1}{1 - i\frac{f_\Theta}{f}}}} \tag{A.5}$$

and

$$k = k_0\sqrt{K\gamma}\sqrt{1 - i\frac{f_\mu}{f}}\sqrt{1 - (1 - \frac{1}{\gamma})\frac{1}{1 - i\frac{f_\Theta}{f}}} \tag{A.6}$$

where $Z_0 = \rho_0 c_0$ is the specific wave impedance in air, $\gamma = c_p/c_v = 1.4$ is the classical ratio of specific heats for air, $f_\mu = \frac{1}{2\pi}\frac{R_s}{\rho_0}\frac{\Omega}{K}$ and $f_\Theta = \frac{1}{2\pi}\frac{R_s}{\rho_0}\frac{1}{N_{pr}}$ – with $N_{pr} = 0.71$ being the Prandtl number for air – are characteristic frequencies associated to the viscous and thermal effects respectively. The theoretical prediction are plotted in Fig. A.2.

The asphalt surface impedance is calculated assuming a rigid-backed layer with a thickness $d = 0.04$m. In the case of non-porous asphalt a flow resistivity value R_s for stone mastic asphalt (SMA) was assumed ($R_s = 6 \cdot 10^7 \frac{\mathrm{Ns}}{\mathrm{m}^4}$). The ground impedance yields

$$Z_c = \rho_0 c_0 \left[1 + 5.5\left(10^3 \frac{f}{R_s}\right)^{-0.632} - j8.43\left(10^3 \frac{f}{R_s}\right)^{-0.632}\right] \tag{A.7}$$

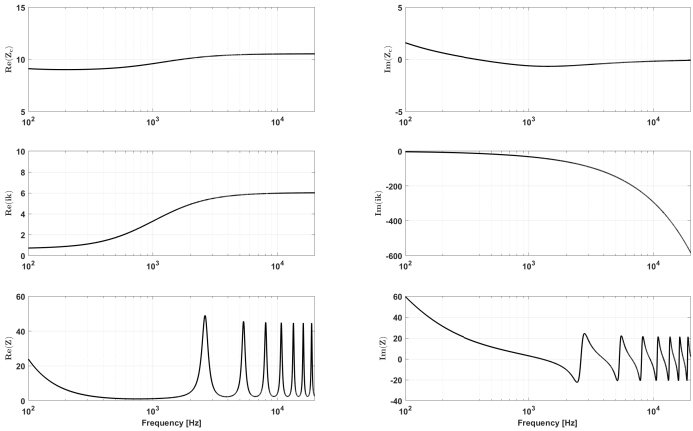


Figure A.2: Theoretical predictions for porous asphalt based on the Hamet model equations for parameters $R_s = 5000\,\frac{\mathrm{Ns}}{\mathrm{m}^4}$, $\Omega = 15\%$, and $q^2 = 2.5$: (a) characteristic impedance Z_c, (b) propagation constant ik, and (c) surface impedance Z_s.

and the complex wave number yields

$$k = k_0 \left[1 + 7.81 \left(10^3 \frac{f}{R_s} \right)^{-0.618} - j11.41 \left(10^3 \frac{f}{R_s} \right)^{-0.618} \right]. \tag{A.8}$$

The modeled frequency-dependent absorption coefficient (Fig. A.3) shows good agreement with empirical results based on impedance tube measurements from literature (c.f. Fig.5 in [176]).

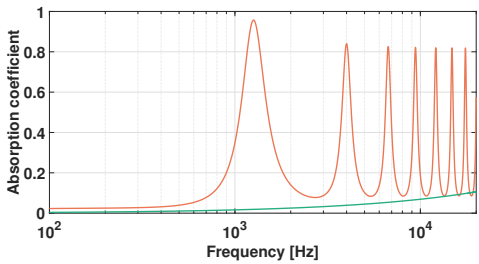


Figure A.3: Absorption coefficients for rigid-backed 4-cm-thick porous asphalt, obtained from the Hamet model (orange curve) for values of $R_s = 5000\,\frac{\mathrm{Ns}}{\mathrm{m}^4}$, $\Omega = 15\%$, and $q^2 = 2.5$, and for non-porous asphalt obtained from the Miki model (green curve) for a flow resistivity of $R_s = 6 \cdot 10^7\,\frac{\mathrm{Ns}}{\mathrm{m}^4}$.

The high absorption coefficient of porous asphalt at 1 kHz might explain the noise-reducing effectiveness of such surfaces since the Horn effect of the tire/road noise component as major contribution in this frequency range might be suppressed.

Based on both absorption coefficients, the impact on the sound propagation transfer functions is calculated (Fig. A.4). From a theoretical point of view, this problem is that of sound propagation along an impedance plane and is implemented as comprehensively discussed in [27]. This approximation – that was obtained by numerical double saddle point method of integration from an analytic form assuming a spherical wave as infinite number of plane waves – is valid for $k_0\, r_2 \gg 1$ where r_2 is the source-receiver distance of the ground-reflected wave, small grazing angles $\cos\Theta \ll 1$ and hard boundaries $|\,\beta\,|^2 \ll 1$ where $\beta = Z_0/Z_s$ is the specific acoustic admittance ratio of the surface. Mathematically, the plane-wave reflection coefficient R_p in equation (A.2) is modified to

$$Q = R_p + (1 - R_p)F(w) \tag{A.9}$$

where Q is the spherical reflection coefficient of the ground. The term $(1-R_p)F(w)$ accounts for a surface wave component, where

$$F(w) = 1 + i\sqrt{\pi}w\exp(-w^2)\,\mathrm{erfc}(-iw). \tag{A.10}$$

In this equation a virtually extended reaction surface is described by

$$w^2 = 2ik_0r_2\,\chi^2\,[Z_{\mathrm{norm}}(1 - R_p)]^{-2} \tag{A.11}$$

using the normalized surface impedance $Z_{\mathrm{norm}} = Z_s/Z_0$. The high absorption coefficient of porous asphalt at multiple harmonic frequencies leads to a theoretical noise reduction of 5−6dB at the receiver, compared to a conventional stone mastic asphalt layer (Fig. A.4).

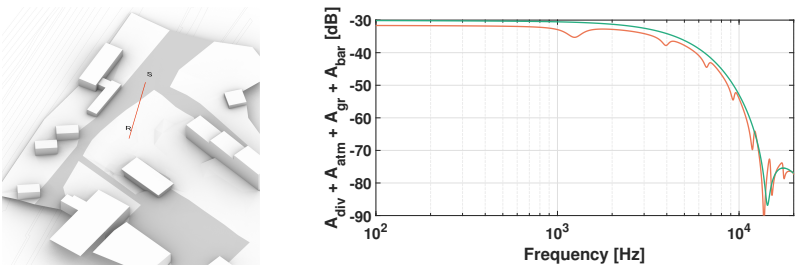


Figure A.4: Influence of different ground materials on excess attenuation between a source S and a receiver R based on the material properties of porous (orange) and non-porous asphalt (green).

B

Wet asphalt auralization filter data

Table B.1: Parametric equalizer parameters for auralization of wet asphalt surfaces by application to dry condition auralizations.

Filter type	Frequency [Hz]	Gain [dB]	Quality factor
Low shelf	9685.77	-1.08	0.28629
Peak	2261.98	5.54	2.14672
Peak	2364.36	-6.02	1.12317
Peak	2973.09	-0.76	6.75985
Peak	3057.59	5.33	1.59375
Peak	4388.94	5.63	2.06575
Peak	6044.31	3.02	1.65248
Peak	7980.59	2.15	2.23261
Peak	10338.40	5.58	1.72285
Peak	12416.61	2.32	3.42560
Peak	19999.99	5.70	0.92439
High shelf	14743.06	5.50	0.73392

C

Vehicle noise sources

Table C.1: Partial sound sources of vehicle exterior noise (compiled from [29]).

Tire/road noise	Powertrain noise	Aerodynamic noise
1) Vibration-related mechanisms • Radial vibrations • Tangential vibrations • Stick-Snap (adhesion) • Sidewall vibrations • Tire belt/Carcass vibrations • Stick-slip (tangential motions) **2) Aerodynamically related mechanisms** • Air pumping • Cavity/Helmholtz resonances • Horn amplification effect • Pipe resonances	• Exhaust outlet • Intake outlet • Induction system • Transmission and gears • Cooling fan	**1) Cavities** • Hood front edge • Wheel wells • Body underside • Windshield base • Windshield/Front door transition (A-pillar) **2) Structure gaps** • Around doors, trunk lid and sunroof • Rain gutters **3) External equipment** • Rear-view mirrors • Windshield wipers • Antennas

D

Semi-empirical aircraft noise model input parameters

Symbol	**Parameter explanation**	**Value**
ΔT	Temperature rise across fan/compressor stage	45K
$\dot{m}$	Mass flow rate through fan or compressor stage	420kg/sec
M_T	Mach number of rotor tip	1.04
M_{TR}	Rotor tip relative inlet Mach number: M_T related to the Mach number of the air in the inlet	1.2
$(M_{TR})_D$	Rotor tip relative inlet Mach number	1.2
RSS	Rotor-stator spacing (rel. to rotor blade span)	200%
RPM_{fan}	Fan shaft rotational speed	100Hz
RPM_{HPC}	Compressor shaft rotational speed	390Hz
B_{fan}	Number of rotor blades of fan	40
V_{fan}	Number of stator blades of fan	90
B_{HPC}	Number of rotor blades of compressor	20
V_{HPC}	Number of stator blades of compressor	35
S_{wing}	Area of both wings	437.3m^2
S_{hTail}	Area of horizontal tail	93m^2
S_{vTail}	Area of vertical tail	47.65m^2
b_{wing}	Span of wing	61.2 m
b_{hTail}	Span of horizontal Tail	21.5 m
b_{hTail}	Span of vertical Tail	9.44 m
ν	Kinematic viscosity of air	$1.5 \cdot 10^{-5}$
/	Type of leading edge lift device (slats of flaps)	slats
/	Number of wheels of nose and main landing gear	2 and 4
D	Wheel diameter of landing gear tires	1.2 m
H	Length of main landing gear struct	2.5 m
S_F	Total area of all flaps combined	21.1m^2
c_F	Chord of trailing edge flaps	0.78m^2
δ_F	deflection angle of flaps (take off/landing)	20/60°
/	number of slotted flaps (1,2, or 3)	1
T_a	Ambient (outside) air temperature	8°C
c_a	Ambient speed of sound	336 m/s
ρ_a	Ambient air density	1.124
ρ_{isa}	Air density of int. standard atmosphere (ISA)	1.225kg/m^3
ρ_j	Density of jet stream	0.515kg/m^3
c_{isa}	Speed of sound of ISA	340 m/s
D_n	Nozzle diameter	1.4 m
v_j	Speed of jet stream	400 m/s
β	Angle of attack	0°
T_j	Jet stream temperature	704°C
R	Distance of observer to source	1 m
V	Velocity of aircraft	240 m/s

Table D.1: List of required parameter for the prediction models

Bibliography

[1] B. Plovsing, J. Kragh: Nord2000. Comprehensive outdoor sound propagation model. Part 1: Propagation in an atmosphere without refraction. DELTA, report AV 1849/00, 2001.

[2] Harmonoise project. Final report of work package 1.1: Source modelling of road vehicles. Available at https://www.imagine-project.org/bestanden/D09_WP1.1_HAR11TR-041210-SP10.pdf, accessed on 2024-10-10.

[3] R. Aalmoes, M. den Boer, and H. Veerbeek. "Virtual reality aircraft noise simulation for community engagement". In: *INTER-NOISE: 47th International Congress and Exposition on Noise Control Engineering*. 2018.

[4] J. Ahrens and C. Andersson. "Perceptual evaluation of headphone auralization of rooms captured with spherical microphone arrays with respect to spaciousness and timbre". In: *The Journal of the Acoustical Society of America* 145.4 (Apr. 2019), pp. 2783–2794. DOI: 10.1121/1.5096164.

[5] K. Ahuja, B. Tester, and H. Tanna. "The free jet as a simulator of forward velocity effects on jet noise". In: *NASA/CR-3056* (1978).

[6] M. Alkmim. "Auralization of moving sources – Application to pass-by and fly-over real-time sound synthesis". PhD thesis. Leuven, Belgium: Katholieke Universiteit Leuven, 2023. URL: https://lirias.kuleuven.be/handle/20.500.12942/718401.

[7] M. Alkmim, J. Cardenuto, E. Tengan, T. Dietzen, T. Van Waterschoot, J. Cuenca, L. De Ryck, and W. Desmet. "Drone noise directivity and psychoacoustic evaluation using a hemispherical microphone array". In: *The Journal of the Acoustical Society of America* 152.5 (2022), pp. 2735–2745. DOI: 10.1121/10.0014957.

[8] K. Attenborough. "Sound propagation in the atmosphere". In: *Springer Handbook of Acoustics*. New York: Springer, 2014. Chap. 4, pp. 117–155. DOI: 10.1007/978-1-4939-0755-7_4.

[9] A. Aumann, B. Tuttle, W. Chapin, and S. Rizzi. "The NASA auralization framework and plugin architecture". In: *INTER-NOISE: 44th International Congress and Exposition on Noise Control Engineering*. 2015.

[10] W. Aures. "Der sensorische Wohlklang als Funktion psychoakustischer Empfindungsgrößen (Sensory pleasantness as a function of psychoacoustic sensations)". In: *Acustica* 58 (1985), pp. 282–290.

[11] A. Avni, J. Ahrens, M. Geier, S. Spors, H. Wierstorf, and B. Rafaely. "Spatial perception of sound fields recorded by spherical microphone arrays with varying spatial resolution". In: *The Journal of the Acoustical Society of America* 133.5 (2013), pp. 2711–2721. DOI: 10.1121/1.4795780.

[12] S. Baldan, H. Lachambre, S. D. Monache, and P. Boussard. "Physically informed car engine sound synthesis for virtual and augmented environments". In: *IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*. 2015. DOI: 10.1109/SIVE.2015.7361287.

[13] H. Bass, L. Sutherland, A. Zuckerwar, D. Blackstock, and D. Hester. "Atmospheric absorption of sound: Further developments". In: *The Journal of the Acoustical Society of America* 97.1 (1995), pp. 680–683. DOI: 10.1121/1.412989.

[14] T. Beckenbauer. "Road Traffic Noise". In: *Handbook of Engineering Acoustics*. Ed. by G. Müller and M. Möser. Springer Berlin Heidelberg, 2013, pp. 367–392. DOI: 10.1007/978-3-540-69460-1_15.

[15] T. Beckenbauer and A. Kuijpers. "Prediction of pass-by levels depending on road surface parameters by means of a hybrid model". In: *INTER-NOISE: The 2001 International Congress and Exhibition on Noise Control Engineering*. 2001, pp. 2528–2533.

[16] L. Beranek. "Muffling the Jet Age". In: *Riding the Waves*. The MIT Press, 2008, pp. 125–139. ISBN: 978-0-26-202629-1. DOI: 10.7551/mitpress/7639.003.0006.

[17] M. Bérengier, M. Stinson, G. Daigle, and J. Hamet. "Porous road pavements: Acoustical characterization and propagation effects". In: *J. Acoust. Soc. Am.* 101.1 (1997), pp. 155–162. DOI: 10.1121/1.417998.

[18] B. Berglund, T. Lindvall, and D. Schwela. *Guidelines for community noise*. Tech. rep. Geneva: World Health Organization, 1999.

[19] L. Bertsch, S. Guérin, G. Looye, and M. Pott-Pollenske. "The Parametric Aircraft NoiseAnalysis Module - Status overview and recent applications". In: *17th AIAA/CEAS Aeroacoustics Conference*. 2011.

[20] J. Billingsley and R. Kinns. "The acoustic telescope". In: *Journal of Sound and Vibration* 48.4 (1976), pp. 485–510. ISSN: 0022-460X. DOI: 10.1016/0022-460X(76)90552-6.

[21] B. Boashash. "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals". In: *Proceedings of the IEEE* 80.4 (1992), pp. 520–538. DOI: 10.1109/5.135376.

[22] C. Bogey, S. Barré, and C. Bailly. "Direct Computation of the Noise Generated by Subsonic Jets Originating from a Straight Pipe Nozzle". In: *Int. J. Aeroacoust.* 7.1 (2008), pp. 1–21. DOI: 10.1260/147547208784079917.

[23] M. Bösing. "Acoustic Modeling of Electrical Drives". PhD thesis. Aachen, Germany: RWTH Aachen University, 2014. ISBN: 978-3-8440-2752-5.

[24] A. Bresciani, J. Maillard, and L. de Santana. "Physics-based scintillations for outdoor sound auralization". In: *The Journal of the Acoustical Society of America* 154.2 (2023), pp. 1179–1190. DOI: 10.1121/10.0020666.

[25] J. E. Bridges and C. A. Brown. "Parametric testing of Chevrons on Single Flow Hot Jets". In: *NASA/TM-2004-213107* (2004).

[26] W. Cheney and D. Kincaid. *Numerical Mathematics and Computing*. 6th. Belmont, USA: Brooks/Cole, 2008, p. 321. ISBN: 78-0-495-11475-8.

[27] C. Chien and W. Soroka. "Sound propagation along an impedance plane". In: *Journal of Sound and Vibration* 43.1 (1975), pp. 9–20. DOI: 10.1016/0022-460X(75)90200-X.

[28] COMSOL Multiphysics. *Acoustics Module User's Guide*. COMSOL AB. Stockholm, Sweden, 2021.

[29] M. J. Crocker. *Handbook of Noise and Vibration Control*. Hoboken: John Wiley & Sons, 2007. ISBN: 978-0-470-20970-7. DOI: 10.1002/9780470209707.

[30] "Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise". In: *Official Journal of the European Union* L 189 (2002), pp. 12–25.

[31] M. Dittrich and X. Zhang. "The Harmonoise/IMAGINE model for traction noise of powered railway vehicles". In: *Journal of Sound and Vibration* 293.3 (2006), pp. 986–994. DOI: 10.1016/j.jsv.2005.12.022.

[32] ICAO Doc 9082. Policies on charges for airports and air navigation services. International Civil Aviation Organization. Quebec, CAN, 2012.

[33] C. Dreier, J. Hahn, J. Heck, J. Llorca-Bofí, and M. Vorländer. "Real-time vehicle pass-by noise synthesis for urban sound auralization". In: *ICA - 24th International Congress on Acoustics*. Gyeongju, South Korea, 2022, pp. 1–8.

[34] C. Dreier and M. Vorländer. "Sound source modelling by nonnegative matrix factorization for virtual reality applications". In: *INTER-NOISE - 50th International Congress and Exposition on Noise Control Engineering*. Washington DC, USA, 2021, pp. 1053–1061. DOI: 10.3397/IN-2021-1742.

[35] C. Dreier and M. Vorländer. "Aircraft noise—Auralization-based assessment of weather-dependent effects on loudness and sharpness". In: *The Journal of the Acoustical Society of America* 149.5 (2021), pp. 3565–3575. DOI: 10.1121/10.0005040.

[36] C. Dreier, X. Vogt, W. Schröder, and M. Vorländer. "Acoustic source characterization of simulated subsonic jet noise using spherical harmonics". In: *The Journal of the Acoustical Society of America* 154.1 (2023), pp. 167–178. DOI: 10.1121/10.0020065.

[37] C. Dreier and M. Vorländer. "Road traffic auralization: Modeling and synthesis of electric drives". In: *Tecniacústica*. Cuenca, Spain, 2023, pp. 187–190. ISBN: 978-84-87985-33-1.

[38] C. Dreier and M. Vorländer. "Drone auralization model with statistical synthesis of amplitude and frequency modulations". In: *Acta Acustica* 13 (2024), p. 13. DOI: 10.1051/aacus/2024026.

[39] C. Dreier and M. Vorländer. "Speed-dependent directivity patterns of road-traffic vehicles". In: *The Journal of the Acoustical Society of America)* 157.4 (2025), pp. 2735–2749. DOI: 10.1121/10.0036382.

[40] Psychoacoustic metrics for ITT equipment – Part 2: Models based on human perception. Geneva, CH, 2022.

[41] A. Erraji, J. Stienen, and M. Vorländer. "The image edge model". In: *Acta Acustica* 5 (2021), p. 17. DOI: 10.1051/aacus/2021010.

[42] R. Ewert and W. Schröder. "Acoustic Perturbation Equations Based on Flow Decomposition via Source Filtering". In: *Journal of Computational Physics* 188.2 (2003), pp. 365–398. DOI: 10.1016/S0021-9991(03)00168-2.

[43] R. Ewert and W. Schröder. "On the simulation of trailing edge noise with a hybrid LES/APE method". In: *Journal of Sound and Vibration* 270.3 (2004), pp. 509–524. DOI: 10.1016/j.jsv.2003.09.047.

[44] A. Farnell. *Designing Sound*. Cambridge, Massachusetts: The MIT Press, 2010. ISBN: 978-0-262-01441-0.

[45] H. Fastl and E. Zwicker. *Psychoacoustics: Facts and Models*. Springer Series in Information Sciences. Springer, 2007. DOI: 10.1007/978-3-540-68888-4.

[46] C. Févotte and J. Idier. "Algorithms for nonnegative matrix factorization with the -divergence". In: *Neural Computation* 23.9 (2011), pp. 2421–2456. DOI: 10.1162/NECO_a_00168.

[47] A. Fiebig, R. Sottek, and E. Kuczmarksi. "Auralization of road traffic noise and its value for environmental noise assessment". In: *AIA-DAGA: Proceedings of the International Conference on Acoustics*. Merano, Italy, 2013, pp. 118–121.

[48] M. R. Fink. "Noise Component Method for Airframe Noise". In: *Journal of Aircraft* (1979). DOI: 10.2514/3.58586.

[49] M. Fisher, M. Harper-Bourne, and S. Glegg. "Jet engine noise source location: The polar correlation technique". In: *Journal of Sound and Vibration* 51.1 (1977), pp. 23–54. ISSN: 0022-460X. DOI: 10.1016/S0022-460X(77)80111-9.

[50] Flightpath 2050 – Europe's vision for aviation: maintaining global leadership and serving society's needs. Europäische Kommission and Generaldirektion Forschung und Innovation and Generaldirektion Mobilität und Verkehr. Publications Office, 2012. doi: www.doi.org/10.2777/15458.

[51] J. Forssén, A. Hoffmann, and W. Kropp. "Auralization model for the perceptual evaluation of tyre–road noise". In: *Applied Acoustics* 132 (2018), pp. 232–240. DOI: 10.1016/j.apacoust.2017.11.023.

[52] Y. Fu. "Auralisation of Traffic Flow using Procedural Audio Methods". PhD thesis. York, England: University of York, 2021.

[53] K. Genuit and A. Fiebig. "Alternative alert signal concepts and their perceptual implications". In: *INTER-NOISE: 45th International Congress and Exposition on Noise Control Engineering*. Hamburg, Germany, 2016, pp. 4207–4216.

[54] F. Georgiou, M. Hornikx, and A. Kohlrausch. "Auralization of a car pass-by using impulse responses computed with a wave-based method". In: *Acta Acustica united with Acustica* 105.2 (2019), pp. 381–391. DOI: 10.3813/AAA.919321.

[55] F. Georgiou, M. Hornikx, and A. Kohlrausch. "Auralization of a car pass-by inside an urban canyon using measured impulse responses". In: *Applied Acoustics* 183 (2021). DOI: 10.1016/j.apacoust.2021.108291.

[56] F. Georgiou. "Modeling for auralization of urban environments: Incorporation of directivity in sound propagation and analysis of a framework for auralizing a car pass-by". PhD thesis. Eindhoven, Netherlands: Technische Universiteit Eindhoven, 2018. ISBN: 978-90-386-4528-5.

[57] M. van der Giet, J. Blum, P. Dietrich, S. Pelzer, M. Müller-Trapet, M. Pollow, M. Vorländer, and K. Hameyer. "Auralization of electrical machines in variable operating conditions". In: *IEEE International Electric Machines & Drives Conference (IEMDC)*. Niagara Falls, Canada, 2011, pp. 1462–1467. DOI: 10.1109/IEMDC.2011.5994824.

[58] W. Graham, C. Hall, and M. Morales. "The potential of future aircraft technology for noise and pollutant emissions reduction". In: *Transport Policy* 34 (2014), pp. 36–51. DOI: 10.1016/j.tranpol.2014.02.017.

[59] D. Griffin and J. Lim. "Signal estimation from modified short-time Fourier transform". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32.2 (1984), pp. 236–243. DOI: 10.1109/TASSP.1984.1164317.

[60] E. Gröschel, W. Schröder, and M. Meinke. "Noise Generation Mechanisms in Single and Coaxial Jets". In: *12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustics Conference)*. 2006, pp. 1–18. DOI: 10.2514/6.2006-2592.

[61] R. Guski. "Personal and social variables as co-determinants of noise annoyance". In: *Noise and Health* 1.3 (1999), pp. 45–56.

[62] D. Hartmann, M. Meinke, and W. Schröder. "A strictly conservative Cartesian cut-cell method for compressible viscous flows on adaptive grids". In: *Computer Methods in Applied Mechanics and Engineering* 200.9-12 (2011), pp. 1038–1052. DOI: 10.1016/j.cma.2010.05.015.

[63] D. Hawkings. "Multiple tone generation by transonic compressors". In: *Journal of Sound and Vibration* 17.2 (1971), pp. 241–250. DOI: 10.1016/0022-460X(71)90458-5.

[64] M. Hayes. "Spectrum estimation". In: *Statistical digital signal processing and modeling*. New York: John Wiley & Sons, 1996. Chap. 8, pp. 391–492.

[65] J. Heck, C. Dreier, and M. Llorca-Bofí Josep Vorländer. "Studying Outdoor Impulse Responses at IHTA park: Simulated vs. Measured". In: *Forum Acusticum – 10th Convention of the European Acoustics Association*. Torino, Italy, 2023, pp. 5605–5608. ISBN: 978-88-88942-67-4.

[66] M. Heckl. "Tyre noise generation". In: *Wear* 113.1 (1986), pp. 157–170. DOI: 10.1016/0043-1648(86)90065-7.

[67] M. F. Heidmann. "Interim Prediction Method for Fan and Compressor Source Noise". In: *NASA/TM X-71763* (1979).

[68] M. Helfer. “Umströmungsgeräusche”. In: *Hucho – Aerodynamik des Automobils*. Wiesbaden: Springer Vieweg, 2013. Chap. 8, pp. 523–549. DOI: 10.1007/978-3-8348-2316-8.

[69] G. Herold. “In-flight directivity and sound power measurement of small-scale unmanned aerial systems”. In: *Acta Acustica* 6.58 (2022), p. 13. DOI: 10.1051/aacus/2022052.

[70] K. Heutschi, B. Ott, T. Nussbaumer, and P. Wellig. “Synthesis of real world drone signals based on lab recordings”. In: *Acta Acustica* 4.6 (2020), p. 24. DOI: 10.1051/aacus/2020023.

[71] A. Hoffmann and W. Kropp. “Auralization of simulated tyre noise: Psychoacoustic validation of a combined model”. In: *Applied Acoustics* 145 (2019), pp. 220–227. DOI: 10.1016/j.apacoust.2018.08.026.

[72] M. Hornikx. “Ten questions concerning computational urban acoustics”. In: *Building and Environment* 106 (2016), pp. 409–421. DOI: 10.1016/j.buildenv.2016.06.028.

[73] IEC 61094-8:2012: Measurement microphones – Part 8: Methods for determining the freefield sensitivity of working standard microphones by comparison. International Electrotechnical Commission. Geneva, CH, 2012.

[74] Institute for Hearing Technology and Acoustics, RWTH Aachen University. *Virtual Acoustics – A real-time auralization framework for scientific research*. http://www.virtualacoustics.org.

[75] ISO 10844:2021. Acoustics – Specification of test tracks for measuring sound emitted by road vehicles and their tyres. International Organization for Standardization. Geneva, CH, 2021.

[76] ISO 2533:1975. Standard atmosphere. Geneva, CH, 1975.

[77] I-group Presence Questionnaire (IPQ). Available at https://www.igroup.org/pq/ipq/download.php (Last viewed 12/16/2024).

[78] ISO 3745:2012. Acoustics – Determination of sound power levels and sound energy levels of noise sources using sound pressure – Precision methods for anechoic rooms and hemi-anechoic rooms. International Organization for Standardization. Geneva, CH, 2012.

[79] ISO 3746:2010. Acoustics – Determination of sound power levels and sound energy levels of noise sources using sound pressure – Survey method using an enveloping measurement surface over a reflecting plane. Geneva, CH, 2010.

[80] ISO 3747:2010. Acoustics – Determination of sound power levels and sound energy levels of noise sources using sound pressure – Engineering/survey methods for use in situ in a reverberant environment. International Organization for Standardization. Geneva, CH, 2010.

[81] ISO 9613-1:1993. Acoustics – Attenuation of sound during propagation outdoors – Part 1: Calculation of the absorption of sound by the atmosphere. International Organization for Standardization. Geneva, CH, 1993.

[82] ISO 9613-2:2024. Acoustics – Attenuation of sound during propagation outdoors – Part 2: Engineering method for the prediction of sound pressure levels outdoors. International Organization for Standardization. Geneva, CH, 2024.

[83] J. Jagla. "Dynamic simulation of urban traffic noise". PhD thesis. Grenoble, France: Université Grenoble Alpes, 2013.

[84] J. Jagla, J. Maillard, and N. Martin. "Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method". In: *The Journal of the Acoustical Society of America* 132.5 (2012), pp. 3098–3108. DOI: 10.1121/1.4754663.

[85] L. Jiang, M. Masullo, L. Maffei, F. Meng, and M. Vorländer. "How do shared-street design and traffic restriction improve urban soundscape and human experience? — An online survey with virtual reality". In: *Building and Environment* 143 (2018), pp. 318–328. DOI: 10.1016/j.buildenv.2018.07.005.

[86] R. F. S. Job. "Community response to noise: A review of factors influencing the relationship between noise exposure and reaction". In: *The Journal of the Acoustical Society of America* 83.3 (1988), pp. 991–1001. DOI: 10.1121/1.396524.

[87] H. Jonasson. "Acoustical source modelling of road vehicles". In: *Acta Acustica united with Acustica* 93.2 (2007), pp. 173–184.

[88] A. Kacem. "Auralisation des transports ferroviaires en milieu urbain (Auralisation of rail transport in urban areas)". PhD thesis. Grenoble, France: Université Grenoble Alpes, 2019.

[89] M. Kamrath, P. Jean, J. Maillard, J. Picaut, and C. Langrenne. "Extending standard urban outdoor noise propagation models to complex geometries)". In: *The Journal of the Acoustical Society of America* 143.4 (2018), pp. 2066–2075. DOI: 10.1121/1.5027826.

[90] J. Kang. *Urban Sound Environment*. 1st. London: CRC Press, 2007. DOI: 10.1201/9781482265613.

[91] J. Kang and B. Schulte-Fortkamp. *Soundscape and the Built Environment*. 1st. Boca Raton: CRC Press, 2016. DOI: 10.1201/b19145.

[92] S. Kephalopoulos, M. Paviotti, and F. Anfosso-Lédée. "Common Noise Assessment Methods in Europe (CNOSSOS-EU)". In: *Publications Office of the European Union* (2012), pp. 1–180. DOI: 10.2788/32029.

[93] M. Klemenz. "Die Geräuschqualität bei der Anfahrt elektrischer Schienenfahrzeuge". PhD thesis. Aachen, Germany: RWTH Aachen University, 2005.

[94] S. Koh, B. Zhou, M. Meinke, N. Gauger, and W. Schröder. "Numerical analysis of the impact of variable porosity on trailing-edge noise". In: *Computers and Fluids* 167 (2018), pp. 66–81. DOI: 10.1016/j.compfluid.2018.02.015.

[95] S. Koh, W. Schröder, and M. Meinke. "Turbulence and heat excited noise sources in single and coaxial jets". In: *Journal of Sound and Vibration* 329.7 (2010), pp. 786–803. DOI: 10.1016/j.jsv.2009.10.012.

[96] R. Kouyoumjian and P. Pathak. "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface". In: *Proc. IEEE* 62.11 (1974), pp. 1448–1461.

[97] K. Kryter. "Perceived noisiness (annoyance)". In: *The Effects of Noise on Man.* Academic Press, 1970, pp. 269–331. DOI: 10.1016/B978-0-12-427450-1.50016-9.

[98] H. Kuttruff. *Acoustics: An Introduction.* 1st. London: CRC Press, 2007. DOI: 10.1201/9780367807696.

[99] J. Laufer, R. Schlinker, and R. E. Kaplan. "Experiments on Supersonic Jet Noise". In: *AIAA Journal* 14.4 (1976), pp. 489–497. DOI: 10.2514/3.61388.

[100] D. Lee and H. Seung. "Learning the parts of objects by non-negative matrix factorization". In: *Nature* 401 (1999), pp. 788–791. DOI: 10.1038/44565.

[101] M. Lee and J. S. Bolton. "Source characterization of a subsonic jet by using near-field acoustical holography". In: *The Journal of the Acoustical Society of America* 121.2 (2007), pp. 967–977. DOI: 10.1121/1.2404626.

[102] T. Leishman, S. Rollins, and H. Smith. "An experimental evaluation of regular polyhedron loudspeakers as omnidirectional sources of sound". In: *The Journal of the Acoustical Society of America* 120.3 (2006), pp. 1411–1422. DOI: 10.1121/1.2221552.

[103] T. Li. "Influencing Parameters on Tire–Pavement Interaction Noise: Review, Experiments, and Design Considerations". In: *Designs* 2.4 (2018). DOI: 10.3390/designs2040038.

[104] T. Li, R. Burdisso, and C. Sandu. "Literature review of models on tire-pavement interaction noise". In: *Journal of Sound and Vibration* 420 (2018), pp. 357–445. DOI: 10.1016/j.jsv.2018.01.026.

[105] M. J. Lighthill and M. H. Newman. "On sound generated aerodynamically I. General theory". In: *Proc.Roy.Soc. (A)* 211.1107 (1952), pp. 564–587. DOI: 10.1098/rspa.1952.0060.

[106] G. Lilley. "On the noise from jets". In: *Proc. AGARD Conf.* 131 (1973), pp. 13.1–13.12.

[107] D. Lincke, T. Schumacher, and R. Pieren. "Synthesizing coherence loss by atmospheric turbulence in virtual microphone array signals". In: *The Journal of the Acoustical Society of America* 153.1 (2023), pp. 456–466. DOI: 10.1121/10.0016847.

[108] A. Lindau and S. Weinzierl. "Assessing the Plausibility of Virtual Acoustic Environments". In: *Acta Acustica united with Acustica* 98.5 (2012), pp. 804–810. DOI: 10.3813/AAA.918562.

[109] A. Liutkus and R. Badeau. "Generalized Wiener filtering with fractional power spectrograms". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 2015, pp. 266–270. DOI: 10.1109/ICASSP.2015.7177973.

[110] J. Llorca-Bofí, C. Dreier, J. Heck, and M. Vorländer. "Urban Sound Auralization and Visualization Framework – Case Study at IHTApark". In: *Sustainability* 14 (2022), p. 2026. DOI: 10.3390/su14042026.

[111] J. Llorca-Bofí, J. Heck, C. Dreier, and M. Vorländer. "Urban background sounds under various weather conditions categorized for virtual acoustics". In: *Journal of Environmental Management* 371 (2024), p. 123081. DOI: 10.1016/j.jenvman.2024.123081.

[112] L. V. Lopes and C. L. Burley. "Design of the Next Generation Aircraft Noise Prediction Program: ANOPP2". In: (2011).

[113] M. Lotinga, C. Ramos-Romero, N. Green, and A. Torija. "Noise from Unconventional Aircraft: A Review of Current Measurement Techniques, Psychoacoustics, Metrics and Regulation". In: *Curr. Pollution Rep.* 9 (2023), pp. 724–745. DOI: 10.1007/s40726-023-00285-4.

[114] J. Maillard and J. Jagla. "Effect of load on engine noise for the auralization of road traffic". In: *Euronoise*. 2015, pp. 763–768.

[115] J. Maillard and A. Kacem. "Auralization Applied to the Evaluation of Pedestrian and Bike Paths in Urban Environments". In: *INTER-NOISE: 45th International Congress and Exposition on Noise Control Engineering*. 2016, pp. 2699–2708.

[116] P. Majdak, F. Zotter, F. Brinkmann, J. De Muynke, M. Mihocic, and M. Noisternig. "Spatially Oriented Format for Acoustics 2.1: Introduction and Recent Advances". In: *J. Audio Eng. Soc.* 70.7 (2022), pp. 565–584. DOI: 10.17743/jaes.2022.0026.

[117] J. Maynard, E. Williams, and Y. Lee. "Nearfield Acoustic Holography: I. Theory of generalized holography and the development of NAH". In: *The Journal of the Acoustical Society of America* 78.4 (1985), pp. 1395–1413. DOI: 10.1121/1.392911.

[118] S. Mazza. "Visualizing jet plume noise using Helmholtz equation least squares (HELS) method". In: *BeBeC: Berlin Beamforming Conference*. 2008.

[119] F. Meng. "Modeling of moving sound sources based on array measurements". PhD thesis. Aachen, Germany: RWTH Aachen University, 2018. DOI: 10.30819/4759.

[120] Y. Miki. "Acoustical properties of porous materials-modifications of Delany-Bazley models". In: *Journal of the Acoustical Society of Japan (E)* 11.1 (1990), pp. 19–24. DOI: 10.1250/ast.11.19.

[121] S. Mitra. "Digital Signal Processing: A Computer-based Approach". In: New York: McGraw-Hill, 2001. Chap. 6.4.2 - Digital Filter Structures - Cascade Realizations, pp. 370–372.

[122] E. Mommertz. "Determination of scattering coefficients from the reflection directivity of architectural surfaces". In: *Applied Acoustics* 60.2 (2000), pp. 201–203. DOI: 10.1016/S0003-682X(99)00057-2.

[123] B. Moore, B. Glasberg, and T. Baer. "A model for the prediction of thresholds, loudness, and partial loudness". In: *J. Audio Eng. Soc* 45.4 (1997), pp. 224–240.

[124] M. Müller-Trapet, M. Pollow, and M. Vorländer. "Spherical Harmonics as a Basis for Quantifying Scattering and Diffusing Objects". In: *Proc. Forum Acusticum* (2011), pp. 1851–1856.

[125] P. Mungur, H. Plumblee, and P. Doak. "Analysis of acoustic radiation in a jet flow enviroment". In: *Journal of Sound and Vibration* 36.1 (1974), pp. 21–52. ISSN: 0022-460X. DOI: 10.1016/S0022-460X(74)80341-X.

[126] M. Nilsson, J. Bengtsson, and R. Klaeboe. *Environmental Methods for Transport Noise Reduction*. London: CRC Press, 2014. DOI: 10.1201/b17606.

[127] M. Nilsson, J. Forssén, P. Lundén, A. Peplow, and B. Hellström. *LISTEN: Auralization of Urban Soundscapes*. Tech. rep. Sweden: Final report to the Knowledge Foundation, 2011.

[128] M. Ochmann. "Die Multipolstrahlersynthese – ein effektives Verfahren zur Berechnung der Schallabstrahlung von schwingenden Strukturen beliebiger Oberflächengestalt (Multipole radiator synthesis - an effective method for calculating the sound radiation of vibrating structures with arbitrary surface shape". In: *Acustica* 72 (1990), pp. 233–246.

[129] OpenDAFF. *An open source file format for directional audio content, Institute for Hearing Technology and Acoustics, RWTH Aachen University*. Available at http://www.opendaff.org/, accessed on 2024-06-12.

[130] J. Panda and R. Seasholtz. "Experimental investigation of density fluctuations in high-speed jets and correlation with generated noise". In: *Journal of Fluid Mechanics* 450 (2002), pp. 97–130. DOI: 10.1017/S002211200100622X.

[131] V. Pauz, A. Niemöller, M. Meinke, and W. Schröder. "Numerical analysis of chevron nozzle noise". In: *23rd AIAA/CEAS Aeroacoustics Conference* (2017). DOI: 10.2514/6.2017-3853.

[132] R. Pieren. "Auralization of environmental acoustical sceneries - Synthesis of road traffic, railway and wind turbine noise". PhD thesis. Delft, Netherlands: Delft University of Technology, 2018. DOI: 10.4233/uuid:8dbfb507-a0b0-4ccd-9772-88e213c69206.

[133] R. Pieren, L. Bertsch, D. Lauper, and B. Schäffer. "Improving future low-noise aircraft technologies using experimental perception-based evaluation of synthetic flyovers". In: *Science of The Total Environment* 692 (2019), pp. 68–81. DOI: 10.1016/j.scitotenv.2019.07.253.

[134] R. Pieren, T. Bütler, and K. Heutschi. "Auralization of Accelerating Passenger Cars Using Spectral Modeling Synthesis". In: *Applied Sciences* 6.1 (2016). DOI: 10.3390/app6010005.

[135] R. Pieren, K. Heutschi, J. M. Wunderli, M. Snellen, and D. G. Simons. "Auralization of railway noise: Emission synthesis of rolling and impact noise". In: *Applied Acoustics* 127 (2017), pp. 34–45. DOI: 10.1016/j.apacoust.2017.05.026.

[136] M. Poletti. "Unified Description of Ambisonics using Real and Complex Spherical Harmonics". In: *Proceedings of Ambisonics Symposium*. 2009.

[137] S. Proskurov, R. Ewert, M. Lummer, M. Mößner, and J. Delfs. "Sound shielding simulation by coupled discontinuous Galerkin and fast boundary element methods". In: *Engineering Applications of Computational Fluid Mechanics* 16.1 (2022), pp. 1690–1705. DOI: 10.1080/19942060.2022.2098827.

[138] I. Proudman and G. Taylor. "The generation of noise by isotropic turbulence". In: *Proc.Roy.Soc. (A)* 214.1116 (1952), pp. 119–132. DOI: 10.1098/rspa.1952.0154.

[139] B. Rafaely, B. Weiss, and E. Bachmat. "Spatial Aliasing in Spherical Microphone Arrays". In: *IEEE Transactions on Signal Processing* 55.3 (2007), pp. 1003–1010. DOI: 10.1109/TSP.2006.888896.

[140] H. Raffaseder. *Audiodesign*. 2nd. Munich, Germany: Carl Hanser Verlag, 2010. ISBN: 978-3-446-41762-5.

[141] R. Rehman, C. Dreier, J. Heck, J. Llorca-Bofí, and M. Vorländer. "Comparison of Virtual Reality and Web-Based Listening Experiments on the Perception in Complex Auralized Environments". In: *The Journal of the Acoustical Society of America* 157.3 (2025), pp. 2001–2017. DOI: 10.1121/10.0036147.

[142] G. Rey Gozalo, P. Aumond, and A. Can. "Variability in sound power levels: Implications for static and dynamic traffic models". In: *Transportation Research Part D: Transport and Environment* 84 (2020), p. 102339. DOI: 10.1016/j.trd.2020.102339.

[143] K. Risse, K. Schäfer, F. Schültke, and E. Stumpf. "Central Reference Aircraft data System (CeRAS) for research community". In: *CEAS Aeronautical Journal* 7.1 (2016), pp. 121–133. DOI: 10.1007/s13272-015-0177-9.

[144] S. A. Rizzi, I. LeGriffon, R. Pieren, and L. Bertsch. "A Comparison of Aircraft Flyover Auralizations by the Aircraft Noise Simulation Working Group". In: *AIAA aviation forum*. 2020. DOI: 10.2514/6.2020-2582.

[145] D. Rocchesso. "Fractionally addressed delay lines". In: *IEEE Transactions on Speech and Audio Processing* 8.6 (2000), pp. 717–727. DOI: 10.1109/89.876310.

[146] A. Sahai. "Consideration of Aircraft Noise Annoyance during Conceptual Aircraft Design". Ph.D. Dissertation. RWTH Aachen University, Germany: Institute of Aerospace Systems, 2016.

[147] A. Sahai, E. Anton, E. Stumpf, F. Wefers, and M. Vorländer. "Interdisciplinary auralization of take-off and landing procedures for subjective assessment in virtual reality environments". In: *18th AIAA/CEAS Aeroacoustics Conference*. 2012. DOI: 10.2514/6.2012-2077.

[148] U. Sandberg. "Noise trailers of the world – Tools for tire/ road noise measurements with the close-proximity method". In: *NOISE-CON: National Conference on Noise Control Engineering*. 1998, pp. 1–8.

[149] L. Sanders, P. Malbequi, and I. LeGriffon. "Capabilities of IESTA-CARMEN to predict aircraft noise". In: *ICSV 2016: 23rd International Congress on Sound & Vibration*. 2016.

[150] L. Savioja and P. Svensson. "Overview of geometrical room acoustic modeling techniques". In: *The Journal of the Acoustical Society of America* 138.2 (2015), pp. 708–730. DOI: 10.1121/1.4926438.

[151] P. Schäfer, L. Reich, and M. Vorländer. "Linking atmospheric and urban auralization models". In: *Acta Acustica* 6.28 (2022). DOI: 10.1051/aacus/2022021.

[152] P. Schäfer and M. Vorländer. "Atmospheric Ray Tracing: An efficient, open-source framework for finding eigenrays in a stratified, moving medium". In: *Acta Acustica* 5.26 (2021). DOI: 10.1051/aacus/2021018.

[153] B. Schäffer, R. Pieren, K. Heutschi, J. M. Wunderli, and S. Becker. "Drone Noise Emission Characteristics and Noise Effects on Humans-A Systematic Review". In: *International Journal of Environmental Research and Public Health* 18.11 (2021). DOI: 10.3390/ijerph18115940.

[154] Z. Schärer and A. Lindau. "Evaluation of Equalization Methods for Binaural Signals". In: *Proc. Audio Engineering Society Convention 126*. Munich, Germany, May 2009, pp. 1–17.

[155] L. Schneiders, C. Günther, M. Meinke, and W. Schröder. "An efficient conservative cut-cell method for rigid bodies interacting with viscous compressible flows". In: *Journal of Computational Physics* 311 (2016), pp. 62–86. DOI: 10.1016/j.jcp.2016.01.026.

[156] D. Schröder. "Physically Based Real-Time Auralization of Interactive Virtual Environments". PhD thesis. Aachen, Germany: RWTH Aachen University, 2011. ISBN: 978-3-8325-3031-0.

[157] B. Schulte-Fortkamp, A. Fiebig, J. Sisneros, A. Popper, and R. Fay. *Soundscapes: Humans and Their Acoustic Environment*. 1st. Springer Handbook of Auditory Research. Boca Raton, 2023. DOI: 10.1007/978-3-031-22779-0.

[158] N. R. Shabtai, G. Behler, M. Vorländer, and S. Weinzierl. "Generation and analysis of an acoustic radiation pattern database for forty-one musical instruments". In: *The Journal of the Acoustical Society of America* 141.2 (2017), pp. 1246–1256. DOI: 10.1121/1.4976071.

[159] J. O. Smith. "Physical modeling using digital waveguides". In: *Computer Music Journal* 16.4 (1992), pp. 74–91.

[160] J. O. Smith. "Digital Waveguide Architectures for Virtual Musical Instruments". In: *Handbook of Signal Processing in Acoustics*. Ed. by D. Havelock, S. Kuwano, and M. Vorländer. New York, NY: Springer New York, 2008, pp. 399–417. ISBN: 978-0-387-30441-0. DOI: 10.1007/978-0-387-30441-0_25.

[161] P. Spalart, M. Shur, and M. Strelets. "Identification of Sound Sources in Large-Eddy Simulations of Jets". In: *13th AIAA/CEAS Aeroacoustics Conference (28th AIAA Aeroacoustics Conference)* 7.1 (2008), pp. 1–25. DOI: 10.2514/6.2007-3616.

[162] D. Stanescu, M. Hussaini, and F. Farassat. "Aircraft engine noise scattering by fuselage and wings: A computational approach". In: *Journal of Sound and Vibration* 263.2 (2003), pp. 319–333. DOI: 10.1016/S0022-460X(02)01126-4.

[163] J. Stienen. "Real-Time Auralisation of Outdoor Sound Propagation". PhD thesis. Aachen, Germany: RWTH Aachen University, 2022. DOI: 10.30819/5629.

[164] J. R. Stone and F. J. Montegani. "An improved prediction method for the noise generated in flight by circular jets". In: *NASA/TM 81470* (1980).

[165] H. Tachibana, T. Iwase, and K. Ishii. "Sound power levels of road vehicles measured by a new method using a reverberant tunnel". In: *Journal of the Acoustical Society of Japan (E)* 2.2 (1981), pp. 117–125. DOI: 10.1250/ast.2.117.

[166] C. Tam. "Jet noise: Since 1952". In: *Theor. Comput. Fluid Dyn.* 10.1 (1998), pp. 393–405. DOI: 10.1007/s001620050072.

[167] C. Tam, K. Viswanathan, K. Ahuja, and J. Panda. "The sources of jet noise: Experimental evidence". In: *Journal of Fluid Mechanics* 615 (2008), pp. 253–292. DOI: 10.1017/S0022112008003704.

[168] H. Tanna and P. Morris. "The noise from normal-velocity-profile coannular jets". In: *Journal of Sound and Vibration* 98.2 (1985), pp. 213–234. DOI: 10.1016/0022-460X(85)90386-4.

[169] U. Tengzelius, F. Bahmani, and I. Karasalo. "SAFT – Simulation of atmosphere and air traffic for a quieter environment". In: *Euronoise - 11th European Congress and Exposition on Noise Control Engineering*. 2018.

[170] B. Tester and C. Morfey. "Developments in jet noise modelling – Theoretical predictions and comparisons with measured data". In: *Journal of Sound and Vibration* 46.1 (1976), pp. 79–103. DOI: 10.1016/0022-460X(76)90819-1.

[171] D. Thompson, B. Hemsworth, and N. Vincent. "Experimental validation of the TWINS prediction program for rolling noise, Part 1: Description of the model and method". In: *Journal of Sound and Vibration* 193.1 (1996), pp. 123–135. DOI: 10.1006/jsvi.1996.0252.

[172] K. Tsukui, Y. Oshino, and H. Tachibana. "Simulation model for road traffic noise taking account of vehicle noise directivity". In: *The Journal of the INCE of Japan* 22.2 (1998), pp. 108–116. DOI: 10.11372/souonseigyo1977.22.108.

[173] S. Ullrich. *Rollgeräuschmessungen der Bundesanstalt für Strassenwesen an Strassen mit unterschiedlichen Decken und an einem Innentrommelprüfstand (Rolling noise measurements of the Federal Highway Research Institute on roads with different ceilings and an indoor drum)*. Mitteilung Nr.57. Institut für Strassen- und Eisenbahn- und Felsbau and der ETH Zürich, 1984, pp. 1–207.

[174] T. Van Renterghem, E. Salomons, and D. Botteldooren. "Parameter study of sound propagation between city canyons with a coupled FDTD-PE model". In: *Applied Acoustics* 67.6 (2006), pp. 487–510. DOI: 10.1016/j.apacoust.2005.09.006.

[175] T. Van Renterghem, M. Hornikx, J. Forssen, and D. Botteldooren. "The potential of building envelope greening to achieve quietness". In: *Building and Environment* 61 (2013), pp. 34–44. DOI: 10.1016/j.buildenv.2012.12.001.

[176] V. Vázquez, F. Terán, J. Luong, and S. Paje. "Functional Performance of Stone Mastic Asphalt Pavements in Spain: Acoustic Assessment". In: *Coatings* 9.2 (2019). ISSN: 2079-6412. DOI: 10.3390/coatings9020123.

[177] M. Vorländer. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. 2nd. Springer Nature Switzerland AG, 2020. DOI: 10.1007/978-3-030-51202-6.

[178] M. Vorländer and E. Mommertz. "Definition and measurement of random-incidence scattering coefficients". In: *Applied Acoustics* 60.2 (2000), pp. 187–199. DOI: 10.1016/S0003-682X(99)00056-0.

[179] Z. Wang and S. F. Wu. "Helmholtz equation–least-squares (HELS) method for reconstructing the acoustic pressure field". In: *The Journal of the Acoustical Society of America* 102.4 (1997), pp. 2020–2032. DOI: 10.1121/1.419691.

[180] P. Welch. "The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms". In: *IEEE Transactions on Audio and Electroacoustics* 15.2 (1967), pp. 70–73. DOI: 10.1109/TAU.1967.1161901.

[181] E. G. Williams. *Fourier Acoustics*. New York: Academic Press, 1999.

[182] I. Witew. "Measurements in room acoustics. Uncertainties and influence of the measurement position". PhD thesis. Aachen, Germany: RWTH Aachen University, 2022. DOI: 10.30819/5529.

[183] R. S. Womersley and I. H. Sloan. "How good can polynomial interpolation in the sphere be?" In: *Advances in Computational Mathematics* 14 (2001), pp. 195–226. DOI: 10.1023/A:1016630227163.

[184] J. M. Wunderli, J. Meister, O. Boolakee, and K. Heutschi. "A Method to Measure and Model Acoustic Emissions of Multicopters". In: *International Journal of Environmental Research and Public Health* 20.1 (2023). DOI: 10.3390/ijerph20010096.

[185] C. Xu, T. Oberman, F. Aletta, H. Tong, and J. Kang. "Ecological Validity of Immersive Virtual Reality (IVR) Techniques for the Perception of Urban Sound Environments". In: *Acoustics* 3 (2020), pp. 11–24. DOI: 10.3390/acoustics3010003.

[186] M. Yang, A. Heimes, M. Vorländer, and B. Schulte-Fortkamp. "Comparison of subjective evaluations in virtual and real environments for soundscape research". In: *The Journal of the Acoustical Society of America* 155.6 (2024), pp. 3715–3729. DOI: 10.1121/10.0026123.

[187] N. Zawodny, A. Christian, and R. Cabell. "A summary of NASA research exploring the acoustics of small unmanned aerial systems". In: *Proc. AHS Int. Tech. Meeting on Aeromech. Des. for Trans. Vert. Fl.* San Francisco, 2018, pp. 1–11.

[188] U. Zölzer. *DAFX: Digital Audio Effects*. 2nd. Chichester, United Kingdom: John Wiley & Sons, 2011. DOI: 10.1002/9781119991298.

[189] F. Zotter. "Analysis and Synthesis of Sound-Radiation with Spherical Arrays". Ph.D. Dissertation. University of Music and Performing Arts, Austria: Institute of Electronic Music and Acoustics, 2009.

[190] F. Zotter and A. Staackmann. “Simulation of room transfer functions with directivity patterns on the basis of modes”. In: *Proc. DAGA - Fortschritte der Akustik* (2017), pp. 1430–1433.

Curriculum Vitae

Christian Dreier

1993	born in Kamp-Lintfort, Germany

Education

28.07.2025	Doctoral examination for Dr.-Ing.
2016–2018	M.Sc. Electrical Engineering, Information Technology and Computer Engineering RWTH Aachen University
2013–2016	B.Sc. Electrical Engineering, Information Technology and Computer Engineering RWTH Aachen University
2012–2013	Electrical Engineering and Information Technology Ilmenau University of Technology
2003 – 2012	General Higher Education Entrance Qualification (Abitur) Gymnasium Adolfinum, Moers

Awards

2012	Federal winner at Jugend forscht Special prize by the Chancellor of Germany for the most original work

Professional Experience

Since 2025	R&D Systems Engineer, Siemens Mobility
2018 – 2024	Research Assistant, Institute for Hearing Technology and Acoustics, RWTH Aachen University
2017 – 2018	Graduate Acoustics Engineer, Dr. Ing. h.c. F. Porsche AG
2016 – 2017	Student Research Assistant, HEAD acoustics GmbH

August 2, 2025

Acknowledgements

This work was made possible through the collaborative efforts of many minds, hearts and hands, and I would like to take this opportunity to express my heartfelt gratitude.

First and foremost, I thank Professor Michael Vorländer for not only sharing his extraordinary talent as an academic teacher during my engineering studies but also for his open-minded, inspiring approach that encouraged me to explore ideas in depth and planted the seeds for new visions.

My thanks also go to Professor Peter Urban for co-supervising this dissertation and offering invaluable feedback.

I am grateful to the administrative and technical teams at IHTA—Emilie Koch, Kim König, Karin Charlier and Ellen Vergöls—for their unfailing support. Special thanks to Marc Eiker and Thomas Schäfer, who built the microphone array and supported me during the measurement campaigns. Without Rolf Kaldenbach and Norbert Konkol, all the calculations in this thesis would have had to be done with an abacus.

I would like to acknowledge the PAAD main station and satellite working group members—Jonas Heck, Dr Josep Llorca-Bofí, Dr Philipp Schäfer and Rouben Rehman—and to thank Dr Ingo Witew for his boundless scientific curiosity and brilliant ideas, which elevated the quality of our work.

I also appreciate the collaborations with all our project partners as well as student assistants—Simon Kamps, Johannes Hahn, Lars Heinrichs and Junjie He—for their hard work and dedication.

My sincere thanks go out to all IHTAliens whose collaboration transformed this research into a truly memorable, enjoyable and enlightening journey.

Finally, and with special emphasis, I want to thank my family—my parents, Caro, my parents- and family-in-law, and above all, Jana and Lauri—for their unwavering support, wise counsel, hugs and love. I am infinitely grateful to each of you!