

**Auralization of Environmental Acoustical Sceneries  
Synthesis of Road Traffic, Railway and Wind Turbine Noise**

Pieren, Reto

**DOI**

[10.4233/uuid:8dbfb507-a0b0-4ccd-9772-88e213c69206](https://doi.org/10.4233/uuid:8dbfb507-a0b0-4ccd-9772-88e213c69206)

**Publication date**

2018

**Document Version**

Final published version

**Citation (APA)**

Pieren, R. (2018). *Auralization of Environmental Acoustical Sceneries: Synthesis of Road Traffic, Railway and Wind Turbine Noise*. [Dissertation (TU Delft), Delft University of Technology].  
<https://doi.org/10.4233/uuid:8dbfb507-a0b0-4ccd-9772-88e213c69206>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# **Auralization of Environmental Acoustical Sceneries**

Synthesis of Road Traffic, Railway and  
Wind Turbine Noise



# **Auralization of Environmental Acoustical Sceneries**

Synthesis of Road Traffic, Railway and  
Wind Turbine Noise

## **Dissertation**

for the purpose of obtaining the degree of doctor  
at Delft University of Technology,  
by the authority of the Rector Magnificus prof.dr.ir. T.H.J.J. van der Hagen,  
chair of the Board for Doctorates,  
to be defended publicly on  
Thursday 6 September 2018 at 15:00 o'clock

by

**Reto PIEREN**

Master of Science in Electrical Engineering and Information Technology,  
ETH Zurich, Switzerland  
born in Basel, Switzerland.

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. D. G. Simons	Delft University of Technology, promotor
Dr. ir. M. Snellen	Delft University of Technology, copromotor

Independent members:

Prof. Dr. rer. nat. M. Vorländer	RWTH Aachen University, Germany
Prof. dr. ir. L. L. M. Veldhuis	Delft University of Technology
Prof. dr. W. B. Kleijn	Delft University of Technology
Prof. dr. ir. C. P. A. Wapenaar	Delft University of Technology

Other members:

Dr. K. Heutschi	Empa, Switzerland
-----------------	-------------------

Dr. Kurt Heutschi from Empa has contributed greatly to the preparation of this dissertation.



*Keywords:* Acoustics, Environmental pollution, Simulation, Noise

*Printed by:* ADAG Print Zürich, Switzerland

*Front & Back:* Anja Lehmann Gestaltung, [anjalehmann.ch](http://anjalehmann.ch)

Copyright © 2018 by R. Pieren

An electronic version of this dissertation is available at

<http://repository.tudelft.nl/>.

# Summary

Environmental noise caused by traffic and industrial facilities is a relevant health factor in urban areas, along major traffic routes and in the vicinity of airports. The technique of auralization helps exploring the relationships between sound and its impact on people under fully controlled conditions. Further, by directly addressing the hearing sensation, it is an intuitive tool for the assessment and communication of environmental noise scenarios. Analogous to visualization, auralization enables us to listen to situations that do not exist in reality.

In this thesis research, methods for the auralization of environmental acoustical sceneries are established. The sceneries are represented by a virtual environment containing virtual sound sources that are arranged in space and time, and within which sound waves propagate to a virtual observer. To that aim, sophisticated calculation models for the synthesis and reproduction of road traffic, railway and wind turbine noise are developed. This requires investigating the relevance of the involved acoustical phenomena for perceived realism. On that basis, calculation models (i.e. synthesizer structures) are proposed that adequately reproduce source characteristics, sound propagation effects, and spatial impression. The models are accompanied by methods to derive the necessary input parameters from own measurements and data analysis.

The presented calculation models are parametric and thus allow for a large versatility with respect to scenarios and sound reproduction. Because the three considered environmental noise sources feature their specific acoustical peculiarities, source-specific models are proposed. These source-specific models have in common that the sound radiated by a source is artificially generated using digital sound synthesis. For wind turbine and road traffic noise, a combination of additive and subtractive synthesis, denoted as spectral modeling synthesis, is applied. A uniqueness of the wind turbine synthesizer is the ability to reproduce and control different types of characteristic amplitude modulation. The synthesizer for road vehicles separately produces tire noise and propulsion noise. The generated propulsion sounds depend on the engine type, the instantaneous engine condition (engine speed and load), and the emission angle. An additional special feature of the propulsion sound synthesis is the fact that, besides amplitude and frequency, the phase of the engine harmonics has to and is considered.

For railway rolling and impact noise, in contrast, a physically-based synthesis approach has been developed that describes the mechanical excitation and the vibration of the dynamic wheel/rail system. The corresponding model considers the microstructure of the wheels and rails, as well as structural resonances of the wheel/rail system to elicit the typical metallic sound character of railway noise. In all models, sound propagation effects, such as geometrical divergence, Doppler effect, atmospheric absorption, ground effect and amplitude fluctuations due to

atmospheric turbulence, from a virtual point source to a virtual observer location are simulated by processing the synthetic source signals with time-variant filters in the time domain.

Auralizations created with the presented models feature a high audio quality and are judged as plausible and realistic by expert listeners. To achieve this realism in the auralizations, it was found that variation with respect to time, frequency, space, and orientation is crucial.

The presented models extend the today's body of existing auralization models and allow for new possible applications.

# Samenvatting

Omgevingsgeluid, veroorzaakt door wegverkeer en industrie, is een relevante gezondheidsfactor in de stedelijke omgeving, langs drukke transport routes en in de directe nabijheid van luchthavens. De auralisatie techniek helpt bij het onderzoeken van de relatie tussen geluid en de impact hiervan op mensen onder volledig gecontroleerde omstandigheden. Daarnaast is auralisatie een intuïtief hulpmiddel voor het beoordelen en communiceren van omgevingsgeluid scenario's, doordat er een directe auditieve gewaarwording wordt gecreëerd. Analoog aan visualisatie, wordt er bij auralisatie een situatie hoorbaar gemaakt die nog niet in het echt hoeft te bestaan.

In deze thesis worden methodes gepresenteerd waarmee akoestische scenario's voor omgevingsgeluid worden gecreëerd. Deze scenario's worden gerepresenteerd door een virtuele omgeving waarin virtuele geluidsbronnen in ruimte en tijd worden gedefinieerd en waarin geluidsgolven van deze bronnen naar een virtuele ontvanger propageren. Om dit te bereiken, zijn geavanceerde rekenmodellen ontwikkeld voor de synthese en reproductie van het geluid van wegverkeer, treinverkeer en windturbines. Een zo realistisch mogelijk waargenomen geluid vergt onderzoek naar de relevantie van de verschillende akoestische fenomenen. Op basis hiervan worden rekenmodellen (d.w.z. de structuur van de geluidssynthese) gepresenteerd die op adequate wijze de geluidskarakteristieken van de bron, de geluidspropagatie effecten en de ruimtelijke indruk weergeven. Deze rekenmodellen worden vergezeld van methodes om de nodige invoergegevens af te leiden uit eigen metingen en de analyse van andere beschikbare data.

De gepresenteerde rekenmodellen zijn parametrisch en hierdoor zeer veelzijdig in het modelleren van veel verschillende scenario's en geluidsreproductie. De drie omgevingsgeluidsbronnen die in deze thesis worden beschouwd hebben allemaal hun eigen akoestische bijzonderheden. Daarom worden er bron-specifieke modellen gepresenteerd met als overeenkomst dat het geluid wat de bron uitstraalt, kunstmatig wordt gegenereerd door middel van digitale geluidssynthese. Voor windturbine- en wegverkeergeluid wordt een combinatie van additieve en subtractieve synthese toegepast, wat ook wel wordt aangeduid als de spectrale modeleringssynthese. Een unieke eigenschap van de geluidssynthese voor windturbines is de mogelijkheid om verschillende types van de karakteristieke amplitudemodulaties te reproduceren en te beheersen. Bij de synthese van wegverkeergeluid worden band- en motorgeluid apart gegenereerd. Het gegenereerde motorgeluid is afhankelijk van het motortype, de momentane motortoestand (motorsnelheid en belasting) en de uitstralingshoek van het geluid. Een bijkomende speciale eigenschap van de synthese van motorgeluid is dat, behalve amplitude en frequentie, de fase van de verschillende harmonischen van het geluid wordt meegenomen.

Voor rol- en impactgeluid van treinverkeer is een fysisch-realistische geluidsynthese ontwikkeld waarbij de mechanische excitatie en de vibraties van het dynamische wiel-rail systeem worden beschreven. Het bijbehorende model beschouwt zowel de microstructuren van de wielen en de rails, alsook de structurele resonanties van het wiel-rail systeem om hiermee het typisch metaalachtige karakter van spoorweggeluid op te wekken.

In alle modellen worden geluidsvoortplanting-effecten tussen de virtuele puntbron en de virtuele ontvanger gesimuleerd door het gesynthetiseerde bronsignaal te verwerken in het tijddomein met een tijdvariabel filter. Geluidsvoortplanting-effecten die worden gesimuleerd zijn geometrische spreiding, Doppler effect, atmosferische absorptie, grondreflectie-effect en amplitudefluctuaties als gevolg van atmosferische turbulentie.

Auralisaties gecreëerd met de gepresenteerde modellen hebben een hoge audiokwaliteit en worden door experts beoordeeld als plausibel en realistisch. Om dit realisme in auralisatie te behalen is gebleken dat variaties in tijd, frequentie, ruimte en oriëntatie cruciaal zijn.

De hier gepresenteerde modellen zijn een uitbreiding en aanvulling op de al bestaande auralisatie modellen en bieden nieuwe toepassingsmogelijkheden.

# Contents

<b>Summary</b>	<b>v</b>
<b>Samenvatting</b>	<b>vii</b>
<b>Nomenclature</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and motivation . . . . .	2
1.2 What is auralization? . . . . .	2
1.3 Auralization in environmental acoustics . . . . .	3
1.3.1 Overview . . . . .	3
1.3.2 Application examples . . . . .	4
1.4 Research objectives . . . . .	5
1.5 Thesis outline . . . . .	6
References . . . . .	6
<b>2 Methodology</b>	<b>11</b>
2.1 Auralization vs. classical noise calculation . . . . .	12
2.1.1 Frequency resolution . . . . .	12
2.1.2 Time resolution - from days to milliseconds . . . . .	13
2.1.3 Phase matters . . . . .	13
2.1.4 Directional information . . . . .	14
2.2 Auralization systems . . . . .	14
2.2.1 Overview . . . . .	14
2.2.2 Sound generation approaches . . . . .	15
2.2.3 Sound recording . . . . .	16
2.2.4 Parametric sound synthesis . . . . .	17
2.2.5 Propagation filtering . . . . .	18
2.2.6 Sound scene description . . . . .	18
2.2.7 Sound reproduction systems . . . . .	19
2.2.8 Loudspeaker reproduction . . . . .	19
2.2.9 Headphone reproduction . . . . .	24
2.2.10 System evaluation criteria . . . . .	25
2.2.11 Loudspeaker vs. headphone reproduction . . . . .	27
2.2.12 Application-specific reproduction system . . . . .	27
2.3 Auralization model structure . . . . .	28
2.3.1 Separation of emission and propagation . . . . .	28
2.3.2 Object-based modeling approach . . . . .	29
2.3.3 Emission synthesizer . . . . .	30
2.3.4 Propagation filter . . . . .	31

2.4	Development strategy for auralization models . . . . .	32
	References . . . . .	33
<b>3</b>	<b>Auralization of Wind Turbine Noise</b>	<b>39</b>
3.1	Introduction . . . . .	40
3.2	Emission synthesis. . . . .	40
3.2.1	Synthesizer structure . . . . .	41
3.2.2	Signal Analysis . . . . .	44
3.2.3	Application to measurements . . . . .	48
3.2.4	Discussion. . . . .	53
3.3	Propagation filtering . . . . .	55
3.3.1	Propagation effects and representation by digital filters . . . . .	55
3.3.2	Doppler frequency shift due to moving receiver . . . . .	56
3.3.3	Geometrical spreading. . . . .	57
3.3.4	Ground effect . . . . .	57
3.3.5	Air absorption. . . . .	60
3.3.6	Attenuation by barriers and foliage . . . . .	61
3.3.7	Additional reflections . . . . .	62
3.3.8	Atmospheric fluctuations . . . . .	62
3.4	Conclusion . . . . .	73
	References. . . . .	74
<b>4</b>	<b>Auralization of Road Traffic - Accelerating Passenger Cars</b>	<b>79</b>
4.1	Introduction . . . . .	80
4.2	Model development. . . . .	81
4.2.1	Overview . . . . .	81
4.2.2	Emission module . . . . .	83
4.2.3	Propagation filtering . . . . .	89
4.2.4	Reproduction rendering . . . . .	95
4.3	Model parameter estimation . . . . .	95
4.3.1	Tire noise . . . . .	95
4.3.2	Propulsion noise . . . . .	96
4.4	Evaluation . . . . .	108
4.5	Conclusions . . . . .	110
	References. . . . .	110
<b>5</b>	<b>Auralization of Railway Noise: Rolling and Impact Noise</b>	<b>115</b>
5.1	Introduction . . . . .	116
5.2	Measurements . . . . .	117
5.3	Model. . . . .	117
5.3.1	Point source model. . . . .	117
5.3.2	Synthesis approach . . . . .	119
5.3.3	Emission synthesis of rolling and impact noise . . . . .	120
5.3.4	Evaluation . . . . .	140
5.4	Conclusions . . . . .	141
	References. . . . .	142

---

<b>6 Conclusion and outlook</b>	<b>147</b>
6.1 Conclusions . . . . .	147
6.2 Future work . . . . .	148
References . . . . .	150
<b>Acknowledgements</b>	<b>151</b>
<b>A Supplementary audio and video files</b>	<b>153</b>
<b>Curriculum Vitæ</b>	<b>155</b>
<b>List of Publications</b>	<b>157</b>



# Nomenclature

$\alpha$	decay rate in $s^{-1}$
$\alpha_i$	experimental tire noise regression parameter in dB
$\alpha_{\text{road}}$	road inclination angle in radians
$\bar{e}$	mean equivalent mass factor
$\beta$	experimental tire noise regression parameter in dB
$\beta_0$	initial angle in degrees
$\chi$	log-amplitude fluctuation
$\Delta f$	frequency resolution in Hertz
$\Delta k$	wavenumber resolution in radians per meter
$\Delta L_{\text{dir}}$	directivity in dB
$\Delta L_{\text{road}}$	road surface correction in dB
$\Delta t$	time difference in seconds
$\Delta$	difference
$\delta$	Kronecker delta function
$\eta$	efficiency factor
$\eta$	random fluctuation function
$\eta$	relative error
$\Gamma$	load in percent
$\gamma$	arbitrary angle in radians
$\kappa$	time lag in seconds
$\lambda$	wavelength in meters
$\mathcal{L}$	Laplace transform
$\mathcal{D}$	Doppler factor
$\mathcal{F}$	propagation filter function

---

$\mu$	fluctuation of acoustical refractive index
$\nu$	bandpass filtered stationary noise
$\omega$	angular frequency in radians per second
$\bar{L}$	arithmetic mean sound pressure level in dB
$\phi$	phase in radians
$\phi$	sound pressure phase with turbulence
$\phi_0$	sound pressure phase without turbulence
$\Pi$	signal period in samples
$\Psi$	random variable with standard normal distribution
$\rho$	spatial separation in meters
$\rho_{ij}$	correlation coefficient
$\sigma$	standard deviation
$\sigma_T$	standard deviation of temperature fluctuations in Kelvin
$\sigma_v$	standard deviation of wind speed fluctuations in m/s
$\tau$	time constant in seconds
Cov	Covariance
imp	impulse
tr	track
Var	Variance
veh	vehicle
$\Theta$	Heaviside function
$\theta$	immission angle in radians
$\varphi$	emission angle in radians
$\vec{s}$	position vector
$\xi$	spatial frequency in $\text{m}^{-1}$
$\xi_s$	spatial sampling frequency in $\text{m}^{-1}$
$\zeta$	damping ratio
$\zeta$	deflection in meters

$A$	signal amplitude
$A$	sound pressure amplitude with turbulence
$a$	acceleration in $\text{m/s}^2$
$a$	correlation length of turbulent refractive index fluctuations in meters
$A_0$	sound pressure amplitude without turbulence
$A_3$	contact filter function in dB
$A_i$	experimental tire noise regression parameter in dB
$A_{\text{atm}}$	attenuation due to atmospheric absorption in dB
$A_{\text{gr}}$	ground effect in dB
$A_{\text{room}}$	room correction in dB
$B$	bandwidth in Hertz
$B$	spatial turbulence correlation function
$B_i$	experimental tire noise regression parameter in dB
$C$	resampling conversion ratio
$c$	speed of sound in still air in $\text{m/s}$
$c_0$	speed of sound in $\text{m/s}$
$c_{\text{eff}}$	effective speed of sound in $\text{m/s}$
$C_{\text{notch}}$	level correction in dB
$C_{\text{off}}$	offset correction in dB
$D$	distance in meters
$d$	distance in meters
$d_0$	reference distance equal to 1 meter
$d_{\text{Ac}}$	distance to acoustical center in meters
$d_{\text{dir}}$	source–receiver distance in meters
$d_{\text{gr}}$	source–ground reflection point–receiver distance in meters
$F$	level fluctuation function in dB
$F$	spatial turbulence spectral density function
$f$	frequency in Hertz

$F_0 - F_2$	coast down parameters
$F_B$	driving resistance in Newton
$f_c$	center frequency in Hertz
$f_s$	audio sampling frequency in Hertz
$f_{\text{ign}}$	ignition frequency in Hertz
$F_{\text{periodic}}$	periodic level fluctuation function in dB
$F_{\text{stochastic}}$	stochastic level fluctuation function in dB
$F_{\text{turb}}$	level fluctuation function in dB
$F_T$	traction in Newton
$f_t$	tone frequency in Hertz
$f_{BP}$	blade passing frequency in Hertz
$f_{c,m}$	modulation 1/3 octave band center frequency in Hertz
$f_o$	engine order frequency in Hertz
$g$	gravity of Earth in $\text{m/s}^2$
$G_{\text{mot}}$	Doppler and convective amplification in dB
$H$	filter frequency response
$h$	filter impulse response
$H_{\text{air}}$	frequency response of air absorption filter
$h_{\text{air}}$	air absorption filter
$H_{\text{hub}}$	hub height in meters
$h_{\text{swrc}}$	spherical wave reflection coefficient filter
$i$	index
$i_{\text{ax}}$	axle ratio
$i_{\text{gear}}$	gear ratio
$j$	imaginary unit
$j$	index
$k$	sample index
$k$	wavenumber in radians per meter

$k_s$	non-integer sample index
$k_x$	wavenumber in $x$ direction in radians per meter
$k_z$	wavenumber in vertical direction in radians per meter
$L$	sound pressure level in dB relative to $p_0$
$l$	index
$L_r$	roughness level in dB relative to $r_0$
$L_e$	emission level in dB
$L_{\text{meas}}$	measured sound pressure level in dB
$L_t$	tone sound pressure level in dB
$M$	Mach number
$M$	torque in Nm
$m$	index
$m$	mass in kg
$M_{\text{max}}$	maximal torque in Nm
$N$	number
$n$	FFT frequency bin
$n$	acoustical refractive index
$n_{\text{rot}}$	rotational speed in revolutions per minute
$o$	engine order
$P$	perimeter in meters
$p$	sound pressure in pascals
$p_0$	reference pressure equal to $20 \mu\text{Pa}$
$P_u$	signal power of signal $u$
$p_e$	sound pressure source signal in pascals
$Q$	spherical wave reflection coefficient
$R$	cross-correlation function
$r$	roughness in meters
$R^2$	coefficient of determination

---

$r_0$	reference roughness equal to $1 \mu\text{m}$
$R_p$	plane wave reflection coefficient
$r_{\text{tire,dyn}}$	dynamic tire radius
$S$	phase fluctuation
$s$	complex variable (frequency) of the Laplace domain
$s_1, s_2$	sound pressure source signals in Pascals
$s_{\text{periodic}}$	standard deviation of periodic level fluctuations in dB
$s_{\text{stochastic}}$	standard deviation of stochastic level fluctuations in dB
$T$	duration in seconds
$T$	momentary temperature in Kelvin
$t$	time in seconds
$T_0$	average temperature in Kelvin
$t_0$	start time in seconds
$T_{60}$	reverberation time in seconds
$u$	index
$u$	prototype function
$V$	vehicle speed in km/h
$v$	momentary wind speed in m/s
$v$	roughness velocity in m/s
$V_0$	reference vehicle speed equal to 70 km/h
$v_0$	average wind speed in m/s
$W$	periodic fluctuation function
$w$	white noise
$X$	Cartesian coordinate along the rail in meters
$x$	Cartesian coordinate in meters
$x$	sample index of spatial roughness signal
$y$	Cartesian coordinate in meters
$z$	Cartesian coordinate in meters

- 
- $z$  path length difference between direct and diffracted sound in meters
- A frequency weighting 'A' according to IEC 61672-1
- ACF autocorrelation function
- AM amplitude modulation
- ARMA autoregressive–moving-average
- BP band-pass
- BPF blade passing frequency
- C Celcius
- CCF cross-correlation function
- cf. confer
- dB decibel
- DC direct current
- DESTINATE Decision supporting tools for implementation of cost-efficient railway noise abatement measures
- DFT discrete Fourier transform
- Empa Swiss Federal Laboratories for Materials Science and Technology
- etc. et cetera
- FDTD finite difference time domain
- FFT fast Fourier transform
- FIR finite impulse response
- HRTF head-related transfer function
- Hz Hertz
- i.e. id est
- IDFT inverse discrete Fourier transform
- IFFT inverse fast Fourier transform
- IIR infinite impulse response
- K Kelvin
- Leq equivalent continuous level

---

MAE	mean absolute error
MDAP	Multiple-Direction Amplitude Panning
NCO	numerically controlled oscillator
ORTF	Office de Radiodiffusion Télévision Française
PA	phase accumulator
PAC	phase-to-amplitude converter
PSD	power spectral density
PSOLA	Pitch Synchronous Overlap and Add
rpm	revolutions per minute
S	point source
SBB	Swiss Federal Railways (German: Schweizerische Bundesbahnen)
SMS	Spectral Modeling Synthesis
SNSF	Swiss National Science Foundation
SPL	sound pressure level
STFT	short-time Fourier transform
TAURA	Traffic Noise Auralisator
TU	Technical University
VBAP	Vector Base Amplitude Panning
VisAsim	Visual-acoustic simulation for landscape impact assessment of wind farms
vs.	versus
W	wheel
WFS	Wave Field Synthesis
WT	wind turbine

# 1

## Introduction

*Disturb the sound of silence.*

Paul Simon

*Noise caused by traffic and facilities is a relevant health factor in urban environments, along major traffic routes and in the vicinity of airports. Still today, only little is known about the relation between a complex sound situation (soundscape) and its effects on a human being. The technique of auralization helps to reduce this knowledge gap between sound and its impact. Analogous to visualization, auralization enables us to listen to situations that do not really exist. By directly addressing the hearing sensation, auralization is an intuitive means for the assessment and communication of environmental noise scenarios. Possible applications of environmental sound auralization involve the generation of stimuli for experimental laboratory studies, the use as a demonstrator or a communication tool for non-experts, or a work bench in the product design.*

---

Parts of this chapter have been published in Applied Sciences **6**(5), 1–27 (2016) [1].

## 1.1. Background and motivation

According to the World Health Organization (WHO), noise pollution is still an underestimated threat that causes a number of short- and long-term health problems [2, 3]. Environmental noise is detrimental for public health, having negative impacts on human health and well-being [3]. Noise, here understood as unwanted sound, acts as a stressor and has been linked to non-auditory health effects including annoyance, sleep disturbance, cardiovascular disease, and impairment of cognitive performance [4]. It is a relevant health factor in urban environments, along major traffic routes and in the vicinity of airports. The major technical environmental noise sources are road, rail and air traffic, and industrial facilities.

Studies on the effects of noise rely on the accuracy of a noise exposure assessment. Noise exposure is typically modeled by calculation results. They are obtained from engineering models that are used to estimate noise indicators, such as the day-evening-night sound level,  $L_{\text{den}}$  [5]. However, noise, in contrast to sound, can principally not be measured or calculated using a device, but has to be assessed. For the most relevant noise sources, objective quantities have been derived that correlate with the annoyance as reported by people. Such examples are exposure-response relationships for noise annoyance [6, 7] or the probability for noise induced awakening reactions [8–10]. However, the confidence intervals of these relationships are usually large. One reason for this is the fact that the describing quantities used so far represent the acoustical situation only in a very simplified manner. In other words, the used noise indicators are rather limited. In general, still today, only little is known about the relation between a complex sound situation (soundscape) and its effects on humans.

A method to further investigate the signal properties relevant to the perception of sound, is to conduct experimental laboratory studies. To that aim, different stimuli are presented to test subjects and rated individually. Relying on audio recordings allows for little variation of different signal aspects only. A more versatile method with a much higher degree of freedom, as well as full control over the influencing signal parameters, is to synthesize the stimuli and, thus, to auralize an acoustical environment. Furthermore, auralization can also be used as a communication instrument for non-experts. In summary, the technique of auralization promises to reduce the knowledge gap between sound and its effects on humans.

## 1.2. What is auralization?

The expression “auralization” was established in analogy to visualization. Auralization is the technique to artificially make an acoustical situation audible. By incorporating prediction models into the auralization process, this technique allows to listen to situations that do not really exist. In literature, different definitions for auralization are found. In 1993, Kleiner stated that [11]:

“Auralization is the process of rendering audible, by physical or mathematical modeling, the sound field of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modeled space.”

In 2008, Vorländer published his book with the title “Auralization”, where he gave the following definition [12]:

“Auralization is the technique of creating audible sound files from numerical (simulated, measured, or synthesized) data.”

In this dissertation, auralization is understood as the technique of creating perceivable sound with the goal to provide audible information about a specific acoustical scene.

Auralization typically involves the fields source description, sound propagation simulation, and spatial audio technology. Today, most auralizations are generated based on computer models, digital signal processing and acoustic transducers. In other words, auralization is a sophisticated acoustical simulation technique that produces information that is easily understandable.

## 1.3. Auralization in environmental acoustics

### 1.3.1. Overview

Auralization has been in use for several decades in architectural acoustics, namely in the fields of room and building acoustics [12–14]. But auralization has only recently been discovered for environmental acoustical applications. Between these fields of application, the individual simulation steps may vary significantly. In room and building acoustical auralizations, it is common to utilize (anechoic) recordings as the source material, whereas in environmental acoustics, it is often required to artificially synthesize the source signals [1, 15–20]. Furthermore, the sound propagation simulation substantially differs. In room acoustical auralizations, the focus lies on the simulation of many room reflections, in particular specular vs. diffuse reflections, and diffraction [12, 14]. In building acoustical auralizations, sound transmission through structures is simulated using sound insulation prediction models [12]. For environmental acoustical auralizations, however, particular emphasis is placed on a detailed simulation of direct sound and ground reflection. Leastwise, in terms of sound reproduction, the same techniques are applicable.

Early attempts in auralizations for environmental noise applications have been made by a group at NASA Langley Research Center, where aircraft flyovers have been binaurally rendered based on monaural recordings [21]. In the same period, a synthesis model for the traction noise of electric rail-bound vehicles was developed at the RWTH Aachen University and used to study sound quality [15]. Newer auralization models synthesize the sounds of aircraft [18, 19, 22–24]. One of these models has already been combined with 3D visualizations to make aircraft noise both heard and seen in immersive virtual reality environments [19]. In the Swiss research project VisAsim<sup>1</sup> auralizations of wind farms were linked to synchronous GIS-based 3D visualizations [17, 25–28]. Within the Swedish project LISTEN [29, 30] and the European project HOSANNA [16, 31], tools for the auralization of road traffic noise were developed. In the recently completed Swiss

<sup>1</sup>Some content of the research project VisAsim is part of this dissertation.

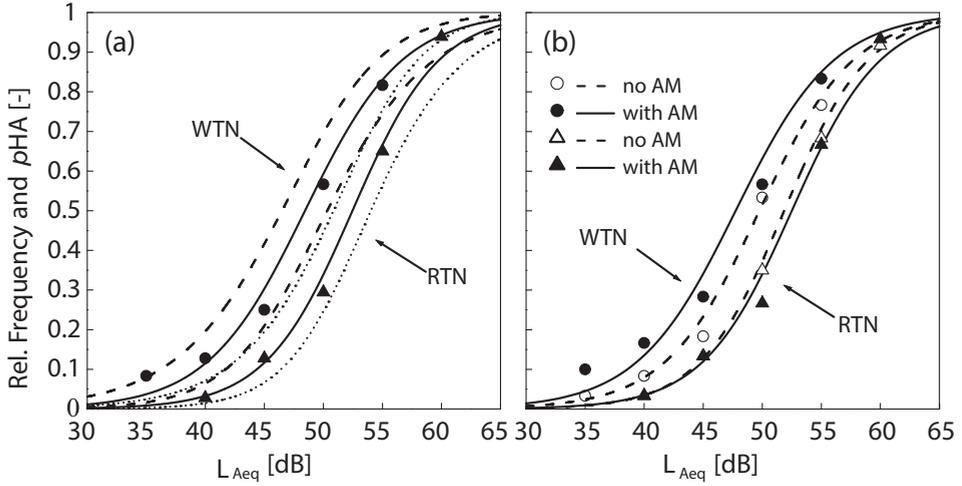


Figure 1.1: Application of auralization in environmental acoustics: Probability for high annoyance ( $p_{HA}$ ) from wind turbine noise (WTN) and road traffic noise (RTN) measured in a psychoacoustic laboratory study [38] that applied models from this dissertation (adapted from [39]). Panel (a) compares the two noise sources, and panel (b) additionally considers the effect of amplitude modulation (AM) in the sounds.

research project TAURA<sup>2</sup>, auralization models for passing cars and railway vehicles were developed [1, 20, 32, 33]. The main motivation is to provide more intuitive information about traffic noise scenarios for city planners, noise consultants and decision makers. There is common agreement in the point that information about noise in the form of dB values is difficult to communicate to the public [34]. Also, there is growing interest in the perceptual aspects of noise abatement measures. While in the past, the quality of noise mitigation measures was understood as the A-weighted sound pressure level reduction, the focus shifts to perceptual efficiency [35]. In this respect the optimal measure is not necessarily the one with the highest dB(A) drop, but the one with the highest annoyance reduction. The idea of describing and subsequently improving the acoustical environment with respect to human perception corresponds to the soundscape concept [36, 37].

### 1.3.2. Application examples

As described in the previous section, there are various types of application of auralization in environmental acoustics. For example, auralization may be used as a stimuli generator for experimental laboratory studies on noise effects on humans. Figure 1.1 shows results from such a psychoacoustic study by Schäffer on short-term annoyance due to wind turbine noise [38]. The wind turbine noise stimuli were auralized using the models developed within this dissertation. The systematic variation of stimuli, which was required by the study design, was only possible due

<sup>2</sup>The content of research project TAURA is part of this dissertation.

to the flexibility of the auralization model. The study showed that wind turbine noise is significantly more annoying than road traffic noise at the same A-weighted sound pressure level (see panel (a) in Figure 1.1). It further is revealed that for wind turbine noise, amplitude modulation (AM) leads to an increase in annoyance, while this is not the case for road traffic noise (see panel (b) in Figure 1.1).

Auralization of environmental sounds can also be used as a communication tool for non-experts. Instead of communicating abstract dB values, a specific situation can be made audible. In the form of a demonstrator for residents or stakeholders, auralization can help to objectify a “noise discussion”. By providing understandable information it can support public opinion forming and help in decision making. Public involvement in the planning process, i.e. participatory planning, may increase the acceptance of local residents. Using auralization as a communication tool may thus help to reduce prejudice against construction projects. In 2015, a mobile wind park demonstrator comprising visualization and auralization, was installed at a Swiss traveling exhibition on wind energy [40]. This approach is also followed within the ongoing European research project DESTINATE<sup>3</sup>, where a demonstrator, to showcase railway noise scenarios with different noise mitigation measures is currently under development by the author [41, 42].

Environmental sound auralization may also be beneficial in the product design as it allows to develop perception driven concepts. In the context of aircraft noise, Rizzi has launched the term “Perception-Influenced Design” as a novel approach for noise control [43]. Vehicle engineers and designers are particularly interested in how their new prototypes may sound like in different outdoor environments. Also in spatial planning, where the visualization of landscapes is common, the inclusion of auralization into the modeling process is demanded and therefore put forward.

## 1.4. Research objectives

As outlined in section 1.3, sophisticated environmental acoustical auralization models have mainly been developed in the context of aircraft noise. For other relevant environmental noise sources, such as road traffic, railway traffic or wind turbine noise, no or rather limited auralization models exist. The primary aim of this dissertation is to fill these gaps.

Three main research objectives comprise the development of computational models that allow the:

- Auralization of wind farms for impact assessment
- Auralization of road traffic for various vehicles, road surfaces and operational conditions
- Auralization of railway noise for various rolling stock, track types and speeds

Each of these main objectives contains three sub-objectives that are:

---

<sup>3</sup>Some content of the research project DESTINATE is part of this dissertation.

1. Investigation on relevance of involved acoustical phenomena for perceived realism
2. Development and implementation of synthesizer structures to reproduce
  - source characteristics,
  - sound propagation effects and
  - spatial impression.
3. Provision of methods to obtain model input parameters (literature, measurements, data analysis).

In order to allow for maximal flexibility with respect to scenarios and reproduction, the developed models are aimed at being parametric and object-based, i.e. separately represent acoustical sources and sound propagation.

## 1.5. Thesis outline

During the thesis research, it was found that the three considered environmental noise sources feature their specific peculiarities regarding sound emission characteristics as well as sound propagation effects at relevant exposure locations. Therefore source-specific auralization models were developed within this thesis, namely separately for wind turbine noise, road traffic noise and railway noise. This development process is also reflected in the chapter structure of this dissertation.

The dissertation is structured as follows: After the introduction in Chapter 1, Chapter 2 presents the methodology and general aspects regarding environmental sound auralization. Chapter 3 is dedicated to the auralization of wind farms. Chapter 4 presents a model for the auralization of accelerating passenger cars. Chapter 5 deals with railway noise, where a physically-based synthesis model for rolling and impact noise is derived. Thereafter follows a final conclusion and outlook in Chapter 6, and the author's Curriculum Vitæ and publication list.

## References

- [1] R. Pieren, T. Büttler, and K. Heutschi, *Auralization of accelerating passenger cars using spectral modeling synthesis*, [Applied Sciences](#) **6**, 5 (2016).
- [2] World Health Organization (WHO), *Night noise guidelines for Europe*, Tech. Rep. (World Health Organization (WHO), Regional Office for Europe, 2009).
- [3] L. Fritschi, L. Brown, D. Schwela, and S. Kefalopoulos, *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*, Tech. Rep. (World Health Organization (WHO), Regional Office for Europe, 2011).
- [4] M. Basner, W. Babisch, A. Davis, M. Brink, C. Clark, S. Janssen, and S. Stansfeld, *Auditory and non-auditory effects of noise on health*, *The Lancet* **383**, 1325 (2014).

- [5] International Organization of Standardization, *ISO 1996-1: Acoustics – Description, measurement and assessment of environmental noise - Part 1: Basic quantities and assessment procedures*, (2016).
- [6] H. M. E. Miedema and C. G. M. Oudshoorn, *Annoyance from transportation noise: Relationships with exposure metrics dnl and denl and their confidence intervals*, *Environmental Health Perspectives* **109**, 409 (2001).
- [7] S. A. Janssen, H. Vos, A. R. Eisses, and E. Pedersen, *A comparison between exposure-response relationships for wind turbine annoyance and annoyance due to other noise sources*, *Journal of the Acoustical Society of America* **130**, 3746 (2011).
- [8] M. Basner, A. Samel, and U. Isermann, *Aircraft noise effects on sleep: Application of the results of a large polysomnographic field study*, *The Journal of the Acoustical Society of America* **119**, 2772 (2006).
- [9] M. Brink, S. Omlin, C. Müller, R. Pieren, and M. Basner, *An event-related analysis of awakening reactions due to nocturnal church bell noise*, *Science of The Total Environment* **409**, 5210 (2011).
- [10] B. Schäffer, R. Pieren, F. Mendolia, M. Brink, and M. Basner, *Noise exposure-response relationships established from repeated binary observations: Modeling approaches and applications*, *Acoustical Society of America Journal* **141**, 3175–3185 (2017).
- [11] M. Kleiner, B.-I. Dalenbäck, and P. Svensson, *Auralization-an overview*, *Journal of the Audio Engineering Society* **41**, 861 (1993).
- [12] M. Vorländer, *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* (Springer, Berlin, 2008).
- [13] M. Kleiner, B.-I. Dalenbäck, and P. Svensson, *Auralization-an overview*, *Journal of the Audio Engineering Society* **41**, 861 (1993).
- [14] L. Savioja and U. P. Svensson, *Overview of geometrical room acoustic modeling techniques*, *The Journal of the Acoustical Society of America* **138**, 708 (2015).
- [15] M. Klemenz, *Sound synthesis of starting electric railbound vehicles and the influence of consonance on sound quality*, *Acta Acustica United with Acustica* **91**, 779 (2005).
- [16] J. Jagla, J. Maillard, and N. Martin, *Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method*, *Journal of the Acoustical Society of America* **132**, 3098 (2012).
- [17] R. Pieren, K. Heutschi, M. Müller, M. Manyoky, and K. Eggenschwiler, *Auralization of wind turbine noise: Emission synthesis*, *Acta Acustica United with Acustica* **100**, 25 (2014).

- [18] M. Arntzen and D. G. Simons, *Modeling and synthesis of aircraft flyover noise*, *Applied Acoustics* **84**, 99 (2014).
- [19] A. Sahai, F. Wefers, P. Pick, E. Stumpf, M. Vorländer, and T. Kuhlen, *Interactive simulation of aircraft noise in aural and visual virtual environments*, *Applied Acoustics* **101**, 24 (2016).
- [20] R. Pieren, K. Heutschi, J. M. Wunderli, M. Snellen, and D. G. Simons, *Auralization of railway noise: Emission synthesis of rolling and impact noise*, *Applied Acoustics* **127**, 34 (2017).
- [21] S. A. Rizzi, B. Sullivan, and C. A. Sondridge, *A three-dimensional virtual simulator for aircraft flyover presentation*, in *International Conference on Auditory Display* (Boston, USA, 2003) pp. 87–90.
- [22] M. Arntzen and D. G. Simons, *Ground reflection with turbulence induced coherence loss in flyover auralization*, *International Journal of Aeroacoustics* **13**, 449 (2014).
- [23] S. A. Rizzi, A. R. Aumann, L. Lopes, and C. L. Burley, *Auralization of hybrid wing-body aircraft flyover noise from system noise predictions*, *Journal of Aircraft* **51**, 1914 (2014).
- [24] F. Rietdijk, K. Heutschi, and C. Zellmann, *Determining an empirical emission model for the auralization of jet aircraft*, in *Euronoise 2015, 10th European Conference on Noise Control* (Maastricht, Netherlands, 2015) pp. 781–784.
- [25] M. Manyoky, U. Hayek Wissen, T. M. Klein, R. Pieren, K. Heutschi, and A. Grêt-Regamey, *Concept for collaborative design of wind farms facilitated by an interactive gis-based visual-acoustic 3d simulation*, in *Digital Landscape Architecture 2012, Peer Reviewed Proceedings at Anhalt University of Applied Sciences, Wichmann* (Berlin and Offenbach, 2012) pp. 297–306.
- [26] M. Manyoky, U. Hayek Wissen, K. Heutschi, R. Pieren, and A. Grêt-Regamey, *Developing a gis-based visual-acoustic 3d simulation for wind farm assessment*, *ISPRS International Journal of Geo-Information* **3**, 29 (2014).
- [27] K. Heutschi, R. Pieren, M. Müller, M. Manyoky, U. Hayek Wissen, and K. Eggenschwiler, *Auralization of wind turbine noise: Propagation filtering and vegetation noise synthesis*, *Acta Acustica United with Acustica* **100**, 13 (2014).
- [28] M. Manyoky, U. Wissen Hayek, R. Pieren, K. Heutschi, and A. Grêt-Regamey, *Evaluating a visual-acoustic simulation for wind park assessment*, [Landscape and Urban Planning](#) **153**, 180 (2016).
- [29] J. Forssén, T. Kaczmarek, J. Alvarsson, P. Lundén, and M. E. Nilsson, *Auralization of traffic noise within the listen project - preliminary results for passenger car pass-by*, in *Euronoise 2009, 8th European Conference on Noise Control* (Edinburgh, Scotland, 2009).

- [30] A. Peplow, J. Forssén, P. Lundén, and M. E. Nilsson, *Exterior auralization of traffic noise within the listen project*, in Forum Acusticum 2011, European Conference on Acoustics (Aalborg, Denmark, 2011) pp. 665–669.
- [31] J. Maillard and J. Jagla, *Real time auralization of non-stationary traffic noise - quantitative and perceptual validation in an urban street*, in AIA-DAGA 2013 Conference on Acoustics (Merano, Italy, 2013).
- [32] R. Pieren, T. Büttler, and K. Heutschi, *Auralisation of accelerating passenger cars*, in Euronoise 2015, 10th European Conference on Noise Control (Maastricht, Netherlands, 2015) pp. 757–762.
- [33] R. Pieren, J. M. Wunderli, A. Zemp, S. Sohr, and K. Heutschi, *Auralisation of railway noise: A concept for the emission synthesis of rolling and impact noise*, in INTER-NOISE 2016, 45th International Congress and Exposition of Noise Control Engineering (Hamburg, Germany, 2016) pp. 3466–3472.
- [34] P. McDonald, H. Rice, and S. Dobbyn, *Auralisation and dissemination of noise map data using virtual audio*, in Euronoise 2009, 8th European Conference on Noise Control (Edinburgh, Scotland, 2009).
- [35] A. Fiebig and K. Genuit, *Development of a synthesis tool for soundscape design*, in Euronoise 2009, 8th European Conference on Noise Control (Edinburgh, Scotland, 2009).
- [36] M. Schafer, *The Soundscape – Our Sonic Environment and the Tuning of the World* (Destiny Books, Vermont, 1994).
- [37] International Organization of Standardization, *ISO 12913-1: Acoustics – Soundscape - Part 1: Definition and conceptual framework*, (2014).
- [38] B. Schäffer, S. J. Schlittmeier, R. Pieren, K. Heutschi, M. Brink, R. Graf, and J. Hellbrück, *Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study*, *Acoustical Society of America Journal* **139**, 2949 (2016).
- [39] K. Eggenschwiler, K. Heutschi, B. Schäffer, R. Pieren, H. Bögli, and M. Bärlocher, *Noise from wind turbines – contributions of switzerland to its assessment*, *Lärmbekämpfung* **5/2016**, 159 (2016).
- [40] U. Wissen Hayek, R. Pieren, K. Heutschi, M. Manyoky, and A. Grêt-Regamey, *Exploring the qualities of gis-based visual-acoustic simulations of wind parks to support public opinion forming*, in *The Virtual and The Real in Planning and Urban Design: Perspectives, Practices and Applications*, edited by C. Yamu, A. Poplin, O. Devisch, and G. de Roo (Taylor & Francis Ltd, UK, 2018).
- [41] R. Pieren, K. Heutschi, R. Aalmoes, and D. G. Simons, *Evaluation of auralization and visualization systems for railway noise scenes*, in INTER-NOISE 2017, 46th International Congress and Exposition of Noise Control Engineering (Hong Kong, 2017) pp. 6555–6566.

- [42] R. Pieren, D. Lauper, and K. Heutschi, *Demonstrator for rail vehicle pass-by events*, in *Euronoise 2018, 11th European Congress and Exposition on Noise Control Engineering* (Crete, Greece, 2018).
- [43] S. Rizzi, *Towards reduced aircraft community noise impact via a perception-influenced design approach*, in *INTER-NOISE 2016, 45th International Congress and Exposition of Noise Control Engineering* (Hamburg, Germany, 2016) pp. 2–26.

# 2

## Methodology

*Prediction is very difficult, especially about the future.*

Nils Bohr

*This chapter introduces the methodology used in this dissertation and presents general concepts for environmental sound auralization. In the first part, the relation between classical noise calculation and auralization is presented and discussed. Then, an overview of the system components of current auralization systems is given and different system design variants are assessed. Thereafter, based on that, the chosen modeling approaches and the model development strategy of this work are elucidated.*

---

Parts of this chapter have been published in Applied Acoustics **127**, 34–45 (2017) [1] and in the proceedings of the International Congress and Exposition of Noise Control Engineering (INTER-NOISE 2017), 6555–6566 (2017) [2]

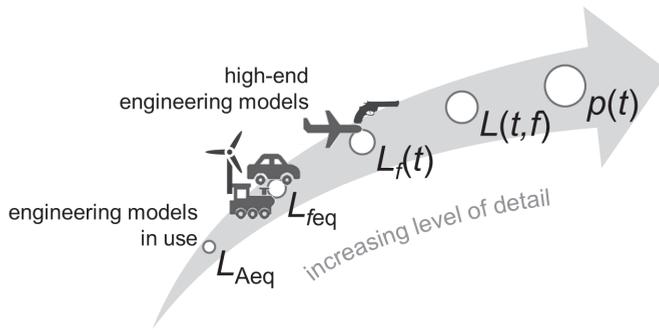


Figure 2.1: Level of detail of engineering models that provide sound pressure levels, e.g.  $L_{Aeq}$  or  $L_{feq,T}$ , and of auralization models where pressure time histories,  $p(t)$ , are created.

## 2.1. Auralization vs. classical noise calculation

As a starting point of this chapter, some general thoughts about similarities and differences between auralization in environmental acoustics and classical noise calculation are set forth. According to personal experience of the author, even professional acousticians who are well familiar with the latter, are often surprised by the disparities and the unfolding challenges involved in auralization.

Today, in environmental acoustics, engineering models are widely used, e.g. in the mapping of noise levels from airports or other environmental noise sources, such as road or railway traffic. The Environmental Noise Directive (END) [3] of the European Union relies on engineering models to identify noise pollution levels and to trigger the necessary action. These so-called engineering models are based on many simplifications that mostly seem justified by their intended application. However, in auralization some of these simplifications are not permitted. In what aspects and to what extent, is outlined in the following sections.

### 2.1.1. Frequency resolution

Still today, environmental noise is almost exclusively quantified by long-term averaged A-weighted sound pressure levels,  $L_{Aeq}$ , or derived indicators such as  $L_{den}$  [4]. Therefore, classical noise calculation engineering models are specifically designed to predict A-weighted average levels. During the past decades, the degree of detail of these models was substantially improved, particularly with respect to frequency resolution. Today, noise calculations in octave or 1/3 octave bands can be considered as state-of-the-art [5–9]. The increase in the level of detail of engineering models is illustrated in Figure 2.1.

However, except at low frequencies, the resolution of the human ear is known to be finer than 1/3 octaves [10]. Additionally, the presence of pure tone components requires a finer frequency resolution to correctly model e.g. pitch or interference patterns. These aspects were particularly investigated within this dissertation and resulted in the auralization models described in the following chapters.

Thus, compared to engineering models, auralization models require a finer fre-

quency resolution. This is particularly the case for the source characterization, as most technical sound sources also generate narrowband components, such as tones (see example of wind turbine noise in Figure 3.2) or resonance peaks (see example of railway rolling noise in Figure 5.3).

### 2.1.2. Time resolution - from days to milliseconds

Besides the frequency resolution, even more striking is the difference in temporal resolution. In classical environmental noise calculations, time averaging is typically performed over one year, which is mainly driven by legislation. The calculation process is often performed for distinct time periods within the 24 hours of the day, e.g. the evening or the night period, with a minimum time window of one hour. Therefore, most environmental noise models only support a minimal time resolution of one hour.

There are however a few rare exceptions that allow for a finer temporal resolution. They are either driven by a country-specific legislation, the nature of the noise type or the considered health outcome. For example, nocturnal aircraft, church bell and shooting noise are sometimes quantified using maximum sound levels [4, 11–13]. There, an accurate level prediction requires a corresponding calculation model to operate on fine temporal resolution. The resolution depends on the type of the involved level smoothing prior to taking the maximum level. In sound level meters, common time constants to smooth the sound pressure level history are 1 s (time weighting Slow [14]) and 125 ms (time weighting Fast [14]). Thus, the calculation resolution in dedicated noise models has to be somewhat smaller. Aircraft noise is commonly measured using the time constant Slow. Therefore, current time-step aircraft noise models resolve time in the order of a second or a bit lower. For shooting noise, for example the Swiss noise abatement legislation ordinance [15] prescribes the Fast time-weighting for civil firing range installations. To account for that, the Swiss shooting noise model sonARMS features a temporal resolution of 10 milliseconds [13].

In contrast to engineering models that predict sound pressure levels, an auralization model creates perceivable sound. Consequently, for an auralization model the human ear has to be seen as the measuring stick. However, the temporal resolution of human hearing is even considerably below 10 milliseconds [10, 16].

In principle, the propagation calculation in today's engineering models does not fundamentally limit the temporal resolution of these models. However, the modest source description constitutes the bottleneck for a finer time scale. The used emission models were developed based on extensive temporal and spatial averaging. This, in return, limits their ability to reproduce the temporal behavior of the emissions, and the source directivity. In the model, these limitations from the source description are inevitably forwarded to the model output.

### 2.1.3. Phase matters

Even if the time resolution is dramatically increased, engineering models only predict sound pressure levels. In contrast to sound pressure, sound pressure level is an energetic quantity that is derived from sound pressure by squaring it in the first

place [14]. In doing so, phase information of the sound wave is lost and cannot be reconstructed. Phase information, however, is needed to create perceivable sound.

The phase relationships within emission signals, between emission signals and between different propagation paths are of relevance for our auditory impression. The lack of such phase information in classical noise calculation is probably *the* major difference to auralization models. Instead of sound pressure levels, an auralization model calculates sound pressure-time histories. These signals implicitly carry detailed time, frequency and phase information. For that, the time resolution is clearly in the sub-milliseconds range. The time resolution is fundamentally determined by the maximum frequency to be represented in the auralized scene (Nyquist theorem). For example, to allow the reproduction of a 10 kHz frequency tone, a temporal resolution of 0.05 ms is required at least.

The generated signals are sometimes denoted as “pseudo-recordings” [17]. Self-evidently, all types of conventional time- and frequency-weighted levels can be derived therefrom. Furthermore, as opposed to engineering models, auralization models allow the calculation of psychoacoustic parameters, such as dynamic loudness, roughness, tonality or fluctuation strength [10].

Due to the much more detailed output obtained from an auralization model, it is evident that also the level of detail within the model is much higher as compared to today’s high-end engineering models, as illustrated in Figure 2.1.

#### 2.1.4. Directional information

In addition to the reported differences in time- and frequency resolution, and the lack of phase information, a forth difference lies in the provision of directional information at the receiver location. Auralization typically demands the evocation of a spatial impression and thus involves the creation of spatial sounds.

## 2.2. Auralization systems

This section describes current auralization systems with respect to applications in environmental acoustics. It gives an overview of the system components, their forms and their respective properties. Evaluation criteria for comparison and assessment are introduced. On that basis, different solutions and system variants are proposed and discussed.

### 2.2.1. Overview

Both visualization and auralization entail two basic processing steps that are

1. content creation and
2. implementation.

Figure 2.2 illustrates that these steps are related to two system components building the basic structure of an auralization system. Firstly, the content is created with a sound signal generation component. Using its output, the auralization is implemented with a sound reproduction component that renders an audible sound field. Both components are interrelated, as illustrated in Figure 2.2.

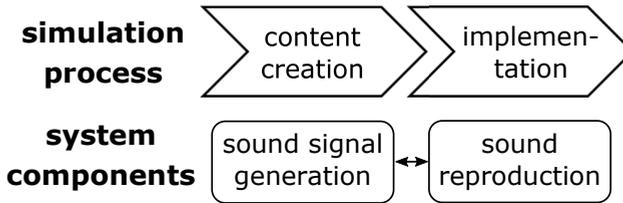


Figure 2.2: Auralization processing steps and corresponding system components.

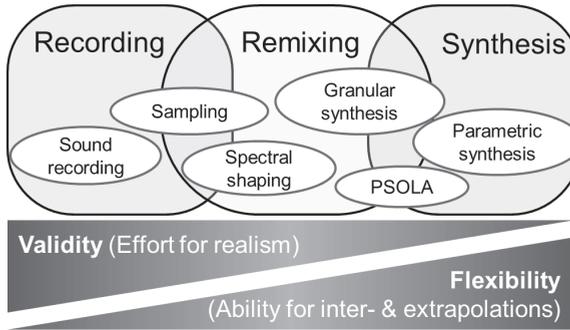


Figure 2.3: Three sound generation approaches and examples of established methods with an assessment of their validity and flexibility.

At the present day, various auralization systems exist, that differ with respect to their sound signal generation, as well as their sound reproduction strategy. The next sections outline these two components.

### 2.2.2. Sound generation approaches

The core of an auralization system is the ability to generate and process audio data. Nowadays, this is primarily achieved by digital signal processing. In this section, different approaches for sound signal generation are presented with respect to environmental sounds.

Audio data may be either based on sound recordings or artificially generated by sound synthesis. These two approaches are elucidated in the following sections. However, also intermediate levels, here denoted as remixing, exist, where recorded sounds are modified in various ways. This trisection is inspired by Lindquist [18], though here understood as partly overlapping approaches, as illustrated in Figure 2.3. Sampling (also denoted sample-based synthesis) is such an intermediate method, where portions of recordings (samples) are shifted in time, scaled in amplitude and summed up. This method was recently followed to generate environmental sound stimuli with predefined level indicators [19]. Also several recent laboratory studies relied on sampling in the auralization of road traffic scenarios that were generated on the basis of recorded single vehicle pass-by events [20–22]. With spectral shaping, the spectral content of a recorded signal is altered by filtering. This method was recently applied in the auralization of railway noise mitigation measures

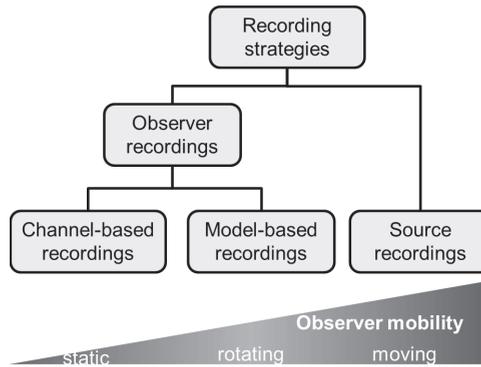


Figure 2.4: Hierarchy of sound recording strategies with three types of virtual observer mobility.

[23]. Granular synthesis and Pitch Synchronous Overlap and Add (PSOLA) are two more sophisticated, but more versatile methods that also rely on recordings. Here recorded signals are dissected into short grains which are then manipulated and mixed together. This has been applied to sounds of passenger cars [24, 25] and trucks [26]. In contrast, with parametric synthesis, sound signals are generated purely artificially.

As illustrated in Figure 2.3, relying on recordings has the advantage of an inherent high degree of realism. Such methods however provide a low flexibility as only existing cases can be adequately represented. In contrast, parametric sound synthesis features a very high degree of flexibility and is very versatile. It allows for interpolations between known states, but also for extrapolations to new, non-existing cases. It is however very challenging to synthesize realistically sounding signals.

In the following, only these two extreme cases of sound generation methods (according to the assessment in Figure 2.3) are outlined, namely sound recordings and parametric synthesis.

### 2.2.3. Sound recording

For sound recordings, a large variety of different microphone characteristics, arrangements and placement strategies exist. The preferred choice depends on the application, e.g. the virtual observer type and the reproduction system. Figure 2.4 shows that recording strategies can be split into two main classes. They feature either the perspective of the source or the observer. In the first class the emitted sound from a source is recorded whereas in the second class the sound at an observer location is captured. The main difference is that in the latter case sound propagation effects are included. Further, in source recordings, emission angles are used as a descriptor, whereas in observer recordings, immission angles are relevant.

Source recordings aim at capturing the sound pressure in close proximity as radiated by a source, at a certain emission angle. To do so, a microphone is typ-

ically placed very close to an individual source, such as implemented in [27] for railway interior noise. This approach is often challenging due to practical reasons such as source motion or multiple interfering sources. Source recordings at larger distances typically consist of two sequential tasks namely source separation and backpropagation [28].

Figure 2.4 shows that the observer recordings can be further split into channel-based and model-based recordings. Model-based recordings aim at capturing the effect of the sound field at a specific location independently of the reproduction system. For that, an abstract representation is used which allows to approximately reproduce sound pressure and particle velocity of that location. A well-known example is Ambisonics where the sound field is described by spherical harmonics.

In channel-based recordings, the microphones are also placed at the observer location but consistent with a specific sound reproduction system. Channel-based recordings require little processing, at the cost of being inflexible with respect to the reproduction system and virtual observer mobility. Examples of channel-based recordings are XY and ORTF stereo [29, 30], the Decca and the Fukada tree [30, 31], and binaural recordings.

When used in an auralization system, these three recording strategies result in different degrees of virtual observer mobility as illustrated in Figure 2.4.

#### 2.2.4. Parametric sound synthesis

In contrast to sound generation methods that rely on recordings, parametric synthesis is very versatile as it gives complete control over the generated signals. Various techniques to synthetically generate audible signals exist [32, 33], most of which were developed and are applied in the context of musical instruments [33], sound design [34] or speech processing [35]. However, not all of these techniques are suitable for the envisaged application in environmental acoustics, as some produce creative but not realistic sounds. In the following only a choice of parametric synthesis techniques with potential application in environmental sound are presented.

Two very basic techniques are additive and subtractive synthesis. Additive synthesis is based on Fourier series where a signal is constructed by the sum of discrete sine waves. Each sine wave has a time-varying amplitude and time-varying phase, and thus a time-varying instantaneous frequency. Subtractive synthesis uses time-varying filters to shape a spectrally rich waveform, typically white noise. The combination of additive and subtractive synthesis is referred to as Spectral Modeling Synthesis (SMS) or "Sines-plus-noise". It consists of the superposition of two signals, one of these generated by additive and the other by subtractive synthesis. This method has already been successfully applied in the context of sounds from aircraft [17, 36, 37], wind turbines [38] and passenger cars [39]. The latter two applications were developed within this dissertation and are presented in chapters 3 and 4 of this thesis.

Besides these signal-based synthesis methods, physically-based synthesis is another class of parametric synthesis methods [33]. This class mainly relies on the separate description of an exciter and a resonator structure, and their interaction. A rather basic representative of physically-based synthesis is modal synthesis. In

this dissertation, modal synthesis was used in the generation of railway rolling and impact noise [1, 2] and is presented in chapter 5.

### 2.2.5. Propagation filtering

If source signals are being synthesized or recorded, sound propagation effects yet have to be added to these signals. The propagation phenomena that have to be considered depend on the specific scene and the application. One fundamental difficulty is that it is *a priori* difficult to determine which sound propagation effects are relevant and thus have to be considered, and which ones can be neglected because they cannot be perceived.

In any case, geometrical spreading must be applied, which for the far-field sound pressure of a point source is a  $1/r$  distance dependency. For quickly varying source–receiver distances the Doppler effect may become relevant and should therefore be simulated. In some situations also the propagation delay may be relevant, e.g. if the auralization is combined with a visualization. For large propagation distances, air absorption attenuates the high frequency content as a function of meteorological conditions. Other meteorological effects are that atmospheric turbulence may introduce random level and phase fluctuations. Reflections at boundaries lead to interferences between sound waves. They have a series of influences on the perception, e.g. on spectral coloration, loudness, localization and spaciousness. The most often encountered reflection phenomenon in environmental acoustics is the ground effect which is the spectral pattern created by the interference of direct sound and ground reflected sound. Cases with many reflections, as they e.g. occur in street canyons, may even lead to room acoustical-like conditions. Also shielding of sources, leading to diffracted sound, may become important, e.g. in the case of a noise barrier. These effects may either be applied in the frequency or in the time domain.

### 2.2.6. Sound scene description

The sound generation component according to Figure 2.2 provides appropriate input for the sound reproduction component. The two components differ in the following aspects: The sound generation component creates audible signals at the observer location which therefore include sound propagation effects. Thus, this component uses data about the virtual environment including information about the angles under which a sound wave reaches the observer point. On the other hand, the reproduction system is blind with respect to the virtual environment.

As an interface between these two components, different sound scene descriptions are being used. The classical interface format is the channel-based description where each output channel from the generation component directly corresponds to a reproduction channel. In contrast, in an object-based description, each virtual source is represented separately by a sound pressure signal at the observer with corresponding time-variant polar angles of the associated incident sound wave. The concept of Ambisonics is to mathematically approximate the sound pressure field at the observer point using spherical harmonics. From an object-based description, the collection of spherical harmonics (denoted as “B-format”) can be created by

an ambisonic encoder. First-Order Ambisonics contains the sound pressure plus one additional signal per spatial dimension. Ambisonics of higher than first order is denoted as Higher-Order Ambisonics (HOA).

### 2.2.7. Sound reproduction systems

The sound reproduction system transforms the generated audio data, as described in section 2.2.2, into an audible sound field. It should produce appropriate and specific sound pressures inside the left and right ear canals of the listener and thus create a credible audible impression of the virtual environment.

An important requirement is the ability to localize sound sources as this increases the credibility of an auralized scene. This task, often referred to as spatial audio or 3D audio, requires multiple acoustic transducers. The generation of the individual transducer channel signals is denoted as spatialization. Two transducer types are available: multiple loudspeakers and headphones. Both solutions have their advantages and disadvantages, which are detailed in the following sections.

### 2.2.8. Loudspeaker reproduction

For multichannel loudspeaker reproduction, three aspects have to be considered that are

- the loudspeaker layouts, i.e. the number and placement of the loudspeakers in the room,
- the reproduction rendering, i.e. the technique to calculate individual loudspeaker feeds and
- room acoustics.

These three aspects are separately described in the following.

#### Loudspeaker layouts

Nowadays, many different speaker layouts are in use. They differ with respect to the number of loudspeakers and where they are placed in the room. Typically multiple satellite speakers are arranged with differing layouts. However, for low frequencies, where sound sources cannot be localized, one or multiple subwoofers are used. Most reproduction systems are optimized for one specific listening point and require that the distances to the satellites are equal. Otherwise, loudspeaker-specific time delays and gains may be used for compensation. According to the ITU recommendation BS.1116-3 [40], a minimal listening distance of 2 meters is recommended for high-quality multichannel sound systems. The zone around the listening point providing an adequate sound experience is denoted as sweet spot or optimum listening area.

Loudspeaker layouts may be divided into three classes depending on their used spatial dimensions:

- frontal setups

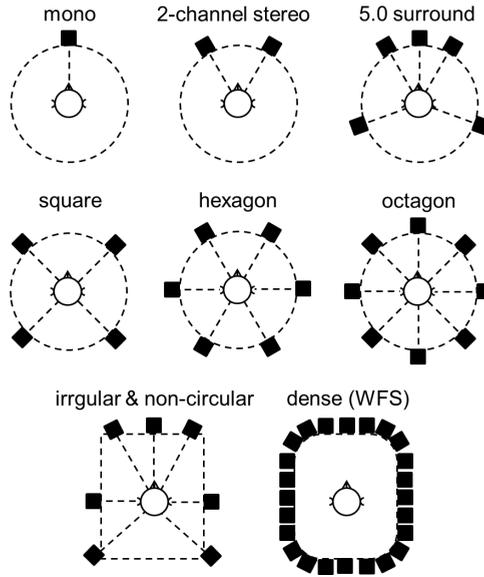


Figure 2.5: A number of common 2D loudspeaker layouts.

- 2D surround setups
- 3D audio

Some 2D loudspeaker layouts are illustrated in Figure 2.5. In frontal setups the loudspeakers are positioned in front of the listener only, e.g. mono, two-channel or three-channel stereo [41]. In 2D surround setups the loudspeakers are located on a horizontal plane around the listener. The height of their acoustical center is at listener's ears, which in most cases is 1.2 m above floor corresponding to a seated person [41]. For ambisonic reproduction, regular layouts are highly preferred (see examples in the second row of Figure 2.5). Some rendering strategies and applications also require very dense speaker arrangements. One such example is the TiME Lab at the Fraunhofer Heinrich-Hertz Institute in Berlin which uses a horizontal layout of 120 loudspeaker channels for a Wave Field Synthesis (WFS) reproduction [42]. Generally, for a given reproduction rendering strategy, the localization capability and stability grows with the number of loudspeakers.

For 3D audio even more variants exist. Full-sphere arrangements contain loudspeakers all around the listener, whereas upper hemisphere arrangements only cover the space on and above the listening plane. Regular layouts are achieved by placing loudspeakers at the corners of a platonic solid or a Lebedev grid [43]. Irregular layouts usually consist of several horizontal loudspeaker layers with typically one being in the listening plane. This also holds for Empa's listening room of the AuraLab which is depicted in Figure 2.6. This lab facility was designed and realized during this research. In the past few years, several 3D audio formats for



Figure 2.6: Photograph of the listening room of the AuraLab at Empa, Switzerland, which was designed and realized during this research to test the developed auralization models.

cinema, such as Auro-3D, Dolby Atmos or DTS:X, have been launched. They describe irregular upper hemisphere arrangements with a total of two or three layers.

Case-specific irregular layouts are used when loudspeakers are used to directly simulate non-moving acoustical sources in-situ. With this approach the interior sound within a train vehicle mock-up was simulated [27].

### Reproduction rendering

By reproduction rendering an input signal for each loudspeaker is derived. Various techniques to calculate signals for multiple loudspeakers, denoted as speaker feeds, exist. These techniques strongly depend on the given input type, which is a certain sound scene description, and their desired output format, i.e. a specific loudspeaker layout.

One rendering strategy may be called virtual microphones, where the responses of microphones within an appropriate arrangement are simulated at the observer location. A possible microphone spacing is simulated by varying time delays between the received signals and directivities by amplitude modulation. Both processes are steered by source-specific immission angles, implying that this strategy requires an object-based sound scene description. In chapter 4 this strategy is described for a virtual ORTF stereo arrangement.

Triplet-wise amplitude panning is a strategy to obtain speaker feeds for 3D audio based on the creation of phantom sources. For that, different panning laws and normalizations exist to determine the speaker gains. The normalization type expresses the assumption about signal superposition at the observer point. The most widely used triplet-wise panning method is Vector Base Amplitude Panning (VBAP) [44], which is a generalization of the (stereo) tangent panning law. Several modifications of the classical VBAP do exist such as a frequency dependent gain normalization [45] or the introduction of virtual speakers [46]. Triplet-wise am-

plitude panning requires an object-based sound scene description and can handle almost any irregular loudspeaker layout. As only a minimum number of feeds are simultaneously active, it produces good source localization, average sized sweet spot and only little coloration.

Ambisonic decoders are another kind of amplitude panners. In contrast to VBAP, they may provide negative speaker gains, which allows for sound wave cancellation at the center. Further, instead of only a few speakers, typically all speakers are simultaneously active and used to recreate the desired sound field from the spherical harmonic components. Different decoding algorithms exist to calculate speaker feeds from an ambisonic description (B-format). The number of loudspeakers must exceed the number of ambisonic components. In Ambisonics, regular speaker layouts are highly preferred.

With Crosstalk-Cancellation (CTC) the sound pressures at the ears are controlled by applying inverse filters to reduce the effect that the signal from the right speaker is picked up by the left ear, and vice versa, denoted as crosstalk. Therewith a spatial impression can be created with only two loudspeakers. This binaural rendering strategy was extended to more than two loudspeakers and dynamic processing [47, 48]. However, in practice compensation of head movements and the estimation of the crosstalk paths are difficult.

With Wave Field Synthesis (WFS) the desired sound field within a horizontal plane is reconstructed using surrounding sound sources. This technique is based on the Kirchhoff-Helmholtz integral and uses typically more than 100 loudspeakers as the upper cut-off frequency is determined by the minimal distance between individual loudspeakers. It allows for a large sweet spot, however at the cost of a high complexity and effort.

### Room acoustics

Listening rooms with controlled room acoustical conditions are mainly achieved by use of sound absorbers to attenuate sound reflections. These rooms typically feature a very low reverberation time. For example, the listening room of Empa's AuraLab (shown in Figure 2.6) has a mid-frequency reverberation time according to ISO 3382-1 [49] of  $T_{\text{mid}} = 0.11$  s. Figure 2.7 shows the frequency dependence of its reverberation time and the background noise at the listening position with the multichannel sound system on. From the shown 1/3 octave band spectrum, single number quantities are determined. The background noise level lies at 7 dBA and the noise ratings are GK0 [50], NC-15 [51] and NR 8 [52]. With this respect, the system thus fulfills the highest requirements on professional listening rooms according to IRT [50], ITU [40], Dolby [53] and THX [54]. These characteristics are the basis for the simulation of different acoustical environments and for the creation of virtual sound sources.

In general, anechoic rooms are ideal. In non-anechoic rooms, a certain amount of spectral coloration can be compensated using loudspeaker-specific equalization, such as reported for NASA's Exterior Effects Room (EER) [46]. Ultimately, each loudspeaker channel has to be calibrated using a sound level meter. The ease of level calibration can be seen as a major advantage of loudspeaker reproduction compared to headphones.

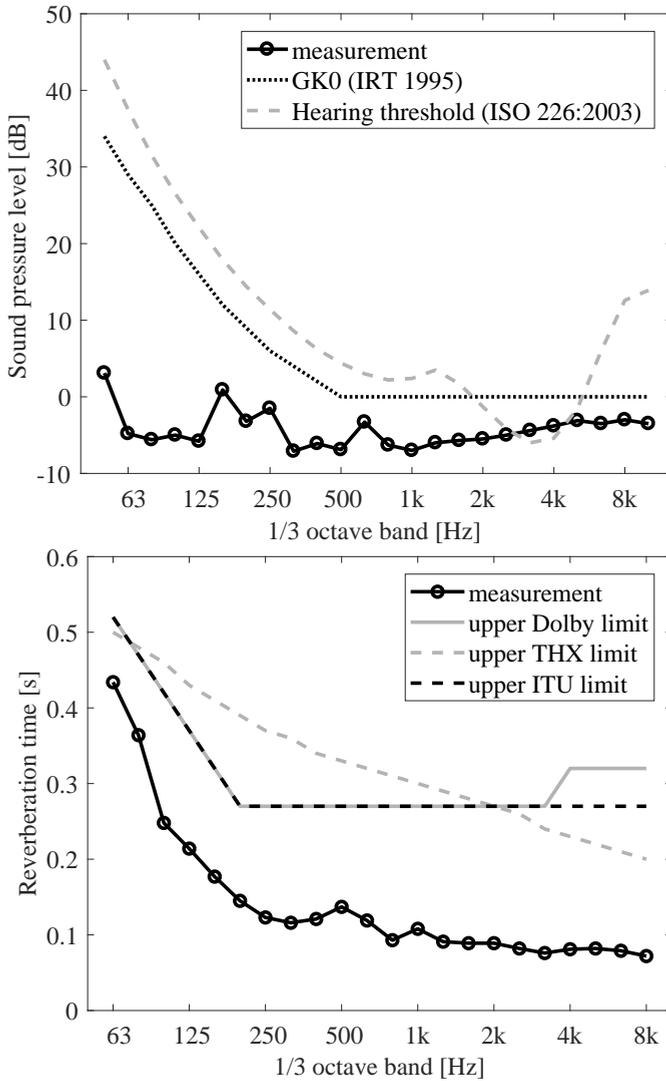


Figure 2.7: Room acoustical characterization of Empa’s AuraLab: Measured background noise and reverberation time in 1/3 octave bands in comparison to upper limit reference data from standards and recommendations.

### 2.2.9. Headphone reproduction

Modern headphone reproduction of spatial audio material is based on the combination of two technologies: Binaural technology and head tracking, both of which are described in the following sections. When listening to a monophonic signal through headphones, the signals which are received at the ear drums lack the influence of the own body on the sound field. Notably the head, torso and pinna significantly affect the received signals by providing shielding, reflections and time delay, which are all a function of the incident angles. This information is essential for the localization of sound sources.

Various headphone types exist that have to be differentiated. Ear-fitting headphones should be avoided due to difficulties regarding calibration and lack of reproducibility of their mounting. Instead circumaural headphones are preferred. Closed-back headphones feature higher attenuation of ambient noise compared to open-back headphones.

#### Binaural technology (HRTF rendering)

The influences of head, pinna and torso on the sound pressure at the location of the ear canal entrance are described by Head Related Impulse Responses (HRIR). For a point source in the far-field, the two HRIRs for left and right ear are a function of the immission angles, e.g. azimuth and elevation. Their Fourier transforms are known as Head Related Transfer Functions (HRTF). By convolving a monophonic signal with the corresponding HRIRs of left and right ear, a binaural (two-channel) signal is rendered.

The three major challenges in binaural technology are HRTF interpolation, individualization and calibration. Due to practical reasons measurements are best made with a spatial resolution of 5°. This is however too rough compared to the localization capabilities of humans. Therefore and in order to realize smooth transitions between the measurement points, a HRIR interpolation strategy is needed. Numerous algorithms have been proposed in the past years.

Most binaural renderers use general HRTFs which are e.g. obtained from acoustical measurements of a head-and-torso simulator or by averaging over many persons. However, HRTFs are not universal, but audibly differ from person to person. Therefore high-quality binaural reproduction requires some sort of individualization. Direct acoustical measurements of HRIR are costly and difficult. Another active field of research is the development of generic models for HRTFs that use anthropometric data as input [55]. HRFTs are also individualized using ear photographs as input.

Instead of rendering monophonic signals in an object-based manner, also intermediate formats are used during binaural rendering. The use of such an intermediate format can be interpreted as a form of interpolating HRTFs. One possibility is Ambisonics, where the B-format is decoded (or transcoded) to a binaural format. Ambisonics has the advantage that virtual head rotations can be easily realized with a matrix mixer. The modified ambisonic channels are then convolved with static HRTF-like filters and summed up. These filters are usually designed based on a virtual loudspeaker setup.

Another possibility is to use a virtual loudspeaker layout as intermediate format. Firstly, the signal for each virtual loudspeaker is obtained as described in

section 2.2.8. Secondly, these non-moving virtual sources are binaurally rendered using HRTFs. Such an emulation of a surround sound system through headphones is often denoted as virtualization and typically covers 5.1 or 7.1 surround sound [41].

### Head tracking

When listening to standard headphones, the reproduced sound field moves with the head. This is highly unnatural and thus weakens the credibility of the scene. To overcome this, an adaptation of the sound scene for (slight) head rotations is required. Binaural rendering thus becomes a function of the current head orientation which is measured by head tracking technologies. A passive measurement can be obtained by optical (“outside-in”) tracking relying on a video camera in the room and a dynamic face recognition algorithm. Active optical systems use infrared lighting, wearable motion tracker markers and cameras. For optical systems unobstructed line-of-sight is necessary for continuous tracking. Another option is electromagnetic motion tracking. Source-less tracking is achieved by wearable inertial sensors, i.e. accelerometers and gyroscopes. However, inertial tracking may suffer from drift errors due to temporal integration. The data transmission from the wearable sensor to a static receiver station follows either via cable or wireless (e.g. Bluetooth).

#### 2.2.10. System evaluation criteria

To differentiate, compare and assess different auralization systems, a list of evaluation criteria was established. The majority of general criteria can as well be applied to visualization systems and are presented in the following.

##### Portability

This criterion describes how easily the system can be moved from one place to another. It also describes to what extent the system depends on the specific location. For the application as a demonstrator it is e.g. important that simulations can be easily showcased at different places. On the other hand, to conduct experimental laboratory studies, a static system has the advantage that it is well defined and controlled.

##### Number of simultaneous participants

The number of people who can simultaneously experience an auralization is usually small. Most presented systems have a small sweet spot and are thus restricted to one or a few persons only. However, the scalability between the systems may vary.

##### Interactivity and user control

The degree of interactivity and the user control by the participant define the minimal required amount of real time signal processing. Virtual observer rotations require a real-time reproduction rendering which implies an appropriate omnidirectional format at the observer point. The option to activate and deactivate acoustical sources requires the separate rendering of their contributions and a real-time mixing. Virtual observer movements or changes of the environment (such as the introduction

of a noise barrier) finally require real-time propagation simulations as well as reproduction, given that the sound sources have simple directivities. Complex sound source modifications or complex directional behavior additionally require real-time sound synthesis.

## 2

### Physical correctness

Self-evidently, a simulation system has to be assessed with respect to its physical correctness. Firstly, it may be heuristically checked with respect to the physical phenomena that are accounted for and how appropriate the chosen models are. Secondly, the performance can be measured and compared, e.g. the angular coverage or the dynamic range of the sound reproduction system. Thirdly, deviations between ideal and reproduced values can be measured and evaluated. This calibration test involves e.g. the sound pressure level at a certain frequency.

### Intrusion

Sound reproduction systems differ with respect to their degree of intrusion. If the participant has to wear some sort of equipment, e.g. headphones and perhaps even with dangling cables, this is unnatural and may have an impact on his assessment of the simulation. Besides technical aspects such as the weight of the equipment, also comfort plays a role.

### Immersion

A high immersion means that the participant believes to be inside the virtual environment. The degree of immersion is influenced by the sound reproduction system, but also by visual aspects. In particular interactivity and spatial cues seem to be important for immersion.

### Perceived realism and plausibility

Generally, the auralizations should be perceived as realistic as possible. Apart from natural sounding signals, also factors such as the degree of immersion and intrusion play a role. In general, this criterion is very difficult to quantify. For some applications, particularly if only relative differences between scenarios are in the focus, merely a plausible simulation is sufficient. Plausibility can be understood as the difference to an inner reference due to experience and expectations [56]. The perceived quality of spatial audio systems may be investigated using the vocabulary developed in [57]. However, it seems even impossible to evaluate to what extent the system allows for a congruent overall perception of the simulated scene.

### Appropriateness

The appropriateness of a system depends on its application. Further, the system must be appropriate for the kind of participants, e.g. experts or laymen, and the type of relevant scenes to be simulated. A multifunctional system is appropriate for several applications, kinds of participants or scenes.

### Effort

Every system requires a certain technical, financial and personnel effort. Technical efforts comprise the equipment, the room and the installation, but also the computational effort. The financial costs for equipment and implementation vary a lot across systems. The personnel effort includes the system development as well as its usage and maintenance.

### Complexity and flexibility

Some fundamental decisions about the modeling approach highly determine the system complexity. However, in return more complex approaches often offer a greater flexibility with respect to scenarios or reproduction systems. Also modifiability, customization and expandability of the system determine its flexibility.

#### 2.2.11. Loudspeaker vs. headphone reproduction

Regarding the sound reproduction system, one central question is whether to use headphones or multiple loudspeakers as transducers to create a spatial audible impression. Both solutions have their specific advantages and disadvantages, which are discussed in the following.

The major advantages of headphones are their portability and lack of room influence. This permits a high flexibility regarding the choice of the location of presentation, which is especially important for a demonstrator.

However, the major disadvantages of headphones are their intrusion, the difficulty of calibration and the need for simulating the subject's influence on the sound field. The intrusion includes discomfort due to acoustical isolation, the device's weight, contact pressure, disturbing cables or warming-up of ears. Headphones are difficult to calibrate as their frequency response often strongly varies between listeners, but also between repeated placings. The influence of head, pinna and torso on the sound field is individual and thus laborious to simulate in a proper way. Otherwise incorrect sound levels and localization confusion including in-head localization occur. And ultimately, binaural rendering requires the use of a head tracking technology.

Conversely, these disadvantages are the major advantages of loudspeaker reproduction, given that a non-binaural rendering is used. Loudspeaker reproduction is non-intrusive and thus more natural. Separate channels can be calibrated easily using standard measurement equipment. Further, non-binaural loudspeaker reproduction nearly perfectly reproduces the individual's influence on the sound field. The main disadvantages however are the limited portability and the severe room influence on the reproduced sound field.

#### 2.2.12. Application-specific reproduction system

The previous section already revealed that there is no single favored reproduction system. It is rather that the specific application determines the requirements and allows for an assessment of the criteria to define a system.

In the following, three different possible applications of an auralization system for environmental noise scenes are discussed, namely

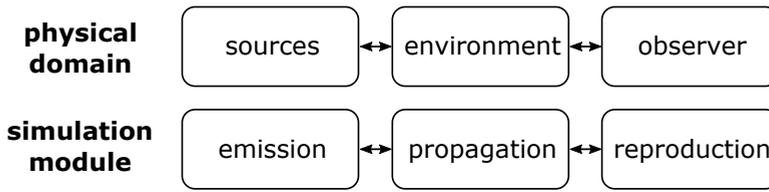


Figure 2.8: Simulation approach with three interconnected simulation modules representing three physical domains.

- (i) Stimuli generator for experimental laboratory study
- (ii) Test bench for experts and engineers
- (iii) Demonstrator for stakeholders, residents and broad public

Experimental laboratory studies are performed in highly controlled environments with well-defined, reproducible stimuli. They thus demand a maximal acoustical quality. As only a limited number of participants are involved, portability of the system is not essential. To avoid mutual influences between participants, usually there will only be one participant at the same time in the experiment and thus the system may operate with a small sweet spot. Based on that, for case (i), a non-binaural loudspeaker reproduction may be recommended.

Also experts and engineers need a high-quality simulation. However, thanks to their expertise they might also deal with some sort of abstraction. A test bench is used during longer time periods and by several participants at the same time, which have to be able to interact and discuss. This sets high requirements on the intrusion of the system. Consequently, for case (ii), a loudspeaker reproduction would be recommended.

A demonstrator needs to be portable and cost-effective in order to be used at different locations and by many different persons. The main goal is to provide information in an easily understandable manner. Therefore a compromise has to be made with respect to the acoustical quality, in particular the physical correctness. Thus, for case (iii), headphone reproduction is recommended. A web-based solution together with private headphones is even widely accessible by a broad public [58].

## 2.3. Auralization model structure

This section introduces the general structure of the auralization models that were established in this research and are explained in detail in the following chapters.

### 2.3.1. Separation of emission and propagation

According to the research objectives in section 1.4, the models should be flexible, by considering a large variety of sound sources and propagation situations. The models should allow for separate changes at the source, as well as on the propagation path. Noise control measures are examples of such changes that typically concern

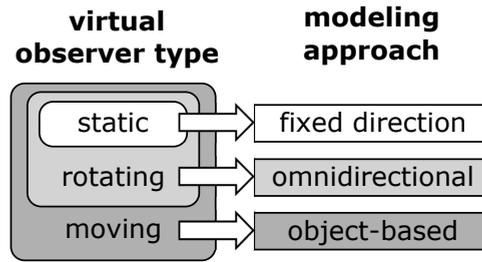


Figure 2.9: Virtual observer type determines the auralization modeling approach.

one of these two domains only, e.g. the installation of a vibration damper, or the installation of a noise barrier.

This requirement motivates to model sound sources and sound propagation separately. This complies with the “fundamentals of auralization” according to Vorländer [59], which is in particular the separation of sound generation and propagation. In acoustics, this is commonly known as the source–path–receiver concept. Figure 2.8 shows the interconnection of the three involved physical domains: Source, environment and observer. In the auralization models, the acoustical effects of these domains are separately represented by three corresponding simulation modules, namely emission, propagation and reproduction. They describe the source signal generation, the sound propagation filtering and the reproduction rendering.

### 2.3.2. Object-based modeling approach

The type of the virtual observer has also to be considered, as this also fundamentally affects the auralization modeling approach. Three virtual observer types with respect to their mobility are distinguished, namely a static, a rotating and a moving observer. A static observer has a fixed position and direction. The rotating observer has a fixed position, but is allowed to rotate. The moving observer enjoys full freedom by changing its position and direction. The observer type has severe implications on the modeling approach and thus determines the model complexity and flexibility. Figure 2.9 illustrates the three virtual observer types and the resulting modeling approaches.

In an auralization system that only supports a static observer with respect to the virtual sound scene, it is sufficient to exclusively model this specific immission situation. This means that the immission signals may be generated in an integral way and channel-based sound scene description (according to section 2.2.6) can be used as an interface between sound generation and reproduction system. For a rotating observer, the system has to allow for virtual rotations of the reproduced sound. This is typically achieved by using an intermediate sound scene description at the observer point. As a result of not using a channel-based format, sound generation and reproduction are decoupled which allows for renderings in various reproduction systems.

Many applications in environmental acoustics demand information for various source–receiver distances. An auralization that principally supports different ob-

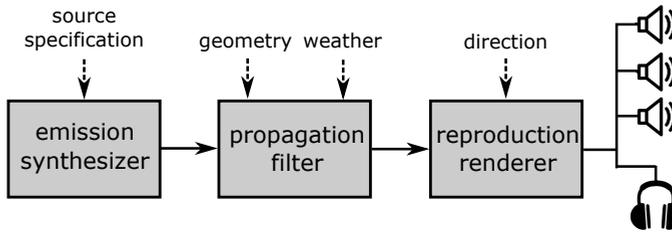


Figure 2.10: Block diagram of the auralization model comprising three simulation modules.

server locations, or even continuously varying propagation distances in the case of a moving virtual observer, requires a more sophisticated modeling approach. This is here referred to as an object-based modeling approach<sup>1</sup>. Therein, objects represent virtual acoustical sources and are interpreted as discrete sound emitting objects that are placed in the 3D modeling space, where also the virtual observer is located in. This approach features the highest degree of flexibility with respect to scenarios and reproduction systems.

Figure 2.10 shows a simple block diagram of the auralization models which contains three separate modules according to section 2.3.1. The first module provides the signals emitted by the sources. For each source object, the sound emissions are synthesized based on source specifications<sup>2</sup>. These signals are subsequently processed by the second module which contains a series of filters that simulate the propagation effects of the sound waves traveling from the sources to the virtual observer point. These propagation filters are generated as a function of the source–receiver geometry, topography, material properties and weather conditions. Also propagation effects due to obstacles, such as buildings [60], barriers [61] or natural objects [13, 62, 63] may be considered here, e.g. reflections, scattering and shielding. The third module is a reproduction renderer, which adequately renders the received signals to headphones (see section 2.2.9) or a multi-channel loud-speaker system (see section 2.2.8). By considering the sound incidence direction, a spatial impression can be created.

### 2.3.3. Emission synthesizer

According to the research objectives in section 1.4, various operating conditions of a noise source, such as a vehicle, shall be represented in the auralizations. Therefore in this work, componential models to describe the sources are striven for. This means that different components of a source are separately represented in the emission models. Besides a larger flexibility, this also makes sense in the light of the fact that within one noise source type, different sound generation mechanisms appear (e.g. aerodynamic and structure-borne). A separate treatment simplifies the corresponding sub-models and thus their development.

<sup>1</sup>This should not be confused with the object-based sound scene description as introduced in section 2.2.6

<sup>2</sup>In the general case, where the source has a directional behavior, also information about the propagation geometry has to be provided (not shown in Figure 2.10).

Section 2.2.2 introduced different methods for digital sound synthesis. According to Figure 2.3 the highest degree of flexibility is attributed to parametric sound synthesis, which is described in section 2.2.4. Flexibility is here understood as the ability to interpolate between known states and to extrapolate to new states. In contrast to other (data-based, non-parametric) synthesis methods, parametric synthesis does not rely on sound recordings to generate audio data. The sound signal is completely artificially generated by a computational algorithm that uses only a few input parameters. These steering parameters (sometimes denoted as control) can be interpolated to create smooth transitions between different sounds. Therewith, interpolations between different source conditions can be realized. Even more fascinating is the fact that the steering parameters can be extrapolated. This allows to synthesize sounds outside the measurement range, e.g. also future scenarios. Although it is challenging to synthesis realistically sounding signals, the strengths of parametric synthesis justify its application in this research.

#### 2.3.4. Propagation filter

Section 2.3.2 introduced the three simulation modules as illustrated in Figure 2.10. The module following the emission synthesizer is the propagation filter which simulates sound propagation effects. This section describes the principle of this module.

Section 2.2.5 introduced the aim and possible ingredients of the propagation filtering. In this work, the propagation model conception originates from the state-of-the-art engineering models that are used in environmental acoustics. However, the propagation effects are implemented differently. As they have to be applied to pressure-time signals, instead of sound energy, filters are used to process the signals. As explained in section 2.1, this process also requires a finer time and frequency resolution as compared to engineering models. Furthermore, some effects that seem irrelevant for environmental noise predictions, but are important for realism in auralizations are included, such as amplitude fluctuations due to atmospheric turbulence or the Doppler effect.

A first idea is to take attenuation data, in dB, obtained from engineering models as input to design corresponding filters [25, 64]. In contrast to this, here sound propagation is simulated as close as possible to the underlying physical mechanisms. For example the Doppler frequency shift is not realized by shifting frequencies, but by changing the propagation delay which is the physical mechanism behind the Doppler effect. Another example is the ground effect which is not realized in the frequency domain, but rather by introduction of an additional signal path that interferes with the direct sound. The ulterior motive therewith is to evoke (and *not* emulate) the correct perception of sound propagation effects. Also for this reason, throughout this research the propagation filtering is performed completely in the time domain. As the propagation conditions change over time—due to source or observer motion—the source signals are processed with a series of time-variant digital filters.

## 2.4. Development strategy for auralization models

This section outlines the strategy followed in this work during the development of the auralization models presented in the following chapters.

The main research objectives in section 1.4 specify the consideration of a variety of different sound sources—they e.g. differ with respect to their area of application, and their mechanisms of sound generation. Also the factors that determine their operational conditions and their functional principles vary widely. The chosen modeling approach described in section 2.3 dictates to model these sources individually. This requires expert knowledge from different technical disciplines and thus already points towards an interdisciplinary approach to develop the intended auralization models. Besides knowledge about the sources, also knowledge about the propagation, reproduction and perception of sound are substantial. In this work, for the model development, the most relevant fields are acoustics, signal processing, psychoacoustics, meteorology, audio engineering and automotive engineering. Besides using knowledge from various fields, in this work also different methods are used, such as numerical simulations, acoustical and non-acoustical measurements, and psychoacoustic experiments, i.e. listening tests.

Section 2.3 revealed that the auralization in this work is based on parametric sound synthesis. The flexibility of this synthesis type allows for the simulation of non-existing situations. It is however *a priori* unclear, how to build an auralization model for future scenarios. To that aim, a model development strategy consisting of several steps is proposed here. The following iterative procedure was established and pursued:

1. Development of model structure (algorithm to compute audio data)
2. Input parameter estimation from existing situations
3. Evaluation by comparison (calculated metrics, listening tests)
4. Feedback to steps 1. and 2.
5. Input parameter estimation with prediction models
6. Auralization of non-existing situations

In the first step, a detailed model structure is designed. This step is *per se* very challenging. Here, the use of expert knowledge is crucial. Detailed information about the sources, i.e. their behavior and operating principle, and the propagation conditions are central. Together with knowledge from acoustics, psychoacoustics and signal processing, this information is used to establish an algorithm to compute audio data with a few input parameters. In step 2, these input parameters are estimated from existing situations. To that aim, different methods can be used, such as literature study, measurements or simulations. In step 3, the model from step 1 is implemented and used with the input from step 2. The model output or interim results are then compared to reference data, which is typically obtained from measurements of existing situations. Initial comparisons may consist of calculated metrics. However, a proper evaluation should be performed with listening

tests. Such a perceptual evaluation is a good means to reveal relevant model deficiencies. Possibly, the results from step 3 suggest the introduction of modifications or adaptations. Thus, in step 4 changes to the model structure and the parameter estimation method are devised. This feedback to steps 1 and 2 is iteratively applied until a satisfying evaluation is achieved with the final model. Based on that, in step 5 the model input parameters for non-existing situations are estimated using prediction models. Finally, in step 6 the corresponding non-existing situations are auralized using these input parameters and the final model from step 1. This procedure, however, is rooted in the premise that all relevant future aspects of the situation have been considered in step 3.

During this research, experience about the feedback in the model development process (step 4) as a result of the evaluation or validation (step 3) was gained. One major finding is that this feedback is essential for the model quality. In many cases, adaptations to the model structure or the input parameter estimation had to be introduced and were motivated by findings from the evaluation, mostly missing aspects that had to be added to the model. Although these cases will be explained in detail in the following chapters, some examples are already stated here: Stochastic level fluctuations in wind turbine noise, rattling sound and the simulation of gear changes in car engine noise, or modal resonances in railway rolling noise are aspects that were only introduced in the proposed models after noticing their perceptual importance by listening to and comparing synthetic and recorded signals. Further, it was found that due to the model complexity, it is effective to perform the model evaluation (step 3) on model components with dedicated measurement data. One example is that in the case of road traffic, the development and evaluation was performed separately for tire noise and for propulsion noise.

## References

- [1] R. Pieren, K. Heutschi, J. M. Wunderli, M. Snellen, and D. G. Simons, *Auralization of railway noise: Emission synthesis of rolling and impact noise*, Applied Acoustics **127**, 34 (2017).
- [2] R. Pieren, K. Heutschi, R. Aalmoes, and D. G. Simons, *Evaluation of auralization and visualization systems for railway noise scenes*, in INTER-NOISE 2017, 46th International Congress and Exposition of Noise Control Engineering (Hong Kong, 2017) pp. 6555–6566.
- [3] European Commission, *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*, Official Journal of the European Union **45** (2002).
- [4] International Organization of Standardization, *ISO 1996-1: Acoustics – Description, measurement and assessment of environmental noise - Part 1: Basic quantities and assessment procedures*, (2016).
- [5] European Commission, *Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive*

- 2002/49/EC of the European Parliament and of the Council, Official Journal of the European Union **58** (2015).
- [6] T. Thron and M. Hecht, *The sonRAIL emission model for railway noise in Switzerland*, Acta Acustica United with Acustica **96**, 873 (2010).
- [7] H. G. Jonasson, *Acoustical source modelling of road vehicles*, Acta Acustica United with Acustica **93**, 173 (2007).
- [8] C. Talotte, P. van der Stap, M. Ringheim, M. Dittrich, X. Zhang, and D. Stiebel, *Railway source models for integration in the new European noise prediction method proposed in Harmonoise*, Journal of Sound and Vibration **293**, 975 (2006).
- [9] C. Zellmann, B. Schäffer, J. M. Wunderli, U. Isermann, and C. O. Paschereit, *Aircraft noise emission model accounting for aircraft flight parameters*, Journal of Aircraft **55**, 682 (2018).
- [10] H. Fastl and E. Zwicker, *Psychoacoustics* (Springer, Berlin, 2007).
- [11] M. Basner, A. Samel, and U. Isermann, *Aircraft noise effects on sleep: Application of the results of a large polysomnographic field study*, The Journal of the Acoustical Society of America **119**, 2772 (2006).
- [12] M. Brink, S. Omlin, C. Müller, R. Pieren, and M. Basner, *An event-related analysis of awakening reactions due to nocturnal church bell noise*, Science of The Total Environment **409**, 5210 (2011).
- [13] J. M. Wunderli, R. Pieren, and K. Heutschi, *The Swiss shooting sound calculation model sonARMS*, Noise Control Engineering Journal **90**, 224 (2012).
- [14] International Electrotechnical Commission, *IEC 61672-1:2013: Electroacoustics - Sound level meters - Part 1: Specifications*, (2013).
- [15] The Swiss Federal Council, *814.41 noise abatement ordinance of 15 december 1986 (nao) (status as of 1 january 2016)*, (2016).
- [16] B. Leshowitz, *Measurement of the two-click threshold*, Acoustical Society of America Journal **49**, 462 (1971).
- [17] S. A. Rizzi, A. R. Aumann, L. Lopes, and C. L. Burley, *Auralization of hybrid wing-body aircraft flyover noise from system noise predictions*, Journal of Aircraft **51**, 1914 (2014).
- [18] M. Lindquist, E. Lange, and J. Kang, *An assessment of the potential of using visual abstraction & sound for inclusive geodesign*, in Peer Reviewed Proceedings of Digital Landscape Architecture 2014 at ETH Zurich (Herbert Wichmann Verlag, VDE VERLAG GMBH, Berlin/Offenbach, 2014) pp. 170–179.

- [19] E. Accolti, F. Miyara, and F. di Sciascio, *Generating sound stimuli with given emergence level and low frequency content by mixing recordings*, *Acta Acustica United with Acustica* **103**, 782 (2017).
- [20] B. Schäffer, S. J. Schlittmeier, R. Pieren, K. Heutschi, M. Brink, R. Graf, and J. Hellbrück, *Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study*, *Acoustical Society of America Journal* **139**, 2949 (2016).
- [21] A. Klein, C. Marquis-Favre, and P. Champelovier, *Assessment of annoyance due to urban road traffic noise combined with tramway noise*, *Journal of the Acoustical Society of America* **141**, 231 (2017).
- [22] F. Rudzik, L. Thiesse, R. Pieren, J. M. Wunderli, M. Brink, M. Foraster, H. Héritier, I. Eze, C. Corrado, D. Vienneau, N. Probst-Hensch, M. Röösli, and C. Cajochen, *Sleep spindle characteristics and arousability from night-time transportation noise exposure in healthy young and older individuals*, *Sleep*, [zsy077](#) (2018).
- [23] T. Koch, Flyer Infopunkt Lärmschutz, Tech. Rep. (Fraunhofer Heinrich-Hertz-Institut, Berlin, 2016).
- [24] J. Jagla, J. Maillard, and N. Martin, *Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method*, *Journal of the Acoustical Society of America* **132**, 3098 (2012).
- [25] J. Maillard and J. Jagla, *Auralization of non-stationary traffic noise using sample based synthesis - comparison with pass-by recordings*, in INTER-NOISE 2012, 41st International Congress and Exposition of Noise Control Engineering (New York City, 2012).
- [26] J. Forssén, P. Andersson, P. Bergman, K. Fredriksson, and P. Zimmermann, *Auralization of truck engine sound – preliminary results using a granular approach*, in AIA-DAGA 2013 Conference on Acoustics (Merano, Italy, 2013).
- [27] Ö. Johansson, S. Schönfeld, and D. Lindfors, *Sound sketch procedure for auralization of the interior sound of a high speed train*, in INTER-NOISE 2012, 41st International Congress and Exposition of Noise Control Engineering (New York City, 2012).
- [28] F. Meng, G. Behler, and M. Vorländer, *A synthesis model for a moving sound source based on beamforming*, *Acta acustica united with Acustica* **104**, 351 (2018).
- [29] G. Ballou, ed., *Handbook for Sound Engineers*, 4th ed. (Elsevier, Amsterdam, 2008).
- [30] M. Dickreiter, V. Dittel, W. Hoeg, and M. Wöhr, eds., *Handbuch der Tonstudioteknik - Band 1*, 7th ed. (ARD.ZDF medienakademie, Nürnberg, 2008).

- [31] S. Weinzierl, ed., *Handbuch der Audiotechnik* (Springer, Berlin Heidelberg, 2008).
- [32] E. R. Miranda, ed., *Computer Sound Design - Synthesis techniques and programming*, 2nd ed. (Focal Press, Oxford, 2002).
- [33] S. Bilbao, ed., *Numerical sound synthesis - Finite difference schemes and simulation for musical acoustics*, 1st ed. (John Wiley & Sons, Sussex, 2009).
- [34] A. Farnell, ed., *Designing sound - Practical synthetic sound design for film, games and interactive media using dataflow* (Applied Scientific Press, London, 2008).
- [35] P. Taylor, ed., *Text-to-Speech Synthesis* (Cambridge University Press, Cambridge, UK, 2009).
- [36] M. Arntzen and D. G. Simons, *Modeling and synthesis of aircraft flyover noise*, *Applied Acoustics* **84**, 99 (2014).
- [37] A. Sahai, F. Wefers, P. Pick, E. Stumpf, M. Vorländer, and T. Kuhlen, *Interactive simulation of aircraft noise in aural and visual virtual environments*, *Applied Acoustics* **101**, 24 (2016).
- [38] R. Pieren, K. Heutschi, M. Müller, M. Manyoky, and K. Eggenschwiler, *Auralization of wind turbine noise: Emission synthesis*, *Acta Acustica United with Acustica* **100**, 25 (2014).
- [39] R. Pieren, T. Bütler, and K. Heutschi, *Auralization of accelerating passenger cars using spectral modeling synthesis*, *Applied Sciences* **6**, 5 (2016).
- [40] International Telecommunication Union, *Recommendation ITU-R BS.1116-3: Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*, (2015).
- [41] International Telecommunication Union, *Recommendation ITU-R BS.775-3: Multichannel stereophonic sound system with and without accompanying picture*, (2012).
- [42] O. Schreer, I. Feldmann, C. Weissig, P. Kauff, and R. Schäfer, *Ultrahigh-resolution panoramic imaging for format-agnostic video production*, *Proceedings of the IEEE* **101**, 99 (2013).
- [43] P. Lecomte, P.-A. Gauthier, C. Langrenne, A. Berry, and A. Garcia, *A fifty-node lebedev grid and its applications to ambisonics*, *J. Audio Eng. Soc* **64**, 868 (2016).
- [44] V. Pulkki, *Virtual sound source positioning using vector base amplitude panning*, *Journal of the Audio Engineering Society* **45**, 456 (1997).

- [45] M.-V. Laitinen, J. Vilkamo, K. Jussila, A. Politis, and V. Pulkki, *Gain normalization in amplitude panning as a function of frequency and room reverberance*, in AES 55th International Conference (Helsinki, Finland, 2014).
- [46] K. J. Faller, S. Rizzi, and A. R. Aumann, *Acoustic Performance of a Real-Time Three-Dimensional Sound-Reproduction System*, report NASA/TM-2013-218004, Tech. Rep. (Langley Research Center, Hampton, Virginia, 2013).
- [47] T. Lentz and C. Renner, *A four-channel dynamic cross-talk cancellation system*, in CFA/DAGA 2004, Annual Conference on Acoustics (Strasbourg, 2004) pp. 37–38.
- [48] E. Röcher, M. Kohonen, J. Stienen, and M. Vorländer, *Dynamic crosstalk-cancellation with room compensation for immersive cave-environments*, in DAGA 2016, Annual Conference on Acoustics (Aachen, Germany, 2016) pp. 1020–1022.
- [49] International Organization of Standardization, *ISO 3382-1: Acoustics – Measurement of room acoustic parameters - Part 1: Performance spaces*, (2009).
- [50] Institut für Rundfunktechnik (IRT), *Akustische Information 1.11-1/1995: Höchstzulässige Schalldruckpegel von Dauergeräuschen in Studios und Bearbeitungsräumen bei Hörfunk und Fernsehen*, (1995).
- [51] Acoustical Society of America (ASA), *ANSI/ASA S12.2: American National Standard Criteria for Evaluating Room Noise*, (2008).
- [52] International Organization of Standardization, *ISO Recommendation R1996: Acoustics – Assessment of noise with respect to community response*, (1971).
- [53] Dolby Laboratories, Inc., *5.1-Channel Music Production Guidelines - Issue 3*, (2005).
- [54] THX, Lucasfilm Ltd., *Professional Multichannel Mixing and Monitoring - pm3 specifications, version 1.3*, (1999).
- [55] R. Bomhardt, M. Lins, and J. Fels, *Analytical ellipsoidal model of interaural time differences for the individualization of head-related impulse responses*, [Journal of the Audio Engineering Society](#) **64**, 882 (2016).
- [56] A. Lindau and S. Weinzierl, *Assessing the plausibility of virtual acoustic environments*, *Acta Acustica united with Acustica* **98**, 804 (2012).
- [57] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkman, and S. Weinzierl, *A spatial audio quality inventory (SAQI)*, *Acta Acustica united with Acustica* **100**, 984 (2014).
- [58] L. Jiang, M. Masullo, L. Maffei, F. Meng, and M. Vorländer, *A demonstrator tool of web-based virtual reality for participatory evaluation of urban sound environment*, *Landscape and Urban Planning* **170**, 276 (2018).

- [59] M. Vorländer, *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* (Springer, Berlin, 2008).
- [60] K. Heutschi, *Calculation of reflections in an urban environment*, *Acta Acustica United with Acustica* **95**, 644 (2009).
- [61] International Organization of Standardization, *ISO 9613-2: Acoustics – Attenuation of sound during propagation outdoors - Part 2: General method of calculation*, (1996).
- [62] J. M. Wunderli, *An extended model to predict reflections from forests*, *Acta Acustica United with Acustica* **98**, 263 (2012).
- [63] R. Pieren and J. M. Wunderli, *A model to predict sound reflections from cliffs*, *Acta Acustica United with Acustica* **97**, 243 (2011).
- [64] J. Forssén, A. Hoffmann, and W. Kropp, *Auralization model for the perceptual evaluation of tyre–road noise*, *Applied Acoustics* **132**, 232 (2018).

# 3

## Auralization of Wind Turbine Noise

*This chapter describes an auralization model for wind turbine noise. It consists of the sound emission synthesis of wind turbines and the corresponding propagation filtering, considering the specific geometry of a highly elevated source and distant receivers. The synthesis of wind turbine sounds considers the spectral content and frequency dependent amplitude modulation (AM). The generated sound pressure signal is composed of discrete tonal components and amplitude modulated broadband noise. The AM functions are synthesized as the superposition of a random process and a periodic function with the blade passing frequency. With the model, the quite complex wind turbine sound can be decomposed into only a few components that can be described by few parameters. As demonstrated by a listening test, the synthesized wind turbine audio signal can be easily confused with the original sound. It is further shown that the propagation filtering can be implemented efficiently as a series of finite impulse response (FIR) filters with a relative small number of taps. Measurements have revealed that a wind turbine has to be modeled as largely extended source in order to correctly simulate the ground effect. Based on numerical simulations, a model to incorporate energy neutral short-time level fluctuation effects due to atmospheric turbulence is derived. As a consequence of the source extension, the fluctuations are significantly smaller than expected for a concentrated point source.*

---

Parts of this chapter have been published in Acta Acustica united with Acustica **100**, 25–33 (2014) [1], in Acta Acustica united with Acustica **100**, 13–24 (2014) [2], in the ISPRS International Journal of Geo-Information **3**, 29–48 (2014) [3] and in the Journal of the Acoustical Society of America **139**(5), 2949–2963 (2016) [4].

### 3.1. Introduction

Planning of new renewable energy installations in the landscape is a complicated matter all over Europe. Although the public generally supports the renewable energy deployment, the implementation of new installations often fails when it comes to the selection of appropriate locations, especially regarding wind farm locations on the local level [5, 6]. One reason is that in the context of wind farms, people perceive wind turbine noise as one of the most prominent annoyance factors [7].

Wind turbine noise is associated with higher annoyance reactions than transportation or industrial noise [8]. Knowledge of the reasons for these annoyance differences, however, is still relatively scarce. In particular, it is not known which acoustical characteristics of wind turbines alone, i.e., without potential effect modifiers such as the visibility of wind turbines, are associated with annoyance. An auralization tool for wind turbine noise is thus helpful for wind farm assessments [9] as well as to increase understanding of the perception of wind turbine noise [4, 10].

A simple approach for the auralization of wind turbine noise starts with an audio recording taken close to the source. Subsequently the signal is filtered with help of a 1/3 octave band equalizer according to damping values obtained with a spectral propagation model [11, 12]. The approach chosen here is based on a purely synthetic emission audio signal [1] and a subsequent filtering [2] that models each propagation effect individually. This allows for a more flexible adaptation of the emission signal characteristics and a more subtle simulation of the propagation, including possible Doppler shifts for moving receivers. The auralization tool presented here is part of the Swiss research project VisAsim [1, 2, 9, 13–15] that has studied the landscape impact assessment of wind farms considering combined visual and acoustical stimuli.

This chapter is structured as follows: In section 3.2 the emission synthesis of wind turbine sounds is presented. Section 3.3 covers the propagation filtering. The chapter ends with a conclusion in section 3.4.

### 3.2. Emission synthesis

Sounds from modern wind turbines generally consist of broadband noise and possible discrete tonal components [16, 17]. An often reported phenomenon related to wind turbine noise is the periodic amplitude modulation (AM) of the sound pressure level, which is related to the blade passing frequency of the turbine. Often found terms describing the hearing sensations of AM are “swishing”, “thumping” or “beating”, which are likely to be caused by different physical mechanisms [18–20]. A good overview of the sound generation mechanisms of wind turbines can be found in the book by Wagner et al. [21].

In this section a pure synthesis model for wind turbine emission sounds is presented in section 3.2.1. The input parameters of the model are obtained by elaborate signal analysis algorithms described in section 3.2.2 which are exemplarily applied to emission recordings of two modern wind turbines in section 3.2.3.

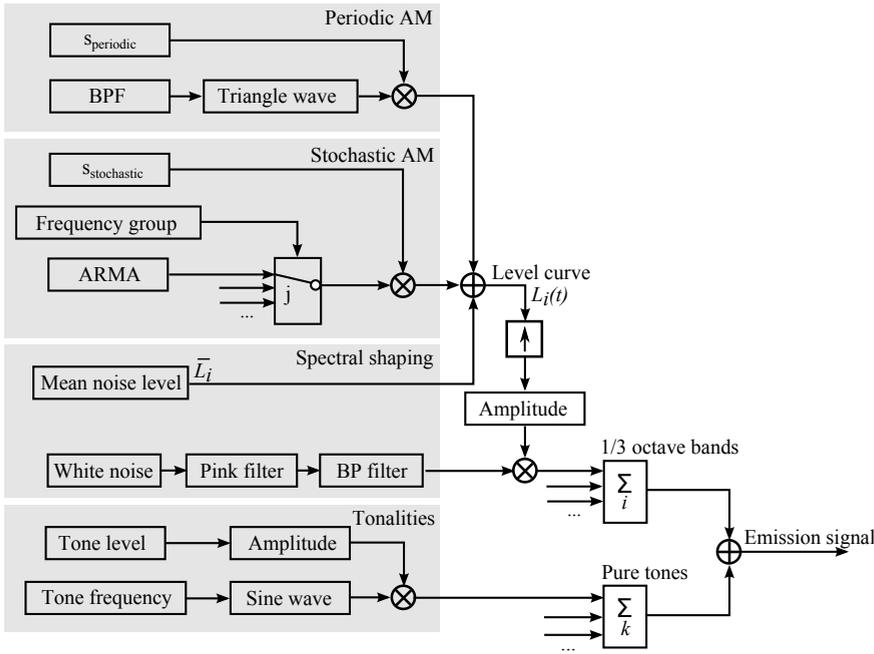


Figure 3.1: Signal flow chart of the wind turbine noise synthesis model. The resulting sound pressure emission signal is obtained by adding a tonal signal, consisting of the sum of pure tones with indices  $k$ , and a noise signal that is the sum of 1/3 octave bands with indices  $i$ . The 1/3 octave band signals are band-pass (BP) filtered noise that is amplitude modulated by a level curve  $L_i(t)$ . Stochastic Amplitude Modulation (AM) functions are generated by an autoregressive–moving-average (ARMA) model for each modulation band  $j$ . Periodic AM functions are calculated using the blade passing frequency (BPF) as input.

### 3.2.1. Synthesizer structure

#### General assumptions and model structure

It is assumed that the wind turbine emission signal has a stationary character in terms of an invariant short-term spectrum where ‘short’ means about 5 seconds. The signal is supposed to be comprised of possible tonal components and broadband noise with frequency dependent amplitude modulation (AM). The AM functions consist of periodic and stochastic parts. Discrete impulsive sounds are not taken into account. With these preconditions follows the flow chart in Figure 3.1, describing the synthesis model structure. The model is subdivided into four blocks (grey areas) that are explained in detail in the following sections.

As illustrated in Figure 3.1, the resulting output sound pressure signal,  $p_e$ , as a function of time  $t$  is modeled by

$$p_e(t) = p_{\text{tonal}}(t) + p_{\text{noise}}(t) \tag{3.1}$$

where the signal  $p_{\text{tonal}}$  contains tonal components and  $p_{\text{noise}}$  is a broadband noise signal, both of which are defined in the following sections.

As input parameters, the model shown in Figure 3.1 uses four figures per 1/3 octave band, two figures per tone plus some general input parameters, such as the blade passing frequency of the turbine. This results in a total number of input parameters of about 120.

### Tonality synthesis

The tonal components are synthesized by  $N_t$  sine wave generators. The output  $p_{\text{tonal}}(t)$  of the tonal synthesis is given by

$$p_{\text{tonal}}(t) = \sum_{k=1}^{N_t} p_0 10^{L_{t,k}/20} \sqrt{2} \sin(2\pi f_{t,k} t) \quad (3.2)$$

with the tone frequencies  $f_{t,k}$  in Hz, the tone sound pressure levels  $L_{t,k}$  and the reference pressure  $p_0 = 20 \mu\text{Pa}$ .

### Spectral shaping

The spectral shaping of the broadband noise component is performed in 1/3 octave bands. For each 1/3 octave band  $i$ , white noise is generated and filtered with a digital pink filter. The pink filter is realized by a cascaded series of three IIR second-order sections, i.e. three digital biquad filters. The output of the pink filter is bandpass filtered by a 8th order Butterworth filter and normalized to unit signal power to obtain the signal  $v_i(t)$ . The broadband component is given by

$$p_{\text{noise}}(t) = \sum_{i=1}^{N_b} p_0 10^{L_i(t)/20} v_i(t) \quad (3.3)$$

with  $N_b$  being the number of considered 1/3 octave bands and  $L_i(t)$  being the sound pressure level of band  $i$ .

### Amplitude modulation

The amplitude modulation (AM) is implemented in Equation (3.3) by the use of 1/3 octave band levels  $L_i(t)$  being a function of time  $t$ . The level curves  $L_i(t)$  are assumed to consist of three additive parts

$$L_i(t) = \bar{L}_i + F_{\text{periodic},i}(t) + F_{\text{stochastic},i}(t) \quad (3.4)$$

where  $\bar{L}_i$  denotes the arithmetic mean sound pressure level,  $F_{\text{periodic},i}(t)$  is a periodic level fluctuation function representing the periodic AM and  $F_{\text{stochastic},i}(t)$  is a stochastic process representing the stochastic AM.  $L_i(t)$  is synthesized at a sampling frequency  $f_{s,L} = 30$  Hz and upsampled to the audio sampling frequency  $f_s = 44.1$  kHz before being applied as amplitude modulation in Equation (3.3).

### Periodic amplitude modulation

The periodic level fluctuation  $F_{\text{periodic},i}(t)$  is a periodic function with zero mean and period  $T_{BP} = 1/f_{BP}$ .  $f_{BP}$  denotes the blade passing frequency (BPF) of the turbine given in Hz and can be calculated from the rotational speed  $n_{\text{rot}}$  in rpm of the turbine by  $f_{BP} = N_{\text{blades}}n_{\text{rot}}/60$  with the number of blades  $N_{\text{blades}}$  being typically = 3. With the periodic level fluctuations the well-known “swishing” and “thumping” sounds are implemented. Regarding phase it is assumed that the maximum level of those sounds occurs at the downstroke when the blades pass the horizontal position [20, 22]. From the initial position  $\beta_0$  of the down stroking blade, defined as the angle between the blade and the vertical axis in degrees ° (up = 0°), the time shift

$$T_h = \frac{90^\circ - \beta_0}{360^\circ} \frac{N_{\text{blades}}}{f_{BP}} \quad (3.5)$$

is calculated. The periodic level fluctuation is a periodic function with a defined amplitude and can be expressed as

$$F_{\text{periodic},i}(t) = s_{\text{periodic},i} W(t - T_h) \quad (3.6)$$

with the standard deviation  $s_{\text{periodic},i}$  of the periodic level fluctuations in dB and the periodic function  $W(t')$  having zero mean, period  $T_{BP}$ , its maximum at time  $t' = 0$  s and unit signal power. Motivated by measured sound pressure level curves (see Figure 3.4 or [23]) and hence based on listening tests where measured signals were compared to synthetic signals, it was decided to use a triangle waveform for  $W$ . In order to satisfy the unit signal power condition, the triangle wave has to be scaled to have minimum and maximum values of  $\pm\sqrt{3}$ , respectively.

### Stochastic amplitude modulation

The stochastic level fluctuations  $F_{\text{stochastic},i}(t)$  are generated by random processes. As especially at high frequencies the stochastic level fluctuations between 1/3 octave bands are highly correlated, groups of 1/3 octave bands steered by the same fluctuation function  $\eta_j(t)$  are used. For every group  $j$  an ARMA model generates a stochastic fluctuation function which is then normalized to unit signal power to obtain the fluctuation signal  $\eta_j(t)$ . The stochastic level fluctuations are then given by

$$F_{\text{stochastic},i}(t) = s_{\text{stochastic},i} \eta_j(t) \quad (3.7)$$

where  $s_{\text{stochastic},i}$  is the standard deviation of the stochastic level fluctuations in dB. In all studied cases, the level fluctuations could be well modeled by a white Gaussian process filtered by a 1st order Butterworth low-pass filter. From about 20 measurements at different wind conditions and different turbine types by means of least squares fits the following approximation formula for the cutoff frequency  $f_{\text{ARMA},j}$  of the low-pass filter was derived

$$f_{\text{ARMA},j} = \frac{1}{N_{m,j}} \sum_{i \in j} \left[ \begin{array}{ll} 10^{0.7 \log_{10}(f_{c,i}) - 1.5} \text{ Hz}, & f_{c,i} < 1.6 \text{ kHz} \\ 5 \text{ Hz}, & f_{c,i} \geq 1.6 \text{ kHz} \end{array} \right] \quad (3.8)$$

with  $N_{m,j}$  being the number of 1/3 octave bands in group  $j$ .

### 3.2.2. Signal Analysis

The basis for the signal analysis is a measured sound pressure signal,  $p(t)$ , (see section 3.2.3) with sampling frequency  $f_s = 44.1$  kHz. On one hand it should be guaranteed that the signal to be analyzed is sufficiently short in terms of constant signal characteristics, e.g. no impulsive sounds, constant blade passing frequency, etc. On the other hand the signal should be long enough to ensure stable statistics. A signal length of around 20 seconds seems to be an appropriate compromise. The signal analysis is structured into four steps:

1. tonality analysis
2. spectral analysis
3. periodic AM and
4. stochastic AM.

#### Tonality analysis

In order to detect possible tonal components, a narrow-band frequency analysis is performed. The power spectral density (PSD) of  $p(t)$  is estimated by using Welch's method (Hann window, 50% overlap, 16 k FFT,  $\Delta f = 2.7$  Hz). In the PSD all local maxima inside a defined band (typically 100 Hz to 5 kHz) are searched. Each local maximum exceeding the arithmetic mean level inside the critical band centered around the detected frequency by more than 4 dB is considered a tone. For each tone the tone level is evaluated by integrating the PSD over a range of 10 Hz around the tone frequency.

In order to suppress the detected tones for the following analysis, a notch filter bank is created. For each tone a 4th order Butterworth band-stop filter at the tone frequency is designed. For frequencies below 1 kHz a relative bandwidth of 5% of the center frequency is applied, whereas for frequencies above 1 kHz a fixed bandwidth of 50 Hz is used. The notch filter bank is then applied to  $p(t)$  resulting in signal  $p_{\text{noise}}(t)$ .

#### Spectral analysis

Following the concept of sub-band coding, the signal  $p_{\text{noise}}(t)$  is decomposed into 1/3 octave bands by applying a 8th order Butterworth bandpass filter bank with center frequencies,  $f_{c,i}$ , from 20 Hz to 10 kHz. Together with the mean levels, the presented synthesis model relies on detailed information about the level fluctuations per 1/3 octave band. Therefore for every 1/3 octave band signal,  $p_{\text{noise},i}(t)$ , the time-weighted sound pressure level

$$L_i(t) = 10 \log_{10} \left( \frac{1}{\tau_i} \int_{-\infty}^t \frac{p_{\text{noise},i}^2(t')}{p_0^2} e^{\frac{t'-t}{\tau_i}} dt' \right) + C_{\text{notch},i} \text{ [dB]} \quad (3.9)$$

with frequency depending time constant

$$\tau_i = \frac{\tilde{a}}{f_{c,i}} \text{ [s]} \quad (3.10)$$

is calculated. Based on listening tests of resynthesized wind turbine sounds, the smoothing parameter  $\tilde{a}$  is chosen as 20. This choice of  $\tilde{a}$  yields a time constant of 20 ms at 1 kHz. The correction term  $C_{\text{notch},i}$  compensates for the influence of the notch filter bank on the broadband noise part of  $p(t)$ .  $C_{\text{notch},i}$  is determined by evaluating the 1/3 octave band level attenuation of the notch filter bank applied to pink noise.

The resulting 1/3 octave band level curves,  $L_i(t)$ , are then downsampled to a sampling frequency  $f_{s,L} = 30$  Hz which was again validated by subjective assessments of resynthesized signals. The arithmetic mean noise level per 1/3 octave band is given by

$$\bar{L}_i = \frac{1}{T - t_0} \int_{t_0}^T L_i(t) dt \quad (3.11)$$

where the integration time offset  $t_0$  is set to 2 seconds in order to avoid edge effects at low frequencies.

### Level fluctuations

Generally, the level fluctuations are obtained by subtracting the DC offset  $\bar{L}_i$  from  $L_i(t)$ . However, together with perceptible level variations, wind turbine sounds also exhibit imperceptible slow level variations which are not intended to be covered by the synthesis model. Therefore after removal of the DC offset, the signals are high pass filtered with a cutoff frequency of 0.1 Hz to obtain the level fluctuations  $F_i(t)$ .

Section 3.2.1 suggests that the level fluctuations are composed of two independent additive terms

$$F_i(t) = F_{\text{periodic},i}(t) + F_{\text{stochastic},i}(t), \quad (3.12)$$

namely a periodic function  $F_{\text{periodic},i}(t)$  with period  $T_{BP} = 1/f_{BP}$  and the blade passing frequency  $f_{BP}$ , and a stochastic term  $F_{\text{stochastic},i}(t)$ . Both functions are assumed to have zero mean.

### Estimation of the blade passing frequency

The acoustical estimation of the blade passing frequency  $f_{BP}$  from level fluctuations can be understood as a pitch detection task which can generally be performed in either the time or the frequency domain. Following other authors [23–26] initial attempts were made in the frequency domain by searching for local maxima in high resolution PSDs. However, it turned out that a time domain based method with the

autocorrelation function yielded more robust results. The autocorrelation function (ACF) for a continuous function  $u(t)$  is defined as

$$R_u(\kappa) \equiv \overline{u(t)u(t + \kappa)} \quad (3.13)$$

with the time lag  $\kappa$ . The application of the ACF to a periodic function preserves its periodicity, whereas the ACF of a noise signal is maximum at  $\kappa = 0$  and steadily declines to both sides. The ACF of  $F_i(t)$  therefore shows local maxima at  $\kappa = lT_{BP}$  with  $l$  an integer. This property is now used to estimate the blade passing frequency from the level fluctuations  $F_i(t)$ . For this, by using a priori knowledge, a search range for the blade passing frequency is defined, i.e.  $\hat{f}_{BP} \in [f_l \dots f_h]$  with e.g.  $f_l = 0.5$  Hz and  $f_h = 1.5$  Hz. Then the 1/3 octave band,  $i'$ , with the largest local maximum of the unbiased ACF within the search range  $1/f_h < \kappa < 1/f_l$  is selected and used to estimate the blade passing frequency  $\hat{f}_{BP} = 1/\hat{T}_{BP}$ , where  $\hat{T}_{BP}$  is the time lag of the largest local maximum within the search range.

### Standard deviation of the periodic AM

The standard deviation,  $s_u$ , of a zero mean signal,  $u(t)$ , is related to the signal power  $P_u$  by

$$s_u^2 = P_u \equiv \overline{u^2(t)} = R_u(0). \quad (3.14)$$

In order to determine the standard deviation,  $s_{\text{periodic},i}$ , of the periodic AM, initial attempts were made by using the PSD. However, it appeared that the spectral separation of periodic and stochastic parts is critical, particularly as harmonics of the periodic part are present and in some cases the stochastic component dominates. A more robust method was found by using the ACF as well

$$s_{\text{periodic},i}^2 = \overline{F_{\text{periodic},i}^2(t)} = R_{F_i} \left( \frac{l}{\hat{f}_{BP}} \right) \quad (3.15)$$

with  $l$  being an integer except 0. In the implementation the unbiased ACF is used and, in order to obtain reliable results also at low frequencies,  $s_{\text{periodic},i}$  is evaluated as an average for  $l = 1, 2$  and 3.

### Stochastic AM

The standard deviation,  $s_{\text{stochastic},i}$ , of the stochastic AM is calculated by

$$s_{\text{stochastic},i}^2 = s_{\text{tot},i}^2 - s_{\text{periodic},i}^2 - s_{\text{pinknoise},i}^2 \quad (3.16)$$

where  $s_{\text{tot},i}$  denotes the total standard deviation of  $F_i$ ,  $s_{\text{periodic},i}$  the standard deviation of periodic level fluctuations from Equation (3.15) and  $s_{\text{pinknoise},i}$  is the standard deviation of level fluctuations of a reference pink noise signal which is processed in the same manner as the wind turbine signal.

By looking at measured level fluctuations it is apparent that there exists a coherence of the stochastic AM between 1/3 octave bands, especially between high frequency bands. Experiments with synthesized sounds have revealed that the

implementation of this finding is very important for a realistic synthesis. The coherence of the stochastic AM between frequency bands can objectively be investigated by using the cross-correlation function (CCF) which for two continuous functions  $u(t)$  and  $\tilde{u}(t)$  is defined as

$$R_{u\tilde{u}}(\kappa) \equiv \overline{u(t)\tilde{u}(t + \kappa)}. \quad (3.17)$$

As a measure of the coherence between the stochastic AM of two 1/3 octave bands  $i$  and  $j$ , the following expression for the correlation coefficient was deduced

$$\begin{aligned} \rho_{ij} &\equiv \frac{\text{Cov}(F_{\text{stochastic},i}, F_{\text{stochastic},j})}{\sqrt{\text{Var}(F_{\text{stochastic},i})\text{Var}(F_{\text{stochastic},j})}} \\ &= \frac{R_{F_{\text{stochastic},i}F_{\text{stochastic},j}}(0)}{\sqrt{R_{F_{\text{stochastic},i}}(0)R_{F_{\text{stochastic},j}}(0)}} \\ &= \frac{R_{F_iF_j}(0) - R_{F_iF_j}\left(\frac{1}{f_{BP}}\right)}{\sqrt{\left[S_{\text{tot},i}^2 - S_{\text{periodic},i}^2\right]\left[S_{\text{tot},j}^2 - S_{\text{periodic},j}^2\right]}}. \end{aligned} \quad (3.18)$$

Now for every 1/3 octave band  $i$  the number  $N_{n,i}$  of adjacent 1/3 bands with  $\rho_{ij} > 0.5$  are counted. The vector  $N_{n,i}$  is then used to build groups of 1/3 octave bands being modulated by the same stochastic AM function.

The spectral content of the level fluctuations is investigated by applying a 1/3 octave band filter bank to  $F_i(t)$  prior to the calculation of the standard deviations by Equation (3.16). Modulation frequency spectra of the stochastic component are evaluated for a typical range of the modulation 1/3 octave band center frequencies from  $f_{c,m} = 0.1$  to 10 Hz, and with  $S_{\text{periodic},i,m}$  being the modulation frequency spectrum of  $F_{\text{periodic},i}$ , which is calculated based on the assumption from Equation (3.6).

### Compensation of sound propagation

As the above extracted levels are immission levels, they have to be corrected for sound propagation effects in order to provide a proper interface to the sound propagation simulation described in section 3.3. Thus, the tone levels and the mean noise levels are converted to a reference distance  $d_0 = 1$  m. The propagation effects geometrical spreading, ground effect and atmospheric absorption are taken into account:

$$L_e = L_{\text{meas}} + 20 \log_{10} \left( \frac{D_1}{d_0} \right) + A_{\text{gr}} + A_{\text{atm}} \quad (3.19)$$

with  $L_{\text{meas}}$  being the extracted sound pressure immission levels,  $D_1$  the distance between microphone and the source,  $A_{\text{gr}}$  the ground effect and  $A_{\text{atm}}$  the atmospheric attenuation.  $D_1$  is approximated by

Identifier	WT1	WT2
Type	Vestas V90-2.0 MW	Enercon E82-2.0 MW
Year of completion	2010	2009
Location	Mont Crosin (CH)	Saint-Brais (CH)
Number of blades	3	3
Hub height [m]	95	78
Rotor diameter [m]	90	82
Measurement distance $D_0$ [m]	140	119
Nominal revolutions [rpm]	14.9	variable
Cut-in wind speed [m/s]	4	none

Table 3.1: Data of the investigated wind turbines

$$D_1 \cong \sqrt{D_0^2 + H_{\text{hub}}^2} \quad (3.20)$$

with the hub height  $H_{\text{hub}}$  and the horizontal distance  $D_0$  between microphone and the wind turbine tower. The measurements were performed on a hard plate according to [16], resulting in a ground effect  $A_{\text{gr}} = -6$  dB.  $A_{\text{atm}}$  is calculated according to the standard ISO 9613-1 [27] with measured temperature and humidity for the tone frequencies and the center frequencies of the 1/3 octave bands, respectively.

### 3.2.3. Application to measurements

In this section the proposed signal processing algorithms are exemplarily illustrated by applying them to two recordings from two different 2.0 MW wind turbines denoted as WT1 and WT2 (see Table 3.1). On September 7, 2011 at 12 h the recording of WT1 was taken at the wind farm on Mont Crosin, Switzerland, with temperature 9 °C and relative humidity 94%. The recording of WT2 was performed at the wind farm in Saint-Brais, Switzerland, on July 23, 2012 at 21 h with temperature 10 °C and relative humidity 80%. Both recordings were taken at strong wind conditions with the turbines operating at nominal power. The sound propagation occurred over flat grassy ground and the microphones were placed on a hard plate of diameter 0.6 m at horizontal distance  $D_0$  (given in Table 3.1) to the turbines. For WT1 the microphone was mounted in downwind direction (position no. 1 according to the standard IEC 61400-11 [16]) and for WT2 in upwind direction (position no. 3 according to [16]). The equipment used consisted of a Sound Devices Two-Track Audio Recorder 702T and a B&K omnidirectional microphone 4006. Weather information was obtained from a nearby ground-based automatic weather stations operated by MeteoSwiss. The recorded audio files are provided in appendix A and at <http://www.visasim.ethz.ch/auralization> [28].

### Spectra

Figure 3.2 shows power spectral densities of sound pressure immission signals of two wind turbines at strong wind conditions. In contrast to WT2, for WT1 three tones have been detected. The corresponding extracted tone levels are indicated

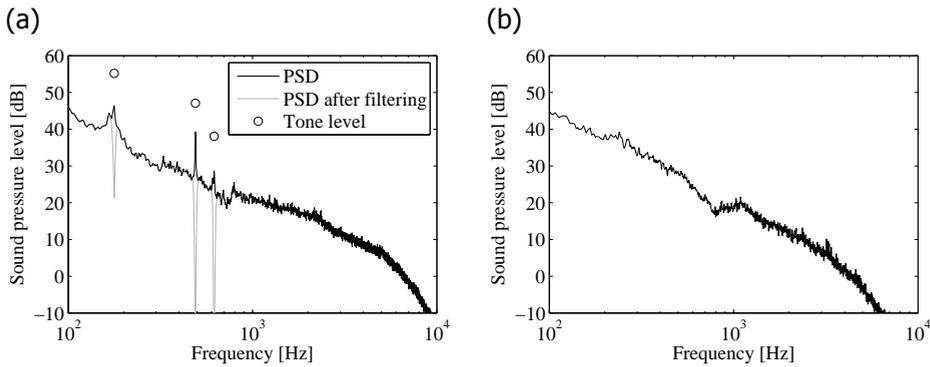


Figure 3.2: Power spectral densities of measured sound pressure signals before (black) and after (grey) applying the notch filter bank and thereof extracted tone levels (circles) of detected tones for WT1 (a) and WT2 (b) at strong wind conditions.

by circles. For WT1 also the power spectral density after applying the notch filter bank is drawn in grey.

Figure 3.3 shows emission spectra,  $\overline{L_{e,ir}}$  in 1/3 octave bands for WT1 and WT2. They were calculated using Equations (3.11) and (3.19).

### Level fluctuations

Figure 3.4 shows level fluctuations,  $F_i(t)$ , of WT1 and WT2 for five 1/3 octave bands. In all plotted bands of WT2 clearly the periodic AM can be observed. In each band the differences of the maximum and minimum level amount to about 10 dB. The sound subjectively exhibits high frequency swishing as well as thumping sound.

### Blade passing frequency

The proposed algorithm to acoustically estimate the blade passing frequency was applied to measured sound pressure signals and compared to corresponding, optically determined rotational speeds. The algorithm proofed to be able to robustly estimate the blade passing frequency even in cases where it could not be subjectively detected by listening to the sounds.

Figure 3.5 shows the autocorrelation functions of the level fluctuations for the 1/3 octave band which was selected by the proposed algorithm to thereof estimate the blade passing frequency. Both curves clearly feature the periodical character of the level fluctuations. In case (a) the 6.3 kHz 1/3 octave band, containing audible swishing sound, was selected. From this curve the blade passing frequency 0.75 Hz, corresponding to 15 rpm, was estimated. In case (b) the 315 Hz band, containing the typical thumping sound, was selected and revealed a blade passing frequency of 0.81 Hz, corresponding to 16 rpm.

### Amplitude modulation

Based on the estimated blade passing frequency,  $\hat{f}_{BP}$ , the separation of periodic and stochastic AM is performed with help of the proposed algorithm (Equations (3.15)

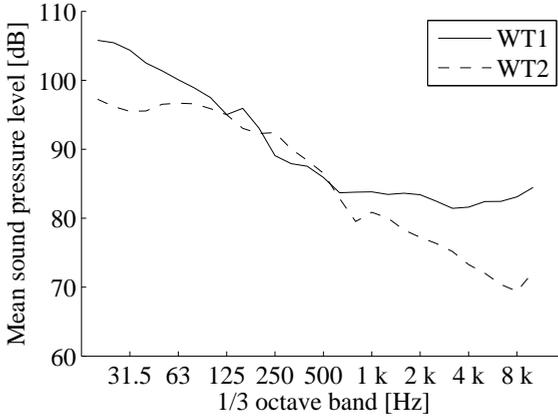


Figure 3.3: Arithmetic mean sound pressure emission spectra at 1 m,  $\overline{L_{e,ir}}$ , for wind turbine WT1 and WT2 at strong wind conditions.

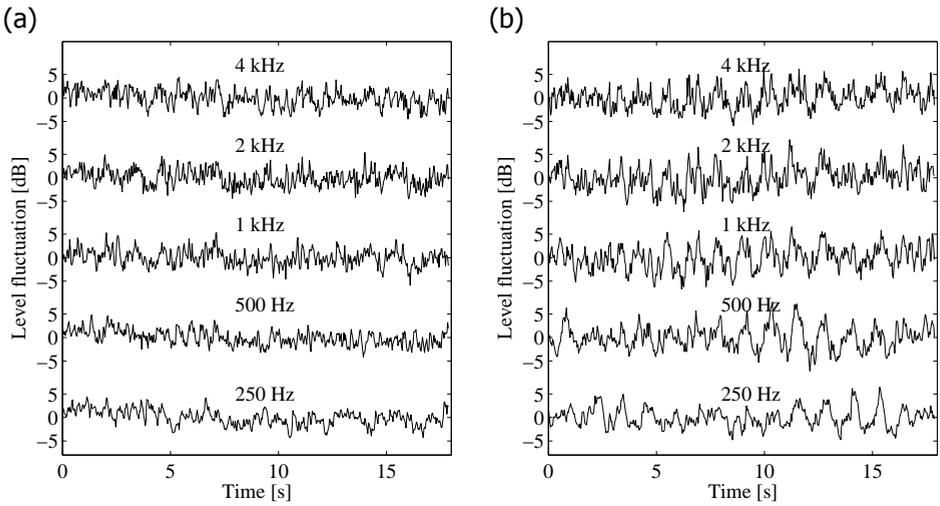


Figure 3.4: Level fluctuations of five 1/3 octave bands for WT1 (a) and WT2 (b) at strong wind conditions.

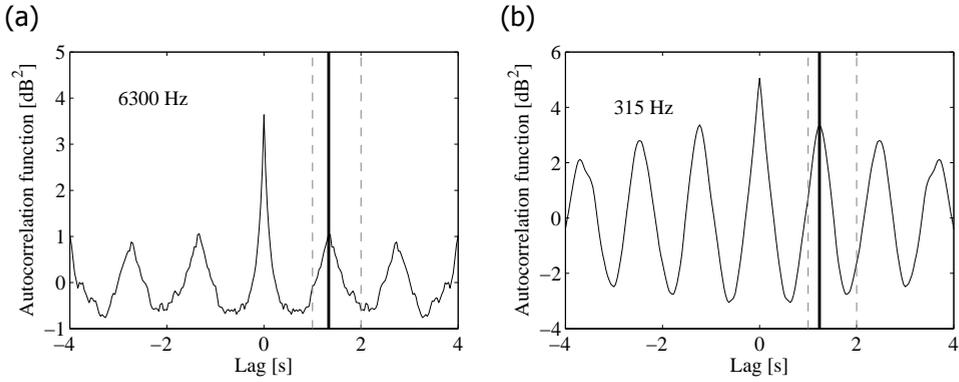


Figure 3.5: Autocorrelation function of level fluctuations with thereof estimated blade passing frequency (vertical black line) and its search range (dashed lines) of WT1 (a) and WT2 (b) at strong wind conditions.

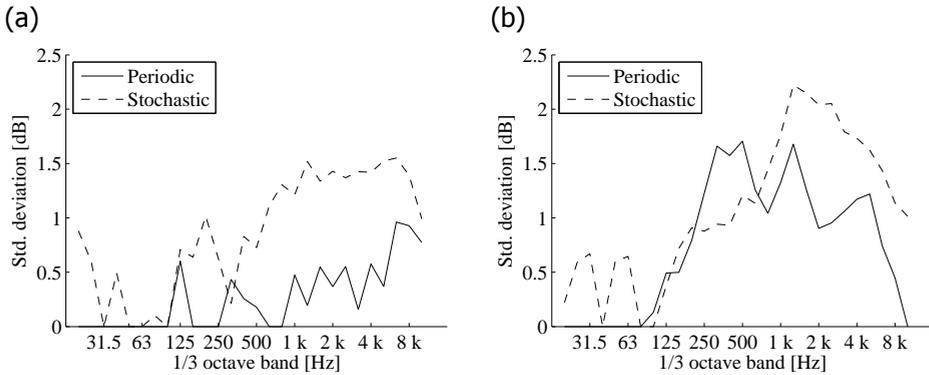


Figure 3.6: Standard deviations of periodic (black) and stochastic level fluctuations (or AM) of WT1 (a) and WT2 (b) at strong wind conditions.

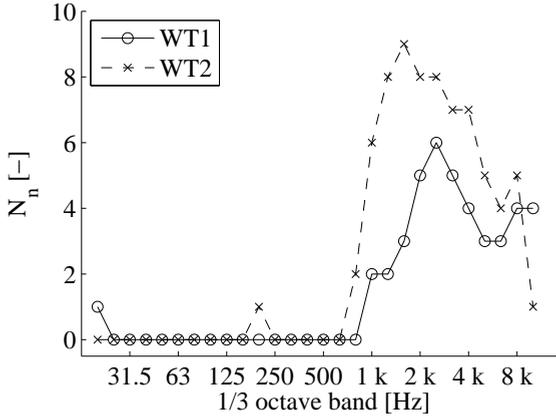


Figure 3.7: Number  $N_{n,i}$  of adjacent 1/3 octave bands with correlation  $\rho_{ij} > 0.5$  of WT1 and WT2 at strong wind conditions.

and (3.16)). The standard deviations of the level fluctuations are shown in Figure 3.6. In both cases at frequencies below 125 Hz no significant level fluctuations occur. At higher frequencies the stochastic AM rises up to  $\approx 1.5$  dB. For WT1 the stochastic AM dominates over the whole frequency range. But it has to be mentioned that a high frequent swishing sound is perceived clearly, which corresponds to the fact that the periodic AM has its maximum standard deviation of 1 dB at 6.3 kHz. Compared to WT1, WT2 shows considerably higher values of the periodic AM. Around 250–500 Hz the periodic AM even dominates and reaches standard deviations  $> 1.5$  dB, which by assuming a triangle waveform of the periodic level fluctuations amounts to differences of maximum and minimum levels  $> 1.5 \times 2 \times \sqrt{3}$  dB = 5.2 dB. This part of the AM corresponds to the thumping sound and is perceived also at greater distances (e.g. at 500 m). In [11] the maximum fluctuation strength of five 0.6 to 2 MW turbines was measured in the frequency range 350–700 Hz which confirms our findings related to WT2.

As an indicator for the coherence of the stochastic AM between frequency bands, the number  $N_n$  of adjacent 1/3 octave bands are counted as described in section 3.2.2. Figure 3.7 exhibits that the stochastic AM for frequency bands below 1 kHz can be modeled independently for each 1/3 octave band. However, above 1 kHz large groups of approximately 5 bands with simultaneous stochastic AM occur. The ten third octave bands above 1 kHz can thus be modeled by just two frequency groups.

Figure 3.8 shows modulation frequency spectra of stochastic AM for 1 kHz and 4 kHz. The maxima of all curves are located around 4 Hz. This coincides with the frequency for which the auditory system is most sensitive to level fluctuations [29]. Lower 1/3 octave bands exhibit a decreasing tendency of the maximum modulation frequency band.

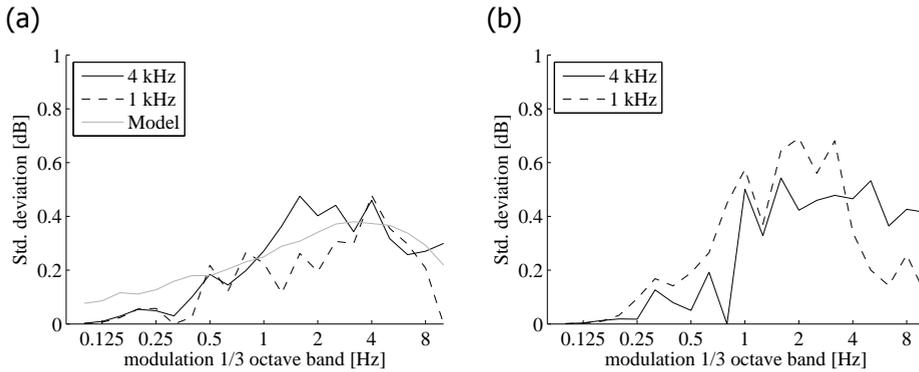


Figure 3.8: Standard deviation of the stochastic level fluctuations in modulation frequency 1/3 octave bands for the 1/3 octave bands 1 and 4 kHz and for WT1 (a) and WT2 (b) at strong wind conditions. The grey curve shows the modulation spectrum of a simulated level fluctuation function generated by an ARMA model ( $f_{\text{ARMA}} = 4 \text{ Hz}$ ) described in section 3.2.1.

### 3.2.4. Discussion

#### Quality of the synthesized sound

The quality of the proposed wind turbine emission synthesis model was evaluated by a test with 12 experienced listeners (11 male, 1 female, age between 25 and 55). For that purpose, four audio files of 8 seconds length were prepared. Two files were the recorded sounds of WT1 and WT2 which were analyzed in section 3.2.3 and two files were the corresponding synthesized sounds created by the proposed model. In order to remove a high frequency hiss sound caused by the recording equipment, 10 kHz low-pass filtering was applied to the recorded sounds. The files were presented monophonically by a single frontal loudspeaker in an anechoic environment at a listening level of 55 dB(A). This corresponds to a realistic sound pressure level in a distance of approximately 150 m from the turbine.

The course of the listening test was controlled exclusively by the listener. At the beginning, the test person had the choice to listen to the recordings of WT1 and WT2 as many times as desired. After feeling familiar with the sounds, the listener could initiate the real test. Hereby each of the four files (2 recordings, 2 synthesized sounds) was played twice in a random sequence. Immediately after the presentation of a file, the listener had to identify the sound as “original recording” or “synthesized sound”. As a third option “I don’t know” could be selected.

Each of the 12 listeners generated 8 answers, summing up to a total of 96 responses. 8 out of the 96 were “I don’t know”. In case of the recording of WT1, 18 out of 24 answers were correct, in case of the synthesized sound of WT1, 8 out of 24 answers were correct. The recording of WT2 was correctly identified 13 times, the synthesized sound of WT2 was correctly identified 17 times. On the other hand, wrong answers were obtained: 4 in case of the recording of WT1, 14 for the synthesized sound of WT1, 9 for the recording of WT2 and 5 for the synthesized sound of WT2. In total, 54% correct answers were given for WT1 and 65% correct answers for WT2. On the other hand, 38% of the answers were wrong in case

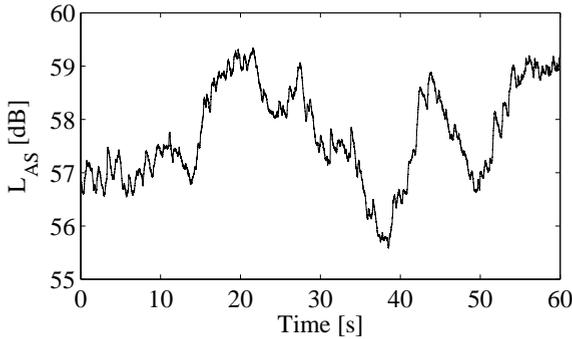


Figure 3.9: A-weighted sound pressure immission level with time constant SLOW of WT1 measured at position no. 1 according to [16] at strong wind conditions. The standard deviation of this 60 seconds period amounts to 0.9 dB.

of WT1 and 29% in case of WT2. The somewhat higher percentage of correct identification in case of WT2 may be explained by the sound of some cow bells contained in the recording of WT2. This may have helped to discriminate between recording and synthesized sound.

Given the fact, that with 54% and 65% correct answers the percentage of proper discrimination between recorded and synthesized sound is only slightly above the result obtained by arbitrary response, and taking into account that at least a few correct decisions were based on the observation of background noise contained in the recording only, the quality of the synthesized sounds can be rated as very good. Indeed, all listeners assessed the synthesized sounds as very plausible. The interested reader is referred to appendix A where the four audio files are provided for individual listening tests. Remark: The recorded sounds still contain the above mentioned high frequency hiss sound caused by the recording equipment.

### Temporal variation of sounds and model parameters

In section 3.2.3 the signal analysis described in section 3.2.2 is exemplarily applied to two recordings in order to obtain the input parameters of the emission model. However, it turned out to be difficult to obtain representative model parameters to generally characterize the sound of the wind turbines. We observed that for identical conditions—from our limited perspective by observing 10 minute averages of horizontal wind speeds only—the emission sounds and thus the model parameters may differ significantly.

Figure 3.9 demonstrates that the sound pressure level considerably fluctuates on a short-term basis. Measurements of WT1 at strong wind conditions revealed standard deviations of the arithmetic mean 1/3 octave band levels,  $\bar{L}_i$ , as well as of the tone levels,  $L_{t,i}$ , of 2–3 dB. At moderate wind speeds even larger variations for some tonal components were observed. However at steady rotational speed the tone frequencies stayed fairly constant.

The measurements confirmed that the depth of periodic AM typically varies a lot,

which was already reported by other authors [18, 24, 30]. For WT1 the standard deviation,  $s_{\text{periodic},i}$  of the periodic level fluctuations exhibited standard deviations of 0.2 dB at mid and 0.3 dB at high frequencies which is in the same order of magnitude as their mean values 0.2 and 0.4 dB, respectively. For WT2, where the thumping sound was present, even larger variations occurred. Within 10 minutes, standard deviations of 0.3 to 0.5 dB at mean values between 0.5 and 1.2 dB were measured. E.g. for the 500 Hz 1/3 octave band, compared to 1.7 dB as shown in Figure 3.6, within the same 10 minutes period also a value of only 0.6 dB was measured. This involves a variation of the difference between maximum and minimum level of the periodic AM of 2 to 6 dB within 10 minutes, which is a well audible change. The model parameters regarding the stochastic AM generally showed minor variations compared to their absolute values.

The temporal variation of the model parameters that were found in the measurements lie within the audible range. They have to be regarded as an inevitable peculiarity of wind turbine sounds.

### 3.3. Propagation filtering

Sound propagation from wind turbines to typical listener positions has the distinction of highly elevated sources and large distances to the receiver. The height of the source strongly reduces the sensitivity of the propagation attenuation with respect to meteorological conditions [31, 32]. In our simulations we therefore ignore a possible curvature of propagation paths due to vertical gradients of the effective speed of sound and assume an attenuation independent of the condition of the atmosphere. However, as wind turbine noise occurs during periods with relative high wind speeds, a simulation of the propagation has to consider atmospheric turbulence which corresponds to a non-stationary, fluctuating channel. A realistic listening impression at a receiver position has to be composed of the signal of the turbines and possible environmental noise contributions. Indeed in many situations vegetation noise plays an important role as it grows with increasing wind speed and thus may mask turbine noise at high and very high wind speeds [25, 33, 34].

In the following sections the propagation filtering of wind turbine noise is discussed in detail for each propagation effect. Hereby the ground effect and fluctuations of the atmosphere that lead to receiver level variations have been investigated extensively. Based on measurements and numerical simulations it can be concluded that a wind turbine has to be modeled as a largely extended source. The point source assumption would significantly overestimate interference effects and level variations.

#### 3.3.1. Propagation effects and representation by digital filters

The emission synthesizer from section 3.2.1 delivers a sound pressure signal  $p_{e,i}$  for each source  $i$ . This signal is scaled as sound pressure in  $d_0 = 1$  m distance, assuming that it is radiated by a point source. The propagation effects from the source to the receiver can be understood as a filter that converts the emission signal  $p_{e,i}$  into a signal  $p_{r,i,j}$  that corresponds to the sound pressure at the receiver position

*j*. The receiver signal  $p_{r,i,j}(t')$  as a function of the receiver time  $t'$  can be written as superposition of direct and ground reflected sound as

$$p_{r,i,j}(t') = \mathcal{F}_{0,i,j}^t \{ \mathcal{F}_{1,i,j} [p_{e,i}(t) + \mathcal{F}_{2,i,j}(p_{e,i}(t))] \} \quad (3.21)$$

where  $\mathcal{F}_2$  is the filter function that represents the reflection characteristics of the ground and the additional delay as a consequence of the longer propagation path compared to the direct sound.  $\mathcal{F}_1$  is the filter function that simulates geometrical spreading, air absorption and possible further attenuation effects introduced by barriers and foliage.  $\mathcal{F}_0^t$  finally is a power-neutral filter function that varies randomly over time to simulate fluctuations due to atmospheric turbulence.  $\Delta t$  corresponds to the propagation delay from source to receiver. In the general case for a moving receiver,  $\Delta t$  and the filter functions  $\mathcal{F}_0$ ,  $\mathcal{F}_1$  and  $\mathcal{F}_2$  vary over time  $t$ .

### 3.3.2. Doppler frequency shift due to moving receiver

If the position of the receiver varies over time, the mapping of the emission signal time axis onto the receiver time axis becomes time-dependent. For a given sound travel time  $\Delta t(t)$  from source to receiver, the emission time  $t$  is related to the receiver time  $t'$  by

$$t = t' - \Delta t(t) \quad (3.22)$$

The index  $k_s$  of the emission signal  $p_e$  that is mapped onto the receiver sample with index  $k'$  is given as

$$k_s = k' - \Delta t(t)f_s \quad (3.23)$$

where  $f_s$  is the sampling frequency. In general,  $k_s$  obtained by Equation (3.23) is not an integer and therefore not directly usable as sample index. In order to avoid discontinuities for varying  $\Delta t$ , an interpolation operation is needed that allows to estimate an emission signal value at arbitrary points in time. The emission signal value  $p_e$  at non-integer index  $k_s$  is determined here as

$$p_{e,i}(k_s) = (1 - \tilde{\delta})p_{e,i}[\lfloor k_s \rfloor] + \tilde{\delta} p_{e,i}[\lfloor k_s \rfloor + 1] \quad (3.24)$$

with the fractional delay part  $\tilde{\delta} = k_s - \lfloor k_s \rfloor$  and the floor function  $\lfloor \cdot \rfloor$ .

The linear interpolation in Equation (3.24) corresponds to a low-pass filter operation and thus introduces an amplitude error at high frequencies. The upper limiting frequency  $f_u$  (corresponding to the  $-3$  dB point) of the low-pass filter is lowest for  $\tilde{\delta} = 0.5$ . With help of numerical simulations, an average value of 14.9 kHz was found for  $f_{iu}$ , assuming a sampling frequency  $f_s$  of 44.1 kHz. In the context of wind turbines, this high frequency error is acceptable as frequencies above 10 kHz play a minor role at distances of several hundred meters. One reason is air absorption that introduces significant high frequency damping and a principal uncertainty due to its high sensitivity with respect to temperature and humidity. At 200 m distance and 14.9 kHz, air absorption varies by 1 dB/°C temperature change and by 0.5 dB/% change in humidity according to ISO 9613-1 [27].

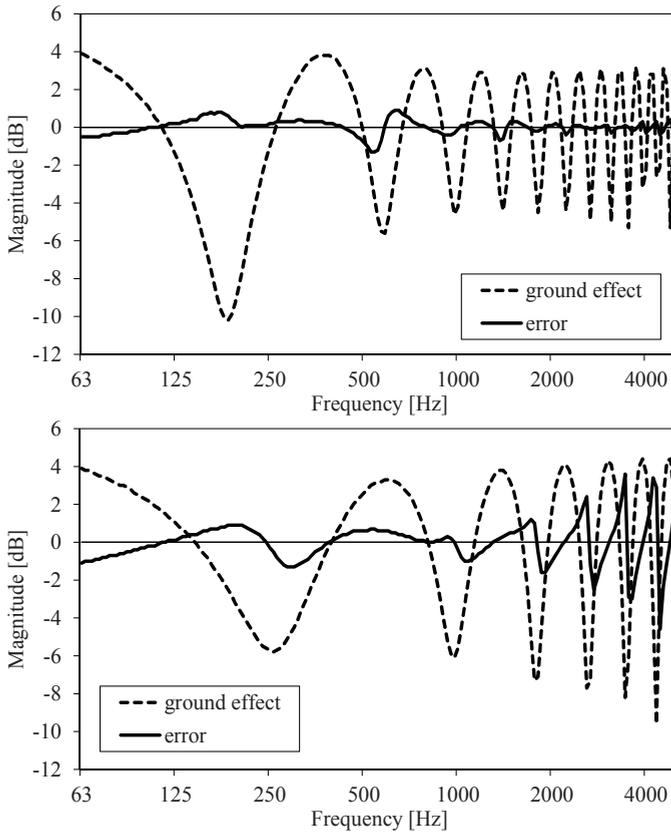


Figure 3.10: Ground effect for grassland (flow resistivity  $20'000 \text{ kPa s m}^{-2}$ ) for a source–receiver distance of 500 m (top) and 1000 m (bottom). The point source is assumed 100 m, the receiver 2 m above ground. The dashed line refers to the exact calculation, the solid line shows the error for a FIR simulation with  $n = 40$  taps.

### 3.3.3. Geometrical spreading

The propagation effect of geometrical spreading is the simplest one to simulate. Depending on the source–receiver distance, a frequency independent scaling of the emission audio signal has to be applied.

### 3.3.4. Ground effect

#### Simulation by a FIR filter

Ground effect is modeled here according to the underlying physical mechanism as interference between direct and ground reflected sound (Equation (3.21)). As a simplification it is assumed that the reflection occurs at an infinitely extended plane defined by the topography at the receiver position. The ground reflection differs from the direct sound by scaling with a complex reflection factor and an additional delay. The ground reflected sound is represented as  $\mathcal{F}_2(p_{e,i}[k])$  with the filter

function  $\mathcal{F}_2$ :

$$\mathcal{F}_2(p_{e,i}[k]) = \frac{d_{\text{dir}}}{d_{\text{gr}}} (p_{e,i} * h_{\text{swrc}} * h_{\text{delay}})[k] \quad (3.25)$$

where  $d_{\text{dir}}$  is the source–receiver distance,  $d_{\text{gr}}$  is the distance source–ground reflection point–receiver,  $h_{\text{swrc}}$  is the impulse response of the spherical wave reflection coefficient  $Q$  [35].  $*$  denotes the discrete linear convolution operation. The filter

$$h_{\text{delay}}[k] = \delta[k - l] \quad (3.26)$$

with  $\delta$  being the Kronecker delta function realizing a delay by  $l$  samples with

$$l = \left\lfloor \frac{(d_{\text{gr}} - d_{\text{dir}})f_s}{c_0} \right\rfloor \quad (3.27)$$

where  $c_0$  is the speed of sound and  $\lfloor \cdot \rfloor$  the rounding function to the nearest integer.

As indicated in Equation (3.25), the linear distortion by the spherical wave reflection coefficient  $Q$  is modeled by a digital filter  $h_{\text{swrc}}$ .  $Q$  depends on the ground type, the reflection angle  $\psi$ , the propagation distance  $d_{\text{gr}}$  and frequency  $f$ .  $Q$  is derived from the plane wave reflection coefficient

$$R_p = \frac{\sin \psi - \frac{1}{Z'}}{\sin \psi + \frac{1}{Z'}} \quad (3.28)$$

by [35, 36]

$$Q = R_p + (1 - R_p)F(w) \quad (3.29)$$

with  $Z'$  being the acoustic impedance of the ground normalized to the impedance of air,  $\rho c_0$ , and  $\psi$  being the angle between the ground plane and the ground reflected path ( $\psi = \pi/2$  for normal sound incidence). The term  $F(w)$  in Equation (3.29) is given by [36]

$$F(w) = 1 + j\sqrt{\pi}we^{-w^2} \operatorname{erfc}(-jw) = 1 + j\sqrt{\pi}w \operatorname{wofz}(w) \quad (3.30)$$

$$w = \frac{1 + j}{2} \sqrt{\frac{2\pi f d_{\text{gr}}}{c_0} \left( \sin \psi + \frac{1}{Z'} \right)} \quad (3.31)$$

with the imaginary unit  $j$ , the complex complementary error function  $\operatorname{erfc}$  and the Faddeeva function  $\operatorname{wofz}$ .  $h_{\text{swrc}}$  is realized here as a FIR filter where the coefficients are obtained by inverse Fourier transformation of  $Q(f)$  and subsequent limitation to a length of  $N$  taps. The FFT length is chosen as 64 k, corresponding to a frequency resolution  $\Delta f = 0.673$  Hz ( $f_s = 44.1$  kHz). It is found that for high elevated sources, already a relatively small number of taps is sufficient to reproduce the ground effect with acceptable accuracy. Figure 3.10 shows the ground effect for the propagation over soft ground (flow resistivity  $200$  kPa s  $\text{m}^{-2}$ ) between a point source  $100$  m

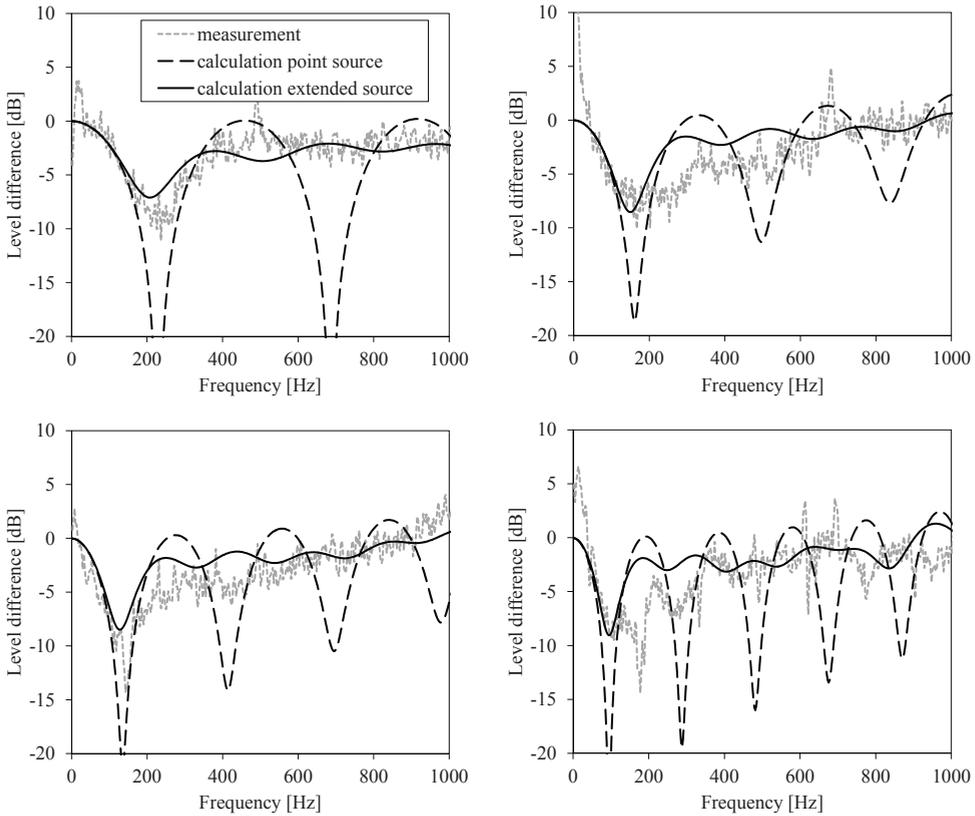


Figure 3.11: Measured and calculated (point source and extended source assumption) level differences of wind turbine noise evaluated as level differences for a microphone position 1.60 m above ground and close to the ground (0.05 m). The source–receiver distances were 400 m (top left), 300 m (top right), 250 m (bottom left) and 150 m (bottom right). The height of the nacelle was 95 m, the propagation was over grassy land.

above ground to a receiver 2.0 m above ground. The exact calculation and the simulation with a FIR filter of  $N = 40$  taps is shown. The accuracy of the FIR simulation is better than 1 dB for the lowest interference dips. The increase of the error at high frequencies in case of the propagation distance of 1000 m is not relevant in an auralization application, as these signal components become inaudible due to strong absorption by the atmosphere.

### Source extension

In order to validate the ground effect simulation, four measurements were performed at Mont Crosin, Switzerland over soft ground (grassland). The source was a Vestas V90-2.0 MW turbine (no. 11) with a hub height of 95 m (see Table 3.1). The receiver was positioned at different distances to the source, firstly close to the ground (0.05 m) and subsequently at a height of 1.60 m. The measurements were

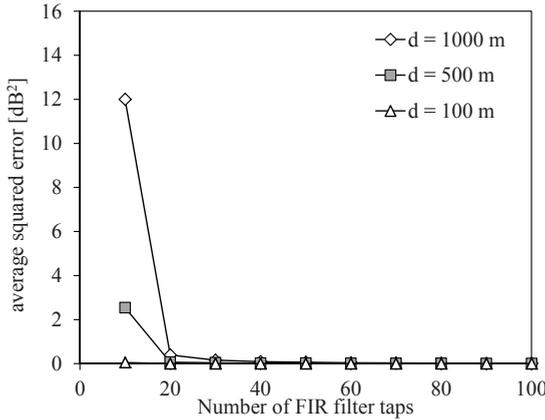


Figure 3.12: Average squared error of the air absorption simulation by an FIR filter of length  $N$ . The distance  $d$  between source and receiver was varied as indicated in the legend.

evaluated as narrow band spectrum differences between the position at 1.60 m and the 'close to ground' position. For the measured geometries, the corresponding level differences were calculated assuming a point source at the hub position. As can be seen in Figure 3.11 the point source assumption leads to pronounced interference patterns that are not confirmed by the measurements. The relatively weak ground effect dips in the measurements suggest an extension of the source. Indeed this is confirmed by observations of the emission of wind turbines by acoustical cameras [22]. In the following it is assumed that the source has a vertical extension that corresponds to the diameter of the rotor. Corresponding calculations with five different point sources distributed equidistantly in height reproduce the measurements in Figure 3.11 reasonably well.

### 3.3.5. Air absorption

The simulation of air absorption is realized by a linear phase FIR filter. The filter coefficients are found by inverse Fourier transformation of the air absorption spectrum. To obtain a linear phase filter, the imaginary part (or phase) of the absorption spectrum  $H_{\text{air}}(f)$  is set to 0:

$$\begin{aligned}\Re\{H_{\text{air}}(f)\} &= 10^{-A_{\text{atm}}(f)/20} \\ \Im\{H_{\text{air}}(f)\} &= 0\end{aligned}\quad (3.32)$$

where  $A_{\text{atm}}(f)$  in dB is the attenuation due to atmospheric absorption at frequency  $f$  according to ISO 9613 [27, 37]. The discrete-time impulse response  $h[k]$  is then determined as:

$$h_{\text{air}}[k] = \text{IFFT}\{H_{\text{air}}(f(n))\} \quad (3.33)$$

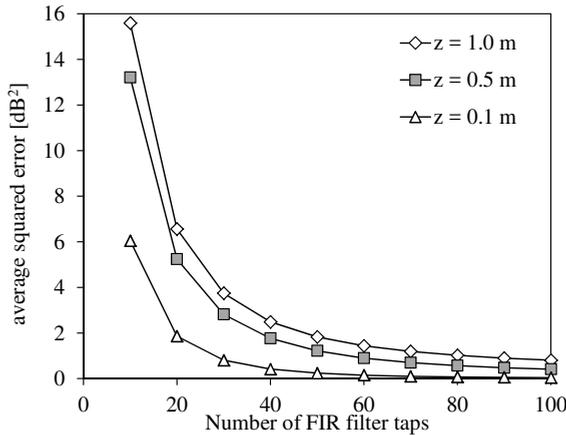


Figure 3.13: Average squared error of the barrier attenuation simulation by an FIR filter of length  $N$ . The path length difference  $z$  around the obstacle and through the obstacle was varied as indicated in the legend.

with the FFT bin frequencies  $n$  and where  $h_{\text{air}}[k]$  is symmetrical around  $k = 0$ .  $h_{\text{air}}[k]$  is symmetrically truncated to a finite length of  $N$  taps and shifted by  $N/2$  samples to obtain a causal filter.

The quality of the air absorption simulation by an FIR filter is expressed as squared difference between the exact solution and the FIR simulation of the air absorption in dB and averaged over the frequency range for which the attenuation due to air absorption is less than 30 dB. Figure 3.12 shows this error as a function of number of taps  $N$  for different source–receiver distances. Air temperature and humidity were set to 10°C and 60%.

As can be seen in Figure 3.12, the necessary number of taps to obtain a good model of the air absorption increases with increasing source–receiver distance. An FIR filter length of  $N = 20$  is sufficient to simulate air absorption with an accuracy better than 1 dB for distances up to 1000 m.

### 3.3.6. Attenuation by barriers and foliage

Similarly to the simulation of air absorption, attenuation by barriers and foliage is implemented as a linear phase FIR filter as well. Figure 3.13 shows the average squared error as a function of the number of taps for the barrier attenuation according to ISO 9613-2 [37] for different values of the path length difference  $z$  around the obstacle and through the obstacle. The error is defined as the difference between the exact solution and the FIR simulation of the barrier attenuation in dB and averaged over the frequency range from 50 Hz to 10 kHz. The necessary number of taps increases for increasing  $z$ . For strongly attenuating barriers ( $z \approx 1.0$ ), an FIR length of  $N = 80$  is needed for an accuracy better than 1 dB. Similar filter lengths have to be used to simulate attenuation by foliage.

### 3.3.7. Additional reflections

Besides the ground, additional reflecting objects or structures may influence sound propagation from source to receiver. If the line-of-sight is not interrupted, direct and ground reflected sound usually dominate. However if the source is shielded, possible reflections at buildings [38], forests [39] or cliffs [40] have to be considered.

### 3.3.8. Atmospheric fluctuations

During propagation over larger distances, significant sound pressure level and phase fluctuations occur due to temporally varying inhomogeneities of the atmosphere. The two relevant properties of the atmosphere are temperature and wind speed. Consequences of these fluctuations are level variations and a loss of coherence between direct and ground reflected sound. The decorrelation between the direct and ground reflected sound path is neglected here as an approximation. In the following, the level fluctuations are discussed in more detail.

#### Measurements of level fluctuations

During two days in September 2011, several measurements were taken at Mont Crosin in Switzerland. The equipment was made up of a Sound Devices 702T recorder and a B&K 4006 microphone. The signals of two Vestas V90-2.0 MW turbines (no. 11 and 12) with a hub height of 95 m were investigated at a microphone height of 1.60 m at different distances. The sound pressure signals were analyzed in third octave bands with help of a Norsonic Environmental Analyzer Nor121 and averaged over periods of 100 ms. From the sequence of these 100 ms values, a moving average over a time interval 5 s was calculated. This slowly varying time history served as reference. Finally the statistics of the level differences of the original 100 ms values and the reference was evaluated and expressed as a standard deviation. The evaluation of the signal power of a bandlimited signal over a finite time window is inevitably associated with an uncertainty that depends on the product of bandwidth and averaging time. Consequently the above described evaluation of the standard deviation can not be attributed exclusively to fluctuations. Figure 3.14 shows the standard deviations of the wind turbine noise measurements and for comparison the standard deviations obtained for stationary pink noise.

The evaluations show at low frequencies comparable standard deviations for the wind turbine noise and the pink noise reference. The turbulences begin to dominate only for frequencies above 500 Hz. Furthermore the measurements demonstrate that the standard deviations increase for increasing frequency, distance and wind speed.

#### Characterization of atmospheric turbulence

In the following, horizontal sound propagation is assumed. Therefore the momentary effective speed of sound  $c_{\text{eff}}$  can be written as

$$c_{\text{eff}} \equiv c(T) + v \quad (3.34)$$

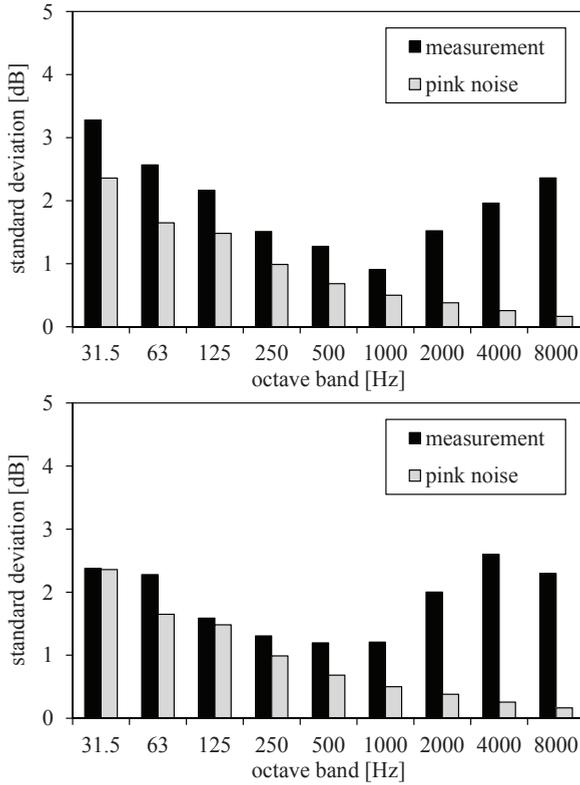


Figure 3.14: Standard deviations of measured wind turbine noise fluctuations 1.60 m above ground at distances 400 m at moderate wind (8 m/s at hub height) (top) and 250 m at strong wind (12 m/s at hub height) (bottom). As reference the standard deviation of stationary pink noise is shown.

where  $c(T)$  is the speed of sound at temperature  $T$  in still air, and  $v$  is the horizontal component of the wind speed vector parallel to the propagation direction from source to receiver.

The characterization of the fluctuations is usually based on the (acoustical) refractive index  $n$ . It can be defined as

$$n \equiv \frac{c_0 + v_0}{c_{\text{eff}}} \equiv 1 + \mu \quad (3.35)$$

where  $\mu$  is fluctuating acoustical index of refraction with  $|\mu| \ll 1$ ,  $c_{\text{eff}}$  is the momentary effective speed of sound,  $c_0$  is the reference speed of sound in a medium at rest at average temperature  $T_0$  and  $v_0$  is the average wind speed component in propagation direction.

The momentary temperature  $T$  and the momentary wind speed component  $v$  in propagation direction can be written as sum of the average values ( $T_0$  and  $v_0$ ) and a fluctuating part ( $\Delta T$  and  $\Delta v$ ):  $T = T_0 + \Delta T$  and  $v = v_0 + \Delta v$ . With the approximation of normally distributed fluctuations and under the assumption that temperature and wind speed fluctuations are not correlated, follows for the variance of the fluctuating acoustical index of refraction  $\mu$  [41]:

$$\langle \mu^2 \rangle \approx \frac{\sigma_v^2}{c_0^2} + \frac{\sigma_T^2}{4T_0^2} \quad (3.36)$$

where  $\sigma_T$  is the standard deviation of the temperature fluctuations and  $\sigma_v$  is the standard deviation of the longitudinal wind speed fluctuations.  $T_0$  is the average air temperature in Kelvin and  $c_0$  is the average speed of sound. Typical values for  $\langle \mu^2 \rangle$  lie between  $10^{-6}$  and  $2 \times 10^{-5}$ .

The simplest case that can be analyzed assumes homogeneous and isotropic turbulence [41–43]. For spherical wave propagation, the average squared log-amplitude fluctuation  $\langle \chi^2 \rangle$  and the average squared phase fluctuation  $\langle S^2 \rangle$  can be expressed as [41]

$$\langle \chi^2 \rangle \equiv \langle (\ln(A/A_0))^2 \rangle = \langle S^2 \rangle \equiv \langle (\phi - \phi_0)^2 \rangle = \frac{\sqrt{\pi}}{2} \langle \mu^2 \rangle \left( \frac{2\pi f}{c_0} \right)^2 d a \quad (3.37)$$

where  $A$  and  $\phi$  are the sound pressure amplitude and phase with turbulence and  $A_0$  and  $\phi_0$  are amplitude and phase without turbulence,  $f$  is frequency,  $c_0$  is speed of sound,  $d$  is the propagation distance and  $a$  is the correlation length. In the vicinity of the ground,  $a$  is in the range of typical 1 to 1.5 m.

The relation in Equation (3.37) expresses that the log-amplitude and phase fluctuations are equal and that they increase with increasing frequency, increasing propagation distance and increasing turbulence. The log-amplitude fluctuation  $\langle \chi^2 \rangle$  can be rewritten as a variance in dB  $\langle \chi'^2 \rangle$  as

$$\langle \chi'^2 \rangle = \langle (20 \log_{10}(A/A_0))^2 \rangle = \langle \chi^2 \rangle (20 \log_{10}(e))^2 = \langle \chi^2 \rangle 75.4 \quad (3.38)$$

The model above predicts for a distance of 400 m with  $\langle \mu^2 \rangle = 10^{-6}$  and  $a = 1.0$  m at 2 kHz a standard deviation of 6 dB and at 4 kHz even 12 dB. These values are much higher than the measured fluctuations in Figure 3.14. This overprediction of the simple model was already observed by Daigle [41]. Furthermore he pointed out that measurements show a saturation of the log-amplitude fluctuations around  $\langle \chi^2 \rangle \approx 0.33$  or  $\langle \chi'^2 \rangle \approx 25$  which corresponds to a standard deviation of 5 dB.

In order to further investigate the amplitude fluctuations for the specific case of wind turbine noise propagation, numerical simulations were carried out. The following section presents these simulations.

### Numerical simulations

Numerical simulations to study amplitude fluctuations due to atmospheric turbulence were performed with the Finite Difference Time Domain (FDTD) method [44]. FDTD allows for solving the wave equation in the time domain. As an initial condition, a spatially and temporally concentrated pressure pulse is introduced at a source location. By an iterative procedure, the temporal evolution of the sound field is calculated and thus the propagation of the pulse through space can be studied. In FDTD, time and space are discretized. For each grid point in the computational region, the speed of sound can be specified individually. By using Equation (3.35), sound propagation through fields with atmospheric turbulence can be simulated, which makes this method particularly suited for this study.

The numerical simulations make usage of the concept of frozen turbulence. Thereby it is assumed that the condition of the atmosphere does not change during a pulse propagation from source to receiver. After many runs with arbitrary fields of the refraction index, statistical quantities can be derived that yield information about the distance and frequency dependency of the global amplitude fluctuations and variations in the ground effect dips due to increasing incoherence between the direct and ground reflected wave. The above observation, that the source has to be regarded as extended over a height that corresponds to the rotor diameter, has to be considered here as well, as this leads to a certain averaging over small-scale local inhomogeneities. The extension of the source is modeled by a series of five point sources distributed equidistantly along a vertical line running through the hub position.

The 2D-FDTD simulation needs a specification of the field of the fluctuating part  $\mu(x, z)$  of the refraction index  $n$  (Equation (3.35)). Hereby,  $x$  is the horizontal coordinate along the propagation direction and  $z$  describes height. The fluctuation statistics is then obtained by evaluating many runs with different  $\mu(x, z)$  realizations.

Different turbulence models have been proposed in the literature, see e.g. [45]. In the following, as a simplifying approximation, a Gaussian turbulence model is assumed.

### Synthesis of an arbitrary field of the fluctuating index of refraction

Following [46], a homogeneous field  $\mu(x, z)$  is assumed in a first step. In this case, the spatial variability of  $\mu$  can be characterized by a correlation function with vector  $\vec{s}$  between the two points in space of interest as the argument. This correlation function  $B(\vec{s})$  is defined as

$$B(\vec{s}) \equiv \overline{\mu(\vec{s}' + \vec{s})\mu(\vec{s}')} \quad (3.39)$$

The bar in Equation (3.39) represents average over time. Under the assumptions of an atmosphere with only temperature fluctuations and a Gaussian spectrum  $B$ , can be written as [46]

$$B(\vec{s}) = B(d) = \langle \mu^2 \rangle \exp\left(-\frac{d^2}{a^2}\right) \quad (3.40)$$

where distance  $d = \|\vec{s}\|$  and  $a$  denotes the correlation length. Equation (3.40) thus describes an isotropic turbulent atmosphere. However, in general, the field has to be considered anisotropic. This is particularly the case in the presence of wind fluctuations, which implicates that the correlation function depends on the direction. In the following, a Gaussian spectrum of the longitudinal wind velocity is assumed. Using Equation (3.36) and Cartesian coordinates with  $\vec{s} = \langle x, y, z \rangle$ , the correlation function is [46]

$$B(x, y, z) = \left[ \frac{\sigma_T^2}{4T_0^2} + \frac{\sigma_v^2}{c_0^2} \left(1 - \frac{\rho^2}{a^2}\right) \right] \exp\left(-\frac{d^2}{a^2}\right) \quad (3.41)$$

where  $\sigma_T$  and  $\sigma_v$  are the standard deviations of the temperature and wind velocity variation. The variables  $d$  and  $\rho$  are defined as distance  $d = \sqrt{x^2 + y^2 + z^2}$ , and the spatial separation transverse to the wave propagation direction  $\rho = \sqrt{y^2 + z^2}$ .

The spectral density  $F$  is defined as the spatial Fourier transform of the correlation function  $B$ . The two-dimensional spectral density function  $F(k_x, k_z)$  of Equation (3.41) is given by [46]

$$F(k_x, k_z) = \frac{a^2}{4\pi} \left( \frac{\sigma_T^2}{4T_0^2} + \frac{\sigma_v^2(k_z^2 a^2 + 2)}{4c_0^2} \right) \exp\left(-\frac{(k_x^2 + k_z^2)a^2}{4}\right) \quad (3.42)$$

with  $k_x$  and  $k_z$  being the wavenumber in  $x$  and vertical direction, respectively. Figure 3.15 shows two spectral density functions for two correlation lengths  $a = 1.2$  m and  $a = 30$  m, respectively.

Based on Equation (3.42), the  $\mu$ -field can be written as [46]

$$\mu(x, z) = \sqrt{4\pi\Delta k} \sum_{i=1}^N \cos(xk_i \cos \gamma_i + zk_i \sin \gamma_i + \tilde{\gamma}_i) \sqrt{F(k_i \cos \gamma_i, k_i \sin \gamma_i)} k_i \quad (3.43)$$

with  $k_i = i\Delta k$ , and  $\gamma_i$  and  $\tilde{\gamma}_i$  being arbitrary angles between 0 and  $2\pi$ .

### Height dependent correlation length

For sound propagation high above ground, the above assumptions of a homogeneous turbulence structure with a constant correlation length  $a$  are no longer fulfilled. In [47] measurements are presented that evaluated  $a$  at different heights  $z$

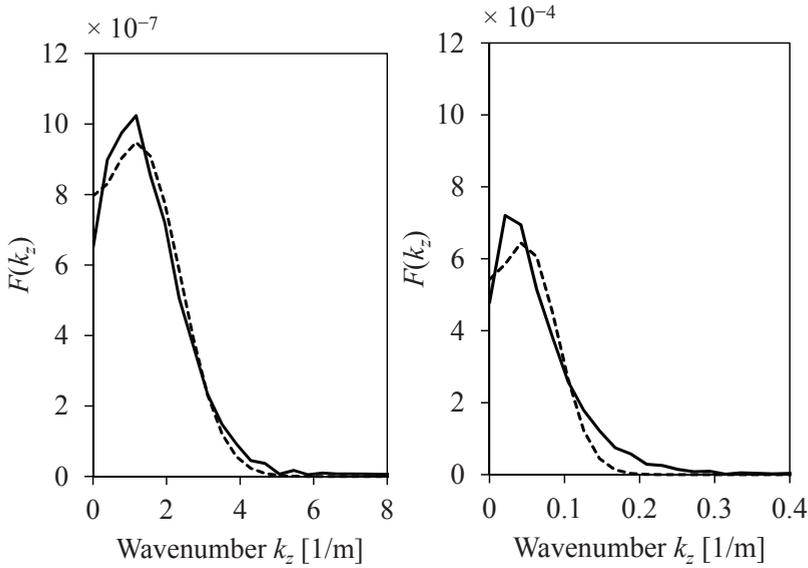


Figure 3.15: Comparison of the spectral density function  $F(k_x = 0, k_z)$  (dashed line) according to Equation (3.42) and the one-dimensional Fourier transform of the correlation function  $B(x = 0, y = 0, z)$  (solid line) for a height-dependent correlation length  $a(z)$ . Left: 2 m above ground, right: 100 m above ground.

above ground. At 1 m above ground, typical values are  $a \approx 1$  m, while at a height of 33 m,  $a$  increases to about 10 m. Here the following functional relation is used:

$$a(z) = 0.3z + e^{-0.3z} \quad (3.44)$$

where  $z$  is the height above ground.

Furthermore, [47] gives typical turbulence parameters for different Turner classes, which is a categorization scheme of atmospheric stability. For class 3 that is representative for moderate to strong wind and thus relevant for wind turbines in operation, the parameters are set as follows:  $\sigma_v = 1.3$  m/s and  $\sigma_T = 0.3^\circ\text{C}$ . With Equation (3.36) follows  $\langle \mu^2 \rangle = 1.5 \times 10^{-5}$ . Thus, for this meteorological condition  $\langle \mu^2 \rangle$  is dominated by wind speed fluctuations (expressed by  $\sigma_v$ ).

Assuming a variation of  $a$  with height  $z$ , the spectral density function  $F(k_x, k_z)$  in Equation (3.42) is no longer strictly valid. To prove exemplarily that Equation (3.42) can still be considered a good approximation, the spectral density function  $F(k_x = 0, k_z)$  was evaluated for the heights  $z = 2$  m ( $\rightarrow a = 1.2$  m) and  $z = 100$  m ( $\rightarrow a = 30$  m) and compared to the Fourier transform of the correlation function  $B(x = 0, y = 0, z)$  with a height dependent  $a(z)$ . As can be seen in Figure 3.15, the differences are rather small, supporting the applicability of Equation (3.42).

Figure 3.16 shows an example of a  $\mu$ -field realization with height dependent correlation length  $a$  where Equation (3.44) was inserted in Equation (3.42).

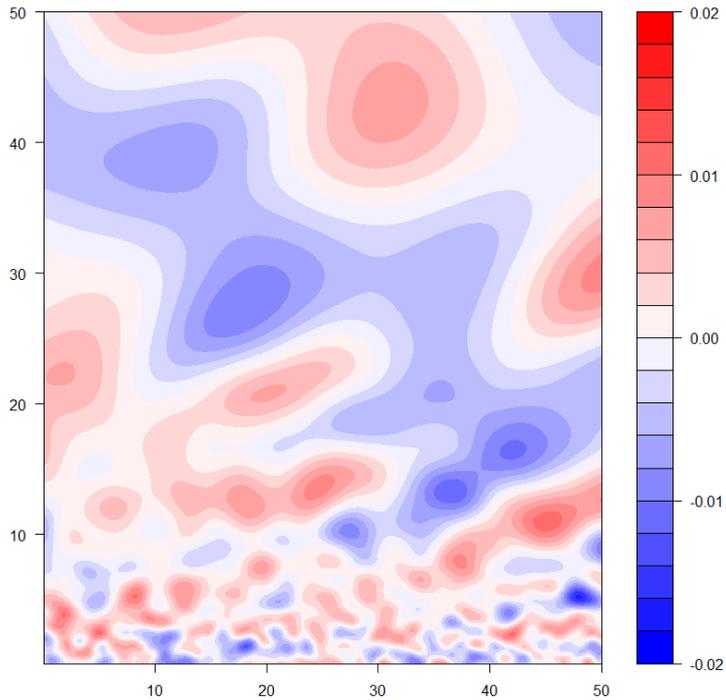


Figure 3.16: Example of a 2D field of the fluctuating part of the refractive index  $\mu(x, z)$  for  $\alpha$  according to Equation (3.44),  $\sigma_v = 1.3$  m/s,  $\sigma_T = 0.3^\circ\text{C}$ ,  $c_0 = 340$  m/s,  $T_0 = 293$  K and  $N = 200$  and  $\Delta k = 0.05$  m $^{-1}$ . The horizontal axis corresponds to the  $x$  coordinate, the vertical axis is  $z$ , both in meters.

### Simulation results

With the above introduced turbulence parameter values  $\sigma_v = 1.3$  m/s,  $c_0 = 340$  m/s,  $\sigma_T = 0.3^\circ\text{C}$ ,  $T_0 = 293$  K,  $\rightarrow \langle \mu^2 \rangle = 1.5 \times 10^{-5}$ , that are considered representative for moderate to strong wind conditions and Equation (3.44), 14 different  $\mu(x, z)$ -field realizations (one of them is shown in Figure 3.16) were determined. For each  $\mu(x, z)$ -field, the sound propagation was calculated with a 2D-FDTD simulation from five different point sources at heights 45 m, 62.5 m, 80 m, 97.5 m and 115 m to four receivers in 100 m, 200 m, 300 m and 400 m at at height of 1.6 m.

### Standard deviation of third octave band level fluctuations

Figure 3.17 shows the fluctuations as a function of frequency for a propagation distance of 400 m. The fluctuations strongly increase with frequency, but show no smooth behavior. This can be attributed to ground effect interferences that lead to frequency dependent sensitivities with respect to an inhomogeneous atmosphere. The standard deviations for the individual point sources range from 5 to 7 dB at 1.6 kHz. However if the already mentioned source extension is taken into account and the average over all five point sources is considered, the resulting standard deviation drops to about 2.5 dB at 1.6 kHz.

Figure 3.18 depicts the influence of distance on the fluctuations. Here the fluctuations are evaluated as average over all five point sources. The simulation results are compared with measured fluctuations as obtained during the campaign mentioned in section 3.3.8. In tendency the fluctuations increase with distance, however there is no monotonous behavior for all frequencies. The reason for this lies again in the nature of ground effect interferences. For varying propagation distances, the frequencies of constructive and destructive interference change. The comparison of the numerically simulated and measured fluctuations yields reasonable agreement if one considers the large spread of the measurement results. Together with the observations about the ground effect, the results support the finding that a wind turbine has to be modeled as significantly extended source.

### Correlation between different third octave bands

The above evaluation yielded the frequency dependency of the fluctuation amplitudes. For the propagation simulation, additional information is needed about the coherence of the fluctuations in different frequency bands.

For that purpose for each  $\mu$ -field realization  $i$ , a reference value  $L_{\text{ref},i}$  was determined as the energetic sum of the 6 third octave band receiver levels from 500 Hz to 1.6 kHz. This band was chosen as it covers the frequency range of highest fluctuations. Subsequently the correlation between this reference and the corresponding receiver levels  $L_{j,i}$  in each third octave band  $j$  was calculated and evaluated as coefficient of determination ( $R^2$ ). As can be seen in Figure 3.19, there is a high correlation of the fluctuations in the frequency bands above 500 Hz. This suggests that, at least in the upper frequency range, the fluctuations can be considered time-synchronous. As the fluctuations below 500 Hz are relatively low, this synchronicity is assumed here over the complete frequency range for simplicity. Consequently the fluctuations can be modeled as a global, time-varying filter function  $\mathcal{F}_0^t()$ .

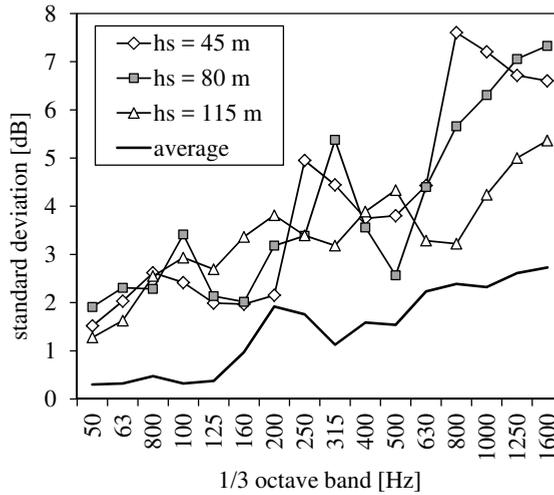


Figure 3.17: Frequency dependent standard deviations of the fluctuations for a point source at three different heights and the average over all five source positions in 400 m distance.

### Implementation of the global fluctuation filter function

Simplifying the frequency dependency of the fluctuations and assuming synchronous fluctuation of all frequency bands, the level variations can be modeled by a time dependent high-shelf filter function  $\mathcal{F}_0^t$  with amplitude response:

$$|\mathcal{F}_0^t(f)| = \frac{F_{\text{turb}}(t) - C_{\text{off}}}{1 + \exp\left(\frac{\ln(81)}{f_1 - f_2} [f - 0.5(f_2 + f_1)]\right)} \quad [\text{dB}] \quad (3.45)$$

The level fluctuation function  $F_{\text{turb}}(t)$  in dB is steered by a random process and  $f_1$  and  $f_2$  are the frequencies for which the high-shelf function reaches 10% and 90% of  $F_{\text{turb}}(t) - C_{\text{off}}$ . As  $F_{\text{turb}}(t)$  in dB varies equally around 0, an additional offset correction  $C_{\text{off}}$  is needed to adjust for energy neutral fluctuations. From the numerical simulations the following parameter setting has been derived:  $f_1 = 2500/\sqrt{d}$  [Hz] and  $f_2 = 20000/\sqrt{d}$  [Hz] with  $d$ : horizontal distance between source and receiver in meters. Here  $F_{\text{turb}}(t)$  is scaled for a resulting standard deviation of 2.0 dB, considering the fact that the synthesized emission signal already contains a significant portion of the propagation fluctuation.

With the above introduced concept of frozen turbulence in mind, information about the spectral content of the random process  $F_{\text{turb}}(t)$  can be derived in principal with help of a series of FDTD simulations. Hereby a random field of refraction index is generated and shifted by a certain offset for each subsequent run. The level history at a given receiver point can then be interpreted as a time function where the time axis is defined by the offset and an assumed representative wind speed.

As this method would be very time consuming, a simplifying approximation was used instead. For that purpose, the local level variation within a single propagation

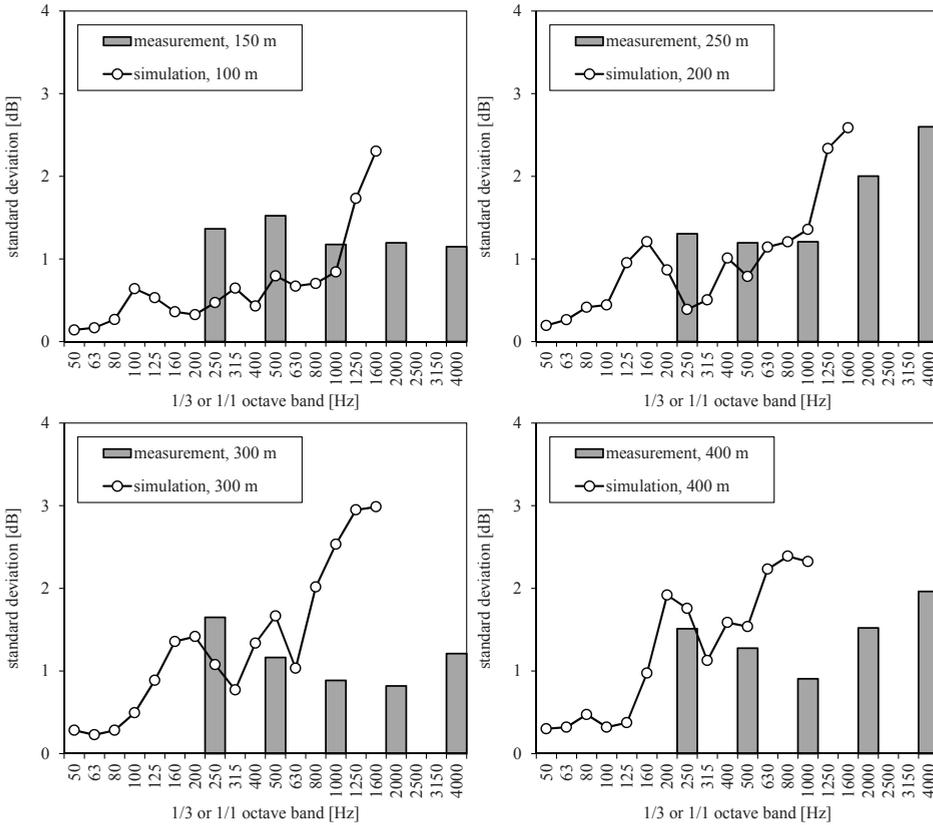


Figure 3.18: Frequency dependent standard deviations of the simulated fluctuations for the average over all five source positions for the propagation distances 100 m (top left), 200 m (top right), 300 m (bottom left) and 400 m (bottom right) in comparison with measured fluctuations.

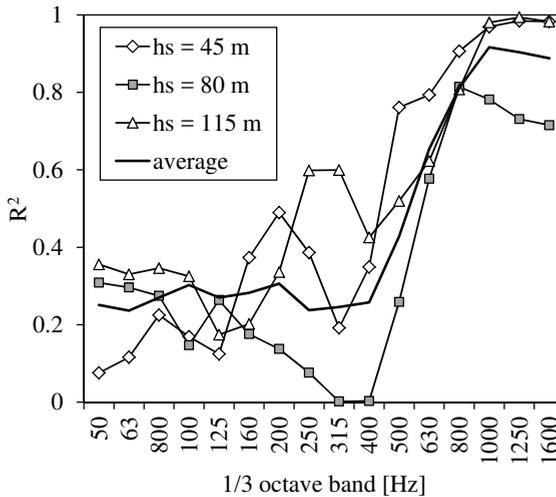


Figure 3.19: Frequency dependent coefficient of determination ( $R^2$ ) between third octave band fluctuations for a point source at three different heights and average over all five source positions in 400 m distance.

simulation was evaluated. To do so, a row of equidistantly distributed receiver points was introduced. 100 receiver points were placed between 200 and 250 m from the source, with a horizontal spacing of 0.5 m. Figure 3.20 shows obtained levels as a function of the receiver distance from the source. As in the simulation a fixed point source was used, these level variations are only caused by changes of the field of refraction near the receiver. However, in a moving medium (in contrast to the frozen medium) a similar effect is expected near the source. So as a first order estimate, one can assume that for a moving medium the level variations as seen in Figure 3.20 would occur at half the indicated source–receiver distance already. Figure 3.20 would thus represent a distance interval of  $\Delta d/2 = 25$  m instead of  $\Delta d = 50$  m. For a certain horizontal wind speed,  $v$ , this distance interval can be translated into a corresponding time interval  $\Delta d/(2v)$ . Therewith, the level variations of Figure 3.20 are interpreted as a level time history.

Due to the short signal lengths, instead of a spectral analysis, the number of turning points (local maxima and minima) are counted in the signals. Assuming a representative wind speed of 6 m/s, the evaluation yields approximately 4 turning points per second.

This information is used to model the random process  $F_{\text{turb}}(t)$  in Equation (3.45). The random process  $F_{\text{turb}}(t)$  is generated by low-pass (order 3) filtered white noise. The upper limiting frequency  $f_u$  is set to 2 Hz in order to match the number of turning points per second as observed in the FDTD simulation. This upper limiting frequency is considered valid for an average distance range of several hundred meters. For larger distances and for lower wind speeds,  $f_u$  is expected to lower accordingly.

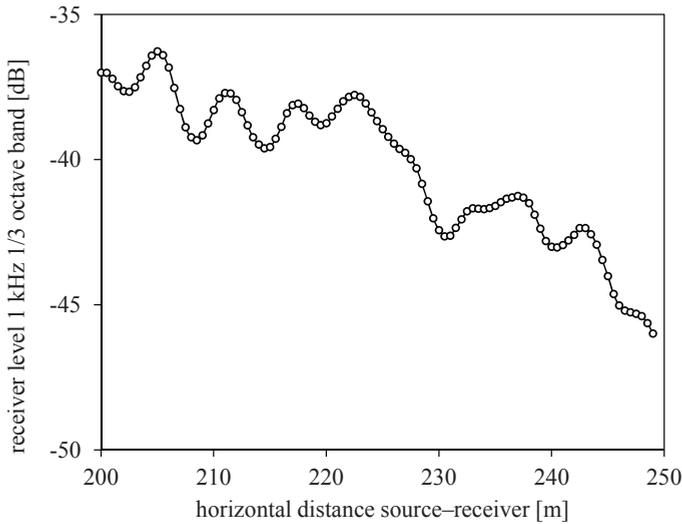


Figure 3.20: Local level variations due to inhomogeneous atmosphere for a row of 100 receivers at horizontal distances to the source between 200 and 250 m.

### 3.4. Conclusion

With the proposed signal analysis method, the quite complex wind turbine sound can be decomposed into only a few components that can be described by few parameters. Depending on the tonal content of the sound, the synthesis model only uses about 120 low-precision figures. As demonstrated by a test with experienced listeners, the presented emission model allows for a synthesis of an audio signal that can be easily confused with the original sound.

A model has been presented to perform the filtering of an audio signal, according to the effects of sound propagation from a wind turbine to a receiver. The model is based on FIR filters that are applied consecutively, each representing a specific propagation phenomenon. Initial measurements have revealed that wind turbines have to be viewed as largely extended sources. This leads to considerable smoothing of the ground effect dip and level variations due to turbulence.

The synthesis model together with the sound propagation filtering and a reproduction system is a reliable basis for psychoacoustic studies in which the effect of specific sound characteristics are systematically modified and hence investigated [4, 10, 15].

### Acknowledgments

The work in this chapter was financially supported by the Swiss National Science Foundation (SNSF).

## References

- [1] R. Pieren, K. Heutschi, M. Müller, M. Manyoky, and K. Eggenschwiler, *Auralization of wind turbine noise: Emission synthesis*, Acta Acustica United with Acustica **100**, 25 (2014).
- [2] K. Heutschi, R. Pieren, M. Müller, M. Manyoky, U. Hayek Wissen, and K. Eggenschwiler, *Auralization of wind turbine noise: Propagation filtering and vegetation noise synthesis*, Acta Acustica United with Acustica **100**, 13 (2014).
- [3] M. Manyoky, U. Hayek Wissen, K. Heutschi, R. Pieren, and A. Grêt-Regamey, *Developing a gis-based visual-acoustic 3d simulation for wind farm assessment*, ISPRS International Journal of Geo-Information **3**, 29 (2014).
- [4] B. Schäffer, S. J. Schlittmeier, R. Pieren, K. Heutschi, M. Brink, R. Graf, and J. Hellbrück, *Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study*, *Acoustical Society of America Journal* **139**, 2949 (2016).
- [5] M. Wolsink, *Wind power implementation: The nature of public attitudes: Equity and fairness instead of "backyard motives"*, *Renew. Sustain. Energy Rev.* **11**, 1188 (2007).
- [6] P. Devine-Wright, *Beyond nimbyism: Towards an integrated framework for understanding public perceptions of wind energy*, *Wind Energy* **8**, 125 (2005).
- [7] D. Shepherd, D. McBride, D. Welch, E. Hill, and K. Dirks, *Evaluating the impact of wind turbine noise on health-related quality of life*, *Noise Health* **13**, 333 (2011).
- [8] S. A. Janssen, H. Vos, A. R. Eisses, and E. Pedersen, *A comparison between exposure-response relationships for wind turbine annoyance and annoyance due to other noise sources*, *Journal of the Acoustical Society of America* **130**, 3746 (2011).
- [9] M. Manyoky, U. Hayek Wissen, T. M. Klein, R. Pieren, K. Heutschi, and A. Grêt-Regamey, *Concept for collaborative design of wind farms facilitated by an interactive gis-based visual-acoustic 3d simulation*, in *Digital Landscape Architecture 2012*, Peer Reviewed Proceedings at Anhalt University of Applied Sciences, Wichmann (Berlin and Offenbach, 2012) pp. 297–306.
- [10] B. Schäffer, R. Pieren, S. J. Schlittmeier, and M. Brink, *Effects of different spectral shapes and amplitude modulation of broadband noise on annoyance reactions in a controlled listening experiment*, *International Journal of Environmental Research and Public Health* **15** (2018), 10.3390/ijerph15051029.
- [11] S. V. Legarth, *Auralization and assessments of annoyance from wind turbines*, in *Second International Meeting on Wind Turbine Noise* (Lyon, France, 2007) pp. 1–14.

- [12] T. Pieter, M. Boes, T. Van Renterghem, D. Botteldooren, M. Hornikx, W. Desmet, and M. Nilsson, *Auralisation of a car pass-by behind a low finite-length vegetated noise barrier*, in *Euronoise 2012, 9th European Conference on Noise Control* (Prague, Czech Republic, 2012) pp. 932–937.
- [13] M. Müller, R. Pieren, and J. Bitzer, *Modell zur Synthese akustischer Emissionen von Windturbinen*, in *Proceedings of the 38th German Annual Conference on Acoustics (DAGA 2012)* (Darmstadt, Germany, 2012).
- [14] M. Manyoky, U. Wissen Hayek, R. Pieren, K. Heutschi, and A. Grêt-Regamey, *Evaluating a visual-acoustic simulation for wind park assessment*, *Landscape and Urban Planning* **153**, 180 (2016).
- [15] R. Ribe, M. Manyoky, U. Wissen Hayek, R. Pieren, K. Heutschi, and A. Grêt-Regamey, *Dissecting perceptions of wind energy projects: A laboratory experiment using high-quality audio-visual simulations to analyze experiential versus acceptability ratings and information effects*, *Landscape and Urban Planning* **169**, 131 (2017).
- [16] International Electrotechnical Commission, *IEC 61400-11:2008: Wind turbines - Part 11: Acoustic noise measurement techniques (IEC 88/324/CD:2008)*, (2008).
- [17] H. Moller and C. S. Pedersen, *Low-frequency noise from large wind turbines*, *Journal of the Acoustical Society of America* **129**, 3727 (2011).
- [18] G. P. van den Berg, *Effects of the wind profile at night on wind turbine sound*, *Journal of Sound and Vibration* **277**, 955 (2004).
- [19] D. Bowdler, *Amplitude modulation of wind turbine noise: A review of the evidence*, *Acoustics Bulletin of the Institute of Acoustics (UK)* **33** (2008).
- [20] S. Oerlemans and J. G. Schepers, *Prediction of wind turbine noise and validation against experiment*, *International Journal of Aeroacoustics* **8**, 555 (2009).
- [21] S. Wagner, R. Bareiss, and G. Guidati, *Wind Turbine Noise* (Springer, Berlin, 1996).
- [22] S. Oerlemans, P. Sijtsma, and B. M. Lopez, *Location and quantification of noise sources on a wind turbine*, *Journal of Sound and Vibration* **299**, 869 (2007).
- [23] S. Lee, K. Kim, and W. Choi, *Annoyance caused by amplitude modulation of wind turbine noise*, *Noise Control Engineering Journal* **59**, 38 (2011).
- [24] G. Lundmark, *Measurement of swish noise. a new method*, in *Forth International Meeting on Wind Turbine Noise* (Rome, Italy, 2011) pp. 1–8.
- [25] A. Bockstael, L. Dekoninck, A. Can, D. Oldoni, B. De Coensel, and D. Botteldooren, *Reduction of wind turbine noise annoyance: An operational approach*, *Acta Acustica United with Acustica* **98**, 392 (2012).

- [26] J. N. McCabe, *Detection and quantification of amplitude modulation in wind turbine noise*, in Forth International Meeting on Wind Turbine Noise (Rome, Italy, 2011) pp. 1–16.
- [27] International Organization of Standardization, *ISO 9613-1: Acoustics – Attenuation of sound during propagation outdoors - Part 1: Calculation of the absorption of sound by the atmosphere*, (1993).
- [28] U. Wissen Hayek, *VisAsim project website*, <http://www.visasim.ethz.ch>, (2017).
- [29] H. Fastl and E. Zwicker, *Psychoacoustics* (Springer, Berlin, 2007).
- [30] C. Larsson and O. Öhlund, *Long-term measurements of sound from wind turbines during different weather conditions*, in Euronoise 2012, 9th European Conference on Noise Control (Prague, Czech Republic, 2012) pp. 789–794.
- [31] J. Forssén, M. Schiff, E. Pedersen, and K. Persson Waye, *Wind turbine noise propagation over flat ground: Measurements and predictions*, *Acta Acustica United with Acustica* **96**, 735 (2010).
- [32] M. Trimpop and R. Bütikofer, *Sound propagation from elevated sources. results of a dedicated measuring campaign*, *Acta Acustica United with Acustica* **98**, 722 (2012).
- [33] K. Bolin, M. Nilsson, and S. Khan, *The potential of natural sounds to mask wind turbine noise*, *Acta Acustica United with Acustica* **96**, 131 (2010).
- [34] K. Bolin, A. Kedhammar, and M. Nilsson, *The influence of background sounds on loudness and annoyance of wind turbine noise*, *Acta Acustica United with Acustica* **98**, 741 (2012).
- [35] C. I. Chessell, *Propagation of noise along a finite impedance boundary*, *Journal of the Acoustical Society of America* **62**, 825 (1977).
- [36] C. F. Chien and W. W. Soroka, *A note on the calculation of sound propagation along an impedance plane*, *Journal of Sound and Vibration* **69**, 340 (1980).
- [37] International Organization of Standardization, *ISO 9613-2: Acoustics – Attenuation of sound during propagation outdoors - Part 2: General method of calculation*, (1996).
- [38] K. Heutschi, *Calculation of reflections in an urban environment*, *Acta Acustica United with Acustica* **95**, 644 (2009).
- [39] J. M. Wunderli, *An extended model to predict reflections from forests*, *Acta Acustica United with Acustica* **98**, 263 (2012).
- [40] R. Pieren and J. M. Wunderli, *A model to predict sound reflections from cliffs*, *Acta Acustica United with Acustica* **97**, 243 (2011).

- [41] G. Daigle, J. Piercy, and T. Embleton, *Line-of-sight propagation through atmospheric turbulence near the ground*, Journal of the Acoustical Society of America **74**, 1505 (1983).
- [42] V. N. Karavainikov, *Fluctuations of amplitude and phase in a spherical wave*, Akust. Zh. **3**, 175 (1957).
- [43] L. A. Chernov, *Wave propagation in a random medium* (Mc Gray-Hill book company, New York, 1960).
- [44] K. Heutschi, M. Horvath, and J. Hofmann, *Simulation of ground impedance in finite difference time domain calculations of outdoor sound propagation*, Acta Acustica United with Acustica **91**, 35 (2005).
- [45] V. Ostashev, *Acoustics in moving inhomogeneous media* (Taylor & Francis, 2002).
- [46] E. Salomons, *Computational atmospheric acoustics* (Kluwer Academic Publishers, 2001).
- [47] M. A. Johnson, R. Respet, and M. T. Bobak, *A turbulence model for sound propagation from an elevated source above level ground*, Journal of the Acoustical Society of America **81**, 638 (1987).



# 4

## Auralization of Road Traffic - Accelerating Passenger Cars

*With road traffic noise being the dominant noise source in most countries, particular interest lies in the synthesis of realistic pass-by sounds. This chapter describes an auralization system for pass-bys of accelerating passenger cars. The key element is a synthesizer that simulates the acoustical emission of different vehicles, driving on different surfaces, under different operating conditions. Audio signals for the emitted tire noise, as well as the propulsion noise are generated using spectral modeling synthesis, which allows for complete control of the signal characteristics. The sound of propulsion is synthesized as a function of instantaneous engine speed, engine load and emission angle, whereas the sound of tires is created in dependence of vehicle speed and emission angle. The sound propagation is simulated by applying a series of time-variant digital filters. To obtain the corresponding steering parameters of the synthesizer, controlled experiments were carried out. The tire noise parameters were determined from pass-by measurements of passenger cars with idling engines. To obtain the propulsion noise parameters, measurements at different engine speeds, engine loads and emission angles were performed using a chassis dynamometer. The chapter describes the identification of the synthesizer parameters by evaluation of audio recordings.*

## 4.1. Introduction

Generally, the auralization process consists of three modules [2]. The first module provides the signal emitted at the source. In its simplest form, this emission module makes use of suitable audio recordings. Obviously, the disadvantage of this strategy is its inflexibility, the limitations with respect to possible emission signal variations and the fundamental difficulty to obtain recordings that are not contaminated by unwanted sound. In addition, due to the Doppler effect, recordings of moving sources can be used for the auralization of situations with similar geometries only. It is therefore promising to use a digital emission synthesizer that is capable of generating the audio signal as radiated by the source. In [3], an enhanced Pitch Synchronous Overlap and Add (PSOLA) method is presented for internal combustion engine noise. This method features low computational costs and is therefore well suited for real-time applications. It is however limited in terms of flexibility, as it does not allow for inter- and extrapolations between measured signals. This limitation can be overcome by the synthesizer presented in this chapter, at the expense of computational costs.

The second module needed in the auralization process is a filter that simulates the sound propagation effects of the wave traveling from the source to the receiver. These effects involve geometrical spreading, propagation delay and atmospheric absorption. In outdoor situations, along with the direct sound path, also reflections occur, particularly ground reflections, leading to the ground effect. When either the source or the receiver moves, additionally, the Doppler effect arises. However, also, due to the time-varying propagation situation, all of the other effects change over time, which means that the propagation filter becomes time variant [4–6].

The third module is a reproduction system, which renders the received signals to headphones or a multi-channel loudspeaker system. In this process, for every received signal, its corresponding emission angle with respect to the observer is required. For the rendering, a variety of different methods exists [7]. Further, it has to be assured that the reproduction system has a linear frequency response and that it is correctly calibrated.

In the research project TAURA: Traffic Noise Auralisator (2014-2016), a traffic noise auralization system was developed that covers road traffic [1, 8] and railway noise [9, 10]. It will form the basis for future experiments to refine the characterization of noise. The key element is a synthesizer that simulates the acoustical emission of a large amount of different vehicles, operating on a wide variety of surfaces and under different operating conditions. In the TAURA model, road traffic noise is created by the superposition of individual vehicle pass-by sounds. The objective of this chapter is to describe how these single pass-by sounds can be generated in the case of passenger cars.

This chapter is structured in two main parts: In section 4.2, the auralization model of accelerating passenger cars is developed step by step and presented. Thereby, an emission synthesizer, which is based on spectral modeling synthesis, and propagation filtering algorithms are elucidated. Section 4.3 shows how the model parameters can be estimated based on controlled measurements. On that account, a series of signal analysis steps to obtain the steering parameters of the

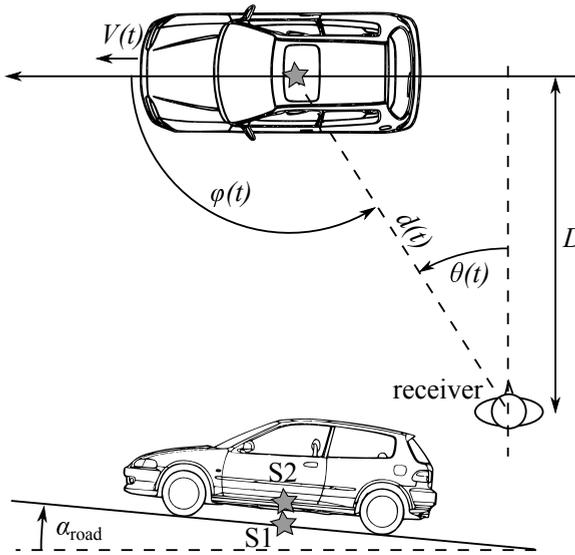


Figure 4.1: Sketch of the geometrical situation showing the two source position S1 and S2, the inclination angle of the road  $\alpha_{\text{road}}$ , the distance  $D$ , the instantaneous vehicle speed  $V$ , emission angle  $\varphi$ , immission angle  $\theta$  and source–receiver distance  $d$ .

synthesizer is proposed. The chapter ends with an evaluation in section 4.4 and conclusions in section 4.5.

## 4.2. Model development

### 4.2.1. Overview

This section presents an overview of the model to auralize accelerating passenger cars. Further, the key assumptions and motivations used for the model development are presented. In the model, each car is represented by two moving point sources. The geometrical situation is depicted in Figure 4.1, in which the distance of the straight driving lane to the receiver is  $D$ , the emission angle  $\varphi$ , the angle of inclination  $\alpha_{\text{road}}$  and the point source positions S1 and S2, respectively. Describing the kinematics of the vehicle, its speed  $V(t)$  in km/h as a function of time  $t$  (at the source) is used throughout this chapter. In correspondence with the Harmonoise model [11], the point sources are vertically stacked and located at heights of 0.01 and 0.3 m above ground. By not attributing separate point sources to each vehicle axle, we limit the applicability of the model to situations with source–receiver distances clearly larger than the axial distance, while in return, saving computational costs.

Road traffic noise is mainly composed of propulsion noise and tire noise [11–14]. Both contributions differ in their relevance for the total noise, depending on the vehicle, its operating conditions and the pavement type. This motivates that

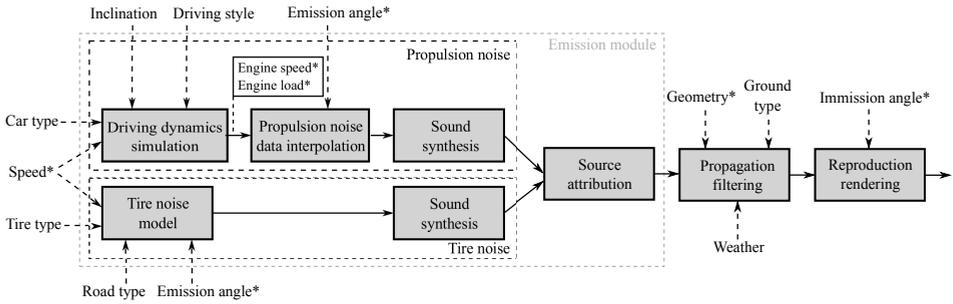


Figure 4.2: Simulation flowchart of the auralization of accelerating passenger cars. The input variables marked by a \* are time dependent.

## 4

in the presented model, the contributions of propulsion noise and tire noise are simulated individually. Their attribution to the point sources S1 and S2 is described in the following section 4.2.2.

In accordance with Vorländer’s definition of the “principle of auralization” [2], the presented model to auralize accelerating passenger cars comprises a separate emission, a propagation and a reproduction module. The emission module is described in section 4.2.2, the propagation module in section 4.2.3 and the reproduction module in section 4.2.4. Figure 4.2 shows the block diagram of the model. The input variables describe the vehicle, driver, road surface, geometry, ground type and the weather; the input variables marked by \* are time dependent.

The emission module describes the emitted sound of an individual passenger car pass-by, i.e. its outside sound. In Figure 4.2, the emission module is edged by a gray dashed line. It contains several simulation blocks creating two separate signal paths. The upper path represents the propulsion noise and the lower path the tire noise simulation. Key elements are two sound synthesizers that artificially generate audible emission signals for propulsion and tire noise, respectively. The synthesizer parameters can be obtained from controlled measurements, as elucidated in section 4.3. Section 4.2.2 describes how the synthesized emission signals are attributed to the two point sources. As a proper interface to the propagation module, the point source signals are defined at a virtual reference distance of  $d_0 = 1$  m. They feature time-varying sound characteristics and already include source directivity.

For both moving source positions, S1 and S2, the sound propagation to the (static) receiver position is simulated in the propagation module. To generate the receiver signals, both corresponding source signals,  $p_{e,1}$  and  $p_{e,2}$ , are filtered by a series of time-varying digital filters, as described in section 4.2.3. These filters depend on the instantaneous propagation geometry, the ground type and the weather conditions (cf. Figure 4.2). Finally, in the reproduction module, the receiver signals are summed up and rendered for multi-channel reproduction using the instantaneous immission angle. Section 4.2.4 exemplifies a possible stereo rendering procedure.

### 4.2.2. Emission module

The emission module describes the emitted sounds of an individual passenger car. Its structure is depicted in Figure 4.2. As described above, the acoustical emission of the passenger car is assumed to consist of the two contributions: tire noise and propulsion noise. Their corresponding emission signals are denoted as  $p_{\text{tire}}$  and  $p_{\text{prop}}$ , respectively.

Tire noise strongly depends on tire type [13, 15], road surface type [11–13, 16] and vehicle speed [11–14]. Further, the horn effect mainly determines the horizontal directivity of tire noise [11, 14, 17]. To model these effects, the tire noise contribution is assumed to depend on the road and tire type, as well as on vehicle speed  $V$  and the emission angle  $\varphi$ . The following section shows how the signals  $p_{\text{tire}}$  are calculated based on these input parameters.

In current noise prediction models, propulsion noise is commonly calculated as a function of vehicle speed, acceleration and road inclination [11, 12, 18]. This is due to the fact that these models are developed and used in cases for which the engaged gear is not known. The gear, however, strongly influences the sound of propulsion [14]. For a given speed, acceleration and road inclination, by changing the gear, the engine speed, as well as the engine load change. From an engine's viewpoint, it is these two parameters that are sufficient to fully describe the engine condition. The section after next explains how in the auralization model, engine speed  $n_{\text{rot}}$  and engine load  $\Gamma$  are calculated by simulating the driving dynamics of the vehicle. These simulations require information on the vehicle and the driving style, the road inclination  $\alpha_{\text{road}}$  and the vehicle speed  $V(t)$  as a function of time  $t$ . Further, propulsion noise features a directivity [14], which is also taken into account in the auralization model. Then it is shown how, based on  $n_{\text{rot}}$ ,  $\Gamma$  and the emission angle  $\varphi$ , the signal  $p_{\text{prop}}$  is calculated.

The audible emission signals  $p_{\text{tire}}$  and  $p_{\text{prop}}$  are generated artificially by two digital sound synthesizers. The synthesizers are based on a combination of additive and subtractive synthesis. In additive synthesis, the signal is constructed by the sum of sinusoids, each having a time-varying amplitude and phase [19–21]. On the other hand, subtractive synthesis uses filters to shape a more complex source signal, e.g., a sawtooth wave or white noise [19, 20]. The combination of both techniques is known as spectral modeling synthesis [21–23]. However, in contrast to the applications presented in [22] and [23], in the presented model, the sounds are not synthesized using the short-time Fourier transform (STFT), but directly in the time domain.

The signal of propulsion noise,  $p_{\text{prop}}$ , is fully attributed to the upper point source S2. However, the sound power of the tire noise contribution is attributed to the point sources by 80%/20% [11]. This translates to a ratio of 2:1 of their respective sound pressure signals. The conditions of incoherent signals and energy conservation yield a normalization factor of  $1/\sqrt{5}$ . Thus, the sound pressure source signals are:

$$p_{e,2}(t) = \frac{1}{\sqrt{5}}p_{\text{tire},2}(t) + p_{\text{prop}}(t) \quad (4.1)$$

$$p_{e,1}(t) = \frac{2}{\sqrt{5}}p_{\text{tire},1}(t) \quad (4.2)$$

at reference distance  $d_0 = 1$  m for source positions S2 and S1, respectively. Indices 2 and 1 indicate that different, uncorrelated signals for the sound of tires are generated for the two source positions.

### Sound of tires

The emission signal of the sound of tires is assumed to consist of broadband noise only, i.e. discrete tones due to, e.g., tire tread resonances or discrete vibrational tire resonances are not taken into account.

The spectral shaping of the broadband noise components is performed in 1/3 octave bands. For each 1/3 octave band  $i$ , white noise is generated and filtered by a digital pink filter. This pre-shaping helps to produce a smoother spectrum of the resulting signal [24]. The output of the pink filter is bandpass filtered by an eighth order Butterworth filter (Class 0 according to the standard IEC 1260:1995 [25]) and normalized to unit signal power to obtain the signal  $v_i(t)$ . For stability reasons, also at low frequencies, the filters are implemented as cascaded second-order sections (SOS).

The sound pressure emission signals of the sound of tires component are thus calculated by [24]:

$$p_{\text{tire}}(t) = \sum_{i=1}^{N_b} p_0 10^{L_{\text{tire},i}(V(t),\varphi(t))/20} v_i(t) \quad (4.3)$$

with  $N_b$  being the number of considered 1/3 octave bands, the reference pressure  $p_0 = 20 \mu\text{Pa}$  and normalized bandpass filtered pink noise signals  $v_i(t)$ . A total of  $N_b = 29$  bands from 20 Hz to 12.5 kHz are used. For the sound pressure level  $L_{\text{tire},i}$  of band  $i$ , a common logarithmic speed relationship [13] with additive correction terms is assumed:

$$L_{\text{tire},i}(t) = A_i + B_i \log_{10} \left( \frac{V(t)}{V_0} \right) + \Delta L_{\text{road},i}(V(t)) + \Delta L_{\text{dir},i}(\varphi(t)) \quad (4.4)$$

with vehicle speed  $V$  in km/h, reference speed  $V_0 = 70$  km/h, regression parameters  $A_i$  and  $B_i$ , the road surface correction  $\Delta L_{\text{road},i}$  and a horizontal directivity  $\Delta L_{\text{dir},i}$ . For the road surface correction, the Swiss "sonRoad" model [12] offers the parameter  $\Delta_{\text{BG}}$  for 10 surface types. However,  $\Delta_{\text{BG}}$  does not depend on frequency or vehicle speed. The recently-published EU directive on establishing common noise assessment methods (CNOSSOS-EU) [18] contains spectral corrections in octave bands in the form of:

$$\Delta L_{\text{road},i}(V(t)) = \alpha_i + \beta \log_{10} \left( \frac{V(t)}{V_0} \right) \quad (4.5)$$

with experimental regression parameters  $\alpha_i$  and  $\beta$ , which are tabulated for 15 different road surface types. The horizontal directivity simulates the horn effect [17] and only applies for signal  $p_{e,1}(t)$ , i.e. for the lower source position (S1). The empirically-obtained relationship [11]:

$$\Delta L_{\text{dir},i}(\varphi(t)) = \begin{cases} -2.5 + 4 |\cos \varphi(t)| + C_{\text{off}}, & 800 \text{ Hz} \leq f_{c,i} \leq 6.3 \text{ kHz} \\ 0, & \text{otherwise} \end{cases} \quad (4.6)$$

with the 1/3 octave band center frequencies  $f_{c,i}$  and the correction  $C_{\text{off}}$  is employed.  $C_{\text{off}}$  accounts for a limited emission angle range during emission measurements. For example, a 90° measurement range with  $45^\circ < \varphi < 135^\circ$  amounts to a correction  $C_{\text{off}} = 0.9$  dB.

### Driving dynamics

Figure 4.2 shows that as a first step of the propulsion noise simulation, the driving dynamics of the car are calculated in order to obtain the instantaneous engine speed  $n_{\text{rot}}(t)$  and engine load  $M(t)$ . The engine speed in engaged mode reads [26]:

$$n_{\text{rot}}(t) = 60 i_{\text{gear}}(t) i_{\text{ax}} \frac{V(t)/3.6}{2\pi r_{\text{tire,dyn}}} \quad [\text{rpm}] \quad (4.7)$$

with the instantaneous vehicle speed  $V(t)$  given in km/h, the gear ratio  $i_{\text{gear}}$ , the axle ratio  $i_{\text{ax}}$  and the dynamic tire radius  $r_{\text{tire,dyn}} \approx 0.3$  m. The traction  $F_T$  is modeled by: [26, 27]

$$F_T(t) = F_B(t) + \bar{e} m a(t) + m g \sin(\alpha_{\text{road}}) \quad [\text{N}] \quad (4.8)$$

$$F_B(t) = F_0 + F_1 V(t) + F_2 V^2(t) \quad (4.9)$$

$$a(t) = \frac{dV(t)/3.6}{dt} \quad (4.10)$$

with the vehicle mass  $m$ , gravity  $g$ , the inclination angle of the road  $\alpha_{\text{road}}$ , the translational acceleration  $a$  of the vehicle and a mean equivalent mass factor  $\bar{e} = 1.15$  for the rotational accelerations for each individual gear. The basic driving resistance  $F_B$  (consisting of rolling resistance and aerodynamic drag) is modeled by the coast-down parameters  $F_0$ ,  $F_1$  and  $F_2$  with units N, N/(km/h) and N/(km/h)<sup>2</sup>, respectively. These parameters have to be provided by the manufacturer during the type approval procedure. The engine load (torque) is formulated by [26]:

$$M(t) = \frac{r_{\text{tire,dyn}} F_T(t)}{\eta i_{\text{gear}}(t) i_{\text{ax}}} \quad [\text{Nm}] \quad (4.11)$$

with a globally-set efficiency factor  $\eta = 0.9$  for the power transmission from the engine to the wheels. The engine load in percent is defined by [28]:

$$\Gamma(t) \equiv \frac{M(t)}{M_{\max}(n_{\text{rot}}(t))} 100 \quad [\%] \quad (4.12)$$

with  $\Gamma = 100\%$  at full load. At idling engine,  $M = 0$  Nm and  $\Gamma = 0\%$ . In engine overrun operation (e.g., while engine braking), the engine delivers a negative torque to the crankshaft, which means that the engine load  $M$  becomes negative. In the model, this state is approximated as idle, i.e.  $M$  is set to zero.

Gearbox shifts are modeled by three consecutive processes:

1. the clutch is disengaged
2. a new gear is put in (at idling engine)
3. the clutch is engaged again.

In dependence of the driving style, these processes vary in their respective durations. In the model, for a sporty driving style, the total gear change takes 0.6 s, whereas for a cozy, economic driving style, the gear change takes 1.3 s. Furthermore, the moments of a gear change strongly depend on the driving style and can be formulated as a function of engine speed and engine load, which is also the basic working principle of an automatic gearbox.

### Sound of propulsion

The structure of the emission synthesizer for the sound of propulsion is depicted in Figure 3.1. The sound pressure emission signal of the sound of propulsion is assumed to consist of a deterministic signal representing the most important engine orders and a quasi-stochastic signal:

$$p_{\text{prop}}(t) = p_{\text{prop,ord}}(t) + p_{\text{prop,noise}}(t) \quad (4.13)$$

Engine order  $o$  corresponds to an event taking place  $o$  times per engine revolution. The engine order signal is composed of the sum of the engine orders  $o$ , which are generated using additive synthesis [20, 21]. The engine order signal is thus calculated by [24]:

$$p_{\text{prop,ord}}(t) = \sum_o p_o 10^{L_{\text{prop,ord},o}^{\dagger}(t)/20} \sqrt{2} \cos(\beta_o(t)) \quad (4.14)$$

A proper selection of the essential orders  $o$  strongly depends on the specific vehicle type and its condition. In the context of sound design, it is known that at least orders up to  $o = 18$  are relevant [14]. Further, the sound characteristics can be influenced by half-orders [14]. In this model, it was decided to synthesize orders  $o = 1$  to 30 in half-order steps, resulting in a total of 59 orders. This somewhat arbitrary, but safe choice leaves room for optimization. In Equation (4.14),  $L_{\text{prop,ord},o}^{\dagger}$  denotes the order level, and  $\beta_o$  is the instantaneous order phase with

$$\beta_o(t) = \phi_o^{\dagger}(t) + 2\pi \int_{-\infty}^t f_o(\tau) d\tau \quad (4.15)$$

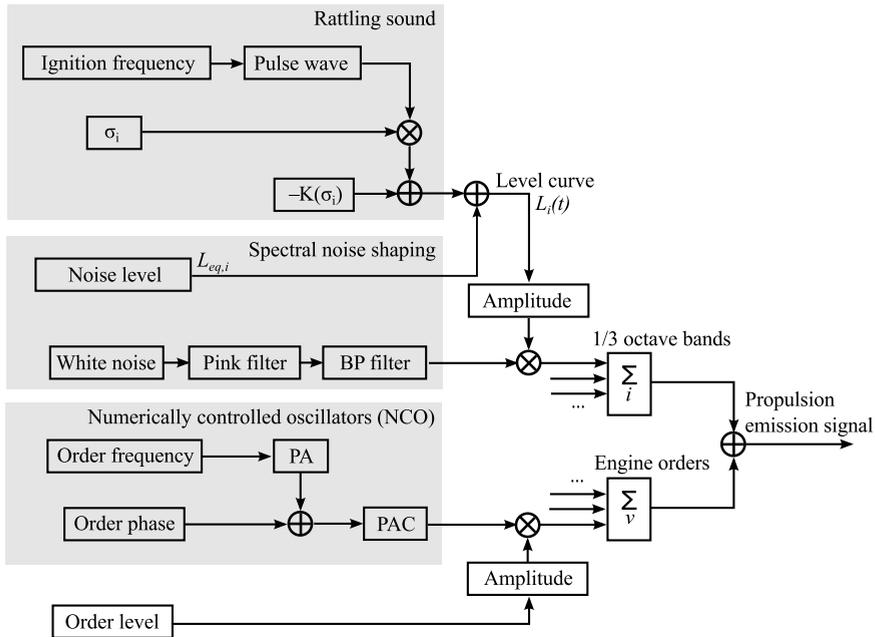


Figure 4.3: Signal flow chart of the synthesizer for the sound of propulsion. Engine order signals are calculated using numerically controlled oscillators (NCO) consisting of phase accumulators (PA) and phase-to-amplitude converters (PAC). The noise signals are generated using band-pass (BP) filtered noise that is modulated in amplitude.

with the order phase  $\phi_o^\dagger$  and the order frequency:

$$f_o(t) = o \frac{n_{\text{rot}}(t)}{60}. \quad (4.16)$$

Listening tests revealed that in this application, the order phase is a relevant synthesizer parameter. For a four-stroke engine with  $N_{\text{cyl}}$  cylinders, the engine order corresponding to the ignition, and mostly the predominant order, is  $o_{\text{ign}} = N_{\text{cyl}}/2$  [14].

For time-discrete signals, Equations (4.14) and (4.15) can be interpreted as a modified numerically-controlled oscillator (NCO) [19], whereas Equation (4.15) corresponds to the phase accumulator (PA), and the phase-to-amplitude converter (PAC) is realized in Equation (4.14). This formulation concurrently implements a frequency modulation by  $f_o$  and a phase modulation by  $\phi_o$ .

The noise signal component of the sound of propulsion is synthesized similarly as the sound of tires (Equation (4.3)) by:

$$p_{\text{prop,noise}}(t) = \sum_{i=1}^{N_b} p_0 10^{L_{\text{prop,noise},i}(t)/20} v'_i(t) \quad (4.17)$$

A total of  $N_b = 29$  bands from 20 Hz to 12.5 kHz are used again. The 1/3 octave band level function is formulated as:

$$L_{\text{prop,noise},i}(t) = L_{\text{eq,prop,noise},i}^\dagger + \sigma_i^\dagger W(t) - C_{\text{off}}(\sigma_i^\dagger) \quad (4.18)$$

with the level  $L_{\text{eq,prop,noise},i}^\dagger$ , a level standard deviation  $\sigma_i^\dagger$  and a periodic fluctuation function  $W(t)$  with zero mean and unit power. The constant  $C_{\text{off}}$  ensures that despite the level fluctuations, the equivalent continuous level (Leq) is not altered. The level modulation  $W(t)$  simulates the rattling sound component that elicits a roughness sensation, which is particularly characteristic for low engine speeds and diesel engines. Measurement data showed that the strongest level fluctuations occur at the ignition frequency

$$f_{\text{ign}}(t) = \frac{N_{\text{cyl}}}{2} \frac{n_{\text{rot}}(t)}{60}. \quad (4.19)$$

Therefore,  $W$  is modeled by a quasi-periodic function with period  $1/f_{\text{ign}}(t)$ . The first half-period of  $W$  is composed of a Hann window, whereas the second half-period is held constant.

In summary, the presented synthesizer needs about 180 input parameters to generate a stationary signal for the sound of propulsion. However, during a pass-by, the sound of propulsion may considerably vary, and so do these parameters. These parameters, which are marked by  $\dagger$  in the above equations, simultaneously depend on the engine speed  $n_{\text{rot}}$ , the engine load  $M$  and the emission angle  $\varphi$  and are hence time dependent. They are calculated by a triangulation-based linear 3D interpolation of measurement data. Measurements were taken on a discrete grid of measuring points, typically  $n_{\text{rot}} \approx \{1000, 2000, 3000, 4000\}$  rpm,  $\Gamma \approx \{0, 40, 70, 100\}$

% and  $\varphi = \{0, 60, 120, 180\}^\circ$ . For the measuring points a Delaunay triangulation is adopted to construct an interpolating surface. The synthesizer parameters are evaluated with a temporal resolution of 20 ms and linearly interpolated to the audio sampling rate,  $f_s$ . For the interpolation of the order phases,  $\phi_o$ , their cyclic behavior has to be considered in order to avoid spurious phase jumps.

### 4.2.3. Propagation filtering

The sound propagation model described in this section incorporates the following effects:

- Propagation delay
- Doppler effect (frequency shift and amplification)
- Convective amplification
- Geometrical spreading
- Ground reflection
- Air absorption

Other outdoor sound propagation effects that may be relevant in certain situations are screening [29–31], foliage attenuation [30], meteorological effects due to an inhomogeneous atmosphere [30–34], as well as reflections at artificial [30, 35] and natural surfaces [34, 36, 37].

Most published environmental noise auralization models simulate some of the above listed effects by applying a 1/3 octave filter bank and adjusting the filter gains [6, 38, 39]. In this model, however, all of these effects are applied in the time domain, i.e. by time-variant digital filters. Sound propagation is modeled by two paths, namely for direct sound and a single ground reflection (in the following account indicated by subscripts *dir* and *gr*, respectively). The sound pressure of a point source has a  $1/r$  distance dependency. Thus, to model geometrical spreading, the emitted sound pressure signals,  $\tilde{p}_e$ , are divided by their path length  $d_{\text{dir}}$  or  $d_{\text{gr}}$ , respectively. The interaction of a sound wave with the ground influences its amplitude and phase as a function of frequency. This effect can be modeled by convolution of the ground-reflected signal with a time-variant filter [4]. Furthermore, the attenuation due to air absorption can be efficiently modeled using a filter as described in section 3.3.5 [4]. Considering these aspects, the receiver signal  $p_r$  is calculated by:

$$p_r(t') = h_{\text{air}}^{t'}(t') * \left( \frac{\tilde{p}_{e,\text{dir}}(t')}{d_{\text{dir}}(t')} + h_{\text{swrc}}^{t'}(t') * \frac{\tilde{p}_{e,\text{gr}}(t')}{d_{\text{gr}}(t')} \right) \quad (4.20)$$

where  $*$  denotes linear convolution,  $t'$  the receiver time,  $d_{\text{dir}}$  is the source–receiver distance,  $d_{\text{gr}}$  is the distance source–ground reflection point–receiver,  $\tilde{p}_{e,\text{dir}}$  and  $\tilde{p}_{e,\text{gr}}$

are delayed versions of the emitted sounds and  $h_{\text{air}}^{t'}$  and  $h_{\text{SWTC}}^{t'}$  denote impulse responses of time-dependent filters described in sections 4.2.3 and 4.2.3. The modeling of effects due to source motion and the propagation delay are explained in the following section. Note that the immission angle  $\theta$ , which is needed for surround reproduction, has to be evaluated on the receiver time axis  $t'$ , as well.

#### Effects due to source motion and propagation delay

Due to the travel time of sound and the movement of the source, the source and the receiver have differing time axis. By neglecting wind and turbulence, the warped time at the receiver,  $t'$ , is given by:

$$t' = t + \Delta t(t) = t + \frac{d_{\text{dir/gr}}(t)}{c_0} \quad (4.21)$$

where  $t$  is the source time which is sometimes denoted as "retarded time",  $d_{\text{dir/gr}}$  denotes the sound propagation distance of the direct sound or the ground reflected sound, respectively, and  $c_0$  is sound speed in still air. A constant sound speed of  $c_0 = 340$  m/s is assumed. Since the receiver signal is supposed to have a constant sampling rate of  $f_s$ , the corresponding times,  $t_s$ , on the emission time axis have to be found. This is achieved by linear interpolation of Equation (4.21). The emission signals  $\tilde{p}_e$  for the direct and the ground-reflected path, respectively, with respect to the receiver time  $t'$  are:

$$\tilde{p}_{e,\text{dir/gr}}(t') = p_e(t) \mathcal{D}^2(t) \quad (4.22)$$

where  $p_e(t)$  is a sound pressure source signal from Equation (4.1) or (4.2), and the Doppler factor  $\mathcal{D}$  is

$$\mathcal{D} = \frac{f'}{f} = \frac{dt}{dt'} \quad (4.23)$$

Equation (4.22) describes the kinematic and the aerodynamic effect of source motion. The former is known as the Doppler effect, i.e. the Doppler frequency shift and amplification. The latter is known as convective amplification. The exponent 2 of the Doppler factor in Equation (4.22) indicates that a subsonically moving volume source (monopole) [40] is assumed.

The change of the time axis in Equation (4.22) realizes the propagation delay, as well as the Doppler frequency shift. For digital signals, this change corresponds to an asynchronous resampling process. It can be implemented using a variable delay-line with delay  $\Delta t$  [41]. If  $\Delta t$  is just rounded to the nearest sample, audible artifacts occur, so-called "zipper noise". Therefore, as in section 3.3.2, an interpolation strategy has to be used. As we are only interested in sequential access to the emission signal, a fractional delay filter can be used [42]. In section 3.3.2 and [12, 41], a linear interpolator is proposed. This however produces high frequency attenuation, as well as strong nonlinear distortions due to aliasing. Therefore, here, we introduce a band-limited interpolation or, respectively, a windowed sinc interpolation [42]:

$$u(k_s) = \sum_{k=\lfloor k_s \rfloor - b + 1}^{\lfloor k_s \rfloor + b} u[k] K_b(|k_s - k|) \quad (4.24)$$

with the floor function  $\lfloor \cdot \rfloor$  used in the bounds of summation, the integer sample index  $k$ , the non-integer sample index  $k_s = t_s f_s$  and the Hamming kernel:

$$K_b(m) = \begin{cases} [0.54 + 0.46 \cos(\pi m/b)] \operatorname{sinc}(m) & \text{if } 0 \leq m < b \\ 0 & \text{otherwise} \end{cases} \quad (4.25)$$

with an integer  $b$  describing the filter length. To keep the computational effort low, in the implementation of Equation (4.24), values of the kernel  $K_b$  are stored in a look-up table. The Doppler factor  $\mathcal{D}$  in Equation (4.22) is implemented by approximating the derivative in Equation (4.23) by finite differences as:

$$\mathcal{D}(k_s[i]) \cong k_s[i + 1] - k_s[i]. \quad (4.26)$$

with index  $i$ .

In order to validate different implementations of Equation (4.22), numerical simulations were performed. Figure 4.4 compares the signal attenuation introduced by different interpolation schemes. The high frequency attenuation of the linear interpolation can be counteracted by a windowed sinc interpolation and controlled by parameter  $b$  in Equation (4.24). Figure 4.5 shows spectrograms of receiver signals calculated by the same three interpolation schemes. As an extreme case, a virtual source emitting a 1 kHz pure tone travels at constant speed  $V = 150$  km/h and passes a static receiver at a distance of  $D = 7.5$  m. Figure 4.5 shows that by introducing a windowed sinc interpolation of sufficient filter length, artifacts due to aliasing can be significantly reduced compared to a linear interpolation (a). The minimal kernel size  $b$  required for a decent sound quality cannot be stated in general, as it strongly depends on the application, i.e. the source signal, the propagation situation and, not least, the sampling frequency. However, in the example of Figure 4.5, already,  $b = 10$  reaches a good sound quality, without audible artifacts. Nevertheless, to be on the safe side, a value of  $b = 100$  was adopted. The careful choice of  $b$ , however, provides the potential for optimization in terms of sound quality and computational cost.

For sound speed  $c_0 = 340$  m/s, the source Mach number  $M \equiv V/c_0 \approx 0.12$ . At times  $t = \pm\infty$ , the received frequencies  $f'$  due to the Doppler shift are given by [2, 43]:

$$f'_{t=\pm\infty} = f \mathcal{D}_{t=\pm\infty} = \frac{f}{1 \pm M} \quad (4.27)$$

with  $f$  being the emitted frequency. In our example, according to Equation (4.27), the received frequency changes by a factor of 1.28 across the pass-by, which corresponds to a musical interval that is larger than a major third.

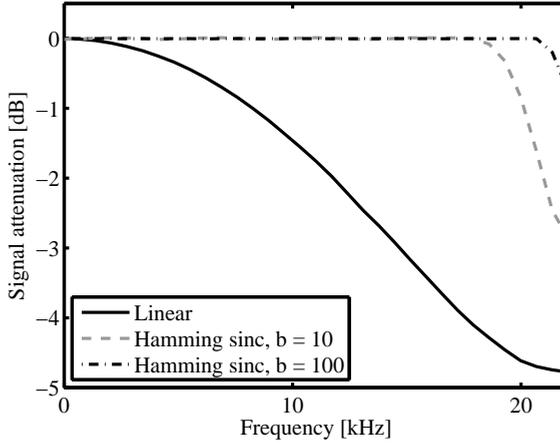


Figure 4.4: Spectral attenuation due to different resampling strategies.

The Doppler amplification is given by the Doppler factor  $\mathcal{D}$ . For a subsonically moving volume source, the combination of the Doppler and the convective amplification in the far-field amounts to  $\mathcal{D}^2$  [40]. At times  $t = \pm\infty$ , the amplification in dB is given by

$$G_{\text{mot}}(t = \pm\infty) = 20 \log_{10} \mathcal{D}^2(t = \pm\infty) = -40 \log_{10} (1 \pm M) \quad [\text{dB}] \quad (4.28)$$

Equation (4.28) yields an amplification of 2.3 dB at  $t = -\infty$  and an attenuation by 2.0 dB at  $t = \infty$ , resulting in a level difference of 4.3 dB across the pass-by. The numerical implementations of Equation (4.22) corresponded well with these theoretical values.

### Ground effect

In Equation (4.20), the ground effect is modeled in a physical way as the interference between direct and ground reflected sound. A flat topography is assumed, i.e. only one ground-reflected path is modeled, which is implemented by adding a second signal path. The ground-reflected sound differs from direct sound by scaling with its propagation distance and a complex reflection factor, as well as an additional delay. The complex reflection factor depends on frequency, geometry and ground surface type and is realized by the filter  $h_{\text{SWRC}}^{t'}$ .  $h_{\text{SWRC}}^{t'}$  is the impulse response of the spherical wave reflection coefficient at an infinite locally-reacting surface. The ground surface is acoustically described by a frequency-dependent surface impedance, for which the widely-used empirical model of Delany and Bazley [44] was used.

In chapter 3 [4], the additional delay of the ground-reflected sound was modeled by a digital delay of integer length. However, in this application, due to the higher relative source speed and short delays, audible artifacts (“zipper noise”) occur.

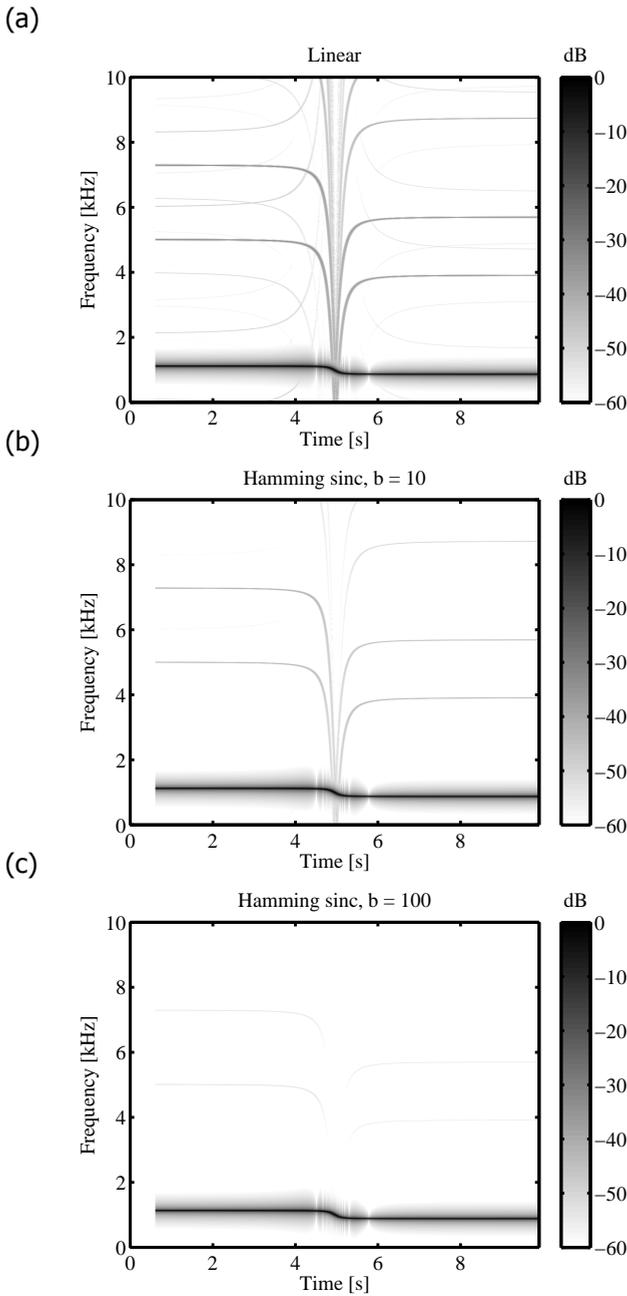


Figure 4.5: Non-linear distortions (aliasing) due to different resampling strategies: Linear interpolation (a) and Hamming sinc interpolations with different filter lengths,  $b = 10$  for (b) and  $b = 100$  for (c), respectively. The simulation was performed for a source emitting a 1 kHz pure tone that travels at constant speed  $V = 150$  km/h and passes a static receiver at a distance of  $D = 7.5$  m.

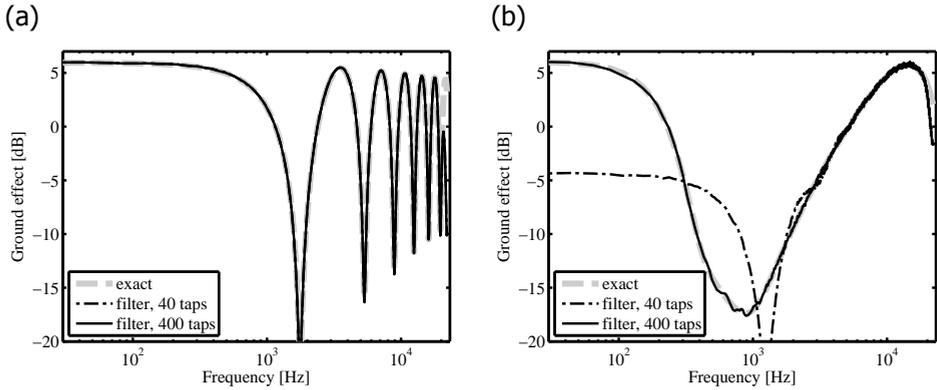


Figure 4.6: Simulated ground effect spectra for a point source at height 0.3 m in the reference situation (a) with a receiver at height 1.2 m at a horizontal distance of  $D = 7.5$  m and propagation over hard ground (flow resistivity 20'000 kPa s m<sup>-2</sup>) and (b) with a receiver at height 2 m at horizontal distance  $D = 100$  m and propagation over grassy ground (flow resistivity 200 kPa s m<sup>-2</sup>).

Therefore, a separate resampling is performed in Equation (4.22) for the ground-reflected sound. Furthermore, this type of processing eliminates the spectrally-fluctuating errors (see Figure 3.10).

The spherical wave reflection coefficient filter  $h_{\text{SWRC}}^{t'}$  is implemented by an FIR filter designed using the inverse FFT, as described in chapter 3 and [4]. It is important that the filter lag is compensated. However, as compared to wind turbines for this application, substantially more filter taps are required to reproduce the correct interference pattern. Figure 4.6 shows simulation results for the standard configuration of road traffic noise emission measurements (a) and a receiver point at distance  $D = 100$  m and height 2 m with sound propagation over grassy ground (b). For the former case a filter with 40 taps is sufficient, as the difference to the simulation with a filter with 400 taps stays well below 1 dB for all frequencies (nearly perfect coincidence of curves in Figure 4.6a). For the latter case, however, Figure 4.6b shows that such a short filter is not able to correctly reproduce the interference pattern and creates large errors at mid and low frequencies. A filter length of 400 taps allows simulations that are in good agreement with the exact solution for both cases. Large errors only occur near the Nyquist frequency. An update interval of the filter coefficients of 200 ms is used.

### Air absorption

For performance reasons, the identical air absorption filter  $h_{\text{air}}^{t'}$  is applied to the direct and ground reflected path in Equation (4.20).  $h_{\text{air}}^{t'}$  are linear-phase FIR filters designed using the inverse FFT, as described in chapter 3 and [4]. The frequency-dependent sound attenuation coefficients for atmospheric absorption as a function of relative humidity and temperature are calculated according to the standard ISO 9613-1 [45]. A filter length of 30 taps is used with an update interval of the

filter coefficients of 200 ms.

#### 4.2.4. Reproduction rendering

The rendering of the immission signals for reproduction strongly depends on the type of reproduction system. For surround reproduction via multiple loudspeakers, techniques, such as Ambisonics [46, 47] or triplet-wise amplitude panning (e.g., Vector Base Amplitude Panning (VBAP) [48] or Multiple-Direction Amplitude Panning (MDAP) [49]), are possible candidates. For binaural reproduction over headphones, generally, head-related transfer functions (HRTF) should be applied. Here, a simulation of the ORTF stereo technique [50, 51] is used. If the listener is facing the road, this allows for a reproduction with sufficient accuracy via headphones and a stereo speaker set-up. The cardioid microphone pattern and the time difference between the left and right channel,  $C_L$  and  $C_R$ , are modeled by:

$$C_L(t') = 0.5 [1 + \cos(\theta(t') - 55^\circ)] p_r(t' + \Delta_{LR}t) \quad (4.29)$$

$$C_R(t') = 0.5 [1 + \cos(\theta(t') + 55^\circ)] p_r(t') \quad (4.30)$$

with the receiver signal  $p_r(t')$  from Equation (4.20) and the time-varying time difference:

$$\Delta t_{LR}(t') = \frac{d_{\text{ORTF}} \sin(\theta(t'))}{c_0} \quad (4.31)$$

with the distance  $d_{\text{ORTF}} = 0.17$  m being related to the size of the human head. The immission angle  $\theta(t')$  has to be evaluated on the receiver time axis (see Equation (4.21)).  $p_r(t' + \Delta_{LR}t)$  is calculated using a windowed sinc interpolation strategy according to Equation (4.24). As a consequence of this interpolation, high-frequency attenuation, as shown in Figure 4.4, and nonlinear distortions are introduced to the left channel.

### 4.3. Model parameter estimation

This section presents procedures to obtain the model parameters of the emission synthesizer described in section 4.2.2. The procedures are based on controlled measurements. The following sections describe the measurements, as well as the signal processing that is applied to the acquired data.

#### 4.3.1. Tire noise

The emission parameters for tire noise were obtained from pass-by measurements with idling engine. For an individual tire type, pass-bys by the same passenger car at different speeds were recorded at a sampling frequency of  $f_s = 44.1$  kHz with a calibrated measurement microphone in a set-up referring to the standard ISO 11819-1 [16] and depicted in Figure 4.7(a). The pass-by speed was measured by radar, and the pass-by time was determined from synchronous video. Under the assumption of constant speed, a time-dependent backpropagation to the source was performed.

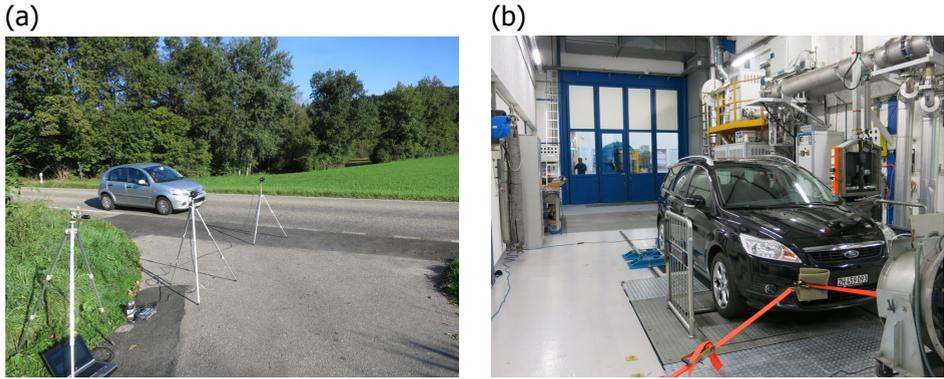


Figure 4.7: Photographs showing the measurements set-ups for tire noise (a) and propulsion noise (b). In (a) the coast-by situation is depicted with two measurement microphones placed at different distances and a camera connected to a laptop. (b) shows the lab with a passenger car on the chassis dynamometer, the airstream fan in front of the car and two microphones on the floor at the left-hand room edge (emission angles  $\varphi = 60^\circ$  and  $120^\circ$ ).

Thereby, two equal incoherent point sources at the nearby wheels were assumed, i.e. placed at the side of the car, horizontally separated by the wheelbase and set on the ground. For the temporal accordance, the sound propagation delay, as well as the filter group delays of the 1/3 octave band filters have to be taken into account. Consequently, emission sound pressure levels at reference distances  $d_0$  were obtained by integration over an emission angle range of  $90^\circ$ . Applying a logarithmic transformation to the measured pass-by speeds, the linear regression parameters  $A_i$  and  $B_i$  of Equation (4.4) were fitted in a least-squares sense. Despite the idling engine, some low 1/3 octave bands were contaminated by the engine sound. To correct for this, in the first step, for each band, a quality criterion based on the correlation coefficient and the slope of the regression line was deployed. Adverse bands were identified and imputed based on the values of adjacent valid bands. In the second step, low-frequency peaks of  $A_i$  were smoothed. Figure 4.8 contains measured parameters  $A$  and  $B$  in 1/3 octave bands of 13 tires.

#### 4.3.2. Propulsion noise

To obtain the emission synthesizer parameters of the propulsion noise, controlled measurements on a chassis dynamometer (see Figure 4.7(b)) and at idling engine under free field conditions were performed. Calibrated audio recordings at a sampling frequency of  $f_s = 44.1$  kHz at different microphone positions around the vehicle and at different engine conditions were taken. During the measurements on the chassis dynamometer, four microphones were placed on the ground at emission angles  $\varphi \approx \{0, 60, 120, 180\}^\circ$  at distances  $d = 1$  to 2 m from the vehicle. During the free field measurements, four additional microphones were placed on the ground at the identical emission angles, but at larger distances of  $d' = 4.5$  to 7 m. The free field measurements were used to correct for the room

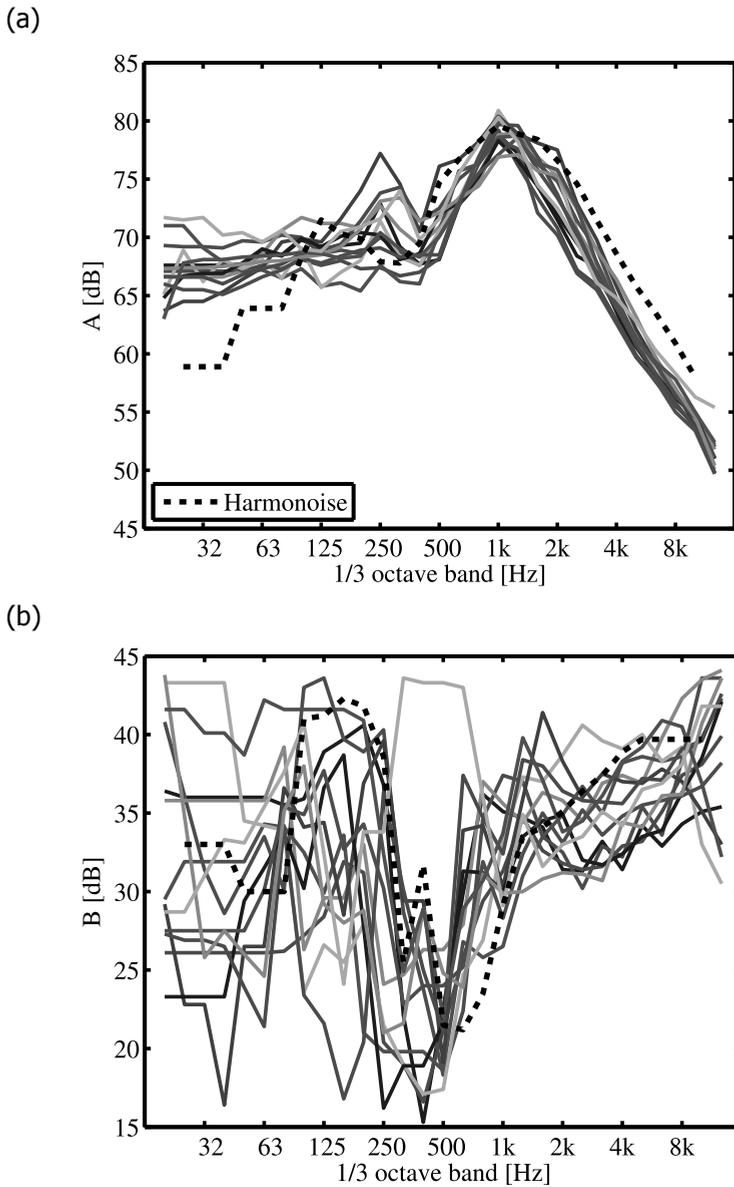


Figure 4.8: Measured tire noise regression parameters  $A_i$  (a) and  $B_i$  (b) of 13 tires and the values according to the Harmonoise model [11] (dotted lines).

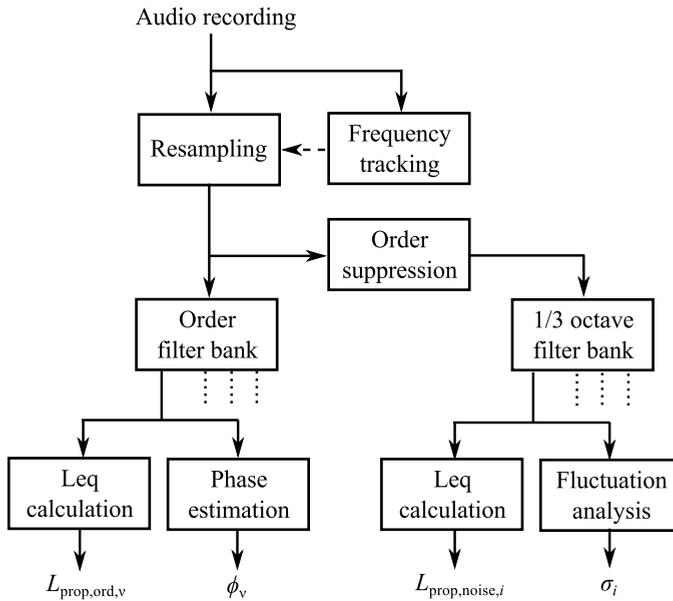


Figure 4.9: Signal analysis flowchart to obtain the synthesizer parameters of propulsion noise as described in section 4.2.2 from audio recordings.

influences of the lab. On the chassis dynamometer, measurements were typically taken at engine speeds of  $n_{rot} \approx \{1000, 2000, 3000, 4000\}$  rpm and engine loads of  $\Gamma \approx \{0, 40, 70, 100\}\%$ . The measuring point pairs  $\langle n_{rot}, M \rangle$  of a measurement performed on a Ford Focus 1.8i are depicted as circles in Figure 4.18. The topmost points at 1000 to 4000 rpm are at full load, i.e.  $\Gamma = 100\%$ . To confine tire noise, low vehicle speeds were aimed for by choosing low driving gears. Mostly, it was the second gear, which resulted in vehicle speeds  $< 50$  km/h.

To these recordings, a series of signal analysis steps were applied, which are outlined in a signal flowchart in Figure 4.9. These steps are further explained in the following sections. For the signal processing, a signal length of 4 s is used.

### Resampling

The emission synthesizer uses detailed information about the engine orders. These parameters are obtained by a narrowband analysis. Although during the measurements, the engine speed was kept fairly constant, the instantaneous order frequencies slightly fluctuate as exemplarily shown in Figure 4.10a. This figure shows the spectrogram of a recording made at the rear of a car with an inline, four cylinder engine idling at 1100 rpm. To be able to separate engine orders and broadband noise by the narrowband analysis, a preceding resampling of the slightly non-stationary signals is performed. In order to actuate the resampling process, the instantaneous ignition frequency,  $f_{ign}(t)$ , of the engine is required. This data are extracted from the audio recordings.

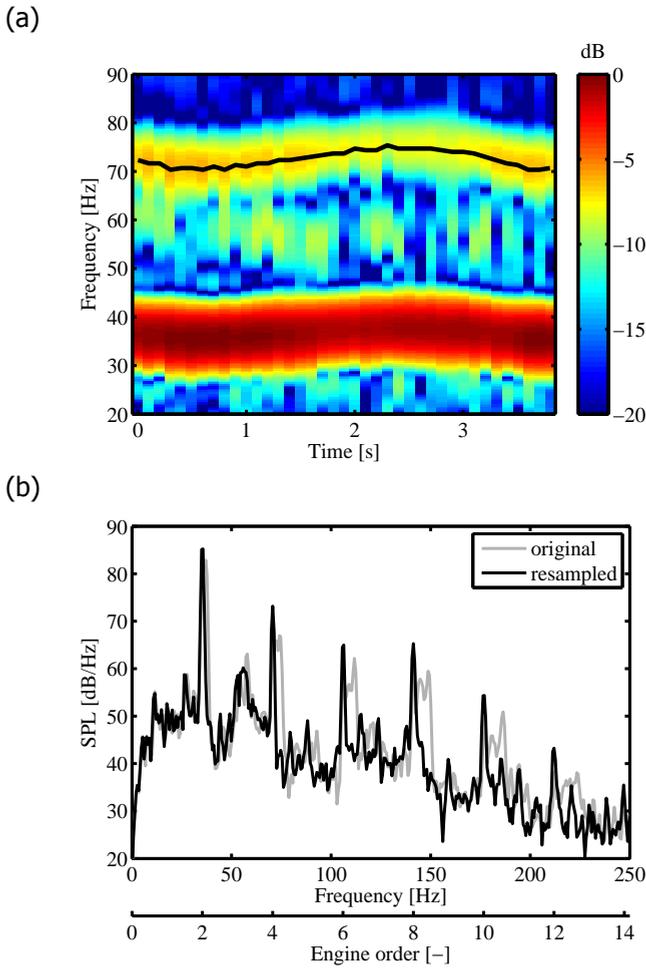


Figure 4.10: Normalized spectrogram (a) of measured sound pressure signal with tracked double ignition frequency (drawn as black line) and power spectral density (b) of the original and asynchronously resampled sound pressure signal, respectively. The recording was conducted at the rear of BMW with an inline, four cylinder engine idling at 1100 rpm.

In the first step, the average ignition frequency is estimated. From the signal taken closest to the exhaust, the power spectral density (PSD) with a frequency resolution  $< 1$  Hz is calculated. Based on the rough indication of the engine speed taken from the car's tachometer, a first estimate of the ignition frequency,  $f_{\text{ign}}$ , is obtained using Equation (4.19). The location of the maximum value of the PSD within a search range around this frequency yields a better, second estimate. Particularly, for low engine speeds, at which the ignition frequency can be as low as 20 Hz, this estimate is still not precise enough due to the low relative resolution at low frequencies. Thus, this estimate is further enhanced by considering the double ignition frequency,  $2f_{\text{ign}}$ , (i.e. engine order  $o = 4$  in Equation (4.16) for a four-cylinder engine) within a smaller range of the PSD.

In the second step, this information is used to track the course of the ignition frequency,  $f_{\text{ign}}(t)$ . This task is generally known as pitch detection [52–54]. A wide variety of algorithms exist that work in the time or frequency domain or a combination of them. In our application, a spectral method was established, in which the course of one discrete frequency component (i.e. an engine order) is tracked in a spectrogram. The spectrogram  $L(t, f)$  expressed in decibels is computed by the short-time Fourier transform (STFT):

$$L(t, f) = 10 \log_{10} \left( \left| \text{STFT}(p(t)) \right|^2 \right). \quad (4.32)$$

The STFT is calculated using the FFT. Windows of 200 ms with a 50% overlap, i.e. a temporal resolution of  $\Delta t = 100$  ms, are multiplied by a Hann window function. To obtain a high frequency resolution of  $\Delta f < 0.5$  Hz, the signals are zero padded. A section of such a spectrogram is depicted in Figure 4.10a. Within the spectrogram, the “maximum cost” path between time  $t = 0$  and  $t = T$ , within a certain frequency range around a reference frequency,  $f_r$ , is sought.  $f_r$  is chosen to be the first multiple of the mean ignition frequency above 55 Hz. This is a compromise between signal power and frequency localization: typically, the power decreases for increasing even orders (see Figure 4.10b), but higher orders exhibit larger absolute frequency variations (see Figure 4.10a). In the example of Figure 4.10,  $f_r$  is 71 Hz (corresponding to the fourth engine order), as the mean ignition frequency lies at 35 Hz.

The optimization task is solved by dynamic programming, which breaks the complex problem down into many simple subproblems. This method prevents taking possible wrong local decisions and guarantees that the best solution is found. A well-known algorithm that uses dynamic programming is dynamic time warping (DTW), which is often applied in, e.g., automatic speech recognition (ASR). Additionally, we make use of the *a priori* knowledge that the engine speed does not change rapidly over time. This is introduced as a requirement on the slope of the optimal path,  $f_{\text{opt}}$ . The algorithm described below is based on an algorithm developed for object tracking in video data [55]. Within the search section of the discrete spectrogram, the local score  $q$  is calculated by:

$$q(m, l) = L(m, l) - \min_{m, l} (L(m, l)) \quad (4.33)$$

for which holds  $q \geq 0$ . From  $q$  for each time step  $m = \{1 \dots \tilde{M}\}$  and frequency bin  $l$ , the global score  $Q(m, l)$  is recursively computed by:

$$Q(m, l) = q(m, l) + \max_{l' \in \{l - \tilde{c} \leq l' \leq l + \tilde{c}\}} Q(m - 1, l') \quad (4.34)$$

with a positive integer  $\tilde{c}$  realizing the requirement:

$$\left| \frac{df_{\text{opt}}(t)}{dt} \right| \leq \tilde{C} \quad (4.35)$$

on the maximal absolute slope of  $f_{\text{opt},r}$  with the chosen parameter  $\tilde{C}$  given in Hertz per second and with

$$\tilde{c} \geq \tilde{C} \frac{\Delta t}{\Delta f} \quad (4.36)$$

with the corresponding time and frequency resolutions,  $\Delta t$  and  $\Delta f$ , respectively.

For the forward processing described by Equation (4.34), the starting condition is that the initial global score,  $Q(1, l)$ , is set to the local score, i.e.  $Q(1, l) = q(1, l)$ . During the evaluation of Equation (4.34), it is essential that the back pointers:

$$B(m, l) = \arg \max_{l' \in \{l - \tilde{c} \leq l' \leq l + \tilde{c}\}} Q(m - 1, l') \quad (4.37)$$

to the optimal predecessors are stored. From the global score  $Q$ , the end point of the optimal path is found by:

$$f_{\text{opt}}(\tilde{M}) = \arg \max_l Q(\tilde{M}, l). \quad (4.38)$$

Using the back pointers  $B$ , the optimal path can be traced back using a recursive procedure known as backtracking:

$$f_{\text{opt}}(m - 1) = B(m, f_{\text{opt}}(m)) \quad (4.39)$$

The annex of [1] provides a simple MATLAB code, which solves Equations (4.34) to (4.39). Figure 4.10 shows the optimal path drawn as a black line following the frequency component around 70 Hz.

In the third step, the sound pressure signal is asynchronously resampled based on the course of the tracked engine order. The warped time axis is calculated by:

$$t_{\text{warp}} = t + \frac{f_{\text{opt}}(t)}{f_r} \quad (4.40)$$

where  $f_{\text{opt}}(t)$  is the linearly-interpolated version of  $f_{\text{opt}}(m)$ . For the resampling of the sound pressure signals, a windowed sinc interpolation, as described by Equation (4.24), is adopted. Figure 4.10 illustrates the effect of the asynchronous resampling on the power spectral density. In contrast to the original signal, for the resampled signal, all even engine orders from two to 12 can be clearly identified as equidistant, narrow peaks.

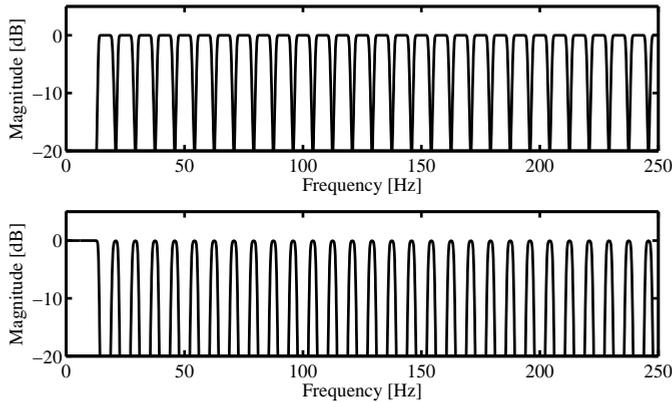


Figure 4.11: Magnitude frequency response of engine order analysis filter bank (top) and engine order suppression filter (bottom) for engine speed  $n_{\text{rot}} = 1000$  rpm, engine orders  $o = 1, 1.5, 2, \dots, 15$  and  $N_{\text{cyl}} = 4$ .

### Order analysis

From the resampled signals, information about the engine orders is extracted. Therefore, a filter bank consisting of one bandpass filter per considered engine order is generated and applied to the signal. Eighth order Butterworth filters centered around the engine order frequency  $f_o$  with a 6 Hz bandwidth are employed. Figure 4.11 shows the magnitude frequency response of the filter bank. At the output of each filter, the corresponding order level,  $L_{\text{prop,ord},o}$  in Equation (4.14), is calculated as an equivalent continuous level (Leq). Figure 4.12 exemplifies measured order levels at idling engine and full load recorded at the rear of a VW Touran running at 1000 and 3000 rpm.

The order phases are detected using the cross-correlation function. Since the above-described infinite impulse response (IIR) filter bank introduces phase shifts, the outputs of the filter bank are time reversed and sent once again through the same filter bank and, finally, time reversed. In doing so, a zero-phase forward and reverse digital IIR filtering is implemented. This signal,  $g_o(t)$ , is cross-correlated with a prototype function  $\cos(2\pi f_o t)$  to obtain the time shift:

$$\kappa_o = \arg \max_{\tau} \left\{ \int g_o(t + \tau) \cos(2\pi f_o t) dt \right\} \quad (4.41)$$

from which the phase shift of Equation (4.15) can be derived as:

$$\phi_o = -2\pi f_o \kappa_o. \quad (4.42)$$

Figure 4.13 compares the sound pressure signals of a recording and the corresponding synthesis consisting of engine orders with constant phases, which were estimated by Equation (4.42).

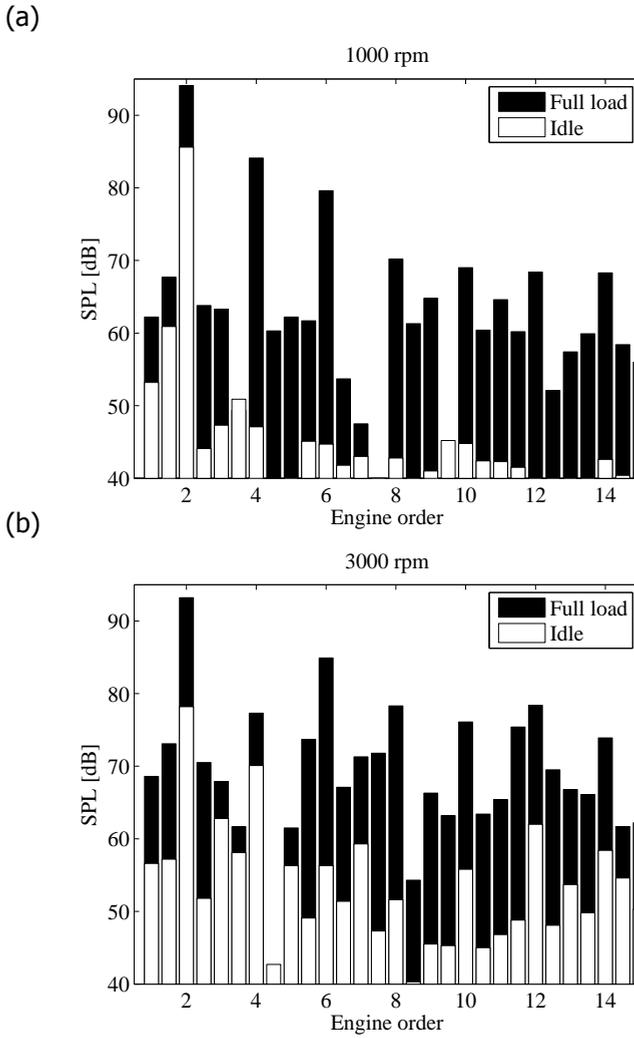


Figure 4.12: Comparison of engine order levels with idling engine (white) and full load (black) at 1000 (a) and 3000 rpm (b). Recorded at the rear of a VW Touran 1.6 FSI.

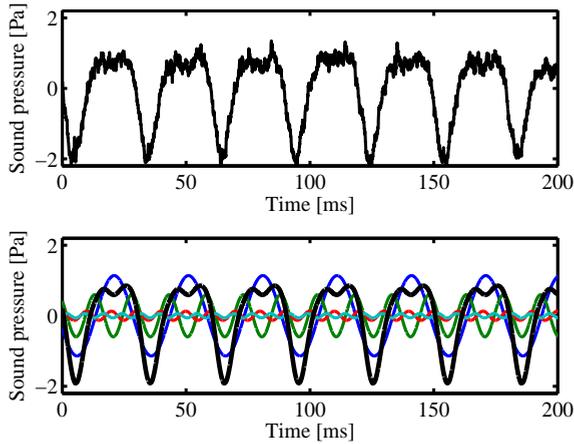


Figure 4.13: Comparison of sound pressure signals of a recording (top) and the corresponding synthesis consisting of engine orders with estimated phases (bottom). For the purpose of illustration only the four dominant engine orders (colored lines, and in black their sum) are used. The recording was conducted at the rear of a Ford Focus 1.8i at 1000 rpm and full load.

### Noise analysis

The noise levels and their short-term level fluctuations are obtained by a series of filtering operations. Starting with the resampled signal, in a first attempt, the engine orders are suppressed using cascaded notch filters. These filters are designed analogously to the engine order filter bank from the previous section, except that instead of bandpass filters, band-stop filters are generated (see Figure 4.11). Figure 4.14 shows two power spectral densities, which illustrate the effect of the order suppression filter. After this operation, the signal is split into sub-bands for further analysis. The signal is therefore decomposed into 1/3 octave bands using a 1/3 octave band filter bank. Each of the  $N_b$  filters yields a signal  $p_{\text{noise},i}(t)$ .

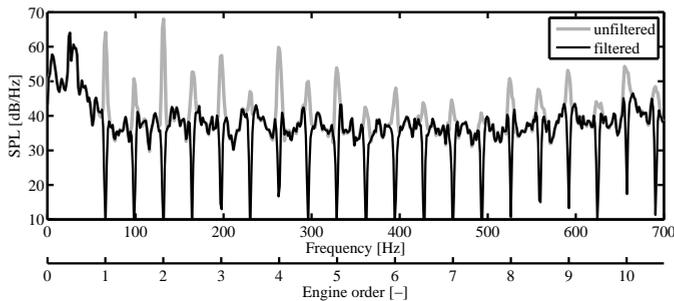


Figure 4.14: Power spectral densities illustrating the effect of the order suppression filter which is applied to a recording of an inline, four cylinder engine idling at 4000 rpm.

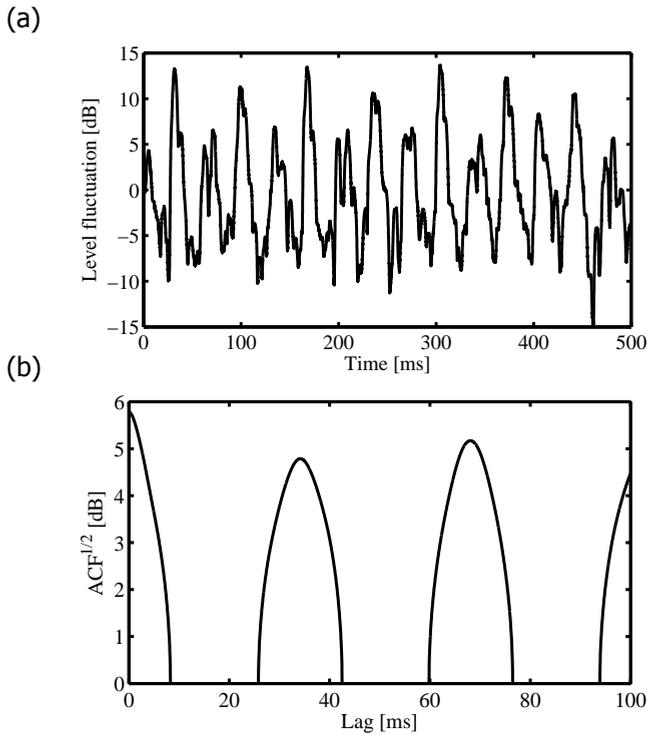


Figure 4.15: Level fluctuation signal,  $F_{\text{noise},i}(t)$ , (a) of the 2.5 kHz 1/3 octave band and its square root of the autocorrelation function (ACF) (b) from a recording taken at the front of an idling four-cylinder diesel engine at 870 rpm, corresponding to an ignition period of 34 ms. Panel (b) indicates that for this band the level standard deviation,  $\sigma_i$ , amounts to about 5 dB.

From  $p_{\text{noise},i}(t)$ , the noise levels,  $L_{\text{eq,prop,noise},i}$  in Equation (4.18), are calculated as Leqs. Moreover, from  $p_{\text{noise},i}(t)$ , using a moving average filter, smoothed sound pressure level-time curves:

$$L_{\text{noise},i}(t) = 10 \log_{10} \left( \frac{1}{T} \int_0^T \frac{p_{\text{noise},i}^2(t + \tau)}{p_0^2} d\tau \right) \quad (4.43)$$

are calculated with the reference pressure  $p_0$  and using a window length of  $T = 4$  ms. Subsequently, from  $L_{\text{noise},i}$ , the mean value is subtracted to obtain a DC-free level fluctuation signal:

$$F_{\text{noise},i}(t) = L_{\text{noise},i}(t) - \overline{L_{\text{noise},i}(t)}. \quad (4.44)$$

Figure 4.15a exemplifies such a fluctuation signal for the 2.5 kHz band recorded at the front of a diesel engine car. The periodic structure is clearly visible. Following section 3.2.2 [24], the autocorrelation function (ACF),  $R$ , is used to estimate the standard deviations  $\sigma_i$  (used in Equation (4.18)) of the level fluctuations with period  $1/f_{\text{ign}}$  by:

$$\sigma_i^2 = R_{F_{\text{noise},i}} \left( \frac{1}{f_{\text{ign}}} \right). \quad (4.45)$$

Figure 4.15b shows the square root of the ACF of the level fluctuation signal depicted in Figure 4.15(a). Clear peaks can be observed at lag zero and multiples of the ignition period of 34 ms. The standard deviation  $\sigma_i$  for the frequency band  $i = 2.5$  kHz amounts to about 5 dB. The fact that a higher peak appears at the double ignition period, at 68 ms, indicates that the signal contains an additional level modulation with a modulation frequency equal to the half ignition frequency. This can also be observed in Figure 4.15a where every second peak is about 5 dB higher than the previous one.

Figure 4.16 shows measured spectra of the standard deviations  $\sigma_i$ . The measurements were performed in front of five cars idling at low engine speeds. It can be seen that diesel cars feature higher values compared to gasoline engine cars. This finding corresponds to the increased rattling sound noticed in the field.

### Background noise corrections

On the test rig, the main background noise sources were the tire noise, the airstream fan, the room ventilation and the dynamometer itself. Firstly, to confine the tire noise, the measurements were performed at low vehicle speeds, i.e. low gears. Secondly, during the measurements, the airstream fan (depicted in Figure 4.7) was briefly switched off for periods of about 10 s. However, the dropping tonal components of the fan still strongly interfered with the propulsion noise of the car (see Figure 4.17). Therefore, several shifted analysis time windows were deployed, and the minimal levels and the maximum level standard deviations,  $\sigma$ , across these windows were exploited. Thirdly, background noise measurements with a switched off engine at different vehicle speeds were performed. For each ordinary measurement, the corresponding background noise was identically analyzed and used for level corrections.

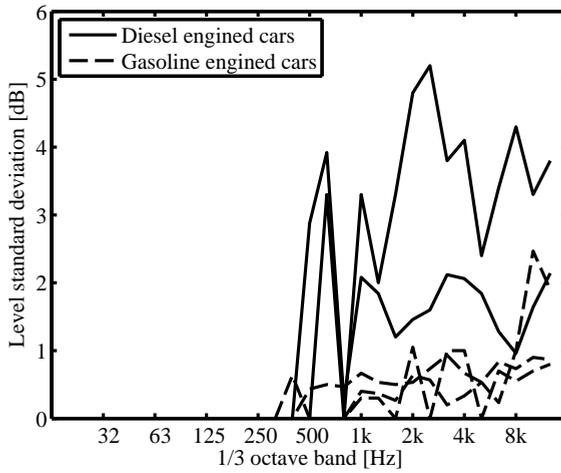


Figure 4.16: Level fluctuation standard deviations,  $\sigma$ , in 1/3 octave bands measured at the front of two diesel engines and three gasoline engines idling at 900 rpm.

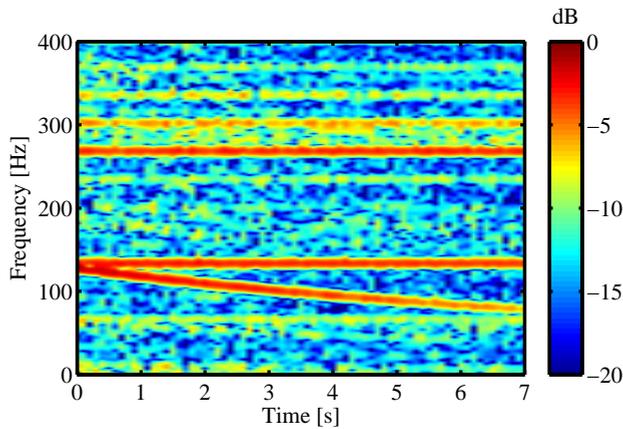


Figure 4.17: Normalized spectrogram of a microphone position in front of a Ford Focus 1.8i at 4000 rpm and full load on the dynamometer. The dropping tonal component around 100 Hz stems from the briefly switched off airstream fan of the lab.

### Backpropagation

As the interface to the propagation model, the emission signals are defined at a (virtual) reference distance of  $d_0 = 1$  meter from the source position. For the measured levels  $L_{\text{meas}}$ , the following inverse sound propagation model is used:

$$L_e = L_{\text{meas}} + 20 \log_{10} \left( \frac{d_{\text{Ac}}}{d_0} \right) + A_{\text{room}} + A_{\text{gr}} \quad (4.46)$$

with the ground effect  $A_{\text{gr}} = -6$  dB, as all microphones were mounted on the ground. For the microphones placed close to the room edge (emission angles  $\varphi = 60^\circ$  and  $120^\circ$ ), the room correction  $A_{\text{room}}$  was set to  $-6$  dB for frequency bands below 1 kHz and to  $-3$  dB otherwise. As for the microphones placed in front and at the back of the car, the distance to the closest wall was about three meters,  $A_{\text{room}}$  was set to 0 dB for these signals.  $d_{\text{Ac}}$  is the distance to the acoustical center, which, by assuming geometrical spreading of a point source, is obtained by simultaneous free field measurements at two points at distances  $d$  and  $d'$  ( $d' > d$ ) with:

$$d_{\text{Ac}} = \frac{d' - d}{1 - 10^{(L_{\text{ff},d'} - L_{\text{ff},d})/20}} \quad (4.47)$$

Parameter  $d_{\text{Ac}}$  is evaluated separately for each emission angle, engine speed and frequency band (or engine order, respectively).

## 4.4. Evaluation

To evaluate and demonstrate the proposed model, two car pass-bys are auralized in this section.

Figure 4.18(a) and 4.18(b) show two simulated engine condition courses within an engine load vs. engine speed diagram. The black lines show the simulated temporal progression of the engine condition during a virtual pass-by. Both simulations start at the same initial engine condition of 900 rpm and 3 Nm, marked by green stars. In both cases, the passenger car starts in first gear and accelerates from  $V = 7$  to 50 km/h, however with differing accelerations  $a$  and driving styles. For the medium driving style and an acceleration of  $a = 1$  m/s<sup>2</sup> (Figure 4.18a), three gear changes occur at around 2000 rpm (Sample 1); whereas for the sporty driving style and an acceleration of  $a = 2$  m/s<sup>2</sup> (Figure 4.18b), only two gear changes happen, but at higher engine speeds of 3000 to 4000 rpm (Sample 2). The temporal behaviors of the engine states for these two examples are published as videos in appendix A and the supplementary file in [1].

Figure 4.18(c) and (d) show normalized spectrograms of two synthesized pass-by sounds. For the synthesis of the sound of propulsion, the respective engine condition courses depicted in Figure 4.18(a) and (b) were used. In both simulations, the car passes the receiver at time 0 at 30 km/h. The receiver is located 1.2 m above a hard ground at a distance  $D = 7.5$  m. As at the pass-by, the engine speed still increases, the Doppler frequency shift is not directly observable in the course of the order frequencies. However, the gear change moments can be well observed

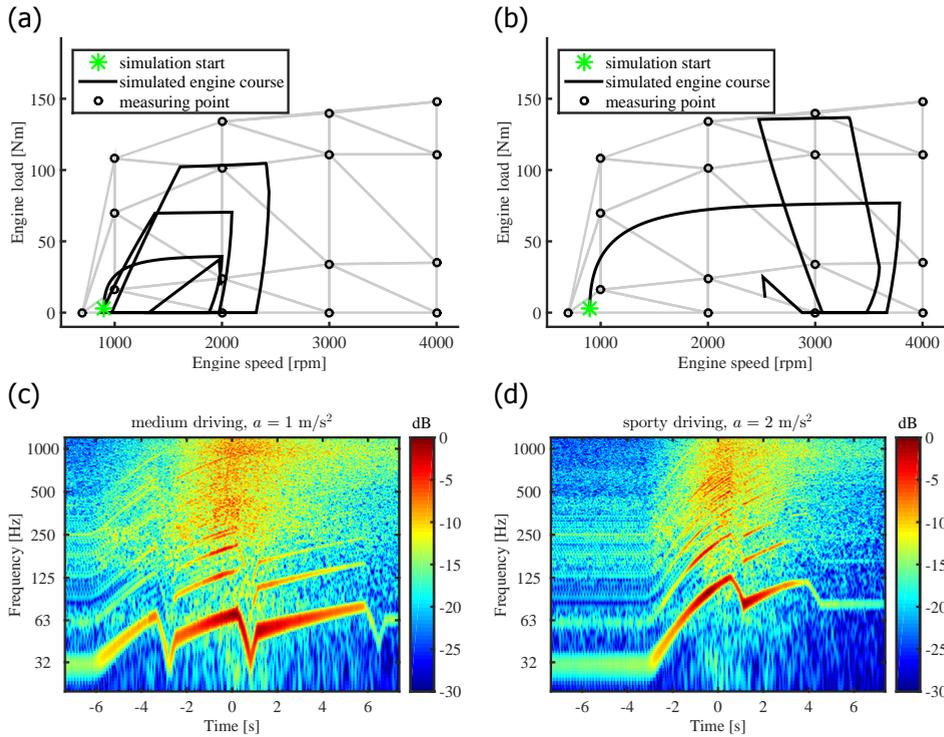


Figure 4.18: Simulation results: The upper graphs, (a) and (b), show two simulated engine condition courses of an accelerating Ford Focus 1.8i with different accelerations and driving styles. The gray triangles show the interpolation grid spanned by the measuring points marked as circles, as introduced in section 4.2.2. The lower two graphs, (c) and (d), show the spectrograms of the corresponding synthesized pass-by sounds (normalized to 0 dB). Their calculation is elucidated in section 4.2.3.

as local decreases of order frequencies. As a consequence of the used engine condition courses, these frequency drops occur at higher frequencies for the sporty driving style (d).

Audio data of ORTF renderings of these two cases are published in the videos in appendix A and the supplementary file in [1].

## 4.5. Conclusions

In the proposed auralization model, emission sounds of accelerating passenger cars are artificially generated based on spectral modeling synthesis. Whereas the sound of tires is synthesized as stationary noise, which is time-dependently shaped in third octave bands, the realistic synthesis of sounds of propulsion requires more subtlety.

It is synthesized as the superposition of a noise component and tones. Frequency-dependent periodic short-term modulations are applied to the noise component in order to create a rattling sound eliciting a roughness sensation. The tones are related to the engine orders. It was found that a large number of engine orders are needed (50...100) to convincingly represent different engine speeds and loads. Moreover auralizations revealed that the order phases have to be included as synthesizer parameters. In conclusion, the presented emission synthesizer gives complete control over the signal characteristics, but is computationally much more demanding than a synthesizer based on the PSOLA method [3] with limited flexibility. However, a hybrid approach could profit from both advantages: Firstly, only the most relevant signal grains are pre-rendered using the described method. Secondly, these grains used as a basis for a granular synthesis or PSOLA approach.

Analysis of the propagation filtering algorithms yielded two main insights. Aliasing, arising from the simulation of the Doppler effect, can be reduced by incorporating a band-limited resampling strategy, such as the windowed sinc interpolation. Furthermore, due to the low source height, a significantly higher number of filter taps is needed to correctly simulate the ground effect in relevant situations, as compared to elevated sources, such as airplanes or wind turbines (see chapter 3).

We conclude that with the presented synthesizer structure, audio signals from vehicle pass-bys can be represented in a compact and elegant manner. To give the reader an impression of the subjective quality of the proposed model, auralizations of two examples are published as supplementary data in the videos in annex A and the supplementary file in [1].

## Acknowledgments

The work in this chapter was supported by the Swiss National Science Foundation (SNSF) (# 200021-149618).

## References

- [1] R. Pieren, T. Bütler, and K. Heutschi, *Auralization of accelerating passenger cars using spectral modeling synthesis*, [Applied Sciences](#) **6**, 5 (2016).

- [2] M. Vorländer, *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* (Springer, Berlin, 2008).
- [3] J. Jagla, J. Maillard, and N. Martin, *Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method*, *Journal of the Acoustical Society of America* **132**, 3098 (2012).
- [4] K. Heutschi, R. Pieren, M. Müller, M. Manyoky, U. Hayek Wissen, and K. Eggenschwiler, *Auralization of wind turbine noise: Propagation filtering and vegetation noise synthesis*, *Acta Acustica United with Acustica* **100**, 13 (2014).
- [5] M. Arntzen and D. G. Simons, *Modeling and synthesis of aircraft flyover noise*, *Applied Acoustics* **84**, 99 (2014).
- [6] A. Sahai, F. Wefers, P. Pick, E. Stumpf, M. Vorländer, and T. Kuhlen, *Interactive simulation of aircraft noise in aural and visual virtual environments*, *Applied Acoustics* **101**, 24 (2016).
- [7] D. Havelock, S. Kuwano, and M. Vorländer, eds., *Handbook of Signal Processing in Acoustics* (Springer, New York, 2008).
- [8] R. Pieren, T. Büttler, and K. Heutschi, *Auralisation of accelerating passenger cars*, in *Euronoise 2015, 10th European Conference on Noise Control* (Maastricht, Netherlands, 2015) pp. 757–762.
- [9] R. Pieren, J. M. Wunderli, A. Zemp, S. Sohr, and K. Heutschi, *Auralisation of railway noise: A concept for the emission synthesis of rolling and impact noise*, in *INTER-NOISE 2016, 45th International Congress and Exposition of Noise Control Engineering* (Hamburg, Germany, 2016) pp. 3466–3472.
- [10] R. Pieren, K. Heutschi, J. M. Wunderli, M. Snellen, and D. G. Simons, *Auralization of railway noise: Emission synthesis of rolling and impact noise*, *Applied Acoustics* **127**, 34 (2017).
- [11] H. G. Jonasson, *Acoustical source modelling of road vehicles*, *Acta Acustica United with Acustica* **93**, 173 (2007).
- [12] K. Heutschi, *Sonroad: New swiss road traffic noise model*, *Acta Acustica United with Acustica* **90**, 548 (2004).
- [13] U. Sandberg and J. A. Ejsmont, *Tyre/Road Noise Reference Book* (INFORMEX, Harg, Kisa, Sweden, 2002).
- [14] P. Zeller, ed., *Handbuch Fahrzeugakustik - Grundlagen, Auslegung, Berechnung, Versuch*, 2nd ed. (Vieweg+Teubner Verlag, Wiesbaden, 2012).
- [15] United Nations, *UN/ECE Regulation No. 117: Uniform provisions concerning the approval of tyres with regard to rolling sound emissions and/or to adhesion on wet surfaces and/or to rolling resistance*, (2005).

- [16] International Organization of Standardization, *ISO 11819-1: Acoustics – Measurement of the influence of road surfaces on traffic noise - Part 1: Statistical Pass-By method*, (1997).
- [17] J. H. Schutte, Y. H. Wijnant, and A. de Boer, *The influence of the horn effect in tyre/road noise*, *Acta Acustica United with Acustica* **101**, 690 (2015).
- [18] European Commission, *Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council*, *Official Journal of the European Union* **58** (2015).
- [19] C. Roads, ed., *The Computer Music Tutorial* (MIT Press, Cambridge, 1996).
- [20] T. D. Rossing, ed., *Springer Handbook of Acoustics* (Springer, New York, 2007).
- [21] J. O. Smith III, ed., *Spectral Audio Signal Processing* (W3K Publishing, Stanford, 2011).
- [22] X. Serra and J. Smith III, *Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition*, *Computer Music Journal* **14**, 12 (1990).
- [23] X. Serra, *Musical sound modeling with sinusoids plus noise*, in *Musical Signal Processing*, edited by C. Roads, S. Pope, A. Piccilli, and G. De Poli (Swets & Zeitlinger, Lisse, The Netherlands, 1997).
- [24] R. Pieren, K. Heutschi, M. Müller, M. Manyoky, and K. Eggenschwiler, *Auralization of wind turbine noise: Emission synthesis*, *Acta Acustica United with Acustica* **100**, 25 (2014).
- [25] International Electrotechnical Commission, *IEC 1260:1995: Electroacoustics - Octave-band and fractional-octave-band filters*, (1995).
- [26] K. Reif, ed., *Bosch Automotive Handbook*, 8th ed. (Wiley, Chichester, 2011).
- [27] United Nations, *UN/ECE/TRANS/WP.29/2014/27: Proposal for a new global technical regulation on the Worldwide harmonized Light vehicles Test Procedure (WLTP)*, (2014).
- [28] International Organization of Standardization, *ISO 15031-5: Road vehicles – Communication between vehicle and external equipment for emissions-related diagnostics - Part 5: Emissions-related diagnostic services*, (2015).
- [29] Z. Maekawa, *Noise reduction by screens*, *Applied Acoustics* **1**, 157 (1968).
- [30] International Organization of Standardization, *ISO 9613-2: Acoustics – Attenuation of sound during propagation outdoors - Part 2: General method of calculation*, (1996).

- [31] van Maercke D. and J. Defrance, *Development of an analytical model for outdoor sound propagation within the harmonoise project*, Acta Acustica United with Acustica **93**, 201 (2007).
- [32] G. Daigle, T. F. W. Embleton, and J. E. Piercy, *Propagation of sound in the presence of gradients and turbulence near the ground*, Journal of the Acoustical Society of America **79**, 613 (1986).
- [33] J. Hofmann and K. Heutschi, *An engineering model for sound pressure in shadow zones based on numerical simulations*, Acta Acustica United with Acustica **91**, 661 (2005).
- [34] J. M. Wunderli, R. Pieren, and K. Heutschi, *The Swiss shooting sound calculation model sonARMS*, Noise Control Engineering Journal **90**, 224 (2012).
- [35] K. Heutschi, *Calculation of reflections in an urban environment*, Acta Acustica United with Acustica **95**, 644 (2009).
- [36] J. M. Wunderli, *An extended model to predict reflections from forests*, Acta Acustica United with Acustica **98**, 263 (2012).
- [37] R. Pieren and J. M. Wunderli, *A model to predict sound reflections from cliffs*, Acta Acustica United with Acustica **97**, 243 (2011).
- [38] A. Peplow, J. Forssén, P. Lundén, and M. E. Nilsson, *Exterior auralization of traffic noise within the listen project*, in Forum Acusticum 2011, European Conference on Acoustics (Aalborg, Denmark, 2011) pp. 665–669.
- [39] J. Maillard and J. Jagla, *Real time auralization of non-stationary traffic noise - quantitative and perceptual validation in an urban street*, in AIA-DAGA 2013 Conference on Acoustics (Merano, Italy, 2013).
- [40] P. M. Morse and K. U. Ingard, *Theoretical acoustics* (Mc Gray-Hill book company, New York, 1968).
- [41] J. Smith, S. Serafin, J. Abel, and D. Berners, *Doppler simulation and the leslie*, in Proc. of the 5th Int. Conference on Digital Audio Effects (DAFx-02) (Hamburg, Germany, 2002) pp. 932–937.
- [42] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, *Splitting the unit delay - tools for fractional delay filter design*, IEEE Signal Processing Magazine **13**, 30 (1996).
- [43] M. Bruneau, *Fundamentals of Acoustics* (ISTE Ltd, London, 2006).
- [44] M. E. Delany and E. N. Bazley, *Acoustical properties of fibrous absorbent materials*, Applied Acoustics **3**, 105 (1970).
- [45] International Organization of Standardization, *ISO 9613-1: Acoustics – Attenuation of sound during propagation outdoors - Part 1: Calculation of the absorption of sound by the atmosphere*, (1993).

- [46] M. Gerzon, *Periophony: With-height sound reproduction*, Journal of the Audio Engineering Society **21**, 2 (1973).
- [47] M. Gerzon, *Ambisonics in multichannel broadcasting and video*, Journal of the Audio Engineering Society **33**, 859 (1985).
- [48] V. Pulkki, *Virtual sound source positioning using vector base amplitude panning*, Journal of the Audio Engineering Society **45**, 456 (1997).
- [49] V. Pulkki, *Uniform spreading of amplitude panned virtual sources*, in Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (New York, 1999).
- [50] G. Ballou, ed., *Handbook for Sound Engineers*, 4th ed. (Elsevier, Amsterdam, 2008).
- [51] M. Dickreiter, V. Dittel, W. Hoeg, and M. Wöhr, eds., *Handbuch der Tonstudientechnik - Band 1*, 7th ed. (ARD.ZDF medienakademie, Nürnberg, 2008).
- [52] L. Rabiner, M. Cheng, A. Rosenberg, and C. McGonegal, *A comparative performance study of several pitch detection algorithms*, IEEE Transactions on Acoustics, Speech and Signal Processing **24**, 399 (1976).
- [53] P. De La Cuadra and A. Master, *Efficient pitch detection techniques for interactive music*, in International Computer Music Conference (La Habana, 2001) pp. 87–90.
- [54] D. Gerhard, *Pitch Extraction and Fundamental Frequency: History and Current Techniques*, Tech. Rep. TR-CS 2003-06 (Department of Computer Science, University of Regina, Canada, 2003).
- [55] P. Dreuw, T. Deselaers, D. Rybach, D. Keysers, and H. Ney, *Tracking using dynamic programming for appearance-based sign language recognition*, in IEEE International Conference on Automatic Face and Gesture Recognition (Southampton, UK, 2006) pp. 293–298.

# 5

## Auralization of Railway Noise: Rolling and Impact Noise

*This chapter focuses on an emission synthesizer for railway noise and presents a concept for rolling and impact noise. The synthesis is based on a physical approach in which the noise generation mechanism is modeled in the time domain. As a starting point, equivalent roughness patterns of each wheel and the rail are generated. These spatial signals are used to implicitly model the mechanical excitation of the wheel/rail system. Transfer paths describing the vibrational behavior and the radiation of wheels and rail are implemented as digital filters. This approach features a high degree of flexibility but requires knowledge of the detailed model parameters.*

## 5.1. Introduction

Railway traffic is a relevant source of environmental noise and leads to noise annoyance and sleep disturbances in the population [2, 3]. Recent sleep laboratory experiments related different acoustical metrics of single train pass-by events to sleep quality and recuperation [4, 5]. Based on auralization, different acoustical scenarios may be compared in terms of their perception using e.g. listening test experiments. For railway noise where noise mitigation measures are diverse and costly, such an assessment of the effectiveness of different measures could be helpful. Measures at the source as well as on the propagation paths are viable options. Therefore an auralization model for railway noise should be able to simulate both intervention types independently, i.e. only measures at the source, on the propagation or on both. This suggests to use separate source and propagation modules which is in line with Vorländer's definition of the auralization process [6].

Auralization models are either based on audio recordings or include a synthesizer that artificially generates audible signals. In contrast to architectural acoustics where mainly speech or music signals are used, in environmental acoustics it is often desirable to synthesize the emission signals instead of relying on audio recordings. The latter only allows for little variation of different signal aspects. A more versatile method with a much higher degree of freedom, as well as full control of the influencing signal parameters is to synthesize the sounds.

Railway noise consists of different contributions which may dominate in dependence on the vehicle and track type, traveling speed, frequency and geometry. It consists of rolling and impact noise as well as noise from secondary sources such as the traction, aggregates and aerodynamic noise [7]. Today, in most noise exposure situations rolling and impact noise dominate the A-weighted sound pressure level. Rolling noise is generated by very small amplitude undulations of the wheel and the rail running surfaces. The resulting varying contact forces excite the wheel/rail structure which consequently vibrates and radiates sound waves. The three main parts contributing to the radiated noise of the wheel/rail interaction are the wheels, the rails and the track sleepers. Impact noise arises due to discrete irregularities on these surfaces [8]. These transient sounds occur notably in the context of wheel flats [9], insulated rail joints or switches [10]. The frequency content of rolling and impact noise is very wide and it covers almost the whole audible frequency range. The maximum sound pressure level lies in the mid frequencies, typically in the range of 500 Hz to 2 kHz [7, 11]. Hence in an auralization, signals with a wide frequency content need to be utilized.

To date, only a few studies have been related to the auralization of railway noise. In [12] the sound quality of traction noise of starting vehicles was assessed using synthesized sounds. In [13] train pass-bys have been auralized based on a combination of filtered and resynthesized binaural recordings. Within the SILENCE project, as software called VAMPPASS was developed which features audio synthesis capabilities for rail vehicle pass-bys [14–16]. The tool models physical acoustic sources on the vehicle by equivalent point sources [17]. Initial attempts to auralize train pass-bys are also indicated in [18]. In [19] conventional beamforming was applied to obtain audio recordings of sub-sources during train pass-bys. These

recordings may be used as input data to synthesize source signals [20].

In the research project TAURA: Traffic Noise Auralisator (2014-2016) an auralization model for traffic noise was developed by the author. It covers road traffic [21, 22] and railway noise [1, 23]. It will form the basis for future listening test experiments to assess different noise mitigation measures. The objective of this chapter is to present the emission synthesizer for the rolling and impact noise components of railway noise. Section 5.2 shortly describes the railway noise measurement campaign that produced expedient data for the model development. In section 5.3 the auralization model is established and presented.

## 5.2. Measurements

In the years 2007 and 2008 a large railway noise measurement campaign was carried out in Switzerland in the context of the sonRAIL project. It involved 15 measurement sites and, along with the regular rail traffic, a dedicated measurement train. At all sites, sound pressure and rail accelerations were synchronously measured and passing axles were detected using light barriers. The measurement sites typically consisted of two-track sections (see [24] for the set-up). For each track, sound pressure was measured on both sides at the reference position 'A' according to the standard ISO 3095 [25], that is at a distance of 7.5 m from the centerline of the track and 1.2 m above rail. Furthermore, direct rail roughnesses, track decay rates and propagation attenuations were determined. A vertical microphone array was used for sound source separation.

In section 5.3, experimental data from one measurement point and the measurement train is shown (Figures 5.2, 5.3 and 5.15). The site was located in Lussy (Swiss National coordinates [CH1903+] 2'562'275/1'173'600) on the route Lausanne-Fribourg. The southern track was built in 1994 and consists of concrete monobloc sleepers on ballast substructure and UIC60 rails. Direct roughness measurements from August 2007 yielded weighted roughness levels  $L_{\lambda,CA}$  [26] of 7.4 and 9.4 dB for the two rails, respectively. On that basis, the rail roughness was classified as average [24].

The measurement train consisted of two locomotives, 7 passenger cars and 6 freight wagons and was composed as listed in Table 5.1. The train passed each measurement site six times in each direction at different speeds, i.e. 1×60 km/h, 3×80 km/h, 2×100 km/h. Subsequent to the pass-by measurements, direct roughness measurements of all 36 freight wagon wheels were conducted.

Parts of this measurement data set were used for the development and validation of the auralization model, which is described in the following section.

## 5.3. Model

### 5.3.1. Point source model

To comply with the fundamentals of auralization [6], in particular the separation of sound generation and propagation, the proposed model follows the source–path–receiver concept and uses distributed point sources.

The train composition as an acoustical source is represented by a series of mov-

Number of wagons	Type	Name	Number of axles/wagon	Wheel diameter [m]
1	locomotive	SBB Re460	4	1.1
3	passenger (D-braked)	SBB EWIV	4	0.92
4	passenger (K-braked)	SBB RIC Bpm	4	0.92
3	flat freight (Ci-braked)	SBB Slmmnps	4	0.92
3	flat freight (Ci-braked)	SBB Kps	2	1.0
1	locomotive	SBB Re420	4	1.23

Table 5.1: Composition of the measurement train of the sonRAIL project. The train consists of wagons with different braking systems, i.e. disc (D) brakes, cast iron (Ci) and composite (K) brake blocks.

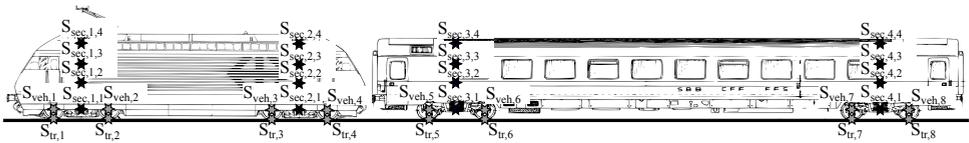


Figure 5.1: Example of equivalent point source locations marked by colored stars along two Swiss rail vehicles of type SBB Re 460 and Bpm RIC.

ing point sources. They all move at the same speed which is equivalent to the traveling speed  $V$  of the train. For most exposure situations, a train has to be viewed as an extended source as its total length is typically larger than or of similar magnitude as the distance to the receiver point. Therefore the point sources have to be horizontally spread across the train composition. To correctly model ground reflections and shielding, the source height is of importance. As railway noise consists of noise sources of different heights, the point sources in the model are thus horizontally and vertically distributed along the train.

Primary point sources, denoted as  $S_{tr}$  and  $S_{veh,i}$  are used to represent rolling and impact noise.  $S_{tr}$  represents the contribution radiated by the track, and  $S_{veh}$  the contribution radiated by the vehicle. For each axle  $i$  two primary point sources,  $S_{tr,i}$  and  $S_{veh,i}$  are located at heights 0 and 0.5 m above rail. Secondary point sources, denoted as  $S_{sec,i,j}$  are introduced for traction noise, aerodynamic noise, aggregates, etc. Secondary point sources are positioned at different heights, i.e. at 0.5, 2, 3, and 4 m above rail according to state-of-the art engineering models [24, 27, 28]. The sonRAIL model [24] describes the sound powers of secondary sources at these predefined heights for different vehicle types. The horizontal distribution of the secondary point sources along the wagons should be defined according to the physical positions of the real sources, i.e. individually for each vehicle type. As a default setting, one stack is located on top of each boogie as shown in Figure 5.1. The point source locations are exemplified in Figure 5.1. However, it can be expected

that in many cases, fewer secondary sources suffice, e.g. at large distances or in unshielded cases. For a wagon with  $N_{\text{ax}}$  axles this yields a total of maximal  $2N_{\text{ax}} + 8$  point sources. Consequently, for a whole train composition, the number of point sources typically lies well above 100.

Note that replacing the emission of the wheel/track system by two stacked equivalent point sources with a prescribed directivity is a simplification of the real system. In fact, the wheel/track system features an extended source, the rail, also radiating at locations away from the point of contact, especially at medium and high frequencies where vibration attenuation along the rail is low. Here, we follow the standard engineering models such as CNOSSOS-EU [29] where it is assumed that, as the wheel moves on the rail, this system is approximated by a single point source.

### 5.3.2. Synthesis approach

Various techniques for digital sound synthesis exist [30]. In the context of sounds from aircraft, wind turbines and road traffic, a combination of additive [31–33] and subtractive synthesis [31, 32], referred to as spectral modeling synthesis [33–35], is commonly used [22, 36–39]. With this synthesis method the emitted sound pressure  $p_e$  at time  $t$  may be generated by

$$p_e(t) = \sum_l a_l(t) \cos\left(\phi_l(t) + 2\pi \int_{-\infty}^t f_l(\tau) d\tau\right) + \sum_j b_j(t) v_j(t) \quad (5.1)$$

where  $a_l$  and  $b_j$  are amplitude modulation functions (with  $a_l \geq 0$  and  $b_j \geq 0$ ),  $\phi_l$  is a phase modulation function (with  $0 \leq \phi_l < 2\pi$ ),  $f_l$  is frequency in Hz,  $v_j$  denotes third-octave band filtered stationary noise, and  $l$  and  $j$  are indices of summation. As implemented in [14, 15], the most obvious approach is to also use spectral modeling synthesis for the emission synthesis of train pass-bys. For secondary sources, this approach seems particularly justified as their sound generation principles resemble other environmental noise sources, e.g. cars or aircraft. However, our initial tests using this synthesizer type (Equation (5.1)) for train pass-bys failed. Two reasons for this failure were impact noise which could not be represented, and the sound characteristics (timbre) of rolling noise which could not be reproduced. However, these subtle temporal and spectral patterns are important for realistic auralizations of railway noise.

Experimental data which illustrates these characteristics is shown in Figure 5.2 and Figure 5.3. The spectrogram in Figure 5.2 contains an excerpt of the sound pressure of the measurement train recorded close to the track (see section 5.2). In this example, a series of typical features of railway noise can be pointed out. Firstly, clear broadband differences between the first part (up to time instant 3 seconds) and the second part can be observed. This is due to different types of wagons, as during the first part, passenger cars (with K-brake blocks), and in the second part, freight wagons (with cast-iron (Ci) brake blocks) are passing by. Secondly, between second 6 and 7, a temporal pattern is present. These vertical lines represent transient sounds which are due to a wheel flat on one of the freight wagon axles. Thirdly, fine spectral patterns consisting of peaks and dips can be

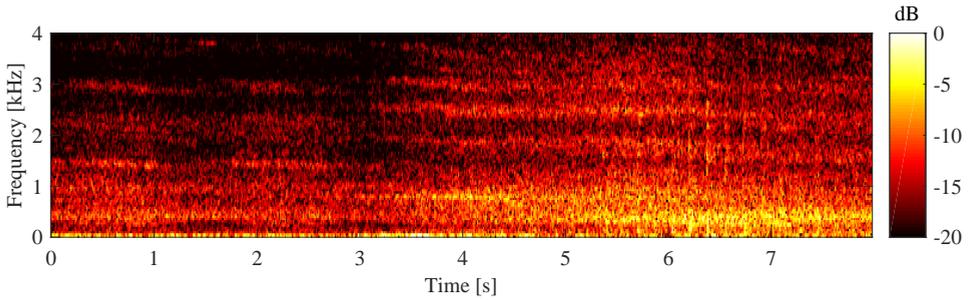


Figure 5.2: Spectrogram of the sound pressure of a train pass-by (normalized to 0 dB). The train speed is 60 km/h and the microphone was at a reference distance of 7.5 m to the track and at height 1.2 m above track. A 1024-point FFT and a Hann window of equal length with 50% window overlap was used.

observed throughout this example. These spectral peaks and dips seem to differ between the two wagon types. For the passenger cars, peaks occur at about 1.5, 2.3 and 3.0 kHz, whereas for the freight wagons peaks occur at somewhat higher values of 1.8, 2.5, 3.2 kHz. Figure 5.3 shows measured narrowband sound pressure spectra of flat freight wagons traveling at different speeds. Identical wagons were measured at the identical location. Flat wagons were chosen for this example as secondary sources are negligible for this vehicle type and rolling noise is dominating. In Figure 5.3 spectral peaks are also clearly apparent. Two different kinds of peaks can be discriminated (see Figure 5.3a, (1) and (2)): (1) Below 1600 Hz, two widely spaced peaks per speed are present. These peaks show an increasing amplitude as a function of the traveling speed as well as a distinct frequency shift. The frequency location of the peaks scales with the speed. (2) Above 1600 Hz, two peaks per speed are present which feature an increasing amplitude with increasing speed. However, their frequency remains constant. Their bandwidth slightly increases with speed which is mainly due to the Doppler frequency shift which leads to a frequency smearing. This indicates that two different physical phenomena are involved in creating these spectral patterns.

Therefore in the next section we propose a physical approach that includes modeling of the mechanical excitation of the wheel/rail system and the vibrational behavior of the system separately. A physical approach on that basis has the advantage that it allows the extrapolation to situations for which no synthesizer parameters have been measured. This approach has already proven to be successful for the synthesis of transient sounds, specifically percussion and bells [40–42] or foot-step sounds [43]. For railway noise, transient sounds occur notably in the context of wheel flats [9], insulated rail joints or switches [10].

### 5.3.3. Emission synthesis of rolling and impact noise

Figure 5.4 illustrates the physical conception behind the proposed emission synthesis model for rolling and impact noise. The sound emission is subdivided into three physical processes that are the mechanical excitation, the vibration and the radiation of airborne sound, which propagates through the environment. The mechanical

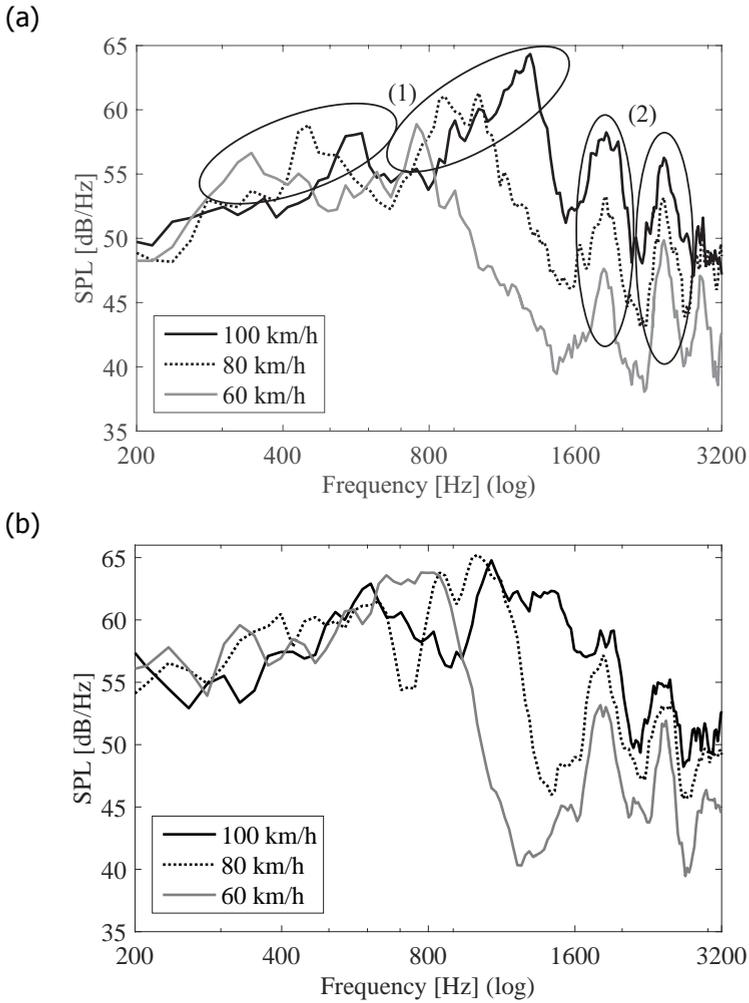


Figure 5.3: Narrowband sound pressure spectra of flat freight wagons (type SBB Slmmpns) at different traveling speeds at a distance of 7.5 m to the track. Panel (a) shows measurement data where some peaks (1) shift their location on the frequency axis as a function of the speed, whereas the frequency locations of other peaks (2) are independent of speed. Panel (b) shows simulated data based on the proposed approach.

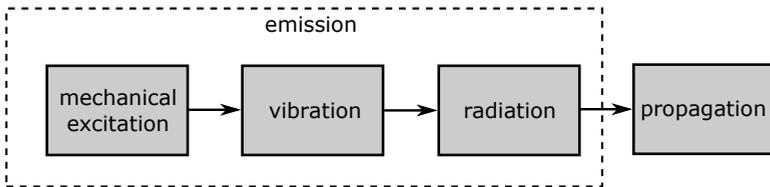


Figure 5.4: Physical emission model of railway rolling and impact noise.

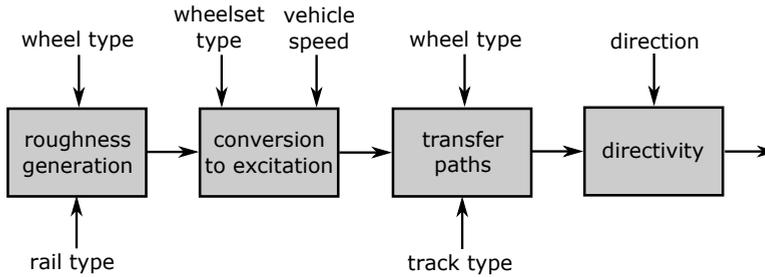


Figure 5.5: Principle of the emission synthesis model for railway rolling and impact noise. Input variables that are related to the vehicle are drawn on top whereas input variables that are related to the track structure are drawn on the bottom.

excitation is driven by the microstructure of the wheel and the rail running surfaces. The same approach is also reflected in theoretical models such as TWINS [44] and state-of-the-art railway noise engineering models such as Harmonoise [27], IMAGINE [28], sonRAIL [24] or CNOSSOS-EU [29]. As an overview, the principle of the proposed emission synthesis model is depicted in Figure 5.5. The shown input variables are either related to the vehicle or to the track. The basis is (in a statistical sense) an estimate of the surface microstructure (i.e. the roughness) of the rail and the wheels. These roughness signals are then combined and processed to obtain the mechanical excitation of the wheel/track structure. Using the vehicle's travelling speed  $V$ , a transformation from the spatial into the time domain is performed. Next, the modal behavior of the structure and the radiation are simulated. These effects are characterized by transfer paths which are applied in the time domain using digital filters.

Figure 5.6 shows a more detailed signal flow of the proposed emission synthesizer for a pass-by of a single axle  $i$ . The upper blocks represent generator modules which produce spatial signals, i.e. depending on the location along the rail axis. These spatial signals are summed up and modified by a contact filter. Subsequently, the resulting spatial signal is converted to the time domain by a vehicle speed dependent resampling. After a differentiation with respect to time to yield vertical speed, the resulting excitation time signal is then fed into two transfer filters modeling the vibration and the radiation efficiency of the system. Their output is amplitude modulated by a directivity function representing the radiation pattern of the source. The resulting signals correspond to the free field sound pressure at a defined reference distance of 1 m, as radiated by the track and the vehicle, respectively. Accordingly these signals are attributed to point sources  $S_{\text{tr},i}$  and  $S_{\text{veh},i}$ .

The synthesizer in Figure 5.6 thus delivers two signals per axle. These are two digital audio signals  $p_{\text{e, tr},i}[k]$  and  $p_{\text{e, veh},i}[k]$  for each axle  $i$  with  $k$  as sample index. Generally an audio sampling rate of  $f_s = 48$  kHz is used. The following sections show in detail how these signal are generated.

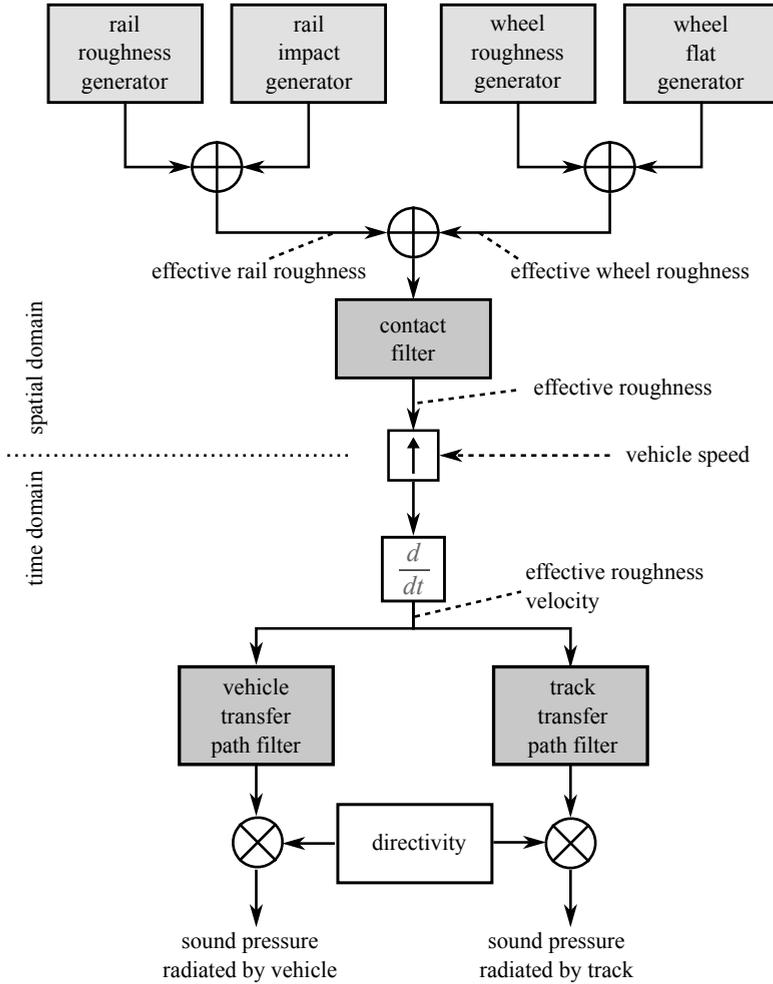


Figure 5.6: Signal flow of the auralization model for rolling and impact noise source signals of a single wheel pass-by.

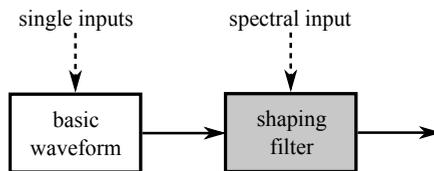


Figure 5.7: Signal flow of the generator modules (subtractive synthesis).

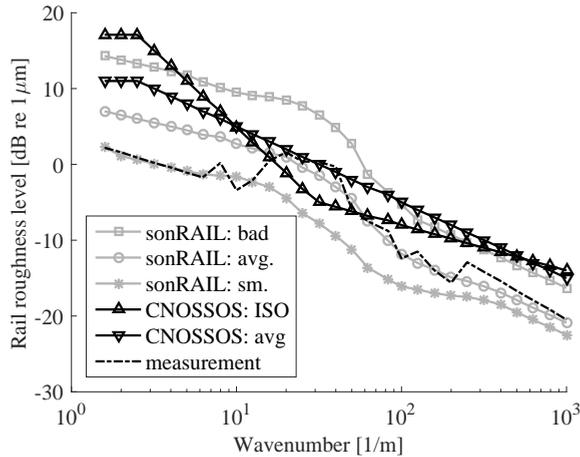


Figure 5.8: Rail roughness level data as a function of the wavenumber in 1/3 octave bands. Three curves for different rail condition classes from the sonRAIL model [24], the two curves from CNOSSOS [29] (denoted as based on ISO 3095:2013 [25] and to represent an average network, respectively) and a single measurement are shown. Note that at low and high wavenumbers most values appear to be (linearly) extrapolated.

### Roughness generators

The four generator modules in Figure 5.6 produce roughness signals as a function of the spatial coordinate  $X$ . In order to do this, they use either measured or predicted roughness spectra as input. Measurement data indicate a strong wavenumber dependency of roughness signals with substantial low wavenumber content (see e.g. Figs. 5.8 and 5.10). These wavenumber dependencies are realized by applying filters to broadband signals. The operating principle of the generators is depicted in Figure 5.7. It is based on the subtractive synthesis technique but differs for each of the four generator modules in terms of input type and basic waveform.

The basis to synthesize rail and wheel roughness profiles is Gaussian white noise. These signals are processed using digital FIR filters. In doing so, the initial constant spectral content is shaped. The corresponding filters are designed in the frequency domain based on spectral input data. In contrast, for impacts and wheel flats, discrete impulses are utilized as basic waveforms. These impulses are spatially shifted to their respective location and filtered in order to bring them into the desired spectral shape. In the following, the details on the computation of the roughness profiles (rail, wheel, impact and flats) are given.

### Rail roughness generator

The rail roughness generator uses a rail roughness spectrum  $L_{r, \text{tr}}$  given in 1/3 octave bands as input. This data may be obtained by measurements on a rail or from a railway noise calculation model. It is generally assumed that roughness data is available in 1/3 octave bands for wavelengths from 63 to 0.1 cm, such as depicted in Figure 5.8. To be able to reproduce this frequency range by digital signals (Nyquist-Shannon sampling theorem), a resolution of  $\Delta X = 0.45$  mm is used for the spatial

coordinate  $X$ , which corresponds to a spatial sampling rate of  $\xi_s = 1/\Delta X \approx 2200$  samples per meter. Based on that, the rail roughness is modeled by a white Gaussian noise signal  $w$  (basic waveform) that is spectrally shaped using a digital FIR filter (shaping filter) with impulse response  $h_{\text{track}}$ . The rail roughness is generated by

$$r_{\text{track}}[x] = (w * h_{\text{track}})[x] \equiv \sum_{m=0}^{N-1} w[x-m] h_{\text{track}}[m] \quad (5.2)$$

where  $*$  denotes the linear discrete convolution,  $x$  is the sample index and  $N$  is the number of filter taps of the filter  $h_{\text{track}}$  with order  $N-1$ . Equation (5.2) is efficiently evaluated by the overlap-add method which uses the fast Fourier transform (FFT). The noise signal  $w$  is scaled to have unit power per wavenumber according to

$$w[x] = \sqrt{\frac{\xi_s}{2}} \Psi(x) \quad (5.3)$$

with the spatial sampling rate  $\xi_s$  and where the random variable  $\Psi$  has a normal distribution with zero mean and variance 1 ( $\Psi \sim \mathcal{N}(0, 1)$ ). For the filter design, i.e. the calculation of the filter coefficients  $h_{\text{track}}[m]$ , the frequency sampling method is adopted. This method uses the inverse discrete Fourier transform (IDFT), or the inverse FFT (IFFT), respectively. Prior to this transformation some inter- and extrapolation steps are necessary:

1. From the rail roughness levels  $L_{r,\text{tr},j}$  given per 1/3 octave band  $j$ , the signal power per unit spatial frequency  $\xi$  is estimated by

$$L_{r,\text{tr}}[\xi = \xi_{c,j}] = L_{r,\text{tr},j} - 10 \log_{10}(B_j) \quad (5.4)$$

where  $\xi_{c,j}$  and  $B_j$  denote the center frequency and the bandwidth of the 1/3 octave band filter  $j$ .

2.  $L_{r,\text{tr}}[\xi]$  is linearly extrapolated to the spatial frequencies  $\xi = 0$  and  $\xi_s/2$ .
3.  $L_{r,\text{tr}}[\xi]$  is interpolated to the FFT bin frequencies  $n$  using piecewise cubic Hermite interpolation to obtain a scaled filter magnitude response  $A[n]$  in dB.
4. From  $A[n]$  a linear phase FIR filter is calculated using the IFFT. A basic version of the filter is obtained by

$$h_{\text{basic}}[m] = \text{IFFT}(a_0 10^{A[n]/20}) \quad (5.5)$$

with variable  $a_0$  being equal to the reference roughness  $r_0$ , i.e.  $a_0 = r_0 = 10^{-6}$  m.

5. As a final step,  $h_{\text{basic}}$  is symmetrically truncated around  $m = 0$  to a total of  $N$  filter taps, shifted by  $N/2$  samples to make the filter causal and multiplied by a  $N$ -point Tukey window with  $\alpha = 0.2$ . The resulting signal corresponds to the filter coefficients of the desired FIR filter.

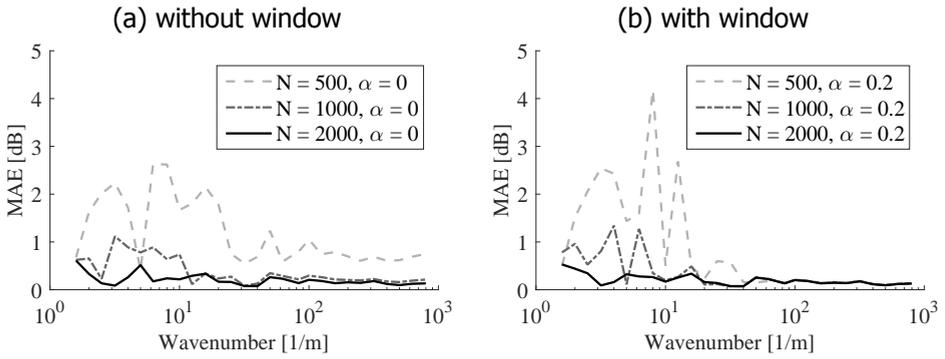


Figure 5.9: Mean absolute errors (MAE) of rail roughness profiles generated with the described procedure for different parameter settings. Signals of length of 200 m using the spectra from Figure 5.8 were generated for different numbers of filter taps  $N$  without (a) and with (b) the Tukey window function. These signals were analyzed using a 1/3 octave band filterbank and subsequently compared to their target functions.

5

The above described filter design procedure was numerically evaluated for the six spectra given in Figure 5.8. Rail roughness signals  $r_{\text{track}}$  of length 200 m were calculated by Equation (5.2) and analyzed in 1/3 octave bands. Figure 5.9 shows the mean absolute errors (MAE) of a parametric study, with

$$\text{MAE} = \frac{1}{6} \sum_{l=1}^6 |L_{r,j,l,\text{target}} - L_{r,j,l,\text{model}}|. \quad (5.6)$$

where the roughness levels denoted by ‘target’ correspond to the data shown in Figure 5.8. It revealed that 2000 filter taps are needed to keep the errors low also at low wavenumbers, meaning that a filter length of about 1 m is required. Further, it was found that filter errors are larger if measurement data is used as input instead of data from current prediction models. This is due to the uneven, jagged shape of a single measurement compared to the smooth shape of model curves (see Figure 5.8) with the latter being easier to model by a filter.

### Wheel roughness generator

Next, the wheel roughness generator produces a periodic signal with the period being equal to the wheel perimeter  $P_i$  in m (i.e., of the running surface) for wheel  $i$ . Along with the wheel perimeter, a wheel roughness spectrum  $L_{r,\text{veh},i,j}$  is needed as input. Figure 5.10 shows wheel roughness data in 1/3 octave bands. The wheel roughness is modeled by a sequence  $s$  of finite length white noise snippets (basic waveform) that is spectrally shaped using a digital FIR filter (shaping filter) with impulse response  $h_{\text{veh},i}$ . The wheel roughness  $r_{\text{veh},i}$  of wheel  $i$  is generated by

$$r_{\text{veh},i}[x] = (s_i * h_{\text{veh},i})[x]. \quad (5.7)$$

The basic waveform  $s_i$  in Equation (5.7) is scaled to have unit power per wavenumber and has a signal period of

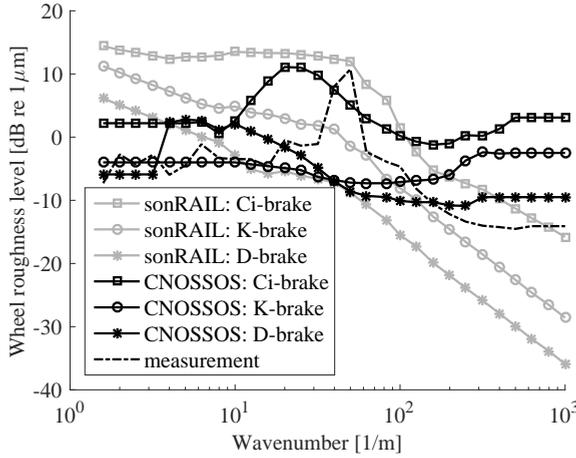


Figure 5.10: Wheel roughness level data as a function of the wavenumber in 1/3 octave bands. Three curves for different braking systems (Cast iron (Ci), composite (K), disc (D)) from the sonRAIL model [24] and from CNOSSOS [29] are shown together with measurement data of a single wheel.

$$\Pi_i = \lfloor P_i \xi_s \rfloor \quad (5.8)$$

where  $\lfloor \cdot \rfloor$  denotes the rounding function to the nearest integer:

$$s_i[x] = \sqrt{\frac{\xi_s}{2}} \Psi_i(x) \text{ for } 0 \leq x < \Pi_i \quad (5.9)$$

$$s_i[x + \Pi_i] = s_i[x] \quad (5.10)$$

where the random variables  $\Psi_i$  have a normal distribution with zero mean and variance 1. The filters  $h_{veh,i}$  are designed based on wheel roughness levels  $L_{r,veh,i,j}$  analogously as for the rail roughness. However, as a relatively short independent signal of  $\Pi$  samples per wheel is generated, larger errors for the roughness of a single wheel as compared to the rail are expected due to the statistical nature of the basic waveform. A typical wheel diameter of 0.92 m results in a perimeter  $P$  of about 2.9 m and a signal period of  $\Pi \approx 6400$  samples. Whereas for rolling highway vehicles, the wheel perimeter is only approx.  $P = 1.1$  m and  $\Pi \approx 2500$  samples. Figure 5.11 shows simulation results where the wheel roughness generation is evaluated for these two wheel sizes. It can be observed that larger errors occur at low wavenumbers and particularly for the small wheel. In these cases, the model underestimates the required roughness levels. Furthermore, the difference between mean and mean absolute errors indicates that a single wheel may exhibit large errors at some frequencies, while the average error over many wheels approaches 0 dB. Thus, for the auralization of a full train consisting of many wheels, this approach seems justified.

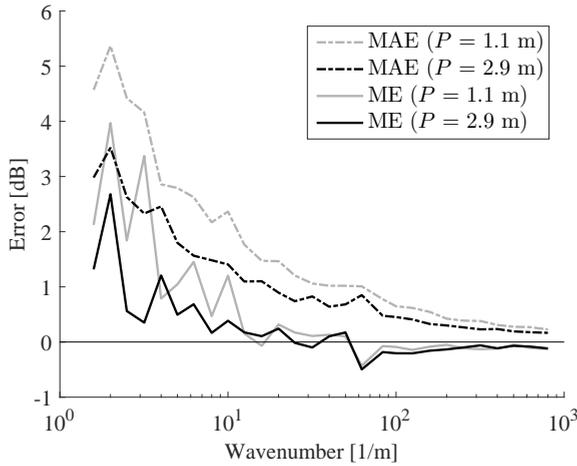


Figure 5.11: Mean (ME) and mean absolute (MAE) errors of wheel roughness profiles generated with the described procedure for two wheel perimeters  $P_i$ . Signals of length 200 m using the spectra from Figure 5.10 were created. For each spectrum and wheel size, twenty individual wheels were synthesized, analyzed and compared to the target spectrum.

5

### Rail impact generator

The rail impact generator uses a spatially shifted impulse as basic waveform and a shaping filter  $h_{\text{imp}}$ . The equivalent impact roughness  $r_{\text{imp}}$  is calculated by

$$r_{\text{imp}}[x] = \begin{cases} \frac{\xi_s}{\sqrt{2}} \sum_I h_{\text{imp}}[x - \check{x}_I + \frac{N}{2}], & \text{if } 0 \leq x < N \\ 0, & \text{otherwise} \end{cases} \quad (5.11)$$

where  $\check{x}_I$  denotes the sample index of the impact location  $I$  on the rail. The scaling factor  $\xi_s/\sqrt{2}$  in Equation (5.11) assures that each Kronecker delta function (basic waveform) has unit energy per wavenumber.  $h_{\text{imp}}$  is designed based on an equivalent roughness spectrum  $L_{r,\text{imp},j}$  for a reference segment of 1 m length given in 1/3 octave bands (see Figure 5.12). However, in the literature, different definitions for the equivalent roughness of impacts exist. For example, CNOSSOS-EU [29] contains an equivalent roughness spectrum for impact noise, which corresponds to the data set from the IMAGINE model [28]. This spectrum is added to the total roughness, i.e. without applying a contact filter. However, the presented auralization model uses an impact roughness spectrum which is added before application of the contact filter and thus requires a compensation. Another example is the sonRAIL emission model that contains impact roughness spectra for wooden and concrete sleepers. The model and thus the data is defined in such a way that the impact roughness only affects the sound power radiated by the track, and not the component radiated by the vehicle. This is also in contrast to the model structure shown in Figure 5.6. A correction spectrum can be estimated as the level difference between the track, and the total (energetic sum of track and vehicle contribution) transfer function, both transformed to wavelengths. Thus, both data bases have to be

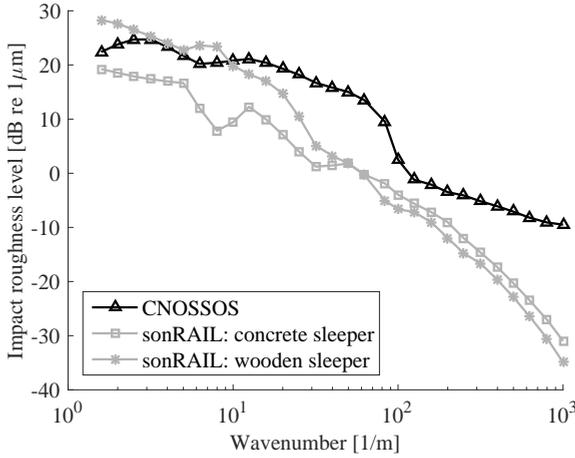


Figure 5.12: Equivalent impact roughness levels as a function of the wavenumber in 1/3 octave bands. The curves show modified data based on the sonRAIL model [24] and CNOSSOS [29] for use in the proposed auralization model.

adapted in order to be used in the described auralization model. Figure 5.12 shows modified data based on the values from CNOSSOS-EU and sonRAIL. The spectra feature a high low wavenumber content with remarkably larger values than the rail roughness spectra from Figure 5.8. This data may be used to design  $h_{\text{imp}}$  with the above described five-step filter design procedure starting with Equation (5.4). This results in symmetrically shaped impulses. Calculations for these spectra showed that also a filter length of 2000 taps is needed in order to keep the filter errors low. Additionally, to attenuate the occurring DC component,  $h_{\text{imp}}$  is high-pass filtered with a cutoff frequency of  $0.8 \text{ m}^{-1}$ , which is well below the lowest 1/3 octave band.

### Wheel flat generator

The roughness corresponding to a wheel flat is modeled using a spectrally shaped impulse train, i.e. the wheel flat generator uses an impulse train with a period equal to the wheel perimeter as basic waveform and a shaping filter  $h_{\text{flat}}$ . The equivalent wheel flat roughness  $r_{\text{flat},i}$  of the wheel  $i$  is calculated by

$$r_{\text{flat},i}[x] = \sqrt{\frac{\Pi_i \xi_s}{2}} \sum_{l=-\infty}^{\infty} h_{\text{flat},i}[x + l\Pi_i + x_{0,i}] \quad (5.12)$$

where  $\Pi_i$  is the signal period from Equation (5.8),  $x_{0,i}$  is a random offset and the scaling factor assures unit power per wavenumber of the basic waveform. The impulse response  $h_{\text{flat},i}$  is calculated based on an equivalent roughness spectrum for the wheel flat  $L_{r,\text{flat},j}$  given in 1/3 octave bands. Here,  $L_{r,\text{flat}}$  is defined analogously as  $L_{r,\text{veh}}$ , i.e. as a signal power, but by only considering the isolated wheel flat and not the remaining roughness of the wheel, which is already represented by  $r_{\text{veh}}$ . Figure 5.13 shows measurement data of two wheels with a wheel flat. The two

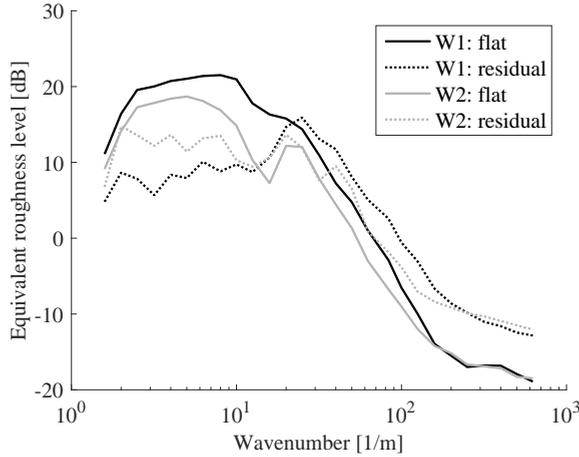


Figure 5.13: Measured equivalent roughness levels of the wheel flat contribution and the residual roughness of two wheels (W1 and W2 from the same wheelset, Ci-braked, 0.92 m diameter).

5

wheels belong to the same wheelset and are responsible for the transient sounds shown in Figure 5.2. Equivalent roughness spectra  $L_{r,\text{flat}}$  and  $L_{r,\text{veh}}$  as used by the presented auralization model are labeled as “flat” and “residual”, respectively, in Figure 5.13. They were derived by spatially windowing roughness raw data before applying a 1/3 octave filterbank. In this data set, the wheel flat dominates the total wheel roughness at low wavenumbers only. This means that the total equivalent roughness (as given in the sonRAIL model) shall not be totally attributed to the wheel flat. Thus, this justifies the chosen modeling approach, where, for a wheel with a wheel flat, the remaining wheel roughness is also considered.

#### From roughness to excitation

The total roughness  $r_{\text{tot},i}$  is modeled by summation of the four signals obtained from Equation (5.2), (5.7), (5.11) and (5.12)

$$r_{\text{tot},i}[x] = r_{\text{track}}[x] + r_{\text{imp}}[x] + r_{\text{veh},i}[x] + r_{\text{flat},i}[x] \quad (5.13)$$

from which the effective roughness  $r_{\text{eff}}$  is obtained by applying a contact filter with impulse response  $h_{\text{contact}}$ . The contact filter models the effect of the finite contact zone of wheel and rail. The contact patch leads to an averaging and thus smoothing of the roughness. Short wavelength variations (as compared to the contact patch size) are attenuated, whereas wavelengths that are much larger than the contact zone remain unaffected. This yields a low-pass behavior in the spatial domain. This effect depends on the wheel diameter and the axle load. The output of this filter corresponds to the effective roughness

$$r_{\text{eff},i}[x] = (r_{\text{tot},i} * h_{\text{contact},i})[x] \quad (5.14)$$

$h_{\text{contact},i}$  is designed for each wheel  $i$  based on a contact filter curve described by a magnitude response in 1/3 octave wavelength bands. This type of data is typically

denoted as an  $A_3$  contact filter [24, 28, 29]. Here,  $A_{3,i,j}$  is the  $A_3$  contact filter of wheel  $i$  for the 1/3 octave band  $j$ . A similar procedure as the above described five-step filter design procedure starting with Equation (5.4) is used for the filter design. However, step 1 is replaced by  $A_{3,i}[\xi = \xi_{c,j}] = A_{3,i,j}$  because  $A_3$  already describes an attenuation, and the reference variable is set to  $a_0 = 1$ . At low wavenumbers,  $A_3$  is 0 dB with a cutoff frequency at a wavelength of about 2.5 cm [7, 28]. Thanks to its broadly constant attenuation at low spatial frequencies, significantly less filter taps (<200) are needed for  $h_{\text{contact}}$  as compared to the filters within the above described roughness generators which require 2000 taps.

Figure 5.14 shows an example of the spatial signals which were generated with the described model for an extreme case where an impact as well as a wheel flat are present. The equivalent roughness of a rail, an impact, a wheel and a wheel flat as well as the combined effective roughness are illustrated. From top to bottom, the first panel shows the random structure of the rail roughness which was generated based on the measurement data from Figure 5.8. At location 7 m an impact is present having a much higher magnitude than the rail roughness. In the second panel, the wheel diameter of 0.92 m leads to signal periods of 2.9 m which is also partially reflected in the effective roughness (third panel). The signals were generated based on the measurement data of the wheel W1 from Figure 5.13. The peaks of the wheel flat are distinctly above the wheel roughness although the underlying equivalent levels at high wavenumbers are lower. In the third panel again, the equidistant impulses due to the wheel flat and the single impulse due to the impact at 7 m as well as the smoothing effect of the contact filter can be observed.

Effective roughness spectra are commonly transformed between spatial and frequency domain using the relation  $\lambda = V/f$  with the wavelength  $\lambda$  in m, vehicle speed  $V$  in m/s and frequency  $f$  in Hz [24, 28]. This translates into the relation  $t = X/V$  for time  $t$ . Applying this transformation to the spatial signal  $r_{\text{eff}}[x]$  yields the deflection time signal

$$\zeta_i[k] = r_{\text{eff},i} [x = V(k/f_s - t_{0,i})\xi_s] \quad (5.15)$$

where  $k$  is the time sampling index,  $f_s$  is the audio sampling rate,  $\xi_s$  denotes the spatial sampling rate and  $t_{0,i}$  is the time instant when axle  $i$  passes the location  $x = 0$ . Equation (5.15) describes a scaling of the argument of the effective roughness by the vehicle speed  $V$ . This leads to a temporal compression of the signal by  $V$  and associated therewith a stretching of its frequency content by  $V$ . In the digital domain, this transformation is known as resampling. In Equation (5.15), it cannot be assumed that  $x$  are integers. Therefore the resampling process needs some interpolation. If the vehicle speed is constant, this process is known as synchronous resampling or sampling rate conversion. From Equation (5.15) it can be seen that the conversion ratio  $C$  depends on  $V$ ,  $\xi_s$  and  $f_s$  and is given by

$$C = \frac{f_s}{V\xi_s} \quad (5.16)$$

In our case, the conversion rate  $C$  takes values between 1.3 and 0.55 for speeds

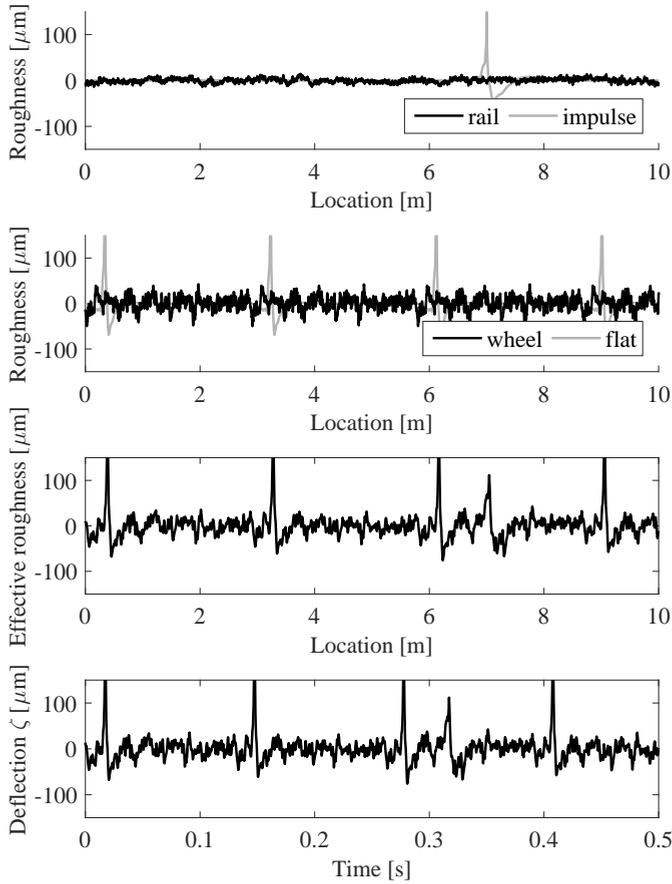


Figure 5.14: Simulated equivalent roughness of a rail  $r_{\text{track}}$ , an impact  $r_{\text{imp}}$ , a wheel  $r_{\text{veh}}$ , a wheel flat  $r_{\text{flat}}$  and the derived effective roughness  $r_{\text{eff}}$  calculated by the presented model. For the deflection signal a vehicle speed of  $V = 80$  km/h was assumed.

between 60 and 140 km/h. At a speed of about 78 km/h  $C = 1$ , which means that no resampling operation is needed. A computationally efficient method to solve Equation (5.15) is rational (or fractional) resampling which involves up- and downsampling operations by integers  $L$  and  $M$ , respectively. This however requires  $C$  to be a fixed, rational number. Therefore in an implementation it is convenient to approximate  $C$  by e.g.

$$C \approx C' = \frac{L}{M} = \frac{\lfloor MC \rfloor}{M} \quad (5.17)$$

where  $\lfloor \cdot \rfloor$  denotes the rounding function to the nearest integer. In choosing the integer  $M$ , a compromise between computational load, and temporal and spectral errors has to be found. The approximation in Equation (5.17) leads to a maximal relative frequency error

$$\eta_{\max} = \frac{1}{2MC} \quad (5.18)$$

Hence, the error is critical at high traveling speeds. For 140 km/h an acceptable value of 2% for the maximal frequency error can be achieved by  $M = 44$ . The lowest panel in Figure 5.14 shows the deflection signal  $\zeta$  which was calculated by Equation (5.15) from the effective roughness  $r_{\text{eff}}$  shown in the panel above for  $t_0 = 0$  s and  $V = 80$  km/h. For the chosen axis limits of 10 m and 0.5 s, respectively, a compression of the deflection signal with respect to the abscissa happens. Theoretically, the 10 m point should be warped to time instant 0.5 s for a speed of 72 km/h.

The subsequent differentiation of  $\zeta_i$  with respect to time  $t$  yields the effective roughness velocity time signal

$$v_i(t) = \frac{d\zeta_i(t)}{dt} \quad (5.19)$$

This signal is the basis for the excitation of the dynamic wheel/rail system. Together with the frequency dependent mobilities of the track, wheel and the contact zone, the contact forces can be derived [7]. However, the derivation of the contact forces requires very detailed knowledge of the complex dynamic system and is thus very ambitious. Therefore in the proposed auralization model, the contact forces are not explicitly identified but rather included in an integral way. The differentiation in Equation (5.19) corresponds to a multiplication with  $j\omega$  in the frequency domain. This operation may be approximated using a FIR filter, where the filter order steers the quality at high frequencies. It is therefore convenient to not implement Equation (5.19) explicitly, but rather implicitly by integrating it into the subsequent transfer path filters.

### Transfer path filters

The two subsequently applied transfer path filters implicitly describe a series of effects: They incorporate the transformations from the velocity signal  $v_i$  to the contact forces which excite the wheel/rail structure. Further, they simulate the

vibrational behavior of the wheel and track and their radiation efficiencies. The outputs correspond to the radiated sound pressure at a reference distance. In today's engineering models, these effects are often summarized by two transfer functions, one for the vehicle and one for the track [24, 28, 29]. Following this approach, we formulate the radiated sound pressure in the axial direction,  $p_{ax,i}$  at a reference distance of 1 m separately for the vehicle and the track using two filter impulse responses  $h_{struc}$  as

$$p_{ax,veh,i}[k] = \frac{p_0}{r_0} (v_i * h_{struc,veh,i})[k] \quad (5.20)$$

$$p_{ax,tr,i}[k] = \frac{p_0}{r_0} (v_i * h_{struc,tr})[k] \quad (5.21)$$

with the reference values for roughness level,  $r_0 = 1 \mu\text{m}$ , and sound pressure,  $p_0 = 20 \mu\text{Pa}$ . Magnitude responses of the transfer filters are typically given in 1/3 octave bands [24, 28, 29]. These transfer functions describe the combinations of Equation (5.19) and (5.20) or (5.21), respectively. However, in the literature, different definitions are used, which e.g. may be formulated for a differing reference distance or for sound power instead of sound pressure. This can be compensated for by either adding a constant to the level data or by using an additional gain correction of the resulting signals,  $p_{ax}$ . The published spectra feature a smooth high-pass behavior as they also include the above stated differentiation (Equation (5.19)) which corresponds to a  $\omega$ -proportionality in the frequency domain.

In a first attempt, the transfer path filters were designed with linear phase and a smooth magnitude response based on 1/3 octave band data. However, listening tests revealed that different filter types with a more complex frequency response are needed. The 1/3 octave band frequency resolution is justified for noise prediction but seems to be too rough for realistic auralizations for which the timbre plays an important role. This is a major difference with previous attempts of railway noise auralization, e.g. VAMPPASS.

Similar to Figure 5.3, Figure 5.15 shows measured narrowband spectra of sound pressure signals recorded close to the passing of axles of the same type. Between 1 and 8 kHz, the fluctuating frequency dependency of rolling noise can be clearly observed. Figure 5.15 further illustrates that the distinct peaks do not depend on the vehicle speed. Thus it appears that they are not linked to the spatial domain, i.e. roughness. For higher speeds, the peaks are still present but somewhat less pronounced due to the Doppler effect. We conclude that the narrowband level variations as shown in Figure 5.15 are most probably attributed to the modal behavior of the structure. Consequently, within the model structure depicted in Figure 5.6, this effect has to be incorporated in the transfer path filters.

The separation into excitation and structural vibration is typical for physical modeling synthesis. To describe the vibration, we propose to use the modal synthesis technique where a resonating structure is described in terms of its modes [41, 45]. The modal resonators are modeled by second order oscillators (also known as damped harmonic oscillators) for an underdamped system. These mechanical os-

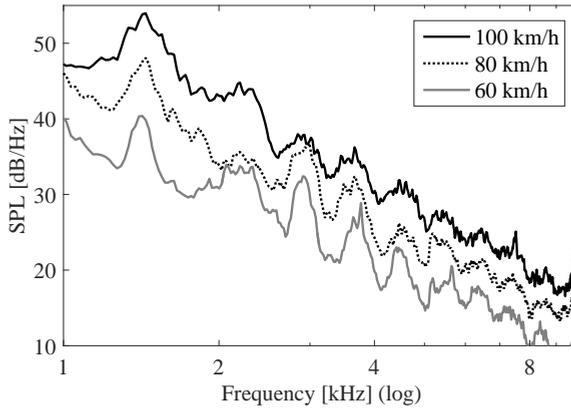


Figure 5.15: Measured sound exposure level of 12 axes with wheel diameter 0.92 m of a SBB Bpm RIC vehicle passing by at different speeds (see legend).

cillators are driven by an external force. The differential equation of the driven harmonic oscillator is given by

$$\ddot{z}(t) + 2\alpha\dot{z}(t) + \omega_0^2 z(t) = \frac{y(t)}{m} \quad (5.22)$$

where the quantities  $\ddot{z}$  and  $\dot{z}$  are the second and first derivative of  $z(t)$  with respect to time  $t$ .  $z$  is the displacement from equilibrium of a point mass of mass  $m$ . Other variables are the decay rate  $\alpha$ , the 'undamped' natural angular frequency  $\omega_0$ , the external (input) force  $y$  applied to the mass, and the condition for underdamping that the damping ratio  $\zeta \equiv \alpha/\omega_0 < 1$ . By applying the Laplace transform  $\mathcal{L}\{.\}$  to Equation (5.22), the transfer function  $H(s)$  between the input force,  $y$ , and the output velocity,  $\dot{z}$ , is derived:

$$H(s) = \frac{\mathcal{L}\{\dot{z}(t)\}}{\mathcal{L}\{y(t)\}} = \frac{sZ(s)}{Y(s)} = \frac{A's}{s^2 + 2\alpha s + \omega_0^2}. \quad (5.23)$$

with the complex variable  $s$  and an amplitude  $A' = 1/m$ . A multiplication in the  $s$ -domain corresponds to a convolution in the time domain and thus  $\dot{z}(t) = (y * h)(t)$ . The inverse Laplace transform  $\mathcal{L}^{-1}\{.\}$  of  $H(s)$  yields the desired impulse response  $h(t)$  of the resonator [46]

$$h(t) = \mathcal{L}^{-1}\{H(s)\} = Ae^{-\alpha t} \sin\left(\sqrt{1 - \frac{\alpha^2}{\omega_0^2}}\omega_0 t - \arctan\left(\frac{\alpha}{\omega_0}\right)\right)\Theta(t) \quad (5.24)$$

with the amplitude  $A$  and the Heaviside function  $\Theta(t)$ . By assuming  $\alpha \ll \omega_0$  Equation (5.24) can be approximated as

$$h(t) \approx Ae^{-\alpha t} \sin(\omega_0 t)\Theta(t). \quad (5.25)$$

Following this approximation, the impulse response  $h_u(t)$  of each mode  $u$  is written as

$$h_u(t) = A_u e^{-\alpha_u t} \sin(2\pi f_{0,u} t) \Theta(t) \quad (5.26)$$

with the amplitude  $A_u$ , the resonance frequency  $f_{0,u}$  and the decay rate  $\alpha_u$  of the respective mode  $u$ . The decay rate  $\alpha_u$  is related to the structural reverberation time  $T_{60,u}$ , which is a quantity that is typically measured in vibroacoustic experiments and is defined as the duration for a 60 dB level drop. Considering the envelope of Equation (5.26), from

$$10 \log_{10} \left( [e^{-\alpha_u 0}]^2 \right) - 10 \log_{10} \left( [e^{-\alpha_u T_{60,u}}]^2 \right) = 60 \quad (5.27)$$

it follows that [47]

$$\alpha_u = \frac{3 \ln(10)}{T_{60,u}}. \quad (5.28)$$

The decay rate  $\alpha_u$  is also directly related to the bandwidth  $B_u$  of the resonance by [47]

$$\alpha_u = \pi B_u. \quad (5.29)$$

$B_u$  is defined as the size of the frequency range over which the power is larger than half the power at the resonance frequency. It is also known as “full width at half maximum”. From the above expressions it can be seen that increasing the decay rate  $\alpha_u$  stronger attenuates the oscillation, decreases the structural reverberation time and widens the resonance peak.

The effect of a mode  $u$  is realized by convolving the input signal with  $h_u(t)$ . This operation is performed by applying a filter with impulse response  $h_u(t)$ . The effect of all modes together is then realized by  $U$  filters that are connected in parallel and by summing their outputs [41, 42]. The total impulse response  $h_{\text{decay}}$  of a transfer path filter is written as

$$h_{\text{decay}}(t) = \sum_{u=1}^U A_u e^{-\alpha_u(t-\tau_u)} \sin(2\pi f_{0,u}(t-\tau_u)) \Theta(t-\tau_u) \quad (5.30)$$

with the time delays  $\tau_u$ . The main challenge is to find an appropriate setting of the parameters  $f_{0,u}$ ,  $\alpha_u$ ,  $A_u$  and  $\tau_u$ . To do so, detailed information about the dynamic behavior is required. The required parameters cannot be determined separately for a single wheel or a track, as they differ for the combined coupled dynamic system. The interaction adds further resonances and damping [7]. It is also known that the rolling of the wheel affects these parameters. For instance, compared to a wheel at rest a rolling wheel possesses additional and differing resonances [7]. Consequently, as isolated measurements or simulations are not expedient and data from current literature is not sufficient, the estimation of the model parameters in

Equation (5.30) may be taken as solving the inverse problem. We propose to use suitable pass-by measurements to fit the required parameters.

In the freight wagon sound pressure spectra of Figure 5.3, distinct peaks which are related to the vibration are located at 1.8, 2.5 and 2.9 kHz. In the passenger wagon spectra of Figure 5.15, distinct peaks may be recognized at different frequencies, namely at 1.4, 2.2, 2.9, 3.6 and 4.4 kHz. For higher frequencies, it is more difficult to discern clear spectral patterns. The data further shows that the peaks may be more easily localized at a low speed of 60 km/h, as compared to 80 or 100 km/h. At lower speeds, spectral smearing due to the Doppler frequency shift is reduced. However, as at lower speeds, rolling noise is reduced, secondary sources increasingly contaminate the data. The measurement data in Figure 5.15 exhibits five peaks between 1 and 5 kHz, which signifies that only slightly less peaks than the number of 1/3 octave bands occur within this frequency range. As modal density increases with frequency [7], at higher frequencies, a higher number of resonances per 1/3 octave band is expected. This might explain the difficulty to recognize distinct peaks at high frequencies. Furthermore, the spectra of Figure 5.15 are averaged over 12 axles with a certain variation in their resonance frequencies. This variation is also expected to be larger at higher frequencies. Therefore at higher frequencies a modal density of more than one mode per 1/3 octave band and some variation of the resonance frequencies between wheels is needed.

The amplitudes  $A_u$  in Equation (5.30) are set for each resonance individually. They steer the gain of each resonance and may be set based on measured 1/3 octave band transfer functions [24, 28, 29]. According to Equation (5.30), the setting of  $A_u$  thus also depends on the decay rate  $\alpha_u$  of each mode which is related to the structural reverberation time by Equation (5.28). Structural reverberation times  $T_{60}$  of freely suspended wheels may exceed 1 second for certain modes [48]. However, our measurements on rails and the literature suggest that structural reverberation times lie well below 1 s for practical situations where the boundaries and the interaction introduce a significant amount of damping. Field measurements revealed values of 0.07–0.23 s for five resonances of a running wheel in the range of 1.7 to 3.7 kHz by observing bandwidths of velocity levels [49]. Similar information might also be obtained by measuring narrowband sound pressure levels. However, this is delicate due to the Doppler effect and as within such measurements, resonance peaks of multiple wheels temporally overlap. A structural reverberation time of  $T_{60} = 0.1$  s results in a resonance bandwidth of  $B = 3 \ln(10)/(\pi T_{60}) = 22$  Hz. For a speed of 100 km/h, the Doppler effect may shift a frequency component by approx.  $\pm 8\%$ . For a frequency of 2 kHz, this results in a difference of 330 Hz. For a speed of 60 km/h, the Doppler effect may shift a frequency component by  $\pm 5\%$  resulting in a difference of 200 Hz at 2 kHz, which is still much higher than the resonance bandwidth.

Figure 5.16 shows impulse and frequency responses of transfer path filters which were designed on the basis of Equation (5.30). Between 1 and 5 kHz, the five resonance frequencies mentioned above (as extracted from Figure 5.15) were introduced.  $T_{60,u}$  is set to 0.1 s for all modes. The delays  $\tau_u$  are set randomly within their respective period to realize random phase relations between the modes. Fur-

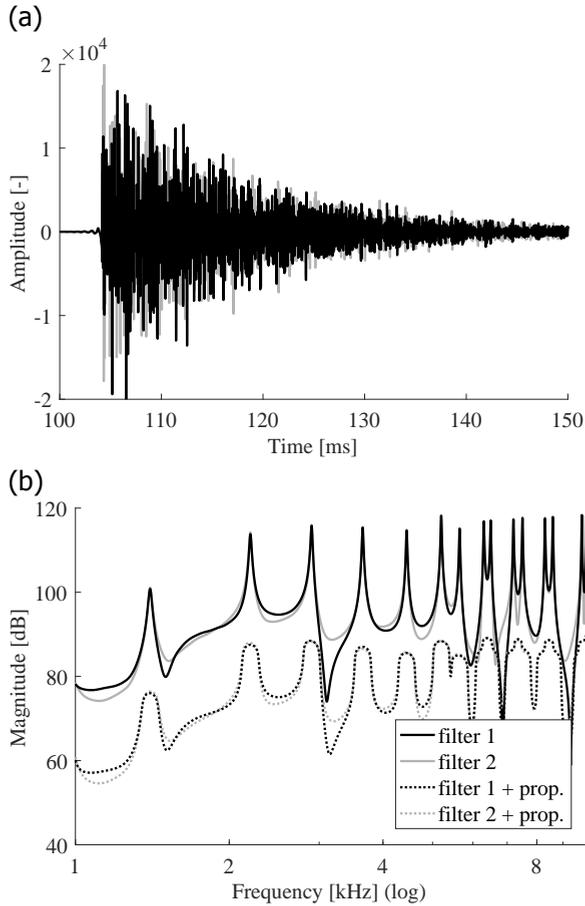


Figure 5.16: Two modeled transfer path filters (black and gray) described by their impulse response (a) and their frequency response (b). The two filters were created for the same resonance frequencies and damping, but differ with respect to the randomly set delays  $\tau_{it}$ . The dashed lines illustrate the spectral smearing effect due to sound propagation at a distance of 7.5 m and for speed 60 km/h.

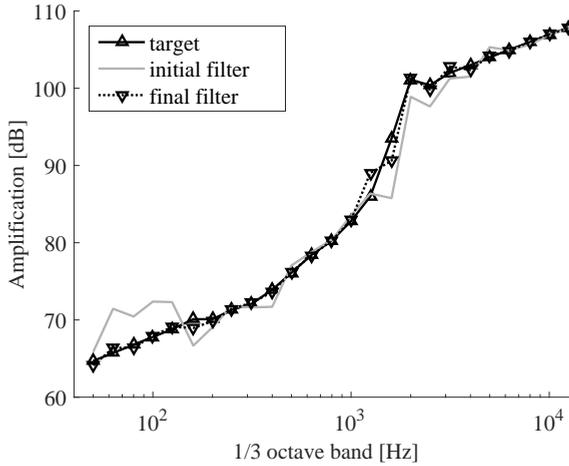


Figure 5.17: A target spectrum for a transfer path from the sonRAIL emission model in comparison with the magnitude responses of a designed initial filter  $h_{\text{init}}$  and a final filter  $h_{\text{TP}}$ .

ther, 1/3 octave data from the sonRAIL model is used. As explained above, it can be clearly seen that the resonance peaks in Figure 5.16 are much narrower than the measured data in Figure 5.15. Furthermore, some variations between the two simulated filters can be observed in the impulse responses as well as in the frequency responses, particularly in the regions between the resonances. This is due to the randomly set delays  $\tau$  which leads to varying interference between the resonances. This has also implications on the accuracy of the filters magnitude response as measured in 1/3 octave bands and thus the filter design process.

Figure 5.16 further illustrates the spectral smearing effect of sound propagation. The simulation considers directivity (see next section), geometrical spreading and the Doppler frequency shift. The shown data was calculated for a distance of 7.5 m, 60 km/h speed and 1 s integration time. Figure 5.16 shows that even for such low speeds, the peak bandwidths are strongly increased, as explained above. This increase is due to the Doppler frequency shift, which is large for short distances and high speeds. However, the applied directivity, as defined in the following section, is maximal at the shortest distance, which leads to an attenuation of the frequency shifted energy and thus decreases spectral smearing.

The filter design process starts by creating a decay filter impulse response  $h_{\text{decay}}$  according to Equation (5.30) with constant amplitudes  $A_u$  for all modes within the same 1/3 octave band. The amplitudes are scaled so that the sum over the corresponding resonances leads to unit magnitude response at the respective band. Secondly, the target response given in 1/3 octave bands, such as shown in Figure 5.17, is linearly extrapolated to a frequency range of 50 Hz to 12.5 kHz. Based on that, an initial shaping filter  $h_{\text{shape},0}$  of 5000 samples length is created as described in section 5.3.3 for the contact filter  $A_3$ . An initial transfer path FIR filter  $h_{\text{init}}$  is obtained by convolving these two impulse responses, i.e.  $h_{\text{init}} = h_{\text{shape},0} * h_{\text{decay}}$ . Due to the absence of resonances, interfering resonances and interpolation,  $h_{\text{init}}$

may exhibit errors of several dB, as exemplified in Figure 5.17. Therefore the deviations of  $h_{\text{init}}$  are analyzed in 1/3 octaves and used to redesign the shaping filter  $h_{\text{shape,corr}}$  which is pre-equalized by modifying the target response. The final transfer path filter is obtained as  $h_{\text{TP}} = h_{\text{shape,corr}} * h_{\text{decay}}$ , applying a temporal truncation after  $4 T_{60}$  and a linear fade out of a few milliseconds. Figure 5.17 illustrates that this filter may closely agree with the target spectrum.

### Directivity

The directivity of the rolling and impact noise sources are assumed to be frequency independent, i.e. shielding effects due to the vehicle body or resonance-specific radiation patterns are not considered here. The radiation pattern can thus be understood as a scaling of the emitted sound pressure as a function of the emission angle. This amplitude modulation is performed based on the horizontal directivity function  $\Delta L_{W,\text{dir,hor},i}$  proposed in CNOSSOS-EU [29]. Converted to sound pressure, the modulation function reads

$$\delta_i(t) = c_{\text{off}} \sqrt{0.01 + 0.99 \cos^2(\varphi_i(t))} \quad (5.31)$$

with a calibration constant  $c_{\text{off}}$  and the emission angle  $\varphi_i$  being equal to 0 on the axis of rotation of axle  $i$  and  $\pi/2$  in the traveling direction.  $c_{\text{off}}$  ensures the correct sound power of the point source. If the used transfer function data (see previous section) was determined in accordance with the directivity pattern in Equation (5.31), no correction has to be made and  $c_{\text{off}} = 1$ . However, if the used transfer functions are based on a sound power formulation where the sound power was estimated based on a directivity pattern differing from Equation (5.31),  $c_{\text{off}}$  has to be set accordingly. E.g. using the sonRAIL emission model, which was developed based on a monopole assumption, requires  $c_{\text{off}} = 1.1$ .

#### 5.3.4. Evaluation

As a verification and demonstration of the presented auralization concept, the measured data from panel i) of Figure 5.3 was reproduced using the proposed emission synthesizer structure and propagation filters. Panel ii) of Figure 5.3 shows sound pressure spectra which were calculated from auralized pass-bys at an observer point close to the track. In the simulation, the only varying parameter was the vehicle speed  $V$ . It was set to 60, 80 and 100 km/h, whereas all the other model parameters were kept constant.

In the simulation, the measured rail roughness shown in Figure 5.8 from the measurement site (see section 5.2) was adopted. Five wagons of the type SBB SImmnps, resulting in a total of 20 axles ( $i = 1, \dots, 20$ ), were modeled. The wheel roughness profiles were generated based on the single wheel measurement spectrum from Figure 5.10 and a wheel perimeter of  $P = 2.9$  m. The transfer path spectra were taken from the sonRAIL model [24]. The wheel modes were modeled as described in section 5.3.3 with a supplemental random variation of the resonance frequencies, i.e. slightly differing resonances for all wheels. The relative frequency shift of each mode had zero mean and a standard deviation of 2%. For the track an exponentially increasing modal density as a function of frequency was used.

Regarding sound propagation effects, geometrical spreading, the Doppler effect and air absorption were considered and modeled according to section 4.2.3 [22]. The ground reflection was neglected as in the specific situation, controlled measurements using a loudspeaker have indicated no significant ground effect above 200 Hz [50]. In general, at close distances to the railway track, the simulation of the ground reflection is challenging due to the extended reaction of the ballast [51]. The receiver point was located at a distance of 7.5 m and 1.2 m above the rail.

In panel ii) of Figure 5.3 the behavior of both types of spectral peaks as observed in panel i) are well reproduced by the model, i.e. peaks due to the excitation and peaks due to the vibration. Above 1.6 kHz three peaks are present which feature an increasing amplitude for increasing speed, but at a constant frequency. These peaks are attributed to the resonances included in the transfer path filters of the vehicle as defined in Equation (5.30), i.e. due to the vibration. For the peak at 2.5 kHz, the speed-related increase of the bandwidth which is due to the Doppler shift is also visible.

For the studied situation, at frequencies below 1.2 kHz the transfer path of the track dominates the simulated sound pressure. The wide peak from 650 to 850 Hz for the speed 60 km/h is shifted upwards in frequency for higher speeds. These peaks are due to the distinct peak at a wavelength of 2 cm of the used wheel roughness as shown in Figure 5.10. These peaks are therefore related to the excitation. Compared to the measurement, however, the magnitude of these peaks stays constant whereas the measurements suggest an increase. Below 800 Hz the pattern is less clear. The peak which shifts from 400 to 600 Hz in the measurements in Figure 5.3 is not well reproduced. This peak might be attributed to the roughness patterns of other wheels within the train. In fact, direct wheel roughness measurements have revealed that some wheels of the same waggon type within the measurement train had a dominant peak at a wavelength of 5 cm [50]. Thus, this demonstrates the variation between axles and underlines the importance of detailed input data for the described model.

In appendix A audio examples of full train pass-bys generated with the presented model are provided.

## 5.4. Conclusions

Our study showed that for a realistic auralization of railway noise, temporal and spectral patterns with a high resolution need to be considered. The temporal patterns arise on the one hand from differences between the vehicles and axles. On the other hand, transient sounds are also characteristic for railway noise. They particularly occur in the cases of wheel flats, or impact noise caused by insulated rail joints or switches. Spectral patterns of rolling noise were observed in narrowband spectra which showed distinct peaks which are due to different kinds of physical phenomena. Empirical data suggests that they can be either attributed to the mechanical excitation or the vibration of the dynamic wheel/track system. These spectral patterns evoke the characteristic "metallic" timbre of rolling noise.

In former studies, a method denoted spectral modeling synthesis has proven successful in the context of wind turbine, road traffic and aircraft noise. However,

in this study it was found that this method is not capable to satisfyingly represent the above described characteristics of railway noise. Further, it is insufficient to rely on 1/3 octave band data only. In addition, modal data of the wheel/rail system has to be included to create the required timbre and thus realistic sounds.

Here, a synthesizer structure was proposed that is based on a physical approach in which the noise generation mechanism is modeled in the time domain. It allows for the reproduction of the above described sound characteristics and concurrently represents rolling and impact noise. The physical model further allows for auralizing the effect of varying wheel and rail roughness as well as different vehicle speeds. Thus, different noise mitigation measures at the source and on the propagation path can be simulated without the need of elaborate measurements.

For the auralization of full train pass-bys, in addition to rolling and impact sounds, secondary sources also need to be considered, such as the noise from the traction system, aggregates or aerodynamic noise. Realistic emission sounds of these sources may probably be generated using spectral modeling synthesis. Further, typical audible railway noise characteristics such as squeal and rattling would require a model with more complexity and subtlety. However, in the assessment and comparison of noise mitigation measures, such as a reduction of roughness, the introduction of dampers or noise barriers, relative differences between scenarios are in the focus and therefore for this purpose these characteristics may possibly be neglected.

The presented auralization model is able to synthesize rolling and impact noise of railway noise without audible artifacts at a high audio quality. It offers a high degree of flexibility, however at the cost of numerous input parameters which are difficult to determine. Future efforts will include better estimation methods for the input data as well as perceptual validation of the proposed auralization system through listening tests.

## Acknowledgments

The work presented in this chapter was funded by the Swiss National Science Foundation (SNSF) (# 200021-149618). I also thank the Swiss Federal Office for the Environment (FOEN) for the data relating to sonRAIL.

## References

- [1] R. Pieren, K. Heutschi, J. M. Wunderli, M. Snellen, and D. G. Simons, *Auralization of railway noise: Emission synthesis of rolling and impact noise*, *Applied Acoustics* **127**, 34 (2017).
- [2] World Health Organization (WHO), *Night noise guidelines for Europe*, Tech. Rep. (World Health Organization (WHO), Regional Office for Europe, 2009).
- [3] L. Fritschi, L. Brown, D. Schwela, and S. Kephelopoulous, *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*, Tech. Rep. (World Health Organization (WHO), Regional Office for Europe, 2011).

- [4] M. Basner, U. Müller, and E. M. Elmenhorst, *Single and combined effects of air, road, and rail traffic noise on sleep and recuperation*, *Sleep* **34(1)**, 11 (2011).
- [5] F. Rudzik, L. Thiesse, R. Pieren, J. M. Wunderli, M. Brink, M. Foraster, H. Héritier, I. Eze, C. Corrado, D. Vienneau, N. Probst-Hensch, M. Röösli, and C. Cajochen, *Sleep spindle characteristics and arousability from night-time transportation noise exposure in healthy young and older individuals*, *Sleep*, [zsy077](#) (2018).
- [6] M. Vorländer, *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* (Springer, Berlin, 2008).
- [7] D. Thompson, *Railway Noise and Vibration - Mechanics, Modelling and Means of Control* (Elsevier, Amsterdam, 2009).
- [8] A. M. Remennikov and S. Kaewunruen, *A review of loading conditions for railway track structures due to train and track vertical interaction*, *Structural Control and Health Monitoring* **15**, 207 (2008).
- [9] T. X. Wu and D. J. Thompson, *A hybrid model for the noise generation due to railway wheel flats*, *Journal of Sound and Vibration* **251**, 115 (2002).
- [10] T. X. Wu and D. J. Thompson, *On the impact noise generation due to a wheel passing over rail joints*, *Journal of Sound and Vibration* **267**, 485 (2003).
- [11] D. J. Thompson, *Wheel-rail noise generation, part i: Introduction and interaction model*, *Journal of Sound and Vibration* **161**, 387 (1993).
- [12] M. Klemenz, *Sound synthesis of starting electric railbound vehicles and the influence of consonance on sound quality*, *Acta Acustica United with Acustica* **91**, 779 (2005).
- [13] S. Khan, M. Ulhan, and V. Geol, *Synthesis of pass-by railway noise*, in *Acoustics'08, 155th ASA Meeting, 7th European Conference on Noise Control* (Paris, France, 2008) pp. 1025–1030.
- [14] S. Molla, E. Bongini, P. Gautier, D. Habault, P. Mattéi, and F. Poisson, *Vehicle passby noise prediction and audio synthesis*, in *19th International Congress on Acoustics (ICA 2007)* (Madrid, Spain, 2007).
- [15] E. Bongini, S. Molla, P. Gautier, D. Habault, P. Mattéi, and F. Poisson, *Synthesis of noise of operating vehicles: Development within silence of a tool with listening features*, in *Noise and Vibration Mitigation for Rail Transportation Systems: Proceedings of the 9th International Workshop on Railway Noise, Munich, Germany, 4 - 8 September 2007*, edited by B. Schulte-Werning and et al. (Springer, Berlin Heidelberg, 2008) pp. 320–326.
- [16] E. Bongini and E. Bonnet, *Railway noise sources definition within the scope of pass-by sound synthesis*, in *Euronoise 2009, 8th European Conference on Noise Control* (Edinburgh, Scotland, 2009).

- [17] E. Bongini and H. Marcus, *Railway equivalent noise sources definition: ESM implementation and optimization*, in *Acoustics 2012* (Nantes, France, 2012) pp. 173–178.
- [18] A. Peplow, J. Forssén, P. Lundén, and M. E. Nilsson, *Exterior auralization of traffic noise within the listen project*, in *Forum Acusticum 2011*, European Conference on Acoustics (Aalborg, Denmark, 2011) pp. 665–669.
- [19] F. Meng, F. Wefers, and M. Vorlaender, *Acquisition of exterior multiple sound sources for train auralization based on beamforming*, in *Euronoise 2015*, 10th European Conference on Noise Control (Maastricht, Netherlands, 2015) pp. 1703–1708.
- [20] F. Meng, F. Wefers, and M. Vorländer, *Moving sound source simulation using beamforming and spectral modelling for auralization*, in *DAGA 2016*, 42nd Annual Conference on Acoustics (Aachen, Germany, 2016) pp. 1–4.
- [21] R. Pieren, T. Bütler, and K. Heutschi, *Auralisation of accelerating passenger cars*, in *Euronoise 2015*, 10th European Conference on Noise Control (Maastricht, Netherlands, 2015) pp. 757–762.
- [22] R. Pieren, T. Bütler, and K. Heutschi, *Auralization of accelerating passenger cars using spectral modeling synthesis*, *Applied Sciences* **6**, 5 (2016).
- [23] R. Pieren, J. M. Wunderli, A. Zemp, S. Sohr, and K. Heutschi, *Auralisation of railway noise: A concept for the emission synthesis of rolling and impact noise*, in *INTER-NOISE 2016*, 45th International Congress and Exposition of Noise Control Engineering (Hamburg, Germany, 2016) pp. 3466–3472.
- [24] T. Thron and M. Hecht, *The sonRAIL emission model for railway noise in Switzerland*, *Acta Acustica United with Acustica* **96**, 873 (2010).
- [25] International Organization of Standardization, *ISO 3095: Railway applications – Acoustics – Measurement of noise emitted by railbound vehicles*, (2005).
- [26] A. von Lier, *The measurement, analysis and presentation of wheel and rail roughness*, Tech. Rep. 8 (AEAT report NSTO/97/9571011/005, 1997).
- [27] C. Talotte, P. van der Stap, M. Ringheim, M. Dittrich, X. Zhang, and D. Stiebel, *Railway source models for integration in the new European noise prediction method proposed in Harmonoise*, *Journal of Sound and Vibration* **293**, 975 (2006).
- [28] M. Dittrich, *The imagine source model for railway noise prediction*, *Acta Acustica United with Acustica* **93**, 185 (2007).
- [29] European Commission, *Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council*, *Official Journal of the European Union* **58** (2015).

- [30] G. de Poli, *A tutorial on digital sound synthesis techniques*, *Computer Music Journal* **7**, 8 (1983).
- [31] C. Roads, ed., *The Computer Music Tutorial* (MIT Press, Cambridge, 1996).
- [32] T. D. Rossing, ed., *Springer Handbook of Acoustics* (Springer, New York, 2007).
- [33] J. O. Smith III, ed., *Spectral Audio Signal Processing* (W3K Publishing, Stamford, 2011).
- [34] X. Serra and J. Smith III, *Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition*, *Computer Music Journal* **14**, 12 (1990).
- [35] X. Serra, *Musical sound modeling with sinusoids plus noise*, in *Musical Signal Processing*, edited by C. Roads, S. Pope, A. Picialli, and G. De Poli (Swets & Zeitlinger, Lisse, The Netherlands, 1997).
- [36] M. Arntzen and D. G. Simons, *Modeling and synthesis of aircraft flyover noise*, *Applied Acoustics* **84**, 99 (2014).
- [37] R. Pieren, K. Heutschi, M. Müller, M. Manyoky, and K. Eggenschwiler, *Auralization of wind turbine noise: Emission synthesis*, *Acta Acustica United with Acustica* **100**, 25 (2014).
- [38] S. A. Rizzi, A. R. Aumann, L. Lopes, and C. L. Burley, *Auralization of hybrid wing-body aircraft flyover noise from system noise predictions*, *Journal of Aircraft* **51**, 1914 (2014).
- [39] A. Sahai, F. Wefers, P. Pick, E. Stumpf, M. Vorländer, and T. Kuhlen, *Interactive simulation of aircraft noise in aural and visual virtual environments*, *Applied Acoustics* **101**, 24 (2016).
- [40] P. R. Cook, *Physically Informed Sonic Modeling (PhISM): Percussive Synthesis*, *Computer Music Journal* **21**, 38 (1997).
- [41] F. Avanzini, M. Rath, and D. Rocchesso, *Physically-based audio rendering of contact*, in *2002 IEEE International Conference on Multimedia and Expo*, Vol. 2 (2002) pp. 445–448.
- [42] K. van den Doel and D. Pai, *Modal synthesis for vibrating objects*, in *Audio Anecdotes III: Tools, Tips, and Techniques for Digital Audio*, edited by K. Greenebaum and R. Barzel (AK Peters, Wellesley, Massachusetts, 2007) pp. 99–120.
- [43] L. Turchet, *Footstep sounds synthesis: Design, implementation, and evaluation of foot–floor interactions, surface materials, shoe types, and walkers' features*, *Applied Acoustics* **107**, 46 (2016).

- [44] 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*, *Journal of Sound and Vibration* **193**, 123 (1996).
- [45] J. M. Adrien, *The missing link: Modal synthesis*, in *Representation of Musical Signal*, edited by De Poli G., Picialli A. and Roads, C. (MIT Press, Cambridge, Massachusetts, 1991) pp. 269–297.
- [46] L. O. Chua, C. A. Desoer, and E. S. Kuh, eds., *Linear and nonlinear circuits* (McGraw-Hill, Singapore, 1987).
- [47] H. Kuttruff, *Room Acoustics*, 5th ed. (Spon Press, 2009).
- [48] R. Stefanelli, *Kurvenkreischen: Untersuchung der Rahmenbedingungen, die zu Kreischen führen und akustische Modellierung der kreisenden Räder*, dissertation, ETH Zürich (2006), no. 16686.
- [49] D. J. Thompson, N. Vincent, and P. E. Gautier, *Validation of a model for railway rolling noise using field measurements with sinusoidally profiled wheels*, *Journal of Sound and Vibration* **223**, 587 (1999).
- [50] D. Sehu, J. M. Wunderli, K. Heutschi, T. Thron, M. Hecht, A. Rohrbeck, and T. Ledermann, *sonRail - Projektdokumentation*, 7. Oktober 2010, Tech. Rep. (Empa, Swiss Federal Laboratories for Materials Science and Technology, Switzerland, 2010).
- [51] K. Heutschi, *Sound propagation over ballast surfaces*, *Acta Acustica United with Acustica* **95**, 1006 (2009).

# 6

## Conclusion and outlook

### 6.1. Conclusions

In this research, environmental acoustical auralization models for wind farms, passenger cars and rail vehicles have been developed. These models complement previously published synthesis models for long-haul aircraft and together largely cover the major environmental noise sources. The models separately describe acoustical sources, sound propagation and sound reproduction. They describe the sources by different components and require only a few model input parameters. Briefly, the auralization models are object-based, componential, and parametric. They are thus very versatile with respect to scenarios and reproduction.

It was found that the development of environmental auralization models not only requires profound knowledge about acoustics and signal processing but also understanding of the operating principles and operating conditions of the considered noise sources. It is challenging to find an adequate and compact representation of the source characteristics and to therefrom synthesize appropriate source signals. As compared to classical environmental noise prediction, many additional aspects have to be considered in auralization in order to produce a realistic hearing impression. To achieve realism in auralizations, it was found that variation with respect to time, frequency, space and orientation is crucial. Examples for such variations are short-term source effects such as rattling sounds from car engines or railway wheel flats, the effects of atmospheric turbulence, narrowband source characteristics, the source extension, and the relevance of directivity. (In the presented models some additional, observed variations remain unaccounted for.)

While the propagation simulation within the presented auralization models relies on physical models, the presented source models are partly physically motivated and partly semi-empiric. The physical components of the source models require non-acoustical input parameters that describe the geometry, kinematics and dynamics of the sources and can be obtained either from measurements or from calculations. In contrast, the semi-empiric components of the source models require

acoustical data to derive the synthesizer parameters. In this work, the synthesizer parameters were extracted from specific acoustical measurements. Whether in an application it suffices to adjust the synthesizer parameters solely based on estimates, predictions, literature or simple measurements, or whether elaborate acoustical measurements have to be performed depends on the specific application, i.e. the scenarios and the objective of the auralization.

The presented models are not purposely designed for real time application. Although their implementations were not profoundly optimized for low computational load, i.e. low calculation time, the immission signal from a single source can be rendered in real time. As with these models the computational load depends linearly on the number of considered sources, the real time capability of the models is limited by the number of considered sources. Concluding, the real time auralization of a wind farm seems feasible. High density road traffic and train pass-bys require more than 100 independent sources which makes their real time rendering unlikely within the next few years. However, for special cases the potential for model simplifications that substantially lower computational load is considered large. An evident example is that for road traffic at high speeds, large parts of the elaborate engine noise synthesis can be neglected due to the dominance of rolling noise.

Auralizations created with the presented models feature a high audio quality. They are judged as moderately to very plausible, and moderately to very realistic by expert listeners. The weakest elements in the synthesis are sounds that are perceived as too monotonous over time. This is the case for tire noise as well as railway rolling noise which both show more temporal level variations in measurements as compared to the synthesis. When synthetic sounds are directly compared to recorded sounds, the former can be correctly identified in most cases. The correct identification is mainly due to non-ideal details in the recordings that are missing in the synthesis, e.g. ambient sounds, wind noise, microphone noise or singular sounds. Nevertheless, the sound characteristics of an exposure situation is generally convincingly reproduced. It was further confirmed that the absence of ambient sounds in environmental auralizations is detrimental for plausibility. Adding recorded ambient sounds to the synthesis increases plausibility and realism by masking unnaturally quiet synthetic sounds, by giving audible anchor points (e.g. known reference sounds such as speech or bird song), by giving familiarity, and by providing context to the virtual scene. It can be concluded that although the effort to achieve realism is considerable, a plausible auralization of environmental noise is feasible.

## 6.2. Future work

Future work should focus both on further development of environmental acoustical auralization models and on their application in studies, communication and planning.

Some of the presented source models are semi-empiric and require specific acoustical measurements to derive their input parameters. This limits their applicability to cases with sources that are similar to known sources with known conditions. The semi-empiric source models are thus potentially inapplicable to future scenarios

with novel technologies, where acoustical measurements are expensive or even impossible. For these cases, physical source models should be developed that allow auralizing acoustically unknown sources and conditions. For example, future research should improve the presented semi-empiric model for tire/road noise. With the presented model, for every combination of road surface type and tire type, the source parameters have to be determined based on controlled pass-by measurements. Instead, a physical source model which uses the tire and road surface geometry, and their material properties to model their dynamic interaction, would be beneficial [1]. This modeling approach allows auralizing tire/road noise during the design phase of new tires or low-noise asphalts.

On the one hand, the presented source models should be extended to describe more complex sound characteristics. For railway noise, such characteristics involve curve squeal and rail singing. Besides the calculation models themselves, also the sets of source parameters should be amended and improved. This is particularly the case for railway rolling and impact noise where it is particularly challenging to determine appropriate synthesizer parameters to create realistic train pass-by sounds. On the other hand, also model simplifications should be addressed, such as the consideration of fewer engine orders for passenger cars. Further, a unified source signal representation should be introduced that could allow exchanging source model outputs and using them by various auralization tools, such as game engines.

Also regarding propagation simulation, some activities should be fostered. Modern outdoor sound propagation engineering models represent the ground effect based on the analytical solution of the sound field generated by a point source above an infinitely extended locally reacting impedance plane. This approach was also followed in the presented as well as in previously published auralization models. However, it was found that this approach is not capable of satisfyingly reproducing the ground effect as observed in measurements of near-ground sources such as road and rail vehicles. Compared to measurements, the ground effect is often spectrally too distinct, and perceived as irritating due to flanging effects. Possible reasons for these deviations are uneven terrain, scattering, extended reacting surface or the source extension violating the model assumption of a concentrated source. These aspects should be clarified in order to adequately reproduce ground effect in simple geometries. Further topics concern the consideration of complex propagation situations in the auralization such as railway line cuttings or street canyons with multiple reflections, scattering and shielding. Also near-ground propagation over long distances with multiple ground reflections or reflections at extended reacting boundaries, such as porous asphalts or track ballast, are open issues. Also current models for effects of atmospheric turbulence on sound propagation, e.g. amplitude fluctuations and coherence loss, should be advanced [2, 3]. A desirable approach to handle geometrically complex situations seems the incorporation of wave-based numerical propagation models into the auralization process [4].

A further aspect is the use of auralization models in experiments. This involves listening tests for model validation but also the application in psychoacoustic studies and sleep studies. The models should also actively be used in communication to

the public, e.g. to demonstrate or compare noise measures. The environmental auralization models should also find their way into the urban planning process. This probably requires easy-to-use models and real time capability.

## References

- [1] J. Forssén, A. Hoffmann, and W. Kropp, *Auralization model for the perceptual evaluation of tyre–road noise*, *Applied Acoustics* **132**, 232 (2018).
- [2] F. Rietdijk, J. Forssén, and K. Heutschi, *Generating sequences of acoustic scintillations*, *Acta acustica united with Acustica* **103**, 331 (2017).
- [3] M. Arntzen and D. G. Simons, *Ground reflection with turbulence induced coherence loss in flyover auralization*, *International Journal of Aeroacoustics* **13**, 449 (2014).
- [4] M. Hornikx, *Ten questions concerning computational urban acoustics*, *Building and Environment* **106**, 409 (2016).

# Acknowledgements

## **Reto PIEREN**

First of all my thanks go to my promotor Dick Simons for his constant support during my PhD. I also would like to thank my copromotor Mirjam Snellen and the courteous employees at TU Delft for their kind help. I further thank the members of the doctoral committee for their valuable time while examining my work and I thank Lothar Bertsch for establishing contacts. Also my superiors Jean Marc Wunderli and Kurt Eggenschwiler at my current employer Empa in Switzerland deserve my sincerest thanks for their encouragement and the great support.

My gratitude also goes to my colleagues Kurt Heutschi and Beat Schäffer for their help, their contribution and their inspiration. Special thanks are due to Anja Lehmann for designing the book cover. An impersonal thank you goes to music—for giving me energy, stimulation and the ability to listen analytically.

Finally, I thank my parents for their assistance and my lovely daughter Alena for being my sunshine. Last but not least, my deepest gratitude and admiration goes to the most important person, my wife Simone who gave me maximum support—even in very difficult times.





## Supplementary audio and video files

Within this dissertation, audio data of many hours in duration has been produced. Most of the content was created during the model development process. This involved audio data from field or laboratory measurements, but considerably more audio data was obtained from calculation models, i.e. as output from different types and stages of sound synthesizers developed and tested during this work. Most of the latter data revealed artifacts, missing sound characteristics and many other surprises which gave insights and were essential for the model development. However, fortunately also some sounds worth presenting were obtained from the presented auralization models.

As an integral part of this dissertation, supplementary audio and video data is provided in an online repository. The aims of the data provision are to demonstrate the capabilities of the presented models to the interested reader, and to make the content available for other researchers. The repository can be accessed through the TU Delft website, where the corresponding files can be downloaded. The online

ID	Source	Content
#1	Wind turbine	WT1: recording (A)
#2	Wind turbine	WT1: synthesis (A)
#3	Wind turbine	WT2: recording (A)
#4	Wind turbine	WT2: synthesis (A)
#5	Passenger car	Pass-by simulation with low acceleration (V)
#6	Passenger car	Pass-by simulation with high acceleration (V)
#7	Rail vehicle	Synthesis of passenger train pass-by (A)
#8	Rail vehicle	Synthesis of freight train pass-by (A)

Table A.1: Supplementary audio (A) and video (V) data of different environmental noise sources.



repository can be found at <http://repository.tudelft.nl> by searching for this dissertation. Besides an electronic version of this dissertation (.pdf file), an archive file (.zip) can be downloaded. It contains the files listed in Table [A.1](#) with their filenames starting with the allocated ID number.

# Curriculum Vitæ

## **Reto PIEREN**

07-10-1982 Born in Basel, Switzerland.

### Education

1998–2001 Grammar School  
Gymnasium Liestal, Switzerland

2002–2008 Undergraduate in Electrical Engineering &  
Information Technology  
ETH Zurich, Switzerland

08-04-2008 Master of Science in Electrical Engineering &  
Information Technology  
ETH Zurich, Switzerland

17-02-2015 Acoustician SGA-SSA  
Swiss Acoustical Society

2017–2018 Doctoral candidate at the faculty of Aerospace Engineering  
Delft University of Technology, Netherlands

### Awards

2015 Prix Colladon, Swiss Acoustical Society

2011 Design Prize Switzerland

### Work Experience

2008–2009 Acoustical engineer  
Applied Acoustics GmbH, Switzerland

2009–today Scientist at the Laboratory for Acoustics / Noise Control  
Empa - Swiss Federal Laboratories for Materials Science and Technology



# List of Publications

## Peer-reviewed journal articles and book chapters

30. **R. Pieren**, B. Schäffer, S. Schoenwald, K. Eggenschwiler, *Sound absorption of textile curtains – Theoretical models and validations by experiments and simulations*, *Textile Research Journal* **88**(1), 36–48 (2018).
29. B. Schäffer, **R. Pieren**, S.J. Schlittmeier, M. Brink, *Effects of different spectral shapes and amplitude modulation of broadband noise on annoyance reactions in a controlled listening experiment*, *International Journal of Environmental Research and Public Health*, **15**(5), 1029, 1–17 (2018).
28. U. Wissen Hayek, **R. Pieren**, K. Heutschi, M. Manyoky, A. Grêt-Regamey *Exploring the qualities of GIS-based visual-acoustic simulations of wind parks to support public opinion forming*, In C. Yamu et al.: *The Virtual and The Real in Planning and Urban Design: Perspectives, Practices and Applications*; Routledge, Taylor and Francis group, London and New York, ISBN 978-1138283480, 233–251 (2018).
27. M. Brink, B. Schäffer, **R. Pieren**, J. M. Wunderli, *Conversion between noise exposure indicators Leq24h, LDay, LEvening, LNight, Ldn and Lden: Principles and practical guidance*, *International Journal of Hygiene and Environmental Health* **221**, 54–63 (2018).
26. R. Ribe, M. Manyoky, U. Wissen Hayek, **R. Pieren**, K. Heutschi, *Dissecting Perceptions of Wind Energy Projects: A Laboratory Experiment Using High-quality Audio-visual Simulations to Analyze Experiential Versus Acceptability Ratings and Information Effects*, *Landscape and Urban Planning* **169**, 131–147 (2018).
25. H. Héritier, D. Vienneau, M. Foraster, I. Eze, E. Schaffner, L. Thiesse, F. Rudzik, M. Habermacher, M. Köpfli, **R. Pieren**, A. Schmidt-Trucksäss, M. Brink, C. Cajochen, J. M. Wunderli, N. Probst-Hensch, M. Rössli, *Diurnal variability of transportation noise exposure and cardiovascular mortality: a nationwide cohort study from Switzerland*, *International Journal of Hygiene and Environmental Health*, **221**(3), 556–563 (2018).
24. F. Rudzik, L. Thiesse, **R. Pieren**, J. M. Wunderli, M. Brink, M. Foraster, H. Héritier, I. Eze, C. Corrado, D. Vienneau, N. Probst-Hensch, M. Rössli, C. Cajochen, , *Sleep spindle characteristics and arousability from night-time transportation noise exposure in healthy young and older individuals*, *Sleep*, in press (2018).
23. L. Thiesse, F. Rudzik, K. Spiegel, R. Leproult, **R. Pieren**, J. M. Wunderli, M. Foraster, H. Héritier, I. Eze, D. Vienneau, M. Meyer, M. Brink, N. Probst-Hensch, M. Rössli, C. Cajochen, , *Adverse impact of nocturnal transportation noise on glucose regulation in healthy young adults: effect of different noise scenarios*, *Environment International*, in press (2018).

22. **R. Pieren**, K. Heutschi, J. M. Wunderli, M. Snellen, D. G. Simons, *Auralization of railway noise: Emission synthesis of rolling and impact noise*, *Applied Acoustics* **127**, 34–45 (2017).
21. B. Schäffer, **R. Pieren**, F. Mendolia, M. Basner, M. Brink, *Noise exposure-response relationships established from repeated binary observations: modeling approaches and applications*, *Journal of the Acoustical Society of America*, **141**(5), 3175–3185 (2017).
20. H. Héritier, D. Vienneau, M. Foraster, I. Eze, E. Schaffner, L. Thiesse, F. Rudzik, M. Habermacher, M. Köpfli, **R. Pieren**, M. Brink, C. Cajochen, J. M. Wunderli, N. Probst-Hensch, M. Rööslü, *Transportation noise exposure and cardiovascular mortality: a nationwide cohort study from Switzerland*, *European Journal of Epidemiology* **32**, 307–315 (2017).
19. I. Eze, M. Foraster, E. Schaffner, D. Vienneau, H. Héritier, F. Rudzik, L. Thiesse, **R. Pieren**, M. Imboden, A. von Eckardstein, C. Schindler, C. Cajochen, M. Brink, J. M. Wunderli, M. Rööslü, N. Probst-Hensch, *Long-term exposure to transportation noise and air pollution in relation to incident diabetes in the SAPALDIA study*, *International Journal of Epidemiology* **46** (4), 1115–1125 (2017).
18. M. Foraster, I. Eze, E. Schaffner, D. Vienneau, H. Héritier, S. Endes, F. Rudzik, L. Thiesse, **R. Pieren**, C. Schindler, A. Schmidt-Trucksäss, M. Brink, C. Cajochen, J. M. Wunderli, M. Rööslü, N. Probst-Hensch, *Exposure to road, railway, and aircraft noise and arterial stiffness in the SAPALDIA study: annual average noise levels and temporal noise characteristics*, *Environmental Health Perspectives* **125** (9), 097004 (2017).
17. I. Eze, M. Imboden, M. Foraster, E. Schaffner, A. Kumar, D. Vienneau, H. Héritier, F. Rudzik, L. Thiesse, **R. Pieren**, A. von Eckardstein, C. Schindler, M. Brink, J. M. Wunderli, C. Cajochen, M. Rööslü, N. Probst-Hensch, *Exposure to night-time noise, melatonin-regulating gene variants and change in glycemia in adults*, *International Journal of Environmental Research and Public Health* **2017** (4), 1492 (2017).
16. **R. Pieren**, T. Bütler, K. Heutschi, *Auralization of Accelerating Passenger Cars using Spectral Modeling Synthesis*, *Applied Sciences* **6**(1), 5, 1–27 (2016).
15. J. M. Wunderli, **R. Pieren**, M. Habermacher, D. Vienneau, C. Cajochen, N. Probst-Hensch, M. Rööslü, B. Brink *Intermittency ratio: A metric reflecting short-term temporal variations of transportation noise exposure*, *Journal of Exposure Science and Environmental Epidemiology* **26** (6), 575–585 (2016).
14. B. Schäffer, S. J. Schlittmeier, **R. Pieren**, K. Heutschi, M. Brink, R. Graf, J. Hellbrück, *Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study*, *Journal of the Acoustical Society of America* **139**, 5, 2949–2963 (2016).
13. M. Manyoky, U. Wissen Hayek, **R. Pieren**, K. Heutschi, A. Gret-Regamey *Evaluating a visual-acoustic simulation for wind park assessment*, *Landscape and Urban Planning* **153**, 180–197 (2016).
12. K. Eggenschwiler, K. Heutschi, B. Schäffer, **R. Pieren**, H. Bögli, M. Bärlocher *Noise from wind turbines – contributions of Switzerland to its assessment*, *Lärmbekämpfung*, **5/2016**, 159–167 (2016).

11. **R. Pieren**, K. Heutschi *Modelling parallel assemblies of porous materials using the equivalent circuit method*, *Journal of the Acoustical Society of America* **137** 2, EL131–136 (2015).
10. **R. Pieren**, K. Heutschi *Predicting sound absorption coefficients of lightweight multi-layer curtains using the equivalent circuit method*, *Applied Acoustics* **92**, 27–41 (2015).
9. **R. Pieren**, K. Heutschi, M. Müller, M. Manyoky, K. Eggenschwiler *Auralization of Wind Turbine Noise: Emission Synthesis*, *Acta Acustica United with Acustica* **100** 1, 25–33 (2014).
8. K. Heutschi, **R. Pieren**, M. Müller, M. Manyoky, U. Wissen Hayek, K. Eggenschwiler *Auralization of Wind Turbine Noise: Propagation Filtering and Vegetation Noise Synthesis*, *Acta Acustica United with Acustica* **100** 1, 13–24 (2014).
7. M. Manyoky, U. Wissen Hayek, K. Heutschi, **R. Pieren**, A. Gret-Regamey *Developing a GIS-Based Visual-Acoustic 3D Simulation for Wind Farm Assessment*, *ISPRS International Journal of Geo-Information* **3**, 1, 29–48 (2014).
6. **R. Pieren**, *Sound absorption modeling of thin woven fabrics backed by an air cavity*, *Textile Research Journal* **82**, 864–874 (2012).
5. J. M. Wunderli, **R. Pieren**, K. Heutschi *The Swiss shooting sound calculation model sonARMS*, *Noise Control Engineering Journal* **60**, 4, 224–235 (2012).
4. M. Brink, M. Basner, S. Omlin, **R. Pieren**, *Do church bells disturb sleep? - A study of the effects of nightly church bell ringing on the population*, *Lärmbekämpfung* **7**, 1, 29–41 (2012).
3. M. Manyoky, U. Wissen Hayek, T. Klein, **R. Pieren**, K. Heutschi, A. Gret-Regamey *Concept for Collaborative Design of Wind Farms facilitated by an Interactive GIS-based Visual-acoustic 3D Simulation*, Peer Reviewed Proceedings of Digital Landscape Architecture, 2012, 297–306 (2012).
2. **R. Pieren**, J. M. Wunderli, *A Model to Predict Sound Reflections from Cliffs*, *Acta Acustica United with Acustica* **97**, 243–253 (2011).
1. M. Brink, S. Omlin, C. Müller, **R. Pieren**, M. Basner, *An event-related analysis of awakening reactions due to nocturnal church bell noise*, *Science of the Total Environment* **409**, 5210–5220 (2011).

## Conference papers and printed abstracts

33. **R. Pieren**, D. Lauper, K. Heutschi, *Demonstrator for rail vehicle pass-by events*, invited paper, In Proceedings of the 11th European Congress and Exposition on Noise Control Engineering (Euronoise 2018), Crete, Greece, 2018.
32. **R. Pieren**, L. Bertsch, J. Blinstrub, B. Schäffer, J. M. Wunderli, *Simulation process for perception-based noise optimization of conventional and novel aircraft concepts*, In Proceedings of the *AIAA Aerospace Sciences Meeting, AIAA SciTech Forum*, (AIAA 2018-0266), Kissimmee, Florida, USA, 2018.

31. M. Rööslü, J. M. Wunderli, M. Brink, C. Cajochen, D. Vienneau, M. Foraster, I. C. Eze, H. H eritier, E. Schaffner, L. Thiesse, F. Rudzik, **R. Pieren**, M. Habermacher, M. K opfli, N. Probst-Hensch, *SiRENE: Kurz- und langfristige Auswirkungen der Verkehrsl armbelastung*, In Proceedings of the 44th German Annual Conference on Acoustics (DAGA 2018), Munich, Germany, 2018.
30. **R. Pieren**, K. Heutschi, R. Aalmoes, D. G. Simons, *Evaluation of Auralization and Visualization Systems for Railway Noise Scenes*, invited paper, In Proceedings of the 46th International Congress of Noise Control Engineering (INTER-NOISE 2017), Hong Kong, 2017.
29. B. Sch affer, **R. Pieren**, S. J. Schlittmeier, M. Brink, K. Heutschi, *Annoyance to wind turbine noise – influence of different acoustical characteristics*, In Proceedings of the 46th International Congress of Noise Control Engineering (INTER-NOISE 2017), Hong Kong, 2017.
28. M. Brink, B. Sch affer, **R. Pieren**, J. M. Wunderli, *Empirical estimation of conversion terms between the noise metrics LDay, LNight, Ldn, and Lden*, In Proceedings of the 46th International Congress of Noise Control Engineering (INTER-NOISE 2017), Hong Kong, 2017.
27. B. Sch affer, **R. Pieren**, F. Mendolia, M. Basner, M. Brink, *On the application of two statistical approaches to establish noise exposure-response relationships from repeated binary observations*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
26. L. Thiesse, F. Rudzik, **R. Pieren**, J. M. Wunderli, K. Spiegel, R. Leproult, M. Foraster, H. H eritier, I. C. Eze, D. Vienneau, M. Brink, N. Probst-Hensch, M. R oosli, C. Cajochen, *Temporal variation of transportation noise during sleep impacts on glucose metabolism*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
25. F. Rudzik, L. Thiesse, **R. Pieren**, J. M. Wunderli, M. Brink, M. Foraster, H. H eritier, I. C. Eze, D. Vienneau, N. Probst-Hensch, M. R oosli, C. Cajochen, *Spindle density predicts arousability from acoustic stimulation during sleep*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
24. M. Foraster, I. C. Eze, D. Vienneau, M. Brink, C. Cajochen, H. H eritier, M. Imboden, A. Jeong, F. Rudzik, L. Thiesse, **R. Pieren**, E. Schaffner, J. M. Wunderli, M. R oosli, N. Probst-Hensch, *Long-term exposure to road, railway, and aircraft noise levels and their association with incidence of obesity and obesity parameters*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
23. M. Foraster, I. C. Eze, E. Schaffner, D. Vienneau, H. H eritier, S. Endes, F. Rudzik, L. Thiesse, **R. Pieren**, C. Schindler, A. Schmidt-Trucks ass, M. Brink, C. Cajochen, J. M. Wunderli, M. R oosli, N. Probst-Hensch, *Exposure to source-specific transportation noise levels and temporal noise characteristics in association with arterial stiffness*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.

22. I. C. Eze, M. Imboden, M. Foraster, E. Schaffner, A. Kumar, D. Vienneau, H. Hérítier, F. Rudzik, L. Thiesse, **R. Pieren**, A. von Eckardstein, C. Schindler, M. Brink, J. M. Wunderli, C. Cajochen, M. Rösli, N. Probst-Hensch, *Association between night-time road traffic noise and perturbations in glucose control may be modified by sleep-related parameters*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
21. M. Rösli, D. Vienneau, M. Foraster, I. C. Eze, H. Hérítier, E. Schaffner, L. Thiesse, F. Rudzik, **R. Pieren**, M. Habermacher, M. K"opfli, M. Brink, C. Cajochen, J. M. Wunderli, N. Probst-Hensch, *Short and long term effects of transportation noise exposure (SIRENE): an interdisciplinary approach*, In Proceedings of the 12th Congress on Noise as a Public Health Problem (ICBEN 2017), Zurich, Switzerland, 2017.
20. F. Rudzik, L. Thiesse, **R. Pieren**, J. M. Wunderli, M. Brink, D. Vienneau, N. Probst-Hensch, M. Rösli, C. Cajochen, *Spindle density predicts arousability from acoustic stimulation during sleep*, [Sleep 40 \(S1\), A36–A37 \(2017\)](#).
19. L. Thiesse, F. Rudzik, **R. Pieren**, J. M. Wunderli, K. Spiegel, R. Leproult, D. Vienneau, M. Brink, N. Probst-Hensch, M. Rösli, C. Cajochen, *Temporal variation of transportation noise during sleep impacts on glucose metabolism*, [Sleep 40 \(S1\), A36–A36 \(2017\)](#).
18. **R. Pieren**, J. M. Wunderli, A. Zemp, S. Sohr, K. Heutschi, *Auralisation of Railway Noise: A Concept for the Emission Synthesis of Rolling and Impact Noise*, invited paper, In Proceedings of the 45th International Congress of Noise Control Engineering (INTER-NOISE 2016), Hamburg, Germany, 2016.
17. J. M. Wunderli, **R. Pieren**, C. Cajochen, N. Probst-Hensch, M. Rösli, M. Brink, *Parameter study on IR, a metric reflecting short-term temporal variations of transportation noise exposure*, invited paper, In Proceedings of the 45th International Congress of Noise Control Engineering (INTER-NOISE 2016), Hamburg, Germany, 2016.
16. M. Brink, **R. Pieren**, M. Foraster, D. Vienneau, I. Eze, E. Schaffner, H. Hérítier, C. Cajochen, N. Probst-Hensch, M. Rösli, J. M. Wunderli, *Do short-term temporal variations of noise exposure explain variance of noise annoyance?*, invited paper, In Proceedings of the 45th International Congress of Noise Control Engineering (INTER-NOISE 2016), Hamburg, Germany, 2016.
15. F. Rudzik, L. Thiesse, **R. Pieren**, J. M. Wunderli, M. Brink, N. Probst-Hensch, M. Rösli, C. Cajochen, *Effects of continuous and intermittent transportation noise on sleep fragmentation*, invited paper, In Proceedings of the 45th International Congress of Noise Control Engineering (INTER-NOISE 2016), Hamburg, Germany, 2016.
14. L. Thiesse, F. Rudzik, **R. Pieren**, J. M. Wunderli, K. Spiegel, M. Brink, N. Probst-Hensch, M. Rösli, C. Cajochen, *Short-term effects of nocturnal transportation noise on glucose metabolism*, invited paper, In Proceedings of the 45th International Congress of Noise Control Engineering (INTER-NOISE 2016), Hamburg, Germany, 2016.
13. **R. Pieren**, T. Büttler, K. Heutschi, *Auralisation of Accelerating Passenger Cars*, invited paper, In Proceedings of the 10th European Conference on Noise Control (Euronoise 2015), Maastricht, The Netherlands, 2015.

12. B. Schäffer, S. Schlittmeier, K. Heutschi, M. Brink, R. Graf, **R. Pieren**, J. Hellbrück, *Annoyance potential of wind turbine noise compared to road traffic noise*, In Proceedings of the 10th European Conference on Noise Control (Euronoise 2015), Maastricht, The Netherlands, 2015.
11. I. Clark, M. Brink, G. Tinguely, M. Röögli, N. Probst-Hensch, J. M. Wunderli, **R. Pieren**, C. Cajochen, *Inter-individual differences in the effects of night-time noise exposure on sleep and cognitive performance during daytime*, Journal of Sleep Research **23 (S1)**, 94, 2014.
10. D. Vienneau, M. Brink, C. Cajochen, I. Eze, H. Héritier, I. Karipidis, D. Keidel, M. Köpfli, **R. Pieren**, N. Probst-Hensch, F. Rudzik, E. Schaffner, L. Thiesse, J. M. Wunderli, M. Röögli, *SiRENE - an interdisciplinary study on the health effects of transportation noise exposure: a study protocol*, In Proceedings of the 11th Congress on Noise as a Public Health Problem (ICBEN 2014), Nara, Japan, 2014.
9. K. Heutschi, **R. Pieren**, *Auralization of Wind Turbines*, invited paper, In Proceedings of the AIA-DAGA 2013, International Conference on Acoustics, including the 39th German Annual Conference on Acoustics (DAGA), Merano, Italy, 2013.
8. **R. Pieren**, *Sound Absorption Modelling of Thin Lightweight Curtains*, In Proceedings of the 9th European Conference on Noise Control (Euronoise 2012), Prague, Czech Republic, 2012.
7. M. Müller, **R. Pieren**, J. Bitzer *Modell zur Synthese akustischer Emissionen von Windturbinen*, In Proceedings of the 38th German Annual Conference on Acoustics (DAGA 2012), Darmstadt, Germany, 2012.
6. K. Eggenschwiler, **R. Pieren**, A. Douglas, M. Lachmann *Leichte und lichtdurchlässige schallabsorbierende Vorhänge*, In Proceedings of the 38th German Annual Conference on Acoustics (DAGA 2012), Darmstadt, Germany, 2012.
5. **R. Pieren**, J. M. Wunderli *Sound Reflections from Cliffs and Reverberation in an Alpine Valley*, In Proceedings of the Forum Acusticum 2011, Triennial European Conference on Acoustics, Aalborg, Denmark, 2011.
4. **R. Pieren**, *Soundscape im Grossraumbüro (Open Plan Space)*, invited paper, In Forum Holz-Bau Spezial Schallschutz & Akustik, Internationale Schall- und Akustiktage (ISA 2011), Bad Wörishofen, Germany, 2011.
3. M. Brink, S. Omlin, M. Basner, **R. Pieren**, *An event-related examination of awakenings due to nocturnal church bell sounds*, In Proceedings of the 10th Congress on Noise as a Public Health Problem (ICBEN 2011), London, UK, 2011.
2. B. Schulte-Fortkamp, K. Genuit, A. Fiebig, V. Acloque, M. Di Gabriele, M. Horvat, M. Masullo, **R. Pieren**, B. Seckin, K. Voigt, M. Yang, *Education in Soundscape – A seminar with young scientists in short term scientific mission „Soundscape - Measurement, Analysis, Evaluation*, invited paper, In Proceedings of the 20th International Congress on Acoustics 2010 (ICA 2010), Sydney, Australia, 2010.
1. **R. Pieren**, *A Passive Method for the Determination of Acoustical Parameters in Occupied Rooms*, In Proceedings of the NAG/DAGA 2009, International Conference on Acoustics, including the 35th German Annual Conference on Acoustics (DAGA), Rotterdam, The Netherlands, 2009.