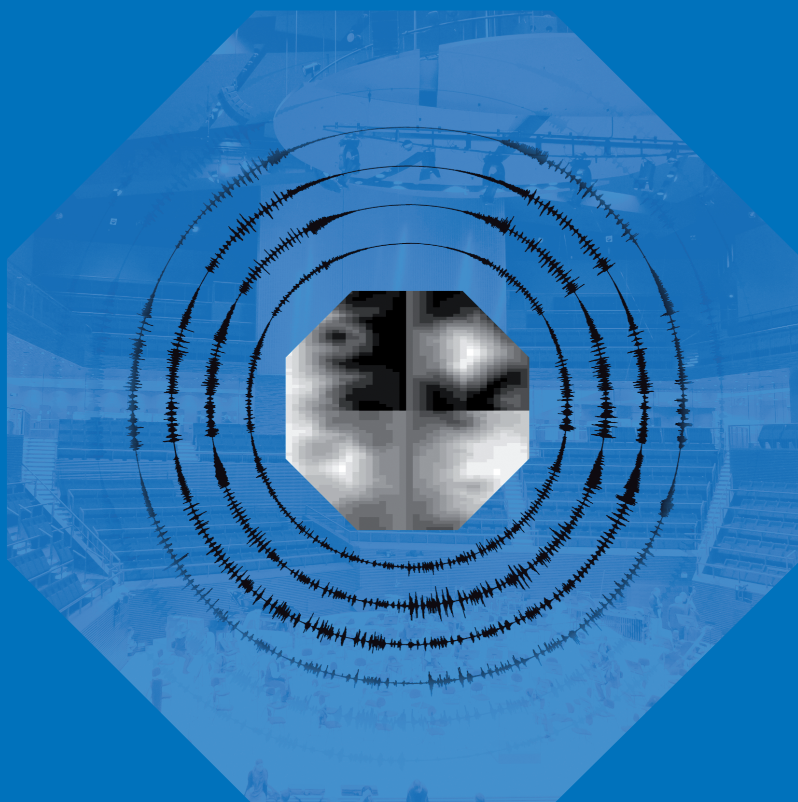
Department of Media Technology

# A virtual symphony orchestra for studies on concert hall acoustics

Jukka Pätynen



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
**DOCTORAL DISSERTATIONS** 86/2011

# A virtual symphony orchestra for studies on concert hall acoustics

**Jukka Pätynen**

Doctoral dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the School of Science for public examination and debate in Auditorium T2 at the Aalto University School of Science (Espoo, Finland) on the 18th of November, 2011, at 12 noon.

**Aalto University**
**School of Science**
**Department of Media Technology**

**Supervisor**
Professor Lauri Savioja

**Instructor**
Adjunct professor Tapio Lokki

**Preliminary examiners**
Dr. Nicola Prodi, University of Ferrara, Italy
Dr. John Bradley, National Research Council, Canada

**Opponent**
Dr. Olivier Warusfel, IRCAM, France

Aalto University publication series
**DOCTORAL DISSERTATIONS** 86/2011



ISBN 978-952-60-4291-6 (pdf)
ISBN 978-952-60-4290-9 (printed)
ISSN-L 1799-4934
ISSN 1799-4942 (pdf)
ISSN 1799-4934 (printed)

Aalto Print
Helsinki 2011

Finland

The dissertation can be read at http://lib.tkk.fi/Diss/

**Aalto University**

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

# Abstract

**Author**
Jukka Pätynen

**Name of the doctoral dissertation**
A virtual symphony orchestra for studies on concert hall acoustics

**Publisher** School of Science

**Unit** Department of Media Technology

**Series** Aalto University publication series DOCTORAL DISSERTATIONS 86/2011

**Field of research** Acoustics

**Manuscript submitted** 20 June 2011 **Manuscript revised** 21 September 2011

**Date of the defence** 18 November 2011 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

**Abstract**

Traditionally, concert hall acoustics is evaluated by listening to live concerts, which makes a direct comparison challenging. This thesis presents new tools and methods in the domain of the room acoustics evaluation, studies, and auralization. Auralization stands for the process of rendering an existing or modeled acoustic space in a way that it can be presented to the listener as he/she was listening to a sound inside the space under study.

An essential topic in this thesis is a framework for studying room acoustics with a wide-area loudspeaker array. The proposed loudspeaker orchestra consists of a number of loudspeakers that are positioned in the shape resembling a symphony orchestra on a stage. The acoustics can be evaluated in-situ by playing back anechoic signals, or in laboratory conditions via convolution of the impulse responses measured from the loudspeaker orchestra. The presented method enables a direct comparison of concert halls and it has been successfully applied in practice in several research articles.

The principal requirement for such a loudspeaker orchestra is anechoic signals of high quality. For this purpose, a method and implementation of a system for recording the symphony orchestra instruments individually is presented. As the result, a selection of anechoic orchestral music is obtained with perfect channel separation. The recordings, intended for advancing the research on acoustics and auralization, are published for academic use. Directivity of the orchestra instruments in performance situation is investigated with anechoic measurements. The results for different instruments can be compared against each other or applied directly into auralizations. Data from the directivity measurements is also applied in the objective analysis of the presented loudspeaker orchestra. Furthermore, the implemented measurement system is utilized in investigating the sound radiation of the balloons, which are often used in room acoustic measurements.

Related to the anechoic recordings, a novel approach to creating an impression of a group of musicians from a single recorded player is proposed. The method is mainly based on the video and audio analysis of the temporal differences between orchestra string players. The method is particularly beneficial with the anechoic recordings, where recording an instrument section is not possible, and recording a large number of musicians individually is time-consuming. The listening test results show that the presented method provides a plausible simulation of an instrument section sound in comparison to an industry-standard method.

**Keywords** room acoustics, anechoic recordings, concert hall, symphony orchestra, directivity

**ISBN (printed)** 978-952-60-4290-9 **ISBN (pdf)** 978-952-60-4291-6

**ISSN-L** 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

**Location of publisher** Espoo **Location of printing** Helsinki **Year** 2011

**Pages** 187 **The dissertation can be read at** http://lib.tkk.fi/Diss/

**Aalto-yliopisto**

**Aalto-yliopisto, PL 11000, 00076 Aalto** www.aalto.fi

# Tiivistelmä

**Tekijä**
Jukka Pätynen

**Väitöskirjan nimi**
Keinotekoinen sinfoniaorkesteri saliakustiikan tutkimusmenetelmänä

**Julkaisija** Perustieteiden korkeakoulu

**Yksikkö** Mediatekniikan laitos

**Sarja** Aalto University publication series DOCTORAL DISSERTATIONS 86/2011

**Tutkimusala** Akustiikka

**Käsikirjoituksen pvm** 20.06.2011 **Korjatun käsikirjoituksen pvm** 21.09.2011

**Väitöspäivä** 18.11.2011 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

**Tiivistelmä**

Saliakustiikkaa arvioidaan yleisimmin konserteissa käymällä. Tästä johtuen akustisten erojen tarkka vertailu on vaikeaa. Tässä väitöskirjassa esitellään uusia menetelmiä ja parannuksia olemassa oleviin saliakustiikan arvioinnin tutkimusmenetelmiin sekä auralisaatioon. Auralisaatiolla tarkoitetaan prosessia, missä mallinnettavan tai jo olemassaolevan tilan akustiikka pyritään esittämään kuuntelijalle niin, että toistettu ääniympäristö kuulostaa siltä, kuin ääntä kuunneltaisiin kyseisessä tilassa.

Työn keskeisenä osana on laaja-alaisesta kaiutinryhmästä koostuva järjestelmä - kaiutinorkesteri - jonka avulla saliakustiikkaa voidaan tutkia paikan päällä konserttisaleissa, tai myöhemmin laboratorio-olosuhteissa tallentamalla saleissa mitatut impulssivasteet. Tällöin menetelmä mahdollistaa suoran vertailun eri akustiikkojen välillä. Kaiutinorkesteria on hyödynnetty menestyksekkäästi tieteellisissä julkaisuissa.

Kaiutinorkesterin yhteydessä käytettäven äänisignaalien on oltava kaiuttomia. Työssä esitellään orkesteriäänitykset, jotka on toteutettu äänittämällä sinfoniaorkesterin soittimet yksitellen kaiuttomassa huoneessa. Tieteelliseen käyttöön julkaisujen äänitysten tavoitteena on edistää yleistä huoneakustiikan tutkimusta sekä auralisaatioiden todenmukaisuutta. Äänitysten yhteydessä tehtyjä orkesterisoittimien suuntaavuusmittauksia käsitellään kattavasti. Tuloksia käytetään myös kaiutinorkesterin objektiiviseen arviointiin. Lisäksi työssä esitellään ilmapallojen räjähdyksen suuntaavuusominaisuuksia aikaisempia tutkimuksia tarkemmin - ilmapalloja sekä muita impulssimaisia äänilähteitä voidaan käyttää sähköakustisten menetelmien sijaan impulssivasteiden mittauksessa.

Kaiuttomiin äänityksiin liittyen työssä esitellään uusi menetelmä, jolla yksittäinen äänitetty jousisoittaja saadaan kuulostamaan siltä kuin soittajia olisi useampia. Menetelmässä simuloitujen soittajien keskinäisiä aikaeroja muutetaan sinfoniaorkesterin jousisoittajien yhtäaikaisuuden analysoinnista saatavien tutkimustulosten perusteella. Lisäksi sävelkorkeuteen, äänensävyyn sekä soittimien keskinäiseen tasapainoon luodaan vaihteluita. Menetelmästä on erityisesti hyötyä kaiuttomien orkesteriäänitysten yhteydessä, sillä soittajat joudutaan yleensä äänittämään yksitellen. Kuuntelukokeella saatujen tulosten perusteella väitöskirjassa esitelty menetelmä luo soitinryhmän ominaisen sointivärin vertailukohtana olevaa yleisesti käytettyä menetelmää paremmin.

**Avainsanat** saliakustiikka, kaiuttomat äänitykset, orkesteri, suuntaavuus

**ISBN (painettu)** 978-952-60-4290-9 **ISBN (pdf)** 978-952-60-4291-6

**ISSN-L** 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

**Julkaisupaikka** Espoo **Painopaikka** Helsinki **Vuosi** 2011

**Sivumäärä** 187 **Luettavissa verkossa osoitteessa** http://lib.tkk.fi/Diss/

# Preface

The research work for the results that are presented in this thesis has been carried out at the Telecommunications Software and Multimedia Laboratory, Department of Media Technology, Helsinki University of Technology during 2007-2009, and in Aalto University during 2010-2011. I want to thank the Finnish Foundation for Technology Promotion and Nokia Foundation for having assessed the presented research worth the financial support. The research leading to these results has also received funding from the Academy of Finland, project no. [119092], and the European Research Council under the European Community's Seventh Framefork Programme / ERC grant agreement no. [203636].

I am deeply indebted to my supervisor Prof. Lauri Savioja and instructor Dr. Tapio Lokki for the positive attitude and fruitful discussions during the course of the research eventually having led to the completion of this thesis. Tapio's inspiring and insightful support has been of tremendous help with the publications. I would also like to express my gratitude to Dr. Brian F.G. Katz and Dr. Ville Pulkki for the collaboration in the articles outside the Department of Media Technology.

I wish to thank the pre-examiners of this thesis, Dr. Nicola Prodi and Dr. John Bradley, for offering their expertise for providing valuable comments and feedback on the manuscript. Special thanks go to Beth Morton for proofreading the manuscript.

I am grateful to my research team fellows — Sampo, Samuel, Raine, Sakari, Antti, Heikki, Hannes, Robert, Alex — and those working at the Laboratory of Acoustics and Signal Processing, for providing an exhilarating and pleasant work atmosphere. I wish to acknowledge Mr. Sakari Tervo for the continuous collaboration in research and exchange of ideas.

Finally, I would like to express my most sincere thanks to my parents and Marika for their uncompromising support for the work and studies in all fields.

Espoo, September 26, 2011,

Jukka Pätynen

# Contents

# List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

**I** J. Pätynen, S. Tervo and T. Lokki. A Loudspeaker Orchestra for Concert Hall Studies. In *The Seventh International Conference On Auditorium Acoustics*, Oslo, Norway. (Published also in *Acoustics Bulletin* 34(6) pp. 32-37, 2009.), October 3-5 2008.

**II** J. Pätynen, V. Pulkki and T. Lokki. Anechoic Recording System for Symphony Orchestra. *Acta Acustica united with Acustica*, Vol. 94(6) pp. 856-865, December 2008.

**III** J. Pätynen and T. Lokki. Directivities of Symphony Orchestra Instruments. *Acta Acustica united with Acustica*, Vol. 96(1) pp.138-167, January 2010.

**IV** J. Pätynen. Directivities of Orchestra Instruments for Auralization. In *Proceedings of the EAA Symposium on Auralization*, Espoo, Finland, June 15-17 2009.

**V** J. Pätynen, B.F.G. Katz and T. Lokki. Investigations on the balloon as an impulse source. *Journal of Acoustical Society of America*, Vol. 129(1) pp. EL27-EL33, January 2011.

# Author’s Contribution

### Publication I: “A Loudspeaker Orchestra for Concert Hall Studies”

A loudspeaker orchestra for studying and comparing concert hall acoustics is introduced. This paper presents the design and implementation of a wide-area acoustic source comprised of a number of loudspeakers that are positioned in the shape resembling a symphony orchestra. Each instrument section of an orchestra is represented by two or three loudspeakers on the stage. With a fixed design, it is possible to build an artificial orchestra that performs identically in different halls. Anechoic orchestral recordings are used as the loudspeaker signals, enabling acoustic evaluation in-situ, and recording the performance for parallel comparison in laboratory conditions. Impulse responses measured from each loudspeaker can be analyzed, or later convolved with any anechoic signal. The loudspeaker orchestra has been utilized in a number of successful studies on concert hall acoustics.

The present author has contributed in the design and building of the loudspeaker orchestra. The shown acoustic measurements have been conducted by the present author. Approximately 80% of the article is written by the present author.

### Publication II: “Anechoic Recording System for Symphony Orchestra”

A method for conducting anechoic recording of a professional symphony orchestra is presented. This paper describes the design, building, and calibration of a system with the purpose of obtaining high-quality anechoic recordings of symphonic music. The instruments of a symphony orchestra are recorded individually in turns in an anechoic chamber. In order to achieve musical synchronization between the players, a reference video track was recorded, containing a conductor and a piano reduction of the orchestra score. The musicians could then follow the conductor with a small video display and listen to the piano track with headphones while playing in the anechoic chamber. The recordings have been highly beneficial for the subsequent acoustics research.

The author has contributed in the design and building of the recording system, and the implementation of the conductor video. Processing and editing of the recordings was conducted solely by the present author. Equalization filters and the conductor video concept were designed by Dr. T. Lokki. The actual recording was conducted by the author with the supervision by Dr. T. Lokki. Approximately 80% of the article is written by the author. Dr. V. Pulkki has provided suggestions for the recording system and the contents of the article.

### Publication III: "Directivities of Symphony Orchestra Instruments"

A thorough investigation on the directivities of symphony orchestra instruments is presented. The purpose was to study the instrument directivities as they appear on the stage of a concert hall. Importantly, the method of measurement kept unchanged during the course of measurements, and the methodology was accurately reported. A comprehensive set of anechoic samples were recorded with the most common orchestral instruments in an anechoic chamber. Twenty-two calibrated microphones were positioned around the musician, providing signals from an equally spaced grid. The results are presented in one-third octave accuracy with the foreseen utilization in auralization and acoustics modeling in mind. The results are compared with the previous studies of similar nature. Additionally, a visualization tool including the measurement dataset in Matlab environment has been made available for download.

The author is responsible for the analysis and the writing of the article. Visualization techniques were developed in collaboration with Dr. T. Lokki.

### Publication IV: "Directivities of Orchestra Instruments for Auralization"

The article presents a conversion process of the measured orchestra instrument directivities from arbitrary data format to CLF (Common Loudspeaker File) format, which is a *de-facto* standard for distributing directivity information in commercial acoustics simulation software. The publication discusses the limitations imposed by the directivity format developed for electroacoustic sources. The by-product of this study, a complete set of CLF format directivity files for orchestra instruments, has been made available for download. The publication also functions as a document for the provided directivity files.

The present author is the sole author of this article.

### Publication V: "Investigations on the balloon as an impulse source"

Investigation on the balloon bursts is presented. This article includes measured directivity data from various balloon types and sizes, and the results are compared to the omnidirectionality requirements imposed upon sound sources in the related standard [78]. The results on the directivity suggest that the balloon directivity is two-folded. First, the principal radiation peak is formed in the direction of the puncture. Second, at a formant-like peak, sound is radiated in the opposite direction. The magnitude and frequencies of the two peaks depend on the balloon size and the level of inflation. The main peak is found to follow the frequency of an adapted Helmholtz resonator. With regard to the ISO3382-1 standard on the directionality criteria, none of the balloon types could provide sufficient omnidirectional radiation suitable for acoustic measurements. Supporting video material is provided with the article, showing high-speed video sequences of the balloon burst process.

The present author has written 75% of the publication. Dr. B.F.G. Katz has participated in the directivity analysis and provided a portion of the code used in the implementation of the directivity visualization.

# List of Abbreviations

| | |
|---|---|
| 3D | Three-dimensional |
| AD | Analog-to-digital |
| CLF | Common loudspeaker format |
| DA | Digital-to-analog |
| DirAC | Directional audio coding |
| ERB | Equivalent rectangular bandwidth |
| FFT | Fast Fourier transform |
| IIR | Infinite impulse response |
| ISO | International Standard Organization |
| MLS | Maximum-length sequence |
| RMS | Root-mean-square |
| SIRR | Spatial impulse response rendering |
| STFT | Short-time Fourier transform |

# 1. Introduction

This thesis aims to improve the acoustics research by presenting new tools for room acoustic studies. The present topics include concert hall measurements, anechoic signals, and source directivity. The topics are closely related to auralization, which stands for the process of creating an audible rendering from a measurement of an existing space or an acoustical computer model, and presenting it as if the listener were listening to sound in the space in question [87].

Auralization can be divided into three principal components: the impulse response, the stimulus signal, and the spatial sound reproduction. The impulse response defines the acoustic characteristics of the space from the sound source to the receiver through a medium, that is, the acoustic properties of the room. It can be measured in-situ in the room under investigation, or calculated with computer models. In concert halls, a symphony orchestra is the typical source, and it is more complex than the sources utilized in the standardized measurements.

The signals applied to the impulse response are required not to include the acoustic effect of any room. That is, the signals have to be anechoic. In addition to a source comparable to an orchestra, the applied signals should contain orchestral music, which in sufficient quality is not trivially available. Regarding the auralization process, this thesis concentrates on the properties of the sound sources and the anechoic stimulus signals. The simulation of the impulse response or the spatial sound reproduction techniques are not included in the scope of the current thesis, but these aspects are briefly reviewed in the literature overview.

## 1.1 Scope of this thesis

The topics related in the scope of this thesis are visualized in Fig 1.1. The aspects in the evaluation of room acoustics and auralization are described in the following.

First, a method for room acoustic evaluation using an array of loudspeakers — a loudspeaker orchestra — is introduced. The acoustic properties of the loudspeaker orchestra are investigated with regard to measurements from authentic instruments.

Second, a detailed description of the anechoic symphony orchestra recordings is presented. The method consists of recording the instruments of a symphony orchestra individually in an anechoic chamber while the synchronization between musicians is assisted by a reference conductor video track. Anechoic recordings are utilized in connection with the loudspeaker orchestra.

Third, the directivities of symphony orchestra instruments are investigated based on the anechoic recordings. While this topic has earlier been under research in several occasions, the study presented in the current thesis contains a comprehensive selection of instruments, a detailed description of the methods, and a constant analysis method applied for all instruments. In addition, transforming the obtained directivity data into a de-facto distribution format is explained. A study of the directional characteristics of balloon bursts expands the scope of this thesis. The balloon burst directivity is evaluated with regard to the standardized requirements for sources in acoustic measurements.

Fourth, a novel method for improving the authenticity of the anechoic string instrument recordings is presented. Here the objective is to synthesize the sound of a string section by introducing variations exhibited by individual musicians to the anechoic signals. For obtaining the data used in the simulation, two different methods for tracking the playing of real orchestras are proposed.

## 1.2 Organization of this thesis

This thesis is organized to the following structure. Chapter 2 introduces the reader briefly to the central concepts in room acoustics and acoustic measurements with a short theoretical overview. The earlier research re-

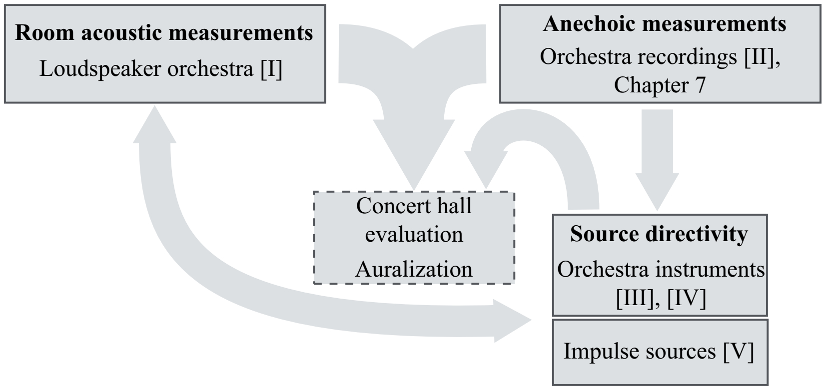



**Figure 1.1.** Scope of the thesis and the relation of the concepts. Included publications discussing the topics are indicated with Roman numerals in the figure.

lated to the topics are reviewed in Chapter 3. Chapter 4 presents the loudspeaker orchestra, which is a tool for evaluating and comparing concert hall acoustics, as proposed in Publication I. The anechoic recording method and the symphony orchestra recordings described in Publication II are introduced in Chapter 5. The results from the directivity measurements in Publications III, IV, and V are summarized in Chapter 6. Chapter 7 presents a method for creating a sound of an orchestra string section from a single recorded instrument. The results of the thesis are finally concluded in Chapter 8, accompanied by suggestions for research directions in the future.

# 2. Background

The topics of this thesis include anechoic measurements, directivity analysis, and their application in the evaluation of room acoustics and auralization. The following sections present the basic physical phenomena and background for the topics under discussion.

## 2.1 Sound sources

Sound is a pressure wave that propagates in a compressible medium. The motion of the particles in the medium is caused by a physical event. Examples of these events are a sudden expansion of material such as an explosion, an impact of objects, or a vibrational motion of a rigid body. The velocity of the propagating wave is determined by the properties of the current medium. The sound velocity in room temperature air is approximately 345 m/s [191, 192].

The harmonic pressure wave propagating along the $x$-axis is defined with the instantaneous pressure as the equation for a plane wave

$$c^2 \frac{\partial^2 p}{\partial x^2} = \frac{\partial^2 p}{\partial t^2} \quad , \tag{2.1}$$

where $c$ denotes sound velocity, $p$ pressure, and $t$ time instant [93, p. 9]. Real-valued solutions to the above equation with regard to $p$ are

$$p(x,t) = p\cos(\omega t - kx) \quad \text{or} \quad p\sin(\omega t - kx) \quad , \tag{2.2}$$

where $k = \omega/c$, $\omega$ denoting the angular frequency. Hence with harmonic motion, the pressure values are equal to wavelength $\lambda = 2\pi/k = 2\pi c/\omega = c/f$ intervals, where $f$ indicates the frequency of the propagating sound wave.

Sound intensity describes the amount of sound energy which passes through an unit area. Intensity is defined as the product of pressure and particle velocity $v = p/(\rho_0 c)$ averaged over time [44, p. 1054]:

$$I = \frac{1}{T}\int_T p(t)v(t)dt = \frac{1}{T}\int_T \frac{p^2(t)}{\rho_0 c}dt \quad , \tag{2.3}$$

where $\rho_0 c$ denotes the impedance of a medium as a product of density $\rho_0$ and sound velocity. This is true for distances that are large compared to the wavelength, where the complex part representing the phase difference between the particle velocity and pressure diminishes asymptotically to zero. The acoustic energy radiated from the source is defined as the intensity flowing through an enclosing surface $S$. Therefore, the acoustic power equals the energy radiated in 1 s time

$$P = \int_S IdS = \frac{1}{\rho_0 c}\int_S \overline{p^2}dS \quad , \tag{2.4}$$

where $\overline{p^2}$ denotes the time-averaged squared pressure. However, if the radius from the source is small, the acoustic power is defined with the volume particle velocity $Q(t)$ and its amplitude $\hat{Q}$ [93, p. 13]:

$$Q(t) = \frac{4\pi C}{jk\rho_0 c}e^{j\omega t} \quad \text{and} \tag{2.5}$$

$$\hat{Q} = \frac{4\pi|C|}{k\rho_0 c} \quad , \tag{2.6}$$

and combined with Eq. (2.4), the power of a point source is

$$P = \rho_0\frac{\hat{Q}^2\omega^2}{8\pi c} \quad . \tag{2.7}$$

The pressure wave emanating from such a point source of infinitesimal size propagates in a spherical pattern, i.e., omnidirectionally. With physically realizable sources of manageable size, the source shape and dimensions themselves become an obstruction and thus a significant factor regarding the pattern of the radiated sound wave. Generally, wavelengths in the scale that are comparable to the source dimensions are affected, and therefore the directivity with physical sources differs from the omnidirectional pattern at high frequencies.

In the physical measurements of acoustic power, surrounding the source completely with intensity probes or pressure sensors, i.e. microphones, is not feasible. For this reason it is suggested that the surface around the source is divided into smaller subsections which are each represented with one microphone at a standard radius. Thus,

$$\hat{P} \sim \frac{1}{N} \sum_{i=1}^{N} \overline{p_i^2} \quad , \tag{2.8}$$

where $i$ denotes one of the $N$ microphones.

Ideal omnidirecitonal sources, that is point sources, are often applied in simulation of sound fields due to their simplicity. In addition, the standardized method for room impulse response measurement requires an omnidirectional source [78]. Some impulsive devices provide repeatedly a directional pattern close to the omnidirectional radiation. With electro-acoustic measurement systems, such as loudspeakers, an exactly omnidirectional pattern cannot be produced at wide bandwidth even with specialized multi-element sources [24, 127]. Ordinary loudspeaker designs are even less omnidirectional, particularly at high frequencies.

Natural sound sources that appear in spaces with great acoustical importance, are musical instruments and the human voice. They exhibit more complicated directional properties, depending on the source type, and the directionality of the sources have a profound effect on the overall sound in a performance.

## 2.2 Room acoustics

The acoustic process is defined objectively by a measured or simulated room impulse response, which is the time-domain transfer function of the investigated space between an ideal source-receiver pair. For instance, music or speech being the source signal $x(t)$ ideally without background noise, the listener receives the convolution $y(t)$ with the impulse response $h(t)$:

$$y(t) = x(t) * h(t) \quad \text{in the time domain, or} \tag{2.9}$$

$$Y(f) = X(f) \cdot H(f) \quad \text{in the frequency domain.} \tag{2.10}$$

Depending on the temporal structure of the impulse response, usually it is roughly divided into three segments in time [44]. These segments are depicted in Fig. 2.1, showing a schematic time-energy diagram in a wide frequency band. The sound emanates from the source in the directions following its directivity pattern. First, sound traveling the shortest free path

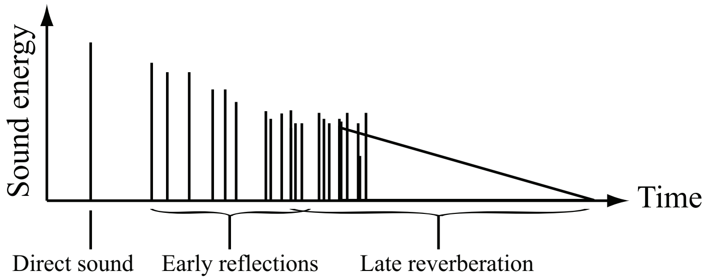



**Figure 2.1.** Time-energy representation of an example room impulse response and its division into three segments.

arrives at the receiver, delayed by the propagation time with the current speed of sound. Second, the sound emitted by the source is reflected from surfaces once or several times before arriving at the receiver. Time taken by the longer path, compared to the direct sound, causes a delay in the reflected sounds. The sound energy of the reflections decreases over time due to the originally emitted sound power being distributed spherically over a larger area (Eq. (2.4)). In addition, the materials on the reflecting surfaces cause losses in the reflected sound due to their absorption [60]. The early reflection part of the impulse response has been found to have a major significance to the perceived acoustical impression [7, 8, 14]. The third segment comprises again of sound reflections, but the increasing number of reflections arrive at the receiver at a rate where individual reflections are not be distinguished as such. This late reverberation stage of the impulse response can be approximated by a linear decay in the logarithmic magnitude scale [66, 159].

## 2.3 Acoustics modeling and auralization

Room acoustics modeling aims to provide an understanding of a current acoustical design and its properties. Usually, the objective of the modeling is to simulate the room impulse response accurately, for its most significant features. The simulated response is used for further estimating the overall acoustical quality or the room acoustic parameters [14, 58, 59, 78]. In addition to a monophonic response, spatial information in the impulse response can be stored binaurally [92, 122] or with spatial sound techniques [87, 140].

One of the most straightforward approaches to auditorium acoustics is to compare the propagating sound to light — as linear rays [142]. This paradigm is cited as geometrical room acoustic, and computationally, it is based on the assumption that sound travels along straight paths. Continuing with the analogy to optics, the sound is reflected from a modeled surface in the same manner as light bounces from a mirror [2, 19, 90]. There are several variations regarding the image-source method. One of the oldest is the ray-tracing method introduced earlier. Here, the source emits a large number of rays which are followed up to a certain reflection order or attenuation level until reaching the receiver [89, 91]. Regarding other geometrical modeling methods, element-based methods, such as the radiosity [163] and the acoustic radiance transfer [164] methods, can provide more efficient means for reflection modeling.

As stated earlier in Section 2.1 the relation of the wavelength to the dimensions of an arbitrary object is relevant regarding obstruction. Similarly, the assumption of the geometrical propagation is only valid when the wavelength is considerably smaller than the dimensions of the reflecting surface. With greater wavelengths, this approach cannot be used, as considerable diffusion and diffraction occurs due to the sound wave motion [35]. Similarly, curved and resonant structures pose problems to geometrical methods [87]. Geometrical methods can be modified to incorporate wave phenomena to some extent [95, 107]. Instead, methods based specifically on solving the wave equation, such as finite-difference time-domain, finite element methods, and boundary element methods provide accurate but a computationally heavier approach. A recent summary including novel techniques is given in [162].

Besides computational modeling a widely used method of designing and estimating room acoustics is the use of scale models. This approach is particularly in use in the field of concert hall acoustics design [39, 129, 130].

In visual design tasks, visualization is a tool for rendering an image of the computer model [54]. Analogously, the technique of rendering an audible version of an acoustic model is called auralization [87, 99, 179]. Usually the term is associated with an acoustic simulation using a computer model and its spatial reproduction with methods such as vector base amplitude panning [146] and Ambisonics [64] or binaural techniques [92, 122]. However, in the scope of this thesis the term refers to the process of convolving spatial impulse responses with anechoic signals (Eq. 2.9) [143].

# 3. Related Research

The previously published research related to the topics of this thesis are presented in this chapter. First, research on the room acoustics evaluation is reviewed followed by a discussion on room acoustic measurement methods. Then, the acoustics of sound sources characteristic of concert halls — musical instruments — are briefly presented with studies related to their directivities. Finally, an overview of the available anechoic recordings is made before concluding with a review on the auralization research.

## 3.1 Evaluation of room acoustics

In general, the evaluation and comparison of the acoustics in different performance spaces pose a multidimensional problem. First, the evaluation is at least to some extent a matter of subjective opinion and second, room acoustics has very little context. In comparison, music or speech contains some point of references. Speech, for one, is a means of exchanging information that can be repeated to someone else. Music contains a rhythm, melody, or harmony which can be sang or hummed after a concert. In contrast, acoustics have an enhancing effect on the ease of understanding or following the performance, but few acoustical features can be relayed or described accurately. [18]

Despite the difficulty of describing acoustics explicitly, opera and concert halls have been compared widely [14]. Traditional comparison methods include attending concerts and making written notes or completing a given questionnaire [71, 84], or interviewing specialized subjects [58, 59, 72].

Problems arise in comparative subjective evaluation with the standardization of the sound sources, signals, and listening environments. With subjective matters even the assessors' emotional state can influence the resulting evaluation, as comparable performances are rarely given on the same day. A reliable comparison of acoustics with authentic concerts is further complicated by the short duration of auditory memory [155]. Therefore, different approaches with recordings have been devised. It is possible to record an authentic or synthetic performance in a concert hall for later reproduction in laboratory conditions [158, 193], or to apply dry recordings to the concert hall measurements in a manner of auralization [29, 47, 63].

Larger arrays of loudspeakers have been mostly used in contemporary music performances, such as that described in [15]. Publication I presents the loudspeaker orchestra, which is used in the studies on concert hall acoustics. As a repeatable wide-spread sound source, the loudspeaker orchestra enhances the possibilities of comparing and analyzing room acoustics.

## 3.2 Sources and techniques for room acoustic measurements

The ISO3382-1 standard states that the sound source should be as omnidirectional as possible in measurement. Directivity variations in a plane are defined by the deviations of acoustic energy in a sliding average window over a 30-degree arc from the average energy over the full circle, while measurements are taken with 5 degree intervals [78]. The standard allows for $\pm$ 1 dB deviation at octave bands below 1 kHz. At higher octave bands the permitted directivity variation increases gradually up to $\pm$ 6 dB at the 4 kHz octave band. A sufficient sound pressure level is also required from a source complying to the standard. Loudspeakers designed for room acoustic measurements fulfill the requirements of omnidirectional radiation [24, 127]. On the other hand, ordinary loudspeakers have an increasing directivity toward the high frequencies.

Measurements with 5 degree rotational intervals do not pose problems with electro-acoustic sources, as the tested device on a turntable is easily driven with a repeatable measurement signal, and the produced sound is recorded with a microphone in a constant position. However, when measuring natural sources requiring human interaction, good repeatability is uncertain.

### 3.2.1 Impulsive sources

A number of natural devices have been applied in room acoustic measurements in order to produce an approximation of the Dirac delta function. Such impulsive sources include pistol shots, explosives, specific impact devices, or balloon bursts. In measurements with a scale model, sparks generated with a special electric device are not uncommon. The advantages in using impulsive devices are the low cost, high portability, and a high signal-to-noise ratio, especially in the case of a starting pistol. Research related to such impulsive sources are reviewed next.

Regarding pistol shots, Lamothe and Bradley have measured five pistols of different types from various angles [96] and compared those with a high-voltage spark source and a loudspeaker. They concluded that a .38-caliber pistol produces a repeatable signal with a high sound pressure level. Directional variation was the lowest with a large-caliber pistol, and in comparison to a spark device, pistol shots were found considerably more effective in exciting the low frequency range. The pistol frequency response was noticed to be far flatter at the high octave bands compared with a loudspeaker. In addition, Bradley has investigated the reverberation times measured with pistol shots [21]. More recently, Sumarac-Pavlovic *et al.* [167] developed and studied a wooden clapper as a portable impulse source, finding that it was more directive at higher frequencies, meeting the ISO3382-1 standard.

Balloon ruptures as impulse sources have been studied more vigorously. The resulting waveform is studied in [41]. Griesinger [67], Nash [125], Horvat *et al.* [75], and Chéenne *et al.* [28] have all presented studies citing directional variations of large magnitudes at low frequencies, good repeatability above the 100 Hz one-third octave band, and the larger balloons containing more low frequency energy. Chéenne *et al.* studied the spectral responses of balloons of various diameters and inflation pressures [28], and in contrast to the theory presented in [125], the spectral responses were found relatively flat. Relations between the sound pressure level and spectral centroid as a function of balloon diameter were proposed although details of the measurement protocol and possible room effect are absent. Recently, the room impulse response excited with a balloon burst has been simulated in [1].

In addition to balloons, Horvat *et al.* have found that small-scale explosives, namely firecrackers, produce a sufficiently strong impulse. Larger firecrackers provide a better response at a low frequency range than smaller ones [75].

Measurements with impulsive sources have been presented by Jambrosic [81], Fausti and Farina [49], and James [82]. These studies have compared the reverberation times of various rooms measured with different source types. In this regard balloon pops were noticed to provide results that are highly comparable to more elaborate methods.

### 3.2.2 Electro-acoustic measurement methods

Several measurement methods employ loudspeakers as sources for a measurement signal. Together with receiver microphones, they can both be easily driven by an ordinary computer and analogue-digital/digital-analogue converters for accurate control and quick assessment of the measurement results. Overviews on the measurement methods with electro-acoustic signals and their properties have been presented in [139, 166]. In the following, the most frequently used methods are briefly discussed.

A very simple method of studying the sound decay in rooms is to use a noise signal which is ended abruptly [33]. With a sufficiently long build-up time, the room response after the signal has been stopped provides the decay curve at the excited frequencies. This technique closely resembles the method first used in estimating the reverberation time [153]. While suitable for analyzing the sound decay in rooms, this approach is not applicable to the measurement of the actual room impulse response.

Maximum-length sequence (MLS) is based on a deterministic, periodic pseudorandom time-series signal with the spectral properties of the white noise [20, 149, 160]. The recorded measurement signal is cross-correlated with the stimulus signal in order to obtain the room impulse response. Benefits of the MLS technique include the relative immunity to the background noise during measurement. With the ability to average multiple measurements, in theory, it is possible to conduct measurements even during performances [160]. However in practice, time-variance in the measured space over long measurements and the loudspeaker distortion limit such measurements.

Currently, a commonly used measurement technique employs swept sinusoids. The room response is obtained from the convolution of the recorded measurement signal and the carefully designed inverse signal. In the measurement, only one frequency is excited at a single time instant. Therefore the sine sweep method is robust against the harmonic distortion compared with MLS. The properties and advantages of the swept-sine technique are discussed in detail in [46].

## 3.3 Orchestra instrument acoustics

Publications II and III are closely related to orchestra instruments. In the following sections an overview on the instrument research is given. The physics of the instruments is important in understanding their directional behavior. The major instrument groups are discussed in an approximate order of complexity in their sound production mechanism. A comprehensive overview concerning all orchestra instruments is presented in [119].

### 3.3.1 Brass instruments

The sound generation of the brass instruments is simple with regard to many other instrument types. Oscillations of the air column in the instrument are induced by the lip vibration in the mouthpiece. The vibration is then amplified by the tube which is nearly closed at the mouthpiece. This results in a series of odd harmonic frequencies. However, the shape of the bell, the mouthpiece, and the pipe bore lower the resonance frequencies to an approximately complete harmonic series [52, 152]. The flared bell increases the efficiency of the sound radiation and the directivity.

The effective length of the pipe is altered by valves that connect extensions to the piping. The trombone employs a slide for changing the pipe length, which obviously requires a cylindrical bore for the slide section. Also the bore of the trumpet is mainly cylindrical. The tuba is conical, while the French horn has a small cylindrical portion [52, 168].

A typical length for the F-tuned French horn is approximately 3.75 m [152]. The trumpet has the shortest length of tubing of the common brass instruments, and the total length is approximately 1.4 m. Equipped with three valves, this results in a playing range of approximately three octaves [3, 152]. The overall tenor trombone length of 2.75 m is twice that of the trumpet, which yields a playing range of an octave lower. Some trombones incorporate a separate valve for connecting additional length to the tubing, providing a lower playing range. The tuba has the lowest playing range of all brass instruments, and its total length is over 5 m

[152].

Research on the directivity of the brass instruments with documented measurements has been published in [132] on the French horn and the trumpet on selected tones. The Cornet and French horn measurements in the horizontal plane are presented in [115]. Loudspeaker-driven trombone measurements have been performed in [183].

### 3.3.2 String instruments

The bowed string instruments form the most substantial portion of a symphony orchestra. Sound is generated by a complex mechanism, consisting of an alternating action of sticking and slipping of the bow on the string. The bridge couples the excited string vibration to the body of the instrument. The vibrating top and back plates have their own set of modal frequencies that change between individual instruments. The side walls and the vertical sound post under the bridge bind the plates together. Moreover, the bass bar increases the top plate strength against the string tension. [36, 52, 152, 170]

The four strings of the violin are tuned with intervals of perfect fifths, beginning from G3 (approximately 196 Hz). Tones around A7 can be played ordinarily, and even higher tones are playable through harmonics. The viola has been stated to have 15% greater dimensions compared to the violins [52]. The strings of the viola are tuned a perfect fifth lower, thus at approximately one third lower frequencies than those of the violin. However, the body resonances are not scaled correspondingly. This results in the viola having a characteristically different tone to the violin [152].

The strings of the cello are tuned an octave lower than the viola, that is, 30% lower than the violin. Overall the frequencies of the vibrational modes are reported being approximately 40% of the corresponding violin modes. Thus, the modes are slightly higher with regard to the string tuning than with the violin [52]. Contrabass, or double bass, differs from the rest of the string family with its flat back and carved shoulders. Typically, the four or five strings are tuned in perfect fourths instead of fifths, up from E1 or B0 respectively. Varying configurations exist but they are rare.

The absence of particular shapes directing the sound as in brass instruments yields considerably more complicated directivity patterns. The violin physics in particular has been under research in a number of studies [16, 17, 36, 77, 83, 114, 156, 181, 185, 186]. The present author is aware only of scarce research on the viola compared to the violin, despite the viola being structurally close to the violin. Research on the cello acoustics is also scarcely found. Input admittance and resonance measurements on different cellos and bass have been presented in [4, 23]. Only limited research concentrating on the contrabass is published.

### 3.3.3 Woodwind instruments

The woodwind group is different from other instrument types in many respects and the mechanism of sound production varies between woodwind instruments. In flutes, an air jet hitting a sharp edge creates oscillations in the pipe. The clarinet uses a single vibrating reed, while the oboe and the bassoon have a double reed to regulate the air flow and produce sound [5]. In each case, the pitch is altered by changing the effective length of the pipe by opening and closing tone holes.

The modern flute is a cylindrical pipe of 66 cm length with open ends. The piccolo is approximately half of the length of the flute, hence, having the range of one octave higher. The sound is produced by forming a Helmholtz resonator between the air jet in the embouchure hole and the closed end of the pipe [165]. A comprehensive study of the flute physics has been presented in [169, 171, 190]. Flute performance technique with dynamics, harmonic structure and blowing pressure have been studied in [51].

The sound of the oboe is produced by a pair of vibrating reeds which causes the air column to oscillate in the pipe. The oboe forms a pipe with one closed end at the mouthpiece, and thus it works as a quarter-wave resonator. As with the cylindrical flute with open ends, the closed conical pipe of the oboe creates a complete harmonic overtone series [152]. The bassoon has a double reed similar to the oboe. The length of the folded pipe of approx. 2.6 m provides the lowest playing range of the woodwinds. The smaller angle of the bore, the metal tube connecting the reed to the pipe, and the long finger hole chimneys in the lower joint are considered to produce the distinct bassoon sound [52, 161]. In contrast, the shape of the clarinet is cylindrical, therefore the even harmonics are attenuated in the low register.

The flute mouthpiece is a considerable source of sound radiation unlike with reed instruments [119]. The far end of the flute, as well as open finger holes, function as radiation sources which are in phase at odd harmonics and in the opposite phase at even harmonics when all finger holes are closed. The open finger holes present considerable radiation at middle frequencies, which makes the total radiation complex [52]. Also with reed instruments, the finger holes and the open end radiate sound. A cutoff frequency is used to roughly define the frequency above which the sound is mostly radiated from the open end [26]. For the oboe and the clarinet a cutoff frequency of 1500 Hz has been reported [12]. The clarinet directivity has been stated to be similar to the oboe below 2000 Hz [119]. The reported bassoon cutoff frequency is considerably lower, around 400–500 Hz [52]. In addition, noticeable formants have been found at 440–500 and 1220–1280 Hz with the bassoon [98].

## 3.4 Instrument directivity

Various approaches to directivity investigations can be found in the literature. First, a number of nationally standardized scalar parameters for indicating directivity have been reviewed in [65]. Statistical directivity factor $Q$ defines a radiation pattern that indicates the relation of the radiated area to the surface of a sphere, e.g., $Q$ for source radiating to a single quadrant equals 4. Directivity factor $\Gamma$ indicates the relative sound pressure between an arbitrary spherical angle pair and a reference direction $\Gamma = p(\theta, \omega)/p_{ref}$. The front-to-random factor $\gamma$ indicates the relation between a reference direction sound pressure and the sound pressure integrated over the measured angles (Eq. (2.8)). The statistical directivity factor, directivity factor, and front-to-random factor values can be expressed in $10\log_{10}$ scale, giving directivity index, directional gain, and front-to-random gain, respectively [65, 79]. These parameters are more often used in describing loudspeaker properties rather than with musical instruments.

The ISO3745 standard [79] defines microphone positions for sound power measurements in fully anechoic and hemi-anechoic environments. A 3-D grid in free field measurements is defined with 20 microphones positioned in a spiral-like formation where the spherical surface surrounding the source is divided into 20 equal areas. However, the symmetry and the equal areas are not intuitively perceived. According to the ISO3382-1 standard [78], the measurement source omnidirectionality in a 2-D plane is confirmed with measurements in 15-degree intervals around the source.

A typical approach to directivity measurements is to use a microphone array distributed around the source or the musician. Singer directivity has been measured with horizontal and vertical microphone arrays [113]. A similar microphone configuration has also been applied in [132, 133].

A large spherical array has recently been applied in [144], where the directivities were measured by playing the instruments inside a grid consisting of 32 microphones. The grid diameter was approximately 4.2 m. Even a denser array with 64 microphones have been constructed for instrument radiation measurements in [74]. Those measurements have been further applied in a system for obtaining the audio signal in arbitrary directions by interpolating the directivity [124]. However, the reported measurement radius with this array was only 1.2 m.

Recently, a recording system with 26 microphones has been applied for capturing the directional sound radiation during violin and guitar playing [112]. The article presents limited directivity analysis for the violin. In addition, the reproduction of the captured directional sound was provided with a custom-built spherical loudspeaker having its elements at the positions corresponding with the recording microphone array.

For musical instrument sound radiation measurements, a reciprocal method of exciting an instrument body externally with a loudspeaker and measuring the response from the violin has been used with violins in [180, 186, 187]. MLS signals have been applied in violin directivity analysis by exciting the violin from the bridge and measuring the response with a microphone in an anechoic chamber [48]. Here, the direct and reciprocal methods are also compared. Such methods enable a high spatial resolution for instrument studies. In addition, the absence of human interaction ensures good repeatability. Therefore the requirements for the amount of equipment are lower than in one-off recordings. Consequently, using electro-acoustic excitation does not represent an authentic performance situation where the player itself has an effect on the sound radiation.

The most widely known series of studies on orchestra instrument directivities were made several decades ago. The results on directivity and tonal analysis are given in detail in [120], but in practice the accurate description of the original measurement setup is not available. A rather large measurement distance of 3.5 m is mentioned in [120]. Given the spatial accuracy of the presented results, the measurements could possibly have been conducted with artificially excited instruments. In addition, the directivity data has been criticized for presenting averaged results over the playing range of the studied instruments [132]. These results are also available in a database for further use [141].

Publication III presents the directivities of the common symphony orchestra instruments measured with a dodecahedron microphone array, and using the same measurement and analysis method for all instruments.

## 3.5 Anechoic Signals

For auralizations of good quality the anechoic signals used for convolution need also to be of high standard. A number of various anechoic recording projects have been undertaken in the past. Regarding the required quality, Buen has discussed the degree of anechoic conditions needed for such recordings [25]. Here the presence of the room in the nearly-anechoic recording has been studied only from the time-domain perspective, i.e., how much the reverberation changes if the recording condition is not entirely anechoic. However, with non-anechoic recording conditions, reflections in the room introduce changes in the captured frequency response of the original source. Hence, excess coloration in the sound may occur when convolved with an impulse response of a small room. In addition, the article states that much of the information in the source signal is lost in anechoic recording conditions. In contrast, if the simulated space is large, the sound level of the direct sound is considerably higher than of the first reflections. Therefore the signal in the direct sound should represent the actual direct sound as well as possible without the effect of the room. Consequently, signals used in auralization and in comparable purposes should be recorded in conditions that can be classified as anechoic [79].

Most of the anechoic recordings have been done commercially. Recordings have been published by Denon [42, 73] and Bang & Olufsen [69, 6]. The former contains a full orchestra and the recording has been made inside an absorbing shell built on a concert hall stage. Therefore it is not entirely comparable to anechoic recordings. The entire orchestra has been recorded at once, which has been stated as not providing sufficient channel separation [106]. Therefore the utility of these recordings for auralization is considered poor. In the latter, individual instruments were recorded in a large anechoic chamber with a single microphone [69]. Moreover, some instruments were additionally recorded in slightly reverberant conditions. Also the equipment, including music stands and notes, were omitted from the recording space. Importantly, it is reported that the musicians recorded in [69] were made aware of the anechoic environment in order to avoid an unnaturally forced tone in their playing.

A commercial, yet available free of charge [55], choral recording has been published by Wenger Corp. [189]. The whole choir was recorded at once in an anechoic chamber equipped with various microphone types and configurations. Finally, there is a varying amount of proprietary instrument recordings for commercial audio production purposes (e.g. [172]).

Only scarce reports of anechoic recordings for research purposes exist. A full symphony orchestra recording with individually recorded parts has been reported in [173], using the methodology presented in [132]. These recordings consist of two excerpts from Mozart's and Brahms' symphonies. Multiple string instrument players were recorded, each playing every part for the corresponding instrument. From all obtained takes, the best takes were selected to represent each part. The recording approach applied to wind instrument parts is not explained in detail [176]. These recordings are not available for free distribution due to copyright restrictions, however, they have been applied in a number of auralization studies [100, 174, 182].

### 3.5.1 Sound of an instrument section and application of the anechoic recordings

A typical symphony orchestra consists of strings, woodwind, and brass instrument groups in addition to percussions. The portion of string instrument players can be over half of the total number of musicians in symphony orchestras. In classical compositions, a string instrument section usually plays in unison, that is, having the same part for all first violins, for example. Hence, the musical tone from a group of players playing in unison is very different to that of a single musician.

The characteristic sound of an instrument section results from the differences in playing technique, individual instruments, and the acoustic conditions. In [43, 116] it is stated that a unison ensemble has a characteristic "very slight pitch, amplitude, and timing randomness among the players." Meyer has stated that the sound characteristic for a section is caused by the broadening of the peaks at harmonic frequencies [120]. The intonation, i.e. nominal pitch of the played notes, is different. With instrumental ensembles, the 3 dB bandwidth of the spectral peaks deviates up to $\pm 20$ cents from the nominal frequencies. The frequency modulation of a violin vibrato has been found to exhibit a frequency range of $\pm 15.2$ cents [117]. Also, individual string instruments exhibit perceivable differences in their frequency response and resonance properties [56].

In commercial recordings a group of violinists can be easily recorded at once in a studio, but with anechoic recordings such an arrangement is not feasible ([173], Publication II). When the instruments are recorded individually, it would be highly time-consuming to record the same part with a large number of individual musicians. For practical reasons, only few musicians are recorded playing every part written for the current instrument. This leads consequently to the lack of the chorus effect due to the use of the same instrument and similar interpretation. Hence, it is beneficial to apply specific signal processing methods in order to improve the impression of a larger group of players in the recordings.

An audio effect, *chorus*, is used as a de-facto standard in audio industry when a single instrument is needed to sound more like an ensemble [194]. The chorus effect is based on a delay line whose tap point is modulated over time, causing variation in tempo and pitch [40]. Multiple delays can be applied for the corresponding number of simulated players. Simulation experiments related to alternative approaches to the chorus effect have been discussed in [85].

In the simulation of a symphony orchestra, constant delays up to 23 ms have been used with a small number of anechoic string instrument tracks [173]. Such an approach does not contribute to the pitch nor amplitude variation. In addition, constant delays between identical signals are manifested in a series of comb filters. The quality of this method has not been evaluated in [173]. Simulating the chorus effect of an ensemble with the phase-synchronous overlap-add algorithm has been studied in [100], where it was stated that the applied processing method gave inconclusive results for the perceived number of players. In addition, the subjective preference varied widely.

Overall, the problem of having a plausible sound of the strings is important for orchestra auralizations of good quality. Chapter 7 concentrates on the discussion on a novel method for simulating the sound of a string section from the recordings reviewed in Publication II.

### 3.6 Auralization

Although auralization is not as such in the direct scope of the present thesis, a majority of the presented research is aimed also at auralization purposes. Publication IV presents a method for converting measured orchestra instrument directivities into a de-facto data format for convenient application in acoustic models with widely used commercial software [38, 128].

Auralization techniques have been under study of varying intensity. A generic overview on the auralization is presented in [87]. Simulation results have been compared with a scale model measurement in [88], and the audible effect of changing acoustic features in rooms was studied with auralization in [37].

A method for simulating directivities in auralization with radiation cones is introduced in [133]. The radiation from a point source was divided into a number of regions, from which the corresponding recorded microphone signal was reproduced. Odeon software [128] was used for simulations in a listening experiment, in which solo clarinet stimuli were presented. Results suggested an improvement in the perceived naturalness in comparison to a monophonic recording. This approach has been further refined and studied in a series of publications by applying varying directivity patterns, source configurations, and hall models [173, 174, 175, 182].

In auralizations with a full orchestra, many sources are needed for the physical authenticity. In most cases the sound sources in the model are defined manually. An increased number of sources yields also an increased computational load. Hence, it is beneficial to use a number of sources that optimizes computation time without sacrificing the quality of the auralization. The effect of the number of the sources on the resulting auralization with a symphony orchestra is investigated in [100]. Different aspects of the methods for evaluating the overall quality of auralization are discussed in [105], whereas a more complete view on the auralization chain is presented in [99, 150].

Several studies have compared in-situ recordings with auralizations using binaural reproduction. Two concert halls were modeled with Odeon in several positions and compared to omnidirectional measurements in [31, 32]. A similar approach was utilized in [188] with the acoustics of churches. Furthermore, two simulation softwares were compared with an in-situ recording of a small room [154]. In [99], a small room was modeled using a custom simulation.

Despite the popularity of binaural listening, the method has a few drawbacks in comparison to ideal spatial reproduction. First, the equalization of headphone responses is required for improved localization and externalization [68]. Second, a common limitation with simulated or measured binaural responses and their convolution is the lack of adaptation to head movements [10]. Hence, the surrounding sound is fixed to the coordinates of the head instead of the surrounding space, reducing the ease of localization. This can be overcome with head tracking techniques [10, 94].

## 3.7 Summary of the related research

The related work in the scope of this thesis has been reviewed. The subjective evaluation of concert hall acoustics in laboratory conditions has been conducted mostly with binaural renderings on room impulse responses. Most importantly, the number of sources used in the measurements has been small, and omnidirectional by their radiation pattern. Such sources do not represent an orchestra which is the characteristic sound source in concert halls. In auralizations, sources with correct directivity are not overly complicated to simulate, and experiments on orchestra auralizations have been published. On the other hand, the auralization quality appears to have been evaluated exclusively with simple source configurations in concert halls or small rooms.

Several acoustic measurement techniques for determining the room impulse response, and further estimating the acoustic parameters, are in use. The traditional methods utilizing impulsive sources have been superseded by electro-acoustic methods relying on synthetic signals repeated with loudspeakers. The research on impulsive sources has mainly concentrated more on the spectral properties, and less on the directivity. On the other hand, studies on loudspeaker directivities have been rare, as the measuring loudspeakers are designed to fulfill the required standards.

Musical instruments and their directivities have been under study already for several decades. Similarly to anechoic recordings, well-documented directivity measurements and results presented in an inter-comparable manner are scarcely found.

Anechoic recordings form an essential part in the auralization chain. However, previous recordings have included only individual instruments, inadequate recording conditions, or they are not freely available. Also the processing and enhancement of anechoic recordings for improved section-like sound has not been extensively studied since the need has been rather small. However, with large-scale simulations, such as with a symphony orchestra, incorporating novel methods are essential for authentic results.

# 4. A loudspeaker orchestra for studies on concert hall acoustics

The acoustics in performance spaces is traditionally evaluated subjectively by listening to a concert, writing notes, organizing interviews, or collecting questionnaires regarding the acoustic impression [14, 58, 59, 158, 84], or objectively by conducting measurements [78]. Standardized acoustic measurements yield quantifiable parameters for various aspects, such as the reverberation time, strength, early decay time, clarity, and lateral energy fraction. Such energy-based parameters provide descriptors for the decaying sound, or the amount and direction of sound arriving at the receiver.

A major challenge in the subjective evaluation with live concerts is that the comparison between halls depends on the human memory on acoustics, which can only partially be improved by writing down notes. In addition, the musical interpretation and playing technique of the professional performers vary due to their ability to adapt to different acoustic conditions. On this basis, comparing live concerts, i.e. on consecutive days, is not an entirely reliable method of collecting the subjective differences between the acoustics in concert halls.

This challenge can be overcome, first, by recording reproduced sounds in the compared halls, or second, by auralization with measured impulse responses. Earlier, room impulse responses have been recorded in concert halls using a pair of sources. A convolution with anechoic signals enabled a comparison in laboratory conditions [158]. However, such a simple source configuration is hardly comparable to the characteristic source in concert halls — a symphony orchestra. Related studies on the sufficient number and configuration of sources for plausible acoustic simulations have been conducted in [100, 173]. Some references of larger-scale loudspeaker setups can be found [15], but they have mostly been applied to contemporary art performances [22, 70, 121]. A system of these magnitudes has not been reported with scientific acoustics research.

The loudspeaker orchestra is presented in Publication I. The basis of the orchestra consists of 24 loudspeakers, each in their individual signal channels. The loudspeakers are positioned on the stage in a shape derived from the American seating arrangement [120]. The use for the calibrated loudspeaker orchestra is two-folded. First, room impulse responses can be measured from individual channels in concert halls and stored for a later convolution with anechoic material. Second, anechoic music can be played back and recorded in-situ for subjective comparison. Naturally, in-situ listening is enabled by an orchestra immune to distractions and playing at will.

Techniques such as the spatial impulse response rendering (SIRR) [118, 147] can be applied to the measured responses for reproduction with arbitrary listening setup. In-situ recordings can be reproduced with ordinary (e.g. binaural head, mono, stereo) or spatial techniques (e.g. DirAC or ambisonics) depending on the microphone configuration [9, 64, 177]

Although the room acoustic parameters can be calculated from the room impulse responses obtained with the loudspeaker orchestra, the directivities of the two-way loudspeakers do not comply with the ISO3382-1 standard as an omnidirectional measurement source [78]. On the other hand, the directivities of orchestra instruments cannot either be regarded as omnidirectional (Publication III). The issues related to directivity are addressed in the following sections.

## 4.1 Loudspeaker orchestra configuration

In the proposed loudspeaker orchestra, a small number of loudspeakers represent the instrument sections of a symphony orchestra. The original layout consists of the loudspeakers marked with numbers 1 through 24 in Fig. 4.1. Loudspeakers 1-3, 4-6, 7-9, 10-12, and 13-14 represent the string instrument sections: I violins, II violins, violas, violoncellos, and double basses, respectively. Woodwinds (flutes, oboes, clarinets, and bassoons) are represented by loudspeakers 15-18 in the center. Loudspeakers 19-20 are dedicated to the French horns, while the furthest row of loudspeakers represent the trumpets (22, 21), trombones (23), and the tuba and the timpani (24). Alternatively, the channel no. 24 has been connected to a loudspeaker representing a soprano soloist in the front of the orchestra.

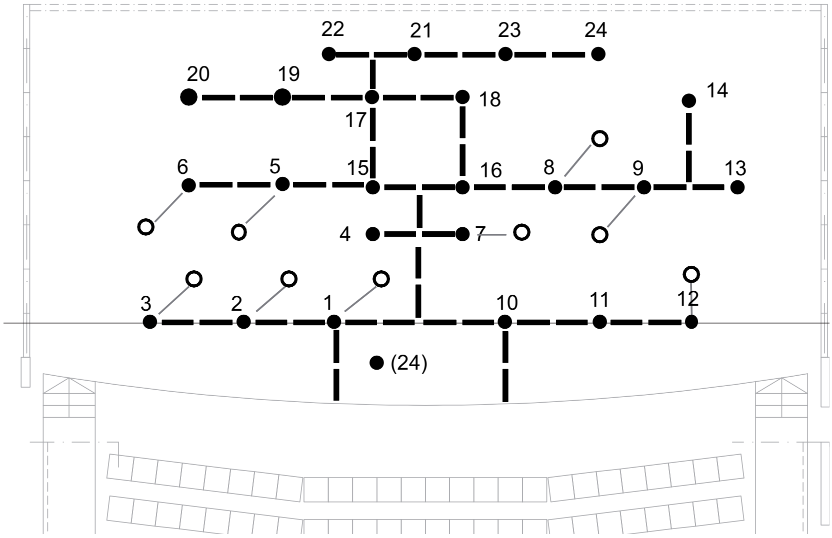



**Figure 4.1.** Plan of the loudspeaker orchestra on a concert hall stage. Each of the thick bars indicate 1 m distance. Hollow symbols indicate a secondary loudspeaker pointing upwards. Channel no. 24 is used alternatively for the tuba and the timpani, or the soprano soloist shown in parentheses.

The position for the soprano source is numbered in parentheses in Fig. 4.1.

The number of channels is limited by the sensible amount of equipment, as 24 channels can currently be ran with a modern laptop computer and a set of AD/DA converters for playback and recording. The described setup is highly portable and it can be fully prepared for use in approximately two hours. In theory, a very high number of channels could be applied, but the time required for the setup would increase correspondingly.

Two-way loudspeakers of traditional design are employed (i.e. Genelec active loudspeakers [61, 62]). In addition to 24 principal loudspeakers, auxiliary loudspeakers are utilized in selected string instrument channels. The purpose of using a combination of two loudspeakers in one signal channel is to improve the high-frequency radiation. Reproduction techniques for measured directivity patterns with special loudspeakers are discussed in the literature [112, 145, 183, 184]. While such sources imitate real instrument directivities in a nearly authentic manner, applying over twenty units of custom-built multi-element loudspeakers is not feasible.

## 4.2 Objective analysis of the loudspeaker orchestra

In the applications of the loudspeaker orchestra anechoic music samples are played back from the loudspeakers. This is achieved either by convolving a measured impulse response with the anechoic samples or by directly driving the loudspeakers with the signal in-situ. In both cases the overall radiated sound energy at different frequencies, i.e. the power response, is determined by the product of the anechoic signal frequency response $X_{\text{inst}}(f)$ and the power response of the loudspeaker $P_{\text{lps}}(f)$. This is analogous to a real instrument radiating sound with certain directivity, albeit more complex than that of a loudspeaker.

One method for evaluating the differences between the loudspeaker orchestra and the real orchestra instruments objectively is the comparison of the resulting power responses. Here, the reproduced power response by the loudspeaker driven with an anechoic instrument signal is subtracted from the measured power response of the instrument $P_{\text{inst}}(f)$. Such comparisons are presented in Figs. 4.2 – 4.5 with the common orchestra instruments. The upper subfigures present two curves. The first curve in thin line shows the measured power response of the instrument calculated with Eq. (2.8). The second, thick, curve represents the frequency response of the instrument in the direction that is used for the loudspeaker reproduction $X_{\text{inst}}^{(\hat{\theta},\hat{\phi})}(f)$. The directions $(\hat{\theta},\hat{\phi})$ of the instrument frequency responses are chosen by a magnitude-weighted least-squares optimization [86], i.e.

$$(\hat{\theta},\hat{\phi}) = \arg\min_{\theta,\phi} \left\{ \int_{-\infty}^{\infty} w(f) \left[ P_{\text{inst}}(f) - X_{\text{inst}}^{(\theta,\phi)}(f) P_{\text{lps}}(f) \right]^2 df \right\}, \quad \text{where} \tag{4.1}$$

$$w(f) = P_{\text{inst}}(f) / \max(P_{\text{inst}}(f)), \quad 0 \leq w(f) \leq 1, \forall f \tag{4.2}$$

and the subjective evaluation of the signal quality in the applied directions. Most importantly, the solid curve in the lower subfigures shows the difference between the power responses of the actual instrument and the loudspeaker driven with the anechoic signal. Ideally, the curve should be flat. That is, the average spectrum of the sound radiated by a loudspeaker driven with a recorded signal equals the average spectrum the actual instrument. The dashed line shows the loudspeaker power response.

In general, the French horn and the tuba exhibit the largest differ-

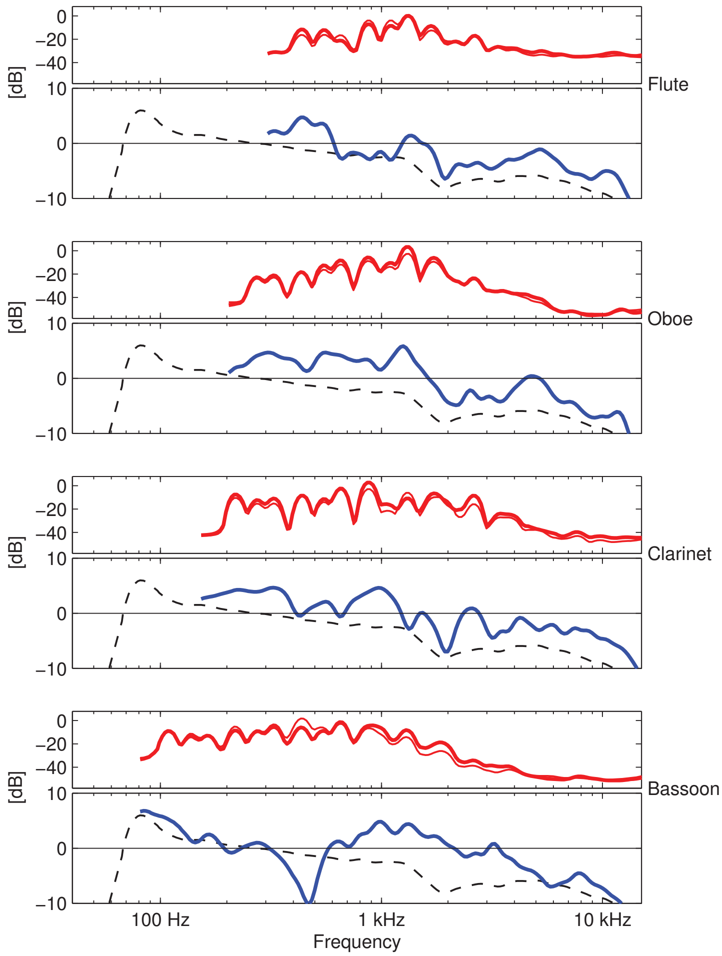



**Figure 4.2.** Power responses for the measured woodwind instruments and a loudspeaker driven with the instrument signal from one direction. Top figures: Comparison of the instrument average power response (thin line) and frequency response at the optimal direction for reproduction (thick line). Lower figures: Power response difference between a real instrument and a Genelec 1029A loudspeaker reproducing the instrument signal recorded in the optimal direction (solid line) [61]. Values below 0 dB indicate less reproduced power in loudspeaker playback. The loudspeaker power response is shown with dashed line. All curves are one-third octave smoothed.

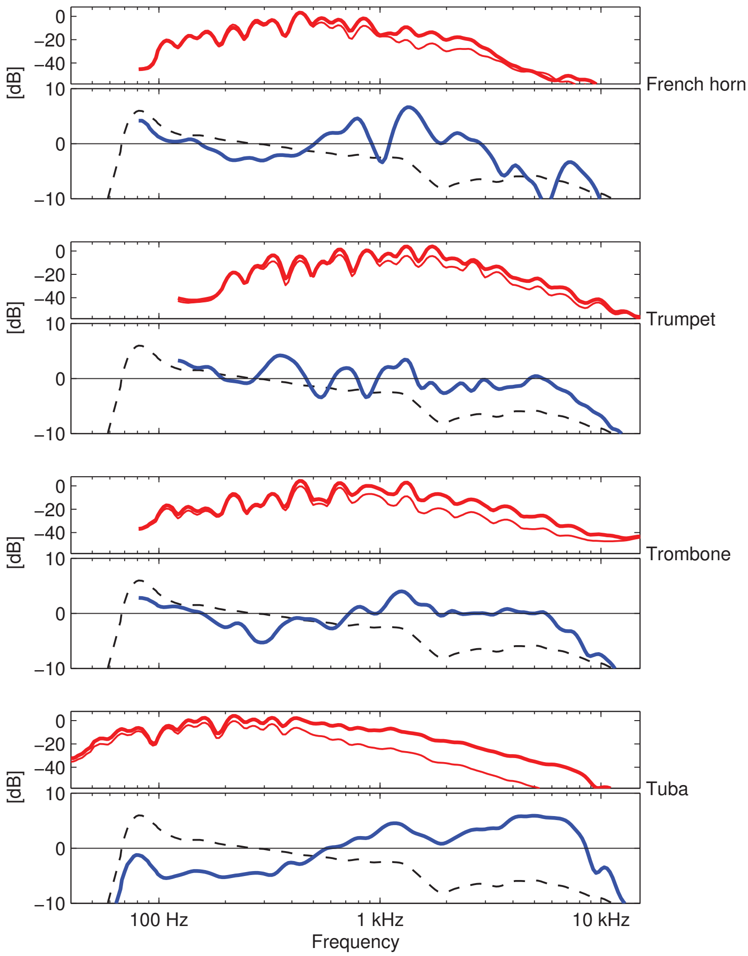



**Figure 4.3.** Comparison of the power responses for the brass instruments to the loudspeaker reproduction. The visualization is similar to Fig. 4.2

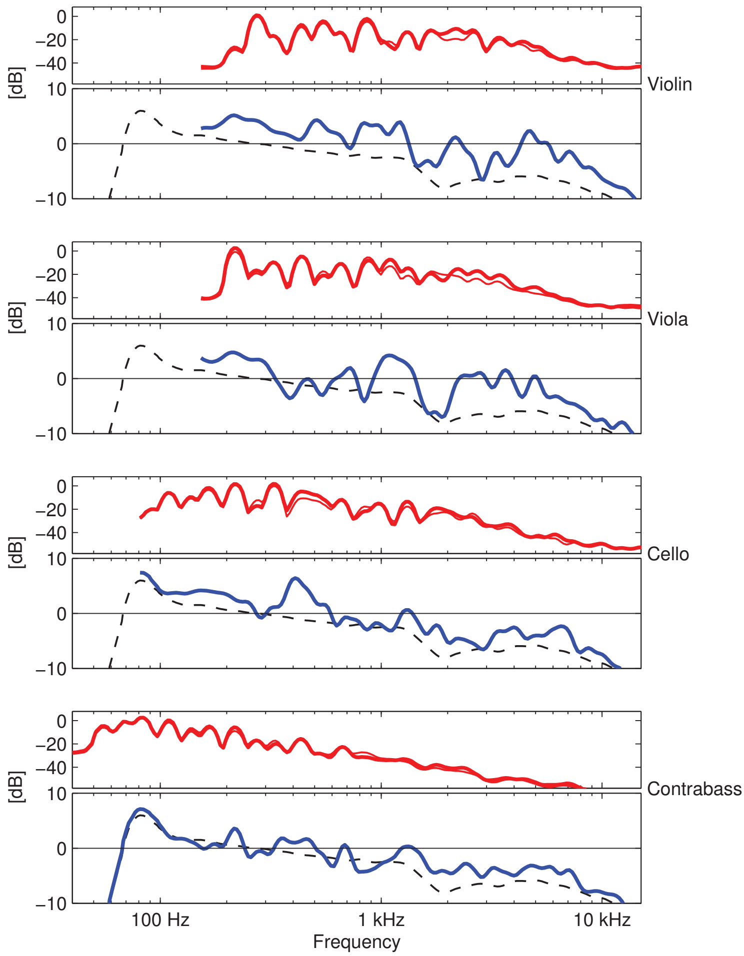



**Figure 4.4.** Comparison of the string instrument power responses to the loudspeaker reproduction. The visualization is similar to Fig. 4.2.

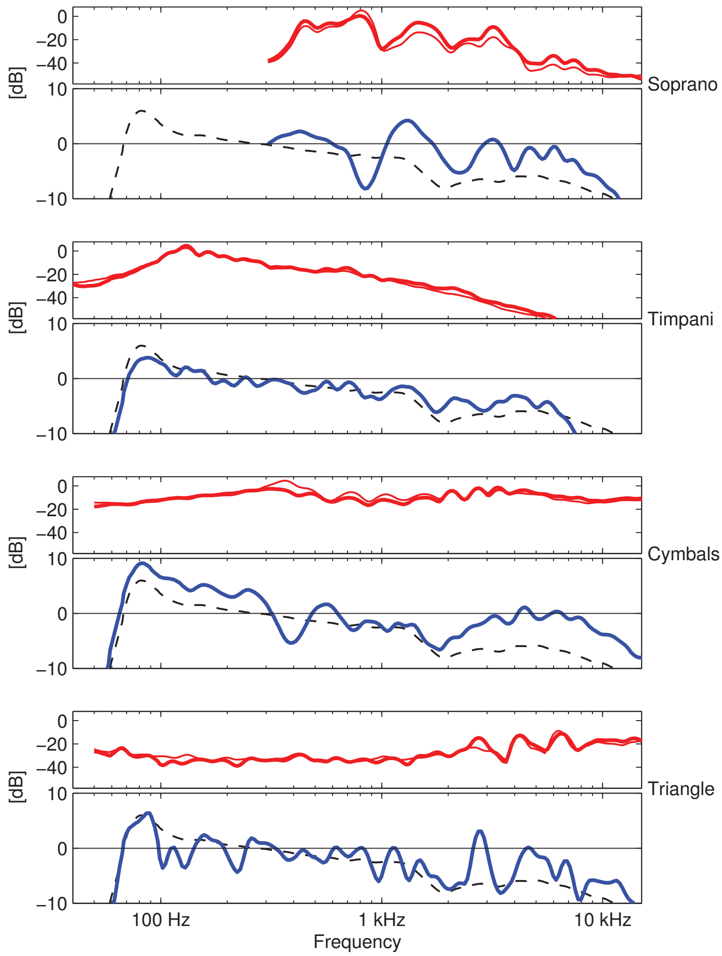



**Figure 4.5.** Comparison of the percussion instruments and soprano power responses to the loudspeaker reproduction. The visualization is similar to Fig. 4.2.

ences in this regard. The reproduced trumpet and trombone power responses are particularly close to the real instrument power responses (see Fig. 4.3). Otherwise, the power response of the instruments is characterized reasonably well by a frequency response in a single direction. The slight low-pass behavior observed e.g. with the woodwind instruments and the cello suggests that the applied loudspeaker is too directional.

Violins account for approximately one-third of all instruments in an orchestra. Thus, reproducing violin directivity and the power response can be considered important. This is supported by the earliest experiments with the loudspeaker orchestra indicating that the violins are possibly the greatest single challenge with regard to the subjective realism of the orchestra sound. Here, an improvement in reproducing the sound of a violin section is proposed by combining two ordinary loudspeakers in different orientations. The main loudspeaker is mounted on a stand in a typical height of a played violin. An auxiliary loudspeaker connected in parallel with the main loudspeaker is positioned on the floor in an upright position. This arrangement is illustrated in Fig. 4.1 with the circles connected to the numbered loudspeakers. The proposed approach is inspired by the insignificant increase in the practical complexity of the orchestra setup. The displacement between the two loudspeakers introduces a comb-filter effect depending on the receiver position. On the other hand, the effect is different for the direct sound and each reflection and similar effects occur to some extent with orchestra violins playing in unison as well. It is thus left disregarded here.

Figure 4.6 illustrates the effect of combining the loudspeakers in unwrapped directivity patterns of CLF coordinate system at the 2 kHz octave band. The directivity of the forward and upward-pointing loudspeakers is shown in Figs. 4.6a-b, respectively. The sum of the directivities is compared to the measured violin directivity in Figs. 4.6c-d. The improvement to the radiation pattern by adding the auxiliary loudspeaker is considerable. This is further investigated with the average differences in the radiated sound levels. Table 4.1 presents statistical values for the directivity and power differences between the violin and the two-loudspeaker configuration. Root-mean-square difference to the violin is calculated from the directivities in CLF type 1 format with 10-degree intervals. The difference is notably smaller with two combined loudspeakers above the 1 kHz octave, which suggests that the proposed approach has the desired effect. $\Delta P_{\mathrm{rel}}$ indicates the difference between the average sound energy

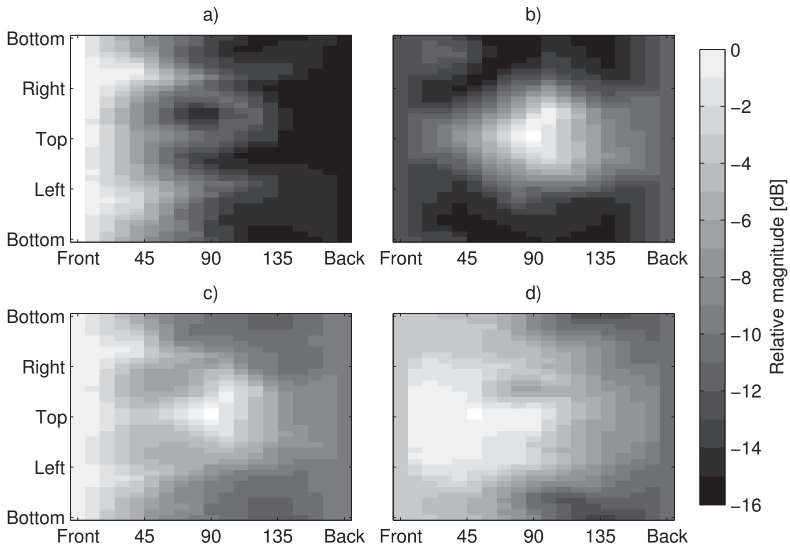



**Figure 4.6.** Comparison of single and combined loudspeaker directivities compared to the average violin directivity at the 2 kHz octave band. Subfigures are in unwrapped CLF coordinate system, which represents the spherical surface around the source as seen from above. Vertical axes indicate the direction through which the arc passes from the front to the back of the source. Horizontal axes indicate the position on the current arc. Labels on the vertical axes denote the direction where the arc is pointing to at 90 degree position (on horizontal axis). a) Measured directivity of Genelec 1029A pointing forward. b) Simulated directivity of Genelec 1029A pointing upward. c) Combined forward and upward radiation pattern. d) Measured violin average directivity pattern.

| | RMS error [dB] | | $\Delta P_{\mathrm{rel}}$ [dB] | | $P$ [dB] |
|---|---|---|---|---|---|
| Octave [Hz] | Single | Combined | Single | Combined | Violin |
| 250 | 1.55 | 1.75 | -0.51 | 0.79 | -2.84 |
| 500 | 1.62 | 1.95 | -0.93 | 0.94 | -3.1 |
| 1000 | 3.22 | 3.25 | -0.38 | 2.37 | -5.59 |
| 2000 | 5.76 | 2.07 | -5.58 | -0.53 | -4.95 |
| 4000 | 3.63 | 2.46 | -3.08 | 1.25 | -6.94 |
| 8000 | 6.24 | 2.8 | -5.35 | -1.02 | -5.38 |

**Table 4.1.** Comparison of the acoustic power output with a single loudspeaker, a combination of two loudspeakers, and a violin. Directivities are based on measurements in Publications III and II. The loudspeaker data is measured with Genelec 1029A [61]. RMS error describes the average difference in the directivity patterns over 10 degrees radiation cones in relation to the violin. $P_{\mathrm{rel}}$ indicates the acoustic power in relation to the violin measurement $P$ at the octave bands. Positive values for the loudspeakers suggest a more omnidirectional radiation than the violin.

| Reference | Type of application |
|---|---|
| [104, 110, 108] | Recording of the loudspeaker orchestra playing anechoic music |
| [103, 109] | Convolution of anechoic music with measured impulse responses |
| [134] | Comparison of convolutions of anechoic music with measured and simulated impulse responses |
| [136, 135] | Convolution of anechoic music with measured impulse responses |
| [102] | Evaluation of acoustic reflections with a simulated loudspeaker orchestra |

**Table 4.2.** A list of studies where the loudspeaker orchestra has been applied in various purposes.

radiated by the loudspeakers and the violin $P$. At low and middle frequencies the single loudspeaker is closer to the violin power response, but at high frequencies, again, the two-loudspeaker combination presents a smaller difference from the violin radiation.

## 4.3 Discussion

A loudspeaker orchestra for the evaluation of room acoustics in-situ or in laboratory conditions has been presented. During its evolution the loudspeaker orchestra has been utilized in several studies on concert hall acoustics. In [134] the loudspeaker orchestra was applied in existing concert halls and their simulated counterparts in order to investigate the reliability of auralization. In [110], the loudspeaker orchestra was used as the sound source for evaluation of the apparent width of the sound field. In [109, 108] several concert halls were assessed subjective with the loudspeaker orchestra. A more comprehensive list of references is given in Table 4.2.

Improvements to the reproduction of the violin sound with the loudspeaker orchestra were proposed above. Regarding the instrument directivities, a more accurate representation could be achieved with purpose-built sources or a combination of different loudspeaker designs. Using specialized hardware would, however, reduce the generic applicability of the loudspeaker orchestra.

The level of realism achieved with the loudspeaker orchestra has been mostly under informal evaluation. The authenticity has received positive verbal feedback from the orchestra musicians or sound engineers who have been present during in-situ listening of the loudspeaker orchestra. Experienced subjects or visitors listening to the loudspeaker orchestra samples used in listening tests (e.g. those given in Table 4.2) have not mentioned about apparent unnaturalness in the orchestra sound. In addition, the quality of the loudspeaker orchestra has been indirectly evaluated with a questionnaire at a listening test which concentrated on the acoustic differences in the measured halls. The overall quality of the convolved loudspeaker orchestra samples used in the test received a mean score of 3.8 with a standard deviation of 0.8 on a scale of 1 to 5 (unsatisfactory — excellent) with 19 subjects. A detailed evaluation is a challenging task for two principal reasons. First, the acoustic impression is affected by the performed music. The currently used anechoic music samples cannot be applied to a live orchestra. Alternatively, the live orchestra cannot assume the directivities of the loudspeakers. Second, the error from approximating individual instruments with loudspeakers can be objectively quantified, but the overall directivity of a real symphony orchestra is very complex. The detailed evaluation of the authenticity with the loudspeaker orchestra is planned for future work.

Independent of applying the loudspeaker orchestra for a convolution with measured responses or in-situ listening, high quality anechoic orchestral music is regarded the most important single requirement. The following chapter discusses the recording of symphonic music in an anechoic chamber.

# 5. Anechoic orchestra recording

Publication II describes the anechoic recordings of four excerpts of orchestral compositions representing different musical styles. The recordings have been made freely available for academic purposes [178]. The following sections summarize the course of accomplishing the anechoic orchestra recordings.

## 5.1 Recording setup

Anechoic conditions are necessary for the reason that any significant sound reflections may modify the measured response. Standing waves affect mostly the decay time of the lowest frequencies in the room. Single reflections could cause coloration by altering the measured frequency response. However, the instruments producing fundamental frequencies significantly below 100 Hz are the timpani, the tuba, the contrabass, and the violoncello. Except for the tuba, these instruments have a noticeable decay time, thus slightly reducing the critical importance of fully anechoic conditions at low frequencies.

The cube shaped anechoic chamber used for the recordings has free dimensions of 4.2 m in each direction, and the absorption wedges are 80 cm deep. The anechoic conditions are assumed at frequencies above 100 Hz. Hence, the measurements below 125 Hz are considered approximate. Large-diaphragm Røde NT1-A type microphones were selected for the measurement in multiple directions. The manufacturer reports a low internal noise in this model ($L_{\text{noise,A}} = 5$ dB) [151].

A dodecahedron shape was selected due to its symmetrical properties as a platonic solid. A dodecahedron can be regarded to sample the spherical surface around the center by the equal distances between microphones positioned at the vertices. In addition, triangular directivity cones, for example multi-channel auralization, can be modeled for future research with its dual polyhedron, the icosahedron (Fig. 5.1). The dodecahedron was oriented to form four horizontal microphone levels, each consisting of five microphones in a regular pentagon. In addition to the 20 microphones in the dodecahedron vertices, two additional microphones of the same kind were positioned at the front and above directions from the center point. The microphone positions are listed in Table 5.1. The recorded musician is facing the direction of $0^o$ az / $0^o$ el.

The distances from the center of the room to the microphones were between 1.81 and 2.49 m, while the average distance was 2.13 m. Top level microphones were the furthest from the center; the average distance was 2.42 m. The microphones at the two middle elevation levels were positioned as far from the center of the room as possible, yet avoiding the proximity of the tips of the absorbing wedges. Due to the loudspeakers present in the room for unrelated purposes, positioning the microphones near the loudspeakers was avoided in order to prevent sound reflections in recordings. The loudspeakers were at least at the same radius from the center of the room as the microphones. Hence it is assumed that the acoustic effect caused by the loudspeaker cabinets is of a diffractive type instead of specular reflections. An existing $1\times1$ m rigid steel grid in the center of the room served as an acoustically transparent floor. For the in-

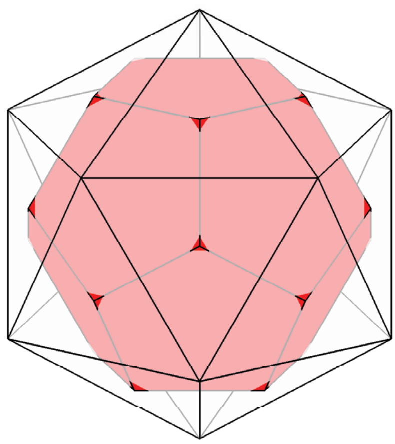


**Figure 5.1.** A dodecahedron shown in color inside an icosahedron. The dodecahedron is rotated to the orientation of the microphone array so that the microphone positions are visualized by dark regions protruding the surrounding icosahedron. The visible icosahedron triangles could be used to model the directivity cones in multi-channel auralization purposes.

**Table 5.1.** Elevation and azimuth angles, and distances of measurement microphones. Distances from the center of the room is denoted with $r$. Microphone no. 14 was aligned off the theoretical position due to the doorway to the anechoic chamber.

| Mic. | Ele [°] | Azi [°] | $r$ [m] | Mic. | Ele [°] | Azi [°] | $r$ [m] |
|---|---|---|---|---|---|---|---|
| **1** | 52.6 | 0 | 2.43 | **11** | -10.8 | 36 | 2.16 |
| **2** | 52.6 | 72 | 2.24 | **12** | -10.8 | 108 | 2.03 |
| **3** | 52.6 | 144 | 2.46 | **13** | -10.8 | 180 | 1.87 |
| **4** | 52.6 | 216 | 2.49 | **14** | -10.8 | 249 | 1.81 |
| **5** | 52.6 | 288 | 2.49 | **15** | -10.8 | 324 | 2.06 |
| **6** | 10.8 | 0 | 2.30 | **16** | -52.6 | 36 | 2.05 |
| **7** | 10.8 | 72 | 1.94 | **17** | -52.6 | 108 | 2.04 |
| **8** | 10.8 | 144 | 1.92 | **18** | -52.6 | 180 | 2.00 |
| **9** | 10.8 | 216 | 2.14 | **19** | -52.6 | 252 | 1.92 |
| **10** | 10.8 | 288 | 2.25 | **20** | -52.6 | 324 | 2.08 |
| | | | | **21** | 0 | 0 | 2.21 |
| | | | | **22** | 90 | 0 | 2.06 |

strument recordings, an additional 2 $m^2$ steel grid was installed on top of rubber dampers for accommodating the larger percussion instruments.

## 5.2 System calibration

The recording setup was equalized with filters designed on the basis of a reference measurement. First, a Genelec 1032A loudspeaker used for the calibration process was measured in an empty, large anechoic chamber with one Brüel & Kjær 4191 (B&K) microphone, which is considered ideal. A laser beam was utilized to align the loudspeaker towards the microphone. In addition, the loudspeaker was equally measured with the Røde microphones to be used for directivity measurements. Results from the reference measurement are depicted in Fig. 5.2. The B&K and Røde responses are shown with dashed and solid lines, respectively. The most apparent feature in the Røde microphones is the pronounced response at high frequencies. The peak visible at 60 Hz is a previously known feature in the large anechoic chamber.

Second, the responses of the 22 Røde microphones in their final positions in the recording room were measured. The same Genelec 1032A

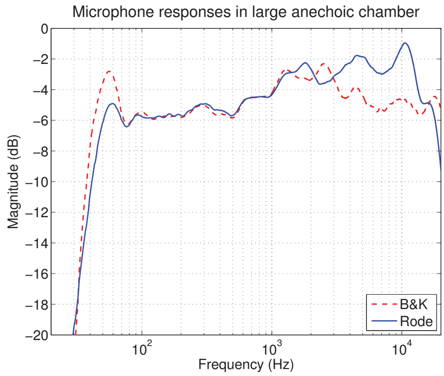



**Figure 5.2.** Comparison of B&K (reference) and Røde (recording) microphone responses with measurement loudspeaker response. Responses are smoothed to 1/3-octave resolution.

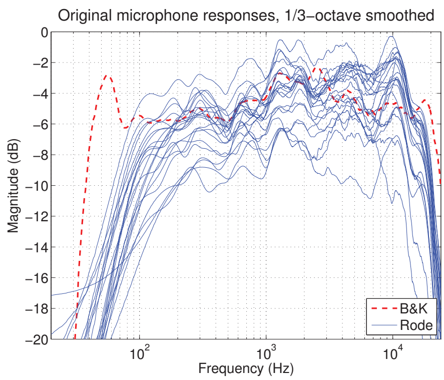



**Figure 5.3.** Original measured magnitude responses of each of the 22 microphone channels in their final positions. The dashed line is the response of the B&K microphone in the large anechoic room.

loudspeaker previously measured with the B&K microphone was used for this purpose. The loudspeaker was mounted on a tiltable and rotatable stand at the center of the room, and again, laser alignment was used. The frequency responses from these measurements are shown in Fig. 5.3 (the B&K response in the large anechoic chamber is plotted with a dashed line). The characteristic response of the loudspeaker is present in all responses, showing similar overall behavior as in Fig. 5.2.

The objective of the equalization was to compensate the differences in sensitivity and frequency response of the Røde microphones compared to the B&K response. The target for the equalization filter design were obtained by deconvolving the Røde measurements in the recording room with the B&K measurement in the large anechoic room.

First, shelving filters were applied to the original difference responses to facilitate a more efficient filter design due to the large differences seen in

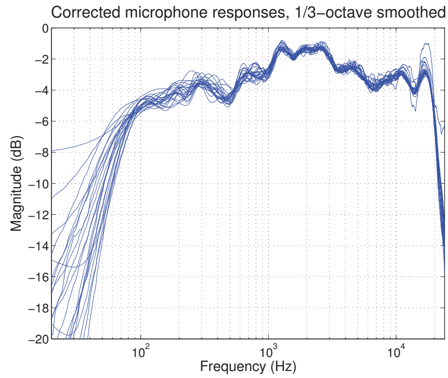



**Figure 5.4.** System corrected magnitude responses of all 22 microphone channels. Note that responses are not totally flat, since the loudspeaker response is included.

Fig. 5.3 at low frequencies. To give more weight to low frequencies in filter design, the 1/3-octave smoothed target frequency responses were sampled according to the ERB resolution [123]. Then, IIR filters of order 50 were fitted to sampled targets with the `yulewalk.m` algorithm in the Matlab environment. Additional, a manually-tuned shelving filter was applied to each filter in order to further equalize the low frequency response. In total, the equalization of each microphone was performed with one IIR and one shelving filter. The equalized responses are depicted in Fig. 5.4. The corrected responses are within 1 dB range between 500 and 10 000 Hz and within 2 dB range from 60 to 20 000 Hz. It is notable, that in the shown responses, the loudspeaker response is still present.

The equipment present in the recording space could cause measurement errors, however, the probability of sound energy reflecting to the microphones from the present loudspeakers was diminutive, and all rigid supporting structures in the room were treated with absorptive padding. On the other hand, the absorbing wedges in the room can cause slight deviation between the microphone positions. The reflections from the outer walls behave differently depending on the position in relation to the absorbing wedges, thus, causing variation in the microphone responses. However, the microphones were not moved from their final calibration positions.

## 5.3 Instrument recording

The orchestra instrument recordings were carried out by recording individual musicians playing in the anechoic environment equipped with the measuring system introduced above. The performers were collected from four professional symphony orchestras. Only one musician per instrument was recorded and hence different parts written for the same instrument were performed by the same musician.

The recorded program consisted of four compositions, each representing a different musical style or complexity. The following excerpts were recorded:

1. *W. A. Mozart* (1756-1791): Soprano aria of *Donna Elvira* from the opera *Don Giovanni* (duration 3 min 47 s)

2. *L. v. Beethoven* (1770-1827): Symphony no. 7, I movement. Bars 1-53 (duration 3 min 11 s)

3. *A. Bruckner* (1824-1896): Symphony no. 8, II movement. Bars 1-61 (duration 1 min 27 s)

4. *G. Mahler* (1860-1911): Symphony no. 1, IV movement. Bars 1-72 (duration 2 min 12 s)

The differing characteristics of the compositions were, e.g., a soloist (Mozart); a typical classical composition with crescendos and varying musical textures (Beethoven); a large, late Romantic period orchestra with a relatively low harmonic complexity (Bruckner); an even larger orchestra with a higher tonal complexity (Mahler). In total, 14 musicians were recorded in sessions with the duration of 1.5–6 h. The largest number of different instruments recorded was 19 in Mahler's symphony, including various percussion instruments, a piccolo, as well as two different clarinets and trombones. For Mozart, Beethoven, and Bruckner works, 9, 11, and 15 different instruments were recorded, respectively.

The synchronization of the individually recorded musicians was guided with a video track showing a conductor conducting a pianist playing a reduction of the current score. The pianist was not visible in the image; only the piano sound was recorded. The video was presented to the musicians with a small LCD monitor and open-air headphones. Thus, the musicians could listen to the piano reduction of the orchestral parts and simultaneously follow the conductor on the video. In addition to the piano track, a self-monitoring was offered via the recording microphones and headphones. A consistent tuning was insured by the piano track with a recorded concert pitch ($a \approx 442$ Hz) and a tuning meter.

The head of a seated musician was at the center of the microphone array. While this causes slight differences between the instruments, this approach was regarded as the most consistent between instruments. Defining the exact acoustic center of the different instruments is not trivial, if at all possible.

In contrast to [176], the microphone gains were kept constant. Suitable levels were estimated by recording drum beats near the microphones for estimating the maximum sound level at the microphones. The benefit here is that the sound level difference between the instruments was not distorted and repeating the time-consuming calibration of the microphones was avoided. As a downside, the adopted approach provides suboptimal signal-to-noise ratio in comparison to adjusting the gains for each instrument individually. Approaches to the background noise issues are discussed in Publication II

## 5.4 Observations on the recording

The anechoic chamber, being an unusual condition in a musical sense, meant the musicians were carefully advised of the environment. Without the acoustic support of a room, they were instructed not to force their playing in order to produce a stronger sound. Such a tendency would cause a notable decrease in the naturalness of the instrument sound, especially regarding the string instruments [69, 176].

The Mozart and Beethoven excerpts were recorded without issues. With Mahler's and Bruckner's symphonies shorter passages were recorded at one time. Most of the brass instrument parts in Mahler's symphony were recorded in multiple segments, since many of the parts include passages where the risk of accidentally playing a wrong note is high. With Bruckner's symphony the continuous passages in fast tremolo and high dynamics were found to be difficult to play accurately in tune in anechoic conditions. Therefore two versions were recorded. First, the passages were recorded in sixteenth-notes with the correct dynamics for keeping accurately in tune. Second recordings were made in tremolo as written in the score, but with lighter dynamics. By combining the takes with two playing styles, an impression of a well-tuned tremolo is obtained.

In [176], the reference track supporting the musicians was provided with MIDI instruments. During the course of the recording process, the instruments in the MIDI reference were replaced by their recorded counterparts. Here, the reference track was kept unchanged during the course of the recordings. For the soprano, recorded last, the piano reference was replaced with the complete orchestra recording.

## 5.5 Post-processing

An editing phase was required to gather takes from all recorded instruments and to form an ensemble playing together. The best takes for each part were selected by subjective listening. If necessary, suitable partial takes were charted for combining one complete part.

The editing of all 22 microphone channels was performed synchronously. First, accidental wrong notes in a full instrument part were replaced. Possible timing inaccuracies in the synchronization against a reference track were adjusted. The recorded piano was used as a timing reference for the few first edited strings and wind instrument parts. Later, the piano track was omitted and the readily edited parts formed the timing reference.

The objective was not to have an unnaturally accurate synchronization. Thus, slight timing discrepancies were left unchanged. All editing operations were delicately performed with automatic crossfades between the editing points. The editing would not be easily perceived even by listening to individual anechoic tracks. With the principal purpose of the recordings in auralization, a cautious editing would be masked by the convolution with a room impulse response.

In comparison of traditional recordings, the current anechoic recordings contain a large amount of microphones — one microphone for each instrument, in practice, when the recordings are applied to auralizations. Hence, the background noise in the recordings must be considered. During pauses the instrument channels contain only noise. Therefore, the application of a noise gate is possible for muting a channel if the signal level remains under a certain threshold. As the noise level in the recordings is moderate, without reverberation suitable parameters can be easily found. This is further discussed in Publication II.

## 5.6 Discussion

The recording of a symphony orchestra in an anechoic chamber is possible by recording the instruments individually. A timing reference is required for retaining a mutual synchronization between the recorded parts. The current recordings of four excerpts of orchestral music have been carried out with 22 microphones around the instruments. The recordings distributed freely for academic purposes have been under considerable interest in the acoustics research field [178].

# 6. Directivity measurements

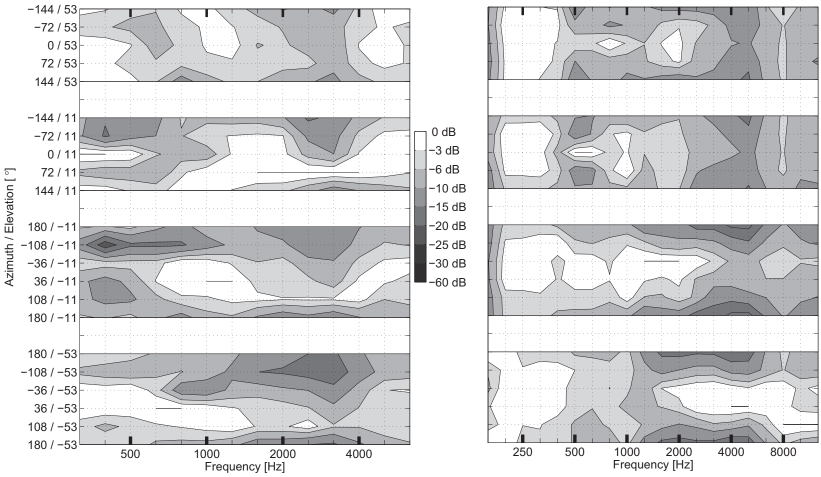
Publications III and V present the measurements of source directivity utilizing the setup described in Chapter 5 and Publication II. The analysis methods and the measurement results are briefly reviewed in the following sections. Furthermore, the obtained orchestra instrument directivities are converted in the Common Loudspeaker Format (CLF) for straightforward application in commercial acoustics modeling software. The conversion process and related discussion are presented in Publication IV.

## 6.1 Musical instrument directivities

A well-known book on the instrument directivity has been first published by Meyer 30 years ago [119], but in practice the measurement details are not available. Instrument directivity is only slightly covered in the acoustics textbooks [52, 152, 44]. A number of papers discussing the directivities of selected instruments have been written, but the differing analysis methods render comparing the results difficult, e.g. [27, 34, 80, 174].

The array used in the measurements consists of 22 microphones. Even the number of microphones in the larger arrays (e.g. 32 microphones in [144]) is fairly low for a truly accurate directivity analysis of natural sources. The main motivation of the present study is to establish a directivity dataset compatible with the anechoic orchestra recordings and to study its properties regarding the application in auralization and room acoustic modeling. An array consisting of 22 microphone is considered providing a sufficient spatial resolution for such purposes.

The measurements were carried out with professional orchestra musicians in the previously discussed recording conditions. All instruments




**(a)** Flute **(b)** Oboe

**Figure 6.1.** Averaged unwrapped directivity plots for a flute and a oboe. The maximum of one-third octave band has been separately normalized to 0 dB.

were played in a manner similar to an actual performance. Hence, the measurements include the effect of the musician itself, and a lightweight music stand in front of the musician. This approach was chosen due to the foreseen application in auralizations and room acoustic modeling, where directivity data reflecting a real performance is preferred to a theoretical radiation of the instrument alone.

The overall directivity is presented in four elevation levels with five microphones each. In the presented directivity visualizations the musician is facing the direction 0 deg azimuth / 0 deg elevation, while the positive angles denote directions to the right and up, respectively.

The instrument directivities were measured using the tones A-major triad played in two octaves in the characteristic playing range of the current instrument. Each tone was played with three instructed dynamics: *p*, *f*, and *ff*.

The following sections summarize the principal results from the detailed directivity analysis also found in Publication III.

### 6.1.1 Wind instrument directivity

The woodwinds exhibit a directivity pattern with generally low predictability, as the instruments can be considered having a set of point sources in a line. As the tone holes are opened and closed, the directivity changes with the played tone. Overall, the high frequencies are radiated in the direction of the bell while the low frequencies are more or less omnidirectional.

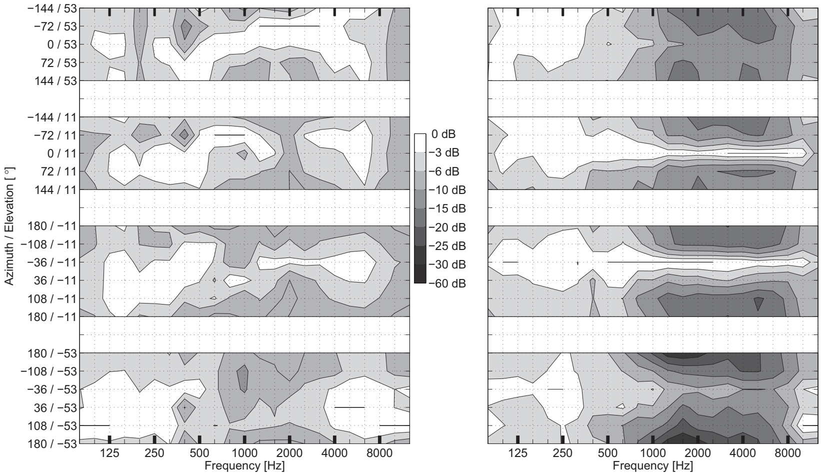
The averaged flute directivity is presented in Fig. 6.1a. Around 500 Hz the sound is radiated to the front region while a substantial attenuation is found on the left side. Above 600 Hz the directivity begins to concentrate on the right side. Very pronounced directivity is noticed below 4 kHz in the direction of the open end.

The effect of the played tone on the directivity is two-folded. First, the radiation patterns of single tones change with respect to different fingerings, and different fundamental frequencies radiate in substantially differing directions. Second, the radiation of the harmonics is similar with tones having the same fingering, hence overblown tones, as the same tone holes remain open. This finding corresponds to the previous results presented in the literature [119]. Importantly, similar effect is found also with tones having nearly similar fingering. An example of this behavior can be seen with tones that have harmonics at the same frequencies, such as the first harmonic of E6 and the second harmonic of E5 (see Figs. 5b-5c in Publication III).

With the oboe, fingerings for the tones in the lowest octaves are similar, as the octave register is changed with a separate tone hole. Hence, the oboe and the flute have comparable fingerings to some extent. This suggests a similar change in the directivity with different tones as with the flute.

The average directivity can be generalized by a more omnidirectional pattern in the lower hemisphere below 400 Hz and a narrowing beam in the bell direction above 1000 Hz, as seen in Fig. 6.1b. The fundamental frequencies of the lowest measured oboe tones radiate mostly omnidirectionally unlike the higher notes. The cutoff frequency is used in the literature to denote the approximate frequency above which the instrument radiates mostly in the bell direction. The present observations are roughly in line with the 1.5 kHz cutoff frequency found in the literature [12, 13]. In contrast to the literature, here nearly omnidirectional radiation is only found at frequencies much below the cutoff. The radiation behind the player attenuates gradually above 500 Hz beginning from higher elevations.

The measured clarinet directivity is partially similar to the oboe with the nearly omnidirectional characteristics below 500 Hz and the narrow radiation to the lower front region above 1000 Hz. Although the even harmonics are missing from the clarinet sound at the lower register tones,




**(a)** Bassoon **(b)** Trombone

**Figure 6.2.** Plots of averaged directivity for a bassoon and a trombone.

the clarinet exhibits the same phenomenon observed with the flute and the oboe, where the radiation patterns are similar at the common harmonic frequencies. The overall results from the analysis are comparable to the directivity values given in [119]. In [131] the presented directivity is more even in the horizontal and median planes at the 1 kHz octave band compared to the results obtained in Publication III.

Also the lowest bassoon frequencies radiate rather omnidirectionally. The two highest elevation levels are the strongest at the middle frequencies, and above 1000 Hz the directivity along the instrument axis is prominent. The decreasing level in the sound spectrum above 3 kHz prevents a reliable analysis at the high frequencies. It is notable that the overall radiation pattern shown in Fig. 6.2a corresponds to the oboe directivity in Fig. 6.1b at frequencies approximately 1.5 octaves lower, when the instrument orientation is taken into account.

Characteristically to the bassoon, the direction closest to the bell receives a considerably low sound pressure level at overtones around 400–500 Hz. This refers to the bassoon formant described in [98], which lies at the specific frequency band. As with other woodwinds, the analysis of individual tones reveals that the sound radiation is similar at harmonic frequencies common to different tones.

Based on the results it can be concluded that with woodwind instruments a directivity pattern averaged over several tones is a good approximation of the actual directivity during performance. Different playing dynamics are not noticed to cause changes in the woodwind instrument directivity, while the obvious change in the harmonic structure related to the increased blowing pressure is observed. The results for the woodwind directivities are generally comparable to the results given in [119], but the differences between the coordinate orientations make detailed comparison complicated.

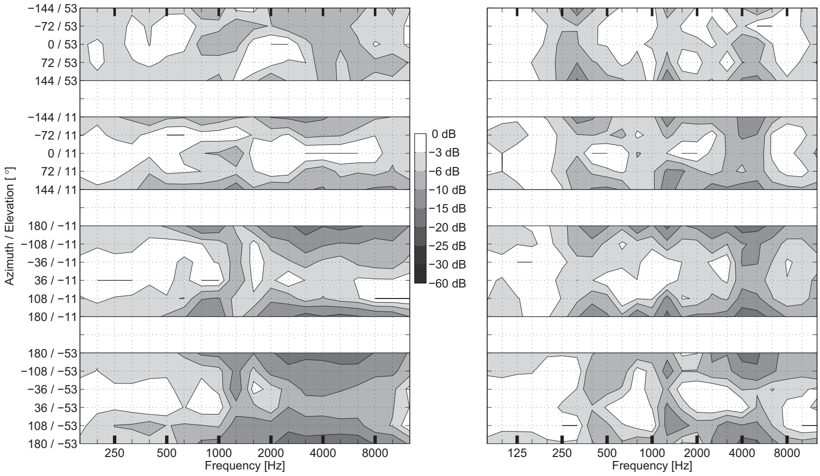
### 6.1.2 Brass instrument directivity

The brass instrument pitch change relies on extending the tubing instead of the tone holes found in the woodwinds. Based on the simple radiation mechanism of the brass instruments, the directivity analysis is straightforward. The tenor trombone results are shown as an example of the brass instrument directivities in Fig. 6.2b.

The French horn radiation is omnidirectional up to the 500 Hz frequency band. At higher frequencies the radiation is concentrated in the direction of the bell. Above 1250 Hz the direction of the dominant radiation varies between the middle and bottom elevations. A possible cause for such a behavior is the typical playing posture, where the right hand of the player is held at the bell opening. Therefore the effective shape of the radiating bell is changed.

These results, particularly the directivity in the lateral plane at lower frequency bands, are in line with the findings in [119]. Compared to the figures at the 1 kHz octave band given in [132], the values on the attenuation at the opposite from the bell direction are in the same magnitude. The measured low-pass power response characteristics (-21 dB/oct. in *p*, -13 dB/oct. in *f*, -9 dB/oct. in *ff*) above 800 Hz are comparable to the rolloff rate of 15 dB/oct. cited in [111].

Also the trumpet directivity becomes narrower as the frequency increases. The radiation remains mostly within -6 dB of the maximum at the frequencies below 400 Hz. Above 1 kHz the directivity in the bell direction is rapidly emphasized in the same manner as in Fig. 6.2b. The apparent cutoff frequency of 1 kHz is consistent with the values in the literature presented in [53]. In [119], it has been reported that the omnidirectional pattern is effective up to 500 Hz within -3 dB range and up to 1100 Hz within -10 dB. Here, differences up to 6 dB are found at the low frequencies. The low-pass characteristics (-21 dB/oct. in *p*, -20 dB/oct. in *f*, -13 dB/oct. in *ff*) are comparable to the values between -15...25 dB/oct.




**(a)** Violin **(b)** Cello

**Figure 6.3.** Plots of averaged directivity for a violin and a cello.

cited in [52].

Figure 6.2b shows the trombone directivity. The pattern remains omnidirectional up to 400 Hz while the directions near the bell remain strong up to the 630 Hz one-third octave band. At the higher frequencies the radiation is straightforward until the noise floor is reached at 10 kHz. In Fig. 6.2b the data is presented without the one-third octave normalization, where the radiated spectrum is close to the average of the tones in *forte*. The sound pressure decreases rather rapidly even in the direction of the bell after 630 Hz, suggesting a cutoff frequency for the instrument.

The tuba directivity follows the previously discussed brass instruments, whereas the sound level of the lowest frequency bands up to 250 Hz is within -6 dB of the maximum direction. Short wavelengths combined to the large bell result in the highest directivity of the measured instruments. Above 4 kHz the second-closest microphones in the bell direction receive sound pressure levels only -20 dB of the maximum, and other directions even less than that. The results comply with the -10 dB limit up to 400 Hz reported in [119].

The overtone spectrum of the brass instruments change considerably with playing dynamics. The directivity is nearly omnidirectional at low frequencies, as the bell does not effectively direct the sound at the long wavelengths. Compared to the discussion in [152], the observed upper limit of omnidirectional patterns are closely proportional to the inverse of the bell diameter.

### 6.1.3 String instrument directivity

The brass instruments are the closest to a single radiating point and the woodwinds can be seen as a line of point sources. The string instruments are closer to volumetric sources, as there is not any single region in the instruments that emit sound. Instead, the body has different vibrational modes that can vary also between individual instruments. Therefore the directivity analysis is more challenging than with wind instruments.

The average violin directivity is shown in Fig. 6.3a. The directivity below 500 Hz is mostly omnidirectional, while in the range of 2–6.3 kHz the sound radiates in the front direction. At this frequency range, the radiation to the lower elevations is weak. The radiated sound is more evenly distributed at above the violinist and the normalized sound levels do not fall below -10 dB in the weakest directions. Around the 1250 Hz one-third octave band, a noticeable drop is detected in the radiation pattern. At this frequency band the sound is radiated relatively more in the direction of the violin neck, that is, on the left side of the player. A direct explanation for this is not found, as the most important vibrational modes are usually located at much lower frequencies [77, 83, 114, 152, 156, 181]. With a more detailed investigation on the directivity, similar behavior that was previously discussed on the constant radiation of the woodwind harmonics, is observed also with the violin tones. The phenomenon has been proposed in [186] as the directional tone color.

In [119] the violin is stated to be omnidirectional up to 600 Hz within -10 dB. This corresponds to the current findings. Also strong attenuations around 1000 Hz in the horizontal plane are reported [119]. While the coordinate system here is different, the results suggest strong fluctuations in that frequency range. The dipole radiation or the major radiation lobes in the bridge plane at 2300 and 3100 Hz, stated in [185], could not be confirmed based on the current analysis.

The viola directivity shows a similar forward directing radiation to the violin (see Fig. 6.3a). Two boundaries in the directivity are observed. The first one is noticed at 630 Hz. The second one, although not as prominent fluctuation in the sound radiation, is present around 2 kHz where the radiation is concentrated in the direction of -36/-11 (az./el.) degrees. Above 2 kHz the directivity remains constant. The lower boundary appears much lower than around 1000-1250 Hz as with the violin. The region between the 500 and 800 Hz one-third octave bands presents strong variations in the directivity also with individual tones while the average radiation is concentrated in the front of the players. As with the violin, the same phenomenon of similarly radiated harmonics with different tones is noticed with the viola.

A comparison of the measured viola and violin power responses lead to findings where comparable features are found at 18–47% lower frequencies with the viola. This corresponds to the values for the relation of the resonance frequencies in the literature. The declining slope of the overall sound level at the high frequencies is also in line with a reported -15 dB/oct. rate [52].

The concept of a volumetric source is manifested with the cello. As the cello plate dimensions are approximately twice larger than in the violin [52], having a corresponding relative measurement distance would turn out as a microphone array with a diameter of 8 m. Hence, the theoretical assumption of a point source is not entirely valid, although the standardized requirements on the sound power measurement is fulfilled [79]. The larger dimensions of the instruments are taken into account by compensating the sound level in the microphones by $1/r$ law following the displacement of the geometrical center of the instrument from the measurement array center.

The average directivity of the cello is presented in Fig. 6.3b. As with the viola, the cello sound radiation can be divided by two boundaries in frequencies where the directivity is concentrated. The cello radiates omnidirectionally up to 300 Hz. At the 315 Hz frequency band the directivity changes rapidly from a figure-of-eight shape in the bottom elevation into a narrow beam at the front. At the 1250 Hz frequency band, the radiation is again concentrated in the front region. The harmonics of the individual tones exhibit constant directivity to some extent, but less prominently than observed with the violin and the viola.

In the literature omnidirectional radiation has been reported below 200 Hz, which is lower than observed here [119]. Pronounced resonances reported around 250–300 Hz could be indicated by the current observations. Also the concentrated directivity in the front at 500 Hz and 2000 Hz is visible in these results. In addition, the front-back ratio is cited to first exceed 10 dB at around 500 Hz, corresponding to current findings.

The contrabass directivity is measured with a similar approach as the cello. Directivity in the forward direction can be characterized in three regions. First, below 200 Hz the sound pressure level is the highest in the lowest elevations. Second, the middle elevation is dominant between 200–750 Hz. Third, the high frontal direction is the strongest above 750 Hz. In addition, the back side of the instrument shows considerable sound levels around 500 Hz. Unlike the smaller string instruments, the contrabass is not noticed to be entirely omnidirectional at any frequency band. A rather ambiguous omnidirectional limit could be observed around 200 Hz.

The directivities at the harmonic frequencies that are common between different tones are again found to be similar. The overtones of the recorded contrabass tones are relatively weak compared to other string instruments. Therefore only a few first harmonics can be compared.

These general observations are fairly well in line with the behavior reported in [119], citing the low frequencies to have varying radiation. The radiation is also noted to be more omnidirectional only around 100 Hz. The lateral directivity to the frontal half-plane at 1000 Hz is also illustrated in [119].

The discussion on the string instruments is concluded by stating that they exhibit more complex directivity patterns than with the wind instruments. In spite, the effect of the harmonic frequencies having a constant radiation pattern is found in several examples. In addition, one or two frequency bands dividing the contiguous regions of the radiation patterns were found depending on the instrument. With the violin, the viola, and the cello, the lower boundary can be conceived as a cutoff frequency for the omnidirectional radiation. Such boundaries were discussed in relation to the woodwind instruments. The phenomenon is further studied by comparing the detected cutoff frequency against the string tuning and the relative resonance frequencies found in the literature [52, 76]. A strong similarity is found between these three relations (see Publication III, Chapter 8.5).

## 6.2 Directivity conversion into Common Loudspeaker Format

The Common Loudspeaker Format (CLF) is a file type for exchanging information on sound sources [38]. While the format is mainly intended for loudspeaker sources, in addition to the included physical features, power response, and electro-acoustic properties, directivity data is included in the file format. Hence, CLF files can be utilized in distributing directivity data of musical instruments. The CLF format is currently supported in various commercial modeling software [38, 128]. Previously, research data from [119] has been published in CLF-compatible text format [141].

Publication IV discusses the conversion of the directivity measurement into CLF format for further application in auralization or acoustics modeling.

### 6.2.1 CLF coordinate system

Three-dimensional directivity data in CLF files is stored in a specific coordinate system which differs from the spherical coordinates. Instead of azimuth and elevation angles, discrete directions are defined with two angles. The first angle indicates the arc of a constant radius from the front to back of the source. The second angle indicates the position on the current arc. The coordinate system is illustrated in Fig. 6.4.

The chosen coordinate system originates from the convenience of measuring loudspeaker directivity with a rotating arc equipped with an array of microphones. It is important to notice that the CLF coordinates do not coincide with equal spherical coordinates in most angle combinations.

Two variants of CLF format exist. In CLF type 1 the directivity is defined in octave bands with 10 degree spatial resolution. CLF type 2 utilized one-third octave bands and 5 degree intervals.

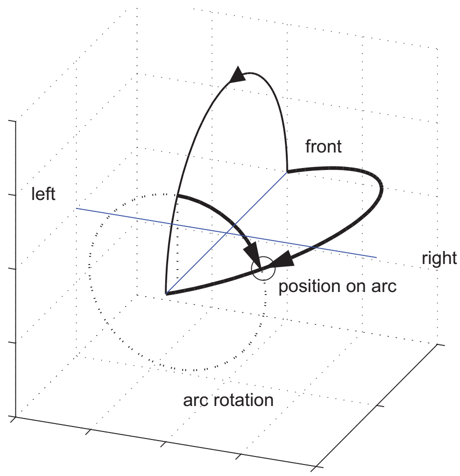



**Figure 6.4.** Illustration of the CLF coordinate system. The reference arc with zero rotation passes through the zenith. The arrows indicate the direction of increasing angles.

### 6.2.2 Conversion

A problem arises when the dense grid of CLF format is to be populated with the measurements from 22 microphones. Two parallel approaches are discussed in Publication IV.

The first method, interpolation mapping method, assumes a unwrapped rectangular grid with azimuth and elevation angles. The measurement microphones are defined on the corresponding positions on the grid. In order to have circularly continuous mesh over the zero azimuth, the microphone positions are replicated to $\pm 360$ degree azimuth. Additionally, the convergence of the directivity at the poles is ensured by replicating the top microphone values densely at the 90-degree elevation over the azimuth angles, and correspondingly the interpolated value from the bottom microphones to the -90-degree elevation. Suitable interpolation is applied to the microphone positions, creating a mesh between the microphone positions filled with interpolated directivity data. A continuous mesh is then obtained by delimiting the interpolation result to 0–360 degree azimuth. This is illustrated in Fig. 2 in Publication IV. For transferring the interpolated directivity data into CLF, a complete set of CLF points is generated in spherical coordinates. Having the interpolated results and the CLF coordinates now defined in the same coordinate system, the nearest positions from the interpolated mesh is sought for each of the CLF points. Finally, the data is then mapped to the best-matching CLF points. Notably, the error introduced by the mapping is decreased when the interpolated mesh is denser. The worst-case angular error of 0.68 is found with 1 degree interpolation density and CLF type 2 target coordinates.

The second, microphone coordinate rotation method, employs transforming the microphone array base from spherical to CLF coordinate system prior to data interpolation. Hence, the interpolation grid is readily in CLF coordinate system. The replication of the microphone positions is performed as described above for ensuring continuous interpolation results. This is illustrated in Fig. 5 in Publication IV. The interpolation result is in a format that does not require additional processing before writing to a CLF file.

Both methods provide an approximation of the 3D directivity pattern. The advantage in the interpolation mapping method is that the microphones are more evenly distributed in the interpolation grid. The front and back regions are better represented in this method. In contrast, the

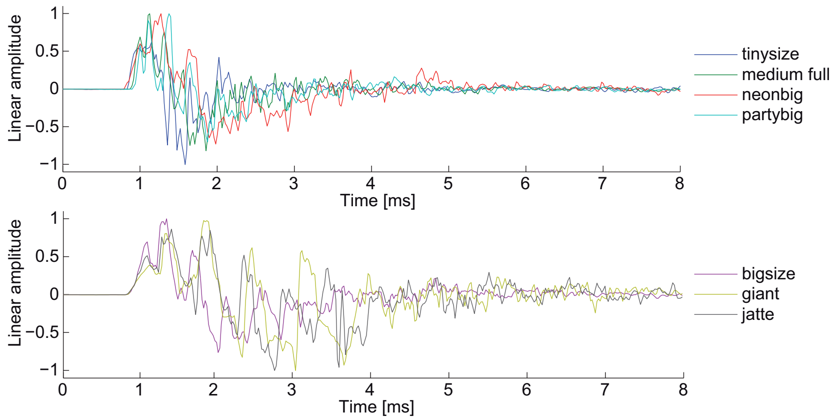



**Figure 6.5.** Examples of impulses produced by the investigated balloon types.

rotation method preserves the characteristic pentagon pattern of the array better (see Fig. 5.1). The observations suggest using the interpolation mapping method for the directivity conversion into CLF format.

## 6.3 Balloon burst directivity

Modern omnidirectional loudspeakers are usually applied in acoustic measurements. Sometimes, however, an inexpensive, portable impulse source provides convenience. The ISO3382-1 standard only imposes specific directivity conditions on loudspeaker sources. The described measurement system was utilized in Publication V to investigate the properties of a balloon burst more thoroughly than in the references found in the literature. Over 160 balloons were measured in the experiment for studies on the power spectrum and directivity for a range of balloon conditions.

Balloons were inflated in the traditional manner: blowing by mouth. Seven different models of balloons were tested. Over 70 examples of *medium* size balloons were tested for the analysis of repeatability. After initial trials, the balloon diameter was not found to be a reliable gauge of inflation level, as some balloons remained considerably under-inflated and others popped prior to reaching the same diameter. As such, balloon groups were inflated to the fullest size until a perceivable threshold in the inflation was reached. This threshold was both in air pressure and surface tension. While no measurement device was used, balloons were inflated by a single person, musically trained on wind instruments, establishing a level of confidence. Subsequently, the white balloons were noticed to behave considerably differently from the other color balloons. For this reason all the measurements with the white balloons were discarded. The material of the white balloon was observed to be less robust than that of colored balloons. For this reason, white balloons could not be inflated safely to the same level than the rest of the *medium* balloons. It is conceivable that the chemical composition of white rubber is different from other colors.

A 1 cm margin of error in the maximum width diameter was allowed for a consistent inflation level. Some variations in shape were observed for different colors in the same package, with some balloons being more elongated. These differences in shape, which were constant for a given color, would result in different volumes for equal diameters. As such, this will add some variance to balloon diameter-based models found in the literature.

The balloons were popped with a custom-built device, which consists of a remote-released arm with a pin attached at the tip. The release of the arm action is controlled with an electrical connection outside the anechoic chamber. An adjustable balloon mount ensured that the balloons remained at the center of the measurement array independent of the investigated balloon dimensions. The small side of the device is regarded not to interfere with the directivity results.

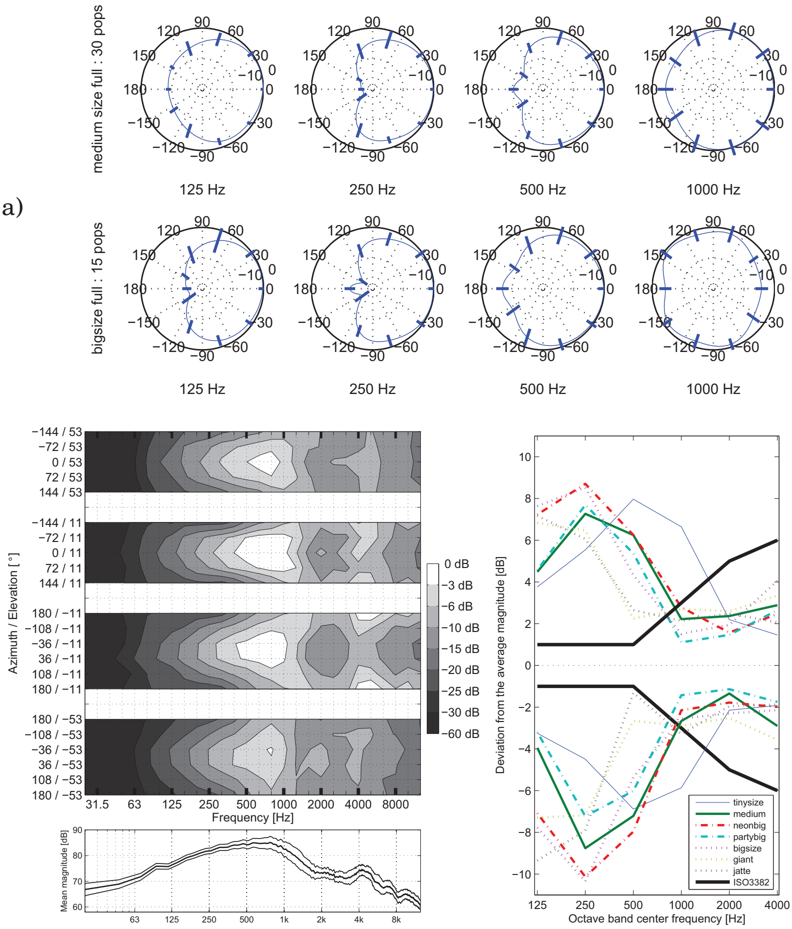
The impulsive nature of a popping explosion is the fundamental property of interest for using balloons in measurements. The shape of excitation with different balloons can be seen in Fig. 6.5. Impulse-like behavior was quantified with an energy rise time $t_{\mathrm{rise}}$. This was calculated as the time to rise from relative –40 to 0 dB from the average radiated energy with 100 Hz high-pass filter. Any variation of the distances from the balloon to the microphones was compensated for. Rise times are shown in Table 6.1, while the average rise time over all types was 1.7 ms (STD 0.3 ms). Smaller balloon types were observed to exhibit shorter impulses.

The directivity results for the balloons are presented in Fig. 6.6. Traditional polar plots on two different balloon sizes are given in Fig. 6.6a. The maximum magnitude at each octave band is normalized at 0 dB. Noticeably the directivity pattern changes between octave bands, where with the medium balloons a considerable cardioid shape is apparent at 250–500 Hz. With a smaller balloon type the phenomenon appears at approximately one octave lower. Figure 6.6b represents the 3D directivity in unwrapped form with four microphone elevation levels. Lighter colors in-

| Balloon type | $d$ [cm] | $PWL_{avg}$ (std) [dB] | $f_0$ (STD) [Hz] | $f_1$ (STD) [Hz] | $t_{\mathrm{rise}}$ [ms] | $n$ |
|---|---|---|---|---|---|---|
| *tiny* | 7–8 | 121.2 (2.2) | 856 (142) | 3106 (871) | 1.0 | 23 |
| *medium* | 18 ± 1 | 131.8 (2.0) | 668 (104) | 4208 (345) | 1.4 | 30 |
| green | " | 131.5 (1.0) | 774 (95) | 4196 (153) | 1.6 | 4 |
| orange | " | 132.0 (0.8) | 774 (23) | 4278 (365) | 1.6 | 6 |
| red | " | 133.1 (1.4) | 775 (62) | 3856 (295) | 1.3 | 6 |
| yellow | " | 131.0 (1.0) | 774 (88) | 3739 (318) | 1.3 | 5 |
| *neonbig* | 24 ± 1 | 130.4 (2.6) | 375 (40) | 2415 (235) | 1.8 | 9 |
| *partybig* | 24 ± 1 | 135.3 (0.8) | 457 (93) | 2813 (505) | 1.9 | 8 |
| *big* | 27 ± 1 | 132.8 (2.2) | 340 (30) | 2555 (234) | 2.5 | 15 |
| *giant* | 39–40 | 133.0 (3.5) | 223 (20) | 1008 (413) | 3.3 | 3 |
| *jatte* | 41–42 | 137.5 (2.1) | 235 (50) | 2274 (182) | 2.6 | 2 |

**Table 6.1.** Balloon types with the measured geometrical and acoustical properties: balloon diameter, $d$; peak sound levels, $PWL_{avg}$; frequencies $f_0$ and $f_1$; and number of samples, $n$. STD stands for standard deviation

a)




b) c)

**Figure 6.6.** Balloon directivity results. a) Traditional interpolated polar plots in horizontal plane in four octave bands for two balloon types. The point of needle impact is at 0 degree azimuth. Standard deviation at the microphone positions are indicated by the perpendicular bars. b) Unwrapped, unnormalized one-third octave band directivity for the medium size balloon. The four strips represent the four elevation levels consisting of five microphones each. The average power spectrum and the corresponding 95% confidence intervals are shown below. c) Comparison of investigated balloon directivity results against the omnidirectional criteria stated ISO3382-1 standard (thick solid line).

dicate higher radiated energy in the corresponding coordinates. Here, one should note that a major radiation peak, denoted $f_0$ in Table 6.1, is formed around the 800 Hz one-third octave band in the direction of the pin impact. At higher frequency bands the radiated energy drops rapidly, until at approximately 4 kHz a second peak, denoted $f_1$, is formed in the opposite direction. This general behavior was noticed with all balloon types in different magnitudes.

The frequency $f_0$ of the first radiation peak is found to correlate with the adapted equation for the Helmholtz resonator; $f_0 \sim \sqrt{r^{(-k)}}$, where constant $k \approx 1.13$.

The particular objective of resolving the compliance of the balloon directivity to the ISO3382-1 standard. This is shown in Fig. 6.6c, where the standardized limit for directivity variation is given in solid line against the investigated balloon types. Overall, none of the balloon types fulfill the requirements below 500 Hz, particularly due to the noticed cardioid effect. At higher frequency bands the directivity pattern is even enough to comply to the standard. The radiated energy is rather small at such frequencies while the wide-band radiated peak sound level surpasses 130 dB on average with most balloon types.

## 6.4 Discussion

Orchestra instruments were measured in an anechoic chamber with a calibrated microphone array in dodecahedron shape. The recording simulated a performance situation. The results indicate that the directivity of an instrument played by a musician can be modeled with the help of an averaged directivity pattern, although an averaging approach has been criticized in the literature [132]. This is supported by the observations where the excited harmonic frequencies tend to radiate in constant directions independent of the fundamental frequency.

A directivity database gathered from the measurements including a visualization tool is freely available for further analysis to accompany the published research articles [178]. In addition, the CLF files for the orchestra instrument directivities are made available for application in acoustic simulations.

# 7. Simulation of section sound for anechoic instrument recordings

The string section in an orchestra produces a broader sound than a string instrument soloist or a chamber ensemble. Individual instruments played in a section are not perceived separately, in contrary, their sounds blend together. The anechoic orchestral recordings discussed in Chapter 5 were performed with a small number of musicians, and the natural variations between the players in a group were not present. Therefore, the simulation of a section sound is necessary.

This chapter discusses a novel method for the string section sound simulation. The proposed method combines pitch-shift and asynchronization using a phase-vocoder technique in the short-time Fourier transfer-domain [45, 97]. In addition, small variations in the playing dynamics are created with amplitude modulation. The simulation of the section sound is based on tracking the temporal differences of a real orchestra during playing. The block diagram of the entire framework of analysis and synthesis is shown in Fig. 7.1. In the following sections the synthesis method is described first. After, two studies on tracking the orchestra performance are summarized. Finally, the simulation method is evaluated with a listening test. The experiments are described by the present author in detail in articles that are in press or yet to appear [137, 138].

## 7.1 A synthesis method for section sound

The time-domain signal $x_m(t)$ for simulated musician $m$ is transformed into short frequency-domain frames $X_m(n,k)$ with short-time Fourier transform (STFT) using a frame length of 2048 samples with 25% frame overlap. A time-base vector $n = 0, 1...N$ indicates the indices of the obtained STFT frames and $k$ denote the frequency bin.

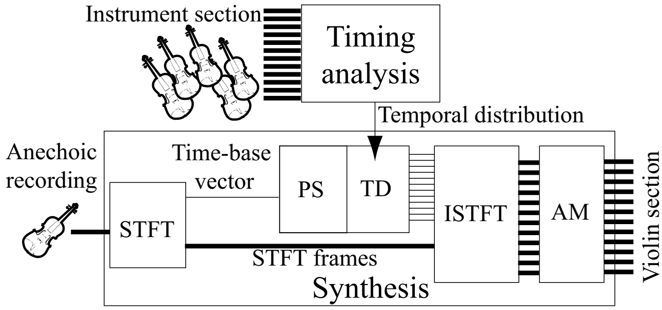



**Figure 7.1.** Block diagram of the section sound synthesis method. STFT, ISTFT, AM, PS, and TD stand for short-time Fourier transform, inverse short-time Fourier transform, amplitude modulation, pitch shift, and time difference, respectively.

The pitch-shift is obtained as the reciprocal effect of time-stretching [195]. The constant pitch-shift of $S^{(m)}_{\text{semitone}}$ semitones is achieved by scaling the time-base vector $n$ for the changed playback rate by the detuning factor of $d^{(m)} \in \mathbb{Q}$ closest to the desired detune in the linear frequency scale $S^{(m)}_{\text{linear}}$, and later resampling the final time-domain signal by the inverse factor $1/d^{(m)}$. The playback rate then is defined by a new time-base vector $\hat{n}^{(m)}$ that is the original vector $n$ resampled by $1/d^{(m)}$ intervals. Consequently, detuning factors of $1/d^{(m)} > 1$ yield a constant negative pitch-shift.

The time-variance is produced by adding fluctuation to the resampled time-base vector $\hat{n}^{(m)}$. The fluctuation is defined by a random vector having the same length as $\hat{n}^{(m)}$. Here, a random Markov chain $r^{(m)}$ following the Random Walk Metropolis-Hastings sampling (function `mhsample` in Matlab) from normal distribution is used [30]. With a low frequency it emulates the effect of a musician playing slightly behind the average rhythm at one moment and at the next moment catching up the tempo, or vice-versa. A 2 Hz frequency was found out to be suitable for a smoothly changing impression during the development of the algorithm ad-hoc. Furthermore, a spline interpolation is applied to the random sequence in order to avoid abrupt changes in the playing position causing audible artifacts. The advantage in using Markov-Hastings sampling is that the values in the random chain follow a normal distribution after the burn-in sequence. Therefore the standard deviation in the temporal distributions between simulated musicians can be adjusted with a single parameter. In addition, the Metropolis-Hastings sampling is not restricted to any par-

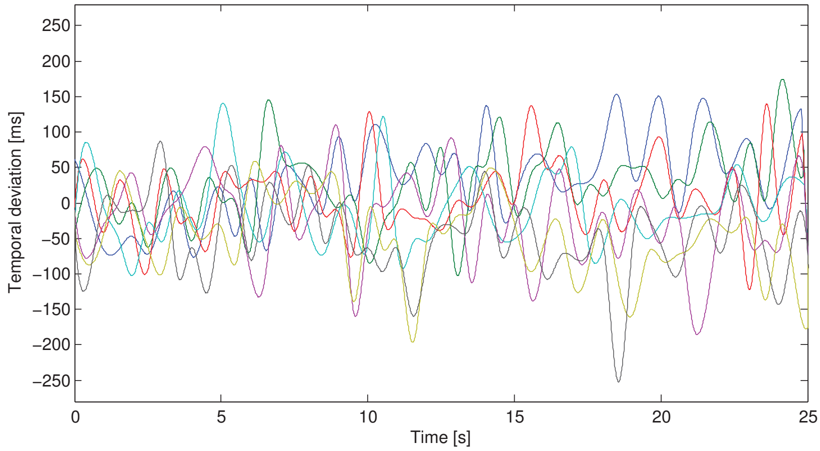



**Figure 7.2.** Visualization of the temporal deviations from the linear tempo with seven simulated violinists.

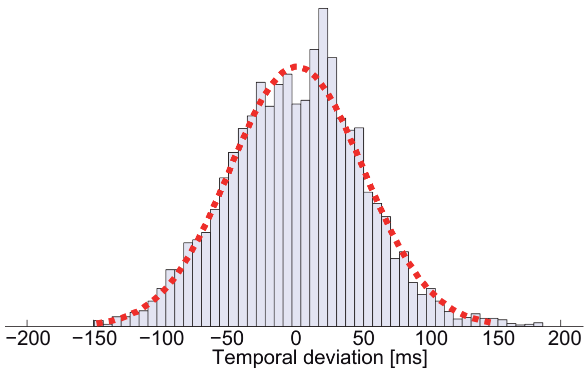



**Figure 7.3.** Example of a histogram showing the temporal deviations of seven simulated players in Fig. 7.2. Metropolis-Hastings sampling from a normal distribution with the standard deviation of 50 ms is used with spline interpolation. A normal distribution fitted to the histogram is shown with the dashed curve.

ticular distribution. An example having an interpolated 2 Hz random sampling frequency with seven violinists having a 50 ms standard deviation is shown in Fig. 7.2. The corresponding distribution for the deviations from the linear time is given in Fig. 7.3.

The time-base vector $\hat{n}_{\mathrm{r}}^{(m)}$ having a constant pitch-shift with a time-varying synchronization is obtained by simply combining $\hat{n}_{\mathrm{r}}^{(m)} = \hat{n}^{(m)} + r^{(m)}$. Finally, STFT frames $X_m(n, k)$ is sampled with the new individual time-base vector $X_m(\hat{n}_{\mathrm{r}}^{(m)}, k)$ and inverse-transformed back to time-domain, and resampled by $1/d^{(m)}$. An example of the positions for three simulated de-synchronized players is given in Fig. 7.4. Here, the temporal deviations are scaled by a factor of ten for improved visibility of the

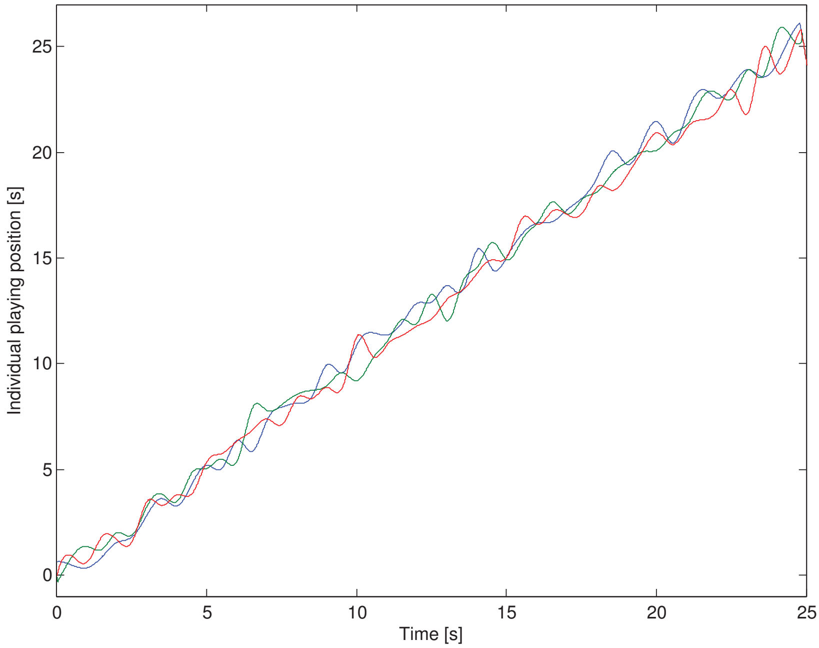



**Figure 7.4.** Temporal positions of three simulated violinists playing a 25-second excerpt. For illustrative purposes the shown synchronization deviation is ten times larger than in reality.

playing position differences. With realistic parameters the curves should be monotonically increasing. It should be noted that the decreasing segments in the shown curves would be reflected as a simulated player momentarily playing the part backwards.

Additionally, amplitude modulation can be applied to the output signals in order to simulate the varying playing dynamics. Suitable modulation curves are obtained similarly to the tempo variations above: low-frequency random series is generated, and the sum of parallel random values are scaled to unity. Low-frequency signals are then interpolated and resampled to the signal sampling rate. Hence, the amplitude modulation does not have an effect on the total signal level. Instead, only the instantaneous balance between the simulated musicians is varied.

In contrast to the simulation method proposed above, some other methods for synthesizing a section sound rely on randomizing the phases of copied signals [157]. With a time difference $\Delta t$ of the observed asynchronous note onsets the actual phase difference $\phi$ at the fundamental frequency $f_0$ is formulated as

$$\Delta t = n \cdot 1/f_0 + \phi \quad \text{, where} \tag{7.1}$$

$$\Delta t >> 1/f_0 \quad . \tag{7.2}$$

The phase being a circular function, the resulting distribution of phase is a wrapped normal distribution [50]. With large $n$, the actual steady-state phase difference can be approximated with a uniform distribution, as the time differences are wrapped to the period of the fundamental frequency.

## 7.2 Tracking of the ensemble playing

The operation of the above method requires a parameter which defines the amount of asynchronicity between simulated musicians. In the literature, an upper limit of 35 ms is suggested for delays that are not detrimental to the ensemble synchronization [120, 148]. Ensemble synchronization has been mostly studied from the perspective of auditory feedback [57]. Propositions for the ensemble timing mechanisms have been discussed in [126].

Tracking of the player synchronization is possible with multiple methods, e.g., with IR markers and cameras, motion sensors, or accelerometers. Here, two methods for tracking the temporal performance of the ensemble are presented. First, the synchronization in a professional symphony orchestra is extracted unobtrusively from a video. Second, the ensemble playing is tracked from an amateur orchestra with contact microphones attached to the violin bridges and employing note onset detection to the recorded signals. Articles yet to appear [137, 138] present these two approaches, respectively.

The approach in ensemble tracking can be thought as an assumption of the musicians playing in perfect synchronization, while the objective of the tracking is to show that there is in fact actual timing differences. For this reason the processing of the tracking data is chosen so that the obtained timing differences would not be exaggerated.

### 7.2.1 Video tracking of orchestra strings

The video tracking method is based on a high-definition video. The video was recorded from the technical bridge running across the ceiling of a concert hall. String instrument sections were framed in the image so that the instruments and the bow hands were visible in the image for all tracked players. The principal idea is to track the movements of the bow hand, and compare the differences between the time instants where the corresponding bow strokes occur. Given that the bow change indicates the intended beginning of a note, the temporal distribution of playing can be deduced.

The video tracking is performed with a kernel-based approach, where the color histogram model is first selected manually from the first frame of the tracked sequence for the back of the bow hand and the violin tailpiece (Fig. 7.5 a). For the actual tracking, the Euclidean distances of the color histogram components are calculated in the pixels surrounding the model location (Fig. 7.5 b-c). The new estimated location of the tracked kernel is at the coordinates that minimize the Euclidean distance, i.e., what area is the most similar to the previous model color histogram (Fig. 7.5 d). YUV color space is chosen due to the color separation of the skin and violin colors from the background (Fig. 7.5 e). The video resolution is $1280 \times 720$ pixels with a progressive frame rate of 59.94 fps.

The tracking model is updated after each frame as the linear combination of the given initial model and the model used in the previous frame. This procedure is selected for the reason that the orientation and shape of the bow hand model as well as the orientation of the violin change gradually during bow strokes and playing. Hence, a static model cannot be applied. Furthermore, violinist's hands can occasionally be close to each other during playing, or a skin color histogram can be close to a bright-colored background. The risk of the model beginning to track an unwanted area is reduced by restricting the deviation of the tracking model to a certain distance from the original model.

The actual bowing action is calculated from the distance between the tracked hand and violin positions. In theory, the bow stroke is indicated by a zero crossing of the first derivative of the intra-hand-violin distance. However, the direct derivative produces excessive number of false detections for the bow direction change. Therefore a heuristic algorithm is applied to the raw distance data. First, local minima and maxima are sought within a 200 ms window for increased robustness against momentary stops in hand movement. Then, the tracked player with the least detected bow changes is selected as the reference. Hence, the number of false detections is minimized. Finally, the bow changes of other tracked players are matched to the reference. It should be noted that with the

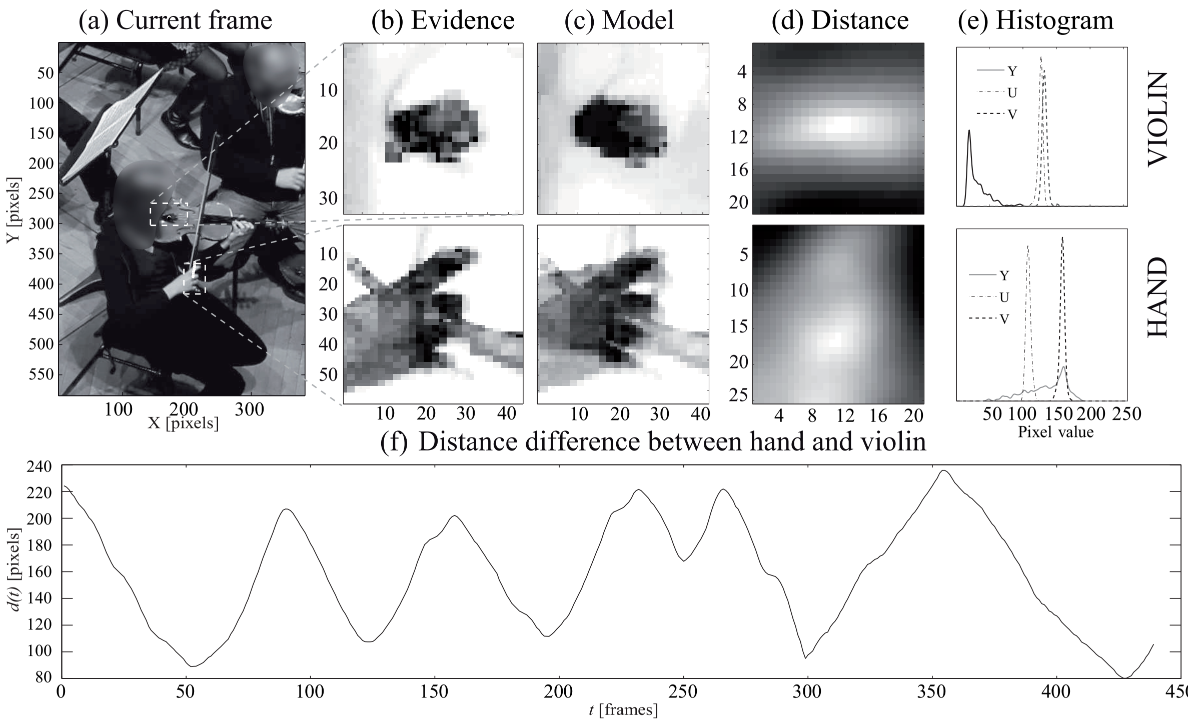



**Figure 7.5.** Tracking example. (a) the video frame; (b) an optimal kernel area in the frame providing the lowest distance measure; (c) the current kernel with the fixed color histogram model; (d) the distance measure map between the model and the current kernel; (e) the color histogram models; and (f) the tracked bow distance from the violin.

matched time indices for the detected bow strokes, the algorithm provides the best-case results. Subsequently, the parameters for the observed temporal distribution of the bow strokes can be calculated.

Excerpts of Prokofiev's Symphony no. 5 was studied for three violins, three violas and three cellos, played by a professional symphony orchestra in a dress rehearsal before the concert. The players were chosen only by their good visibility that allowed error-free tracking. The results presented in Table 7.1 suggest that the data for the first violins and celli is normally distributed, as the Anderson-Darling test values are lower than $1.092$ with a 99% confidence interval. Approximately half of the bow strokes occur within a 60 ms time window. The standard deviation for the temporal bowing differences is approximately 40-50 ms. In addition, an excerpt of Dvořák's Cello concerto were investigated with a wider view spanning the entire first violin section, where seven violinists could be tracked without obstructions. Anderson-Darling test yields a value of 62.3, which could partially result from the lower spatial resolution per player. The temporal standard deviation of the analyzed bow strokes is higher than with other excerpts, 82.1 ms. On the other hand, being longer than the Prokofiev excerpt, the Dvořák excerpt contains more bow strokes of varying speed.

Overall, from the video tracking results it can be concluded that the standard deviation of the strings synchronization is in the magnitude of 40-80 ms.

### 7.2.2 Audio tracking with contact microphones

An alternative method for estimating the timing differences is to use note onset detection. This is done by attaching contact microphones to the bridge of each instrument in a section and detecting the note onset from the captured signals. Contact microphones are beneficial for greatly reducing crosstalk between recorded channels. This method is explained in detail in [137].

The onset detection is implemented with the spectral difference (SD) method presented in [11]. Spectral difference is calculated by

$$f(n)^{(m)} = \sum_k [H(\|Y^{(m)}(n,k)\| - \|Y^{(m)}(n-1,k)\|)]^2 \qquad (7.3)$$

where $H(y) = (y + \|y\|)/2$, $Y(n,k)$ is the short-time Fourier transform of the signal at time index $n$ and discrete frequency bin $k$. The onset is detected by using 85 ms time windows with 0.5 ms steps. Averaging smoothing filter of a 25 ms window is applied, as suggested in [11]. The local maxima of the SD detection function $f(n)^{(m)}$ yield the detected onsets.

Similarly to the video tracking described above, a reference and grouping of the detected note onsets is required. The timing reference is created from the detected onsets with individual musicians as the combined likelihood

$$F(n) = \Pi_{m=1}^{M} f(n)^{(m)} \tag{7.4}$$

where $M$ is the number of analyzed musicians. The combined likelihood function is also filtered as the individual detection functions, and the local maxima of the function are selected as the reference. The detected onsets are grouped with respect to the reference. Detected onsets within $\pm 150$ ms of the reference are considered to belong to the same group, that is, the same note. If multiple onsets exist for a single musician, the closest one to the reference is selected, giving the minimum temporal variance.

Each detected onset $o_t$ at reference onset $t$ is normalized with respect to the group normal, i.e.,

$$\hat{o}_t^{(m)} = o_t^{(m)} - 1/M \sum_{m=1}^{M} o_t^{(m)} \quad . \tag{7.5}$$

The final distribution is calculated over all the normalized groups. Missing data, i.e. undetected onsets, are simply neglected from the results.

The onsets are calculated for a section of 12 violins of an amateur orchestra. The selected passage is from Sibelius' Symphony No. 3, I movement, bars 167-181. Although the investigated orchestra was not professional, the chosen passage is relatively easy, and it was played with a good tone and intonation. The statistical values for the obtained distribution are shown in the last row in Table 7.1. The results for the tracked passage show that the temporal differences between the note onsets of the violin players are approximately normally distributed with a 40 ms standard deviation. While the resulting distribution depends on the chosen limits for the onset grouping, standard deviations from the video and audio tracking methods are approximately 40-50 ms.

## 7.3 Subjective evaluation of the string section synthesis

The authenticity of the proposed method for simulating the violin section sound was subjectively evaluated with a listening test where eleven sub-

**Table 7.1.** Statistical properties of the analyzed excerpts with video and audio tracking. STD and AD stand for the standard deviation Anderson-Darling test, respectively. The respective percentiles are indicated in in milliseconds. Approximate tempi are given in quarter notes per minute. Results from the Dvořák excerpt are not entirely comparable due to the lower video resolution.

| Method | Excerpt | Section | Tempo | STD | AD | 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|---|
| Video | Prokofiev | I Violins | 90 | 40.1 | 0.22 | -98 | -36 | 0 | 29 | 102 |
| | | Violas | 75 | 48.7 | 1.48 | -133 | -39 | 3 | 30 | 179 |
| | | Cellos | 37.8 | 45 | 0.95 | -86 | -33 | 3 | 25 | 103 |
| | Dvořák | I Violins | 105 | 82.1 | 62.3 | -140 | -24 | 0 | 27 | 142 |
| Audio | Sibelius | I Violins | 118 | 40.4 | 5.0 | -102 | -18 | 0 | 19 | 84 |

jects rated the section sound processed with methods. All subjects had a background on acoustics and/or signal processing.

### 7.3.1 Test setup

Two excerpts from the anechoic violin recordings discussed earlier were utilized for evaluating the processing method. Short, six-second violin passages from Mahler's and Beethoven's works were selected to represent typical orchestral repertoire (1st Symphony, IV movement, II violin, bars 57-61, and 7th Symphony, I movement, I violin, bars 14-15, respectively).

Five conditions were created from the anechoic signals. First, an unprocessed recording in one direction was taken to represent a solo violin performance as a reference. Second, 11 copies of the same recording were processed with individual chorus effects for creating an impression of a violin section. Third, the currently proposed method was similarly applied to the original recording. The last two conditions incorporated methods for altering the frequency response, i.e. timbre, in a manner of unique violins [83]. Fourth, instead of one microphone signal, recordings from 11 different directions were processed with the proposed method. Fifth, the differences in the frequency responses of different violins were experimentally simulated with 11 filters whose magnitudes at the four Dünnwald bands were randomized between $\pm 6$ dB [56]. After filtering the anechoic violin signal in one direction, the proposed processing was applied also here. In total, one solo performance and four violin sections with one original and 11 simulated players were obtained.

The chorus effect for each violin copy in the second condition was implemented as a linearly interpolated variable-length delay line without feedback, as presented in [194]. The parameters for the chorus effect were chosen iteratively in order to produce a desired impression within the abilities of the algorithm. The randomized delay lengths for the individual copies were between 0-25 ms. Modulation signals were low-pass filtered white noise with the cutoff frequency at 3 Hz. Modulation depth was 1.3 ms. Lower values were considered introducing too small differences between the copies and higher values for the modulation yielded unnaturally fuzzy results. These values fall within the guidelines in the literature [40, 194].

In the proposed processing method, the pitch shifts of the individual copies were distributed within $\pm$ 10 cents as presented in [101]. Temporal variation followed a normal distribution having a 45 ms standard deviation, as suggested by the tracking results above. For the normally distributed amplitude modulation, the standard deviation was 1 dB with 5 Hz modulation frequency, which corresponds approximately to eighth notes in moderate tempo.

The final stimuli were created by convolving the dry signals with impulse responses. The spatial impulse responses were measured in an unoccupied concert hall with the loudspeaker orchestra and a GRAS 3-D microphone probe. Source numbers 1-3 were utilized (see Fig. 4.1). For each source, the spatial response was rendered into two virtual cardioid microphones as a coincident XY pair with 90 degree separation for headphone listening. The processed copies of the anechoic recording were distributed evenly to the three source channels and convolved with the corresponding impulse responses. The levels of the convolved signals were equalized with A-weighting.

The subjects were asked to assess the perceived impression of a string section on a continuous linear scale. The end points of the scale were "one or few individual instruments" and "large section with many instruments". The subjects were instructed before the test that in an authentic section the individual instruments are not perceived as such. Instead, they are blended together, yet without artifacts or artificial coloration. The subjects were allowed to familiarize themselves with the signals and the test procedure before the test. The test for each condition and signal was repeated three times in a fully random order. The test was conducted in a quiet, acoustically treated listening room. The convolved stimuli were presented to subjects with Sennheiser HD650 headphones.

### 7.3.2 Results

Each processing condition was evaluated 66 times (2 signals $\times$ 3 repeats $\times$ 11 subjects). The results were analyzed with ANOVA having four factors: method, music, repeat, and subject. The results for the processing method are shown in Fig. 7.6. Higher location on the vertical axis indicates a more convincing simulation of the section sound. The differences between all five conditions were significant ($F(4,329) = 182.18, p = 0$). Solo violin condition received expectedly the lowest rating. All three variations of the proposed method were assessed to give an impression of a larger instrument section than the applied chorus effect. Utilizing dif-

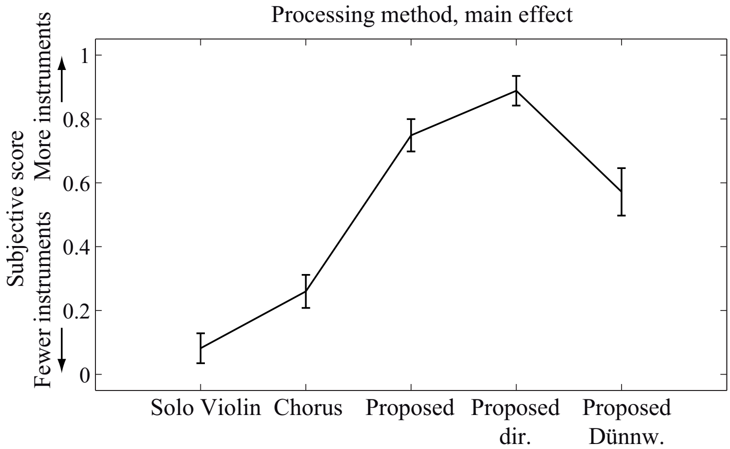



**Figure 7.6.** Results of the listening test shown with means and 95% confidence intervals. Higher score indicates a section perceived having more players. ”Proposed dir.” stands for the proposed method where the simulated players employ anechoic violin signals recorded in different directions. ”Proposed Dünnw.” stands for the method where the signal frequency responses have been altered with filters following the Dünnwald bands.

ferent microphone directions improved the impression compared to the processing applied to a signal from one microphone. However, altering the violin frequency responses at Dünnwald bands was not assessed better. The variations between the subjects, repetitions, or signals did not show significant differences.

## 7.4 Discussion

The discussed method for simulating section sound modifies the recorded signal by introducing a constant pitch shift, varying time differences, and amplitude modulation. The proposed method has been applied to the recordings described in Publication II and the results from the section simulation have been used in numerous auralizations of measured concert halls and simulated spaces. During the course of earlier studies in [101, 108, 110] and Publication I, the applied processing has advanced gradually to the method discussed in previous sections. Compared with the preceding stages in simulating the section sound, the present method has been noticed to considerably increase the warmth and softness naturally exhibited by a string section. This is supported by the listening test results. Furthermore, the utilization of different recording directions for the instrument signals provides the strings simulation with another noticeable improvement.

The spread in the pitch-shift has been adjusted by ad-hoc listening. Iteratively, suitable settings were found with a maximum detuning of $\pm$ 10 to 12 cents. For the Mozart and Beethoven excerpts, the $\pm$ 10 cent pitch-shift range can be considered adequate. In contrast, the late Romantic period works by Bruckner and Mahler can support a wider spread up to approximately $\pm$ 12 cents. Excess values for the detuning are easily heard as an orchestra of low quality. The found values correspond to the same magnitude as cited in the literature for the frequency spread in a string section [117, 120] or in a choir [157].

Temporal variations that have a standard deviation in the magnitude of the obtained tracking results are not sensible to be used with the traditional chorus effect. While there is no limit for the constant delay length, the usable magnitudes for the tap point modulation are considered too low to introduce naturally occurring timing differences.

# 8. Summary

Studies on the sources and signals related to the room-acoustic evaluation and performance of music have been presented in this thesis. The loudspeaker orchestra has been proposed as a method for evaluating concert halls with a repeatable, orchestral-like source. Comprehensive anechoic recordings of orchestral music have been presented as the signals for the loudspeaker orchestra. The recording system has been utilized for thoroughly investigating the directivities of the orchestra instruments. Additionally, the directivities of balloon bursts have been presented with regard to their applicability to a source in impulse source measurements. Also, the temporal distributions of the orchestra string players have been studied. Utilizing the temporal information, this thesis has presented a method for enhancing the section sound of string instrument recordings and improving the naturalness of the loudspeaker orchestra.

## 8.1 Main results of the thesis

A summary of the main results and findings in the current thesis are listed as follows:

- Woodwind and string instrument directivities with different tones correspond to their average radiation patterns at the exited harmonic frequencies. The directivity changes considerably with the played tone. The directivities can be modeled with directivity filters for obtaining the correct power response, but time-invariant filtering would not take into account the directional tone color effect.

- A combination of two two-way loudspeakers facing selected directions decreases the directivity difference between a single loudspeaker and a violin.

- Anechoic recordings of orchestral music can be carried out with good quality by recording the instruments one at a time. A reference video of a conductor and piano track containing the reduced score enables professional musicians to hold a mutual synchronization between individual recording sessions.

- The temporal deviation in the synchronization of bow strokes in orchestras follows the normal distribution in most cases. The standard deviation of the deviations from the group average is in the magnitude of 40-80 ms and the (25 75)-percentiles approximately $\pm$ 20-30 ms.

- The proposed method for simulating the sound of an instrument section by introducing pitch-shift, time-varying temporal variation, and amplitude modulation is more efficient than the traditional chorus effect. The naturalness of the section simulation can be further improved by utilizing the directivity properties of the instruments.

- Balloon bursts exhibit a repeatable impulse whose directivity changes as a function of frequency. The frequency-dependency follows the balloon size, but the shape of the frequency response curve remains nearly constant between similar balloons.

## 8.2 Future work

The topics presented in this thesis suggest a number of avenues for future research, some of which are the following:

- The calculation of the total directivity of a symphony orchestra using the data from the instrument directivity measurements.

- Further development of the loudspeaker orchestra by calculating the total radiation pattern with regard to the measured instruments, and distributing the anechoic signals to the loudspeakers with a more advanced method.

- Supplementary investigations of the methods for improving the naturalness of the anechoic recordings, e.g., by applying directivity filters as a function of a played note.

- Measurement of the proposed loudspeaker orchestra array with a loudspeaker having an adjustable directivity pattern, and the detailed evaluation of the loudspeaker orchestra performance.

# Bibliography


[1] J.S. Abel, N.J. Bryan, P.P. Huang, M. Kolar, and B.V. Pentcheva. Estimating room impulse responses from recorded balloon pops. In *the 129th Convention of the Audio Engineering Society*, Nov 2010. paper no. 8171.

[2] J.B. Allen and D.A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am*, 65(4):943–950, Apr. 1979.

[3] J. B. Arban. *Complete conservatory method for trumpet*. Carl Fischer, L.L.C., London, 1st edition, 1982.

[4] A. Askenfelt. Eigenmodes and tone quality of the double bass. *Dept. for Speech, Music and Hearing, Quarterly progress and status report*, 23(4):149–174, Jun. 1982.

[5] J. Backus. Input impedance curves for the reed woodwind instruments. *J. Acoust. Soc. Am*, 56(4):1266–1279, 1974.

[6] Bang & Olufsen. CD B&O 101 Music for Archimedes, 1992.

[7] M. Barron. The subjective effects of first reflections in concert halls–The need for lateral reflections. *J. Sound. Vib.*, 15(4):475–494, 1971.

[8] M. Barron and A.H. Marshall. Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. *J. Sound. Vib.*, 77(2):211–232, 1981.

[9] B. Bauer. Some techniques toward better stereophnic perspective. *IEEE Transactions on Audio*, 11(3):88–92, 1963.

[10] D.R. Begault, E.M. Wenzel, and M.R. Anderson. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *J. Audio Eng. Soc.*, 49(10):904–916, 2001.

[11] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler. A tutorial on onset detection in music signals. *IEEE Transactions on Speech and Audio Processing*, 13(5):1035–1047, 2005.

[12] A. H. Benade. *Fundamentals of music acoustics*. Oxford Universty Press, London, 1st edition, 1976.

[13] A. H. Benade. Front instrument to ear in a room: direct or via recording. *J. Audio Eng. Soc.*, 33(4):218–233, Apr. 1985.

[14] L. Beranek. *Concert and Opera Halls: music, acoustics, and architecture*. Springer, New York, NY, USA, 2004.

[15] F. W. Bianchi and R. H. Campbell. The virtual orchestra: technical and creative issues. *Journal of Sound and Vibration*, 232(1):275–279, 2000.

[16] G. Bissinger. A0 and A1 coupling, arching, rib height, and f-hole geometry dependence in the 2 degree-of-freedom network model of violin cavity modes. *J. Acoust. Soc. Am*, 104(6):3608–3615, 1998.

[17] G. Bissinger and J. Keiffer. Radiation damping, efficiency, and directivity for violin normal modes below 4 kHz. *Acoustics Research Letters Online*, 4(1):7–12, 2003.

[18] J. Blauert. *Spatial Hearing. The psychophysics of human sound localization*. MIT Press, Cambridge, MA, USA, 2nd edition, 1997.

[19] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc Am*, 75(6):1827–1836, 1984.

[20] J. Borish and J.B. Angell. An efficient algorithm for measuring the impulse response using pseudorandom noise. *J. Audio Eng. Soc.*, 31(7):478–488, 1983.

[21] J. S. Bradley. Auditorium acoustics measured from pistol shots. *J. Acoust. Soc. Am.*, 80(1):199–205, Jul. 1986.

[22] Bristol University, Deparment of Music. Loudspeaker Orchestra, 2010. URL: http://www.bristol.ac.uk/music/facilities/studios/bulo.html [Cited Dec. 8, 2010].

[23] A. W. Brown. *Acoustical studies on the flat-backed and round-backed double bass*. Doctoral thesis, Universität für Music und darstellende Kunst Wien, 2004. URL: http://iwk.mdw.ac.at/Forschung/pdf_dateien/2004_Brown_WEB_Dissertation.pdf .

[24] Bruel & Kjaer SV. Type 4292 and Type 4295 datasheets, 2010. URL: http://www.bksv.com/Products/SoundSources/ [Cited Jan. 20, 2011].

[25] A. Buen. How dry do the recordings for auralization need to be? In *Proc. Institute of Acoustics*, volume 30, pages 107–114, 2008.

[26] M. Campbell, C. Greated, and A. Myers. *Musical Instruments*. Oxford University Press, New York, NY, USA, 2004.

[27] R. Causse, J. Bresciani, and O. Warusfel. Radiation of musical instruments and control of reproduction with loudspeakers. In *Proc. of ISMA*, Tokyo, Japan, 1992.

[28] D. J. Chéenne, M. Ardila C. G. Lee, and B. Bridgewater. A qualitative and quantitative analysis of impulse responses from balloon bursts. In *Proc. Acoustics '08*, Paris, France, Jun. 29 - Jul. 4 2008. paper no. 5pAAh8.

[29] W.H. Chiang and J.C. Huang. Subjective Evaluation of Acoustical Environments for Solo Performance. *J. Building Acoustics*, 6(1):17–36, 1999.

[30] S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *American Statistician*, 49(4):327–335, 1995.

[31] Y. J. Choi and F. R. Fricke. A Comparison of Subjective Assessments of Recorded Music and Computer Simulated Auralizations in Two Auditoria. *Acta Acustica united with Acustica*, 92(4):604–611, 2006.

[32] Y.J. Choi and F.R. Fricke. Evaluation of the Relative Acoustic Performance of Two Auditoria Using Measurements and Auralization. *Acta Acustica united with Acustica*, 91(6):1051–1062, 2005.

[33] W.T. Chu. Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. *J. Acoust. Soc. Am*, 63(5):1444–1450, 1978.

[34] P.R. Cook and D. Trueman. A database of measured musical instrument body radiation impulse responses, and computer applications for exploring and utilizing the measured filter functions. In *Proc. Int. Symposium on Musical Acoustics*, pages 303–308, Woodbury, NY, USA, 1998.

[35] T.J. Cox and P. D'Antonio. *Acoustic absorbers and diffusers: theory, design and application*. Spon Press, London, United Kingdom, 2004.

[36] L. Cremer. *The physics of the violin*. The MIT Press, Cambridge, Massachusetts, USA, 1st edition, 1984.

[37] B.-I. Dalenbäck, M. Kleiner, and U. P. Svensson. Audibility of changes in geometric shape, source directivity, and absorptive treatment – experiments in auralization. *J. Audio Eng. Soc.*, 41(11):905–913, Nov. 1993.

[38] D.-I. Dalenbäck. CATT-Acoustic. URL: http://www.catt.se/, Feb. 2011. (Cited Feb. 4 2011).

[39] J.J. Dammerud. *Stage acoustics fo symphony orchestras in concert halls*. PhD thesis, University of Bath, Bath, United Kingdom, Sep. 2009.

[40] J. Dattorro. Effect Design, Part 2: Delay-Line Modulation and Chorus. *J. Audio Eng. Soc.*, 45(10):764–788, 1997.

[41] D.T. Deihl and F.R. Carlson Jr. "N Waves" from Bursting Balloons. *American Journal of Physics*, 36(5):441, 1968.

[42] Denon. Anechoic orchestral music recording. Audio CD, Denon Records, 1995. ASIN: B0000034M9.

[43] J. Eargle. *Handbook of recording engineering*. Springer Verlag, New York, NY, USA, 4 edition, 2002.

[44] T.D. Rossing (ed.). *Handbook of acoustics*. Springer, New York, NY, USA, 2007.

[45] D. P. W. Ellis. A phase vocoder in Matlab, 2002. URL: http://www.ee.columbia.edu/ dpwe/resources/matlab/pvoc/.

[46] A. Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *the 108th Audio Engineering Society Convention*, Paris, France, Feb. 19-22 2000. preprint no. 5093.

[47] A. Farina and R. Ayalon. Recording concert hall acoustics for posterity. In *the 24th Audio Engineering Society International Conference*, Jun 2003. paper no. 38.

[48] A. Farina, A. Langhoff, and L. Tronchin. Realisation of “virtual” musical instruments: measurements of the Impulse Response of Violins using MLS technique. In *Proc. of CIARM 95*, Ferrara, Italy, May 1985.

[49] P. Fausti and A. Farina. Acoustic measurements in opera houses: comparison between different techniques and equipment. *Journal of Sound and Vibration*, 232(1):213–229, 2000.

[50] N.I. Fisher. *Statistical analysis of circular data*. Cambridge Univ Press, Australia, 1993.

[51] N. H. Fletcher. Acoustical correlates of flute performance technique. *J. Acoust. Soc. Am.*, 57(1):233–237, Jan. 1975.

[52] N. H. Fletcher and T. D. Rossing. *The Physics of Musical Instruments*. Springer, New York, 2nd edition, 1998.

[53] N. H. Fletcher and A. Tarnopolsky. Blowing pressure, power, and spectrum in trumpet playing. *J. Acoust. Soc. Am.*, 105(2):874–881, Feb. 1999.

[54] J.D. Foley. *Computer graphics: principles and practice*. Addison-Wesley, Reading, MA, USA, 1995.

[55] R. Freiheit. Creating an anechoic choral recording. In *the International Symposium on Room Acoustics (ISRA2010)*, Melbourne, Australia, Aug 29-31 2010.

[56] C. Fritz, I Cross, B.C.J. Moore, and J. Woodhouse. Perceptual thrsholds for detecting modifications applied to the acoustical properties of a violin. *J. Acoust. Soc. Am.*, 122(6):3640–3650, 2007.

[57] A. Gabrielsson. The performance of music. *The psychology of music*, 2:501–602, 1999.

[58] A.C. Gade. Investigations of musicians’ room acoustic conditions in concert halls. Part I: Methods and laboratory experiments. *Acoustica united with Acta Acustica*, 69:193–203, 1989.

[59] A.C. Gade. Investigations of musicians’ room acoustic conditions in concert halls. Part II: Field experiments and synthesis of results. *Acoustica united with Acta Acustica*, 69:249–262, 1989.

[60] M. Garai. Measurement of the sound-absorption coefficient in situ: the reflection method using periodic pseudo-random sequences of maximum length. *Applied Acoustics*, 39(1-2):119–139, 1993.

[61] Genelec Oy. 1029A specifications, 2011. URL: http://www.genelec.fi/documents/datasheets/DS1029a.pdf [Cited Dec. 9, 2010].

[62] Genelec Oy. 1030A specifications, 2011. URL: http://www.genelec.fi/documents/datasheets/DS1030a.pdf [Cited Feb. 24, 2011].

[63] M.A. Gerzon. Recording concert hall acoustics for posterity. *J. Audio Eng. Soc*, 23(7):569–571, 1975.

[64] M.A. Gerzon. Ambisonics in multichannel broadcasting and video. *J. Audio Eng. Soc*, 33(11):859–871, 1985.

[65] F. Giron. *Investigations about the Directivity of Sound Sources*. PhD thesis, Ruhr-Universität Bochum, Verlag Shaker, Aachen, Germany, 1996.

[66] D. Griesinger. How loud is my reverberation? In *the 98th Audio Engineering Society Convention*, Feb 1995. paper no. 3943.

[67] D. Griesinger. Beyond MLS, occupied hall measurement with FFT techniques. In *the 101st Audio Engineering Society Convention*, 1996. Preprint n. 4403.

[68] D. Hammershøi and H. Møller. Methods for binaural recording and reproduction. *Acta Acustica united with Acustica*, 88(3):303–311, May/June 2002.

[69] V. Hansen and G. Munch. Making recordings for simulation tests in the Archimedes project. *J. Audio Eng. Soc.*, 39(10):768–774, Oct 1991.

[70] Harvard University Studio for Electroacoustic Composition. Loudspeaker orchestra, 2010. URL:
http://huseac.fas.harvard.edu/4-hydra/hydra.html [Cited Dec. 8, 2010].

[71] R.J. Hawkes and H. Douglas. Subjective acoustic experience in concert auditoria. *Acustica*, 24:235–250, May 1971.

[72] T. Hidaka and L.L. Beranek. Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas. *J. Acoust. Soc. Am*, 107:368–383, 2000.

[73] T. Hidaka, K. Kageyama, and S. Masuda. Recording of anechoic orchestral music and measurement of its physical characteristics based on the autocorrelation function. *Acoustica united with Acta Acustica*, 67:68–70, 1988.

[74] F. Hohl. *Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten*. Master's Thesis, Institut für Elektronische Musik und Akustik, Universität für Musik und darstellende Kunst, Graz, Austria, 2009.

[75] M. Horvat, K. Jambrosic, and H. Domitrovic. A comparison of impulse-like sources to be used in reverberation time measurements. In *Proc. Acoustics '08*, pages 4503–4508, Paris, France, Jun. 29 - Jul. 4 2008.

[76] C. M. Hutchins. A 30-year experiment in the acoustical and musical development of violin-family instruments. *J. Acoust. Soc. Am.*, 92(2):639–650, Aug. 1992.

[77] C. M. Hutchins. The air and wood modes of the violin. *J. Audio Eng. Soc.*, 46(9):751–765, Sep. 1998.

[78] ISO Standard 3382-1:2009. *Acoustics — Measurement of the room acoustic parameters – Part I: Performance spaces*. International Standards Organization, 2009.

[79] ISO Standard 3745:2003. *Acoustics — Determination of sound power levels of noise sources using sound pressure — Precision methods for anechoic and hemi-anechoic rooms*. International Standards Organization, 2003.

[80] R. Jacques, B. Albrecht, D. de Vries, F. Melchior, and H.-P. Schade. Multichannel source directivity recording in an anechoic chamber and in a studio. In *Proc. Forum Acusticum 2005*, pages 479–484, Budapest, Hungary, Aug. 29 - Sept. 2 2005.

[81] K. Jambrosic, M. Horvat, and H. Domitrovic. Reverberation time measuring methods. In *Proc. Acoustics '08*, Paris, France, Jun. 29 - Jul. 4 2008. paper no. 4pAAh10.

[82] A. James. Results of the NPL study into comparative room acoustic measurement techniques part 1, reverberation time in large rooms. In *Proc. Institute of Acoustics*, volume 25, 2003.

[83] E. V. Jansson. Admittance measurements of 25 high quality violins. *Acta Acustica united with Acustica*, 83(2):337–341, Mar. 1997.

[84] E. Kahle and J.-P. Juillien. Subjective listening tests in concert halls: methodology and results. In *The Int. Congr. Acoustics*, pages 441–444, Trondheim, Norway, Jun. 1995.

[85] D. Kahlin and S. Ternstrom. The chorus effect revisited-experiments in frequency-domain analysis and simulation of ensemble sounds. In *Proc. 25th EUROMICRO Conference*, volume 2, pages 75–80, 1999.

[86] S. M. Kay. *Fundamentals of statistical signal processing, Volume I. Estimation theory*, chapter 8. Least squares. Prentice-Hall, Eaglewood Cliffs, NJ, USA, 1st edition, 1993.

[87] M. Kleiner, B.-I. Dalenbäck, and U. P. Svensson. Auralization – an overview. *J. Audio Eng. Soc.*, 41(11):861–875, Nov. 1993.

[88] M. Kleiner, R. Orlowski, and J. Kirszenstein. A comparison between results from a physical scale model and a computer image source model for architectural acoustics. *Applied Acoustics*, 38(2-4, Special Issue on Computer Modelling and Auralisation of Sound Fields in Rooms):245–265, 1993.

[89] A. Korkstad and U. Kristiansen. Statistical models of sound fields in rooms based on calculating early reflections by ray tracing or image sources. In *Proc. 14th Int. Congr. Acoust.*, Beijing, China, July 1992. paper F2-11.

[90] U.R. Kristiansen, A. Krokstad, and T. Follestad. Extending the image method to higher-order reflections. *Applied Acoustics*, 38(2-4):195–206, 1993.

[91] A. Krokstad, S. Strom, and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. *J. Sound Vib.*, 8(1):118–125, 1968.

[92] H. Kuttruff. Auralization of impulse responses modelled on the basis of ray-tracing results. In *the 91st Audio Engineering Society Convention*, New York, NY, USA, Oct. 4-8 1991. preprint no. 3122.

[93] H. Kuttruff. *Room Acoustics*. Elsevier Science Publishers, Essex, England, 3rd edition, 1991.

[94] M.V. Laitinen and V. Pulkki. Binaural reproduction for directional audio coding. In *IEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 337–340, 2009.

[95] Y.W. Lam. A comparison of three diffuse reflection modeling methods used in room acoustics computer models. *J. Acoust. Soc. Am*, 100(4):2181–2192, Oct 1996.

[96] M. J. R. Lamothe and J. S. Bradley. Acoustical characteristics of guns as impulse sources. *Canadian Acoustics*, 13(2):16–24, Apr. 1985.

[97] J. Laroche and M. Dolson. New phase-vocoder techniques for pitch-shifting, harmonizing and other exotic effects. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 91–94, Oct. 17-20 2002.

[98] P. R. Lehman. Harmonic structure of the tone of the bassoon. *J. Acoust. Soc. Am.*, 36(9):1649–1653, Sept. 1964.

[99] T. Lokki. *Physically-based Auralization - Design, Implementation, and Evaluation*. PhD thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, report TML-A5, 2002. Available at http://lib.hut.fi/Diss/2002/isbn9512261588/.

[100] T. Lokki. How many point sources is needed to represent strings in auralization? In *the International Symposium on Room Acoustics (ISRA'2007)*, Seville, Spain, Sep. 10-12 2007. Paper P11.

[101] T. Lokki and J. Pätynen. Applying anechoic recordings in auralization. In *The EAA Symposium on Auralization*, Espoo, Finland, Jun. 15-17 2009.

[102] T. Lokki, J. Patynen, S. Tervo, S. Siltanen, and L. Savioja. Engaging concert hall acoustics is made up of temporal envelope preserving reflections. *J. Acoust. Soc. Am.*, 129(6):EL223–228, Jun. 2011.

[103] T. Lokki, J. Pätynen, and H. Vertanen. Suomalaisten konserttisalien vertailu aistinvaraisin menetelmin. In *Proceedings of Akustiikkapäivät 2011*, pages 1–6, Tampere, Finland, May 11-12 2011.

[104] T. Lokki and K. Puolamäki. Canonical analysis of individual vocabulary profiling data. In *Proc. Second International Workshop on Quality of Multimedia Experience*, pages 152–157, Trondheim, Norway, Jun. 21-23 2010.

[105] T. Lokki and L. Savioja. Evaluation of auralization results. In *Forum Acusticum 2005*, Budapest, Hungary, Aug. 29 - Sep. 2 2005. Paper 471-0.

[106] T. Lokki and L. Savioja. State-of-the-art in auralization of concert hall models — What is still missing? In *Proc. Joint Baltic-Nordic Acoustical Meeting*, Reykjavik, Iceland, Aug. 17-19 2008.

[107] T. Lokki, UP Svensson, and L. Savioja. An efficient auralization of edge diffraction. In *the 21st Audio Engineering Society Internation Conference*, pages 317–325, St. Petersburg, Russia, Jun 1-3 2002.

[108] T. Lokki, H. Vertanen, A. Kuusinen, J. äatynen, and S. Tervo. Concert hall acoustics assessment with individually elicited attributes. *J. Acoust. Soc. Am.*, 130(2):835–849, 2011.

[109] T. Lokki, H. Vertanen, A. Kuusinen, J. Pätynen, and S. Tervo. Auditorium acoustics assessment with sensory evaluation methods. In *the International Symposium on Room Acoustics (ISRA2010)*, Melbourne, Australia, Aug 29-31 2010.

[110] T. Lokki, H. Vertanen, and S. Siltanen. Intuitive hand gestures in measurement of the perceived size of an auditory image of a symphony orchestra. In *the 38th International Audio Engineering Society Conference*, pages 111–120, Piteå, Sweden, Jun 13-15 2010.

[111] D. Luce and M. Clark. Physical correlates of brass-instrument tones. *J. Acoust. Soc. Am.*, 42(6):1232–1243, Jun. 1967.

[112] K. Maki, T. Kimura, and M. Katsumoto. Reproduction of sound radiation directivities of musical instruments by a spherical loudspeaker with multiple transducers. In *Proc. 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry*, pages 85–88, 2010.

[113] A.H. Marshall and J. Meyer. The directivity and auditory impressions of singers. *Acustica*, 58(3):130–140, 1985.

[114] K. D. Marshall. Modal analysis of a violin. *J. Acoust. Soc. Am.*, 77(2):695–709, Feb. 1985.

[115] D. W. Martin. Directivity and the acoustics spectra of brass wind instruments. *J. Acoust. Soc. Am.*, 13(3):309–313, Jan. 1942.

[116] Stijn Mattheij. String ensemble vibrato: A spectroscopic study. In *Audio Engineering Society Convention 128*, London, United Kingdom, May 22-25 2010. paper no. 8016.

[117] M. Mellody and G.H. Wakefield. The time-frequency characteristics of violin vibrato: Modal distribution analysis and synthesis. *J. Acoust. Soc. Am.*, 107:598–611, 2000.

[118] J. Merimaa and V. Pulkki. Spatial impulse response rendering I: Analysis and synthesis. *J. Audio Eng. Soc.*, 53(12):1115–1127, 2005.

[119] J. Meyer. *Acoustics and the Performance of Music*. Verlag das Musikinstrument, Frankfurt/Main, 1978.

[120] J. Meyer. *Acoustics and the Performance of Music*. Springer, New York, NY, USA, 2009.

[121] Miso Music Portugal. Loudspeaker Orchestra, 2010. URL: http://www.misomusic.com/ingl/circul/concprop/ol_main.htm [Cited Dec. 8, 2010].

[122] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36(3-4):171–218, 1992.

[123] B.C.J. Moore, R.W. Peters, and B.R. Glasberg. Auditory filter shapes at low center frequencies. *J. Acoust. Soc. Am.*, 88(1):132–140, July 1990.

[124] C. Nachbar, G. Nistelberger, and F. Zotter. Listening to the direct sound of musical instruments in freely adjustable surrounding directions. In *Proc. 2nd International Symposium on Ambisonics and Spherical Acoustics*, Paris, France, May 6-7 2010.

[125] A. Nash. On the acoustical characteristics of a balloon. In *International Symposium on Room Acoustics 2007*, Seville, Spain, Sep. 10-12 2007. paper no. s07.

[126] G. M. Naylor. The achievement of ensemble. *Applied Acoustics*, 23:109–120, 1988.

[127] Norsonic. Nor276 Dodecahedron Loudspeaker datasheet, 2010. URL: http://www.norsonic.com/ [Cited Jan. 20, 2011].

[128] Odeon A/S. Odeon Room Acoustics Software. URL: http://www.odeon.dk/, Feb. 2011. (Cited Feb. 4 2011).

[129] K. Oguchi, S. Ikeda, and M. Nagata. Application of computer simulation and scale model testing to room acoustical design. In *the 89th Audio Engineering Society Convention*, Sep 1990. preprint no. 2991.

[130] K. Oguchi, S. Ikeda, and M. Nagata. Application of binaural hearing to scale-model testing. *J. Audio Eng. Soc.*, 41(11):931–938, Nov. 1993.

[131] F. Otondo and J. H Rindel. A new method for the radiation representation of musical instruments in auralizations. *Acta Acustica united with Acustica*, 91(5):902–906, 2005.

[132] F. Otondo and J.H. Rindel. The influence of the directivity of musical instruments in a room. *Acta Acustica united with Acustica*, 90(6):1178–1184, 2004.

[133] F. Otondo and J.H. Rindel. A new method for the radiation representation of musical instruments in auralizations. *Acta Acustica united with Acustica*, 91(5):902–906, 2005.

[134] J. Pätynen and T. Lokki. Evaluation of concert hall auralization with virtual symphony orchestra. In *the International Symposium on Room Acoustics (ISRA2010)*, Melbourne, Australia, Aug 29-31 2010.

[135] J. Pätynen and T. Lokki. Investigations on the development of the frequency response over time in concert halls. In *IOA Auditorium Acoustics 2011*, pages 1–10, Dublin, Ireland, May 20-22 2011. paper no. P22.

[136] J. Pätynen and T. Lokki. Konserttisalien akustiikan taajuusvaste ajan funktiona. In *Proceedings of Akustiikkapäivät 2011*, pages 1–6, Tampere, Finland, May 11-12 2011.

[137] J. Pätynen, S. Tervo, and T. Lokki. Simulation of the violin section sound based on the analysis of orchestra performance (in press). In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011)*, pages 1–4, New Paltz, New York, USA, Oct. 16-19 2011.

[138] J. Pätynen, S. Tervo, and T. Lokki. Temporal differences in string bowing of symphony orchestra players (conditionally accepted). *Journal of New Music Research*, 1(1):1–8, Mar. 2012.

[139] T. Peltonen. A multichannel measurement system for room acoustics analysis. Master's thesis, Helsinki University of Technology, Faculty of Electrical Engineering, Laboratory of Acoustics and Audio Signal Processing, 2000. 119 p.

[140] T. Peltonen, T. Lokki, B. Gouatarbes, J. Merimaa, and M. Karjalainen. A system for multi-channel and binaural room response measurements. In *the 110th Audio Engineering Society Convention*, Amsterdam, the Netherlands, May 12-15 2001. preprint no. 5289.

[141] Physikalisch-Technische Bundesanstalt. Directivities of musical instruments. http://www.ptb.de/en/org/1/16/163/directivity/richtchar.htm. (Cited Feb. 1 2011).

[142] J.D. Polack. Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics. *Applied Acoustics*, 38(2-4):235–244, 1993.

[143] J.D. Polack, X. Meynial, and V. Grillon. Auralization in scale models: Processing of impulse response. *J. Audio Eng. Soc.*, 41:939–939, 1993.

[144] M. Pollow, G. Behler, and B. Masiero. Measuring directivities of natural sound sources with spherical microphone array. In *Proc. Ambisonics Symposium 2009*, Graz, Austria, Jun. 25-27 2009.

[145] M. Pollow and G.K. Behler. Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization. *Acta Acustica united with Acustica*, 95(6):1082–1092, 2009.

[146] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.

[147] V. Pulkki and J. Merimaa. Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests. *J. Audio Eng. Soc.*, 54(1):3–20, 2006.

[148] R.A. Rasch. Synchronization in performed ensemble music. *Acustica*, 43(2):121–131, 1979.

[149] D.D. Rife and J. Vanderkooy. Transfer-function measurement with maximum-length sequences. *J. Audio Eng. Soc.*, 37(6):419–444, 1989.

[150] J.H. Rindel and C.L. Christensen. Auralisation of a symphony orchestra with Odeon – The chain from musical instruments to the eardrums. In *The EAA Symposium on Auralization*, Espoo, Finland, Jun. 15-17 2009.

[151] Røde Microphones. NT1-A datasheet, 2011. URL: http://www.rodemic.com /nt1-a/ [Cited Feb. 20, 2011].

[152] T. D. Rossing, R. F. Moore, and P. A. Wheeler. *The Science of Sound*. Benjamin Cummings, 3rd edition, 2002.

[153] W.C. Sabine. *Collected papers on acoustics*. Harvard university press, 1922.

[154] K. Saher, J.H. Rindel, and L. Nijs. Comparisons between Binaural In-situ Recordings and Auralizations. In *120th Convention of Audio Engeneering Society*, Paris, France, May 20-23 2006. Paper no. 6744.

[155] M. Saks, R. Hari, J. Rif, and J. Knuutila. The Human Auditory Sensory Memory Trace Persists about 10 sec: Neuromagnetic Evidence. *Journal of Cognitive Neuroscience*, 5(3):363–370, 1993.

[156] H. O. Saldner, N.-E. Molin, and E. V. Jansson. Vibration modes of the violin forced via the bridge and action of the soundpost. *J. Acoust. Soc. Am.*, 100(2):1168–1177, Aug. 1996.

[157] N. Schnell, G. Peeters, S. Lemouton, P. Manoury, and X. Rodet. Synthesizing a choir in real-time using Pitch Synchronous Overlap Add (PSOLA). In *Proc. Int. Computer Music Conference*, pages 102–108, Berlin, Germany, Sep. 2000.

[158] M. R. Schroeder, D. Gottlob, and K. F. Siebrasse. Comparative study of european concert halls: correlation of subjective preference with geometric and acoustic parameters. *J. Acoust. Soc Am*, 56(4):1195–1201, Oct. 1974.

[159] M.R. Schroeder. Natural sounding artificial reverberation. *J. Audio Eng. Soc*, 10(3):219–223, 1962.

[160] M.R. Schroeder. Integrated-impulse method measuring sound decay without using impulses. *J. Acoust. Soc. Am*, 66(2):497–500, 1979.

[161] D.B. Sharp, H.A.K. Wright, and W. Ring. An acoustical investigation into the effect of the crook profile on the sound produced by the bassoon. *Acta Acustica united with Acustica*, 89(1):137–144, 2003.

[162] S. Siltanen. *Efficient physics-based room-acoustics modeling and auralization*. PhD thesis, Aalto University School of Science and Technology, Department of Media Technology, Espoo, Finland, Jan. 2010.

[163] S. Siltanen, T. Lokki, S. Kiminki, and L. Savioja. The room acoustic rendering equation. *J. Acoust. Soc. Am*, 122(3):1624–1635, Sep. 2007.

[164] S. Siltanen, T. Lokki, and L. Savioja. Frequency domain acoustic radiance transfer for real-time auralization. *Acta Acustica united with Acustica*, 95(1):106–117, 2009.

[165] J. Smith, J. Wolfe, and M. Green. Head joint, embouchure hole and filtering effects on the input impedance of flutes. In *Proc. Stockholm Music Acoustics Conference SMAC 03*, pages 295–298, Stockholm, Sweden, Aug. 6-9 2003.

[166] G.B. Stan, J.J. Embrechts, and D. Archambeau. Comparison of different impulse response measurement techniques. *J. Audio Eng. Soc.*, 50(4):249–262, 2002.

[167] D. Sumarac-Palvolic, M. Mijic, and H. Kurtovic. A simple impulse sound source for measurements in room acoustics. *Applied Acoustics*, 69(4):378–383, Apr. 2008.

[168] The University of New South Wales, Sydney, Australia. Brass instrument acoustics, 2010. URL: http://www.phys.unsw.edu.au/jw/brassacoustics.html [Cited Dec. 2, 2010].

[169] The University of New South Wales, Sydney, Australia. Flute acoustics, 2010. URL: http://www.phys.unsw.edu.au/music/flute/ [Cited Dec. 2, 2010].

[170] The University of New South Wales, Sydney, Australia. Violin acoustics, 2010. URL: http://www.phys.unsw.edu.au/jw/violin.html [Cited Dec. 2, 2010].

[171] M. P. Verge, A. Hirschberg, and R. Caussé. Sound production in recorder-like instrument: II. A simulation model. *Acta Acustica*, 101(5):2925–2939, May 1997.

[172] Vienna Symphonic Library. Vienna instruments, website, 2011. URL: http://www.vsl.co.at/ [Cited Feb. 8, 2011].

[173] M.C. Vigeant, L.M. Wang, and J. H. Rindel. Investigations of orchestra auralizations using the multi-channel multi-source auralization technique. *Acta Acustica united with Acustica*, 94(6):866–882, Dec. 2008.

[174] M.C. Vigeant, L.M. Wang, and J.H. Rindel. Investigations of multi-channel auralization technique for solo instruments and orchestra. In *Proc. 19th International Congress on Acoustics*, Madrid, Spain, September 2-7 2007. paper no. RBA-15-004.

[175] M.C. Vigeant, L.M. Wang, and J.H. Rindel. Objective and subjective evaluations of the multi-channel auralization technique as applied to solo instruments. *Applied Acoustics*, 72(6):311–323, 2011.

[176] M.C. Vigeant, L.M. Wang, J.H. Rindel, C.L. Christensen, and A.C.Gade. Multi-channel orchestral anechoic recordings for auralizations. In *Proc. Int. Symposium on Room Acoustics (ISRA 2010)*, Melbourne, Austalia, Aug 29-31 2010. Paper RBA-15-004.

[177] J. Vilkamo, T. Lokki, and V. Pulkki. Directional Audio Coding: Virtual Microphone-Based Synthesis and Subjective Evaluation. *Journal of the Audio Engineering Society*, 57(9):709–724, 2009.

[178] Virtual acoustics research group, Aalto University School of Science. Website, 2011. URL: http://auralization.tkk.fi/ [Cited Feb. 24, 2011].

[179] M. Vorländer. *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer-Verlag, Berlin, Germany, 2008.

[180] H.J. Vos, O. Warusfel, N. Misdariis, and D. de Vries. Analysis and reproduction of the frequency spectrum and directivity of a violin. *J. Acoust. Soc. Neth.*, 167:1–11, 2003.

[181] L. M. Wang and C. B. Burroughs. Radiation from bowed violins. *J. Acoust. Soc. Am.*, 110(1):543–555, Jul. 2001.

[182] L.M. Wang and M.C. Vigeant. Evaluations of output from room acoustic computer modeling and auralization due to different sound source directionalities. *Applied Acoustics*, 69(12), Dec 2008.

[183] O. Warusfel, P. Derogis, and R. Caussé. Radiation synthesis with digitally controlled loudspeakers. In *the 103rd Audio Engineering Society Convention*, New York, USA, Sep. 26-29 1997. preprint no. 4577.

[184] O. Warusfel and N. Misdariis. Directivity synthesis with a 3D array of loudspeakers application for stage performance. In *Proc. COST G-6 Conference on Digital Audio Effects (DAFx-01)*, Limerick, Ireland, Dec. 6-8 2001.

[185] G. Weinreich. Sound hole sum rule and the dipole moment of the violin. *J. Acoust. Soc. Am.*, 77(2):701–718, Feb. 1985.

[186] G. Weinreich. Directional tone color. *J. Acoust. Soc. Am.*, 101(4):2338–2346, Apr. 1997.

[187] G. Weinreich and E.B. Arnold. Method for measuring acoustic radiation fields. *J. Acoust. Soc. Am.*, 68(2):404–411, 1980.

[188] C.A. Weitze, C.L. Christensen, J.H. Rindel, and A.C. Gade. Comparison between In-situ recordings and Auralizations for Mosques and Byzantine Churches. In *Proc. Joint Baltic-Nordic Acoustical Meeting*, Aug. 26-28 2002.

[189] Wenger Corp. CD: Anechoic choral recordings, 2003.

[190] J. Wolfe, J. Smith, J. Tann, and N.H. Fletcher. Acoustic impedance of classical and modern flutes. *Journal of Sound and Vibration*, 241(1):127–143, May 2001.

[191] G.S.K. Wong. Speed of sound in standard air. *J. Acoust. Soc. Am.*, 79:1359–1366, 1986.

[192] G.S.K. Wong and T.F.W. Embleton. Variation of the speed of sound in air with humidity and temperature. *J. Acoust. Soc. Am.*, 77:1710–1712, 1985.

[193] K. Yamaguchi. Multivariate analysis of subjective and physical measures of hall acoustics. *J. Acoust. Soc. Am*, 52(5A):1271, 1972.

[194] U. Zölzer, editor. *DAFX:Digital Audio Effects. Chapter 2.6: Delay-based audio effects*. John Wiley & Sons, Chichester, United Kingdom, 2nd edition, May 2011.

[195] U. Zölzer, editor. *DAFX:Digital Audio Effects. Chapter 7: Time-frequency processing*. John Wiley & Sons, Chichester, United Kingdom, 2nd edition, May 2011.

Aalto-DD 86/2011

ISBN 978-952-60-4291-6 (pdf)
ISBN 978-952-60-4290-9
ISSN-L 1799-4934
ISSN 1799-4942 (pdf)
ISSN 1799-4934

**Aalto University**
**School of Science**
**Department of Media Technology**
**www.aalto.fi**

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

**DOCTORAL**
**DISSERTATIONS**