



COMMUNICATION ACOUSTICS

AN INTRODUCTION TO SPEECH,
AUDIO AND PSYCHOACOUSTICS

VILLE PULKKI
MATTI KARJALAINEN

WILEY

COMMUNICATION ACOUSTICS

AN INTRODUCTION TO SPEECH,
AUDIO AND PSYCHOACOUSTICS

Ville Pulkki and Matti Karjalainen

Aalto University, Finland

WILEY

This edition first published 2015
© 2015 John Wiley & Sons, Ltd

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought

MATLAB® is a trademark of The MathWorks, Inc. and is used with permission. The MathWorks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by The MathWorks of a particular pedagogical approach or particular use of the MATLAB® software.

Library of Congress Cataloging-in-Publication Data

Pulkki, Ville.

Communication acoustics : an introduction to speech, audio, and psychoacoustics / Ville Pulkki and Matti Karjalainen.

pages cm

Includes index.

ISBN 978-1-118-86654-2 (hardback)

1. Sound-Recording and reproducing. 2. Hearing. 3. Psychoacoustics. 4. Sound. I. Karjalainen, Matti, 1946-2010.

II. Title.

TK7881.4.P85 2015

620.2-dc23

2014030757

A catalogue record for this book is available from the British Library.

Set in 10/12pt Times LT Std by SPi Publisher Services, Pondicherry, India

Printed in Singapore by C.O.S. Printers Pte Ltd

COMMUNICATION ACOUSTICS

To Vappu, Sampo, Raisa, and HUT people

To Sari, Veikko and Sannakaisa, and Aalto Acoustics people

About the Authors

Ville Pulkki has been working in the field of audio from 1995. In his PhD thesis (2001) he developed a method to position virtual sources for 3D loudspeaker set-ups after researching the method using psychoacoustic listening tests and binaural computational models of human hearing. Later he worked on the reproduction of recorded spatial sound scenarios, on the measurement of head-related acoustics and on the measurement of room acoustics with laser-induced pressure pulses. Currently he holds a tenure-track assistant professor position in Aalto University and runs a research group with 18 researchers. He is a fellow of the Audio Engineering Society (AES) and has received the AES Publication Award. He has also received the Samuel L. Warner memorial medal from the Society of Motion Picture and Television Engineers (SMPTE). He has a background in music, having received teaching from the Sibelius Academy in singing and audio engineering alongside instruction in various musical instruments. He has also composed and arranged music for many different ensembles. He enjoys being with his family, renovating his summerhouse and dancing to hip hop.

Matti Karjalainen (1946–2010) began his career as an associate professor of acoustics at the Helsinki University of Technology (TKK) in the 1980s. He maintained a long and prolific career as a researcher and visionary leader in acoustic and audio signal processing, both as a pioneer of Finnish language speech synthesis and developer of the first portable microprocessor-based text-to-speech synthesizer in the world. For ten years he was Finland's only university professor of acoustics, leading the Laboratory of Acoustics and Audio Signal Processing at the TKK (now Aalto University) until 2006. Some of his groundbreaking work included applying his expert knowledge of psychoacoustics to computational auditory models, as well as sophisticated physical modelling of stringed instruments utilizing the fractional delay filter design for tuning, a now standard technique in this field. Later in life, augmented reality audio and spatial audio signal processing remained among Matti's greatest research interests. For his achievements in audio signal processing Matti received the Audio Engineering Society fellowship in 1999, the AES silver medal in 2006 and the IEEE fellowship in 2009. On his 60th birthday he founded the Matti Karjalainen Fund, supporting young students into studying acoustics. Matti's share of the revenues from this book are routed to the fund. In May 2010, Matti passed away at home, survived by his wife, daughter and son Sampo, a well-known software designer living in the US.

Preface

The book *Kommunikaatioakustiikka* by Matti Karjalainen (1946–2010) has always been around during my research career in audio and psychoacoustics, starting from my PhD studies (1996–2001), through the periods when I was a postdoc (2001–2005), a senior researcher (2005–2012), and now during my tenure track professorship (2012–). I first used the book as a reference, as it summarized many relevant topics and provided good pointers on where to find more information. I have also been teaching the corresponding course at Aalto University (the university formerly called Helsinki University of Technology), first during Matti's sabbatical years, then sick leaves, and regularly after his passing away. It was my and many other people's opinion that the book was great, but it did not have a counterpart written in English. Matti himself knew this, and he worked on a translation, a of which he completed about 30% in 2002, including the preface that follows.

For a long time I thought that I should finish Matti's work, as it would benefit people in the fields of audio and speech. However, I also understood that it would be quite a hard job. The final motivation came from my university, which stated that all MSc-level teaching should have course material in English as well, starting from autumn 2015. So, in autumn 2013, I decided to complete the book. To ensure that I would really do it, I proposed the book to Wiley, since I understood that I needed a deadline. I also thought that international distribution would benefit the propagation of the book. The book project meant a period of 10 months where I worked so much that I felt that my hands were stuck permanently to my laptop.

The book grew by about 30% from the original Finnish book, as I added quite a bit more material on audio techniques and updated many parts of the book. Consequently, the subject matter of the book might be too large for a single-semester course. Teachers are encouraged to leave some chapters out, as the whole book might be too much information to be digested in one go. I shall be updating the companion web page of this book with sound examples and other material to help teachers of such courses.

The book covers many fields within acoustics, and without great help from many professionals in the field, the book would be less detailed and less complete. First of all, I received very kind help in translating and updating the text from my PhD students Marko Takanen (Chapter 13), Teemu Koski (Chapter 19), and Olli Rummukainen (Section 11.7). Juha Vilkkamo and Marko Takanen also provided text and figures from their PhD theses. The following professionals have read and commented on, or otherwise helped with, the project: Paavo Alku, Brian C. J. Moore, Mikko Kurimo, Ville Sivonen, Nelli Salminen, Ilkka Huhtakallio, Cleopatra Pike, Catarina Hiipakka, Alessandro Altoè, Mikko-Ville Laitinen, Søren Bech, Archontis Politis,

Olli Santala, Sascha Disch, Tapio Lokki, Lauri Savioja, Hannu Pulakka, Richard Furse, Unto K. Laine, Vesa Välimäki, Javier Gómez Bolaños, Cumhur Erkut, Damian Murphy, Simon Christiansen, Jesper Ramsgaard, Bastian Epp, Athanasios Mouchtaris, Nikos Stefanakis, Antti Kelloniemi, Kalle Koivuniemi, Ercan Altinsoy, Lauri Juvela, Symeon Delikaris-Manias, Tapani Pihlajamäki, Antti Jylhä, Tuomo Raitio, Martti Vainio, Gaëtan Lorho, Mari Tervaniemi, Antti Kuusinen, Jouni Pohjalainen, Christian Uhle, Torben Poulsen, Davide Rocchesso, Nick Zacharov, and Thibaud Necciari. Luis R. J. Costa worked on removing the worst Finglishisms in the book, making them into more readable English expressions.

I, of course, hope that the book is successful, and new editions come out in time. With that in mind following the tradition started by Brian C. J. Moore in his *Introduction to Psychology of Hearing*, I would hereby like to open a similar contest. A prize of a box of Finnish chocolate confections will be awarded to the reader who spots the most errors in this edition, and writes to me to point them out. Game on!

I hope you will enjoy reading the book, and that you will find it beneficial in your research work and studies.

Ville Pulkki
Otaniemi, Espoo, Finland
May 2014

Preface to the Unfinished Manuscript of the Book

The origins of this book date back to 1980 at the Helsinki University of Technology (HUT). I started lecturing the course 'COMMUNICATION ACOUSTICS' (in Finnish), based on a collection of material from previous lecturers, in the tradition of engineering psychoacoustics influenced by some US and German textbooks.

The first encounter with principles on how we perceive sound was a very inspiring one, not only because of the need to traverse through a vast amount of experimental results on the functioning of hearing and gradually structure the underlying principles, but also because of having to start to generate ideas on how this could be simulated by computational means. This knowledge has gradually developed in various research projects, and the present course on 'COMMUNICATION ACOUSTICS' is a much more mature version than the first one.

Finishing a Finnish textbook for the course after 18 years was the proper time to look at the possibility of rewriting and extending it into an English version. This task was motivated by finding that there was no modern textbook covering the wide field of communication by sound and voice, especially from the point of view of engineering psychoacoustics. To fill a bit of this gap, this book was written. Sabbatical leave during 1999–2000 from my HUT duties allowed me to do most of this writing.

The importance of engineering psychoacoustics has been growing rapidly since the late 1980s. There are several reasons for this. The rapid development of digital audio is one of the most important driving forces. Audio coding and related questions of sound quality showed engineers and scientists how essential knowing the principles of human auditory perception is.

Not only hearing-related knowledge and know-how but also knowledge of a wide variety of other disciplines involving sound and voice are needed to work successfully in advanced research and development. Three cornerstones, in my view, are essential for such success: (physical) acoustics, auditory perception, and (digital) signal processing. The first tells us how physical systems behave, the second how we behave when capturing sound and its content, and the last is the primary tool to implement modern sound technology.

This textbook is written mainly for engineering-oriented readers as a general tutorial to a wide range of topics important in R&D on modern sound technology. On the other hand, it is written with the objective of providing this knowledge to a much wider audience as well. The reader may focus on the subtopic of interest, in most cases without having to digest all the prior material in full detail. Also, formal theories from mathematics are not extensive and may be skipped, at least in the first reading. One important function of the book is to provide

pointers to relevant literature on each subtopic. In this sense, it may also have a modest role as a handbook of references.

As I have already said, the scope of the book is wide, and it tries to be comprehensive but not deep in every subject. The first chapters present an overview of acoustics, signal processing, speech, and audio. These are more like introductions to the most important concepts. Readers who are familiar with them may skip them. To some degree, understanding these topics and concepts is necessary in the rest of the book. The main part of the book is related to the human auditory system, its function and properties, as well as to modelling for research and applications.

I have been lucky in having help and support from many people. My students have provided feedback on my lectures and the Finnish version of this book.

This is the first English version of the book, and I hope, in addition to proving useful to as many readers as possible, that readers are willing to provide feedback to help me improve this book in the future.

Matti Karjalainen
Fiskars, Finland
January 2002

Contents

About the Authors	xix
Preface	xxi
Preface to the Unfinished Manuscript of the Book	xxiii
Introduction	1
1 How to Study and Develop Communication Acoustics	7
1.1 Domains of Knowledge	7
1.2 Methodology of Research and Development	8
1.3 Systems Approach to Modelling	10
1.4 About the Rest of this Book	12
1.5 Focus of the Book	12
1.6 Intended Audience	13
References	14
2 Physics of Sound	15
2.1 Vibration and Wave Behaviour of Sound	15
2.1.1 <i>From Vibration to Waves</i>	16
2.1.2 <i>A Simple Vibrating System</i>	16
2.1.3 <i>Resonance</i>	18
2.1.4 <i>Complex Mass–Spring Systems</i>	19
2.1.5 <i>Modal Behaviour</i>	20
2.1.6 <i>Waves</i>	21
2.2 Acoustic Measures and Quantities	23
2.2.1 <i>Sound and Voice as Signals</i>	23
2.2.2 <i>Sound Pressure</i>	24
2.2.3 <i>Sound Pressure Level</i>	24
2.2.4 <i>Sound Power</i>	25
2.2.5 <i>Sound Intensity</i>	25
2.2.6 <i>Computation with Amplitude and Level Quantities</i>	25

2.3	Wave Phenomena	26
2.3.1	<i>Spherical Waves</i>	26
2.3.2	<i>Plane Waves and the Wave Field in a Tube</i>	27
2.3.3	<i>Wave Propagation in Solid Materials</i>	29
2.3.4	<i>Reflection, Absorption, and Refraction</i>	31
2.3.5	<i>Scattering and Diffraction</i>	32
2.3.6	<i>Doppler Effect</i>	33
2.4	Sound in Closed Spaces: Acoustics of Rooms and Halls	34
2.4.1	<i>Sound Field in a Room</i>	34
2.4.2	<i>Reverberation</i>	36
2.4.3	<i>Sound Pressure Level in a Room</i>	37
2.4.4	<i>Modal Behaviour of Sound in a Room</i>	38
2.4.5	<i>Computational Modelling of Closed Space Acoustics</i>	39
	Summary	41
	Further Reading	41
	References	41
3	Signal Processing and Signals	43
3.1	Signals	43
3.1.1	<i>Sounds as Signals</i>	43
3.1.2	<i>Typical Signals</i>	45
3.2	Fundamental Concepts of Signal Processing	46
3.2.1	<i>Linear and Time-Invariant Systems</i>	46
3.2.2	<i>Convolution</i>	47
3.2.3	<i>Signal Transforms</i>	48
3.2.4	<i>Fourier Analysis and Synthesis</i>	49
3.2.5	<i>Spectrum Analysis</i>	50
3.2.6	<i>Time–Frequency Representations</i>	53
3.2.7	<i>Filter Banks</i>	54
3.2.8	<i>Auto- and Cross-Correlation</i>	55
3.2.9	<i>Cepstrum</i>	56
3.3	Digital Signal Processing (DSP)	56
3.3.1	<i>Sampling and Signal Conversion</i>	56
3.3.2	<i>Z Transform</i>	57
3.3.3	<i>Filters as LTI Systems</i>	58
3.3.4	<i>Digital Filtering</i>	58
3.3.5	<i>Linear Prediction</i>	59
3.3.6	<i>Adaptive Filtering</i>	62
3.4	Hidden Markov Models	62
3.5	Concepts of Intelligent and Learning Systems	63
	Summary	64
	Further Reading	64
	References	64
4	Electroacoustics and Responses of Audio Systems	67
4.1	Electroacoustics	67
4.1.1	<i>Loudspeakers</i>	67
4.1.2	<i>Microphones</i>	70

4.2	Audio System Responses	71
4.2.1	<i>Measurement of System Response</i>	71
4.2.2	<i>Ideal Reproduction of Sound</i>	72
4.2.3	<i>Impulse Response and Magnitude Response</i>	72
4.2.4	<i>Phase Response</i>	74
4.2.5	<i>Non-Linear Distortion</i>	75
4.2.6	<i>Signal-to-Noise Ratio</i>	76
4.3	Response Equalization	76
	Summary	77
	Further Reading	78
	References	78
5	Human Voice	79
5.1	Speech Production	79
5.1.1	<i>Speech Production Mechanism</i>	80
5.1.2	<i>Vocal Folds and Phonation</i>	80
5.1.3	<i>Vocal and Nasal Tract and Articulation</i>	82
5.1.4	<i>Lip Radiation Measurements</i>	84
5.2	Units and Notation of Speech used in Phonetics	84
5.2.1	<i>Vowels</i>	86
5.2.2	<i>Consonants</i>	86
5.2.3	<i>Prosody and Suprasegmental Features</i>	88
5.3	Modelling of Speech Production	90
5.3.1	<i>Glottal Modelling</i>	92
5.3.2	<i>Vocal Tract Modelling</i>	92
5.3.3	<i>Articulatory Synthesis</i>	94
5.3.4	<i>Formant Synthesis</i>	95
5.4	Singing Voice	96
	Summary	96
	Further Reading	97
	References	97
6	Musical Instruments and Sound Synthesis	99
6.1	Acoustic Instruments	99
6.1.1	<i>Types of Musical Instruments</i>	99
6.1.2	<i>Resonators in Instruments</i>	100
6.1.3	<i>Sources of Excitation</i>	102
6.1.4	<i>Controlling the Frequency of Vibration</i>	103
6.1.5	<i>Combining the Excitation and Resonant Structures</i>	104
6.2	Sound Synthesis in Music	104
6.2.1	<i>Envelope of Sounds</i>	105
6.2.2	<i>Synthesis Methods</i>	106
6.2.3	<i>Synthesis of Plucked String Instruments with a One-Dimensional Physical Model</i>	107
	Summary	108
	Further Reading	108
	References	108

7	Physiology and Anatomy of Hearing	111
7.1	Global Structure of the Ear	111
7.2	External Ear	112
7.3	Middle Ear	113
7.4	Inner Ear	115
7.4.1	<i>Structure of the Cochlea</i>	115
7.4.2	<i>Passive Cochlear Processing</i>	117
7.4.3	<i>Active Function of the Cochlea</i>	119
7.4.4	<i>The Inner Hair Cells</i>	122
7.4.5	<i>Cochlear Non-Linearities</i>	122
7.5	Otoacoustic Emissions	123
7.6	Auditory Nerve	123
7.6.1	<i>Information Transmission using the Firing Rate</i>	124
7.6.2	<i>Phase Locking</i>	126
7.7	Auditory Nervous System	127
7.7.1	<i>Structure of the Auditory Pathway</i>	127
7.7.2	<i>Studying Brain Function</i>	129
7.8	Motivation for Building Computational Models of Hearing	130
	Summary	131
	Further Reading	131
	References	131
8	The Approach and Methodology of Psychoacoustics	133
8.1	Sound Events versus Auditory Events	133
8.2	Psychophysical Functions	135
8.3	Generation of Sound Events	135
8.3.1	<i>Synthesis of Sound Signals</i>	136
8.3.2	<i>Listening Set-up and Conditions</i>	137
8.3.3	<i>Steering Attention to Certain Details of An Auditory Event</i>	137
8.4	Selection of Subjects for Listening Tests	138
8.5	What are We Measuring?	138
8.5.1	<i>Thresholds</i>	138
8.5.2	<i>Scales and Categorization of Percepts</i>	140
8.5.3	<i>Numbering Scales in Listening Tests</i>	141
8.6	Tasks for Subjects	141
8.7	Basic Psychoacoustic Test Methods	142
8.7.1	<i>Method of Constant Stimuli</i>	143
8.7.2	<i>Method of Limits</i>	143
8.7.3	<i>Method of Adjustment</i>	143
8.7.4	<i>Method of Tracking</i>	144
8.7.5	<i>Direct Scaling Methods</i>	144
8.7.6	<i>Adaptive Staircase Methods</i>	144
8.8	Descriptive Sensory Analysis	145
8.8.1	<i>Verbal Elicitation</i>	147
8.8.2	<i>Non-Verbal Elicitation</i>	148
8.8.3	<i>Indirect Elicitation</i>	148