UNIVERSITY OF SOUTHAMPTON

Advances in Sound Field Analysis and Control based on Cylindrical Coordinates

by

Falk-Martin Hoffmann

A thesis submitted for the degree of

Doctor of Philosophy

in the
Faculty of Engineering and the Environment
Institute of Sound and Vibration Research

June 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

ADVANCES IN SOUND FIELD ANALYSIS AND CONTROL BASED ON CYLINDRICAL COORDINATES

by Falk-Martin Hoffmann

This Ph.D. thesis concerns advances in acoustic transducer array technology for improved sound field analysis and control performance. Four principal investigations are presented, which address specific performance limitations of microphone arrays and loudspeaker arrays. The basic model, on which these investigations are founded, is the general solution of the Helmholtz equation in cylindrical coordinates. The individual acoustical investigations draw on the analysis of the eigenvalues and eigenfunctions of the forward operator, providing information on the robustness of the inverse solutions against non-uniqueness, ill-conditioning and spatial aliasing. A circular microphone array design based on tangentially aligned pressure gradient sensors is studied. The theoretical analysis is complemented by a simulation study, comparing the new design to conventional arrays built from pressure sensors. It is shown that the proposed design can provide an improved performance at low frequencies, while performing worse at high frequencies due to spatial aliasing. The effects of the latter can be compensated if the Direction-of-Arrival (DOA) of the incoming waves is known. A novel DOA estimation method for sound fields measured with circular microphone arrays is proposed to address this. Using analytical expressions to model the sound fields of point sources and plane waves, it is studied for which sound fields the method is applicable and how robust it is against model imperfections. The estimation accuracy for different numbers of sources and different levels of background noise is investigated in a simulation study and the method is tested against real data, obtained through acoustic measurements. The estimation results achieved in simulations and with experimental data compare well. The general solution to the Helmholtz equation is then applied as a model for acoustic radiation in wedge-shaped spaces. This investigation aims to improve the performance of loudspeaker arrays in restricted propagation spaces, e.g. rooms. By introducing boundary conditions to the general model, different sets of basis functions are implemented in the solution and it is shown that the model enables Nearfield Acoustical Holography (NAH). Using the same propagation model, a technique for sound field control with arrays in wedge spaces is developed. The inverse problem is solved by means of a mode-matching approach, leading to an expression for the driving signals based on a target beam pattern. Both simulations and experiments with a hemi-cylindrical loudspeaker array prototype confirm the applicability of the model for both NAH and beamforming with loudspeaker arrays in wedge spaces. Different beam patterns are considered and the model is tested through simulations and experiments. The implications of the findings, how they are linked and what future developments they may lead to is discussed.

Contents

List of Figures xi

Nomenclature xvii

Acronyms xix

Declaration of Authorship xxi

Acknowledgements xxiii

1 Introduction 1
- 1.1 Historical Background and Placement of this Thesis into the Scientific Context . 1
- 1.2 Contributions and Structure of the Thesis 2
  - 1.2.1 Journal Paper: Theoretical Study of Circular Arrays with Tangential Pressure Gradient Sensors 3
  - 1.2.2 Journal Paper: DOA Estimation Performance with Circular Arrays in Sound Fields with Finite Rate of Innovation 4
  - 1.2.3 Journal Paper: A General Radiation Model for Sound Fields and Nearfield Acoustical Holography in Wedge Propagation Spaces 6
  - 1.2.4 Journal Paper: Beamforming with Wedge-Shaped Acoustic Arrays 7
  - 1.2.5 Conclusions and Appendices 8

2 Literature Review 11
- 2.1 Wave Field Models for Acoustic Transducer Arrays 12
  - 2.1.1 Solution of the Homogeneous Helmholtz Equation in Cylindrical Coordinates 13
  - 2.1.2 Solution to the Homogeneous Helmholtz Equation in Spherical Coordinates 14
  - 2.1.3 Solution of the Homogeneous Helmholtz Equation in Infinite Wedges . . . 16
  - 2.1.4 The Herglotz Wave Function 17
  - 2.1.5 The Single Layer Potential 18
- 2.2 Inverse Problems 18
  - 2.2.1 Non-Uniqueness of a Solution 19
  - 2.2.2 Ill-Conditioning, Transducer Noise, Misalignment and Variation of Environmental Parameters 20
  - 2.2.3 Spatial Aliasing 21
  - 2.2.4 Nearfield Acoustical Holography 22
- 2.3 Microphone Arrays & Signal Processing Strategies 23
  - 2.3.1 Types of Transducers 23
  - 2.3.2 Signal Processing Strategies 24
    - 2.3.2.1 Direction-of-Arrival Estimation 24
    - 2.3.2.2 Compressed Sensing 25
    - 2.3.2.3 Finite Rate of Innovation Theory 25
- 2.4 Compact Loudspeaker Arrays 26
  - 2.4.1 Designs & Performance 26

2.4.2 Applications in Sound Field Control . . . 27

3 Mathematical Notation, Definitions and Commonly Used Functions 29
3.1 Vectors, Matrices and Norms . . . 29
3.2 Complex Numbers . . . 31
3.3 Sets . . . 31
3.3.1 Defining Sets . . . 31
3.3.2 A Cylinder of Infinite Height . . . 31
3.4 Function Spaces . . . 31
3.4.1 L^2 Spaces . . . 31
3.4.2 Scalar Product . . . 32
3.4.3 Total Orthonormal Sets . . . 32
3.4.4 Series Representations in $L^2(V)$ . . . 33
3.5 Special Functions and Distributions . . . 33
3.5.1 Dirac Pulse . . . 33
3.5.2 Rectangular Function . . . 34
3.5.3 Sinc Function . . . 34
3.5.4 Circular Sinc Function . . . 34
3.5.5 The Kronecker Delta . . . 35
3.6 Fourier Series . . . 35
3.6.1 Definition . . . 35
3.6.2 Parseval's Theorem . . . 36
3.7 Operators . . . 36
3.8 Fourier Transform . . . 36
3.8.1 Definition . . . 36
3.8.2 Important Properties and Special Transformation Pairs . . . 37
3.8.2.1 Convolution Theorem . . . 37
3.8.2.2 Differentiation Property . . . 37
3.8.2.3 Plancherel Theorem . . . 38
3.8.2.4 Dirac Pulse . . . 38
3.8.2.5 Rectangular Function . . . 38
3.9 Coordinate Systems . . . 38
3.9.1 Cartesian Coordinates . . . 38
3.9.2 Cylindrical Coordinates . . . 39
3.9.3 Spherical Coordinates . . . 39

4 Theoretical Study of Acoustic Circular Arrays With Tangential Pressure Gradient Sensors 41
4.1 Introduction . . . 41
4.2 Model of the Sound Field . . . 45
4.2.1 Plane-Wave Decomposition . . . 45
4.2.2 Analysis of the Integral Operator G . . . 48
4.2.3 Eigenvalue Decomposition of G . . . 49
4.2.4 The Eigenvalues λ_n . . . 50
4.2.5 Solution of the Inverse Problem . . . 51
4.3 Discretisation and Aliasing Analysis . . . 53
4.3.1 The Aliasing Pattern . . . 54
4.3.2 Mode Recovery using Linear Algebra . . . 55
4.3.3 Recovery of the 0th Mode using an additional pressure sensor . . . 56
4.3.4 Robustness of the 0th Order . . . 58
4.3.5 HD Coefficients of the Sound Field of a Plane Wave . . . 59
4.3.6 Example of Mode Recovery With Spatial Aliasing . . . 60
4.4 Simulations and Quantitative Error Analysis . . . 61

4.4.1 The Noise Model . . . 62
4.4.2 Regularisation Against Ill-Conditioning . . . 63
4.4.3 The Quality Measure . . . 63
4.4.4 Simulation Results . . . 63
4.4.5 An Example . . . 65
4.5 Conclusions . . . 65

5 DOA Estimation Performance with Circular Arrays in Sound Fields with Finite Rate of Innovation 67
5.1 Introduction . . . 67
5.2 Sound Field Model . . . 69
5.2.1 Plane Wave Expansion in Cylindrical Coordinates . . . 69
5.2.2 The Herglotz Wave Function . . . 70
5.2.3 Sound Field of L Plane Waves . . . 70
5.3 Circular Microphone Array Measurement . . . 71
5.3.1 General PW Sound Field Measurement . . . 71
5.3.2 HD Coefficients of PWs travelling parallel to $z = z_M$ . . . 72
5.3.3 HD Coefficients of a Single PW with $\theta_l \neq \frac{\pi}{2}$ . . . 73
5.3.4 HD Coefficients of a Single Point Source . . . 73
5.3.5 Practical Limitations of Microphone Arrays . . . 73
5.4 Finite Rate of Innovation Analysis . . . 74
5.4.1 The System Inherent Sampling Kernel . . . 74
5.4.2 The Rate of Innovation of the HD . . . 74
5.4.3 Recovering the DOA Information . . . 74
5.4.4 A Measurement Noise Model . . . 76
5.4.5 Choosing K with Noisy Coefficients q_n . . . 78
5.4.6 Interpretation of the Annihilating Filter . . . 78
5.5 Simulations . . . 79
5.5.1 Estimation Behaviour for $L > N$ . . . 80
5.5.2 The DOA Estimation Error (DEE) . . . 81
5.5.3 Mean & Median DEE . . . 81
5.5.4 Sound Fields with PWs not travelling parallel to the Measurement Plane 82
5.5.4.1 A Single OOP-PW . . . 83
5.5.4.2 One PW and one OOP-PW . . . 83
5.5.5 Sound Fields with Point Sources . . . 84
5.5.5.1 Estimation Accuracy for a Single PS . . . 84
5.5.5.2 Estimation Accuracy for Two Point Sources . . . 85
5.5.6 Discussion of Simulation Results . . . 86
5.5.6.1 Single OOP-PW . . . 86
5.5.6.2 One PW and one OOP-PW . . . 86
5.5.6.3 Single Point Source . . . 87
5.5.6.4 Two Point Sources . . . 87
5.5.6.5 Practical Implications of the Results for Point Sources and OOP-PW 87
5.5.6.6 Influence of the Range of Coefficients . . . 88
5.6 Experimental Results . . . 88
5.6.1 Measurement Setup . . . 88
5.6.2 Measured DOA of the Sources . . . 89
5.6.3 Mean & Median DEE for Linear Combinations of Individual Measurements 90
5.6.4 Example: Measurement of Six Sources . . . 91
5.6.5 Discussion . . . 91
5.7 Conclusions . . . 91
5.8 Appendix I: General Plane Wave Expansion in Cylindrical Coordinates . . . 92
5.9 Appendix II: Series Expansion for the Sound Field of a Point Source . . . 93

5.10 Appendix III: HD Coefficients for the Sound Field of a Point Source 93

6 A General Radiation Model for Sound Fields and Nearfield Acoustical Holography in Wedge Propagation Spaces 95
6.1 Introduction 95
6.2 Sound Field Model 97
6.2.1 General Solution of the Helmholtz Equation in the Infinite Wedge 97
6.2.2 General Solution in the Finite Wedge 100
6.2.3 Pressure and Velocity Spectrum 101
6.3 Rayleigh-like Formulae and Neumann-Green Functions 102
6.3.1 Pressure Radiated from the Boundary V for Infinite Wedges 103
6.3.2 Pressure Radiated from the Boundary V for Finite Wedges 104
6.3.3 Rayleigh-like Formula and Neumann-Green Function for Infinite Wedges . 104
6.3.4 Rayleigh-like Formula and Neumann-Green Function for Finite Wedges . 105
6.4 Radial Functions 106
6.4.1 Identifying the Radial Functions 106
6.4.2 Properties 106
6.4.3 Truncation and Limitation of the Solutions 107
6.5 Holography in Infinite and Finite Wedges 108
6.6 Sound Field Simulations 109
6.6.1 Point Source Model 110
6.6.2 Piston Source Model 110
6.7 Measurements 111
6.7.1 Measurement Rig 112
6.7.2 Angular Sampling and Resolution 112
6.7.3 Vertical Sampling and Resolution 113
6.7.4 Analysis Equations 114
6.7.5 Result for a Single Driver 114
6.7.6 Result for 15 Active Drivers 116
6.7.7 Discussion 118
6.8 Summary & Outlook 118

7 Beamforming with Wedge-Shaped Acoustic Arrays 121
7.1 Introduction 121
7.2 Sound Field Model 123
7.2.1 General Pressure Model 124
7.2.2 The General Solution to the Inverse Problem 125
7.2.3 Velocity Spectrum of an Arbitrary Wedge Array 126
7.2.4 The Array-Specific Forward Problem 127
7.3 Loudspeaker Driving Functions 127
7.3.1 Diaphragm Velocity as a Function of Driving Voltage 127
7.3.2 Model for the Mode-Matching Approach 128
7.3.3 Loudspeaker Positioning 128
7.3.4 Orthonormality and Aliasing 129
7.3.5 The Arc Radial Functions 130
7.3.6 Truncation of Higher Order Modes 133
7.3.7 Regularised Driving Functions 133
7.4 Beamforming 134
7.4.1 Beam Pattern I 134
7.4.2 Beam Pattern II 135
7.5 Simulations 135
7.5.1 Case Study: Quadrant Cylindrical Array with $L = 15$ 136
7.5.2 Case Study: Hemi Cylindrical Array with $L = 15$ 138

7.6 Measurements . . . 142
7.6.1 Setup . . . 142
7.6.2 Normalised Mean Squared Measurement Error . . . 142
7.6.3 Results for Different HCA Beamforming Filters . . . 143
7.6.4 Discussion . . . 143
7.7 Summary & Outlook . . . 144
7.8 Appendix I: Stationary Phase Approximation of the Arc Radial Functions . . . 146

8 Conclusions 147
8.1 Tangential Pressure Gradient Array . . . 147
8.2 DOA Estimation with a Finite Rate of Innovation Sound Field Model . . . 149
8.3 A General Radiation Model and Nearfield Acoustical Holography in Wedge Propagation Spaces . . . 150
8.4 Loudspeaker Arrays for Beamforming in Infinite Wedge Spaces . . . 151
8.5 General Outlook . . . 153

A Three Dimensional Sound Field Acquisition Based On The Tangential Component Of The Pressure Gradient 155
A.1 Model of the Sound Field . . . 155
A.1.1 Angular Derivative in θ . . . 156
A.1.2 Angular Derivative in ϕ . . . 156
A.1.3 The Tangential Pressure Gradient . . . 157
A.2 Solution to the Inverse Problem . . . 158
A.2.1 Singular Value Decomposition of G_s . . . 158
A.2.2 Solution Using Linear Algebra . . . 159

B A De-Aliasing Strategy Based On Signal Processing 161
B.1 Sound Field Model . . . 162
B.2 A New Set of Basis Functions For Sound Field Decomposition . . . 163
B.2.1 Band Limited HD of a Single Plane Wave . . . 164
B.2.2 Band Limited HD of M Plane Waves . . . 166
B.2.3 Reconstructing the HD With The New Basis . . . 166
B.2.4 Autocorrelation Function and Orthogonality Property of the Circular Sinc Function . . . 167
B.3 Methods to Identify the Parameters of the New Basis . . . 168
B.3.1 Method I: Modified CLEAN Algorithm . . . 169
B.3.2 Method II: Finite Rate of Innovation (FRI) Approach . . . 170
B.3.3 Method III: Compressive Sensing . . . 173
B.3.4 Performance of the CLEAN and the FRI Method for $M > N$ . . . 174
B.4 Sound Field Decomposition with the New Basis . . . 174
B.4.1 Solving the Inverse Problem with the New Basis . . . 175
B.4.2 Average Plane Wave Identification Error . . . 175
B.5 De-Aliasing Stage . . . 177
B.5.1 Predicting Aliasing of a Single Plane Wave . . . 177
B.5.2 De-Aliasing of Sound Fields with M Plane Waves . . . 181
B.5.3 Position of the Side Lobes in the Aliased HD of a Single Plane Wave . . . 183
B.6 Performance Study of the Combination of the SFA Stage and the De-Aliasing Stage 184
B.6.1 Definition of the Error-To-Signal Ratio . . . 185
B.6.2 Statistical Performance Analysis . . . 185
B.6.2.1 Results for the De-Aliasing Strategy With CLEAN Parameter Estimation . . . 185
B.6.2.2 Results for the De-Aliasing Strategy With FRI Parameter Estimation . . . 186

B.7 Conclusions . . . 187

C Cross-Correlation of the HD and the Weighted Circular Sinc Function 191

D A Multichannel Low-Noise Signal Conditioning Amplifier for Microphone Arrays 193
D.1 Circuit Design & Specifications . . . 193
D.1.1 Full Electronic Circuit . . . 193
D.1.2 Microphone Conditioning Amplifier: Specifications . . . 193
D.1.3 Power Supply Unit (PSU): Specfications . . . 194
D.1.4 Sub-D-78 Connector: Pin-Layout . . . 195
D.2 Performance Specifications . . . 195
D.2.1 Conditioning Amplifier . . . 195
D.2.2 Frequency Response & THD+Noise Measurements . . . 196
D.2.3 Power Supply Unit (PSU) . . . 197

E Design & Realisation of a High-Quality Hemi-Cylindrical Array 199
E.1 Design Sketches . . . 199
E.2 Construction of the Array . . . 203
E.2.1 Enclosure . . . 203
E.2.2 Driver Units . . . 203
E.2.3 Front Panels . . . 204
E.2.4 The Assembled HQ Hemi-Cylindrical Array . . . 204

F Plane Wave Identification With Circular Arrays By Means Of A Finite Rate Of Innovation Approach 207
F.1 Introduction . . . 207
F.2 Sound Field Model . . . 208
F.3 Recovering the Herglotz Density from Circular Array Measurements . . . 209
F.4 Finite Rate of Innovation Approach . . . 210
F.4.1 HD of a Single Plane Wave . . . 210
F.4.2 New Model for the Herglotz Density . . . 211
F.4.3 Definition of the Annihilating Filter . . . 211
F.4.4 Calculation of the Annihilating Filter . . . 213
F.4.5 Recovering the Complex Amplitudes of the Individual Plane Waves . . . 213
F.5 Performance Evaluation . . . 214
F.5.1 An Example . . . 214
F.5.2 Simulation Study . . . 215
F.6 Conclusion . . . 216
F.7 Acknowledgements . . . 216

List of Figures

1.1 Illustration of the concept to observe the tangential component of the pressure gradient rather than the pressure with a circular array on the surface of a rigid cylindrical scatterer. 3
1.2 Example of a plane wave (or Herglotz) density $q(\phi)$ with a finite number of Dirac pulses, i.e. a finite rate of innovation, within the period of 2π. The periodicity is inherent to the circular domain. 5
1.3 Birds-eye view into the infinite wedge propagation space with the curved radiating surface. A finite wedge propagation space has additionally a floor and ceiling at specific values along the z-axis. 6
1.4 Cross section of the infinite wedge propagation space with the loudspeakers arranged on an arc with radius r_S. 8

2.1 Sound field scenario where the HHE is valid inside the region Λ, which is centred around the coordinate system's origin, and with acoustic source/scatterer distributions S_1 and S_2. 13
2.2 Sound field scenario where the HHE is valid outside the region V, which is centred around the coordinate system's origin and which contains the source/scatterer distributions S_1 and S_2. 13
2.3 Infinite wedge scenario, with the source distribution S located inside the radius r_V, and boundary conditions at ϕ_1 and ϕ_2. 16

3.1 Cartesian coordinates with the three orthogonal coordinate axes. 38
3.2 Cylindrical coordinates in the context of Cartesian coordinate axes. 39
3.3 Spherical coordinates in the context of Cartesian coordinate axes. 40

4.1 (left) Schematic description of the volumes and boundaries involved in the sound field model, (right) schematic description of an infinite cylinder of radius r_V. . . 45
4.2 Magnitude of the eigenvalues $|\lambda_n|$ of G for $n = 1 \dots 7$, $f = 10 \dots 20 \times 10^3$ Hz for an array on a rigid cylindrical structure at the origin with $r_V = 0.1$ m. 50
4.3 Magnitude of the eigenvalues $|\nu_n|$ of H of a conventional array based on pressure sensors on a rigid cylindrical structure with $r_V = 0.1$ m for $n = 0 \dots 7$, $f = 10 \dots 20 \times 10^3$ Hz. 51
4.4 Visualisation of the orthogonality matrix $A_{m,n}$ for a circular array with $L = 15$ microphones. Each cell represents whether two modes a_m and a_n are orthogonal (white cell) or not (black cell), when sampled at L points on a circle. The two light grey lines indicate the mode range ($\pm N$) of the array. 54
4.5 Schematic visualisation of the new orthogonality matrix $\tilde{A}_{m,n}$ for a circular array with $L = 15$ HD sensors and one pressure sensor. The two light grey lines indicate the mode range ($\pm N$) of the array. 59
4.6 Mode recovery performance of a circular array composed of $L = 15$ TPG sensors and one pressure sensor with $r_V = 0.1$ m. The incoming sound field has a frequency $f = 5$ kHz and the results are shown for $N_s = [7, 8, 12, 14]$. The dashed grey line indicates the theoretical values for the magnitude of the HD coefficients, while the black stems display the recovered coefficients' magnitude. 60

4.7 Visualisation of the sound field with $\phi_i = 0$, $f = 5\,\text{kHz}, c = 343\,\frac{\text{m}}{\text{s}}$, limited to order $N_s = 8$, (A) original sound field, (B) sound field as reproduced from the observation of $\tilde{q}_n$, using the microphone array described above, (C) absolute amplitude error of the reproduced field, given in dB. 61
4.8 Plots of the average ESR of the recovered coefficients $\tilde{q}_{n,R}$ from a simulated TPG sensor array measurement (black line) and from a simulated pressure sensor array measurement (light grey line) within the frequency range from 100 Hz to 10 kHz. 64
4.9 Visualisation of a plane wave approximation with $\phi_i = \frac{\pi}{7}, f = 800\,\text{Hz}, c = 343\,\frac{\text{m}}{\text{s}}$, limited to order $N_s = 15$, (A) original sound field, (B) synthesis based on TPG array measurement with $\beta = \sqrt{10^{\frac{-50}{20}}}$, (C) synthesis based on pressure array measurement with $\beta = \sqrt{10^{\frac{-50}{20}}}$, (D) synthesis based on TPG array measurement with $\beta = 0$, (E) synthesis based on pressure array measurement with $\beta = 0$. . . . 66

5.1 Cylindrical coordinate system with $\mathbf{r} = (r, \phi, z)^T$, the elevation angle θ and other entities as they are used in this work. 69
5.2 HD for $L = 5$ distinct plane waves, where the HD is periodic with $q(\phi) = q(\phi + u2\pi), \forall u \in \mathbb{Z}$. 72
5.3 Radial functions for a microphone array with radius $R = 0.075\,\text{m}$ and $f_c \approx 5.1\,\text{kHz}$ ($N = 7$, $c = 343\,\frac{\text{m}}{\text{s}}$). 78
5.4 Conceptual z-domain spectrum of the HD coefficients for the sound field of a PW and a PS. 79
5.5 Results of different FRI estimation performances for $L = 5$ and $N = K = 4$. . . . 80
5.6 Mean performance of the FRI estimator over 10,000 samples for different combinations of SNR and L. 82
5.7 DEE for a single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen φ_l. 83
5.8 DEE for a single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen DOA, as obtained with a greater coefficient range. 83
5.9 DEE for one OOP-PW with $\theta_2 \in [0.01\pi \ldots 0.99\pi]$ and one with $\theta_1 = \frac{\pi}{2}$, averaged over 100 samples with randomly chosen φ_1 and φ_2. 84
5.10 Mean DEE for a randomly positioned PS as a function of the radius, averaged over 100 samples per evaluation radius. 84
5.11 Mean DEE for two randomly positioned PSs as a function of r_2, averaged over 100 samples per evaluation radius. 85
5.12 Sketch of the arrangement of microphone array and sound sources inside the anechoic chamber. 89
5.13 (Left) Angle laser gauge with Genelec 8020C, (Right) Close-up view of tweeter with laser reflection on the diaphragm. 89
5.14 Mean and median of the DEE in %, averaged over 1,000 randomised iterations for various number of sources. 90

6.1 Cylindrical Coordinates with radius r, azimuth angle ϕ and height z, as well as further boundaries and specific angles needed in this work. 98
6.2 Concept of the radiation from a curved boundary V with the given velocity profile $v_r(r_S, \phi, z)$ inside the wedge defined by the rigid boundaries S'. 102
6.3 Magnitude of the radial functions $\Gamma_n(k_z, r, r_S, \kappa)$ in [dB] with reference to unity, $r = 2\,\text{m}$, $r_S = 0.15\,\text{m}$, $\kappa = \pi$, $f = 5\,\text{kHz}$. 107
6.4 Pressure field of a point source located at $(0.2\,\text{m}, \frac{\pi}{4}, 0)^T$, $Q = 1.5$, $\Delta k_z = 0.01\,\frac{\text{rad}}{\text{m}}$. 110
6.5 Pressure field of a piston located at $(0.2\,\text{m}, \frac{\pi}{4}, 0)^T$, $Q = 1.5$, $\Delta k_z = 0.01\,\frac{\text{rad}}{\text{m}}$, $\alpha = \frac{\pi}{6}$, $B = 0.1\,\text{m}$. 111
6.6 The measurement setup with a measurement array consisting of 30 *G.R.A.S. 40PL* array microphones fitted on an arch over the HCA with radius $r_H = 0.3\,\text{m}$. . . . 112

6.7 Modal decay from source surface ($r_S = 0.1426\,\mathrm{m}$) to hologram surface ($r_H = 0.3\,\mathrm{m}$) computed from $\Gamma_N^{(C)}$ within the audible frequency band. The data is normalised to the model strength on the surface of the loudspeaker array. 113
6.8 Magnitude and phase of both the measured data and the piston simulation for a driver at $(r_S, \frac{\pi}{2}, 0.0315\,\mathrm{m})^T$, $f = 8\,\mathrm{kHz}$. 115
6.9 Spatial spectrum of the measured data vs. simulated measurement. 115
6.10 (A) Velocity Reconstruction on the prototype surface, (B) Detailed excerpt with outline of driver, $f = 8\,\mathrm{kHz}$. 116
6.11 Measured pressure data vs. simulated data from piston model for all 15 drivers active, $f = 8\,\mathrm{kHz}$. 117
6.12 Spatial spectrum of the measured vs. simulated pressure, $f = 8\,\mathrm{kHz}$. 117
6.13 Reconstruction of v_r on the prototype surface with overlaid circles to indicate the driver positions from both measured (left) and simulated data (right), $f = 8\,\mathrm{kHz}$. Note that the horizontal axis is given as $x(\phi) = \phi r_S$. 117
6.14 Normalised driving signal strength of the 15 drivers, $f = 8\,\mathrm{kHz}$. 118

7.1 Propagation space with wedge boundaries described in cylindrical coordinates. . 123
7.2 Wedge boundaries projected into the xy-plane with the loudspeaker array mounted on the curved baffle. 124
7.3 Visualisation of the orthogonality matrix, $L = 15$. 129
7.4 Magnitude of Γ_n^{Ψ}, $\kappa = \frac{\pi}{2}$, $L = 15$, $z_S = z_B = 0\,\mathrm{m}$. 131
7.5 Magnitude Γ_n^{Ψ} with $\kappa = \pi$, $L = 15$, $r_B = 2\,\mathrm{m}$, $r_S = 0.2\,\mathrm{m}$, $z_S = z_B = 0\,\mathrm{m}$. 132
7.6 Magnitude of $\Gamma_{n,\mathrm{FF}}^{\Psi}$, $\kappa = \frac{\pi}{2}$, $L = 15$, $R = r_B = 2\,\mathrm{m}$, $\theta = \frac{\pi}{2}$, $z_S = 0\,\mathrm{m}$. 132
7.7 Examples for Beam Pattern I with different steering angles ϕ_B, $\kappa = \pi$. 135
7.8 Examples for Beam Pattern II with different steering angles ϕ_B and $D = 0.2\pi$, $\kappa = \pi$. 135
7.9 Simulated beam patterns for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\mathrm{m}$ and $r_S = 0.2\,\mathrm{m}$. 136
7.10 Simulated beam patterns for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\mathrm{m}$ and $r_S = 0.2\,\mathrm{m}$. 137
7.11 Simulated beam pattern for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\mathrm{m}$ and $r_S = 1\,\mathrm{m}$. 138
7.12 Prototype of a HCA with $r_S = 14.2\,\mathrm{cm}$ and $L = 15$ *Tang Band W1-1070SG* full-range drivers. 138
7.13 Visualisation of the orthogonality matrix $A_{n,m}$ for the full prototype and the bottom layer only. 139
7.14 Simulated beam patterns for an HCA with $L = 15$, $N = 10$, $\beta = 0.1$, $\phi_B = \frac{\pi}{2}$ and $r_B = 2\,\mathrm{m}$. 140
7.15 Simulated beam patterns for an HCA with $L = 8$, $N = 7$, $\beta = 0.1$, $\phi_B = \frac{\pi}{2}$ and $r_B = 2\,\mathrm{m}$. 140
7.16 Theoretically achievable beam patterns for an array with $L = 15$ speakers, with $N = 10$, $\beta = 0.1$ and $r_B = 2\,\mathrm{m}$. 141
7.17 The measurement setup with a measurement array consisting of 30 *G.R.A.S. 40PL* array microphones fitted on an arch over the HCA with 30 cm radius. 142
7.18 HCA performance for different steering angles ϕ_B and beam patterns. 144
7.19 HCA performance for different steering angles ϕ_B, beam patterns and arc radial filters. 145

B.1 Plot of the circular sinc function defined in (B.14) of order $L = 15$ with $\theta = \pi$. . 165
B.2 Plot of the autocorrelation function of $\mathrm{csinc}_{2N+1}(\phi)$. 168
B.3 Set of graphs describing the iterative basis acquisition of Method I: a) original HD, b) visualisation of iteration step $m = 1$, c) visualisation of iteration step $m = 2$, d) visualisation of iteration step $m = 3$, e) visualisation of iteration step $m = 4$, f) residual after the algorithm terminated. 170

B.4 Example of an original HD (top) and the basis functions obtained through Method II (bottom) for $M = 4$ and $N = 7$. 173
B.5 a) Plot of the original HD, b) reconstructions of the HD based on the two bases gained through Method I and II, c) basis functions identified with Method I, d) basis functions identified with Method II, e) reconstruction error of Method I, f) reconstruction error of Method II. 176
B.6 a) Sound field corresponding to the original HD in Fig. B.5, b) sound field corresponding to the HD reconstructed with Method I, c) sound field corresponding to the HD reconstructed with Method II, d) sound field error with Method I, e) sound field error with Method II, $f = 1$ kHz. 177
B.7 Plots of the recovered Herglotz coefficients $\tilde{q}_n$ of a single plane wave with $\theta = \frac{\pi}{2}$ that were recovered in different frequency bands ($f = 2, 6, 10, 16$ kHz) from a simulated circular array with $L = 15$, $N = 7$, $r_V = 10$ cm and a maximum sound field order of $|n| = N_s = 40$. 178
B.8 Plots of the recovered Herglotz Densities $\tilde{q}(\phi)$ of a plane wave travelling in the direction $\theta_m = \frac{\pi}{2}$ from coefficients $\tilde{q}_n$ that were recovered in different frequency bands ($f = 2, 6, 10, 16$ kHz) from a simulated circular array with $L = 15$, $N = 7$, $r_V = 0.1$ m and a maximum sound field order of $|n| = N_s = 40$. 179
B.9 a) HD of a single plane wave approximation up the 40th order with $\theta = \frac{4\pi}{7}$ at the frequency $f = 16$ kHz, b) HD as recovered from a simulated circular array with $L = 15$, $N = 7$ and $r_v = 0.1$ m, c) recovery error $q(\phi) - \tilde{q}(\phi)$, d) predicted aliased HD for $N_s = 7$, e) predicted aliased HD for $N_s = 20$, f) predicted aliased HD for $N_s = 40$, g) recovery performance for $N_s = 7$, h) recovery performance for $N_s = 20$, i) recovery performance for $N_s = 40$. 180
B.10 a) HD of a sound field with $M = 4$ plane wave approximations up the 40th order with $\theta = \frac{4\pi}{7}$ at the frequency $f = 16$ kHz, b) HD as recovered from a simulated circular array with $L = 15$, $N = 7$ and $r_v = 0.1$ m, c) recovery error $q(\phi) - \tilde{q}(\phi)$, d) predicted aliased circular sinc functions for $N_s = 7$, e) predicted aliased circular sinc functions for $N_s = 20$, f) predicted aliased circular sinc functions for $N_s = 40$, g) recovery performance for $N_s = 7$, h) recovery performance for $N_s = 20$, i) recovery performance for $N_s = 40$. 182
B.11 Plots of the aliased Herglotz Densities $q_{N_s}^{(a)}(\phi)$ of a plane wave travelling in the direction $\theta = \frac{4\pi}{7}$ for $f = 16$ kHz, $L = 15$, $N = 7$, $r_V = 0.1$ m and different maximum sound field orders $N_s \in [7, 10, 15, 20, 25, 30, 35, 40]$. 184
B.12 Histograms (or estimated Probability Density Functions (PDF)) of the Error-to-Signal-Ratio (ESR) for the De-Aliasing Strategy based on the CLEAN parameter estimation for various M and different aliasing model parameters N_s, for $f = 16$ kHz and the array parameters $L = 15$, $N = 7$, $r_V = 0.1$ m. 186
B.13 Histograms (or estimated Probability Density Functions (PDF)) of the Error-to-Signal-Ratio (ESR) for the De-Aliasing Strategy based on the FRI parameter estimation, where the simulations encompassed 10000 randomly generated Herglotz Densities, which have been taken out for various M and different aliasing model parameters N_s, for $f = 16$ kHz and the array parameters $L = 15$, $N = 7$, $r_V = 0.1$ m. 187

D.1 Schematic of the conditioning amplifier. 194
D.2 Schematic of the gain control stage. 194
D.3 Schematic of the gain switching circuit. 195
D.4 Pin layout for the Sub-D-78 connector. 196
D.5 Frequency responses of channels 1 and 16 of all four gain settings for an input signal with 1.25 mV RMS ($\equiv -55.85$ dBu) within the frequency band from 20 Hz up to 40 kHz. 197
D.6 THD+Noise of channels 1 and 16 of all four gain settings for an input signal with 1.25 mV RMS ($\equiv -55.85$ dBu) within the frequency band from 20 Hz up to 40 kHz. 198

E.1 Design of the lower layer of the Mirrorsys prototype. 200

E.2 Design of the upper layer of the Mirrorsys prototype, with filler wedges for both, hemi-cylinder array and full array applications. 201
E.3 Design of the front plate for the speaker compartments in the upper and lower layer. 202
E.4 Enclosure of the HQ hemi-cylindrical array with damping material (upholstery cotton felt) in place. 203
E.5 Visible part of the Tang Band W1-1070SG driver. 204
E.6 Aluminium panel with driver in place and silicone sealing along the edges. 204
E.7 Fully assembled HQ hemi-cylindrical array 205

F.1 Example of the results achieved with the proposed method for $M = 4$ and $N = 7$, (Top) real- and imaginary part of the original HD, (Middle) real- and imaginary part of the reconstructed HD, (Bottom) absolute reconstruction error as an indicator for the parameter estimation performance. 214

Nomenclature

Operators and General Symbols

$:=$	definition
$\in$	...is an element of...
$:$	...such that...
$\forall$	...for all...
$\{x : \ldots\}$	definition of a set by stating the properties of its elements x
$\{\cdot, \ldots, \cdot\}$	definition of a set by a list of its members
$\partial\cdot$	boundary of a set
$< \cdot \vert \cdot >_D$	scalar product as defined in the domain D
$\Vert \cdot \Vert_m^{(D)}$	l_m norm ($m = [1, 2]$) as defined in the domain D
$*$	convolution
i	imaginary unit $i = \sqrt{-1}$
$(\cdot)^*$	complex conjugate of a complex number or adjoint operator
$\arg(\cdot)$	argument (or polar angle) $\varphi = \arg z$
	of a complex number $z = \vert z \vert e^{i\varphi}, z \in \mathbb{C}$
∇	Nabla-operator
Δ	Laplacian-operator

Special functions

$\delta(\cdot)$	Dirac delta distribution
δ_n	Kronecker delta
$\Pi(\cdot)$	Rectangular Pulse
$J_n(\cdot)$	n-th order Bessel function
$Y_n(\cdot)$	n-th order Neumann function
$H_n^{(1)}(\cdot)$	n-th order Hankel function of the first kind
$H_n^{(2)}(\cdot)$	n-th order Hankel function of the second kind
$j_n(\cdot)$	n-th order spherical Bessel function
$y_n(\cdot)$	n-th order spherical Neumann function
$h_n^{(1)}(\cdot)$	n-th order spherical Hankel function of the first kind

$h_n^{(2)}(\cdot)$	n-th order spherical Hankel function of the second kind
$P_n(\cdot)$	n-th order Legendre polynomial
$P_n^m(\cdot)$	n-th order m-th degree associated Legendre function
$Y_n^m(\cdot)$	n-th order m-th degree spherical harmonic
$\mathrm{sinc}(\cdot)$	sinc function
$\mathrm{csinc}(\cdot)$	circular sinc function

Further functions, Coefficients and physical quantities

q	Herglotz Density
$p(\mathbf{x})$	acoustic pressure $\left[\frac{\mathrm{kg}}{\mathrm{m}\cdot\mathrm{s}^2}\right]$
$G(\mathbf{x},\mathbf{y})$	Green function of Helmholtz equation of the free space $[\mathrm{m}^{-1}]$
$G_N(\mathbf{x},\mathbf{y})$	Neumann-Green function $[\mathrm{m}^{-1}]$
t	time in s
c	speed of sound propagation $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
ρ_0	static density of the medium of acoustic wave propagation $\left[\frac{\mathrm{kg}}{\mathrm{m}^3}\right]$
f	acoustic frequency in Hz
ω	angular frequency $[s^{-1}]$
k	acoustic wave number $\left[\frac{\mathrm{rad}}{\mathrm{m}}\right]$
$C_n(k_z,\omega)$	Coefficients for the interior domain problem (cylindrical coordinates)
$A_n(k_z,\omega)$	Coefficients for the exterior domain problem (cylindrical coordinates)
$C_{mn}(\omega)$	Coefficients for the interior domain problem (spherical coordinates)

Sets, Vectors and Matrices

Λ	region where the homogeneous wave equation is satisfied
$\partial\Lambda$	boundary of Λ
Ω	unit sphere and unit circle, respectively ($\Omega := \{\mathbf{x} \in \mathbb{R}^m : m = [2,3], \lvert\mathbf{x}\rvert = 1\}$)
$\mathbb{N}$	the set of natural numbers including zero
$\mathbb{N}^+$	the set of natural numbers excluding zero
$\mathbb{Z}$	the set of integer numbers
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^m$	the m-dimensional Euclidean space
	$\mathbb{R}^m := (\mathbf{x} = (x_1, x_2, \ldots, x_m)^T : x_1, x_2, \ldots, x_m \in \mathbb{R})$
$\mathbb{C}$	the set of complex numbers
$\mathbf{x}$	vector describing a point in $\mathbb{R}^m$
$\hat{\mathbf{y}}$	unitary direction vector in $\mathbb{R}^m$
$\mathbf{I}$	identity matrix with $I_{ij} = \delta_{ij}$
$\mathrm{diag}(\cdot)$	diagonal matrix

Acronyms

B&K	Brüel & Kjær
CICS	Circular Integrated Cross Spectrum
CS	Compressed Sensing
DEE	Direction Estimation Error
DOA	Direction-of-Arrival
ESR	Error-to-Signal-Ratio
EVD	Eigenvalue Decomposition
FDM	Finite Difference Method
FEM	Finite Element Method
FFT	Fast Fourier Transform
FRI	Finite Rate of Innovation
HD	Herglotz Density
HHE	Homogeneous Helmholtz Equation
HOA	Higher Order Ambisonics
HWE	Homogeneous Wave Equation
HWF	Herglotz Wave Function
KIE	Kirchhoff Integral Equation
MEMS	Microelectromechanical System
NAH	Nearfield Acoustical Holography
NGF	Neumann-Green Function
OOP-PW	Out-Of-Plane Plane Wave
PS	Point Source
PW	Plane Wave
SLP	Single Layer Potential
SNR	Signal-to-Noise-Ratio
SVD	Singular Value Decomposition
THD	Total Harmonic Distortion
TLS	Total Least-Squares
TOS	Total Orthonormal Set
TPG	Tangential Pressure Gradient
WFS	Wave Field Synthesis

Declaration of Authorship

I, Falk-Martin Hoffmann, declare that this thesis entitled 'Advances in Sound Field Analysis and Control based on Cylindrical Coordinates' and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
    - F.-M. Hoffmann, F.M. Fazi. Circular microphone array with tangential pressure gradient sensors. In *4th Joint Workshop Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014*, pages 97-101, May 2014.
    - F.-M. Hoffmann, F.M. Fazi. Theoretical study of Acoustic Circular Arrays with Tangential Pressure Gradient Sensors. IEEE/ACM Transactions on Audio, Speech, and Language Processing, pages 1762-1774, Vol. 23, No. 11.
    - F.-M. Hoffmann, F.M. Fazi, P.A. Nelson. Plane Wave Identification with Circular Arrays by Means of a Finite Rate of Innovation Approach, In *140th Convention of the Audio Engineering Society, Paris 2016, May.*
    - F.-M. Hoffmann, F.M. Fazi, S. Fontana. Sound field control with hemi-cylindrical loudspeaker arrays. In *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*, 2016.

Signed:

Date:

Joint Authorship on Manuscripts presented in Chapters 4 to 7

I, Falk-Martin Hoffmann, declare that the chapters 4, 5, 6 and 7 of this thesis are the manuscripts of journal papers that, at the time of submission of this thesis, have either been published, have been submitted to a journal or are in the stage of final preparation for submission to a reputable journal.

I confirm that:

1. The research that led to the presented manuscripts was done wholly as part of this PhD project;
2. At the beginning of the corresponding chapters, all co-authors of these manuscripts are given with full name;
3. All manuscripts were written by myself as principal author in my own words;
4. Dr. Filippo Maria Fazi and Prof. Philip Arthur Nelson were forming the supervisory team in this PhD project and have second and/or third authorship on these manuscripts where appropriate;
5. Dr. Earl G. Williams of the Naval Research Laboratory, Washington D.C., has contributed to the manuscripts shown in Chapters 6 and 7 through his vast experience in this field, pointing out relevant previous work and providing valuable feedback on the manuscripts during the writing phase;
6. Dr. Simone Fontana of the HUAWEI European Research Center, Munich, Germany, was the collaboration lead on a research project funded by the company on compact loudspeaker arrays. He contributed to the scientific discussions related to the work presented in Chapters 6 and 7 and thus had fourth authorship on the corresponding manuscripts;

Signed:

Date:

Certification by the PhD Project Supervisor

I, Dr. Filippo Maria Fazi, certify the author's above declaration through my signature.

Signed:

Date:

Acknowledgements

I would like to express my profound gratitude to my supervisor Dr. Filippo Maria Fazi, his support during this project has been truly invaluable. His encouragement and guidance will continue to help and drive me onwards, far beyond the completion of my doctoral training. He invested what feels like uncountable hours discussing and revising my work, never ceasing to inspire me with his own curiosity to look even deeper into the matter of what I was working on, or to look into something entirely new. Dr. Fazi has played a crucial role in my scientific and personal development. Over the course of this Ph.D. project working with him has been a truly enriching experience.

I would also like to thank my supervisor Prof. Philip Arthur Nelson, for his guidance, experience and encouraging support. Despite being involved in commitments of much greater importance, he managed to find time to provide helpful, sustainable and also critical feedback, whenever it was needed. He never held back on expressing enthusiasm about the progress I had made and often helped me to see my achievements in a more positive light.

I am also extremely grateful for my collaboration with Dr. Earl G. Williams from the Naval Research Laboratory, Washington D.C., with whom I had great scientific discussions about the research on Nearfield Acoustical Holography in Wedge Propagation Spaces. His book *Fourier Acoustics* has been a fountain of knowledge as a basis for my work and research.

For his hint in direction of the Finite Rate of Innovation sampling theory, I would like to acknowledge Dr. Thomas Blumensath. The discussion we had after my 9-month report led to an entire chapter in this thesis. The research on this part of my Ph.D. project has been extremely interesting until the very end.

A big "thank you" goes to my friend and colleague, Fábio Casagrande Hirono, for unhesitantly helping with my experiments during weekends, evenings and even more for collaborating on the development (and construction) of the multichannel microphone conditioning amplifier used in this thesis. The quality of the final product need not shun comparisons.

Many thanks also to all my other colleagues from the Virtual Acoustics and Audio Engineering group, Dr. Andreas Franck, Dr. Keith Holland, Dr. Dylan Menzies-Gow, Dr. Marcos Simon-Galvez and Michael Cousins for their help and patience with various 'quick questions', but

particularly to Dr. Ferdinando Olivieri for sharing his experience with compact arrays and for his amazing collaboration during the HUAWEI project.

I also want to acknowledge Tom Bell from *Bowers & Wilkins* for kindly helping with the electrical performance measurements of the developed microphone conditioning amplifiers.

On a non-professional level, I want to deeply thank my family for their emotional support during all of this time. Most of all my mother, father and my deeply missed uncle. Their absolute support and their trust in me has been, and will always be, an invaluable gift.

Another big "thank you" goes to all my friends, particularly to Guido Zarini, Nikhil Mistry, my 'Wilton housemates' and everyone from the 'coffee room'. Their friendship has helped me stay positive during the most demanding period of my life.

Chapter 1

Introduction

1.1 Historical Background and Placement of this Thesis into the Scientific Context

The engineering challenge to capture a sound field and control or even reproduce it at another place and time presented itself with the invention of the first microphones and loudspeakers. A single pressure sensitive microphone is evidently able to capture sound at that point in space, while a single loudspeaker can radiate a pressure field that is proportional to the input signal. As such, both a single microphone and a single loudspeaker are a means to solve the corresponding problems, however only in a wider sense of interpretation. Being able to reconstruct a wave field implies that the data observed by the capturing system carries sufficient information on the original wave field. More rigorously speaking, the capture of an entire sound field requires the unrestricted knowledge of the pressure and particle velocity at *any* point within the considered volume. Equally, controlling a sound field implicates the ability to match both the pressure and particle velocity within a control volume to an arbitrary target field. Such ambitious objectives necessitate a sound mathematical model for the physics of any possible wave field under consideration and the transducer technology to keep the practical limitations of the model down to an acceptable level. It is known from well-established acoustic models that a system solving the problem, strictly speaking, requires ideally continuous scanning and a continuous layer of acoustic sources, respectively, on the boundary of the region of interest.

Substantial theoretical and practical work has been done in this field over the past 140 years since Thomas Edison invented and patented his phonograph. The quality of transducers for capture and reproduction was continuously improved, the devices became smaller and more robust. In parallel the signal processing power increased and the precision of (industrial) machine-assisted manufacture improved, surpassing levels of human ability. Theoretical ground work on sound fields inside undisturbed volumes and the acoustic scattering off geometrically simple objects became the foundation of theoretically powerful and well-researched models. These have been used and refined to form the basis of the various sound field capture, analysis and reproduction techniques proposed over the past four decades. Particularly research on sound field analysis combined signal processing with acoustic modelling, spawning powerful algorithms that identify

the arrival directions of wave fronts impinging into the region of interest from sources located outside. For certain sound fields, such information can be used to dramatically reduce the data describing a sound field comprehensively and even provide upscaling mechanisms. The latter serve to push a system's spatial resolution beyond the theoretical bounds of its underlying model. It follows that a model should not only be judged by its baseline accuracy but also by the potential it bears for further signal processing strategies to be applied as a means of post-processing to further improve the overall system performance.

This thesis is concerned with advancing developments for sound field analysis and control applications that can be described through an acoustic wave field model in cylindrical coordinates. These advances address microphone array designs, signal processing strategies for plane wave identification, and acoustic field modelling for *Nearfield Acoustical Holography* (NAH) and beamforming in acoustic propagation spaces with specific boundary conditions. These developments are motivated by both theoretical and practical limitations as well as modelling mismatches of state-of-the-art transducer arrays, and their associated methods and techniques. These were well-investigated and described over the past two decades so that now, with a sound understanding of the shortcomings, the performance and resolution of the associated array systems can be improved through substantial design changes or advanced signal processing strategies. Especially the latter have experienced a surge of interest in the research community over the past ten years.

The conducted investigations have a common anchor point in the general solution of the Helmholtz Equation in cylindrical coordinates as it can be obtained through a variable separation approach. Among the published research (which often focuses on the audio frequency range), problem descriptions in cylindrical coordinates predominantly find applications for two-dimensional approximations of acoustic scenarios but rarely for three-dimensional considerations. It will be seen in the course of this thesis that cylindrical coordinates also enable convenient three-dimensional modelling of practically relevant wave propagation spaces, combining high accuracy with a compact mathematical representation. The option to simplify a consideration to two dimensions of course remains, since the required steps come inherently trouble-free with the chosen coordinate system.

The nature of this specific general solution to the Helmholtz Equation facilitates a mathematical analysis of the solution through methods of functional decomposition, e.g. *Eigenvalue Decomposition* (EVD) and *Singular Value Decomposition* (SVD). These provide a powerful framework to pin-point effects that determine the performance of acoustic transducer array systems to a physical cause or explanation. From a mathematical point of view, functional analysis provides a holistic approach to consider the acoustical forward and inverse problems that describe the applications of sound field capture and control.

1.2 Contributions and Structure of the Thesis

This thesis is conceived as a *thesis-by-papers*, where the four principal contributions are presented in four separate chapters, each of which forms the manuscript of a journal paper. These have either been published, submitted and are currently undergoing revision, or are in the final stage of preparation for submission to a reputable journal. Each of these chapters thus encompasses an

individual introduction, a literature review, the main contribution and conclusions. They can therefore be understood as independent scientific contributions.

The remainder of the thesis is to be understood as a framework that puts the main chapters into mutual perspective, concludes the overall findings and gives an outlook onto prospective future research. Additional work that was done through the course of this Ph.D. project is presented in the appendices. The content of the subsequent chapters and the scientific contributions of this thesis are summarised in the following.

Chapter 2 gives an overview on previous work and findings that were used as the basis upon which the presented investigations are founded. It covers wave field models, acoustical inverse problems, microphone array designs and signal processing strategies, and compact loudspeaker arrays for sound field control.

The mathematical notation, definitions, functions and coordinate systems used throughout this thesis are introduced in Chapter 3.

1.2.1 Journal Paper: Theoretical Study of Circular Arrays with Tangential Pressure Gradient Sensors

Chapter 4 is the manuscript of a paper that was published in the *IEEE/ACM Transactions on Audio, Speech and Language Processing*. The principal contribution of this chapter is the investigation of the theoretical performance of a circular microphone array based on sensors that observe the component of the pressure gradient that is tangential to the array's rigid corpus (see Figure 1.1). Such a design with a spherical geometry was proposed by Peter Craven, Chris Travis

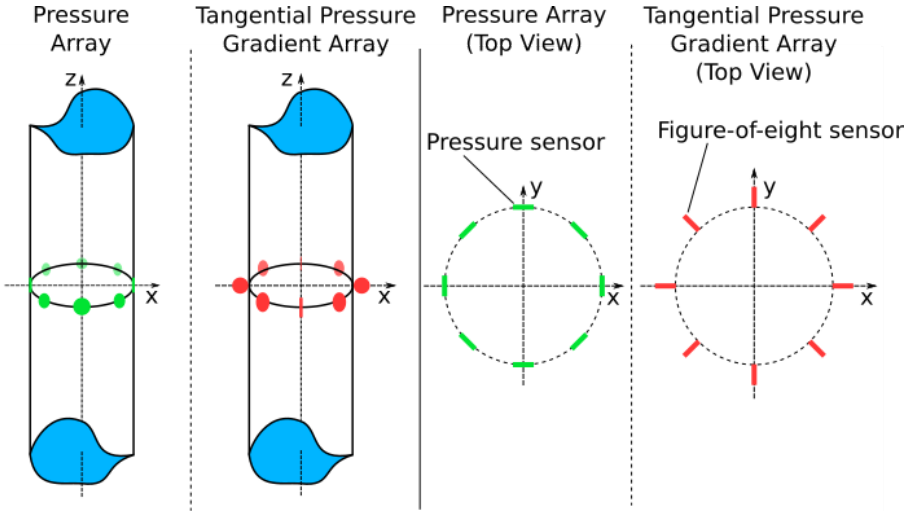


FIGURE 1.1: Illustration of the concept to observe the tangential component of the pressure gradient rather than the pressure with a circular array on the surface of a rigid cylindrical scatterer.

and Malcolm Law in 2009, claiming that their design reduces the adverse effects of transducer noise on the spatial resolution at low frequencies. With their idea, Craven et al. were addressing the problem that transducer arrays typically have a very narrow frequency band in which they perform best. If their claims were confirmed, one could hope that such an array would have

a wider optimal frequency band than conventional arrays based on pressure sensors. To the knowledge of the author, a rigorous modelling and analysis of the capturing mechanism based on the observation of the *Tangential Pressure Gradient* (TPG) had not been presented, which motivated the work that led to the publication of the results from this Ph.D. project.

The mathematical operator describing the capturing mechanism of such an array was analysed on the basis of an EVD. An inspection of the eigenvalues that map a generic quantity describing the sound field to the TPG serves to assess the capturing resolution of the array at different frequencies. The properties that are responsible for spatial aliasing, ill-conditioning and non-uniqueness of the inverse solution are thoroughly studied, resulting in the presentation of analytical expressions for the inverse solution and the spatial aliasing patterns of the design, and the discovery that at least one pressure sensor is required to solve a non-uniqueness problem that otherwise occurs. Furthermore, it is shown that the quality of the important sound field component extracted from the additional pressure sensor depends on information recovered from the TPG sensors. The discovery of this dependency necessitates the tracing of error propagation in the inverse solution to predict compromised captured information due to perturbations caused by measurement noise or spatial aliasing. A simulation study comparing the TPG design to a conventional pressure sensor design was conducted and discussed.

The manuscript concerns designs for the capture of two-dimensional sound fields, where the transducers are arranged in a circular layout on an infinite cylindrical baffle. An initial modelling of a spherical design based on TPG sensors that is fit for three-dimensional sound field capture is presented in Appendix A.

1.2.2 Journal Paper: DOA Estimation Performance with Circular Arrays in Sound Fields with Finite Rate of Innovation

Chapter 5 is the manuscript of a paper investigating the performance of a novel *Direction-of-Arrival* (DOA) estimation method for circular microphone arrays, which was initially proposed to the research community in a conference paper by the author, Filippo Fazi and Philip Nelson in [HFN16]. It is based on a work from 2002 by Vetterli, Marziliano and Blu in which they initially proposed a sampling strategy for signals in the time domain that assumes the signal has a *Finite Rate of Innovation* (FRI). This innovative way of approaching the sampling of continuous, non-band limited signals has since seen further refinement and variants for increased robustness against measurement noise were proposed.

One contribution of this Ph.D. project is finding that the sampling method for time domain signals by Vetterli et al. can also be applied to signals in the spatial domain. It was shown by the author et al. in [HFN16] that the function defining a two-dimensional sound field through the superposition of plane waves (i.e. the plane wave or Herglotz density) also has a finite rate of innovation if the sound field consists of a finite number of plane waves (see Figure 1.2). The DOA can then be estimated from an annihilating filter method, which was also presented in the work by Vetterli et al. Unlike some recently proposed DOA estimation techniques, the one studied in this work neither requires the assumption that only one source is present nor that the individual sources have at least one region in the spectrogram where they are isolated.

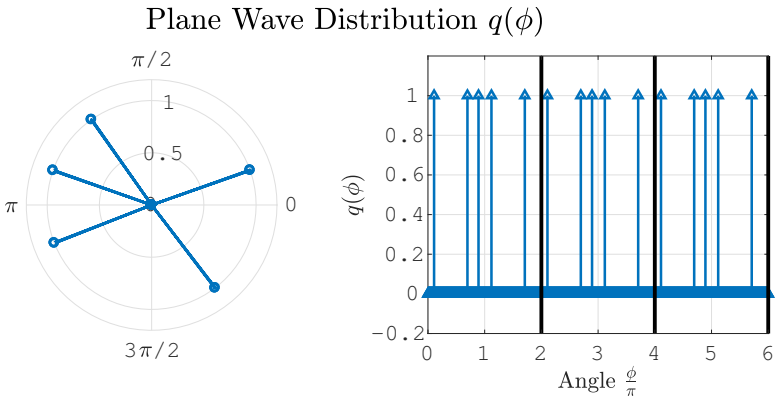


FIGURE 1.2: Example of a plane wave (or Herglotz) density $q(\phi)$ with a finite number of Dirac pulses, i.e. a finite rate of innovation, within the period of 2π. The periodicity is inherent to the circular domain.

The findings presented in Chapter 5 are the results of a performance study for the proposed DOA estimation method. Using well-established basis function expansions for the sound field of a finite number of plane waves, the accuracy and robustness of the proposed method against measurement noise and non-ideal acoustic sources is evaluated on the basis of both simulation studies and experimental data. The basis function expansion decomposes the plane wave density into circumferential modes, where each mode's complex amplitude depends on the number and strength of the plane waves in the sound field. It is shown that the modal coefficients for a single (ideal) plane wave can be interpreted as the discrete analytic signal $e^{in\phi_0}$ whose single frequency is determined by the DOA.

The modal coefficients of non-ideal sources are shown to also form a sequence of a monochromatic analytic signal but with modulated amplitude. An explanation for the mechanisms behind the perturbing effect of non-ideal acoustic sources can then be given in analogy to the masking effect of wide spectra of amplitude-modulated carrier frequencies on (weaker) neighbouring carrier frequencies in radio transmission systems.

Both the presented simulations and the experimental results are used to validate the method and provide an initial prospect on the performance that can be expected of it.

In the context of the other three manuscripts presented in this thesis, it shall be pointed out that the task of DOA estimation is conceptually an extension of the inverse problem of the microphone array that aims to recover the plane wave density. For this manuscript, it is assumed that the inverse problem leading to the recovery of the plane wave density can be solved without compromise, which is practically the case within the optimal frequency band of the microphone array used for the capture. From the perspective of the DOA estimation problem, the parameters that define the individual plane wave (i.e. the DOA and the complex amplitude) map to the plane wave density, which thus defines the forward problem. With the coefficients of the plane wave density available, the estimation of the DOA poses the corresponding inverse problem.

1.2.3 Journal Paper: A General Radiation Model for Sound Fields and Nearfield Acoustical Holography in Wedge Propagation Spaces

Chapter 6 is the manuscript of a paper that introduces a novel radiation model for the sound propagation and NAH in wedge propagation spaces. The main contributions of this chapter are the model for the radiation from a curved surface with a given velocity profile and the associated mechanisms that enable NAH in wedge propagation spaces (see Figure 1.3).

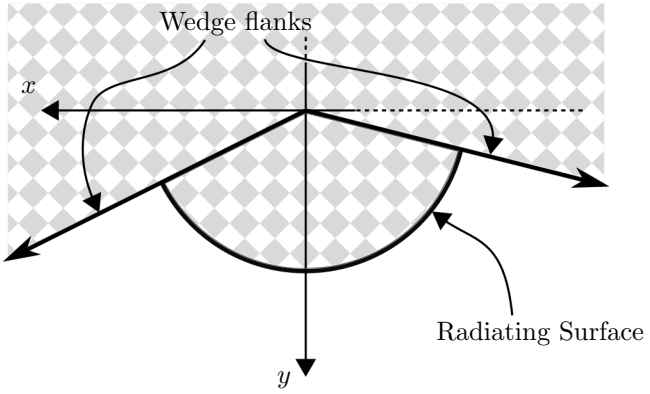


FIGURE 1.3: Birds-eye view into the infinite wedge propagation space with the curved radiating surface. A finite wedge propagation space has additionally a floor and ceiling at specific values along the z-axis.

Both finite and infinite wedges can be understood as segments of an infinite cylinder, making the geometrical description in cylindrical coordinates convenient. Solutions for sound fields in wedge spaces with acoustic radiation away from the origin (i.e. the vertical coordinate axis) are obtained from the general solution of the Helmholtz Equation in cylindrical coordinates by introducing a boundary condition at two different azimuth angles and, in the case of the finite wedge, at two different planes orthogonal to the vertical axis.

The radiation model proposed in Chapter 6 assumes a curved surface with constant radius around the origin and a given normal velocity, i.e. another boundary condition. The general solution for outgoing waves in the wedge space is therefore amended accordingly so that it always satisfies the boundary condition on the curved surface. The obtained solution is given in the form of an orthonormal basis function expansion that maps the spatial spectrum of the surface velocity to the pressure inside the wedge space. Using this model, a *Neumann-Green Function* (NGF) is derived for the infinite and the finite wedge, respectively. These NGFs enable the reformulation of the model for the radiation from the curved surface in a form akin to Rayleigh's first integral formula. As a result, two versions of the radiation model are presented, one that links a given normal velocity spectrum on the curved baffle to the pressure inside the wedge and one that links the normal velocity itself to the pressure inside the wedge.

The developed radiation models are used to simulate the sound field in infinite and finite wedges for point and piston sources arranged on the curved rigid surface. These results are later used for comparison with experimental data.

To enable NAH in wedge spaces, spatial transforms are formulated that link the hologram of the pressure at a given radius to the corresponding pressure spectrum in the modal (or wave number)

domain. Mathematically speaking, the spectra of these holograms can be easily propagated back or forth to different radii, so that the pressure can be calculated at the new radii. The combination of this spatial translation of the spectra and the exploitation of Euler's Equation then forms the basic mechanism enabling NAH in wedge propagation spaces.

Since NAH poses an inverse problem and is prone to ill-conditioning (see Chapter 2), the radial functions that map the normal spatial velocity spectrum of the curved boundary to the spatial pressure spectrum at an arbitrary radius are first identified. For the formulation of stable inverse propagators, these radial functions are analysed to identify those modal combinations that have an evanescent nature and fall into the region outside the "radiation circle". Tikhonov regularisation is used to stabilise the inverse propagator.

The developed tools that enable NAH are tested on the basis of experimental data obtained with a prototype of a hemi-cylindrical loudspeaker array. The modal resolution that can be expected from the measurements is predicted based on the angular and vertical sampling of the hologram area. Normal velocity reconstructions are performed on both simulated and measured data, to demonstrate the performance developed radiation model. Experimental data is used for both a single active loudspeaker and all loudspeakers active at the same time.

The motivation for the modelling of acoustic radiation in wedge propagation spaces lies in practical relevance. Due to the beneficial effects on the radiation impedance at low frequencies, loudspeakers are often installed in the corner of a room. The corner of a room with rigid floor, walls and ceiling is essentially a wedge and can thus be described with the developed model. Using a scanning aperture as the one presented in Chapter 6 for a loudspeaker installed in the corner of a room then enables a more accurate description for the radiation pattern of the loudspeaker at different frequencies.

1.2.4 Journal Paper: Beamforming with Wedge-Shaped Acoustic Arrays

Chapter 7 is the manuscript of a paper that uses the propagation model developed for the manuscript presented in Chapter 6 and applies it to develop beamforming techniques for sound field control with loudspeaker arrays in infinite wedge propagation spaces. To simplify the considerations, only arrays were considered that enable a control of the beam pattern in the azimuth direction, i.e. loudspeaker arrays that are arranged on an arc (see Figure 1.4).

The loudspeaker array is modelled as a finite number of loudspeakers that are distributed on a curved rigid baffle with constant radius around the vertical axis of the wedge. A general model for the normal velocity profile of the entire baffle (including the array) is formulated as a function of the driving functions for the individual loudspeakers. This model poses the forward problem, mapping the individual driving functions to the radiated pressure field inside the wedge.

Due to the arc-based layout, the array can only control modes along the azimuth direction to begin with. This means that any target beam pattern is described by a one-dimensional function, and thus by a set of modal coefficients. Finding the loudspeaker driving functions that yield a desired target radiation pattern requires solving the inverse problem. The expedient actions to obtain the solutions for the driving functions produce a set of intermediate results that determine

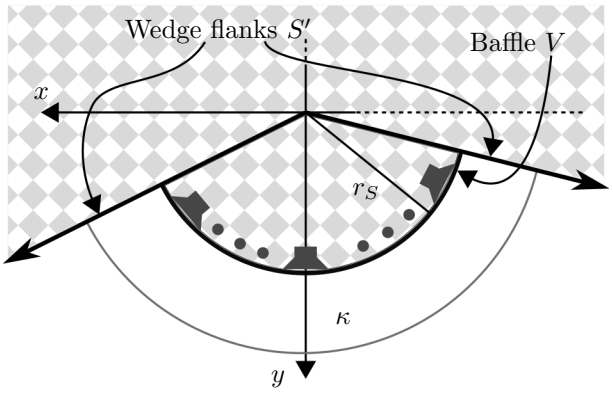


FIGURE 1.4: Cross section of the infinite wedge propagation space with the loudspeakers arranged on an arc with radius r_S.

the performance of the loudspeaker array. The relation between the target sound field modes and the actually radiated modes results from a non-exact orthogonality relation at high orders due to the finite number of loudspeakers. The resulting aliasing pattern describes the deviation of the radiated beam pattern from the target pattern at high frequencies. The radiation of the individual modes to the target distance is determined by the newly introduced *Arc Radial Functions*. These are defined to describe the modal propagation from the plane that contains the loudspeaker array to the arc formed by the intersection of the hologram at the target radius and a target plane that is orthogonal to the vertical axis of the coordinate system. Thorough evaluation of these arc radial functions reveals the existence of a near field, within which modes decay at a different rate than in the far field. This is confirmed by a comparison of the arc radial functions to their stationary phase approximation, i.e. the far-field arc radial functions.

The evanescent behaviour of arc radial functions is analysed for the different modes as a function of frequency. Strong decay of a mode effects ill-conditioning of the inverse solution for the loudspeaker driving functions, which may compromise the stability of the system. The effect of the regularisation for the calculation of the inverse on the directivity of the beam pattern is predicted from the analysis of the arc radial functions. Furthermore, it is discussed how the dependency of the arc radial functions on frequency justifies an order truncation for the modelling of the forward problem when calculating the inverse solution.

The beamforming performance of the described arrays was initially tested by means of simulations. To validate the model, experimental data was acquired by scanning the hologram of the radiated sound pressure from a hemi-cylindrical loudspeaker array. The measured beam patterns were reconstructed for evaluation using the NAH mechanisms presented in Chapter 6.

The experimental results were used to validate the presented model and to confirm the accuracy of the simulations.

1.2.5 Conclusions and Appendices

The final chapter provides more detailed concluding remarks on the results of the four principal chapters of this thesis and highlights the connections between the seemingly different topics. The

final section of this chapter contains a more general outlook on future research work and trends in the light of the findings in this work and that by other researchers in the field.

The appendices present different aspects and topics that were pursued over the course of this Ph.D. project, including

- a mathematical formulation of the forward problem describing a spherical microphone array composed of TPG sensors,
- a study on a signal processing strategy for circular microphone arrays that uses DOA information acquired in the optimal frequency band to upscale and de-alias data acquired in corrupted frequency bands,
- the design of a low-noise multichannel conditioning amplifier for microphone arrays,
- the design of a hemi-cylindrical compact loudspeaker array.
- the preprint of the publication: F.-M. Hoffmann, F.M. Fazi, P.A. Nelson. Plane Wave Identification with Circular Arrays by Means of a Finite Rate of Innovation Approach, In *140th Convention of the Audio Engineering Society, Paris 2016, May.*

Chapter 2

Literature Review

The spatial analysis and control of acoustic wave fields requires a multitude of sensors and actuators, respectively. The expedient evaluation of sensor data and control of acoustic sources, respectively, requires both accurate yet, from a computational point of view, efficient models of the acoustic wave fields in the underlying propagation spaces. Once a suitable model is found, the performance of a loudspeaker array is largely determined by the acoustic properties of the system itself. With microphone arrays, however, the situation is quite different. The final spatial resolution of a sound field analysis system is not only dependent on the sound field model, and the number and distribution of sensors, but also on the post processing of the observed data. The field of (acoustic) signal processing itself is vast (refer to [Ell01, Idi01, KB02, VT02, Hay02, HBC08, HL09, Can09, EK12, BIB13]) and only a small fraction is concerned with the *spatial* analysis of wave fields.

This chapter is divided into four subsections, where the first two lay crucial ground work that is equally important to sound field analysis as well as sound field control applications.

Section one introduces the models of acoustic wave fields most commonly used in combination with both microphone and loudspeaker arrays. The scope for this work is limited to acoustic modelling in the frequency domain, assuming a steady state of the system.

The second section is dedicated to inverse problems and their relevance in both sound field analysis and synthesis. Dominant factors adversely affecting inverse solutions are non-uniqueness, spatial aliasing, ill-conditioning, and transducer noise and misalignment.

The third section focuses on acoustical field analysis with microphone arrays and signal processing strategies, both developed to improve the performance and/or extend the functionality of microphone arrays. Particularly, parametric signal processing strategies for *Direction-of-Arrival* (DOA)-estimation are considered.

Section four refers to existing work on the design and performance of compact loudspeaker arrays. In this context, the descriptor ‘compact’ refers to arrays that form sound sources with controllable directivity, as opposed to large and spatially distributed loudspeaker arrays, which aim to control the sound field inside the volume they envelop; However, this is only of minor importance in this work.

2.1 Wave Field Models for Acoustic Transducer Arrays

It has already been stated in this chapter's introduction that the practical value of a transducer array is directly linked to the aptitude of the sound field model it is combined with. The models used frequently today are almost all derived from the wave field propagation model given through the acoustic *Homogeneous Wave Equation* (HWE) [FLS63, ABC$^+$70, Wil99, KFCS00, BX08]

$$\Delta p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{2.1}$$

where p denotes the pressure. Its derivation can be found in [FLS63, BX08]. The exact expression for the Laplacian operator Δ depends on the coordinate system chosen to describe the propagation space [AW05]. It is obvious from Equation (2.1) that its solution must be a function of both space and time. An early example of an analytical solution for the one dimensional wave equation was given by d'Alembert [lRD47, BX08] with the expression for the pressure

$$p(x,t) = p_+\left(t - \frac{x}{c}\right) + p_-\left(t + \frac{x}{c}\right). \tag{2.2}$$

The pressure is composed of two components travelling in opposite directions, where c denotes the speed of sound. Even though it implicates few major aspects of sound fields described by the HWE, d'Alembert's solution is hardly tangible and poses more of a solution in principle, since it lacks an explicit analytical form for p_+ and p_-. Typically, Equation (2.1) is solved using numerical methods, such as the *Finite Difference Method* (FDM) and the *Finite Element Method* (FEM) or more specialised methods that are less computationally demanding than the FDM or FEM [JKPS11].

It appears that the HWE in its given form needs to either be further simplified or combined with additional assumptions in order to find a more useful analytical solution. At this point the problem is typically translated into the frequency domain, where the wave field can be assumed in a steady state at every single frequency [Wil99]. Applying the Fourier transform (see Chapter 3, Equation (3.42)) to both sides of the equation results in the transformation pair

$$\Delta p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \Delta p = -k^2 p,\ k = \frac{\omega}{c}. \tag{2.3}$$

The equation on the right is widely known as the *Homogeneous Helmholtz Equation* (HHE) [ABC$^+$70, Wil99, AW05]. A collection of solutions of the HHE in various coordinate systems, and for different boundary conditions, can be found for electromagnetic and acoustic wave fields in [ABC$^+$70] and [FM03]. Williams thoroughly derived and introduced analytical solutions in Cartesian, cylindrical and spherical coordinates for acoustic applications on the basis of an expansion through orthonormal basis functions, which he obtained through a separation of variables approach [Wil99].

In particular, the solutions in cylindrical and spherical coordinates have found major interest in the field of audio related array research. The specific form of a solution differs for loudspeaker arrays and microphone arrays. The former can be based on a model that describes radiation from the device (in a given propagation space), where the latter often requires a combination of the solutions for the case of the undisturbed field and for the radiation case. This is further discussed in the following subsections.

2.1.1 Solution of the Homogeneous Helmholtz Equation in Cylindrical Coordinates

For many considerations in this work it is convenient to consider an undisturbed region Λ around the origin, with no sources or scattering objects inside an infinite cylinder of radius r_Λ (see Figure 2.1). The HHE shall be satisfied inside Λ and the pressure is finite at the origin. The

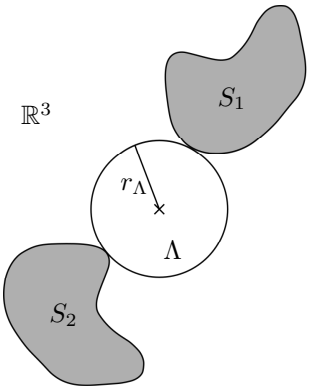


FIGURE 2.1: Sound field scenario where the HHE is valid inside the region Λ, which is centred around the coordinate system's origin, and with acoustic source/scatterer distributions S_1 and S_2.

general solution of the HHE for the pressure inside of Λ is given through [Wil99]

$$p(r,\phi,z,\omega) = \sum_{n=-\infty}^{\infty} e^{in\phi}\frac{1}{2\pi}\int_{-\infty}^{\infty} C_n(k_z,\omega)J_n(k_r r)e^{ik_z z}dk_z, \quad r < r_\Lambda. \tag{2.4}$$

Williams denotes Eq. (2.4) as the solution to the interior domain problem. The exterior problem refers to scenarios in which all sound sources are located *inside* a cylindrical region V with radius r_V around the origin and infinite length (see Figure 2.2), so that the solution [Wil99]

$$p(r,\phi,z,\omega) = \sum_{n=-\infty}^{\infty} e^{in\phi}\frac{1}{2\pi}\int_{-\infty}^{\infty} A_n(k_z,\omega)H_n^{(1)}(k_r r)e^{ik_z z}dk_z, \quad r > r_V, \tag{2.5}$$

satisfies the HHE in the remaining part of $\mathbb{R}^3$, namely Λ.

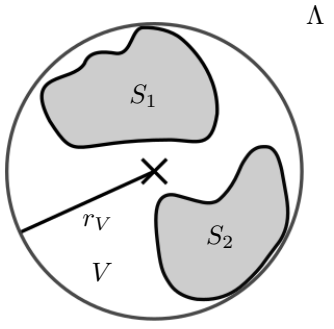


FIGURE 2.2: Sound field scenario where the HHE is valid outside the region V, which is centred around the coordinate system's origin and which contains the source/scatterer distributions S_1 and S_2.

The above solutions of the HHE have applications in the field of compact transducer arrays [TK06, PBA15], underwater acoustics and sonar technology [Wil99, KAH04, KA08]; but they are also used to simplify sound field models for analysis and synthesis applications by reducing the spatial dimensionality to two, describing a sound field that is invariant along the z-axis [Pol00, TK06, AS08, KRWB10, FNW12, AR14]. In that case, the integral in Equations (2.4) and (2.5), respectively, simplifies considerably. In fact, the two-dimensional model appears to be more frequently used than the three dimensional model, presumably because the rather complex integrand and the infinite sum in Eqs. (2.4) and (2.5) typically do not permit to write a closed-form expression for the right hand side[1]. Along with the practically inevitable discretisation of the integration domain along the k_z-axis and the resulting limitations, this peculiarity makes the cylindrical model and the associated arrays appear as a highly specified tool and therefore restricted in its applicability. Given that, it is not surprising that a large part of the research community focuses on the spherical solution instead, which is presented in the next subsection.

2.1.2 Solution to the Homogeneous Helmholtz Equation in Spherical Coordinates

Treating all spatial directions in a more homogeneous way, spherical coordinates often enable a decomposition of the three-dimensional sound field that is less likely to require simplifying assumptions or numerical approximations than that of cylindrical coordinates. The main difference compared to the solution in cylindrical coordinates is the replacement of plain Bessel-, Neumann- and Hankel-functions by their spherical variants [AW05, Wil99]. Just like the solution in cylindrical coordinates, this solution can be specified for sound fields that have no sources or scatterers inside a region Λ spreading around the origin with radius r_Λ, yielding [Wil99]

$$p(r, \phi, z, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_{mn}(\omega) j_n(kr) Y_n^m(\theta, \phi), \quad r < r_\Lambda, \tag{2.6}$$

with the spherical harmonics

$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}$$

as given by Williams in [Wil99], where $P_n^m(x)$ as the associated Legendre Functions.

This solution, or variants of it, are widely used in the field of acoustical array research and form the foundation of the theory on which the *Ambisonics* technique is based. Originally, Michael Gerzon formulated and introduced Ambisonics as one of the first three-dimensional recording techniques [Ger73, Ger75] and patented it together with Peter Craven [CG74]. Gerzon did not rigorously link his work on Ambisonics to the solution of the Helmholtz equation given in Eq. (2.6), although he made use of spherical harmonics when describing the gain structure he proposed to control the directivity of the microphone arrays associated with the technique. While the theory was not limited to first order sound fields, it was not until the mid 1990s and early 2000s that researchers refined both theory and microphone array designs to recover higher sound field orders using the *Higher Order Ambisonics* (HOA) technique in combination with the solution

[1] In his PhD-dissertation, Teutsch derived a solution for the scattering off a truncated cylinder, by neglecting the boundary condition on the cylinder's end caps [Teu05, Teu07]

of the Helmholtz equation in Eq. (2.6) [BV95, Mal99, ME02, DMN03, LDGD04, Raf05, Pol05a, MDB06, AS08, FNW12].

The performance of the HOA technique for sound field recording depends largely on the corresponding acquisition systems. The performance of spherical microphone arrays was theoretically analysed by Rafaely [Raf05] based on the solution given in Equation (2.6). He investigated the effects of sampling the aperture with a finite number of microphones, spatial aliasing as a result of spatial sampling, transducer noise and inaccurate transducer positions on the array's performance, as well as the properties of different spatial sampling schemes. In a more detailed analysis of spatial aliasing with spherical microphone arrays, Rafaely et al. [RWB07] studied the aliasing structure between the modes, the influence of the spatial sampling pattern on aliasing, and theoretically investigated the performance of anti-aliasing filters. In an earlier work [Raf04], Rafaely described the plane wave decomposition of a sound field based on the recorded pressure on a sphere, e.g. obtained from a spherical microphone array. In the context of that work, he also investigated the directivity of the decomposition method based on the order up to which the coefficients for the spherical expansion of the pressure can be recovered from the microphone array.

Ambisonics systems are not the only ones making use of spherical coordinates with the objective to record all three dimensions of a sound field; Meyer and Elko [ME08] proposed a circular array with an additional sensor at the origin to achieve modal beamforming with some degree of control in the elevation direction. Their design is based on a three dimensional sound field model using spherical coordinates.

The solution presented in Equation (2.6) not only serves as a model for sound field acquisition but also for sound field synthesis. Ahrens and Spors published a number of articles on the sound field reproduction using spherical (and circular) loudspeaker arrays based on the above equation [AS08, AS12b, AS12a]. They formulated analytical models for the sound field reproduction problem through the HOA approach and the *Wave Field Synthesis* (WFS) approach [Ber88], where both models draw on the sound field expansion given in Eq. (2.6) [AS08]. In [AS12b], they proposed a method to convert the coefficients obtained from an HOA recording into the driving functions of a WFS system. With that work, they enabled the use of a WFS system to reproduce a sound field that was recorded rather than virtually created. Addressing the issue of discretisation, which is a problem with sound field reproduction as much as with sound field acquisition, they investigated the influence of a discrete secondary source distribution as they are used in sound field reproduction systems on the synthesis performance. Their analysis draws on the modal structure of the sound field model, which enables them to describe the impact of the discretisation through the repetition and potential overlap of the spatial spectra in the modal domain. In that sense, their work corresponds directly to the work by Rafaely et al. on microphone arrays [RWB07]. The presented findings address analytical reproduction systems based on the methods of WFS, Near-field Compensated HOA and Spectral Division.

Both the cylindrical and the spherical solution make no assumptions as to what type of sources or elementary wave fields form the observed sound field. Consequently, they retain a large degree of generality. The two most elementary entities in acoustics are the field of a point source and that of a plane wave. The basis functions used in Eqs. (2.4), (2.5) and (2.6) are also found as

part of specific yet popular sound field models that are based on either a superposition of plane waves (see Section 2.1.4) or a distribution of point sources (see Section 2.1.5).

2.1.3 Solution of the Homogeneous Helmholtz Equation in Infinite Wedges

The diffraction of electromagnetic and acoustic waves from the surface of infinite wedges has long been modelled on the basis of the cylindrical coordinate system. Asvestas et al. [ABC$^+$70] presented a collection of analytical solutions for the HHE describing the acoustic wave field originating from point and line sources in the vicinity of infinite wedges with either acoustically soft or hard boundaries. By choosing the angle of the wedge arbitrarily larger than 180 degrees, the sources can also be modelled *inside* the infinite wedge (see Figure 2.3 with the wedge given in grey). Using the results presented in [ABC$^+$70], the solution of the HHE with acoustically

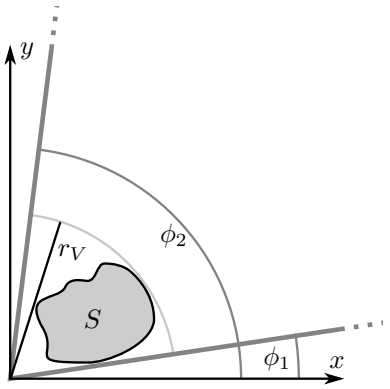


FIGURE 2.3: Infinite wedge scenario, with the source distribution S located inside the radius r_V, and boundary conditions at ϕ_1 and ϕ_2.

rigid boundary conditions at ϕ_1 and ϕ_2 is given by

$$p(r,\phi,z,\omega) = \sum_{n=0}^{\infty} \nu_n \cos\left(\frac{n\pi}{\kappa}\phi\right) \int_{-\infty}^{\infty} A_n(k_z) H^{(1)}_{\frac{n\pi}{\kappa}}(k_r r) \frac{e^{ik_z z}}{\sqrt{2\pi}} dk_z, \quad r > r_V. \tag{2.7}$$

with $\kappa = \phi_2 - \phi_1$, $\phi_1 < \phi < \phi_2$ and $\nu_n = \sqrt{\frac{2-\delta_n}{\kappa}}$, where δ_n denotes the Kronecker delta. The boundary conditions can be changed from rigid to acoustically soft by replacing the cosine with a sine in the above expression [ABC$^+$70].

The solution in Eq. (2.7) has applications in the field of shoreline acoustics. Buckingham [Buc84, Buc89, BT90], and Luo and Zhang [LZ15] previously used this solution to model the sound field of point sources located between the ocean surface and the sloping sea bed. The basis functions in Eq. (2.7) pertain to the case where the boundaries at ϕ_1 and ϕ_2 are rigid, i.e. perfectly reflective. Along the shoreline below the water surface, only the seabed may be considered (approximately) rigid, while the surface poses a pressure-release boundary. In the case of the latter, the basis functions in the ϕ-domain change from cosine to sine, but that as well would only be a valid model if both boundaries have pressure-release characteristics. To model the case of mixed boundaries, Buckingham proposed to use the basis functions $\sin\left((n+\frac{1}{2})\frac{\pi}{\kappa}\phi\right)$ which satisfy the rigid boundary condition on the seabed ($\phi = \kappa$) and the pressure-release boundary condition at $\phi = 0$ [Buc89].

The wedge solution is effectively a solution of the HHE in cylindrical coordinates that satisfies a set of boundary conditions. A conceptually similar solution was presented for spherical coordinates by Pomberger et al. [PZ13, PP14], where the spherically expanding propagation space is limited by adding a rigid boundary condition. The shape of the surface where the boundary condition applies is formed by two infinite cones aligned along the z-axis, where the top one is facing down and the bottom one facing up. Both cones intersect only at the tip point. The resulting propagation space is then expanded in spherical Slepian functions.

2.1.4 The Herglotz Wave Function

The sound field inside a region Λ (see Figure 2.1) can often be modelled with good approximation as the superposition of plane waves. The *Herglotz Wave Function* (HWF) [FNW12]

$$p(\mathbf{x},\omega) = \int_{\Omega} H(\mathbf{x},\hat{\mathbf{y}},\omega) q(\hat{\mathbf{y}},\omega) d\Omega(\hat{\mathbf{y}}), \quad \mathbf{x} \in \Lambda, \tag{2.8}$$

describes the composition of a sound field through the superposition of plane waves on the basis of two main components. The propagation model of a plane wave is given by the Herglotz kernel $H(\mathbf{x},\hat{\mathbf{y}},\omega)$. It models a free field propagation by definition [CK98], but it can be altered by replacing the kernel to model a combination of the incident field and the scattering off a (from the perspective of the coordinate system) primitive object, e.g. as given in [PR05]. The phase and amplitude of each plane wave is determined by the *Herglotz Density* (HD) $q(\hat{\mathbf{y}},\omega)$. The integration domain Ω denotes the unit sphere, i.e. it covers all possible directions-of-travel or directions-of-arrival, respectively. As such, the HWF is very general and surprisingly not as widely used as its simplified variant, which assumes only a single plane wave [Pol00, ME02, LDGD04, Raf04, PR05, Raf05, Pol05a, MDB06, TK06, LD07, ME08, AR12, AR14, AR16] while invoking the principle of superposition. Since the linearity of the model is given, this simplification is not a restriction to generality of the approach but makes analysis simpler and thus potentially clearer. Especially, works focussing on microphone arrays in beamforming applications prefer this approach of modelling only a single plane wave [ME08, LD07, AR12, AR14, AR16]. It enables a convenient theoretical analysis of practically relevant performance aspects, e.g. aliasing due to spatial discretisation and beam patterns.

However, the use of the HWF as a model must be considered carefully because it suffers from two major practical limitations [FNW12]:

- The HWF cannot model the sound field of a single point source with a bounded Herglotz Density $\|q\|_2 < \infty$, and
- the Herglotz Density may require up to an infinite number of plane waves to represent an arbitrary sound field, with the exception of sparse sound fields.

With the plane wave being the theoretical equivalent of a point source in the far field ($r \to \infty$), it is not surprising that a sound field can also be modelled by superposition of the sound fields coming from a number of point sources. This will be introduced in the next subsection.

2.1.5 The Single Layer Potential

The *Single Layer Potential* (SLP) is an integral operator that superposes the individual sound fields originating from a distribution of point sources on a given boundary $\partial\Lambda$ [CK98, FNW12].

$$p(\mathbf{x},\omega) = \int_{\partial\Lambda} G(\mathbf{x},\mathbf{y},\omega)\gamma(\mathbf{y},\omega)dS(\mathbf{y}), \quad \mathbf{x}\in\Lambda, \mathbf{y}\in\partial\Lambda \tag{2.9}$$

In this expression, *Green's function* $G(\mathbf{x},\mathbf{y},\omega)$ models the sound field of a single point source located at $\mathbf{y}$ [Wil99] and $\gamma(\mathbf{y},\omega)$ assigns the point source its strength. The integral over all point sources on the boundary $\partial\Lambda$ then determines the sound field.

The SLP is of particular practical interest since it overcomes the first limitation of the HWF mentioned in the previous subsection, and at the same time enables modelling of both the analysis problem *and* the synthesis problem [Pol05a, SRA08, Faz10, FNW12, FN12].

The sound field models introduced above have been the basis of many transducer array designs introduced over the recent decades, each of which connects an observable physical quantity, i.e. the sound pressure, to a model-specific quantity describing the entire sound field. Inferring the latter from the observed pressure is an *inverse problem* and will be described in more detail in Section 2.2.

2.2 Inverse Problems

The field of inverse problems is an entire research area by itself and only a small fraction has so far been used or shown relevant for the problems of sound field analysis and control. Tarantola [Tar04] presented a mathematically very general work on the theory behind inverse problems and approaches to identify stable solutions. That includes an analysis of the properties inherent to least-squares (l_2-norm minimisation of the error) and sparse (l_1-norm minimisation of the solution, while minimising the l_2-norm of the error) solutions, which are both particularly popular in the field of acoustics (see below).

The mathematical theory behind inverse problems and inverse operators in acoustic scattering considerations has been described in detail by Colton and Kress [CK98]. The theory of inverse operators and their limitations have been investigated w.r.t. the problem of sound field reproduction by Fazi [Faz10]. The focus of his work was to determine the validity of inverse operators, preservation of the uniqueness of the solution to the inverse problem and spatial aliasing, from the theory down to application level. Considering inverse problems in sound field analysis, it can be said that every Ambisonics, HOA or non-Ambisonics related sound field measurement system makes use of inverse operators or their discrete counterparts, respectively, in order to recover comprehensive information (e.g. the HD or the SLP) about the sound field from the array observation.

2.2.1 Non-Uniqueness of a Solution

Non-uniqueness of a solution to the inverse problem has been a factor with significant impact on the designs and the performance of both microphone arrays for sound field analysis and loudspeaker arrays for sound field reproduction applications [Pol00, Pol05a, MDB06, EF07, AS08, FN10, FN12, ZS13, BP14]. The causes of non-unique inverse solutions can be better understood by considering the forward problem in the form of a mathematical operator that maps a quantity describing the sound field (e.g. the HWF or SLP) to the observable pressure. An *Eigenvalue Decomposition* (EVD) (or *Singular Value Decomposition* (SVD) where necessary) of this forward operator yields a decomposition into a set of eigenfunctions (or left- and -right singular functions) and corresponding eigenvalues (or singular values) [Kre78, CK98, Faz10]. The eigenvalues (or singular values) describe how the forward operator maps the orthonormal eigenfunctions (or left- and -right singular functions) from one domain to the other. Should these eigenvalues (or singular values) become zero, it means that the corresponding components span the null space of the operator [Kre78, FN10], i.e. these components do not contribute to the final result. Whether or not an eigenvalue (or singular value) is zero depends highly on the evaluation point of the operator, e.g. the location of the sound field observation and the frequency. The inverse solution then suffers from non-uniqueness if the null space is non-trivial, at all observation points, as a result of the zero eigenvalue in the forward operator.

A non-unique inverse solution affects sound field analysis and sound field reproduction differently. For spherical and circular microphone arrays with an open sphere or open cylinder design, respectively, a non-unique inverse solution occurs when the boundaries containing microphones fall into zero-crossings of the radial functions in the field expansion [Pol00, AW02, Pol05a, Raf05, MDB06, EF07, FN10, ZS13]. As a consequence, the component associated with the radial function that has a zero at the boundary of the microphones may take any value and contribute to the overall sound field, but it would not be observed by the microphone array. An identification of the strength of that component is then evidently impossible. The inverse solution can take any value for that component, and the forward operator would still yield the same observation at the points of the microphones. Several ways to avoid non-uniqueness issues with microphone arrays have been proposed. Designs that avoid non-uniqueness issues have either rigid corpora, use directional transducers or several staggered measurement boundaries [AW02, ME02, PR05, Raf05, Pol05a, MDB06].

Non-uniqueness also affects the sound field reproduction performance of loudspeaker arrays, yet in a different manner. It was found that, for this application, the uniqueness of an inverse solution depends on certain boundary conditions that are assigned during the modelling of the reproduction problem [AS08, FN10, ZS13]. If the sound field reproduction system can be constructed entirely in accordance with the *Kirchhoff Integral Equation* (KIE) then the system does not suffer from non-uniqueness. However, in practice the quality of the required dipole speakers does not match that of state-of-the-art 0th order loudspeakers [ZS13]. Therefore, the sound field reproduction system is commonly described on the basis of the SLP (see Section 2.1.5), which is sometimes also referred to as the simple source formulation [Faz10, ZS13]. The non-uniqueness arises from the definition of the control volume *inside* the reproduction system. Controlling the sound field in a defined control volume corresponds to finding a solution to the HHE with a Dirichlet boundary condition [Wil99, FN10]. The inverse problem becomes non-unique at the control volume's 'cavity resonances' or forbidden frequencies [Wil99, AS08].

An interesting overview on methods to overcome with this problem was presented by Zotter and Spors [ZS13].

2.2.2 Ill-Conditioning, Transducer Noise, Misalignment and Variation of Environmental Parameters

Ill-conditioning is one aspect of inverse problems where the effects and their severity depend highly on the measurement error in the acquired data, and the dynamic range of the non-zero eigen/singular values of the forward problem, i.e. the condition number of the system [Han98]. If the magnitude of some eigen/singular values take very small values, they may project some components to pressure levels below the noise floor of the measurement system. That means the information on the sound field corresponding to that component is effectively lost. Accordingly, a recovery of that component from the inverse solution can only lead to a random amplified contribution that stems from the measurement noise. As a consequence, the norm of the inverse solution is considerably larger than expected [Han98], i.e. the inverse solution is unstable [Faz10]. Whether or not an ill-conditioned system suffers from the effects of ill-conditioning is mathematically determined through the abidance or violation of Picard's two conditions [Han98, Faz10]. Based on the input data of the inverse problem, Picard's second condition determines whether the effects of ill-conditioning can be mitigated or not.

Rafaely [Raf05] and Poletti [Pol05b] both identified noise from the microphone capsules as significant sources of error in spherical and circular microphone array designs. Poletti has shown that transducer noise becomes a more severe problem at low frequencies, as a consequence of the ill-conditioning of the inverse problem. Therefore, even with state-of-the-art microphone capsules, transducer noise is still a limiting factor in the field of three (and two) dimensional sound field recordings.

Transducer noise observed at the electrical terminals of a microphone is generated by thermal agitation of molecules in front of the diaphragm (Brownian motion) [SF74] and electronic noise [HBR95, Oak97]. The electronic noise itself is composed of:

- Brownian motion of charges in the conductors [Nyq28],
- shot noise (in vacuum tubes and semiconductors) [Sch18, Zie53],
- $1/f$ noise (from the internal impedance converters) [Off70] and
- flicker noise (as it occurs in MOSFETs, which are typically used as internal impedance converters) [Sch26].

Further sources of error that shall be mentioned are

- the inaccuracy of sensor positioning in the array, which can result from a limited production tolerance or inadequate assembly,
- the introduction of phase errors due to the temperature dependency of the speed of sound, and

- boundary conditions that are not as ideal in practice as in their model.

Rafaely [Raf05] found that the error induced through inaccurate positioning of the sensors can be kept at a relatively low level (-40 dB) with practical limitations on the manufacturing tolerance.

A mismatch between the real speed of sound and the one used in the model can however have a significant impact on the performance of microphone arrays. The expressions given in Eqs. (2.4), (2.5), (2.6) and (2.7) all depend on the acoustic wave number, $k = \frac{\omega}{c}$, and thus also on the speed of sound. The use of k as an argument to different basis functions can lead to a phase mismatch between the model and reality, of the real wave numbers deviate from the modelled ones due to a temperature difference. Particularly with transducer arrays, where the phase becomes critical for accurate exploitation of acoustic interference, such a mismatch can have a significant effect on the system performance.

Elko et al. [EDG03] investigated the impact of room impulse variation due to temperature changes on the performance of an echo-cancellation system. They found that even a deviation of 1 °C from the temperature associated with the assumed speed of sound can significantly change the room impulse response and thus affect the echo-cancellation performance. Chen et al. [CSY07] proposed a wideband beamformer for the near field of a microphone array that is robust against the phase changes introduced by the mismatched between the assumed speed of sound and the actual one. Annibale et al. [AFNR13] proposed a method to identify the speed of sound through the analysis of the time-differences-of-arrival in rooms directly from the microphone data. If accurately identified, the obtained information can be used to update the array processing accordingly, thus maintaining a robust system performance. A more general work on the influence of ambient atmospheric conditions on acoustic measurements was presented by Chapman [Cha14].

2.2.3 Spatial Aliasing

With microphone arrays, spatial aliasing is a result of the (so far) inevitable sampling of the sound field below the Nyquist rate. It was shown in the past that the specific aliasing pattern of an array can be anticipated and derived from the array design [LDGD04, LD07, RWB07, Faz10, KRWB10, AR12, AR14, AR16]. Determinant factors are the chosen coordinate system of the model, the chosen orthonormal basis functions, the sampling scheme and the array corpus. Dmochowski et al. [DBA09] presented a broadband beam pattern analysis to evaluate the spatial aliasing with uniform linear arrays. A thorough analysis of aliasing phenomena in spherical arrays was presented by Rafaely [Raf05, RWB07], Li and Duraiswami [LD07] and Fazi [Faz10]. Rafaely also suggested to implement a spatial version of an anti-aliasing filter to suppress higher order components, e.g. by integrating over consecutive measurements of just one microphone in stationary sound fields or by using pressure sensors with relatively large diaphragms, which is similar to spatially integrating over the local sound field. But despite extensive ongoing research in that field, aliasing remains one of the most prominent problems in microphone array technology. Nevertheless, further potential ways to deal with aliasing arose quite recently. Alon and Rafaely suggested de-aliasing strategies for single plane waves observed by circular and spherical microphone arrays [AR12, AR14, AR16]. The proposed method is based on the a priori knowledge of the aliasing pattern. Other techniques from the field of signal processing have been

applied to approach the problem of spatial aliasing in a similar manner [EJvS09] (see also Section 2.3.2.2).

With loudspeaker arrays, spatial aliasing is a result of the finite number of sources used to control the sound field. Effectively an ambiguity problem in the modal domain, spatial aliasing always occurs, but with high order modes being more prone to evanescent behaviour, the effects only become prominent beyond the spatial aliasing frequency of the system [AWS10, PB13], which is determined by the spacing of the loudspeakers in the array [PB13]. Spors and Rabenstein presented a study on the spatial aliasing with linear and circular arrays that are used with wave field synthesis systems [SR06]. Afterwards, Ahrens and Spors presented a more in-depth analysis of the spatial aliasing with sound field reproduction systems within the reproduction area [AS08]. Epain et al. [EJvS09] proposed to use compressed sensing methods (Section 2.3.2.2) to reduce reproduction artefacts that occur due to spatial aliasing. Ahrens et al. [AWS10] presented an analysis of spatial aliasing not from a monochromatic point of view, but in the temporal domain instead. They found that the discretisation through the finite number of loudspeakers leads to a series of 'echoes' from different directions following the reconstructed wave form in the time domain. The spectral composition of these echoes was restricted to frequencies above the loudspeaker system's Nyquist frequency. Poletti and Betlehem [PB13] presented a work on compact cylindrical loudspeaker arrays as sources with directivity control, and derived a formula for the system's Nyquist frequency for the circumferential modes. The result is a function of the radius, the modes' order and the number of loudspeakers. Beyond its (spatial) Nyquist frequency, a circumferential mode can no longer be controlled and the system produces aliasing artefacts.

2.2.4 Nearfield Acoustical Holography

A prominent specific example of an inverse problem in acoustics is *Nearfield Acoustical Holography* (NAH). Maynard et al. proposed the technique of NAH in their original work from 1985 [MWL85]. In an earlier work, Weinreich and Arnold proposed a method the identify the radiation behaviour of an acoustic source (e.g. a violin) in a non-anechoic environment through measurement of the pressure on two concentric spheres of different radius [WA80]. Their work can be understood as a precursor of the NAH work presented by Maynard et al. The mathematical model used by Weinreich and Arnold was based an expansion similar to that in Eq. (2.6), only did their method require two sets of coefficients to describe the weights of two sets of radial basis functions, namely the Hankel functions of the first and the second kind. These are needed to describe the incoming and outgoing waves as they would occur in the sound field of a non-anechoic room around an acoustic source.

Williams presented a comprehensive work on NAH in Cartesian, cylindrical and spherical coordinates [Wil99]. Focussing on the ill-conditioned inverse problems that occur in the field of NAH, Williams presented a work on regularisation methods for stable solutions [Wil01]. For situations where there is only a small patch of measurement data available, Williams et al. proposed a method to reconstruct and thus extend the measurement aperture on the basis of an iterative *Fast Fourier Transform* (FFT)-based algorithm that projects the measured data back and forth, while updating the initially unknown section of the now larger patch [WHH03]. Jacobsen and Jaud [JJ07], and Zhang et al. [ZJBC09] proposed variants of NAH based on the measurement of pressure and velocity. Jacobsen and Jaud performed the NAH without a

transform of the data taken in the spatial domain, whereas Zhang et al. based their variant on the equivalent source method [KSF89]. Both methods enable (within limits) a distinction of waves that originate from both directions of the plane of the measurement aperture, i.e. when the free field assumption is not valid on the side of the measurement plane opposite to the source. Klippel and Bellmann proposed to use NAH to analyse the near-field and far-field radiation behaviour of loudspeakers [KB16]. During his 2017 workshop held at Dresden University from 13th - 15th March, Klippel extended their original concept by measuring two holograms at different distances from the loudspeaker inside a reflective room [Kli17]. This data is then combined with the model for the exterior and interior domain in spherical coordinates (compare Sections 2.1.1 and 2.1.2). The inverse problem is then solved through matrix inversion. The extension of the method enables NAH from data where sound waves are travelling both into and out of the hologram volume.

2.3 Microphone Arrays & Signal Processing Strategies

Most of the aspects concerning the limitations of microphone arrays were already discussed in the previous section, therefore this section focuses on microphone array designs and signal processing strategies.

Microphone arrays for sound field acquisition applications are typically (but not exclusively) found in circular/cylindrical [Pol00, Mey01, KAH04, Pol05b, TK06, ME08, KA08] or spherical sensor arrangements [AW02, ME02, LDGD04, Pol05a, Raf05, PR05, MDB06, RWB07, LD07, ZDG10]. Pomberger et al. proposed a spherical segment array [PZ13, PP14], where the aperture was truncated along the elevation direction by the introduction of two rigid cones.

Note that linear microphone arrays are not considered in this work due to their inherent ambiguity between the fields on either side of the array, i.e. their axially symmetric directivity pattern [VT02].

2.3.1 Types of Transducers

While most of the designs today apply pressure sensors, the earliest design for an array in the field of airborne acoustics, presented by Alan D. Blumlein [Blu31, Blu35], is ideally (and usually) implemented in the form of two coincident pressure gradient (figure-of-eight) microphone capsules, also referred to as a "Blumlein-pair". This recording technology is still used by recording engineers, yet it enables only a two-dimensional analysis of a sound field.

Poletti [Pol00, Pol05a], and Rafaely [Raf05] both found that the use of directional sensors in microphone arrays can overcome non-uniqueness problems (see 2.2). Daniel et al. [DMN03] and also Poletti [Pol05a] found that such a design can reduce the effects noise at low frequencies.

Craven, Law and Travis [CLT09] proposed a spherical array based on tangential velocity sensors.

2.3.2 Signal Processing Strategies

The limitations affecting the performance of microphone arrays can sometimes be mitigated or avoided by applying additional signal processing. This subsection gives an overview on the signal processing strategies proposed in the past for applications relevant to this work.

2.3.2.1 Direction-of-Arrival Estimation

The sampling of a sound field at different spatial locations, as it is done with microphone arrays, permits the estimation of the DOA for an incoming wave front. Capon presented a DOA-estimation method that is purely based on the exhaustive search of the maximum of the array's angle-dependent output energy, which is calculated on the basis of the beamforming vector and correlation matrix of the input signals [Cap69]. Capon's method does not employ a parametric model of the wave field.

To improve the resolution of the DOA estimation, Schmidt proposed the MUSIC algorithm, which performs an analysis of the space spanned by the input signal data and uses a subspace approach to determine the DOA [Sch86]. MUSIC is a parametric method, i.e. it requires a model of the wave field and exact knowledge of the array aperture. To increase robustness and decrease computational intensity, Roy and Kailath proposed the ESPRIT method as an alternative to MUSIC [RK89]. Unlike MUSIC, ESPRIT constrains the structure of the sensor array by imposing what the authors call *translational invariance*. Both, MUSIC and ESPRIT are subspace-based algorithms that make a narrowband assumption for the sources in the sound field. Teutsch and Kellermann proposed a variant of the ESPRIT algorithm, EB-ESPRIT (EB - Eigenbeam), which does not apply ESPRIT to the sensor signals but rather to the modal components obtained from the analysis of circular array data [TK06].

Högbom proposed the CLEAN method, which iteratively subtracts the aperture signature of an elementary source located at the angle corresponding to the maximum of the aperture response [H$\ddot{7}$4]. It can be seen as a variant of Mallat and Zhang's Matching Pursuit algorithm [MZ93], with the difference that it does not choose from a predefined dictionary of elementary responses. Instead, CLEAN identifies the best fitting aperture signature of an elementary source by a predetermined criterion (e.g. maximum of the aperture response or the maximum of the correlation between the aperture response and the aperture signature of an elementary source).

Karbasi and Sugiyama proposed a DOA estimation method for circular arrays based on the *Circular Integrated Cross Spectrum* (CICS) [KS07]. Their algorithm enables the identification of a single source in slightly reverberant environments, with low computational complexity and robustness against noise. The use of a circular array helps to resolve issues related to localisation ambiguities, as they occur with linear arrays. Pavlidi et al. proposed a modified version of Karbasi and Sugiyama's algorithm that detects several sources [PPGM12]. Their method requires that the microphone signals contain 'constant-time analysis zones' for each source in the field, i.e. a tempo-spectral region in the signals where only one source is active. Stefanakis and Mouchtaris presented a variant of the algorithm by Pavlidi et al. for a circular microphone array in front of a reflective plane [SM16].

A method for high resolution plane wave identification called HARPEX was proposed by Berge and Barrett [BB10b, BB10c]. It uses a parametric decomposition of the B-format signal that yields two signals, referring to the two dominant directions from which sound waves are incident.

2.3.2.2 Compressed Sensing

It was in the early 2000s that researchers in the field of sampling-based data acquisition started pushing the boundaries of the well-known Nyquist-Shannon sampling theorem. One of the methods to reconstruct a signal which was sampled below the Nyquist frequency is based on the *Compressed Sensing* (CS) theory (sometimes also referred to as compressive sampling) [CW08]. A very thorough insight into CS from an information theory point of view was given by Donoho in [Don06]. In order for CS methods to be applicable, two major assumptions on the signal and its representation domain need to be satisfied:

- the signal itself needs to be *sparse* and
- the elements in the signal domain should ideally be incoherent to the elements in the representation domain.

In other words, sparsity is given for signals that can be represented by a smallest possible finite set of basis components, and the incoherence requirement states that the correlation between the signal space and the representation space should not exceed a certain threshold. The group around Epain and Jin applied CS successfully with various types of arrays to improve the sound field acquisition performance of the acoustical system [EJvS09, WEMJ11, WEJ12]. They also showed that CS theory can be used to upscale data for use with an HOA reproduction system with order higher than that of the recorded HOA data. Through this, they achieved a wider 'sweet-spot' and more distinct acoustic source positioning.

2.3.2.3 Finite Rate of Innovation Theory

Although it is theoretically more related to a sampling theory than to a signal processing approach, the *Finite Rate of Innovation* (FRI) theory provides an interesting new approach to sampling, signal processing and wave field modelling for sound field acquisition purposes. The main principle behind FRI theory is based on a periodic signal assumption (which is valid for the observation of a circular array and presumably also for spherical arrays) and a finite number of parameters (i.e. 'innovations') defining the continuous signal within the period. The sampling interval that still allows for perfect reconstruction is now determined by the rate of innovations per period. A sound theoretical description of the theory and associated sampling method was given by Vetterli [VMB02]. Dragotti and Homann analysed and discussed the performance in noise in a related work [DH09]. An interesting overview of the connection between sparseness and FRI is presented in [BDV$^+$08]. The FRI based sampling method is particularly interesting for sound fields composed of a finite number of plane waves, as will be shown later in this work.

A review of the FRI theory and the associated annihilating filter method for signal reconstruction is given in Chapter 5.

2.4 Compact Loudspeaker Arrays

Generally speaking, loudspeaker arrays are sound reproduction systems that are composed of more than one independently controllable acoustic source, i.e. loudspeaker. That definition encompasses stereo systems and other multichannel layouts, e.g. commercially available surround sound systems (in 5.1, 7.1, and similar configurations). While they match the definition, these types of arrays are mainly expansions of the stereo concept and can only marginally draw upon the potential of more elaborate array designs. Widely known examples of such are large linear arrays used in professional audio to optimise the systems directivity for both improved efficiency and quality of public address systems. Exploiting the nature of sound as a wave phenomenon, these systems control the individual loudspeakers in such a way that the radiated wave fields interfere with each other to create/approximate a desired directivity pattern of the combined system.

Over the past 30 years, the scientific community proposed and investigated large loudspeaker array systems of linear, spherical and circular geometry with the goal to control entire sound fields over a preferably large reproduction area. This research effort lead to the extensively studied WFS and HOA systems. Both aim to synthesise or reproduce a sound field within a bounded control area using a finite number of loudspeakers. WFS systems were originally limited to synthesising virtual sound scenes from a model, while HOA systems are capable of both, reproducing a previously recorded scene or synthesising a virtual sound scene. Ahrens and Spors proposed a method to overcome this limitation of WFS systems, by relating the WFS driving functions to the sound field coefficients obtained from an HOA recording [AS12b]. Notable and important work on WFS systems and HOA and sound field reproduction can be found in [Ber88, BVV93, WA01, Dan03, DMN03, SR06, SRA08, AS08, EJvS09, AWS10, Faz10, Ahr12, AS12b, AS12a].

The general downside of arrays that make for well-performing HOA or WFS systems is the vast number of channels and space they require. While there is potential in the market for a dedicated commercial installation, built to address large audiences, the domestic application of WFS systems currently faces negligible interest. This lead researchers to shift their focus to more compact arrays, small enough to be suitable for domestic or even hand-held applications.

This section gives an overview on published work proposing compact array designs based on spherical and cylindrical geometries and their applications for sound field control and synthesis.

2.4.1 Designs & Performance

Unlike large loudspeaker arrays that aim to reproduce a sound field in the reproduction area they envelop, compact loudspeaker arrays aim to provide an acoustic source with controllable directivity. Early designs based on cylindrical coordinates were proposed by Bédard and Berry [BB05], by Møller [MOA$^+$10], by Kolundzija et al. [KFV10, KFV11], and by Poletti and Betlehem et al. [PB13]. Zotter et al. proposed a directional acoustic source, based on spherical coordinates and a spherical harmonics-based beamforming [ZSH07]. Fazi et al. compared the pressure matching and mode matching performance of compact circular arrays [FSO$^+$14] and investigated their low-frequency beamforming performance [FSOF15].

2.4.2 Applications in Sound Field Control

While full sound field reproduction techniques aim to reproduce a sound field accurately over a predefined reproduction area, sound field control simplifies the problem by limiting the sound field control objective to a few (spatially separate) zones. Poletti [Pol08] proposed this for circular loudspeaker arrays to create a multizone two-dimensional surround sound system through a least-squares pressure matching approach [Pol07]. Betlehem and Poletti et al. then proposed to use acoustic sources with controllable directivity (see Section 2.4.1) as the elements in a circular loudspeaker array for an advanced sound field control/reproduction system that accounts for the reflections of the room [PAS12, BP14, PBA15].

More suitable for the application of compact loudspeaker arrays are scenarios in which different zones around the system shall be created, or, more plainly speaking, the sound field surrounding the compact array shall be controlled. Rafaely investigated the option of using a spherical loudspeaker array with controllable directivity for advanced active control of the local sound field by means of simulation [Raf09]. He found that a directional source yields a larger quiet zone than a standard monopole source. Olivieri et al. investigated the sound zone control performance of compact loudspeaker arrays driven through a pressure matching approach on the basis of both measured and analytical transfer functions [OSF+13]. To counter the ill-conditioning affecting the low-frequency performance of compact loudspeaker arrays, they proposed a control point selection strategy that is frequency-dependent [OFSN15]. In a more recent work, Olivieri et al. refined the pressure matching method by adding mechanisms for the creation and flexible control of private and silent sound zones [OFNF16].

A good overview on the recent advances in the general area of sound field control was provided in various works. Olik et al. presented a study on the comparison of the performance of different sound zoning methods (e.g. pressure matching [Pol07], acoustic contrast control [CK02] and delay-and-sum beamforming [VB88]) in reflective environments [OFC+13]. Betlehem et al. presented a broad overview on the problem of sound zone control, covering different control methods, loudspeaker arrangements, and arrays of sources with controllable directivity [BZPA15].

Chapter 3

Mathematical Notation, Definitions and Commonly Used Functions

This chapter serves the purpose to introduce this work's general mathematical notation and applied concepts. Many of those are already well known in the field of acoustics and transducer arrays, but they are listed here to simplify the access to the presented work.

3.1 Vectors, Matrices and Norms

Both, vectors and matrices are written in bold font, where non-capitals are assigned to vectors (e.g. $\mathbf{x}$) and capitals are reserved for matrices (e.g. $\mathbf{H}$). Unless stated otherwise, vectors are always arranged in columns. A vector of dimension $m \times 1$ with the identifier $\mathbf{x}$ describes a point in the space $\mathbb{R}^m$ where

$$\mathbf{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \tag{3.1}$$

with $x_1, \ldots, x_m \in \mathbb{R}$. The $m \times 1$ vector $\hat{\mathbf{y}}$ depicts a vector that also lies in $\mathbb{R}^m$ but has unitary length. These are typically used to indicate directions. A superscript letter $(\cdot)^T$ denotes the transposition of a vector and a superscript $(\cdot)^H$ denotes the hermitian transposition, which is the transposition and element-wise complex conjugation. The aforementioned two identifiers are reserved for vectors in $\mathbb{R}^m$ where $m \in \{2, 3\}$. Vectors with other identifiers will be defined and specified as they are introduced.

A number of norm definitions are useful in the context of this work. The most commonly known norm is the L_2-norm, which is sometimes also referred to as the Euclidean norm of a vector

$\mathbf{x} \in \mathbb{R}^m$. It is usually used to calculate the length of a vector and is defined as follows

$$\|\mathbf{x}\|_2 = |\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{n=1}^{m} x_n^2}. \tag{3.2}$$

The scalar product of two vectors in $\mathbb{R}^m$ as it has been used in the equation above is defined as

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \ldots + a_m b_m. \tag{3.3}$$

The vector product of two vectors in $\mathbb{R}^3$ is defined as

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - b_3 a_2 \\ a_3 b_1 - b_1 a_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \tag{3.4}$$

In relation to compressed sensing theory, two further norm definitions are of importance to assess the sparseness of a solution. The l_0-'norm' (the term is commonly used in the field of engineering) can be thought of as the l_1-norm of the Hamming distance [BSMM06] of the vector $\mathbf{x}$ to a zero-vector of the same dimensions. Simply speaking, the l_0-'norm' is the number of non-zero elements in a vector $\mathbf{x}$. The l_0-'norm' as such is, technically speaking, not a norm. Firstly, it does not satisfy the homogeneity relation [BSMM06]

$$\|a\mathbf{b}\|_p = |a| \cdot \|\mathbf{b}\|_p. \tag{3.5}$$

Secondly, it can neither be expressed through the general definition of the p-norm in *Banach* Spaces with finite dimensions [Don06, BSMM06]

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{m} |x_n|^p \right)^{\frac{1}{p}}, 1 \leq p < \infty \tag{3.6}$$

nor through the general definition for the integral norm of a function f in L^p (also *Lebesgue*) spaces [BSMM06]

$$\|f(\mu)\|_p = \left(\int_D |f(\mu)|^p d\mu \right)^{\frac{1}{p}}, 1 \leq p < \infty. \tag{3.7}$$

The l_1-norm is covered by the definitions above and can be simplified to

$$\|\mathbf{x}\|_1 = \sum_{n=1}^{m} |x_n|. \tag{3.8}$$

Both, the l_1-norm and the l_0-'norm' are used as a measure to optimise sparse solutions to a problem in the field of signal processing, although it should be pointed out that in practice, the l_1-norm is used almost exclusively.

3.2 Complex Numbers

With complex numbers being a substantial tool in nearly all considerations hereunder, a few notational aspects will be provided in the following. The symbol $i = \sqrt{-1}$ denotes the imaginary unit. A complex number

$$c = a + ib = |c|e^{i\phi} \in \mathbb{C} \tag{3.9}$$

has a real part $a \in \mathbb{R}$, an imaginary part $b \in \mathbb{R}$ or magnitude $|c|$ and a phase angle ϕ. Its complex conjugate is given as $c^* = a - ib = |c|e^{-i\phi}$, such that $cc^* = |c|^2$.

3.3 Sets

3.3.1 Defining Sets

Even though not of major interest and not within the focus of this work, sets are however needed to clearly define the different domains involved in modelling and designing microphone arrays. A set D is either defined directly

$$D := \{x_1, x_2, \ldots\} \tag{3.10}$$

or by providing a criterion for each element in D to satisfy, e.g.

$$D := \{x \in \mathbb{N} : x < 6\} = \{0, 1, 2, 3, 4, 5\}. \tag{3.11}$$

3.3.2 A Cylinder of Infinite Height

Equation (3.11) shows that sets can sometimes also be defined in dependence of other sets. An important *closed* set (i.e. it contains also its boundary) used in this work is that of all points $\mathbf{x}$ within a cylindrical volume of radius r_V around the origin of radius r_V and of infinite height

$$V := \{\mathbf{x} \in \mathbb{R}^3 : \sqrt{x_1^2 + x_2^2} \leq r_V\}. \tag{3.12}$$

3.4 Function Spaces

3.4.1 L^2 Spaces

Another set whose definition depends on V is that of all square-integrable functions $f(\mathbf{x})$ on the integration domain V

$$L^2(V) := \{f(\mathbf{x}) : \|f(\mathbf{x})\|_2 = \sqrt{\int_V |f(\mathbf{x})|^2 dS(\mathbf{x})} < \infty\}. \tag{3.13}$$

In other words, $L^2(V)$ is the Banach or Lebesgue space, where all elements have a finite integral norm as given in (3.7) for $p = 2$. As a Banach space, $L^2(V)$ is also *complete* [BSMM06].

3.4.2 Scalar Product

The completeness property of Banach spaces is a very powerful one and all it takes to show that $L^2(V)$ is also a Hilbert space is to find a scalar product (or inner product) that induces the norm used to define $L^2(V)$. The definition of a suitable scalar product is given by

$$\langle f|g\rangle_V = \int_V f(\mathbf{x})g(\mathbf{x})^* dS(\mathbf{x}) \tag{3.14}$$

and it satisfies the following axioms for every element in $L^2(V)$:

$$\langle f|f\rangle_V \geq 0, \text{ with } \langle f|f\rangle_V = 0 \text{ when } f = 0, \tag{3.15}$$

$$\langle \alpha f|g\rangle_V = \alpha \langle f|g\rangle_V, \alpha \in \mathbb{C}, \tag{3.16}$$

$$\langle f + h|g\rangle_V = \langle f|g\rangle_V + \langle h|g\rangle_V, \tag{3.17}$$

$$\langle f|g\rangle_V = (\langle g|f\rangle_V)^*. \tag{3.18}$$

It is now clear that the integral norm and the scalar product are related as follows.

$$\sqrt{\langle f|f\rangle_V} = \sqrt{\int_V f(\mathbf{x})f(\mathbf{x})^* dS(\mathbf{x})} = \sqrt{\int_V |f(\mathbf{x})|^2 dS(\mathbf{x})} \tag{3.19}$$

Therefore, with this scalar product inducing its norm, $L^2(V)$ is a Hilbert space. Further information on Hilbert spaces can be found in [Kre78, CK98, BSMM06] and other textbooks on functional analysis.

3.4.3 Total Orthonormal Sets

The introduction of a scalar product also provides the concept of orthogonality. Two elements $f, g \in L^2(V)$ are orthogonal if

$$\langle f|g\rangle_V = 0. \tag{3.20}$$

Just like for the vector space $\mathbb{R}^3$, every Hilbert space $L^2(V)$ different from $\{0\}$ has at least one *Total Orthonormal Set* (TOS) [Kre78]. The latter refers to a subset $M \subset L^2(V)$ whose *span* is dense in $L^2(V)$, i.e.

$$\overline{\text{span}\{M\}} = L^2(V). \tag{3.21}$$

The span itself is defined as the set of all possible linear combinations of the elements of M.

For this work, only those TOSs that are countable are of interest. In *separable* Hilbert Spaces, every orthonormal set is countable [Kre78]. As the name already implies, every element in the set (or family) $M = (p_n), n \in \mathbb{Z}$ is orthonormal, i.e. the elements are pairwise orthogonal

$$\langle p_m|p_n\rangle_V = \int_V p_m(\mathbf{x})p_n(\mathbf{x})^* dS(\mathbf{x}) = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n, \end{cases} \tag{3.22}$$

and are of unitary norm

$$\|p_n(\mathbf{x})\|_2 = 1, \forall n \in \mathbb{Z}. \tag{3.23}$$

Equation (3.22) is also sometimes called *orthonormality relation* and is a very basic tool of geometrical sound field analysis, as will be seen later. Another relation for the elements p_n of a TOS is the *completeness* or *closure relation* [Wil99]

$$\sum_{n\in\mathbb{Z}} p_n(\zeta')^* p_n(\zeta) = \delta(\zeta - \zeta') \tag{3.24}$$

with $\zeta, \zeta' \in V$. The Dirac Pulse $\delta(x)$ used in the above equation is defined in Section 3.5.1.

3.4.4 Series Representations in $L^2(V)$

With M as a TOS, any element $f(\mathbf{x}) \in L^2(V)$ can be represented as an infinite series of products of unique coefficients $c_n \in \mathbb{C}$ and their corresponding basis function p_n.

$$f(\mathbf{x}) = \sum_{n\in\mathbb{Z}} c_n p_n(\mathbf{x}) \tag{3.25}$$

The orthonormal nature of the elements in M ensures that the coefficients c_n are unique [Kre78].

3.5 Special Functions and Distributions

3.5.1 Dirac Pulse

The *Dirac pulse* $\delta(x)$ is a distribution (or generalised function) with the following definition [Wil99]:

$$\delta(x) = \begin{cases} +\infty & , x = 0, \\ 0 & , x \neq 0. \end{cases} \tag{3.26}$$

with

$$\int_{\mathbb{R}} \delta(x - x')dx = 1, x' \in \mathbb{R}. \tag{3.27}$$

For the case of finite integration limits, one obtains

$$\int_{-\infty}^{x_0} \delta(x - x')dx = \begin{cases} 0 & , \text{ if } x_0 < x', \\ \frac{1}{2} & , \text{ if } x_0 = x', \\ 1 & , \text{ if } x_0 > x'. \end{cases} \tag{3.28}$$

Also worth mentioning is the very useful sifting property

$$\int_{\mathbb{R}} f(x)\delta(x - x')dx = f(x'), x' \in \mathbb{R}. \tag{3.29}$$

3.5.2 Rectangular Function

The definition for the *rectangular function* $\Pi(\frac{x}{T})$ is given by

$$\Pi\left(\frac{x}{T}\right) := \begin{cases} 1 & , |x| < \frac{T}{2} \\ \frac{1}{2} & , |x| = \frac{T}{2} \\ 0 & , |x| > \frac{T}{2}. \end{cases} \tag{3.30}$$

3.5.3 Sinc Function

The sinc function is defined defined as follows:

$$\operatorname{sinc}(x) := \begin{cases} \frac{\sin x}{x} & , \text{ if } x \neq 0 \\ 1 & , \text{ if } x = 0. \end{cases} \tag{3.31}$$

The value for $x = 0$ results from the application of *L'Hopital's rule*, where

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\frac{d}{dx}\sin x}{\frac{d}{dx}x} = \lim_{x\to 0} \frac{\cos x}{1} = 1 \tag{3.32}$$

3.5.4 Circular Sinc Function

While the sinc function is defined for $x \in \mathbb{R}$, the circular sinc function [Faz10] of the order L is only defined on the interval $[-\pi, \pi]$:

$$\operatorname{csinc}_L(x) := \begin{cases} \frac{\sin(L\frac{x}{2})}{\sin(\frac{x}{2})} & , x \neq 0 \\ L & , x = 0. \end{cases} \tag{3.33}$$

Once more, the value for $x = 0$ can be derived from L'Hopital's rule. Alternatively, $\operatorname{csinc}_L(x)$ can be defined as a finite series of complex exponentials

$$\operatorname{csinc}_L(x) := \sum_{n=-N}^{N} e^{inx} \tag{3.34}$$

with $N = \frac{L-1}{2}$.

Furthermore, using the completeness relation for TOS in (3.24) and using $p_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}$ for the elements of the TOS, it can be shown that

$$\lim_{L\to\infty} \operatorname{csinc}_L(x) = \sum_{n=-\infty}^{\infty} e^{inx} = 2\pi\delta(x). \tag{3.35}$$

3.5.5 The Kronecker Delta

The Kronecker delta δ_n is a sequence that is zero everywhere but for $n = 0$, where it assumes the value 1.

$$\delta_n = \begin{cases} 1, & \text{if } n = 0, \\ 0, & \text{if } n \neq 0, \end{cases} \tag{3.36}$$

It is typically given in the notation δ_{ab} for two integers a and b and is zero for $a \neq b$ and one for $a = b$ [Wil99, BSMM06]. In this work however, the notation in (3.36) will be used for the sake of brevity.

3.6 Fourier Series

In the field of engineering, a very well known Hilbert space is the one spanned by the orthogonal basis functions $\{e_n = e^{inx}, n \in \mathbb{Z}\}$

$$L^2([-\pi, \pi]) := \left\{ f(\mathbf{x}) : \|f(x)\|_2 = \sqrt{\int_{-\pi}^{\pi} |f(x)|^2 dx} < \infty \right\}. \tag{3.37}$$

Every element in $L^2([-\pi, \pi])$ can be represented by an infinite series.

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e_n = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{3.38}$$

This particular series expansion is also known as a *Fourier series* [BSMM06] and it is typically used to represent a T-periodic function

$$g(t) = g(t - kT), k \in \mathbb{Z}.$$

When the interval of the period $[0, T]$ is mapped onto the interval $[-\pi, \pi]$, then $g(t)$ can be written in the form of a Fourier series

$$g(t) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi n \frac{t - T/2}{T}}.$$

3.6.1 Definition

Even though typically used in the form given above, in this work, the Fourier series will be used in a fully normalised form, so that the basis functions form a TOS. The Fourier series is thus redefined to

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e_n = \sum_{n=-\infty}^{\infty} c_n \frac{e^{inx}}{\sqrt{2\pi}}. \tag{3.39}$$

Since the basis functions e_n now form an orthonormal set, the scalar product as defined in (3.14) for $L^2([-\pi, \pi])$ can be used to calculate the *Fourier coefficients* $c_n \in \mathbb{C}$. Given a function f(x)

defined by (3.39), the Fourier coefficients are calculated as follows.

$$c_n = \langle f|e_n\rangle_{[-\pi,\pi]} = \int_{-\pi}^{\pi} f(x)\frac{e^{-inx}}{\sqrt{2\pi}}dx \tag{3.40}$$

For the considerations in this work, the Fourier series expansion as specified by (3.39) and (3.40) will be used and the various variables x fall in the interval $[-\pi, \pi]$.

3.6.2 Parseval's Theorem

For the fully normalised Fourier series introduced, the relation

$$\sum_{n=-\infty}^{\infty} |\langle f|e_n\rangle_{[-\pi,\pi]}|^2 = \|f\|_2^2, \tag{3.41}$$

also known as *Parseval's Theorem* (also Parseval Relation or Parseval Equation), holds [Kre78, BSMM06]. It can also be understood as an equation which underlines that all the energy in a function $f(x)$ is preserved in its Fourier coefficients. This is also due to the orthonormality of the basis (e_n).

3.7 Operators

Operators will be represented by capital letters (e.g. H or G) and are individually defined as they are introduced. The notation for the application of an operator H to a function q is $(Hq)(\mathbf{x})$, where the variables in $\mathbf{x}$ belong to the projection domain of the operator H. A superscript asterisk on an operator denotes its corresponding adjoint operator.

All operators occurring in this work are mapping operators for functions in an L^2 space A to either the same or a different L^2 space B.

3.8 Fourier Transform

3.8.1 Definition

For the transition from time domain considerations to the frequency domain and vice versa, the *Fourier transform* (FT) [Wil99, BSMM06] is typically applied. Following the strict normalisation of the individual Fourier components presented in Section 3.6, the Fourier transform of a time domain function $a(t), t \in \mathbb{R}$ is given by

$$(\mathcal{F}a)(\omega) = \int_{\mathbb{R}} a(t)\frac{e^{i\omega t}}{\sqrt{2\pi}}dt, \tag{3.42}$$

using the operator notation introduced in the previous section. The corresponding inverse transform of a frequency domain function $A(\omega)$ is then given by

$$\left(\mathcal{F}^{-1}A\right)(t) = \int_{\mathbb{R}} A(\omega)\frac{e^{-i\omega t}}{\sqrt{2\pi}}d\omega. \tag{3.43}$$

The transformation pair is then given by

$$\left(\mathcal{F}^{-1}A\right)(t) = a(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad A(\omega) = \left(\mathcal{F}a\right)(\omega) \tag{3.44}$$

Note that the definition of the Fourier transform and the inverse transform given in Eqs. (3.42) and (3.43), respectively, is in accordance with the definition used by Williams [Wil99].

3.8.2 Important Properties and Special Transformation Pairs

The definitions and properties of the FT presented in the following are generally well known in the field, but can all be found in [Wil99, BSMM06] as well. Note that the normalisation of the basis function in the definition of the FT and its inverse (see (3.42) and (3.43) may in some cases lead to a difference by a factor of $\sqrt{2\pi}$.

3.8.2.1 Convolution Theorem

The convolution theorem for two function $f(t)$ and $g(t)$ is given by the transformation pair

$$f(t) * g(t) := \int_{\mathbb{R}} f(t')g(t-t')dt' \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \sqrt{2\pi}F(\omega)G(\omega), \tag{3.45}$$

where $F(\omega) = \left(\mathcal{F}f\right)(\omega)$ and $G(\omega) = \left(\mathcal{F}g\right)(\omega)$. Accordingly, the convolution theorem for the frequency domain is given by

$$F(\omega) * G(\omega) := \int_{\mathbb{R}} F(\omega')G(\omega-\omega')d\omega' \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \sqrt{2\pi}f(t)f(t). \tag{3.46}$$

3.8.2.2 Differentiation Property

Taking the inverse FT given by (3.43) while considering (3.44)

$$a(t) = \int_{\mathbb{R}} A(\omega)\frac{e^{-i\omega t}}{\sqrt{2\pi}}d\omega$$

and applying the differential operator $\frac{d}{dt}$ to both sides of the equation yields the transformation pair

$$\frac{d}{dt}a(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad -i\omega A(\omega). \tag{3.47}$$

3.8.2.3 Plancherel Theorem

The generalised version of Parseval's theorem in (3.41) for the FT is called *Plancherel's theorem* [Pla10, Tit25].

$$\begin{aligned} \int_{\mathbb{R}} f(t)^* f(t) dt &= \int_{\mathbb{R}} F(\omega)^* F(\omega) d\omega \\ \iff \int_{\mathbb{R}} |f(t)|^2 dt &= \int_{\mathbb{R}} |F(\omega)|^2 d\omega \end{aligned} \tag{3.48}$$

Despite its proof was not given until Plancherel's work, this theorem is very often still referred to as Parseval's theorem. The equality can be shown by exploitation of the convolution theorem and the frequency domain representation of the Dirac pulse (see below) and its shifting property.

3.8.2.4 Dirac Pulse

$$\delta(t - t_0) \quad \overset{\mathcal{F}}{\iff} \quad \frac{e^{i\omega t_0}}{\sqrt{2\pi}} \tag{3.49}$$

3.8.2.5 Rectangular Function

$$\Pi\left(\frac{t}{T}\right) \quad \overset{\mathcal{F}}{\iff} \quad \frac{T}{\sqrt{2\pi}} \mathrm{sinc}\left(\omega \frac{T}{2}\right) \tag{3.50}$$

3.9 Coordinate Systems

3.9.1 Cartesian Coordinates

Every point $\mathbf{x} \in \mathbb{R}^3$ is expressed through the triple

$$\mathbf{x} = [x_1, x_2, x_3]^T. \tag{3.51}$$

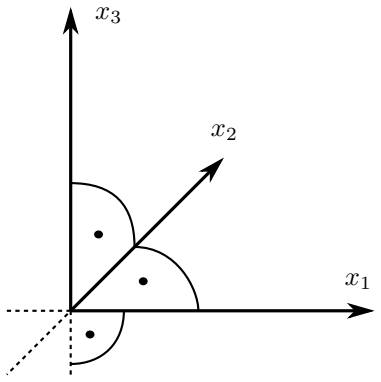


FIGURE 3.1: Cartesian coordinates with the three orthogonal coordinate axes.

3.9.2 Cylindrical Coordinates

Every point $\mathbf{r} \in \mathbb{R}^3$ is expressed by the triple $[r, \phi, z]$, with the following relations to Cartesian coordinates.

$$r = \sqrt{x_1^2 + x_2^2} \tag{3.52}$$

$$\phi = \arctan \frac{x_2}{x_1} \tag{3.53}$$

$$z = x_3 \tag{3.54}$$

and

$$x_1 = r \cos \phi \tag{3.55}$$

$$x_2 = r \sin \phi \tag{3.56}$$

$$x_3 = z, \tag{3.57}$$

respectively.

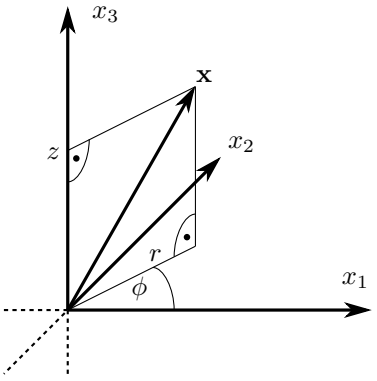


FIGURE 3.2: Cylindrical coordinates in the context of Cartesian coordinate axes.

3.9.3 Spherical Coordinates

Every point $\mathbf{x} \in \mathbb{R}^3$ is expressed by the triple $[r, \phi, \theta]$, with the following relations to Cartesian coordinates.

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{3.58}$$

$$\phi = \arctan \frac{x_2}{x_1} \tag{3.59}$$

$$\theta = \arctan \frac{\sqrt{x_1^2 + x_2^2}}{x_3} \tag{3.60}$$

and

$$x_1 = r \cos \phi \sin \theta \tag{3.61}$$

$$x_2 = r \sin \phi \sin \theta \tag{3.62}$$

$$x_3 = r \cos \theta, \tag{3.63}$$

respectively.

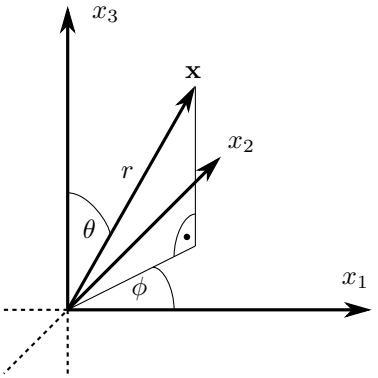


FIGURE 3.3: Spherical coordinates in the context of Cartesian coordinate axes.

Chapter 4

Theoretical Study of Acoustic Circular Arrays With Tangential Pressure Gradient Sensors

Falk-Martin Hoffmann and Filippo Maria Fazi

Abstract

Microphone arrays as a means of sound field acquisition have been the topic of extensive research for more than eight decades now. A number of designs have been suggested, each trying to overcome difficulties that are inherent to either the decomposition of the sound field, the transducers in use or the presence of the array itself. This work presents a theoretical analysis of circular microphone arrays that do not measure the sound pressure but the component of its gradient that is tangential to a given boundary. Its performance is compared to that of a conventional pressure sensor array as a benchmark. The focus of the analysis and subsequent assessment lies on spatial aliasing and performance in the presence of noise. It is shown that a TPG sensor array has improved spatial resolution at low frequencies, provided that the TPG sensors are of comparable quality as pressure sensors with respect to *Signal-to-Noise-Ratio* (SNR). The analysis also reveals that such an array is more vulnerable to spatial aliasing at high frequencies than pressure sensor arrays.

4.1 Introduction

The general idea behind microphone arrays is to record sound not only at one point in space, but to capture information about the sound field, such as the direction of travel of incoming wave fronts. Although it is not necessarily the main objective, one ambitious application of microphone arrays is to make a recording of *the entire sound field*. While in theory this goal can be achieved easily [Wil99], there are various limitations in practice that will degrade the

accuracy of the results. Examples of limiting factors are acoustic transducer noise, impractical requirements on the measurement resolution and assumptions on the sound field that cannot be satisfied in practice. Therefore, the focus of the latest research has been on finding means and ways to overcome those limitations.

The field of array technology is vast and certainly not limited to acoustical applications. For example, Van Veen and Buckley presented a general work on array technologies for beamforming applications in wave fields [VB88], among which are e.g. antennas, sonar systems and microphone arrays. Krim and Viberg presented beamforming with sensor arrays as a problem in the field of signal processing, based on a wave propagation model [KV96] and considering different array geometries. Even though not conceptually limited, this work focuses on airborne wave fields and circular (or cylindrical) arrays.

The earliest approach to a pressure sensor array in the field of airborne acoustics that was applied to make recordings for stereophonic purposes was presented by A. Blumlein [Blu31, Blu35], which had later been modified into using two coincident pressure gradient (figure-of-eight) microphone capsules (also referred to as a "Blumlein-pair"). This recording technology is still used by today's recording engineers, yet it only allows for a two dimensional analysis of the sound field. The first approach to achieve a three dimensional analysis was based on decomposition of the sound field by means of spherical harmonics and was presented by Craven and Gerzon [CG74], thereby founding the ambisonics technique. The ambisonics approach and also the *Higher Order Ambisonics* (HOA) approach involve the representation of a sound field as a weighted series of elementary functions in the spatial domain. These functions are determined by the underlying spherical coordinate system, leading to an infinite set of both radial (Bessel- and Hankel-functions) and angular (spherical harmonics) functions. Their form is independent of the sound field itself, ergo the individual weights of the series are sufficient information to know the sound field in its entirety. This is similar to the Fourier analysis presented in [Wil99]. However, it has proven itself difficult to find reliable ways to obtain these weights from measurements.

The general approach of preceding works [ME02, Raf05, LD07, ZDG10] relies on the exploitation of the orthogonality relation of the underlying basis functions and on its application to the array's measured data. One variation of this approach was presented by Rafaely and Park [Raf04, PR05], who proposed to decompose the sound field into plane waves, which however still involves the exploitation of the orthogonality relation of spherical harmonics. The latter plays a major role in many microphone array designs, but in order for these relations to hold without exception, the sound field must be observed at every single point on the observation boundary. This is so far not feasible in practice, since the number of observation points would need to be infinite. The practical solution is to sample the sound field on that boundary instead. Sampling is however bound to cause aliasing if the observed object's complexity is not within the limitations imposed by the sampling scheme. This is already well known in the field of digitalisation of time domain signals. A thorough analysis of aliasing phenomena in spherical arrays is presented in references [Raf05, RWB07, LD07]. However, aliasing still remains one of the most prominent problems in microphone array design.

Rafaely also identified noise of the array's sensors and inaccuracies in their positioning as significant sources of error, and presented a detailed analysis of these issues in [Raf05]. Supporting these results, Poletti has shown that noise induced by the capsules of the array is a problem, especially

at low frequencies, because the recovery of the information describing the sound field from the array's observation requires considerably high gain factors for higher order modes. He showed that, for a spherical array with pressure sensors on a rigid sphere, the presence of noise may already pose a fundamental problem for the recovery of the fourth order mode, even with a large dynamic range of 120 dB for each capsule [Pol05a].

Fazi and Nelson presented the theory and analysis of the problem of nonuniqueness [FN12], as it occurs in the context of sound field acquisition and reproduction, significantly affecting the design of microphone arrays.

For applications in which it is sufficient to consider the sound field in two dimensions only, circular arrays have been investigated as an alternative to spherical arrays. One such application would be a teleconference scenario where one part of the participants is situated at a table and the array serves to separate the individual speakers into individual audio signals. Another example is that of a humanoid robot that needs to distinguish between people addressing him from different directions, so it can then turn towards them. Meyer presented a work on beamforming in combination with circular microphone arrays mounted on spherical objects and already considered the use of pressure sensors as well as dipole sensors [Mey01]. Teutsch and Kellermann have investigated the theory and practicability of a circular array fitted into a cylindrical baffle for source detection and localisation purposes [TK06]. Kleider et al. [KRWB10] investigated the aliasing behaviour of circular arrays based on a two dimensional analysis of the sound field, while Meyer and Elko [ME08] used a circular array to achieve modal beamforming, still assuming a spherical (three dimensional) sound field model. Poletti [Pol05b] investigated the performance of circular arrays w.r.t. noise and transducer variability.

Ever since Blumlein proposed the recording technique based on pressure gradient sensors, the choice of sensors used in microphone arrays was mainly that of pressure sensors. There are however a few exceptions. Meyer investigated the use of dipole sensors in a circular array arranged on the equator of a sphere [Mey01]. His work considered radially aligned dipoles as well as dipoles aligned in the circumferential direction. Hulsebos et al. presented a theoretical study of a circular microphone array with a large number of cardioid microphones to perform high directivity beamforming for incoming wave fields [HSdVB03]. Their study is largely based on simulation and the use of cardioid sensors was motivated by the prospect of a reduced number of sensors while avoiding zeros in the modal responses of the array (compare to [Pol05a] where this is related to nonuniqueness of the inverse problem). Using a large number of sensors was proposed to push spatial aliasing effects to very high frequencies, however, the effective channel number was further reduced by electrically combining sets of microphones into one virtually large microphone capsule. This can be related to a work by Rafaely et al. in which the spatial aliasing behaviour in spherical microphone arrays is studied [RWB07]. The study by Hulsebos et al. also investigated the system's directivity for plane waves impinging at an angle to the plane of the microphone array. Poletti has considered the application of directional sensors in an open sphere design pointing radially outward, and he found that the problem of nonuniqueness as it occurs with pressure sensors can be overcome [Pol05a]. Melchior et al. proposed a dual radius design of a spherical array with cardioid microphones for use in a binaural auralisation system [MTDG$^+$09]. The dual radius design increases the width of the optimal frequency range. Another more recent design based on differential sensors was presented by Craven, Law and Travis, who proposed a spherical array based on tangential velocity sensors [CLT09]. They found that such

a design allows for a reduced effect of noise at lower frequencies. As theirs was a conceptual study, a theoretical analysis of the design has so far not been presented. Zaunschirm and Zotter proposed a circular table-top microphone array with cardioid microphones and a single pressure sensor for beamforming with both horizontal and vertical steering ability [ZZ14]. They propose a microphone signal decomposition method that is obtained through the (regularised) inversion of a MIMO (Multiple Input Multiple Output) system, extracting the coefficients of a spherical harmonics sound field representation. This enables modal beamforming in the spherical harmonics domain with their proposed array design. Chen et al. proposed a planar microphone array with a hybrid configuration of both pressure sensors as well as first-order (dipole) sensors to enable three-dimensional sound field analysis [CAZ15]. Despite the two-dimensional array aperture, the mathematical model of the sound field is derived from a spherical harmonics decomposition. From their theoretical considerations, they propose a specific design procedure for the considered type of array. They found that the performance of their proposed design compares well to that of spherical microphone arrays.

On the background of the work of Craven, Law and Travis, it should be noted that, in homogeneous sound field conditions, the pressure gradient relates to the velocity by Euler's equation. On the surface of a rigid object, however, this relation is no longer valid and needs to be replaced by a set of equations describing the laminar sound field behaviour within the acoustic boundary layer. These can be derived from the Navier-Stokes equations. A detailed analysis of this is not within the scope of this work, but should be mentioned for the sake of completeness.

This study aims to mathematically confirm the low frequency improvement predicted by Craven et al. and perform a full analysis of the spatial aliasing behaviour of tangential pressure gradient arrays on the basis of a circular array aperture. A theoretical analysis of a circular microphone array is presented, where the array is composed of sensors that measure the tangential component of the pressure gradient. An initial study of this design was presented by the authors in reference [HF14]. This work complements the initial work with further theoretical and numerical in-depth analysis, as well as with an investigation of the noise and aliasing performance and the array's optimal frequency band. In Section 4.2, a mathematical model of the sound field is presented that is based on the Herglotz Wave Function, modelling a sound field as a superposition of plane waves. The relation between the sound field and the observation of the microphone array is expressed by means of an integral operator. This leads to an inverse problem, which is dealt with in a functional analysis framework. This mathematical approach is not very common in the field of transducer array research, but it is a very neat and effective tool to perform the type of analysis this work presents. It has been applied in previous work by Colton and Kress [CK98] and by Fazi [Faz10]. In Section 4.3, the presented array model is discretised and a thorough analysis of the array's aliasing behaviour is undertaken. Section 4.4 provides a simulation based study of the array's performance for the measurement of the approximated sound field of a single plane wave, comparing the recovered sound field information to the theoretical results, evaluating the overall error of the recovery and discussing the significance of aliasing and transducer noise as a problem. The final section summarises the findings of this work and gives a brief overview on upcoming research. It is shown that the predicted improvement of the low frequency performance is confirmed, provided that the sensors are of comparable quality to state-of-the-art pressure sensors, particularly w.r.t. their self-noise and achievable SNR. The spatial aliasing behaviour is shown to be worse than that of conventional arrays with pressure sensors.

4.2 Model of the Sound Field

For the subsequent analysis, a mathematical model describing the pressure of an arbitrary sound field *within* a given region of space (or area, in the two-dimensional case) Λ is needed. The model

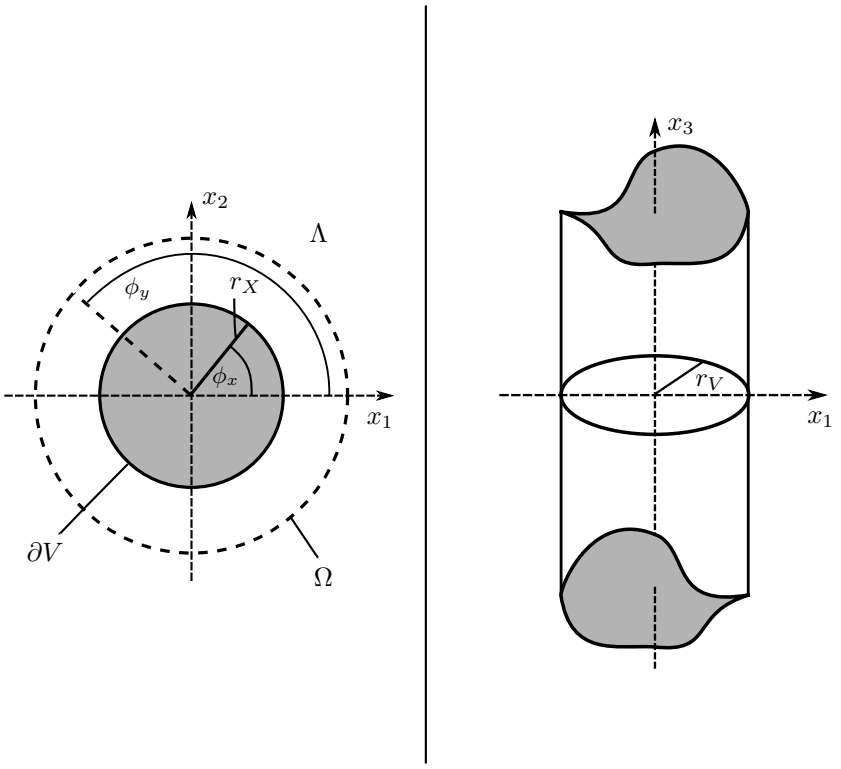


FIGURE 4.1: (left) Schematic description of the volumes and boundaries involved in the sound field model, (right) schematic description of an infinite cylinder of radius r_V.

relies on the assumption that the measured sound field can be represented as the superposition of an infinite number of plane waves. This model is particularly useful when considering a limited number of incoming plane waves (see Section 4.4). When dealing with spherical or cylindrical geometries, it is convenient to represent the plane wave model by means of Fourier expansion. This is the approach chosen in this work.

4.2.1 Plane-Wave Decomposition

The equation

$$p(\mathbf{x},\omega) = \int_{\Omega} H(\mathbf{x},\hat{\mathbf{y}},\omega) q(\hat{\mathbf{y}},\omega) d\Omega(\hat{\mathbf{y}}),\ \mathbf{x} \in \Lambda, \tag{4.1}$$

as given in [FNW12], describes the acoustic pressure at an arbitrary position $\mathbf{x}$ as a linear superposition of an infinite number of plane waves travelling into all possible directions $\hat{\mathbf{y}}$, where $\hat{\mathbf{y}}$ is a unitary vector. $\omega = 2\pi f$ denotes the angular frequency corresponding to the acoustic frequency f and Ω represents the unit sphere or circle in $\mathbb{R}^3$ or $\mathbb{R}^2$, respectively. The above integral equation is also referred to as the *Herglotz Wave Function* (HWF) [FNW12] where the *Herglotz Density* (HD) $q(\hat{\mathbf{y}},\omega)$ describes the complex amplitude of the different plane waves. These are expressed by the kernel $H(\mathbf{x},\hat{\mathbf{y}},\omega) = e^{ik\mathbf{x}\cdot\hat{\mathbf{y}}}$, where $i = \sqrt{-1}$ is the imaginary unit and

k is the acoustic wave number[1]. In principle, equation (4.1) is valid for $\Lambda = \mathbb{R}^3$ (or $\Lambda = \mathbb{R}^2$), provided that the sound field satisfies the homogeneous wave equation in that domain. In practice, this equation is often used to represent a sound field that satisfies the homogeneous wave equation only in a bounded domain Λ.

This representation is used as a foundation for the theory of the baseline microphone array design, using pressure sensors only. For the sake of brevity, the argument ω is omitted in all equations used hereunder, since all calculations are derived for a single frequency ω.

In the subsequent analysis, all considerations are limited to a two-dimensional scenario, assuming that the sound field is constant along the x_3-axis with

$$p(x_1, x_2, x_3) = p(x_1, x_2). \tag{4.2}$$

This is a common means of simplification when a *height-invariant* sound field is observed on the boundary δV of an infinitely long cylinder extending along the x_3-axis, as depicted in Fig. 4.1. The assumption given in equation (4.2) then allows for another simplification from cylindrical coordinates to polar coordinates, so that

$$x_1 = r_x \cdot \cos\phi_x, \text{ and} \tag{4.3}$$

$$x_2 = r_x \cdot \sin\phi_x, \tag{4.4}$$

where $\phi_x = \arctan\frac{x_2}{x_1}$ is the polar angle of a vector $\mathbf{x} = [x_1, x_2]^T$.

Since this study aims at an array design observing the tangential component g of the pressure gradient, an expression similar to (4.1), relating the HD and the *Tangential Pressure Gradient* (TPG) component, needs to be found. With the assumptions made in the previous paragraph, the latter is defined by the scalar product

$$g(\mathbf{x})\Big|_{\mathbf{x}\in\partial V} = \nabla p(\mathbf{x}) \cdot \mathbf{b}(\mathbf{x}). \tag{4.5}$$

∂V denotes the boundary of the microphone array which embeds all the observation points[2] or sensors, respectively. $\mathbf{b}(\mathbf{x})$ represents the unit vector pointing in the direction tangential to ∂V at the observation point $\mathbf{x}$. It is important to realise that, depending on the chosen coordinate system and on the shape of ∂V, the analytical expression for g given by (4.5) may become considerably more complicated.

Let ∂V be a circle of radius r_V. Using the Jacobi-Anger expansion [CK98]

$$e^{ik\mathbf{x}\cdot\hat{\mathbf{y}}} = \sum_{n=-\infty}^{\infty} i^n J_n(kr_x) e^{in\phi_x} e^{-in\phi_y}, \tag{4.6}$$

to replace the kernel, equation (4.1) can be reformulated in polar coordinates, thus obtaining

$$p(r_x, \phi_x) = \sum_{n=-\infty}^{\infty} i^n J_n(kr_x) e^{in\phi_x} \int_0^{2\pi} e^{-in\phi_y} q(\phi_y) d\phi_y. \tag{4.7}$$

[1] Note that the Herglotz Wave Function as given in [FNW12] uses a complex conjugate kernel $e^{-ik\mathbf{x}\cdot\hat{\mathbf{y}}}$ instead, since Fazi et al. define $\hat{\mathbf{y}}$ as the direction of arrival, instead of the direction of propagation.

[2] This denomination is compliant with the work of Williams [Wil99], chapter 8, page 258.

ϕ_x and ϕ_y correspond to the polar angles of $\mathbf{x}$ and $\hat{\mathbf{y}}$, respectively. $J_n(kr_x)$ denotes the Bessel function of order n and describes the radial component of the sound field. To allow for modelling either free field conditions (FF) or the sound field in the presence of a rigid infinite cylindrical scatterer (CS) of radius r_s at the origin, $J_n(kr_x)$ needs to be replaced by a more general *radial function* $R_n(kr_x)$. For reasons of brevity, the complex factor i^n in (4.7) is also included in this radial function, so that the latter is defined by

$$R_n(kr_x) = i^n \begin{cases} J_n(kr_x) & \text{, FF} \\ J_n(kr_x) - \frac{J_n'(kr_s)}{H_n^{(1)\prime}(kr_s)} H_n^{(1)}(kr_x) & \text{, CS} \end{cases} \tag{4.8}$$

where $H_n^{(1)}(x)$ denotes the Hankel function of the first kind of order n, and $J_n'(kr_s)$ and $H_n^{(1)\prime}(kr_s)$ are the derivatives of the Bessel function and the Hankel function, respectively, evaluated at the boundary of the rigid cylindrical scatterer. The derivation of $R_n(x)$ can be found in [Wil99], Section 6.10. For $r_s = r_V$, the tangential component of the pressure gradient on ∂V is now a function of ϕ_x only and is given by [Wil99]

$$g(\phi_x) = \frac{1}{r_x} \frac{\partial p(r_x, \phi_x)}{\partial \phi_x}\bigg|_{r_x = r_V}. \tag{4.9}$$

Equations (4.7) and (4.9) then provide the relation between the HD $q(\phi_y)$ and the pressure on ∂V and its gradient, respectively. In order to obtain $q(\phi_y)$ from the observed pressure $p(\phi_x)$ or from its gradient $g(\phi_x)$, respectively, the corresponding integral equations need to be solved for $q(\phi_y)$.

A common approach [Raf05, Pol05a, MDB06] to obtain a solution, based on the observation of the pressure, is to represent the HD by a Fourier series

$$q(\phi_y) = \sum_{m=-\infty}^{\infty} q_m \frac{e^{im\phi_y}}{\sqrt{2\pi}}, \tag{4.10}$$

where

$$q_m = \int_0^{2\pi} q(\phi_y) \frac{e^{-im\phi_y}}{\sqrt{2\pi}} d\phi_y. \tag{4.11}$$

Replacing $q(\phi_y)$ in (4.7), replacing $i^n J_n(kr_x)$ with $R_n(kr_x)$ and using the orthogonality relation

$$\int_0^{2\pi} e^{im\phi} e^{-in\phi} d\phi = 2\pi\delta_{m,n}, \tag{4.12}$$

leads to the expression of the Fourier coefficients

$$q_m = \frac{1}{R_m(kr_V)(2\pi)^{\frac{3}{2}}} \int_0^{2\pi} e^{-im\phi_x} p(\phi_x) d\phi_x, \forall m \in \mathbb{Z}. \tag{4.13}$$

It is important to realise that this approach leads to a *unique solution* only if $R_m(kr_V) \neq 0, \forall m \in \mathbb{Z}, \forall k$. Similarly, as $R_m(kr_V)$ can converge to zero for high values of m, a solution may exist and be unique, yet be potentially unstable. As shown below, such a solution based *only* on the observation of the TPG on a boundary ∂V recovering all coefficients $q_m, m \in \mathbb{Z}$ cannot be found. This poses a significant problem to the intended array design.

Nevertheless, a derivation similar to that shown in the previous paragraph and based on equation (4.9) would lead to a very similar solution, however, only for this particular geometry. For the sake of a more general approach, the recovery of $q(\phi_y)$ from $g(\phi_x)$ is derived in the following subsection, using functional analysis as a tool.

4.2.2 Analysis of the Integral Operator G

The two quantities of major interest in the given sound field model are the HD $q(\phi)$ and the pressure gradient $g(\phi)$. Both $q(\phi)$ and $g(\phi)$ are assumed to be square-integrable and are considered as elements of an open Hilbert space A. Furthermore, it is assumed that q and g can be described through an infinite weighted sum of orthonormal basis functions $a_n(\phi) \in A$.

$$q(\phi) = \sum_{n=-\infty}^{\infty} q_n a_n(\phi),\ q_n \in \mathbb{C},\ \text{and} \tag{4.14}$$

$$g(\phi) = \sum_{n=-\infty}^{\infty} g_n a_n(\phi),\ g_n \in \mathbb{C}. \tag{4.15}$$

As for the requirement of all functions $q(\phi)$ and $g(\phi)$ being square-integrable, it is convenient to make an exception in the case of the Dirac pulse, which is not square-integrable but can be expressed as an infinite weighted sum of orthonormal basis functions (see Chapter 3, Section 3.24). The Dirac pulse is useful for the considerations made hereunder, e.g. to model the field of a single plane wave.

Note that, since both functions are elements of A, the indices x and y of both r and ϕ have been dropped in the following. The operator H is defined by evaluating equation (4.7) only for $\mathbf{x} \in \partial V$ and by replacing $J_n(x)$ by $R_n(x)$, yielding

$$(Hq)(\mathbf{x}) := p(\phi, r_V) = \sum_{n=-\infty}^{\infty} R_n(kr_V) e^{in\phi} \int_0^{2\pi} e^{-in\phi'} q(\phi') d\phi', \mathbf{x} \in \partial V, \tag{4.16}$$

with $H : A \to A$, which is of no further interest in this study. However, replacing p in (4.9) by (4.16) leads to the integral equation

$$g(\phi) = \sum_{n=-\infty}^{\infty} R_n(kr_V) \frac{1}{r_V} \frac{\partial}{\partial \phi} \left(e^{in\phi}\right) \int_0^{2\pi} e^{-in\phi'} q(\phi') d\phi', \tag{4.17}$$

that is similar to (4.16). This equation defines a new integral operator $G : A \to A$ that maps a given HD $q(\phi)$ to the corresponding pressure gradient $g(\phi)$ on ∂V:

$$(Gq)(\mathbf{x}) := g(\phi) = \sum_{n=-\infty}^{\infty} R_n(kr_V) \frac{in}{r_V} e^{in\phi} \int_0^{2\pi} e^{-in\phi'} q(\phi') d\phi', \mathbf{x} \in \partial V. \tag{4.18}$$

This equation perfectly describes what a TPG sensor located at ϕ observes in a sound field that is defined by the HD $q(\phi)$. This result was presented before in [HF14] and is also very similar to what Meyer described in principle for dipole sensors aligned with a circumferential orientation on the equator of a sphere [Mey01].

Aiming at a more general approach to aliasing analysis, the following subsection briefly introduces the eigenvalue decomposition of G.

4.2.3 Eigenvalue Decomposition of G

G has been identified as a mapping operator from a function $q(\phi) \in A$ to a function in $g(\phi) \in A$. For a better understanding of the relation between q and g, G is analysed w.r.t. how the strength of a mode (see below) of q is transmitted to the corresponding mode of g.

The eigenvalue decomposition of the operator G is based on the equation

$$(Ga_n)(\phi) = \lambda_n a_n(\phi), \tag{4.19}$$

where λ_n denotes the eigenvalue associated to the corresponding eigenfunction $a_n(\phi)$ of G, where the latter is hereafter referred to as a mode. The eigenvalues and eigenfunctions of G for the case under consideration are

$$\lambda_n = 2\pi \frac{inR_n(kr_V)}{r_V}, \tag{4.20}$$

$$a_n(\phi) = \frac{e^{in\phi}}{\sqrt{2\pi}}. \tag{4.21}$$

When comparing equation (4.18) to the results of the *Eigenvalue Decomposition* (EVD) (equations (4.20) and (4.21)), it can be assumed that the operator G can be written as follows:

$$(Gq)(\phi) = \sum_{n=-\infty}^{\infty} a_n(\phi)\lambda_n \langle a_n|q\rangle_\Omega = g(\phi). \tag{4.22}$$

This expression may be interpreted as the spectral decomposition of the compact operator G [CK98], where $\langle f|g\rangle_\Omega = \int_\Omega f^*(x)g(x)d\Omega(x)$ describes the scalar product of two functions in A. This scalar product serves to extract the coefficient q_n of the nth mode a_n as a component of q. The extracted mode strength q_n is then weighted by the corresponding eigenvalue λ_n and multiplied with the mode itself again. This formulation allows for the interpretation of the eigenvalues as coupling factors that describe how the mode strength in $q(\phi)$ is transformed to the resulting function $g(\phi)$ [Faz10].

It is evident that *not all eigenvalues are non-zero*. The most trivial case is that of $\lambda_0 = 0$, which corresponds to the mode $a_0(\phi) = \frac{1}{\sqrt{2\pi}}$. This result implies that a constant

$$q(\phi) = \kappa \in \mathbb{C}, \tag{4.23}$$

is an element of the null space $N(G)$ of the operator G. This can be easily proven when evaluating equation (4.18) for a constant q. In fact, the latter is still of the form given in (4.14) and, hence, an element of A, where $q_n = 0, \forall n \neq 0$. Further zero eigenvalues can arise from an open cylinder array design when $R_n(kr_V) = 0$. The null space of G has a crucial influence on the existence of a unique solution to the inverse problem, as shown in subsection 4.2.5.

4.2.4 The Eigenvalues λ_n

A deeper analysis of the eigenvalues provides a theoretical insight into the system's vulnerability to both aliasing and transducer noise. The latter becomes a problem when the observation $g(\phi)$ is heavily amplified for the recovery of a coefficient q_n, while the former is an inherent product of discretisation. The influence of the eigenvalues on the aliasing behaviour is investigated in more detail in Section 4.3.

In this section, the noise performance is analysed. First, it is necessary to analyse how the strength q_n of a mode $a_n(\phi)$ in $q(\phi)$ is transformed into the corresponding mode strength g_n in $g(\phi)$. With $g(\phi) = (Gq)(\phi)$ and comparing the different factors in equations (4.22) and (4.15), it can be seen that the following equation must hold:

$$g_n = \lambda_n \langle a_n | q \rangle_\Omega = \lambda_n q_n. \tag{4.24}$$

As the scalar product simply extracts the strength of the nth mode from $q(\phi)$, it is evident that the eigenvalues serve as complex gain factors. When assessing the theoretical performance of microphone arrays, it is more interesting to study the above equation after rearranging it for q_n. This yields

$$q_n = \frac{1}{\lambda_n} g_n, \text{ for } \lambda_n \neq 0. \tag{4.25}$$

This shows that the desired mode strength q_n is calculated by extracting g_n from the array observation and then weighting it by $\frac{1}{\lambda_n}$. However, for very small λ_n, the fraction in (4.25) becomes very large. In such a case, when g_n happens to be slightly corrupted by measurement noise, the resulting q_n is very likely to significantly deviate from the actual value, as the noise component in g_n is heavily amplified.

To define if the problem introduced above is of any practical significance, the behaviour of the eigenvalues as a function of frequency needs to be analysed. Fig. 4.2 shows the graphs of the magnitude of the first[3] seven non-zero eigenvalues ($n = 1 \dots 7$) within the audible frequency range. It can be observed that the development of eigenvalues of different orders for a TPG array

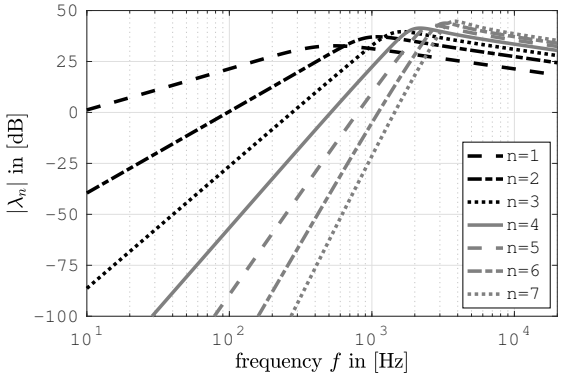


FIGURE 4.2: Magnitude of the eigenvalues $|\lambda_n|$ of G for $n = 1 \dots 7$, $f = 10 \dots 20 \times 10^3$ Hz for an array on a rigid cylindrical structure at the origin with $r_V = 0.1$ m.

[3]Since the eigenvalues have not been ordered so far, the term 'first' simply refers to the index n.

roughly compares to that of the eigenvalues of the operator H

$$\nu_n = 2\pi R_n(kr_V) \tag{4.26}$$

for a conventional pressure sensor array, which is presented in Fig. 4.3. An analogous plot was already presented by Meyer [Mey01] for a circular array mounted on the equator of a sphere and by Elko and Meyer [ME02] for the case of spherical arrays. Poletti presented an equivalent plot for a spherical array with radially aligned first order sensors [Pol05a].

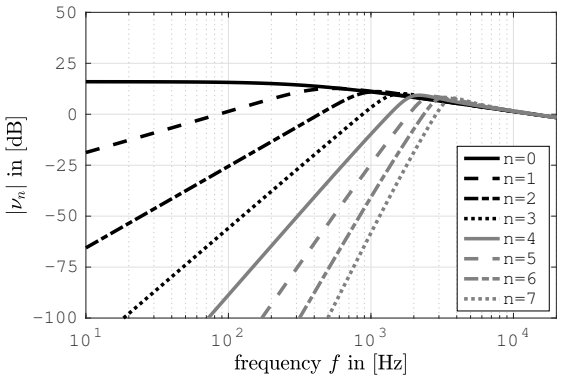


FIGURE 4.3: Magnitude of the eigenvalues $|\nu_n|$ of H of a conventional array based on pressure sensors on a rigid cylindrical structure with $r_V = 0.1\,\text{m}$ for $n = 0 \ldots 7$, $f = 10 \ldots 20 \times 10^3\,\text{Hz}$.

Two major differences between the two figures are that, firstly, Fig. 4.2 does not include λ_0 and, secondly, the additional component $\frac{|n|}{r_V}$, which does not affect the shape of the individual graph (in a dB scale) but translates them vertically.

In conclusion, similar to an array using radially outward pointing gradient sensors [Pol05a], the design investigated in this work is expected to achieve a better noise performance than arrays with pressure sensors [CLT09]. This is not only the case at low frequencies but overall because of the additional gain component mentioned above. Of course, this statement relies on the assumption that pressure gradient sensors and pressure sensors are of equal quality. However, for $n = 7$, the gain that needs to be applied to the measured g_7 for the recovery of q_7 is approximately $+25\,\text{dB}$ at $f = 1\,\text{kHz}$ and $r_V = 0.1\,\text{m}$ with the TPG sensor array, while it is more than $+50\,\text{dB}$ for the same scenario with a pressure sensor array. This leads to the assumption that a TPG sensor array allows for an increased spatial resolution in scenarios with transducer noise (see Section 4.4).

While the overall development of the eigenvalues appears to make the array more robust against noise, the additional component $\frac{|n|}{r_V}$, which depends linearly on the order n, reduces the system's robustness against aliasing. This is further investigated in Section 4.3.

4.2.5 Solution of the Inverse Problem

In order to retrieve $q(\phi)$ from the observation of g, it is necessary to invert G. Considering equation (4.22) and taking into account that the basis functions are orthonormal, the solution to

the inverse problem is

$$\tilde{q}(\phi) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} a_n(\phi)\frac{1}{\lambda_n}\langle a_n|g\rangle_\Omega = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} a_n(\phi)\tilde{q}_n. \tag{4.27}$$

From equation (4.27), it is evident why, in order for $\tilde{q}$ to be bounded, the mode $a_0(\phi)$ corresponding to the eigenvalue $\lambda_0 = 0$ must be excluded from the solution. As a matter of fact, the mode $a_0(\phi)$ is in the null space of the adjoint operator G^* and therefore does not satisfy the first Picard condition for the uniqueness of a solution [CK98, Faz10]. Furthermore, considering equation (4.20) it is also evident why the solution does not exist when

$$R_n(kr_V) = 0. \tag{4.28}$$

The limitation implied by (4.28) is a well-known problem in the field of microphone array research [Pol05a, FN12], which is often overcome in practice by choosing designs based on a rigid array structure.

Another requirement for the solution to be bounded is that the expression

$$\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{|\langle a_n|g\rangle_\Omega|^2}{\lambda_n^2} < \infty$$

is satisfied, which corresponds to the second condition of Picard's theorem. In theory, this is potentially not satisfied; however, the order truncation discussed in Section 4.3 ensures that this condition is always satisfied. Finally, the explicit expression of the solution is

$$\tilde{q}(\phi) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} e^{in\phi}\frac{-ir_V}{4\pi^2 nR_n(kr_V)}\int_0^{2\pi} e^{-in\phi'}g(\phi')d\phi'. \tag{4.29}$$

It follows from (4.27), (4.29) and (4.12) that the series coefficients are given by

$$\tilde{q}_n = \frac{-ir_V}{(2\pi)^{\frac{3}{2}} nR_n(kr_V)}\int_0^{2\pi} e^{-in\phi'}g(\phi')d\phi', \forall n \in \mathbb{Z}\backslash\{0\}. \tag{4.30}$$

Equations (4.29) and (4.30) have already been presented in [HF14] but they were derived from a Singular Value Decomposition instead. The result in (4.30) is very similar to that given in (4.13). The minor difference in the denominator of the fraction is due to the use of the TPG instead of the pressure itself.

The solution presented above leads to a valid HD. However, it is crucial to realise that, since the nullspace of G (see eq. (4.23)) is non-trivial, the solution $\tilde{q}(\phi)$ recovered from the observation of g may differ from the original HD $q(\phi)$ by an additional constant factor α, because any solution

$$q(\phi) = \tilde{q}(\phi) + \alpha,\ \alpha \in \mathbb{C} \tag{4.31}$$

is also a solution. This must be the case since $\alpha = q_0$ (compare eq. (4.11) for $m = 0$), which can also be extracted from $q(\phi)$ by exploiting orthogonality (see eq. (4.12)).

The zero-order mode represents the direct component q_0 in (4.14). The physical interpretation of this mode is most easily understood, considering the source-receiver reciprocity, as the sound field generated by a breathing cylinder, i.e. an outgoing or incoming wave with a constant magnitude and phase for all angles at a fixed radius r_V. This can be easily proven by evaluating (4.7) for $q(\phi) = q_0 a_0(\phi) = \frac{q_0}{\sqrt{2\pi}}$. As a consequence, the TPG g is not affected by a direct component q_0 of q, which unfortunately implies that q_0 *cannot be recovered* from the knowledge of g. Moreover, the implicit assumption made in this section, that $q_0 = 0$, imposes a significant restriction to generality.

These findings are rather discouraging, considering the original objective to capture an entire sound field from an array consisting of TPG sensors only. Nevertheless, it is shown below that this problem can be overcome in certain conditions by adding one or more pressure sensors to the array.

4.3 Discretisation and Aliasing Analysis

Similar to the sampling of a time domain signal, the sampling of the pressure gradient on a circle with radius r_V leads to spatial aliasing effects, if the sound field contains modes of higher order than those captured by the array. These higher order modes are bound to corrupt the observed modes. An odd number L of TPG sensors distributed uniformly on the circle allows for the recovery of modes $a_n(\phi), |n| \leq N$ with

$$N = \frac{L-1}{2}. \tag{4.32}$$

Note that one might be led to assume that it would suffice to have $L = 2N$ sensors in order to solve a fully determined linear system of equations to recover the coefficients q_n for $1 \leq |n| \leq N$. While from a linear algebra point of view that may be true, such a configuration results in sampling below the Nyquist rate that requires more than two samples within a period of the highest modal component. As a consequence, the choice of $L = 2N$ would affect the spatial aliasing behaviour of the system, i.e. the highest order in particular would likely be compromised.

The reconstruction of a sound field containing modes of order higher than N is investigated in the following. For reasons of brevity, the argument of the radial functions $R_n(kr_V)$ has been omitted hereunder.

The integral in (4.30) is discretised with L sampling points and the equation is then rewritten as

$$\tilde{q}_n = \frac{-ir_V}{nR_n(2\pi)^{\frac{3}{2}}} \sum_{l=1}^{L} e^{-in\Delta\phi l} g(\Delta\phi l)\Delta\phi, \forall n \in \mathbb{Z}\backslash\{0\} \tag{4.33}$$

where $\Delta\phi = \frac{2\pi}{L}$ is the angular spacing between neighbouring sensors. Note that a uniform sampling is not strictly necessary and other sampling schemes (e.g. Golden-Ratio sampling [KRWB10]) could be considered in principle. In this work, however, only uniform sampling and the corresponding aliasing pattern (see next section) is considered.

4.3.1 The Aliasing Pattern

Replacing g by the series expansion given in (4.15) and using (4.21) leads to

$$\tilde{q}_n = \frac{1}{\lambda_n} \sum_{m=-\infty}^{\infty} g_m A_{m,n}, \tag{4.34}$$

where the factors

$$A_{m,n} = \frac{1}{L} \sum_{l=1}^{L} e^{-in\Delta\phi l} e^{im\Delta\phi l} = \begin{cases} 1\ , m = n + uL, u \in \mathbb{Z} \\ 0\ ,\ \text{otherwise} \end{cases} \tag{4.35}$$

are the elements of a matrix that describes the aliasing pattern of the system. The results in (4.34) and (4.35) have already been found in a similar form by Poletti in [Pol05b]. Fig. 4.4 shows the values of (4.35) evaluated for a circular array with $L = 15$ microphones and $n \in [-7, \ldots, 0, \ldots, 7]$ and $m \in [-22, \ldots, 0, \ldots, 22]$. Black cells indicate values of $A_{m,n}$ that are different from zero,

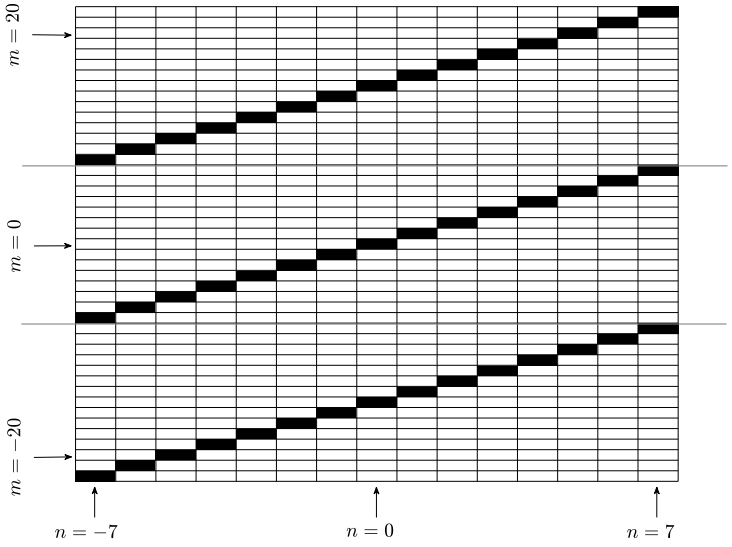


FIGURE 4.4: Visualisation of the orthogonality matrix $A_{m,n}$ for a circular array with $L = 15$ microphones. Each cell represents whether two modes a_m and a_n are orthogonal (white cell) or not (black cell), when sampled at L points on a circle. The two light grey lines indicate the mode range ($\pm N$) of the array.

hence indicating a pair of modes that are *not orthogonal* to each other. As such, Fig. 4.4 is a good indicator of which modes n recovered by the array are corrupted by modes of order m of the sound field - in other words, aliasing. For example, a recovered coefficient $\tilde{q}_7$ of the mode corresponding to $n = 7$ is corrupted by aliasing, if the sound field contains modes of order $m = 7 + uL, u \in \mathbb{Z}$, i.e. $m = -8$ and $m = 22$ as given in Fig. 4.4.

Repeating the same steps as above, starting from equation (4.13) leads to the same orthogonality matrix $A_{m,n}$ for an array composed of pressure sensors. This shows the aliasing scheme to be the same for both array types.

Using eq. (4.24) in (4.33) and exploiting (4.35) leads to the following equation describing precisely the aliasing behaviour

$$\tilde{q}_n = q_n + \sum_{\substack{u=-\infty \\ u\neq 0}}^{\infty} \frac{\lambda_{n+uL}}{\lambda_n} q_{n+uL}. \tag{4.36}$$

Following mathematical passages similar to those presented above, it can be easily shown that equations (4.24) and (4.36) are the same for a pressure sensor array but with different eigenvalues ν_n [Faz10].

The ratio of two eigenvalues $\frac{\lambda_{n+uL}}{\lambda_n}$ in (4.36) has a major effect on the vulnerability to aliasing. The reason for this can be seen from the graphs of the eigenvalues in Fig. 4.2. The additional factor $|n|$ leads to an *amplified* contribution of aliased modes on the TPG observation. It can therefore be expected that for the proposed design the aliasing induced at high frequencies by high order modes is worse compared to that of pressure sensor arrays for which the ratio $|\frac{\nu_{n+uL}}{\nu_n}|$ remains close to one (compare to Fig. 4.3 and the results in [Pol05a]). However, since the maximum of λ_n shifts towards higher frequencies with increasing n (see Fig. 4.2), this effect is only relevant up to a certain u_a where $\frac{\lambda_{n+u_aL}}{\lambda_n} < 1$.

Considering the case of $n = 0$, Fig. 4.4 does not give any information on the aliasing pattern for the 0th order mode since q_0 cannot be recovered from an array using pressure gradient sensors only. To overcome this limitation, the array needs to be extended by at least one pressure sensor, leading to an overall amount of $L + 1$ sensors.

A linear algebra formulation of the mode recovery problem is introduced in the following subsection and then one additional pressure sensor is included into the system.

4.3.2 Mode Recovery using Linear Algebra

In the following, the objective is to recover modes up to order N from the data acquired with L sensors. For a given HD with limited order N (aliasing free case)

$$q(\phi) = \sum_{n=-N}^{N} q_n a_n(\phi), \tag{4.37}$$

the result of equation (4.18) can be rewritten as a function of the coefficients q_n. Using (4.15), (4.21) and (4.24) yields a solution for the TPG with limited order N, given by

$$g(\phi) = \sum_{n=-N}^{N} \frac{in}{r_V} R_n \sqrt{2\pi} e^{in\phi} q_n. \tag{4.38}$$

The expression of (4.38), when evaluated for all L observation points, leads to a system of linear equations, which can be written using matrix notation:

$$\mathbf{g} = \begin{bmatrix} g(\Delta\phi \cdot 1) \\ \vdots \\ g(\Delta\phi \cdot L) \end{bmatrix}_{L\times 1} = \mathbf{HJ} \begin{bmatrix} q_{-N} \\ \vdots \\ q_{-1} \\ q_1 \\ \vdots \\ q_N \end{bmatrix}_{2N\times 1} = \mathbf{HJq}, \tag{4.39}$$

where

$$\mathbf{J} = \frac{i\sqrt{2\pi L}}{r_V} \cdot \mathrm{diag}(-NR_{-N}, \ldots, -R_{-1}, R_1, \ldots, NR_N) \tag{4.40}$$

and

$$\mathbf{H} = \frac{1}{\sqrt{L}} \begin{bmatrix} e^{-iN\Delta\phi 1} & \cdots & e^{-i1\Delta\phi 1} & e^{i1\Delta\phi 1} & \cdots & e^{iN\Delta\phi 1} \\ \vdots & & \vdots & \vdots & & \vdots \\ e^{-iN\Delta\phi L} & \cdots & e^{-i1\Delta\phi L} & e^{i1\Delta\phi L} & \cdots & e^{iN\Delta\phi L} \end{bmatrix}. \tag{4.41}$$

Note that the mode coefficient q_0 is not included in any of these equations. In order to recover the mode vector $\mathbf{q}$ from the observations $\mathbf{g}$, the linear equation system needs to be solved by matrix inversion. $\mathbf{J}$ is a diagonal matrix is therefore trivial to invert. However, with $L > 2N$, the overall equation system is overdetermined, hence $\mathbf{H}$ needs to be inverted by its Moore-Penrose Pseudoinverse [Hay02]

$$\mathbf{H}^\dagger = (\underbrace{\mathbf{H}^H\mathbf{H}}_{\mathbf{I}_{2N}})^{-1}\mathbf{H}^H = \mathbf{H}^H, \tag{4.42}$$

where $\mathbf{I}_{2N}$ denotes the $2N \times 2N$ unity matrix. The least-squares estimate $\tilde{\mathbf{q}}$ of $\mathbf{q}$ based on the observations $\mathbf{g}$ is then given by

$$\tilde{\mathbf{q}} = \mathbf{J}^{-1}\mathbf{H}^H\mathbf{g} = \underbrace{\mathbf{J}^{-1}\mathbf{H}^H\mathbf{HJ}}_{\mathbf{I}_{2N}}\mathbf{q} = \mathbf{q} \tag{4.43}$$

and hence even an *exact solution*. This is because the chosen $\mathbf{g}$ has no components in the nullspace of $\mathbf{H}^H$ and is order-limited (aliasing free case). An alternative way to obtain the weights q_n is given by equation (4.33).

In the next subsection, the recovery of the 0th mode is described.

4.3.3 Recovery of the 0th Mode using an additional pressure sensor

An additional pressure sensor located at $\phi = \phi_p$ is used as a means to recover the coefficient q_0, when all other sensors are TPG sensors.

Similar to the TPG in (4.38), the pressure at ϕ_p can be expressed as a function of q_n, where the HD $q(\phi)$ has still a limited order N. Combining (4.16) and (4.10) leads to

$$p(\phi) = \sum_{n=-N}^{N} R_n \sqrt{2\pi} e^{in\phi} q_n. \tag{4.44}$$

This sum can also be expressed using linear algebra:

$$p(\phi_p) = \begin{bmatrix} \sqrt{2\pi} R_0 & \mathbf{b} \end{bmatrix} \begin{bmatrix} q_0 \\ q_{-N} \\ \vdots \\ q_{-1} \\ q_1 \\ \vdots \\ q_N \end{bmatrix} \tag{4.45}$$

with

$$\mathbf{b} = \sqrt{2\pi} \begin{bmatrix} R_{-N} e^{-iN\phi_p} \cdots R_{-1} e^{-i1\phi_p} & R_1 e^{i1\phi_p} \cdots R_N e^{iN\phi_p} \end{bmatrix}.$$

With the recovery of all $2N+1$ coefficients but q_0 from the TPG observation, the only unknown in equation (4.45) is q_0. Using the result for $\tilde{\mathbf{q}}$, the solution for $\tilde{q}_0$ is then given as

$$\tilde{q}_0 = \frac{1}{\sqrt{2\pi} R_0} \left[p(\phi_p) - \mathbf{b}\tilde{\mathbf{q}} \right] = q_0, \tag{4.46}$$

which is *also an exact solution* when (4.38) is satisfied.

An equivalent formulation of the same result is to comprise the full array observation $\mathbf{g}$ and $p(\phi_p)$ in an extended equation system

$$\mathbf{g}_e = \begin{bmatrix} p \\ \mathbf{g} \end{bmatrix} = \mathbf{Y} \begin{bmatrix} q_0 \\ \mathbf{q} \end{bmatrix} = \mathbf{Y}\mathbf{q}_e, \tag{4.47}$$

$$\mathbf{Y} = \begin{bmatrix} \sqrt{2\pi} R_0 & \mathbf{b} \\ \mathbf{0} & \mathbf{HJ} \end{bmatrix}_{(L+1)\times(2N+1)}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{L\times 1}.$$

The system matrix $\mathbf{Y}$ is a composition of different matrices and vectors, which clearly limits the contribution of the $0th$ mode to the observed pressure p. The subscript e serves to distinguish between the original and the extended vectors. In the light of the results found in (4.42), (4.43) and (4.46), the Moore-Penrose Pseudoinverse of $\mathbf{Y}$ is of the form

$$\mathbf{Y}^{\dagger} = \begin{bmatrix} \frac{1}{\sqrt{2\pi} R_0} & \mathbf{v} \\ \mathbf{0} & \mathbf{J}^{-1}\mathbf{H}^H \end{bmatrix}, \tag{4.48}$$

which has been confirmed by numerical results. As already implied by eq. (4.48), it is clear that p in $\mathbf{g}_e$ is not used to recover any coefficients other than q_0, because the contribution of q_0 to p cannot be compensated for, in general, by a linear combination of the elements of $\mathbf{g}$.

In fact, the analytic expression of the vector $\mathbf{v}$ can be derived from the findings above. From (4.43), (4.46) and (4.48) it is evident that $\mathbf{v}$ needs to satisfy

$$\frac{-1}{\sqrt{2\pi}R_0}\mathbf{bq} = \frac{-1}{\sqrt{2\pi}R_0}\mathbf{bJ}^{-1}\mathbf{H}^H\mathbf{g} = \mathbf{vg} \tag{4.49}$$

$$\implies \mathbf{v} = \frac{-1}{\sqrt{2\pi}R_0}\mathbf{bJ}^{-1}\mathbf{H}^H. \tag{4.50}$$

The L components of $\mathbf{v}$ are therefore

$$v_l = \frac{ir_V}{LR_0}\sum_{\substack{n=-N\\n\neq 0}}^{N}\frac{1}{\sqrt{2\pi}n}e^{in\phi_P}e^{-in\Delta\phi l}, l \in [1,\ldots,L]. \tag{4.51}$$

The expression in (4.49) defines $\mathbf{v}$ uniquely, since $\mathbf{g}$ can be any element of an L dimensional vector space.

The final result for all recovered mode coefficients is

$$\tilde{\mathbf{q}}_e = \mathbf{Y}^\dagger\mathbf{g}_e. \tag{4.52}$$

It has been shown that, provided the HD $q(\phi)$ defining the sound field leading to the observation of $\mathbf{g}_e$ does not contain any modes $a_n(\phi), |n| > N$, the application of the Moore-Penrose Pseudoinverse $\mathbf{Y}^\dagger$ leads to an exact solution for $\tilde{\mathbf{q}}_e$, and all coefficients q_n, including the 0th order coefficient q_0, are recovered correctly.

Unfortunately, sound fields are typically not order-limited. The aliasing pattern for an array observing the TPG only has already been presented in subsection 4.3.1; the consequences of aliasing on the recovery of q_0 are investigated in the following subsection.

4.3.4 Robustness of the 0th Order

It has been shown in the previous subsection that the mode coefficient q_0 can be recovered successfully if the condition

$$q_n = 0, \forall |n| > N \tag{4.53}$$

is satisfied (compare to eq. (4.46)). This subsection deals with the recovery of the coefficients q_n in the presence of spatial aliasing.

For a HD of infinite order, equation (4.36) describes exactly how the elements of $\tilde{\mathbf{q}}$ are distorted, yet it does not predict how $\tilde{q}_0$ is affected. Evidently, from equation (4.46) it can be seen that a recovered coefficient vector $\tilde{\mathbf{q}}$ corrupted by aliasing also results in an inaccurate recovery of q_0, namely

$$\tilde{q}_0 \neq q_0.$$

As a consequence, q_0 can only be recovered accurately if (4.53) holds. Whenever this condition is not satisfied, not only are the observed coefficients $\tilde{\mathbf{q}}$ degraded in accordance with the scheme given by orthogonality matrix $A_{m,n}$, but q_0 is subject to aliasing.

In conclusion, the aliasing pattern of the array with the additional pressure sensor can only partially be described by $A_{m,n}$ as the corruption of the 0th order mode is not accounted for. Fig. 4.5 qualitatively indicates the resulting orthogonality matrix $\tilde{A}_{m,n}$ obtained after combining $A_{m,n}$ with the effect that corrupted higher orders $\tilde{q}_n, 0 < |n| \leq N$ have on the 0th order.

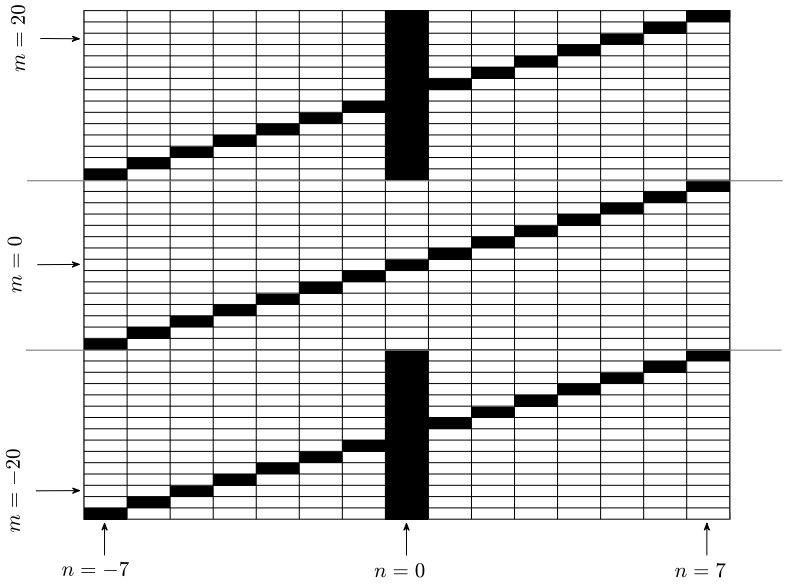


FIGURE 4.5: Schematic visualisation of the new orthogonality matrix $\tilde{A}_{m,n}$ for a circular array with $L = 15$ HD sensors and one pressure sensor. The two light grey lines indicate the mode range ($\pm N$) of the array.

When comparing the aliasing patterns in Fig. 4.4 and Fig. 4.5, it can be seen that the TPG sensor array with an additional pressure sensor is more vulnerable to aliasing than a system based only on pressure sensors.

4.3.5 HD Coefficients of the Sound Field of a Plane Wave

The HD of a single plane wave of unity magnitude travelling in the direction ϕ_i is given by

$$q(\phi) = \delta(\phi - \phi_i). \tag{4.54}$$

This leads to the following TPG

$$g(\phi) = \frac{1}{r_V} \sum_{n=-\infty}^{\infty} inR_n e^{in\phi} e^{-in\phi_i}. \tag{4.55}$$

The Fourier coefficients of the HD given in (4.54) are

$$q_n = \int_0^{2\pi} q(\phi) \frac{e^{-in\phi}}{\sqrt{2\pi}} d\phi = \frac{1}{\sqrt{2\pi}} e^{-in\phi_i}. \tag{4.56}$$

For comparison with the work of Williams [Wil99], the following equation describes the relation between the coefficients used in this work and the helical wave coefficients C_n in Williams' work

(compare [Wil99], Section 4.3, pp. 121 ff).

$$q_n = \frac{i^{-n}}{(2\pi)^{\frac{3}{2}}} C_n \tag{4.57}$$

Equation (4.56) shows that for a single plane wave, the magnitude of all coefficients is $\frac{1}{\sqrt{2\pi}}$.

4.3.6 Example of Mode Recovery With Spatial Aliasing

Fig. 4.6 shows the mode recovery performance of an array of $L = 15$ TPG sensors and one pressure sensor. The simulated sound field has a limited order $N_s \geq N$ ($q_n = 0, \forall |n| > N_s$) and the coefficients q_n are specified by (4.56). The wave field is an approximation of a plane wave within a radius $r_{PW} \approx \frac{N_s}{k}$ around the origin [WA01, Ahr12]. The theoretical magnitude of the

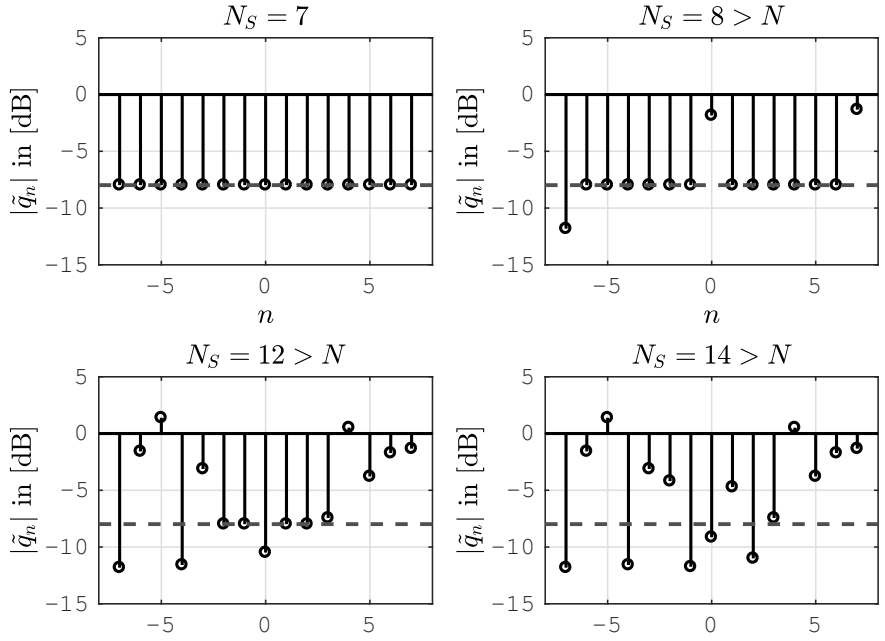
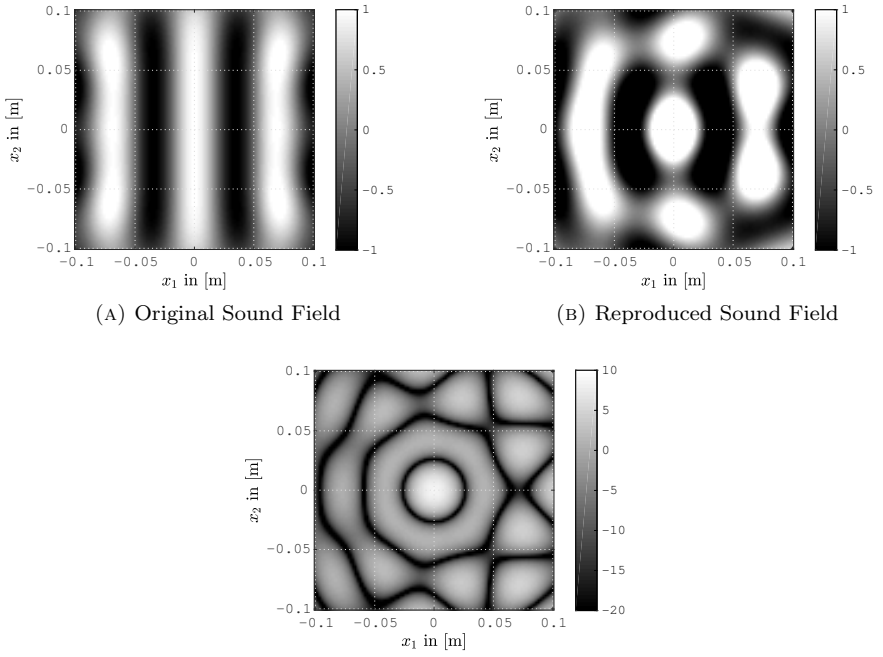


FIGURE 4.6: Mode recovery performance of a circular array composed of $L = 15$ TPG sensors and one pressure sensor with $r_V = 0.1\,\text{m}$. The incoming sound field has a frequency $f = 5\,\text{kHz}$ and the results are shown for $N_s = [7, 8, 12, 14]$. The dashed grey line indicates the theoretical values for the magnitude of the HD coefficients, while the black stems display the recovered coefficients' magnitude.

HD coefficients is indicated by a dashed grey line in all four graphs. The top left graph shows that for $N_s = 7 = N$, the observed mode coefficients are recovered perfectly. The other three graphs however confirm the findings of subsection 4.3.4 and the aliasing scheme shown in Fig. 4.5. The graph for $N_s = 8$ shows that q_{-7}, q_7 and q_0 are affected by aliasing. For $N_s = 12$, only the coefficients $q_n, n \in [-2, -1, 1, 2]$ are still recovered correctly, while for $N_s = 14$ all coefficients are distorted by aliasing. This is consistent with the pattern given in Fig. 4.5. The simulated sound field ($\phi_i = 0, f = 5\,\text{kHz}, c = 343\,\frac{\text{m}}{\text{s}}, N_s = 8$) and its reproduction based on the observed information $\tilde{\mathbf{q}}$ can be seen in Fig. 4.7 for $N_s = 8$. Despite the order limitation, it still resembles that of a plane wave. The middle picture shows that the shape of the incoming wave fronts are significantly deformed, even though only three modes have been corrupted by aliasing (compare



(A) Original Sound Field

(B) Reproduced Sound Field

(C) Absolute Amplitude Error in dB

FIGURE 4.7: Visualisation of the sound field with $\phi_i = 0$, $f = 5\,\text{kHz}, c = 343\,\frac{\text{m}}{\text{s}}$, limited to order $N_s = 8$, (A) original sound field, (B) sound field as reproduced from the observation of $\tilde{q}_n$, using the microphone array described above, (C) absolute amplitude error of the reproduced field, given in dB.

to Fig. 4.5 and Fig. 4.6). Especially the influence of the corrupted 0th order mode is clearly visible in the form of the concentric circle around the origin in the bottom graphic depicting the difference between the original and the reconstructed sound field.

4.4 Simulations and Quantitative Error Analysis

As it has been argued in subsection 4.2.4, it can be expected that TPG sensor arrays are more robust against measurement noise than pressure sensor arrays. This is due to the considerable difference in the magnitude of the eigenvalues and is most significant at low frequencies. At the same time, the findings in subsections 4.3.1 and 4.3.4 indicate that the TPG sensor design is more vulnerable to spatial aliasing than conventional designs. In order to confirm both hypotheses, a vast number of simulated measurements of an arbitrary sound field specified by (4.56) with different ϕ_i have been performed in MATLAB. These involve a simulated TPG array with $L_G = 15$ uniformly spaced TPG sensors plus an additional pressure sensor at $\phi_p = \frac{\pi}{6}$, yielding a total amount of 16 sensors. The simulated conventional array consists of $L_P = 16$ uniformly spaced pressure sensors. That leaves both arrays with the same number of sensors overall and both had a radius of $r_V = 0.1\,\text{m}$. They were both set to recover all the HD coefficients q_n for $|n| \leq 7$. The signals observed at the various array sensors have been computed using the analytical solutions in equation (4.38) for the TPG and (4.44) for the pressure, both accounting for orders up to $|n| = N_s = 40$. Note that this will lead to spatial aliasing with the given array at high frequencies. The needed HD coefficients q_n are specified by equation (4.56).

The following three subsections describe the noise model, the regularisation of the inverse matrix $\mathbf{Y}^{\dagger}$ and the quality measure used in the simulations. Subsection 4.4.4 then presents the quantitative performance results for both array types.

4.4.1 The Noise Model

In order to compare the robustness of the two different arrays to measurement noise, the simulated measured signals were corrupted with a noise signal d. The noise in (ultrasonic) transducers has been identified to originate predominantly from the electronic circuitry [HBR95, Oak97]. Electrical noise consists mainly of four components:

- thermal noise due to thermal agitation of charges in conductors [Nyq28],
- shot noise (as it typically occurs in semiconductors and vacuum tubes),
- $1/f$ noise (e.g. caused by impedance fluctuations in combination with DC currents occurring in the sensor's internal impedance converters) and
- flicker noise [Sch26] (e.g. as it occurs in MOSFETs, which are frequently used as internal impedance converters).

The last three kinds of noise are decaying in level towards higher frequencies, where they are covered by the thermal noise, which is spectrally white. Assuming transducers with carefully designed impedance converters in combination with high quality signal pre amplifiers, it is deemed valid to model the overall noise with a white spectrum. Additionally, it is assumed that the conversion principle between the acoustic and the electrical quantities as well as the thermal agitation of the molecules in the fluid before the transducer are negligible compared to the electrical noise. Therefore, the overall noise can be modelled as a random signal with the same constant average magnitude at all frequencies for both types of transducers. For an individual sensor, the additive noise component is then specified by

$$d = \sigma_n e^{i2\pi\mu}, \tag{4.58}$$

where μ is a *uniformly* distributed random variable on the interval $[0, 1]$ and σ_n is the standard deviation of the noise. To further randomise the nature of the noise, σ_n was chosen to also be a random variable

$$\sigma_n = 1 \cdot 10^{-60/20}\xi, \tag{4.59}$$

where ξ is a *normally* distributed random variable with unit variance. The *average* magnitude of the noise is then set to 60 dB below that of a pressure sensor's output signal in the undisturbed field of a plane wave with unit amplitude. Note that on a rigid cylinder, the average magnitude of the measured signals for either type of sensors depends on the sensor's position, the frequency and the direction of travel of the incoming plane wave.

The final signal model for the two sensors is then given as follows:

$$\tilde{g}(\phi_s) = g(\phi_s) + d, \tag{4.60}$$

$$\tilde{p}(\phi_s) = p(\phi_s) + d. \tag{4.61}$$

4.4.2 Regularisation Against Ill-Conditioning

The recovery of the coefficients q_n from noisy measurements poses an ill-conditioned problem at low frequencies due to the excessive gains applied as a consequence of the inversion of the eigenvalues λ_n and ν_n. Therefore, these simulations require a regularised Moore-Penrose Pseudoinverse $\mathbf{Y}_R^\dagger$ to solve the inverse problem. In this work, a Tikhonov regularisation [Mal08] has been applied. The regularised pseudoinverse is then defined as

$$\mathbf{Y}_R^\dagger = (\mathbf{Y}^H\mathbf{Y} + \Gamma^T\Gamma)^{-1}\mathbf{Y}^H, \tag{4.62}$$

where $\Gamma = \beta\mathbf{I}$. The matrix $\mathbf{I}$ is an $(L_G+1)\times(L_G+1)$ identity matrix for the TPG sensor array and an $L_P \times L_P$ identity matrix for the pressure sensor array, respectively. The value for β was chosen empirically and set to $\sqrt{10^{\frac{-50}{20}}}$, which is equivalent to impose a maximum bound of 50 dB to the largest eigenvalue of the inverse matrix. An alternative way of choosing β on the basis of the underlying data is through analysis of the L-curve [Han98, Han00]. A second alternative based on the measurement SNR was proposed by Alon et al. [ASR15]. In both cases, β is likely to be frequency dependent. The coefficients $\tilde{q}_{n,R}$ are then calculated from the noisy array observation $\tilde{\mathbf{g}}_e$ using

$$\tilde{\mathbf{q}}_{e,R} = \mathbf{Y}_R^\dagger\tilde{\mathbf{g}}_e. \tag{4.63}$$

The elements of $\tilde{\mathbf{g}}_e$ are computed for orders up to $|n| = 40$ and contain noise. The coefficients in $\tilde{\mathbf{q}}_{e,R}$ are thus corrupted by measurement noise, regularisation and spatial aliasing effects.

4.4.3 The Quality Measure

The accuracy of the recovered coefficients $\tilde{q}_{n,R}$ can be evaluated based on the energy of the error between the original HD $q(\phi)$ and the recovered HD $\tilde{q}(\phi)$. The *Error-to-Signal-Ratio* (ESR) in dB for a given frequency is then given by

$$\mathrm{ESR} = 10\cdot\log_{10}\left(\frac{\sum_{n=-N}^{N}|\tilde{q}_{n,R}-q_n|^2}{\sum_{n=-N}^{N}|q_n|^2}\right)\ \mathrm{dB}. \tag{4.64}$$

4.4.4 Simulation Results

The ESR is used to compare the performance of the TPG array and the pressure array in a statistical analysis for a finite number of frequencies in the range between 100 Hz and 10 kHz with a step size of 10 Hz. For each considered frequency f, 100 simulated measurements were conducted: i.e. 10 iterations each for 10 randomly selected approximated plane wave fields (see subsection 4.3.6). This allows for a good approximation of the noise model and the aliasing performance is not biased by the choice of specific directions of incidence for the incoming waves.

Fig. 4.8 shows the result of the analysis, where the black line depicts the performance of the TPG sensor array and the light grey line the performance of the pressure sensor array. The plots clearly show the effects of spatial aliasing and noise. While the error at high frequencies is dominated by the effect of spatial aliasing, the effect of transducer noise is dominant at low frequencies.

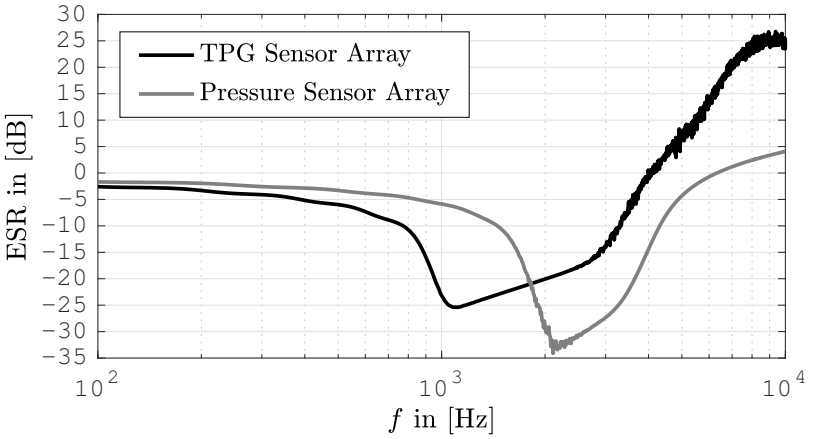


FIGURE 4.8: Plots of the average ESR of the recovered coefficients $\tilde{q}_{n,R}$ from a simulated TPG sensor array measurement (black line) and from a simulated pressure sensor array measurement (light grey line) within the frequency range from 100 Hz to 10 kHz.

As suggested by the findings in subsections 4.3.1 and 4.3.4, the error due to spatial aliasing at high frequencies is higher with the TPG sensor array than with the pressure sensor array. Above $f = 1.8\,\text{kHz}$ the conventional array clearly outperforms the TPG array with the ESR reaching values below $-30\,\text{dB}$, yet it should be mentioned that the conventional array with $L_P = 16 > 2N + 1$ sensors is more robust to aliasing from higher orders than a conventional array with $L_P = 15 = 2N + 1$ sensors.

At lower frequencies, the performance of the TPG sensor array is theoretically significantly better than that of the conventional array. This result needs to be seen with caution as gradient sensors are known to suffer from high self-noise at low frequencies due to comparably low SNR caused by the small gradient magnitude.

Judging by the width of the respective frequency band in which the two arrays' ESR falls below $-15\,\text{dB}$, both arrays achieve a band that is around 2.3 kHz wide. However, w.r.t. the lowest ESR value, the conventional array outperforms the TPG sensor array by around 6.5 dB.

It is worth noticing that the optimum band of the TPG sensor array is shifted towards lower frequencies. This can be achieved with a conventional array by increasing its radius, which shifts the curves of the eigenvalues in Fig. 4.2 and Fig. 4.3 towards lower frequencies due to the change of the argument in the radial functions. However, comparing a TPG sensor array with a conventional array of the same radius, it is clear that the former performs better at low frequencies. Therefore, considering audio applications, the two the arrays combined on one structure would extend the usable bandwidth of the TPG sensor array by 45 % or that of the conventional array by 33 %, respectively. This would gain a substantially increased accuracy within a frequency band where the human ear is very sensitive (compare to the curves of equivalent loudness in ISO 226) and which is very important for localisation (due to interaural time difference and spectral cues from head and torso reflections [Bla97]).

The following simulation example visualises the effect of the improvement over conventional designs at low frequencies.

4.4.5 An Example

With $L_P = 16$ pressure sensors and $L_G = 15$ TPG sensors plus the additional pressure sensor, the respective arrays allow for the recovery of the coefficients $q_n, |n| \leq 7$. From Fig. 4.2 and Fig. 4.3 it can be seen that the gain applied to the array observation for the recovery of q_{-7} and q_7 at a frequency of $f = 800\,\mathrm{Hz}$ is about 35 dB for the TPG array and about 70 dB for the conventional array. With a given Signal-to-Noise-Ratio of 60 dB, it can be expected that the noise is going to affect the measurement of the conventional array considerably more than that of the TPG array.

Fig. 4.9 shows the reconstruction of a plane wave approximation ($\phi_i = \frac{\pi}{7}, f = 800\,\mathrm{Hz}, c = 343\,\frac{\mathrm{m}}{\mathrm{s}}, N_s = 15$) on the basis of the coefficients $\tilde{\mathbf{q}}$ recovered from a TPG sensor array and from a pressure sensor array for the cases of $\beta = \sqrt{10^{\frac{-50}{20}}}$ and $\beta = 0$. It shows that the synthesis based on the measurement of the TPG array with and without regularisation is only marginally different from the original sound field. With the pressure array, regularisation leads to a suppression of higher orders, resulting in a reduced size of the region of accurate reconstruction [WA01, Ahr12] and without regularisation, the noise leads to a spatially distorted synthesis outside the region where 6th order contributions become significant. This result supports the results in subsection 4.4.4 that TPG sensor arrays are more robust against transducer noise than conventional arrays.

4.5 Conclusions

The performance of a circular microphone array composed of TPG sensors has been investigated by means of theoretical analysis and numerical simulations. The integral operator describing the sound field model has been introduced and decomposed by means of the eigenvalue decomposition. This allows for an entirely separate analysis of transducer noise induced errors and aliasing effects, since both can be associated with different components of the EVD.

It has been shown that, at low frequencies, a TPG sensor array is more robust against transducer noise than conventional arrays, while, at high frequencies, it is significantly more vulnerable to spatial aliasing. In the simulations presented in this work, the lowest achievable ESR of the TPG sensor array is surpassed by that of a comparable conventional array by approximately 6.5 dB. Furthermore, it was found that the usable optimum frequency band is not extended by the new design, but it is shifted towards lower frequencies. Finally, even though it has been found that this new design can potentially gain on former designs in terms of transducer noise, this is only true under the condition that both transducer types are of equal quality. The costs of a gradient sensor that matches the quality (w.r.t. noise, symmetry of its directivity, etc.) of a state-of-the-art pressure sensor are however expected to be considerably higher. It is therefore possible that what the TPG sensor array gains in robustness at low frequencies is not enough to compensate for the generally higher transducer noise.

It has also been shown that for the recovery of the 0th mode, at least one pressure sensor needs to be added to the array. This leads to a change in the spatial aliasing scheme of the TPG array, making the 0th mode particularly vulnerable to aliasing.

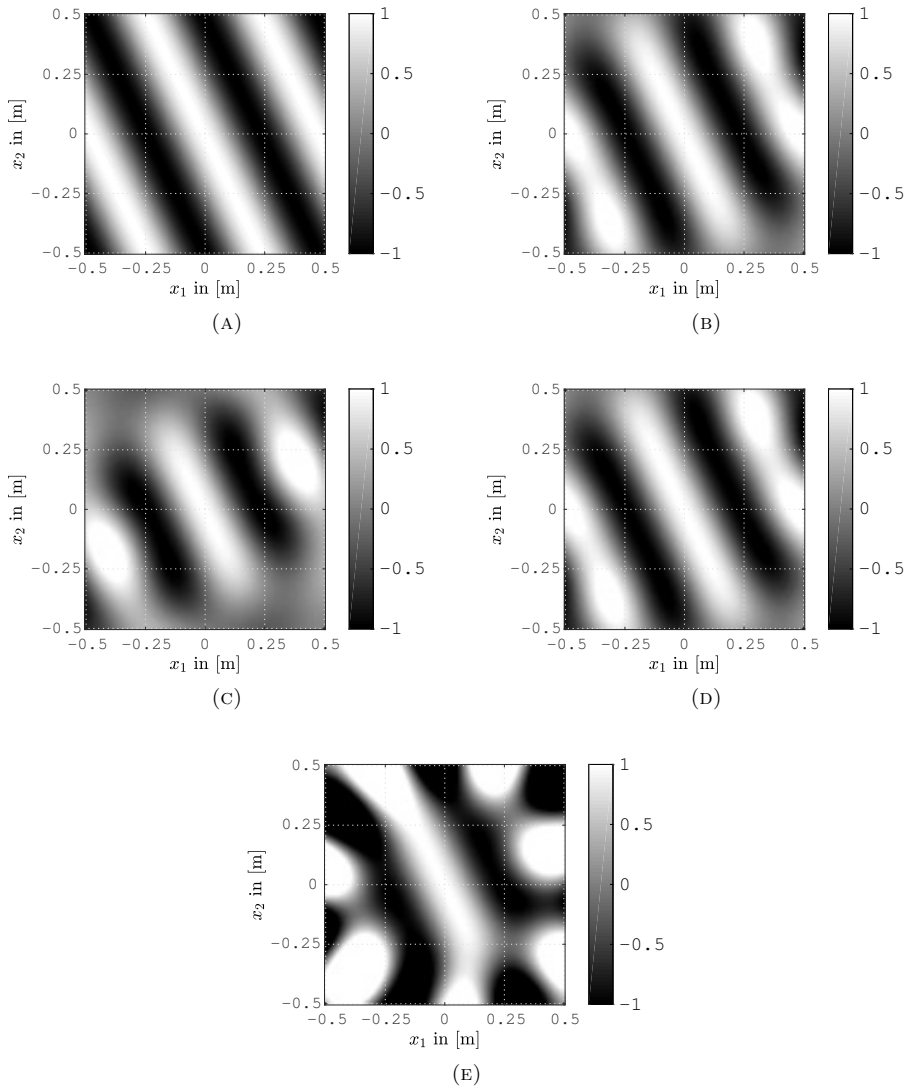


FIGURE 4.9: Visualisation of a plane wave approximation with $\phi_i = \frac{\pi}{7}, f = 800\,\text{Hz}, c = 343\,\frac{\text{m}}{\text{s}}$, limited to order $N_s = 15$, (A) original sound field, (B) synthesis based on TPG array measurement with $\beta = \sqrt{10^{\frac{-50}{20}}}$, (C) synthesis based on pressure array measurement with $\beta = \sqrt{10^{\frac{-50}{20}}}$, (D) synthesis based on TPG array measurement with $\beta = 0$, (E) synthesis based on pressure array measurement with $\beta = 0$.

Chapter 5

DOA Estimation Performance with Circular Arrays in Sound Fields with Finite Rate of Innovation

Falk-Martin Hoffmann, Philip Arthur Nelson and Filippo Maria Fazi

Abstract

A plane wave identification method based on a *Finite Rate of Innovation* (FRI) model for the wave field has been recently proposed for circular microphone arrays [HFN16] (see Appendix F). This work presents a more detailed analysis of the method that was originally conceived for the identification of distinct pulses in a time domain signal. A general model is presented for sound fields with a finite rate of innovation and the data that can be measured with circular arrays in such fields. Using different theoretical acoustic source models, an investigation is presented of the estimation accuracy and robustness of the Direction-of-Arrival estimation method against noisy data and deviations from the model. The performance is assessed on the basis of the analysis of results obtained from both simulations and experimental data.

5.1 Introduction

Direction-of-Arrival (DOA) estimation with sensor arrays is a problem of continued relevance, independent of the kind of wave field. Whether the objective is underwater source localisation in a submarine or the optimisation of the beam pattern of a controllable source or receiver (e.g. wireless LAN router, teleconference systems with microphone arrays, etc.), a high accuracy DOA estimate for a given wave field has become a key element in many transmission systems.

The associated information can be used as a parameter for a beamformer that optimises the *Signal-to-Noise-Ratio* (SNR) in the transmission system.

In applications where the *exact* source position (i.e. direction *and* distance) does not provide any information that is useful to the system beyond the DOA alone, one may want to make the additional assumption that the receiver is in the far field of the source. If that is indeed the case, it is common to model the wave field around the receiver as a number of plane waves [RK89, Raf04, TK06, ZDG10, BB10a]. A *Plane Wave* (PW) is fully described by its complex amplitude and DOA, yielding four parameters in 3D space. That makes a wave field composed of a finite number of PWs eligible for a sparse representation [CW08, WEMJ11]. Sparse wave fields are particularly suitable for DOA estimation problems as they do not require an infinite resolution of the estimator.

The DOA estimator presented in [HFN16] (Appendix F) falls into the field of parametric approaches. Parametric DOA estimation methods rely on a sound field model. Prominent examples of well established parametric methods are MUSIC [Sch86] and ESPRIT [RK89]. Both provide a high-resolution DOA estimation performance yet only for a sound field generated by narrowband sources. To overcome this limitation and enable DOA estimation with wideband sources, Teutsch and Kellermann derived the EB-ESPRIT algorithm[TK06]. This evolution of the ESPRIT algorithm takes the latter's subspace nature and instead of applying it to the sensor signals directly, it is applied to the modal domain of a circular microphone array. Using the EB-ESPRIT algorithm, Teutsch and Kellermann reported estimates with less than three degrees error.

In previous work [HFN16] (Appendix F), the authors found that a sparse sound field with a finite number of plane waves travelling in the same plane can be described through a periodic function with a *finite rate of innovation* [VMB02, BDV$^+$08]. It was shown that the sampling method for signals consisting of a finite number of Dirac pulses proposed by Vetterli et al. [VMB02] can be directly transferred to solve the problem of DOA estimation with circular arrays. The work presented in [HFN16] was based on a perfect observation of an ideal sound field and the proposed method requires further investigation to assess its realistic performance. Blu et al. [BDV$^+$08] published further work on FRI signal analysis that addressed, among other things, how to improve robustness against noise and model imperfections.

This work investigates the robustness of the DOA estimation method based on FRI signal analysis for sound fields with plane waves travelling in the same plane. The method is tested against noise, sound field imperfections (i.e. point sources and plane waves that travel in a different plane) and properties of the circular measurement array. The remainder of this work is organised in six further sections. The following section introduces the sound field model. Section Three derives expressions for the data gained from a circular microphone array while observing the different source types. Section Four recapitulates the robust FRI signal analysis methods proposed by Blu et al. [BDV$^+$08] and presents a noise model based on the nature of the sensor array observation. A simulation study of the DOA estimation performance and its results with respect to the estimation error are described and discussed in Section Five. The results of an experimental study are presented and discussed in the sixth section.

5.2 Sound Field Model

This work uses predominantly the cylindrical coordinate system, but sometimes also the elevation angle θ. These and a basic sketch of the geometric entities relevant for this work are depicted in Figure 5.1. All considerations pertain to the frequency domain and assume a (quasi-)stationary

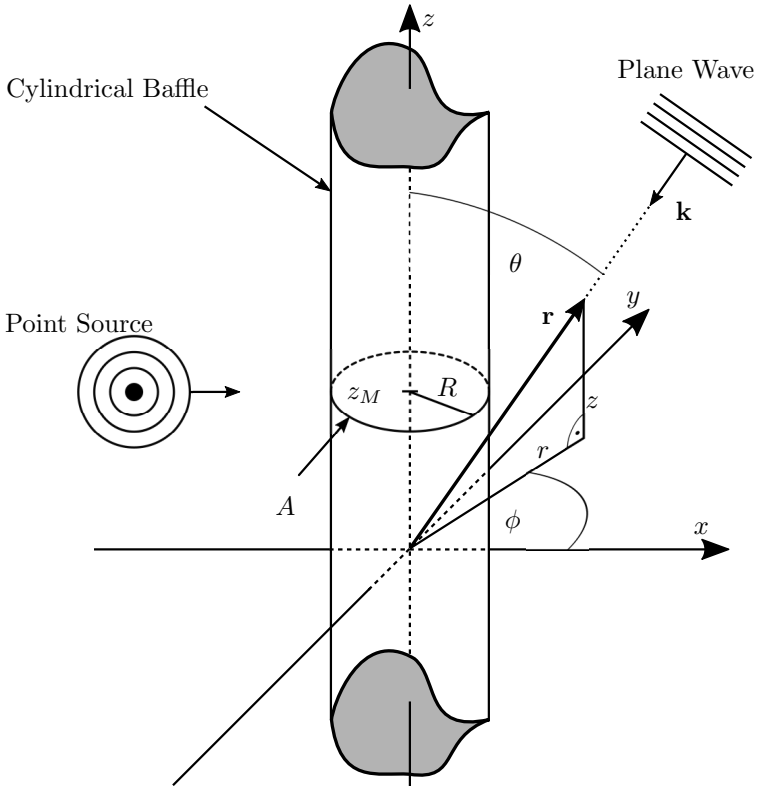


FIGURE 5.1: Cylindrical coordinate system with $\mathbf{r} = (r, \phi, z)^T$, the elevation angle θ and other entities as they are used in this work.

field. The factor $e^{-i\omega t}$ and the various quantities' explicit dependency on the angular frequency ω are omitted for reasons of brevity.

It is furthermore assumed that the sound field within a bounded area under consideration, V, is in the far field of isolated point sources and satisfies the homogeneous wave equation, so that the pressure is fully defined through a bounded *Herglotz Density* (HD) [Faz10, FNW12]. An expression for the sound field of a number of plane waves is derived hereunder.

5.2.1 Plane Wave Expansion in Cylindrical Coordinates

A single PW of unitary magnitude and zero-phase is defined by its wave number vector $\mathbf{k}$ (see Figure 5.1). In cylindrical coordinates, the latter can be written in the form

$$\mathbf{k} = -\begin{bmatrix} k_r \cos\varphi \\ k_r \sin\varphi \\ k_z \end{bmatrix}, \tag{5.1}$$

where φ is the azimuth angle associated with the DOA of the PW. The l_2-norm of $\mathbf{k}$ provides the wave number

$$k = \|\mathbf{k}\|_2 = \sqrt{k_r^2 + k_z^2} = \frac{\omega}{c} \tag{5.2}$$

at the angular frequency ω and the speed of sound c. The series expansion of a single PW in cylinder basis functions is given through (see Appendix 5.8)

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik\cos\theta z} \sum_{n=-\infty}^{\infty} i^{-n} J_n(kr\sin\theta) e^{in(\phi-\varphi)}. \tag{5.3}$$

The above expression represents the sound field of a PW through the product of angular and axial modes, where only one of the latter is required for a single PW. It represents the kernel for the *Herglotz Wave Function* (HWF) [CK98] that expresses a sound field as the superposition of PWs.

5.2.2 The Herglotz Wave Function

The PWs forming the sound field can come from any direction associated with the points on the sphere

$$\Omega := \{\mathbf{r} : \|\mathbf{r}\|_2 = 1\} \tag{5.4}$$

around the origin. Each PW has an individual complex amplitude $q(\mathbf{k})$, $\mathbf{k} \in \Omega$. The sound field is then given through the HWF [CK98]

$$p(r, \phi, z) = \int_{\Omega} e^{i\mathbf{kr}} q(\mathbf{k}) d\Omega(\mathbf{k}), \tag{5.5}$$

where $q(\mathbf{k})$ is known as the aforementioned HD. The HWF poses an expression for any sound field that satisfies the assumptions made at the beginning of this section.

5.2.3 Sound Field of L Plane Waves

Let the sound field consist of a finite set of L distinct far-field sources, i.e. PWs. Each PW is specified through a complex amplitude b_l, and a DOA (φ_l, θ_l). The integral in (5.5) can then be replaced by a sum, yielding the expression for the pressure

$$p(r, \phi, z) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{L} b_l R_n(k_r^{(l)} r) e^{-ik_z^{(l)} z} \Phi_n^*(\varphi_l) \Phi_n(\phi), \tag{5.6}$$

where the notation

$$k_r^{(l)} = k\sin\theta_l \tag{5.7}$$

$$k_z^{(l)} = k\cos\theta_l \tag{5.8}$$

$$R_n(x) = 2\pi i^{-n} J_n(x) \tag{5.9}$$

and

$$\Phi_n(\phi) = \frac{e^{in\phi}}{\sqrt{2\pi}} \tag{5.10}$$

was introduced for reasons of brevity.

Equation(5.6) is used to describe the measurement signals for a circular microphone array. In order to test the robustness of the FRI method against non-plane wave components in the sound field, the expression in Eq. (5.6) can be complemented by adding the series expansion for the pressure field of a *Point Source* (PS) (see Appendix 5.9).

5.3 Circular Microphone Array Measurement

Using the developed model, it is now possible to provide expressions for the signals obtained from circular microphone array measurements.

5.3.1 General PW Sound Field Measurement

Let a circular microphone array sense the pressure of the wave field continuously in the plane $z = z_M$ on the circle (see Figure 5.1)

$$A := \{\mathbf{r} : r = R, \phi \in [0, 2\pi], z = z_M\}. \tag{5.11}$$

The pressure measured on A can be expressed as a Fourier Series [HF15, HFN16]

$$p(R, \phi, z_M) = \sum_{n=-\infty}^{\infty} p_n \Phi_n(\phi). \tag{5.12}$$

Multiplying both sides of (5.6) with $\Phi^*_{n'}(\phi)$ and integrating over A w.r.t. ϕ yields

$$p_n = \sum_{l=1}^{L} b_l R_n(k_r^{(l)} R) e^{-ik_z^{(l)} z_M} \Phi^*_n(\varphi_l),\ n' = n, \tag{5.13}$$

The definition of the factors $R_n(k_r^{(l)} r)$ in (5.9) differs for a microphone array with a rigid scattering corpus [Wil99, TK06, HFN16]. For an infinite cylindrical scatterer of radius R (see Figure 5.1), these radial functions become a combination of Bessel functions, $J_n(\cdot)$, Hankel functions of the first kind, $H_n(\cdot)$, and their derivatives:

$$R_n(k_r^{(l)} r) = \frac{2\pi}{i^n} \left(J_n(k_r^{(l)} r) - \frac{J'_n(k_r^{(l)} R)}{H'_n(k_r^{(l)} R)} H_n(k_r^{(l)} r) \right). \tag{5.14}$$

For a more thorough analysis of the radial functions of circular microphone arrays, the reader is referred to the corresponding literature, e.g. [TK06, HF15].

5.3.2 HD Coefficients of PWs travelling parallel to $z = z_M$

In the case where all L PWs are travelling parallel to the plane $z = z_M$, i.e. $\{\theta_l\}_{l=1}^{L} = \frac{\pi}{2}$, the expression in (5.13) can be further simplified to $(n = n')$

$$p_n = R_n(kR) \sum_{l=1}^{L} b_l \Phi_n^*(\varphi_l) = R_n(kR) q_n. \tag{5.15}$$

The right hand side of the above equation indicates that the expression for p_n can be separated into a radial function, $R_n(kR)$, and a component describing the plane waves

$$\sum_{l=1}^{L} b_l \frac{e^{-in\varphi_l}}{\sqrt{2\pi}} = q_n. \tag{5.16}$$

These HD coefficients, q_n, fully define the sound field through the corresponding HD

$$q(\phi) = \sum_{n=-\infty}^{\infty} q_n \frac{e^{in\phi}}{\sqrt{2\pi}} = \sum_{l=1}^{L} b_l \delta(\phi - \varphi_l). \tag{5.17}$$

Note that there are no requirements as to the complex amplitudes b_l, which implies that the sources can even radiate coherently. This poses a significant advantage of the proposed DOA estimation method over established methods such as MUSIC [Sch86] that are based on the analysis of the autocorrelation matrix of the microphone signals and require the source signals to be uncorrelated. An example of a HD with $L = 5$ distinct plane waves of unit magnitude is shown in Figure 5.2 and it can be seen that $q(\phi)$ must be periodic with a period of 2π. The HD

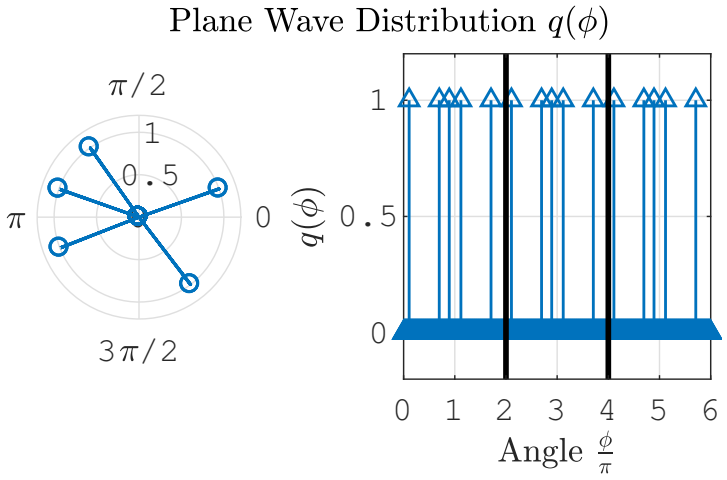


FIGURE 5.2: HD for $L = 5$ distinct plane waves, where the HD is periodic with $q(\phi) = q(\phi + u2\pi), \forall u \in \mathbb{Z}$.

is fully defined by the set of tupels $\{(\varphi_l, b_l)\}_{l=1}^{L}$. The extraction of the q_n from the array signals is crucial for the application of the FRI DOA estimation method.

At a later point it will be shown that a wave field specified by the HD coefficients in (5.16) is ideal for the application of the FRI DOA estimation method with circular arrays. The next two subsections present expressions for the HD coefficients of two source types that are likely to

corrupt the performance of the DOA estimation method in combination with the measurement aperture under consideration.

5.3.3 HD Coefficients of a Single PW with $\theta_l \neq \frac{\pi}{2}$

For a single plane wave impinging on the measurement aperture at an angle $\theta_l \neq \frac{\pi}{2}$, the corresponding pressure coefficients are given through (5.13), yielding

$$p_n = b_l R_n(k_r^{(l)} R) e^{-ik_z^{(l)} z_M} \Phi_n^*(\varphi_l) \tag{5.18}$$

A factor separation as shown in Equation (5.15) of the result in (5.18) yields HD coefficients of the form

$$q_n = b_l \underbrace{\frac{R_n(k_r^{(l)} R)}{R_n(kR)} e^{-ik_z^{(l)} z_M}}_{\Psi_n(k, k_r^{(l)}, R)} \frac{e^{-in\varphi_l}}{\sqrt{2\pi}} \tag{5.19}$$

for the sound field of the single PW with $\theta_l \neq \frac{\pi}{2}$. The main difference to Eq. (5.16) is the additional factor $\Psi_n(k, k_r^{(l)}, R)$.

5.3.4 HD Coefficients of a Single Point Source

The derivation of the HD coefficients for the sound field of a single PS

$$q_n = b' \underbrace{\frac{i^{n+1}}{8\pi} \frac{\int_{-\infty}^{\infty} R_n(k_r R) H_n^{(1)}(k_r r') dk_z}{R_n(kR)}}_{\zeta_n(r')} \frac{e^{-in\phi'}}{\sqrt{2\pi}} \tag{5.20}$$

is given in Appendix 5.10. Similar to the source type discussed in the previous subsection, the key difference between Eqs. (5.16) and (5.20) is an additional factor, namely $\zeta_n(r')$. For both types of sources, these additional factors are later shown to negatively affect a DOA estimation using the FRI method in combination with circular measurement apertures.

5.3.5 Practical Limitations of Microphone Arrays

In practice, microphone arrays can only sample the pressure on A and therefore cannot recover all coefficients q_n needed to fully reconstruct the HD. A uniformly distributed circular microphone array with M sensors can recover the coefficients q_n for $|n| \leq N$, where $N = \lfloor \frac{M-1}{2} \rfloor$, if no angular modes with $n > N$ contribute to the pressure field on A (i.e. no spatial aliasing). This is the case (to a good approximation) for all frequencies that satisfy $N \geq kR$ [Wil99, AR14, HF15]. At higher frequencies, the recovered coefficients q_n are corrupted by spatial aliasing [AR14, HF15]. It shall for now be assumed that the coefficients $q_n, |n| \leq N$ can be recovered perfectly from the array measurement, and that the frequency range in questions ensures that no spatial aliasing occurs.

5.4 Finite Rate of Innovation Analysis

The form of the HD given in (5.17) represents a set of distinct Dirac-pulses, for which the FRI theory was originally developed [VMB02, BDV+08]. The key difference between previous work and that presented here is that the observed signal is not in the time domain but in the spatial domain.

5.4.1 The System Inherent Sampling Kernel

While the method proposed in [VMB02] requires a sampling kernel (i.e. a lowpass filter) for the observed time domain signal, that is not always required when applied to the sampled pressure in the spatial domain. This stems from the fact that the HD cannot be sampled directly but needs to be extracted from the sampling of the pressure.

It is known that modes of orders higher than a frequency dependent threshold N_A hardly contribute to the pressure observed on the measurement aperture A. Their corresponding radial functions $R_n(kr)$ act as a system inherent lowpass filter. Therefore, at least for frequencies at which the array does not suffer from spatial aliasing, the coefficients $\{q_n\}_{n=-N}^{N}$ can be recovered practically uncompromised, so that a lowpass filtered version of the HD can be reconstructed from them.

5.4.2 The Rate of Innovation of the HD

The coefficients q_n of the form given by (5.16) describe a (periodic) HD with finite rate of innovation. Every PW is described through its DOA, φ_l, and its complex amplitude, b_l. In other words, every PW is an entity with two 'innovations' for the HD. It follows that a HD with L PWs incorporates $2L$ innovations. Due to its 2π periodicity, $q(\phi)$ has the *rate of innovation*

$$\rho = \frac{2L}{2\pi} = \frac{L}{\pi}$$

that is *finite* for $L < \infty$ (see also Fig. 5.2). The value of ρ or rather the number of innovations within the period, $2L$, determines how many consecutive coefficients q_n are needed to apply the annihilating filter method [VMB02, BDV+08]. The exact relationship is derived in the following subsection.

5.4.3 Recovering the DOA Information

The exact procedure for the DOA estimation with ideal coefficients q_n is described in [HFN16]. In practice, however, measured data may be compromised due to imperfections of the measurement setup and noise. To make the estimator more robust against such, Blu et al. [BDV+08] proposed the use of a 'total least-squares approach'. Their method is described hereunder.

The global objective of the DOA estimator mechanism is to find an annihilating filter $\{a_k\}_{k=0}^{K}$, $K \leq N$, that satisfies

$$q_n * a_k = 0, \tag{5.21}$$

where $*$ denotes the discrete convolution

$$\sum_{k=0}^{K} a_k q_{n-k} = 0,\ (-N+K) \leq n \leq N. \tag{5.22}$$

The above equation defines a system of $2N+1-K$ equations that can be expressed in the algebraic form

$$\mathbf{Q}\mathbf{a} = 0, \tag{5.23}$$

with the Toeplitz matrix

$$\mathbf{Q} = \underbrace{\begin{bmatrix} q_{-N+K} & q_{-N+K+1} & \cdots & q_{-N} \\ q_{-N+K+1} & q_{-N+K} & \cdots & q_{-N+1} \\ \vdots & \ddots & \ddots & \vdots \\ q_{-N+2K} & \ddots & q_{-N+K+1} & q_{-N+K} \\ \vdots & \ddots & \ddots & \vdots \\ q_{N} & q_{N-1} & \cdots & q_{N-K} \end{bmatrix}}_{(2N+1-K)\times(K+1)} \tag{5.24}$$

and $\mathbf{a} = [a_0, a_1, \ldots, a_K]^T$. The vector $\mathbf{a}$ denotes an annihilating filter that suppresses the exponential components in the coefficient matrix $\mathbf{Q}$ that are associated with the individual plane waves in the field. For an ideal sound field with no measurement noise, the system in Eq. (5.23) is overdetermined for $N > K$ and the rank of $\mathbf{Q}$ cannot exceed L. This is due to the nature of the coefficients q_n. They can only lead to $K - L + 1$ independent, non-trivial solutions of the annihilating filter $\mathbf{a}$, one of which has zeros for the last $K - L$ entries [BDV$^+$08]. It follows that

$$\text{rank}(\mathbf{Q}) = L, \tag{5.25}$$

i.e. the number of PWs in the sound field can be identified from the rank of the Toeplitz matrix $\mathbf{Q}$.

When the coefficients q_n are corrupted by noise, there may not be a non-trivial solution that exactly satisfies Eqs. (5.21) and (5.23). As an alternative, Blu et al. [BDV$^+$08] proposed to identify a *Total Least-Squares* (TLS) estimate for the annihilating filter $\mathbf{a}$ by solving the optimisation problem

$$\tilde{\mathbf{a}} = \underset{\mathbf{a}}{\text{argmin}} \|\mathbf{Q}\mathbf{a}\|_2,\ \|\mathbf{a}\|_2 = 1. \tag{5.26}$$

With the constraint for $\mathbf{a}$, the solution to the above problem can be solved through a *Singular Value Decomposition* (SVD) of $\mathbf{Q} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. This provides the right-singular vectors $\{\mathbf{v}_k\}_{k=0}^{L}$ (i.e. the columns of $\mathbf{V}$) that all satisfy $\|\mathbf{v}_k\|_2 = 1$. The right-singular vector $\mathbf{v}_{\min}$ corresponding to the smallest non-zero singular value $\sigma_{\min}$ then provides a solution to the total least-squares problem that yields

$$\mathbf{Q}\tilde{\mathbf{a}} \approx 0,\ \text{with}\ \|\tilde{\mathbf{a}}\|_2 = 1, \tag{5.27}$$

where $\tilde{\mathbf{a}} = \mathbf{v}_{\min}$. Unlike the annihilating filter presented in [HFN16], that obtained from the TLS method can have a leading factor ψ [BDV+08], yielding

$$\tilde{a}_n = \psi(\delta_n + \sum_{k=1}^{K} \alpha_k \delta_{k-n})$$

with the z-transform

$$\tilde{A}(z) = \psi \prod_{k=1}^{K} (1 - u_k z^{-1}). \tag{5.28}$$

The roots $\{u_k\}_{k=1}^{K}$ of the characteristic polynomial of $\tilde{A}(z)$ contain information on the DOA [HFN16]. For a given polynomial $\tilde{A}(z)$, the roots can be calculated through numerical factorisation, e.g. using the integrated *MATLAB*-function *roots*. From [HFN16] it is known that these roots are of the form $u_k = e^{-i\varphi_k}$, so that the DOA can be calculated through

$$\varphi_k = i \ln(u_k), \forall k = 1 \dots K, \tag{5.29}$$

where $\ln(\cdot)$ denotes the principle value of the complex logarithm. The full process for the DOA estimation is summarised below.

DOA Estimation Procedure

1. Choose a value $K \leq N$, if available $K = L$,
2. construct $\mathbf{Q}$ according to Eq. (5.24),
3. solve the TLS problem in (5.26) to obtain an estimate $\tilde{\mathbf{a}}$,
4. calculate the roots, $\{u_k\}_{k=1}^{K}$, of the z-transform of $\tilde{\mathbf{a}}$,
5. convert the K roots into DOA estimates using Eq. (5.29),
6. use $\{\tilde{\varphi}_l\}_{l=1}^{N}$ to calculate the complex amplitudes, $\{\tilde{b}_k\}_{k=1}^{K}$, by solving the linear equation system defined (5.16) for a set of coefficients q_n.

With respect to this work, the described procedure is the most general approach to the FRI DOA estimation method, yielding estimates for exactly K PWs, even if $L < K$ or $L > N$. This is clearly not ideal and requires further investigation.

For particularly noisy data, additional Cadzow-Denoising [BDV+08] can be applied before the DOA estimation to fit the data closer to the model. This was done for all applications of the DOA estimation method presented in this work, where the steps of the Cadzow-Denoising method were iterated 20 times.

5.4.4 A Measurement Noise Model

The applied noise model is a spatial transform of the transducer noise model for microphone arrays used in [HF15]. A pressure measurement on a circular aperture around the origin would lead to a noisy observation

$$\tilde{p}(\phi) = p(\phi) + d(\phi),$$

where d denotes the measurement noise. Following the same steps as those that link Eqs. (5.12) and (5.13) yields

$$\tilde{p}_n = p_n + \sigma_\epsilon \frac{e^{i2\pi\mu}}{\sqrt{2\pi}}. \tag{5.30}$$

Dividing by the radial functions $R_n(kR)$ provides the result for the HD coefficients as they can be recovered from a circular microphone array

$$\tilde{q}_n = q_n + \epsilon_n, \ \forall |n| \leq N. \tag{5.31}$$

Each of the ideal coefficients q_n can be corrupted with the additive component

$$\epsilon_n = \frac{\sigma_\epsilon}{R_n(kR)} \frac{e^{i2\pi\mu}}{\sqrt{2\pi}} \tag{5.32}$$

that reflects the statistical properties of the measurement noise in microphone arrays. σ_ϵ denotes the standard deviation of the noise in the observed pressure signal and μ is a random variable that is *uniformly* distributed on the interval $[0, 1]$.

The expected value for the magnitude of ϵ_n, $E\{|\epsilon_n|\}$, can be useful to estimate the number of PWs for the case that $L < N$. The model assumes that the noise distribution in the pressure domain is Gaussian, therefore the expected value of the magnitude of the measurement noise in the pressure domain, $|d(\phi)|$, must be

$$E\{|d(\phi)|\} = \frac{\sigma_\epsilon}{\sqrt{2\pi}}. \tag{5.33}$$

To obtain a useful expected value for the measurement noise in the HD coefficient domain, it is necessary to consider the influence of the radial functions $R_n(kR)$. From the literature on microphone arrays, it is known that radial functions as given by (5.14) for $k_r^{(l)} = k$ and $|n| \leq N$ have very different magnitudes for frequencies higher than

$$f_c = \frac{Nc}{2\pi R}$$

[Wil99, TK06, HF15]. This is rather unfavourable, since this property of the radial functions would be inflicted on the coefficients ϵ_n as a consequence of Eq. (5.32). The components of q_n that are associated with the individual PWs are constant in magnitude across the range $|n| \leq N$. Hence it would be difficult to define a clear 'noise floor' if the expected value for the noise magnitude is a function of n. However, at frequencies higher than f_c, the magnitude of the radial functions can be approximated by $|R_0(kR)|$, which can be seen in Figure 5.3. It follows that the expected value for the magnitude of ϵ_n can be approximated by

$$E\{|\epsilon_n|\} \approx \frac{\sigma_\epsilon}{\sqrt{2\pi}|R_0(kR)|} \tag{5.34}$$

for frequencies higher than $f_c \approx 5.1\,\text{kHz}$ with $R = 0.075\,\text{m}$. This value provides an estimate of the noise floor.

To acquire a value of σ_ϵ for a specific array, which may also be frequency dependent [HF15], one could simply perform a measurement without active sources in an anechoic environment.

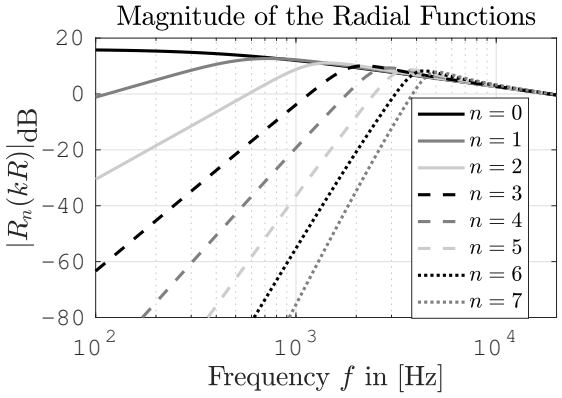


FIGURE 5.3: Radial functions for a microphone array with radius $R = 0.075\,\text{m}$ and $f_c \approx 5.1\,\text{kHz}$ ($N = 7$, $c = 343\,\frac{\text{m}}{\text{s}}$).

5.4.5 Choosing K with Noisy Coefficients q_n

If the number of PWs is known a priori, then the value for K should be simply chosen to be L. Otherwise it can always be chosen to be the maximum number of DOAs that can be estimated from the available array data, i.e. $K = N$. Three cases can then be distinguished for a spatially stationary PW distribution when the individual PWs' energy exceeds that of the background noise:

1. $\mathbf{L < N}$: The DOAs corresponding to the L largest values of $|\tilde{b}_k|$ would be consistent irrespective of the sample in time, where the $N - L$ smallest estimates must be expected to vary with every estimation attempt due to noise.
2. $\mathbf{L = N}$: All DOA estimates describe an individual PW.
3. $\mathbf{L > N}$: The method provides N least-squares DOA estimates.

In the case of $L < N$, the number of PWs, L, can be estimated from the number of magnitudes $\tilde{b}_k$ that exceed the estimate of the noise floor given by Eq. (5.34). This is however only possible for frequencies higher than f_c.

In the case of $L > N$, the rate of innovation is too high for the available range of coefficients $|n| \leq N$, i.e. the sampling of the sound field is insufficient. The estimator can only provide $K = N$ DOA estimates, which will be shown to become increasingly inaccurate as L approaches N. One explanation for this is that the TLS solution for the annihilating filter 'groups' the $L - N$ surplus PWs together with the closest neighbours within the set of the remaining L PWs. This is more thoroughly investigated in Section 5.5.1.

5.4.6 Interpretation of the Annihilating Filter

The structure of the annihilating filter given in Eq. (5.28) describes a system that cancels out distinct 'spatial frequencies' in the z-domain of the sequence of HD coefficients. From Eq. (5.16) it follows that the sequence given by the HD coefficients of a sound field with L distinct PWs travelling in parallel to the plane of the measurement aperture have L distinct spatial frequency

components in the z-domain. The annihilating filter precisely cancels these spatial frequencies from the sequence.

In this light, it can now be understood why both a PS and a PW with $\theta \neq \frac{\pi}{2}$ are likely to compromise the performance of the FRI method. The sequences of the HD coefficients given in (5.19) and (5.20) both describe an amplitude modulation of the carrier sequence $e^{-in\phi'}/\sqrt{2\pi}$ with the sequence $\Psi(k, k_r^{(l)}, R)$ and $\zeta_n(r')$, respectively. From the theory of the spectra of AM signals, it follows that the corresponding HD coefficient sequences contain more than one distinct spatial frequency in the z-domain. The difference between the spatial spectrum of the sequence of a PW and that of a PS is conceptually depicted in Figure 5.4. While the PW is responsible

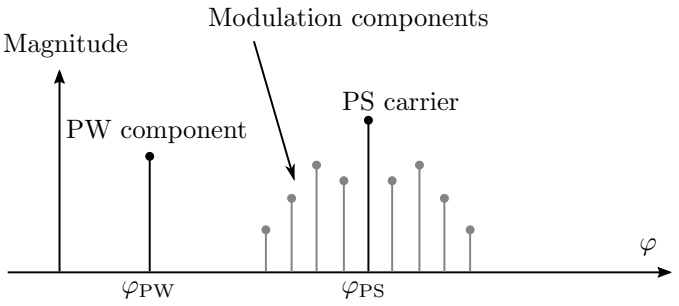


FIGURE 5.4: Conceptual z-domain spectrum of the HD coefficients for the sound field of a PW and a PS.

for just a single spatial frequency component, the PS has a 'carrier' component that refers to the position of the PS **and** modulation components resulting from the term $\zeta_n(r')$ described in Eq. (5.20). The same can be shown for the PW with $\theta \neq \frac{\pi}{2}$, where the term $\Psi_n(k, k_r^{(l)}, R)$ is responsible for the modulation components.

Nevertheless, depending on the modulating sequence, the carrier sequence $e^{-in\phi'}/\sqrt{2\pi}$ may still be the dominant frequency component. In that case, the FRI method may still be able to detect the DOA of these types of sources. The TLS solution would then provide a filter that attempts to suppress the strongest spatial frequency components in the sequence, i.e. that of the individual carriers. The directions of arrival can then be identified as usual from the roots of the annihilating filter's z-transform. However, it must be expected that the estimation accuracy is adversely affected for these types of sources, e.g. if the different sources significantly differ in magnitude or for very close PSs. This is confirmed by means of simulation results given in Section 5.5.

5.5 Simulations

For all simulations discussed hereunder, the DOA estimation method was set to identify $K = N$ plane waves by choosing the matrix $\mathbf{Q}$ accordingly. When the number of plane waves L is known a priori, only those DOA estimates associated with the largest magnitudes $\tilde{b}_l$ were taken into account. For those cases when $L > N$, only N DOA estimates are provided due to the inherent limitation of the method. Note that in this section, a DOA estimate refers to a value for φ_l.

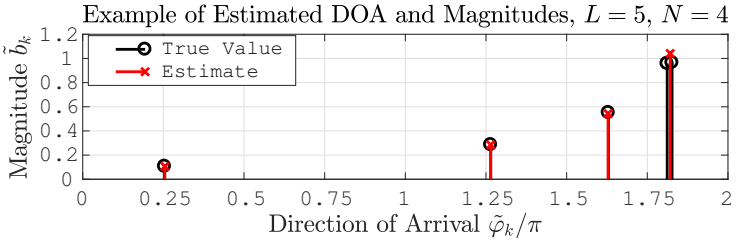
All sound fields were calculated directly in the cylindrical harmonics domain for $f = 1\,\mathrm{kHz}$, using the equations presented in the preceding sections. That way the estimation method's accuracy

and robustness can be studied without the influence of effects that stem from the microphone array (e.g. spatial aliasing, ill-conditioning etc.).

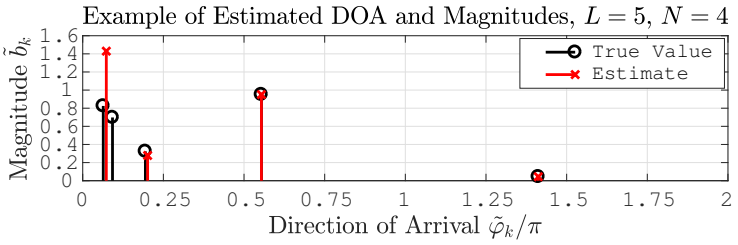
Before looking more closely at the simulation performance for the unrestricted case, $L \leq N$, the estimation behaviour for $L > N$ is investigated by means of a number of selected samples.

5.5.1 Estimation Behaviour for $L > N$

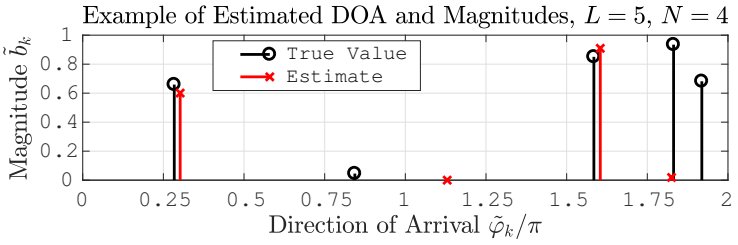
The effects occurring when estimating $K = N$ directions of arrival from the HD coefficients of a sound field with $L > N$ PWs were briefly mentioned in Section 5.4.5. Figure 5.5 shows a selection of estimation results for $L = 5$ and $N = K = 4$.



(A) Two PWs grouped (Example I)



(B) Two PWs grouped (Example II)



(C) Interpolation

FIGURE 5.5: Results of different FRI estimation performances for $L = 5$ and $N = K = 4$.

Figure 5.5a shows an example of the aforementioned grouping behaviour. The estimates of the plane waves with a DOA at around 0.25, 1.25 and 1.625 (on the abscissa) are very accurate, as are the estimated magnitudes. Only for the second estimate on the right it can be noticed that its magnitude is slightly lower than the true value. The estimate beyond 1.75 on the abscissa 'groups' two true PWs with very small difference in their individual DOA around itself. That is likely to explain why its estimated magnitude is larger than that of the two true plane waves.

A similar behaviour of the method can be observed from the result shown in Figure 5.5b. The estimates of magnitude and DOA beyond 0.5 on the abscissa show good accuracy. The leftmost

estimate 'groups' two plane waves, not quite as closely aligned as those in the previous example, and shows a significantly higher magnitude. In this case, the estimate to the right of 0.25 not only shows a magnitude mismatch but also a DOA estimate deviation from the true value.

The result in Figure 5.5c is very different from the others. A grouping behaviour cannot be detected. Only three of the DOA estimates are close to true plane waves, but the magnitude estimate of one of them is significantly low. The fourth estimate is significantly inaccurate.

While there are cases where some estimates may still be of value, it becomes clear that the estimation accuracy decreases when $L > N$. That makes the method unreliable in such cases. For predominantly isotropic sound fields with $L < N$ distinct sources of sufficiently high level, the method may prove more robust, as will be seen in the next subsection.

5.5.2 The DOA Estimation Error (DEE)

As a measure to evaluate the quality of the estimated directions of arrival, $\{\tilde{\varphi}_k\}_{k=1}^{K}$, this work uses the *Direction Estimation Error* (DEE) in percent, which is defined as

$$\text{DEE} = (|\tilde{\varphi}_k - \varphi_k| \mod \pi) \cdot 100\,\%, \tag{5.35}$$

where mod denotes the modulo operator. A DEE of 100 % corresponds to the maximum difference between estimate and true value of $\pi \equiv 180°$.

5.5.3 Mean & Median DEE

In order to assess the estimation accuracy of the FRI DOA estimation method, a simulation study was conducted for different SNR values and different number of PWs. For each combination of SNR and L, 10,000 samples of randomised PW fields were simulated.

The clean HD coefficients were calculated through Eq. (5.16). The complex magnitudes b_l of the individual plane waves were calculated through

$$b_l = (0.01 + 0.99\nu)e^{-i2\pi\mu}, \tag{5.36}$$

where ν, μ are uniformly distributed on the interval $[0, 1]$. The maximum difference between the magnitudes of the individual plane waves is thus 40 dB. The corresponding DOA parameters, φ_l, are uniformly distributed on the interval $[0, 2\pi]$.

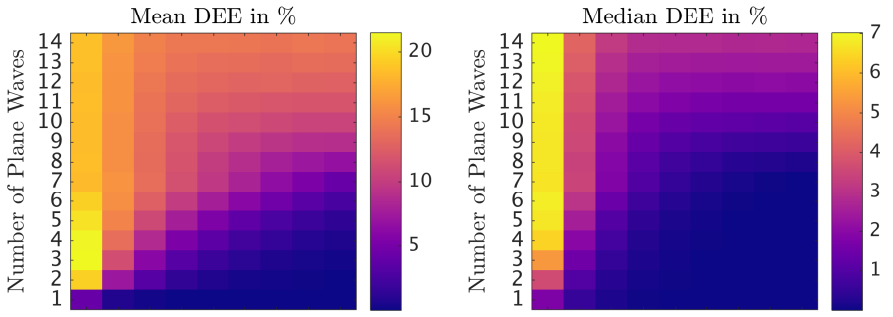
The noise was generated directly in the HD coefficient domain using

$$\epsilon_n = \sigma_\epsilon \xi e^{i2\pi\mu},$$

where ξ is a normally distributed random variable and μ is uniformly distributed on the interval $[0, 1]$. The value σ_ϵ was chosen so that the noisy coefficients $\tilde{q}_n = q_n + \epsilon_n$ have a specified SNR given by

$$\text{SNR} = 10 \log_{10} \left(\frac{\sum_{|n| \leq N} |q_n|^2}{\sum_{|n| \leq N} |\epsilon_n|^2} \right). \tag{5.37}$$

Figure 5.6 shows the results of a simulation study for the mean/median DEE in percent. The



(A) Mean DEE (B) Median DEE

FIGURE 5.6: Mean performance of the FRI estimator over 10,000 samples for different combinations of SNR and L.

algorithm was applied to the noisy coefficients

$$\tilde{q}_n = q_n + \epsilon_n$$

within the range $|n| \leq 7$. This corresponds to the HD coefficient data that can be obtained from a microphone array measurement with $M = 15$ sensors. The maximum number of plane waves whose DOA can be estimated is $K = N = 7$. For the conducted simulations, the algorithm parameter for the number of plane waves was chosen so that $K = \min(L, N)$. Note that with the given dynamic range of the PW magnitudes specified in Eq. (5.36), some of them may be covered by the noise floor.

The result in Fig. 5.6a shows that the mean accuracy of the DOA estimates increases with the SNR and decreases with the number of sources. At 0 dB SNR the values for more than one PW range around a DEE of 20 %, with a maximum at $L = 3$ PWs . At high values of SNR, the DEE increases significantly for more than $L = 7$ plane waves, as a result of the limitation due to the available range of coefficients q_n.

The results for the median DEE given in Fig. 5.6b are similar w.r.t. the overall trend, however the maximum values are significantly lower than those shown in the mean results. This suggests that, for a given combination of SNR and value L, 50 % of the estimates' DEE are lower than the displayed value. Accordingly, the other 50 % of estimates may suffer from 'outliers' that are significantly wrong compared to the true value. These outliers cause the mean to be significantly higher than the median DEE values.

5.5.4 Sound Fields with PWs not travelling parallel to the Measurement Plane

The coefficients describing the sound field of a PW that is not travelling parallel to the measurement plane $z = z_M$ is given in Eq. (5.19). Such a plane wave is henceforth referred to as an *Out-Of-Plane Plane Wave* (OOP-PW). It was postulated in Section 5.4.6 that the effect of the additional factor $\Psi_n(k, k_r^{(l)}, R)$ in the specific form of the coefficients q_n would have a negative effect on the

accuracy of the DOA estimates for more than one PW. It was also argued that the FRI method may provide an adequate estimate of the DOA for a single such plane wave regardless.

5.5.4.1 A Single OOP-PW

This subsection serves to investigate the influence of the elevation angle θ_l of a single OOP-PW on the DOA estimation performance. Figure 5.7 shows the mean DEE for a single OOP-PW for randomly chosen DOA component φ_l. The elevation angle associated with the OOP-PW

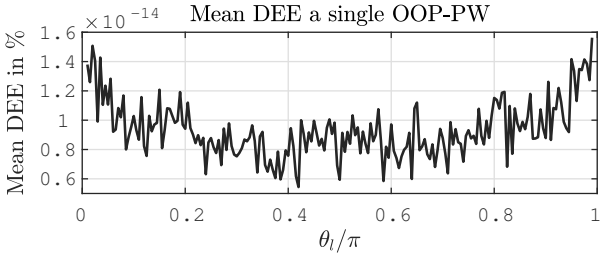


FIGURE 5.7: DEE for a single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen φ_l.

was tested for $0.01\pi \leq \theta_l \leq 0.99\pi$. The DOA estimation would naturally fail for values $\theta_l = 0$ and $\theta_l = \pi$, which were thus not tested. The estimates were obtained from the coefficients $\{q_n\}_{|n|\leq 1}$, which suffice to estimate the DOA of $N = 1$ plane wave. The average result is below 1.6×10^{-14} %.

The same simulation was repeated for a greater range of coefficients, $\{q_n\}_{|n|\leq 5}$, and the result is shown in Figure 5.8. While for some values of θ_l there appears to be no difference, the mean DEE

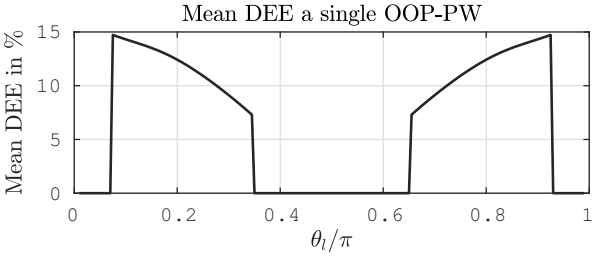


FIGURE 5.8: DEE for a single OOP-PW for $0.01\pi \leq \theta_l \leq 0.99\pi$, averaged over 100 samples with randomly chosen DOA, as obtained with a greater coefficient range.

is significantly higher for angles $0.07\pi \leq \theta_l \leq 0.35\pi$ and $0.65\pi \leq \theta_l \leq 0.93\pi$. This suggests that the estimation corrupting effect of the additional factor $\Psi_n(k, k_r^{(l)}, R)$ increases towards higher values of $|n|$.

5.5.4.2 One PW and one OOP-PW

The DEE performance for two PWs, where one is an OOP-PW, is shown in Figure 5.9. Following the findings in Section 5.5.4.1, the coefficient range was chosen to $|n| \leq 2 = L$ so that both PWs can be identified but the coefficient range is kept as low as possible. Both PWs had the same

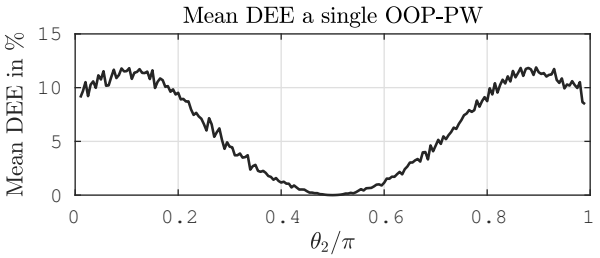


FIGURE 5.9: DEE for one OOP-PW with $\theta_2 \in [0.01\pi \dots 0.99\pi]$ and one with $\theta_1 = \frac{\pi}{2}$, averaged over 100 samples with randomly chosen φ_1 and φ_2.

magnitude. One plane wave is impinging at $\theta_1 = \frac{\pi}{2}$ and the other is simulated at different angles θ_2. The two parameters φ_1 and φ_2 were randomly chosen for every iteration. It can be observed that the accuracy of the DOA estimates decreases up to a DEE of approximately 12 % as the parameter θ_2 deviates from $\frac{\pi}{2}$.

5.5.5 Sound Fields with Point Sources

As with the plane waves discussed in Section 5.5.4, the specific form of the coefficients q_n describing the sound field of a single PS in the plane of the microphone array were predicted to have a negative effect on the accuracy of the DOA estimates associated with the waves impinging from more than one PS. This is now investigated further through a set of simulations.

5.5.5.1 Estimation Accuracy for a Single PS

Figure 5.10 shows the mean DEE for the DOA estimates obtained over 100 iterations of the sound field with a randomly positioned (w.r.t. the angle ϕ') single PS as a function of the position parameter r'. The set of coefficients $\{q_n\}_{|n|\leq 1}$ were calculated from Eq. (5.20), i.e. the point

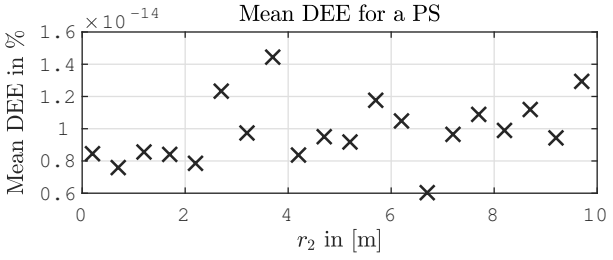


FIGURE 5.10: Mean DEE for a randomly positioned PS as a function of the radius, averaged over 100 samples per evaluation radius.

sources lie in the same plane as the microphones. The simulated array radius was $R = 0.1$ m and the FRI algorithm was set to detect one source only ($K = 1$). It can be seen that a very high estimation accuracy was obtained that supports the reasoning given in Section 5.4.6. Note that the simulation did not include any noise.

The influence of the compromising factor $\zeta_n(r_l)$ in the coefficients q_n was found to increase with the range of coefficients used for the estimation, as it was observed for the OOP-PW in the previous subsection, thus negatively affecting the performance.

5.5.5.2 Estimation Accuracy for Two Point Sources

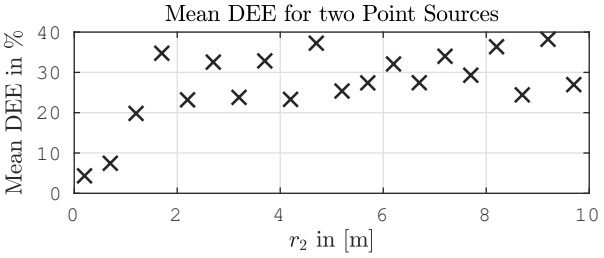
A similar simulation was conducted for two PSs, where the coefficients are given through Eq. (5.20) and superposition as

$$q_n = b_1 \frac{e^{-in\varphi_1}}{\sqrt{2\pi}} \zeta_n(r_1) + b_2 \frac{e^{-in\varphi_2}}{\sqrt{2\pi}} \zeta_n(r_2).$$

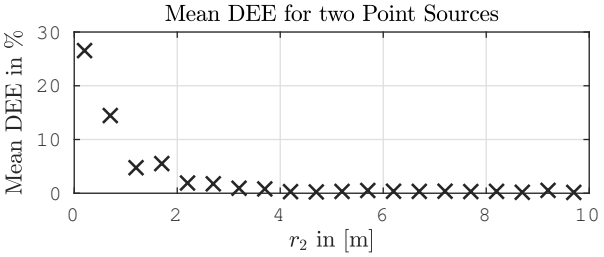
Three different scenarios were investigated using the parameters $b_1 = b_2 = 1$, $r_2 = 0.2 \ldots 9.7\,\mathrm{m}$ and:

1. First source located closely ($r_1 = 0.2\,\mathrm{m}$),
2. first source located distantly ($r_1 = 9.7\,\mathrm{m}$), and
3. both sources at the same distance, i.e. $r_1 = r_2$.

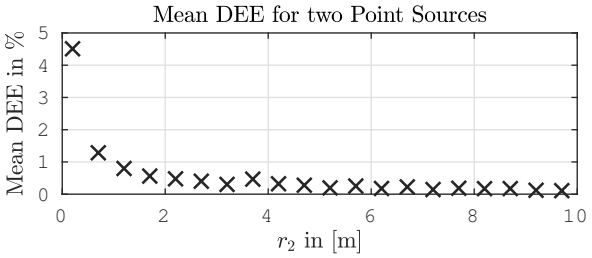
The FRI algorithm was set to detect two sources, i.e. $K = 2$, using the set of coefficients $\{q_n\}_{|n|\leq 2}$. The results are shown in Figure 5.11. It can be observed that the average DEE has drastically



(A) One close source, $r_1 = 0.2\,\mathrm{m}$



(B) One distant source, $r_1 = 9.7\,\mathrm{m}$



(C) $r_1 = r_2$

FIGURE 5.11: Mean DEE for two randomly positioned PSs as a function of r_2, averaged over 100 samples per evaluation radius.

increased just by adding a second PS. The severity however differs between the three scenarios.

In the first scenario, the results in Fig. 5.11a show that for all radii $r_2 > 1\,\mathrm{m}$ the mean DEE is higher than 20 %. When $r_2 < 1\,\mathrm{m}$ then the accuracy for both sources remains below 10 %.

In the second scenario, the results in Fig. 5.11b show that the DEE is larger than 10 % for radii smaller than 1 m. For larger radii, the DEE drops significantly below 10 %, reaching values less than 1 % for $r_2 > 3.2\,\mathrm{m}$. The mean DEE appears to have a decreasing trend as the sources move to larger radii. This can be expected since the wave field of a PS that is far away approximates that of a plane wave.

As in the second scenario, the results of the third scenario also describe a decreasing trend for the DEE as shown in Fig. 5.11c. At $r_1 = r_2 = 0.2\,\mathrm{m}$ the mean DEE is already clearly below 10 % and dropping further. The DEE drops significantly below 1 % for $r_1 = r_2 > 2\,\mathrm{m}$.

5.5.6 Discussion of Simulation Results

The simulation study of the DOA estimation performance for different values of SNR and numbers of PWs confirm the degrading influence of noise on the estimation accuracy. The number of sources in the field also affect the accuracy. As stated by Blu et al. [BDV$^+$08], using a surplus of coefficients (i.e. $N > L$) increases the robustness against noise for an observation, i.e. the HD, with distinct Dirac pulses (compare Section 5.5.6.6). The significant difference between the mean and the median results suggest that 50 % of the estimates have an accuracy with less than 7 % error, while the other 50 % may suffer from outliers with DEE values larger than 20 %.

The relatively high DEE values would be lower if the difference in magnitude between the individual PWs were smaller (it was simulated as 40 dB). Due to the large differences in magnitude, some of the plane waves may either vanish below the noise floor or the estimates would be degraded due to a low SNR for that specific source.

5.5.6.1 Single OOP-PW

As it was argued in Section 5.4.6, the modulating sequence $\Psi_n(k, k_r^{(l)}, R)$ had little effect on the DOA estimation accuracy, but only when using a small range of coefficients, $|n| \leq 1$. When increasing the latter to $|n| \leq 5$, the performance degraded significantly for a distinct range of angles. This suggests that the overall modulation due to the factor $\Psi_n(k, k_r^{(l)}, R)$ is so strong that it overpowers the carrier in the extended sequence $q_{-5}, \ldots, q_5$ and thus leads to a high DEE (see also Section 5.5.6.6).

5.5.6.2 One PW and one OOP-PW

In light of the theoretical considerations made in Section 5.4.6, the severe loss of accuracy with the obtained estimates suggests that the modulating sequence $\Psi_n(k, k_r^{(2)}, R)$ is very strong in magnitude and its contribution cloaks both carriers $e^{-in\varphi_1}/\sqrt{2\pi}$ and $e^{-in\varphi_2}/\sqrt{2\pi}$, and strongly alters the roots of the annihilating filter. The result is compromised DOA estimates.

5.5.6.3 Single Point Source

Similar to the results for the single OOP-PW, the estimation performance for the single point source have shown results of high accuracy. Following the performance dependency on the range of coefficients q_n found for the single OOP-PW, it was found that the same applies for the field of a single point source, though the results were not presented for reasons of brevity. Accordingly, the modulation of the additional factor $\zeta_n(r')$ is also strong enough to affect the estimation accuracy when using a wider range of coefficients.

5.5.6.4 Two Point Sources

The scenario where the first PS is located at a small radius ($r_1 = 0.2\,\mathrm{m}$) and the second source is moved from $r_2 = 0.2\,\mathrm{m}$ to $r_2 = 9.7\,\mathrm{m}$ shows the worst performance among the results for two PSs. This suggests that the modulation component around the 'carrier' of the close PS (see Sec. 5.4.6) in the spatial frequency spectrum of q_n is cloaking the 'carrier' of the far PS, even at radii where r_2 is only slightly larger than or equal to r_1. The DOA estimate for φ_2 therefore becomes strongly inaccurate, leading to the comparably high mean DEE of more than 20 %.

In the scenario where the first PS is located far away ($r_1 = 9.7\,\mathrm{m}$), the DEE improvement that occurs as the distance of the second source increases can be explained by looking at the source strength. As discussed in Section 5.4.6, a strong difference in strength of the PSs' components in q_n might cause the modulation contributions to mask the 'carrier' associated with the weaker source. That would lead to a bad DOA estimate for the more distant PS, hence causing an elevated mean DEE for $r_2 < 2\,\mathrm{m}$. Once the levels become more comparable, the mean DEE is similar to that obtained in the third scenario.

The third scenario had both PSs located at the same distance from the origin. While still above 4 % for $r_1 = r_2 = 0.2\,\mathrm{m}$, the mean DEE drops quickly towards very small values as the source radii increase. The field of the PSs eventually effectively resembles that of plane waves. Since their source strength is equal, their estimates would on average be of similar quality. The high mean DEE at very close range, $r_1 = r_2 < 1\,\mathrm{m}$, suggests that the modulation components occurring in the coefficient sequence are strong enough to cloak the dominance of their carriers, leading to elevated DEE results.

5.5.6.5 Practical Implications of the Results for Point Sources and OOP-PW

Point sources and OOP-PW are very relevant for any practical implementation. The point source can model any source whose dimensions are small compared to the wave length that is located within close distance to the measurement aperture. However, even if a sound source is far enough away for its wave fronts to resemble those of a plane wave travelling parallel to the measurement plane, there may be floor or ceiling reflections that lead to an OOP-PW in the sound field.

5.5.6.6 Influence of the Range of Coefficients

As stated by Blu et al. [BDV+08], using the highest possible number of coefficients as input to the estimation stage increases the robustness of the estimates against noise through redundancy. It was however found in this work that this only applies for sound fields with distinct plane waves that propagate in the plane of the measurement aperture.

The simulations of both the point sources and the plane waves travelling out of the measurement plane have revealed that the imperfection introduced by the additional factors $\zeta_n(r')$ and $\Psi_n(k, k_r^{(l)}, R)$, respectively, affect the performance of the DOA estimates. For single sources, the influence of the additional factors can be overridden by keeping the range of coefficients q_n fed into the DOA estimation algorithm as small as possible. As the number of sources increases, the estimator requires a higher number of coefficients (see Section 5.4.3). At the same time, however, the influence of the modulating factors compromises the performance of the estimator (see Section 5.5.4.1).

5.6 Experimental Results

This section investigates the DOA estimation accuracy based on measured sound fields. The sound fields were created using ten different sources, both individually and in various combinations.

One goal of the experiments was to test the dependency of the accuracy of the DOA estimates as a function of the number of sources in the sound field. In order to do that, the data of the individual sources were linearly combined to generate arbitrary fields with multiple sources. The data that was measured with several sources active at the time was used to validate the results obtained through the linear combination of the individual data.

5.6.1 Measurement Setup

The measurement array consists of 15 electret microphones arranged in a circle with uniform distribution on an acrylic cylindrical baffle with radius $R = 7.5\,\text{cm}$ (see Figure 5.1). The array was set up in the anechoic chamber of the ISVR. Nine *Genelec 8020C* loudspeakers were arranged at different angles to the array at a distance larger than $r = 2.5\,\text{m}$, while one was deliberately put at a one meter distance to simulate a point source. A sketch of the arrangement is shown in Figure 5.12. The floor of the anechoic chamber consists of a steel grid over a volume of absorbing fibre glass wedges. To reduce the amount of reflections additional wedges were used to cover the floor between the loudspeakers and the microphone array.

Exponential sweeps of 10 seconds duration were used as measurement signals to exclude non-linear effects from the data and obtain a high measurement SNR, using the method proposed by Farina [Far00, Far07]. The impulse responses obtained from the inverse filtering were then transformed into the frequency domain by using the *MATLAB* implementation of the *Fast Fourier Transform* (FFT). As an alternative, the same measurements were performed with band-limited noise (1 to 8 kHz) for comparison. All data was evaluated at the optimal frequency $f_c \approx 5.1\,\text{kHz}$ (compare Eq. 5.4.4).

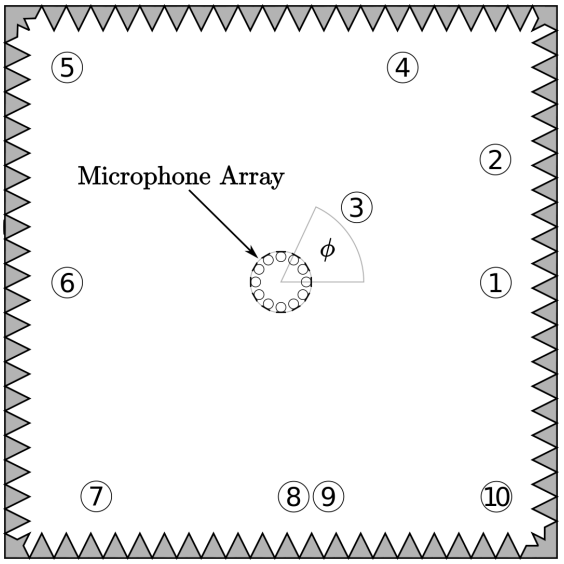


FIGURE 5.12: Sketch of the arrangement of microphone array and sound sources inside the anechoic chamber.

5.6.2 Measured DOA of the Sources

The DOA of the ten loudspeakers was individually measured using a laser angle gauge, located at the origin of the coordinate system. The loudspeakers were adjusted so that the laser pointed at the middle of the tweeter (see Fig. 5.13), ensuring all tweeters are in the same plane with the 15 microphone capsules. The distance of each tweeter from the origin was obtained with a tape

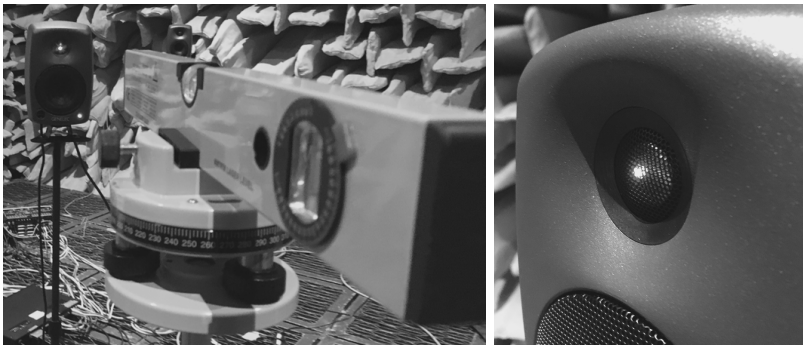


FIGURE 5.13: (Left) Angle laser gauge with Genelec 8020C, (Right) Close-up view of tweeter with laser reflection on the diaphragm.

measure. The results are noted down in Table 5.1.

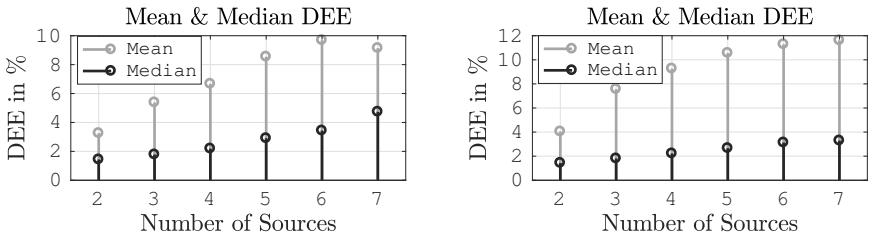
The depicted values will be used as reference for the testing of the DOA estimation performance based on the linear combinations of individual sound fields. The SNR was estimated using Eq. (5.37), where the value of the measured data $\tilde{q}_n$ was used in the numerator. The noise coefficients ϵ_n were estimated from data taken from the tail of the estimated impulse responses as proposed by Stan et al. [SEA02]. The SNR data suggests that the level difference between the strongest and weakest source is approximately 19.5 dB.

Source No.	Dist. to Origin	φ_l	$\tilde{\varphi}_l$	SNR/dB
1	2.88 m	0°	2.5°	73.0
2	3.47 m	30°	31.6°	72.9
3	1.04 m	45°	46.1°	80.6
4	3.45 m	60°	55.6°	69.1
5	4.20 m	135°	139.4°	65.9
6	3.08 m	180°	179.6°	61.1
7	3.24 m	230°	234°	73.4
8	2.52 m	275°	276.7°	68.0
9	2.56 m	280°	280.8°	77.3
10	3.35 m	315°	314.7°	72.1

TABLE 5.1: Reference location information for all ten loudspeakers, estimated ($\tilde{\varphi}_l$) DOA given in degrees and SNR obtained from measurements of the individual sources at $f \approx 5.1\,\text{kHz}$, $N = 7$.

5.6.3 Mean & Median DEE for Linear Combinations of Individual Measurements

The mean and median results for the DEE of DOA estimates obtained from measurements are shown in Figure 5.14a. The results for each number of sources were averaged over 1,000 iterations,



(A) Without Source No. 3 (B) With Source No. 3

FIGURE 5.14: Mean and median of the DEE in %, averaged over 1,000 randomised iterations for various number of sources.

during which sources were randomly chosen, where the near-field source (No. 3) was excluded. The algorithm was set to detect $N = 7$ sources, but only the L sources with the largest magnitude $|b_l|$ were chosen for these results.

It can be seen that both the mean and the median have a rising trend as L increases. The mean rises to approx. 9.7 % for $L = 6$ and the median reaches slightly under 4.7 % for $L = 7$. That makes for an uncertainty of $\pm 17.46°$ and 8.5°, respectively.

Figure 5.14b shows the results of the same experiment, only the near source was included in pool of sources this time. Both mean and median have a rising trend as L increases. The mean peaks at approx. 11.6 % for $L = 7$ and the median reaches slightly under 3.3 %. That makes for an uncertainty of $\pm 20.9°$ and $\pm 5.9°$, respectively.

The results of Table 5.1 indicates that the reference position of sources 4,5 and 7 may suffer from setup imperfections. The influence of these were investigated by repeating the analysis using the estimated values $\tilde{\varphi}_l$ in Table 5.1 as reference. It was found that the mean and median DEE results then improve slightly for $L < 5$, compared to the results given in Figure 5.14.

5.6.4 Example: Measurement of Six Sources

To confirm that the DOA estimation also works for fields with several sources playing (uncorrelated) noise at the same time, measurement data with six active sources were analysed. The results are given in the table below.

Source	1	3	5	7	9	10
$\tilde{\varphi}_l$ (est.) in [deg]	-0.2	46.3	139.1	220.9	270.8	310.7
φ_l (true) in [deg]	0	45	135	234	280	315

TABLE 5.2: Estimates from measurement signals obtained in a field with six sources active at the same time, with a mean DEE of 2.6 %.

5.6.5 Discussion

The values obtained for the mean and median of the DEE compare well to those obtained from the simulations with additional noise presented in Section 5.5.3. The discrepancy between mean and median results suggest that the estimates' accuracy sometimes suffers from outliers. The effect of the near-field source was reflected in the slightly degraded mean performance, yet the median remains nearly entirely unaffected by it; it even improves at high values of L, probably due to a relatively high SNR. This suggests that real sources are sufficiently different from ideal point sources to not entirely compromise the DOA estimation performance.

It must be expected that the close angular distance of sources 8 and 9 has an negative effect on the estimation performance. If the resolution of the method is not high enough to resolve the difference between the two sources, then this would potentially lead to worse DEE results whenever these two sources were randomly selected into the same set. The relatively good result of the multi-source measurement supports this hypothesis.

Another aspect that needs to be considered is the influence of the semi-anechoic nature of the measurement environment. The metal mesh of the floor may produce floor reflections, resulting in a wave field with plane waves that are not travelling parallel to the measurement plane z_M. It was shown in Section 5.5 that OOP-PW can degrade the performance of the DOA estimation mechanism due to the additional modulation of the sequence of coefficients q_n. It must thus be expected that the estimation performance with the measured data may have been suffering from degradation due to floor reflections. While this type of vulnerability of the estimation method is undesirable, it should be seen with relation to the fact that the measurement aperture is designed for two-dimensional measurements to begin with. It may be argued that this is actually not a fault of the method but of the two-dimensional measurement aperture.

5.7 Conclusions

A thorough analysis of a previously proposed DOA estimation method was presented. The work is based on theoretical considerations, simulation and experimental results. The DOA estimation method uses data obtained from circular arrays in sound fields with a finite rate of innovation.

A set of theoretical models for the array output in such sound fields was derived, considering both plane waves and point sources. It was found through theoretical considerations that the DOA estimates may suffer from point sources located within close range of the array and plane waves that do not travel parallel to the plane of the measurement aperture (OOP-PW).

The sound source models were applied in a simulation study to investigate the robustness and accuracy of DOA estimates against measurement noise and model imperfections. Percentages of average estimation errors were given, indicating the dependency of the DOA accuracy on the SNR and the number of plane waves in the field. It was found that the DOA estimation error for ideal sound fields increases significantly for sound fields where the rate of innovation exceeds what can be captured by the array.

Experimental results confirmed the findings on the performance in ideal sound fields of the simulation study. It was observed in both simulation and experimental results, that the mean accuracy is notably higher than its median. This suggests that the estimation method produces outliers that deviate from the median by a significant step, while 50 % of the time, the estimates have a much higher accuracy than is suggested by the mean accuracy.

Point sources and OOP-PW were shown to severely degrade the DOA estimation performance if their number in the field is larger than one. These types of sources were argued to be practically relevant and thus pose an important factor to consider in practical implementations. It can be concluded that, since it captures only a two-dimensional subset of the sound field, a circular array and the data obtained from it are particularly vulnerable to any sound field components that deviate from a two-dimensional scenario.

It is noteworthy that the proposed method does not require the signals associated with the incoming plane waves to be mutually uncorrelated. In fact, the developed sound field model even allows for them to be coherent. This is a major advantage over well established methods like MUSIC.

5.8 Appendix I: General Plane Wave Expansion in Cylindrical Coordinates

Using Eq. (5.1), the propagation term of the PW is given through

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik_r r(\cos\varphi\cos\phi+\sin\varphi\sin\phi)}e^{-ik_z z}, \tag{5.38}$$

where $\mathbf{r} = (r\cos\phi, r\sin\phi, z)^T$. Using the product theorems of the sine and cosine functions [BSMM06] and exploiting the symmetry of the latter yields

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik_r r\cos(\phi-\varphi)}e^{-ik_z z}. \tag{5.39}$$

The z-component of $\mathbf{k}$ can alternatively be written as a function of the elevation angle θ associated with the DOA, so that $k_z = k\cos\theta$. It then follows from (5.2) that $k_r = k\sin\theta$.

Using the Jacobi-Anger expansion [CK98] to expand the first term of Eq. (5.39) provides the expression

$$e^{-ik\sin\theta\ r\cos(\phi-\varphi)} = \sum_{n=-\infty}^{\infty} i^n J_n(-kr\sin\theta)e^{in(\phi-\varphi)}. \tag{5.40}$$

It follows from the series expansion of the Bessel function $J_n(z)$ [Wil99] that $J_n(-z) = (-1)^n J_n(z)$. The conjunction of this and the result in (5.40) applied to Eq. (5.39) leads to the series expansion of a single PW in cylinder basis functions [Wil99]

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-ik\cos\theta z} \sum_{n=-\infty}^{\infty} i^{-n} J_n(kr\sin\theta)e^{in(\phi-\varphi)}. \tag{5.41}$$

5.9 Appendix II: Series Expansion for the Sound Field of a Point Source

In cylindrical coordinates, the Green function of the Helmholtz equation can be expressed through cylinder basis functions [ABC+70, FM03], yielding

$$G(\mathbf{r}|\mathbf{r}') = \frac{i}{4}\sum_{n=-\infty}^{\infty} i^n \Phi_n^*(\phi')\Phi_n(\phi)\int_{-\infty}^{\infty}\frac{e^{ik_z(z-z')}}{2\pi}R_n(k_r r)H_n^{(1)}(k_r r')dk_z \tag{5.42}$$

for $r < r'$.

Note that the integral in (5.42) is not trivial to solve and typically needs to be either calculated numerically or approximated through a stationary phase approximation [Wil99, KFCS00].

5.10 Appendix III: HD Coefficients for the Sound Field of a Point Source

The pressure field of a single PS measured on the circle A in the plane $z = z_M$ yields the result of the form in Eq. (5.12). A comparison of the factors with those in Eq. (5.42) yields the expression for the pressure coefficients

$$p_n = b'\frac{i^{n+1}}{4}\Phi_n^*(\phi')\int_{-\infty}^{\infty}\frac{e^{ik_z(z_M-z')}}{2\pi}R_n(k_r R)H_n^{(1)}(k_r r')dk_z.$$

If the PS is located within the plane, i.e. $z' = z_M$, this expression simplifies to

$$p_n = b'\Phi_n^*(\phi')\frac{i^{n+1}}{8\pi}\int_{-\infty}^{\infty}R_n(k_r R)H_n^{(1)}(k_r r')dk_z. \tag{5.43}$$

Equating the right hand side expression with $R_n(kR)q_n$ (compare Eq. (5.15)) and rigorous reformulation of Eq. (5.43) yields the expression for the HD coefficients:

$$q_n = b'\frac{e^{-in\phi'}}{\sqrt{2\pi}}\frac{i^{n+1}}{8\pi}\frac{\int_{-\infty}^{\infty}R_n(k_r R)H_n^{(1)}(k_r r')dk_z}{R_n(kR)} \tag{5.44}$$

Chapter 6

A General Radiation Model for Sound Fields and Nearfield Acoustical Holography in Wedge Propagation Spaces

Falk-Martin Hoffmann, Earl G. Williams, Filippo Maria Fazi and Simone Fontana

Abstract

This work introduces a solution to the acoustic Helmholtz equation for wedge propagation spaces that is suitable for *Nearfield Acoustical Holography* (NAH). Such propagation spaces represent realistic scenarios for many acoustical problems where a free field assumption is not eligible. The proposed sound field model is derived from the general solution of the wave equation in cylindrical coordinates, using sets of orthonormal basis functions. The latter are modified to satisfy several boundary conditions representing the reflective behaviour of wedge-shaped propagation spaces. It is shown that the obtained solutions are suitable for NAH as well as the formulation of Neumann Green functions.

The model and its suitability for NAH are demonstrated through both numerical simulations and measured data, where the latter was acquired for a specific case with several individually controlled sources of vibration.

6.1 Introduction

Analytical sound field models exist in a variety of mathematical forms that can differ to suit different types of acoustic environments. Most rigorous analytical models are solutions to the

acoustic wave equation (at least within a problem-specific region of validity). Some models are simply based on a specific coordinate system that best describes a given problem. Some more specific models employ boundary conditions to include diffraction of scattering objects within the sound field. However, adding boundary conditions to a problem quickly makes it difficult to formulate an analytical model, which is where numerical methods can be applied instead. Especially models that are based on functional analysis are widely used for both sound field capture/analysis [Buc84, Wil99, Mey01, AW02, ME02, LDGD04, MDB06, TK06, LZ15] and reproduction/control [DMN03, AS08, FNW12, Ahr12, SAP14]. A typical approach to modelling a specific acoustic scenario is to take the solution pertaining to the free and undisturbed field (within a bounded volume) for the most suitable coordinate system for the given problem. All sound sources are typically assumed to be located outside of a bounded volume of interest and modelled on the basis of the superposition of either plane waves or point sources [CK98, Wil99, Faz10]. These solutions may be sufficient for many acoustic problems, yet those where scattering has a significant impact on the sound field require more accurate solutions. For example, the sound field models for acoustic transducer arrays often need to be modified to account for the scattering due to the array's own physical presence in the sound field. This is achieved by superimposing the specific solution for the sound field radiated from a source distribution to the free field solution, so that the sum satisfies the boundary condition [Wil99, Mey01, AW02, ME02, TK06, KFV10, KFV11].

The example of the model for the sound field of acoustic transducer arrays shows how introducing boundary conditions to a problem can serve to better describe the characteristics of a specific acoustic environment. Further examples for acoustic environments that require several boundary conditions are ducts and transmission lines [Web19, BX08], rooms [AB79, KFCS00] and underwater shorelines [Buc84, Buc89, LZ15]. Asvestas et al. [ABC$^+$70] described the wave equation as separable inside infinite wedges for both the Dirichlet and the Neumann boundary condition, yielding analytical expressions for the field of plane waves, and point and line sources through an infinite series expansion and infinite integration in the wavenumber domain. Buckingham [Buc84, Buc89] adopted their point source model to the field of shoreline underwater acoustics for different boundary conditions on the faces of the wedge. Luo and Zhang [LZ15] recently published exact solutions for both the Dirichlet and Neumann boundary conditions in the infinite wedge, including a numerical approach to integrate over potentially occurring singularities.

Boundary conditions also play a role in the application of Rayleigh's First Integral formula to describe the radiation from a piston integrated into an infinite baffle [KFCS00, BX08]. While this formula is a well established tool to model the radiation from panels, loudspeaker drivers or similar structures, it is still derived for a practically non-existent environment. Williams [Wil99] introduced the concept of Rayleigh-like formulae to transfer the principle to other coordinate systems and geometries.

In principle, these Rayleigh-like formulae project the behaviour of one specific boundary to any other point in space. NAH extends this concept to link two boundaries within a given sound field through propagators, which makes NAH particularly well suited for the use with coordinate systems in which the acoustic wave equation is separable [MWL85, Wil99, Wil01, WHH03]. Their solutions enable propagation of holograms either further away from, or back towards their source. These forward or inverse problems, respectively, can be solved through either an eigenvalue decomposition or a singular value decomposition of the solutions of the wave equation. Williams et al. presented an overview for planar and cylindrical NAH in [Wil01]. One prominent application

of NAH is to identify the vibration pattern on the surface of an object [Wil99, Wil01, WHH03], which makes it conceptually the inverse of Rayleigh-like formulae. Another application is to investigate the directivity of sound sources by looking at the pressure field directly [KB16] or by evaluating the intensity field from the reconstructed pressure and velocity holograms [WVHK06].

One key issue with NAH is that inverse problems are known to suffer from ill-conditioning when very small eigen/singular values meet measurement noise. In that case, regularisation of the inverse yields a stable solution at the cost of reducing spatial resolution [Han98, Wil01, FN07].

This work introduces a radiation model for propagation spaces that are relevant for many practical acoustical applications. It adopts the solution of the wave equation in acoustical wedges by Asvestas et al. [ABC$^+$70] as the basic model and extends it by yet another boundary condition. The final model enables the deduction of a *Neumann-Green Function* (NGF) and the formulation of a Rayleigh-like formula. Both provide tools to describe the acoustic radiation from sources in rooms or other wedge-shaped propagation spaces, e.g. a corner between two walls. Furthermore, it is shown that the model can be used for NAH. The findings are confirmed through both simulated and measured data.

The remainder of this work is organised as follows. Section 6.2 introduces the sound field model in the form of the solution to the acoustic Helmholtz equation in finite and infinite wedge propagation spaces. On the basis of the two models, Rayleigh-like formulae and NGFs are derived in Section 6.3. Section 6.4 defines the radial functions associated with the combination of the spatial modes, enabling the study of the propagation behaviour of individual wave components. Using these radial functions, the propagators needed for the reconstruction of both velocity and pressure from a measured pressure hologram are defined in Section 6.5. Section 6.6 presents simulated sound fields for two elementary types of source, one of which is later used to provide a comparison to measured data. The results of NAH with measured holograms obtained from a specially designed measurement rig are presented and discussed in Section 6.7 before the findings of this work are summarised in the final section.

6.2 Sound Field Model

6.2.1 General Solution of the Helmholtz Equation in the Infinite Wedge

The reference geometry used throughout this paper is based on the cylindrical coordinate system shown in Fig. 6.1. This enables a compact description of the boundary conditions for wedge-shaped propagation spaces. The model is developed in two stages, starting from the sound field in the free space.

The well-known general solution of the Helmholtz equation (without wedge) in cylindrical coordinates is determined through a separation of variables approach in the frequency domain [Wil99], yielding the series-integral representation

$$p(\mathbf{r},\omega) = \sum_{n=-\infty}^{\infty} \Phi_n(\phi) \int_{-\infty}^{\infty} A_n(k_z,\omega) R_n(k_r,r) Z(k_z z)\, dk_z, \tag{6.1}$$

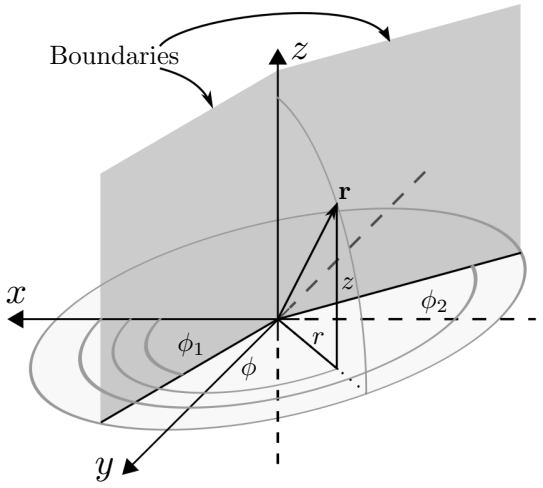


FIGURE 6.1: Cylindrical Coordinates with radius r, azimuth angle ϕ and height z, as well as further boundaries and specific angles needed in this work.

with $\mathbf{r} = (r, \phi, z)^T$, the basis functions

$$\Phi_n(\phi) = \frac{e^{in\phi}}{\sqrt{2\pi}}, \qquad \text{(angular basis functions)} \tag{6.2}$$

$$Z(k_z z) = \frac{e^{ik_z z}}{\sqrt{2\pi}}, \qquad \text{(axial basis functions)} \tag{6.3}$$

$$R_n(k_r, r) = H_n^{(1)}(k_r r), \qquad \text{(radial basis functions)} \tag{6.4}$$

and

$$k_r = \sqrt{k^2 - k_z^2}, \tag{6.5}$$

where $k = \frac{\omega}{c}$ denotes the wave number as a function of the angular frequency, ω, and the speed of sound, c. Since this work is devoted to the radiation of sound away from the coordinate origin, only the Hankel functions (of the first kind) are considered as radial basis functions in the following [Wil99]. Accordingly, the coefficients $A_n(k_z, \omega)$ describe a sound field radiated from sources located within a cylinder of radius r_S around the origin[Wil99]. The parameter ω is omitted hereunder for the sake of brevity.

The propagation space shall now be restricted to an *infinite wedge* by introducing two planes whose intersection coincides with the z-axis. The distinct angles ϕ_1 and ϕ_2 indicate the positions of the two planes (see Fig. 6.1).

Two types of boundary conditions are considered below. A *pressure release boundary* requires for the acoustic pressure to vanish at the angles ϕ_1 and ϕ_2 and is formally given by

$$p(\mathbf{r})\Big|_{\phi=\phi_1,\phi_2} = 0. \tag{6.6}$$

A *rigid boundary* affects the velocity vector field

$$\mathbf{v}(\mathbf{r}) = v_r(\mathbf{r}) \cdot \mathbf{e}_r + v_\phi(\mathbf{r}) \cdot \mathbf{e}_\phi + v_z(\mathbf{r}) \cdot \mathbf{e}_z \tag{6.7}$$

by forcing the velocity component coincident with and normal to the boundary to zero, i.e.

$$v_\phi(\mathbf{r})\big|_{\phi=\phi_1,\phi_2} = 0.$$

The above result can also be expressed as a boundary condition for the pressure gradient, using Euler's equation in the frequency domain

$$i\rho_0 ck\vec{v}(\mathbf{r}) = \nabla p(\mathbf{r}), \tag{6.8}$$

to obtain

$$\frac{\partial}{\partial\phi}p(\mathbf{r})\Big|_{\phi=\phi_1,\phi_2} = 0, \tag{6.9}$$

where $\phi_1 < \phi_2$.

The two boundary conditions specified in (6.6) and (6.9) individually lead to a new countable Total Orthonormal Set (TOS) [Kre78] of basis functions

$$\Phi_n(\phi) = \nu_n \begin{cases} \sin\left(\frac{n\pi}{\kappa}(\phi-\phi_1)\right), & \text{(Pressure Release BC)} \\ \cos\left(\frac{n\pi}{\kappa}(\phi-\phi_1)\right), & \text{(Rigid BC)} \end{cases} \tag{6.10}$$

with

$$\nu_n = \sqrt{\frac{2-\delta_n}{\kappa}}, \quad \text{and} \quad \kappa = \phi_2 - \phi_1,$$

where δ_n denotes the Kronecker Delta, $n \in \mathbb{N}$ and $\phi_1, \phi_2 \in [0, 2\pi)$. In either case it holds that the basis functions $\Phi_n(\phi)$ are all real-valued and hence describe exclusively standing wave components along the ϕ-direction. The corresponding solutions to Bessel's now modified equation are Hankel functions of non-integer order

$$R_n(k_r r) = H^{(1)}_{\frac{n\pi}{\kappa}}(k_r r), \quad n \in \mathbb{N}_0. \tag{6.11}$$

Note that the solution for the radial functions is the same for both types of boundary conditions.

The solution for sound fields in the infinite wedge is then given by

$$p(\mathbf{r}) = \sum_{n=0}^{\infty} \Phi_n(\phi) \int_{-\infty}^{\infty} A_n(k_z) R_n(k_r r)\, Z(k_z, z) dk_z. \tag{6.12}$$

The solution in (6.12) is identical to the one presented in [Buc84, Buc89, LZ15].

The model in Eq. (6.12) requires the uncountable set $\{Z(k_z, z)\}_{k_z\in\mathbb{R}}$ to enable the description of any arbitrary pressure field. This is in analogy to the inverse Fourier Transform, mapping from the frequency domain to the time domain. An infinitely spreading propagation space however hardly applies to real life conditions; On the contrary, most of the time it is necessary to consider the sound field in a room with a floor at $z = z_1$ and a ceiling at $z = z_2$, i.e. there is an additional boundary condition for those values of z.

The following subsection introduces a pair of boundary conditions along the z-axis to restrict the model to *finite wedges*.

6.2.2 General Solution in the Finite Wedge

The propagation space is now further restricted to a finite wedge by introducing two parallel planes $z = z_1$ (the floor) and $z = z_2$ (the ceiling), effecting a wedge-shaped waveguide. As for the angular boundary conditions, the model should allow for either a pressure release or a rigid boundary condition at the floor and ceiling. Following similar steps to those described in Section 6.2.1, the expression for the pressure release boundary condition is

$$p(\mathbf{r})\Big|_{z=z_1,z_2} = 0 \tag{6.13}$$

and for the rigid boundary condition

$$\frac{d}{dz}p(\mathbf{r})\Big|_{z=z_1,z_2} = 0. \tag{6.14}$$

Accordingly, the TOSs of basis functions that satisfy Eqs. (6.13) and (6.14), respectively, are given by

$$Z\left(k_z = \frac{m\pi}{\zeta}, z\right) = Z_{m,\zeta}(z) = \mu_m \begin{cases} \sin\left(\frac{m\pi}{\zeta}(z - z_1)\right), & \text{(Pressure Release BC)} \\ \cos\left(\frac{m\pi}{\zeta}(z - z_1)\right), & \text{(Rigid BC)} \end{cases} \tag{6.15}$$

with

$$\mu_m = \sqrt{\frac{2 - \delta_m}{\zeta}} \quad \text{and} \quad \zeta = z_2 - z_1,$$

where $z_1, z_2 \in \mathbb{R}$, $z_1 < z_2$ and $m \in \mathbb{N}$. Note that the TOS $\{Z_m(z)\}_{m\in\mathbb{N}}$ is now also countable, likewise yielding an infinite yet countable set of axial wave numbers

$$\{k_z^m\}_{m\in\mathbb{N}_0} := \left\{ k_z^m = \frac{m\pi}{\zeta} \;:\; m \in \mathbb{N} \right\} \tag{6.16}$$

associated with the propagation along the z-axis. As a consequence, Eq. (6.12) can now be rewritten without the integral but with a double sum instead:

$$p(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} A_n^m R_n\left(k_r^m r\right) \Phi_n(\phi)\, Z_{m,\zeta}(z). \tag{6.17}$$

Note that Eq. (6.5) now needs to be reformulated in accordance with (6.16), so that the radial wave numbers

$$k_r^m = \sqrt{k^2 - \left(\frac{m\pi}{\zeta}\right)^2}, \; m \in \mathbb{N} \tag{6.18}$$

form the now countable set $\{k_r^m\}_{m\in\mathbb{N}}$. The new basis functions $\Phi_n(\phi) \in \mathbb{R}$ and $Z_m(z) \in \mathbb{R}$ no longer describe propagating waves along the circumferential and axial direction, respectively, but exclusively standing wave components instead.

Eqs. (6.12) and (6.17) represent the general solutions to the acoustic Helmholtz equation inside infinite and finite (in z) wedge spaces, respectively, and the coefficients $A_n(k_z)$ and A_n^m describe sound fields in the respective propagation spaces. The next subsection describes how these

coefficients are related to the pressure and velocity spectrum and how the latter two can be identified from pressure *holograms*.

6.2.3 Pressure and Velocity Spectrum

Let $\{p(\mathbf{r})\}_{\mathbf{r}\in\Lambda}$ be a hologram of the pressure at a given radius r_H with

$$\Lambda := \begin{cases} \{\mathbf{r} : r = r_H, \phi \in [\phi_1, \phi_2], z \in \mathbb{R}\}, & \text{infinite wedge,} \\ \{\mathbf{r} : r = r_H, \phi \in [\phi_1, \phi_2], z \in [z_1, z_2]\}, & \text{finite wedge.} \end{cases} \tag{6.19}$$

Taking the scalar product [Kre78, CK98] of the hologram and the model's basis functions defines the spatial analysis transform, the result of which provides the holograms' spectrum for an infinite wedge as

$$P_n(k_z, r_H) = \int_{\phi_1}^{\phi_2} \int_{-\infty}^{\infty} p(r_H, \phi, z)\Phi_n(\phi)\, Z^*(k_z z)\; dz\, d\phi, \tag{6.20}$$

and for a finite wedge as

$$P_n^m(r_H) = \int_{\phi_1}^{\phi_2} \int_{z_1}^{z_2} p(r_H, \phi, z)\Phi_n(\phi)\, Z_m\,(z)\; dz\, d\phi. \tag{6.21}$$

Exploiting the orthonormality relation of TOSs to invert the expressions in (6.20) and (6.21) provides the corresponding spatial synthesis transforms

$$p(r_H, \phi, z) = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} P_n(k_z, r_H)\Phi_n(\phi)\, Z(k_z z)\, dk_z \tag{6.22}$$

and

$$p(r_H, \phi, z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P_n^m(r_H)\Phi_n(\phi)\, Z_m\,(z)\,, \tag{6.23}$$

respectively.

Replacing the expression for the pressure in Eqs. (6.20) and (6.21) by the expressions in (6.12) and (6.17), respectively, and exploiting of the orthonormality relation of orthonormal sets [Kre78, CK98] yields the relation between the sound field coefficients and the pressure spectrum for the infinite wedge

$$A_n(k_z)R_n(k_r r_H) = P_n(k_z, r_H), \tag{6.24}$$

and the finite wedge

$$A_n^m R_n\,(k_r^m r_H) = P_n^m(r_H). \tag{6.25}$$

Both equations can be rewritten for $A_n(k_z)$ and A_n^m, respectively, if

$$|R_n(k_r r_H)\,| \neq 0 \quad \text{and} \quad |R_n\,(k_r^m r_H)\,| \neq 0,$$

which applies for the Hankel functions of the first kind for $0 < |\alpha r_H| < \infty$ and $r_H \in \mathbb{R}$, where $\alpha = k_r$ and $\alpha = k_r^m$ as defined in (6.5) and (6.18), respectively [CS82].

In accordance with what is shown for the pressure hologram above, the expression for the spectrum of a particle velocity hologram in the infinite wedge is given by

$$\mathbf{V}_n(k_z, r_H) = \int_{\phi_1}^{\phi_2} \int_{-\infty}^{\infty} \mathbf{v}(r_H, \phi, z)\Phi_n(\phi)\, Z^*(k_z z)\; dz\, d\phi. \tag{6.26}$$

and for the finite wedge by

$$\mathbf{V}_n^m(r_H) = \int_{\phi_1}^{\phi_2} \int_{z_1}^{z_2} \mathbf{v}(r_H, \phi, z)\Phi_n(\phi)\, Z_m\,(z)\; dz\, d\phi. \tag{6.27}$$

Since $\mathbf{v}(\mathbf{r})$ is a vector, it follows that the resulting spectra must have the same dimensions.

At this point, the model for sound propagation in infinite and finite wedges is complete, however, some applications can benefit from some more convenient solutions. The following section presents the derivation of Rayleigh-like formulae [Wil99] and NGFs on the basis of the developed model.

6.3 Rayleigh-like Formulae and Neumann-Green Functions

Rayleigh-like formulae describe the pressure field radiated from the velocity distribution on a given rigid boundary

$$V := \begin{cases} \{\mathbf{r} : r = r_S, \phi \in [\phi_1, \phi_2], z \in \mathbb{R}\}, & \text{infinite wedge,} \\ \{\mathbf{r} : r = r_S, \phi \in [\phi_1, \phi_2], z \in [z_1, z_2]\}, & \text{finite wedge,} \end{cases} \tag{6.28}$$

inside the propagation space and are thus conceptually similar to Rayleigh's first integral formula [Wil99]. Since the velocity at V causes acoustic radiation, the entire boundary effectively becomes a source. The developed models both already support this premise through the choice of the radial basis functions in (6.11). It can be assumed that the velocity of V is limited to a radial component, i.e. $\mathbf{v}(r_S, \phi, z) = v_r(r_S, \phi, z) \cdot \mathbf{e}_r$ (see Fig. 6.2) [Wil99].

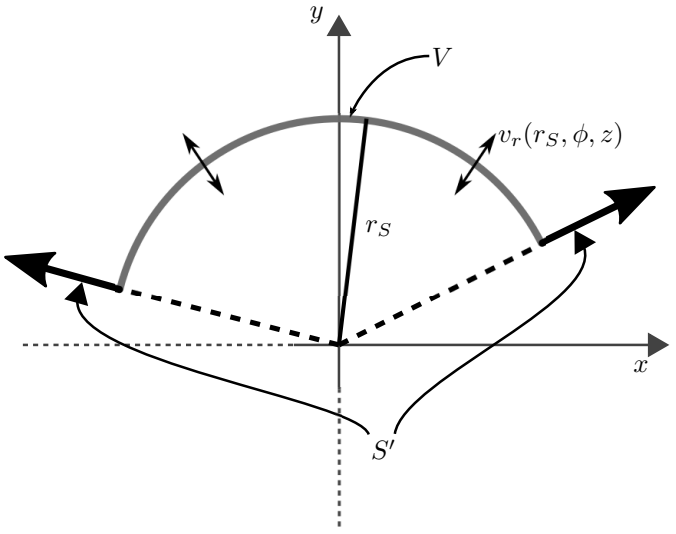


FIGURE 6.2: Concept of the radiation from a curved boundary V with the given velocity profile $v_r(r_S, \phi, z)$ inside the wedge defined by the rigid boundaries S'.

6.3.1 Pressure Radiated from the Boundary V for Infinite Wedges

The nature of the velocity $\mathbf{v}(r_S, \phi, z)$ implies that

$$\mathbf{V}_n(k_z, r_s) = V_n(k_z, r_S) \cdot \mathbf{e}_r.$$

In this case, Euler's equation reduces to

$$V_n(k_z, r) = \frac{-i}{\rho_0 ck} \frac{\partial}{\partial r} P_n(k_z, r). \tag{6.29}$$

Since the coefficients $A_n(k_z)$ fully describe a sound field inside the infinite wedge, it can be derived from Eq. (6.24) that the relation between two pressure spectra at different radii is

$$P_n(k_z, r) = \frac{R_n(k_r r)}{R_n(k_r r_S)} P_n(k_z, r_S). \tag{6.30}$$

Using the above to replace $P_n(k_z, r)$ in (6.29) yields

$$V_n(k_z, r) = \frac{-ik_r}{\rho_0 ck} \frac{R'_n(k_r r)}{R_n(k_r r_S)} P_n(k_z, r_S), \tag{6.31}$$

where

$$R'_{n,\kappa}(x) = \frac{\partial}{\partial x} H_{\frac{n\pi}{\kappa}}(x). \tag{6.32}$$

Provided that $k_r R'_n(k_r r) \neq 0$, transposing Eq. (6.31) for the pressure spectrum $P_n(k_z, r)$ and exchanging r with r_S yields

$$P_n(k_z, r) = i\rho_0 ck \frac{R_n(k_r r)}{k_r R'_n(k_r r_S)} V_n(k_z, r_S), \; r \geq r_S. \tag{6.33}$$

The synthesis of the pressure $p(\mathbf{r})$ at radius r from the result in (6.33) through (6.22) leads to the expression for the pressure as a function of the velocity spectrum on the boundary V

$$p(\mathbf{r}) = i\rho_0 ck \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{R_n(k_r r)}{k_r R'_n(k_r r_S)} V_n(k_z, r_S) \Phi_n(\phi) \, Z(k_z z) \, dk_z. \tag{6.34}$$

An analytic expression for (6.34) is hard to find due to the rather complex integral. Therefore it is common to limit the bounds of integration and solve it numerically instead. When attempting that, it is important to deal with the case when $k_r = 0$ as it poses a singularity. To solve that issue, Luo and Zhang proposed to marginally shift the integration path into the complex plane in order to avoid the singularity [LZ15]. Furthermore, it is necessary to limit the integration bounds for the numerical calculation. The error introduced by this limitation can be controlled by considering the evanescent behaviour of higher order components [Wil99]. This is more thoroughly discussed in Section 6.4.

6.3.2 Pressure Radiated from the Boundary V for Finite Wedges

The expression corresponding to (6.34) for the finite wedge can be derived in a similar way. Following the same steps of the previous section but with the model for the finite wedge yields

$$p(\mathbf{r}) = i\rho_0 ck \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{R_n\left(k_r^m(\zeta) r\right)}{k_r^m(\zeta) R_n'\left(k_r^m(\zeta) r_S\right)} V_n^m(r_S) \Phi_n(\phi)\, Z_m\left(z\right). \tag{6.35}$$

The dependency of k_r^m on ζ is omitted in the remainder of this work for the sake of brevity.

One major advantage of the expression in (6.35) over the one in (6.34) is that it does not require (numerical) integration. The summations should be infinite, however, the series truncation error can be kept below a reasonable limit since the higher order components of $\Phi_n(\phi)$ and $Z_m\left(z\right)$ become evanescent [Wil99] when

$$\frac{n\pi}{\kappa} > k_r^m r,$$

potentially leading to a negligible contribution of that particular component, depending on the observation radius r. This is further discussed in Section 6.4.

6.3.3 Rayleigh-like Formula and Neumann-Green Function for Infinite Wedges

Equation (6.34) can be used to derive both the Rayleigh-like formula and the NGF for radiation problems based on the normal velocity distribution v_n on a boundary $S = V \cup S'$ (see Fig. 6.2), where

$$S' := \{\mathbf{r} : r > r_S\ ,\ \phi = [\phi_1, \phi_2]\ ,\ z \in \mathbb{R}\} \tag{6.36}$$

The defining condition of an NGF is that it satisfies

$$\left.\frac{\partial G_N}{\partial n}\right|_S = 0, \tag{6.37}$$

where n denotes the geometrical dimension that locally corresponds to the normal vector on S. However, for the infinite and the finite wedge, this condition is inherent to the flanking walls for the developed sound field model (compare Eq. (6.9)), and in the case of the finite wedge also for the roof and floor (compare Eq. (6.14)) of the propagation space, respectively. This entails that the NGF for infinite and finite wedges depends solely on the boundary condition

$$\left.\frac{\partial G_N}{\partial r}\right|_V = 0. \tag{6.38}$$

Referring to the *Kirchhoff Integral Equation* (KIE) as given in [Wil99], the Rayleigh-like formula for the infinite wedge is based on the NGF and must be of the form

$$p(\mathbf{r}) = i\rho_0 ck \int_{\phi_1}^{\phi_2} \int_{-\infty}^{\infty} G_N(\mathbf{r}|\mathbf{r}') v_n(\mathbf{r}') r_S dz' d\phi' \tag{6.39}$$

with $\mathbf{r}' = (r_S, \phi', z')$. Using (6.26) to compare the Rayleigh-like formula to the expression in (6.34) yields the Neumann-Green function for the infinite wedge

$$G_N(\mathbf{r}|\mathbf{r}') = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{1}{r_S} \frac{R_n(k_r r)}{k_r R_n'(k_r r_S)} \Phi_n(\phi')\, Z^*(k_z z')\, \Phi_n(\phi)\, Z(k_z z)\, dk_z. \tag{6.40}$$

To confirm that this results satisfies (6.38), (6.40) is derived w.r.t. r and the result evaluated for $\mathbf{r} \in V$, yielding

$$\frac{\partial G_N(\mathbf{r}|\mathbf{r}')}{\partial r}\Big|_{\mathbf{r}\in V} = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{1}{r_S} \Phi_n(\phi')\, Z^*(k_z z')\, \Phi_n(\phi)\, Z(k_z z)\, dk_z. \tag{6.41}$$

From the completeness relation of total orthonormal sets [Kre78][Wil99], it follows that

$$\frac{\partial G_N(\mathbf{r}|\mathbf{r}')}{\partial r}\bigg|_{\mathbf{r}\in V} = \frac{1}{r_S}\delta(\phi - \phi')\delta(z - z'). \tag{6.42}$$

This confirms that the normal derivative of the NGF is indeed equal to zero everywhere on the boundary V but at the position of the singularity/point source $\mathbf{r}' = (r_S, \phi', z')^T$ (compare to Chapter 8 in [Wil99]).

6.3.4 Rayleigh-like Formula and Neumann-Green Function for Finite Wedges

Following a similar procedure as in the previous subsection for finite wedges yields the expression for the NGF

$$G_N(\mathbf{r}|\mathbf{r}') = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{r_S} \frac{R_n\left(k_r^m r\right)}{k_r^m R_n'\left(k_r^m r_S\right)} \Phi_n(\phi')\, Z_m\left(z'\right) \Phi_n(\phi)\, Z_m\left(z\right) \tag{6.43}$$

that likewise satisfies the boundary condition

$$\frac{\partial G_N(\mathbf{r}|\mathbf{r}')}{\partial r}\Big|_{\mathbf{r}\in V} = \frac{1}{r_S}\delta(\phi - \phi')\delta(z - z') \tag{6.44}$$

everywhere but at the position of the singularity/point source $\mathbf{r}' = (r_S, \phi', z')^T$. The Rayleigh-like formula for the finite wedge is then given by

$$p(\mathbf{r}) = i\rho_0 ck \int_{\phi_1}^{\phi_2} \int_{z_1}^{z_2} G_N(\mathbf{r}|\mathbf{r}') v_n(\mathbf{r}') r_S dz' d\phi'. \tag{6.45}$$

With both Rayleigh-like formulae given in explicit form through (6.39) and (6.45), it is now possible to describe the radiated sound field from arbitrary source distributions on the boundary V.

6.4 Radial Functions

The practical implementation of Eqs. (6.34) and (6.35), as well as Eqs. (6.39) and (6.45) require limitation of the integration bounds or truncation of the sum, respectively. The following considerations are equally relevant for both the Rayleigh-like formulae and the NGFs, but their meaning becomes more intuitive when evaluating the expressions in (6.34) and (6.35) instead.

The determining factor for the limitation/truncation is the transfer function of the velocity modes on the boundary to the corresponding pressure modes at a given observation radius r. These transfer functions are also referred to as 'radial functions' and can be identified from Eqs. (6.34) and (6.35) as shown in the following subsection.

6.4.1 Identifying the Radial Functions

Considering that both axial and angular basis functions are orthonormal, the radial functions for the infinite wedge are identified as

$$\Gamma_n(k_z, r, r_S, \kappa) = i\rho_0 ck \frac{R_n(k_r r)}{k_r R_n'(k_r r_S)}. \tag{6.46}$$

The dependency on k_z stems from Eq. (6.5).

Similarly, the radial functions for the finite wedge are given by

$$\Gamma_n^m(r, r_S, \kappa, \zeta) = i\rho_0 ck \frac{R_n(k_r^m r)}{k_r^m R_n'(k_r^m r_S)}. \tag{6.47}$$

In both cases the radial functions describe a discrete spectrum in the angular domain; However, Eq. (6.46) is continuous in the k_z-domain while (6.47) also has an entirely discrete spectrum. Comparing the model descriptions at hand to the interrelations of Fourier Transform and Fourier Series, this is not at all surprising.

6.4.2 Properties

Figure 6.3 shows an example of Γ_n for $r = 2\,\mathrm{m}$, $r_S = 0.15\,\mathrm{m}$, $\kappa = \pi$ and $f = 5\,\mathrm{kHz}$. To help with legibility, the plot is continuous along the n-axis, but only discrete points of n are relevant. It can be observed that modal combinations (n, k_z) with axial wave numbers $k_z > k$ are heavily attenuated, i.e. they decay as evanescent wave components. The same applies for evanescent angular modes $n > k\frac{r_S \kappa}{\pi}$, only their descent is not as steep as for the axial wave numbers [Wil99].

The radial functions describe the physics of the propagation inside wedge spaces: While purely axial modes ($n = 0$) cannot transition from evanescent into propagating modes as the radius increases, purely angular modes ($k_z = 0$) can. The reason is simply that a mode n may oscillate too densely in space at a smaller radius, so that across a small arc section it 'shuffles' air between its maxima and minima, while at a larger radius, the same mode now oscillates across a longer arc length, no longer effecting an acoustic 'short circuit'. This principle does not hold for axial modes, but it does for modal combinations (n, k_z). With view on the NAH performance prospect,

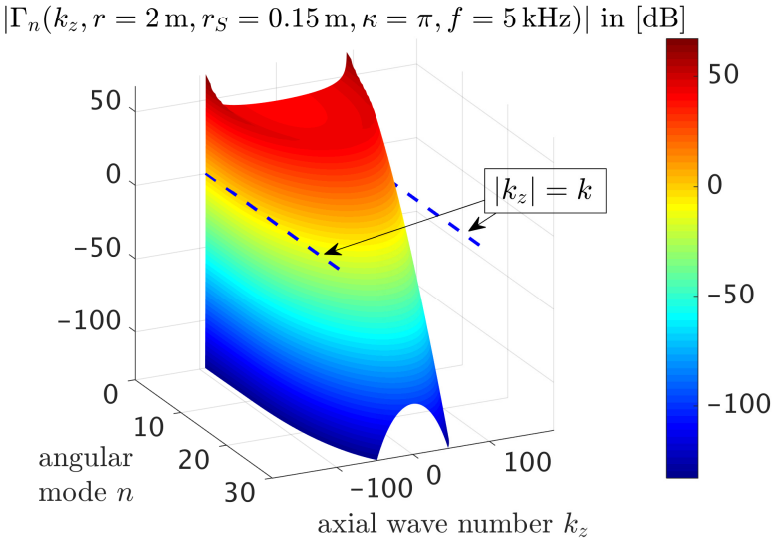


FIGURE 6.3: Magnitude of the radial functions $\Gamma_n(k_z, r, r_S, \kappa)$ in [dB] with reference to unity, $r = 2\,\mathrm{m}$, $r_S = 0.15\,\mathrm{m}$, $\kappa = \pi$, $f = 5\,\mathrm{kHz}$.

this finding suggests that having a small standoff distance, i.e. the gap between hologram (measurement) plane and V, is less critical for the NAH resolution in the ϕ-domain than for the resolution in the z-domain (see Section 6.5).

These properties equally apply to the radial functions for the finite wedge, since they are but a sampled version with $k_z^m = \frac{\pi m}{\zeta}$, $n \in \mathbb{N}$.

6.4.3 Truncation and Limitation of the Solutions

As a consequence of the propagating modes accumulating around the origin of the n-k_z-plane or the n-m-plane, respectively, the integrations and summations in (6.34) and (6.35) can be truncated while maintaining a good approximation of the actual sound field, provided that $A_n(k_z)$ and A_n^m have a similar, finite magnitude for all combinations (n, k_z) and (n, m), respectively.

In the case of (6.34), the integration limit $K_z > k$ and the final term of the truncated sum $n = N$ can be chosen such that

$$\tilde{p}(\mathbf{r}) = \sum_{n=0}^{N} \int_{-K_z}^{K_z} \Gamma_n(k_z, r, r_S, \kappa) V_n(k_z, r_S) \Phi_n(\phi)\, Z(k_z, z) dk_z \tag{6.48}$$

approximates the actual $p(\mathbf{r})$ adequately. This requires however the a priori knowledge of *all* coefficients $V_n(k_z, r_H)$ to determine the contribution of evanescent modes in comparison to the propagating modes.

In the case of (6.35), suitable limits N and M can be found that truncate the double sum so that

$$\tilde{p}(\mathbf{r}) = \sum_{n=0}^{N} \sum_{m=0}^{M} \Gamma_n^m(r, r_S, \kappa, \zeta) V_n^m(r_S) \Phi_n(\phi)\, Z_m(z) \tag{6.49}$$

approximates $p(\mathbf{r})$ with reasonable error.

Not only is the evaluation of the radial functions useful for the numerical calculation of sound fields, but they also provide valuable information when attempting to solve inverse problems [CK98, Wil99]. Inverse problems are a key component of NAH, whose methods are adapted for both infinite and finite wedges in the subsequent section.

6.5 Holography in Infinite and Finite Wedges

This section introduces the equations that enable acoustic holography, i.e. the reconstruction of the pressure (and also the particle velocity) at any point within the wedge space from the knowledge of a hologram $\{p(\mathbf{r})\}_{\mathbf{r}\in\Lambda}$ (see Section 6.2.3).

Equation (6.30) links pressure spectra obtained from different radii for infinite wedges. An equivalent version can be formulated for the finite wedge, yielding

$$P_n^m(r) = \frac{R_n\left(k_r^m r\right)}{R_n\left(k_r^m r_S\right)} P_n^m(r_S). \tag{6.50}$$

Equivalently, replacing the expression for the pressure in Eqs. (6.20) and (6.21) by Eqs. (6.34) and (6.35) (evaluated at $r = r_H$), respectively, shows that the radial functions in (6.46) and (6.47) link the pressure spectra of the hologram at radius r_H to the normal velocity spectra at the radius r_S, yielding

$$P_n(k_z, r_H) = \Gamma_n(k_z, r_H, r_S, \kappa) V_n(k_z, r_S) \tag{6.51}$$

and

$$P_n^m(r_H) = \Gamma_n^m(r_H, r_S, \kappa, \zeta) V_n^m(r_S), \tag{6.52}$$

respectively. A key component of NAH is to reformulate both the above equations as an expression to calculate the velocity spectra, which can be obtained by multiplying both sides of the equations by the reciprocals of the radial functions. However, the evanescent nature of the higher order modes reflects in an extremely small magnitude of the radial functions for modes outside the "radiation circle" defined by $k = \sqrt{k_z^2 + (\frac{\pi n}{\kappa r_S})^2}$ (see Section 6.4.2). Accordingly, their reciprocals take extremely large values, effecting heavy amplification of the corresponding pressure modes. While no problem in theory, the measurement noise of the hologram data in practice also manifests itself in the derived pressure spectra. When heavily amplified, the noise in the spectra can severely corrupt the results for $V_n(k_z, r_S)$ and $V_n^m(r_S)$, respectively. This common issue in the field of inverse problems can be addressed by introducing a Tikhonov regularisation to the reciprocal of the radial functions. The regularised solutions for the velocity spectra $\tilde{V}_n(k_z, r_S)$ and $\tilde{V}_n^m(r_S)$ are then given by

$$\tilde{V}_n(k_z, r_S) = \frac{\Gamma_n^*(k_z, r_H, r_S, \kappa)}{|\Gamma_n(k_z, r_H, r_S, \kappa)|^2 + \beta_V^2} \hat{P}_n(k_z, r_H) \tag{6.53}$$

and

$$\tilde{V}_n^m(r_S) = \frac{\Gamma_n^{m*}(r_H, r_S, \kappa)}{|\Gamma_n^m(r_H, r_S, \kappa)|^2 + \beta_V^2} \hat{P}_n^m(r_H), \tag{6.54}$$

respectively. The $\hat{\cdot}$ symbol on the pressure spectra denotes that they were calculated from holograms measured at r_H by evaluating equations (6.20) and (6.21), respectively. A suitable regularisation parameter β_V that is optimal for minimising the impact of noise in the measured data can be chosen through the L-Curve analysis [Han98, Han00] or other techniques.

From the literature it is known that regularisation of the inverse problem is equivalent to a spatial lowpass filtering of the synthesised sound field quantity (pressure and velocity). Examples of such surface velocity reconstructions from measured data and further practical aspects of the mechanisms described above are presented in the Section 6.7.

Once the normal velocity spectra have been obtained from (6.53) or (6.54), respectively, the velocity on V can be reconstructed using the synthesis equation

$$\tilde{v}_r(r_S, \phi, z) = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \tilde{V}_n(k_z, r_S) \Phi_n(\phi)\, Z(k_z z)\, dk_z \tag{6.55}$$

for the infinite wedge or

$$\tilde{v}_r(r_S, \phi, z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \tilde{V}_n^m(r_S) \Phi_n(\phi)\, Z_m\,(z) \tag{6.56}$$

for the finite wedge (compare to Eqs. (6.22) and (6.23)).

6.6 Sound Field Simulations

The numerical calculation of the pressure field was accomplished on the basis of the truncated synthesis equations (6.48) and (6.49). In the case of (6.48) it is additionally necessary to numerically solve the integration. For the results in this work, the truncated integration was numerically solved using the trapezoidal rule [BSMM06] (1^{st} order approximation) with the integration limits

$$K_z = \left\lceil \frac{Qk}{\Delta k_z} \right\rceil \Delta k_z, \quad Q \in \mathbb{R} \geq 1, \tag{6.57}$$

and the step size Δk_z for the integration variable, where $\lceil \cdot \rceil$ denotes the ceiling operator. The ceiling operator ensures that all propagating wave components are definitely part of the synthesis. The parameter Q can be adjusted to control the amount of evanescent components included in the calculation. Alternatively, the integral in (6.48) can be solved using the IFFT [Wil99, WHH03], which is commonly done in the field of NAH as it may increase the speed of the calculation. The limits for the truncation of the sums were chosen as

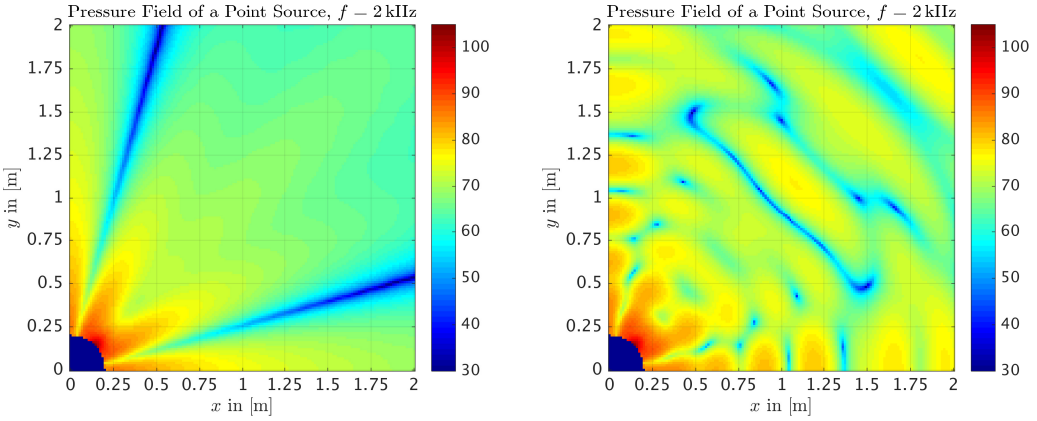
$$N = \lceil Qkr_S \rceil \tag{6.58}$$

for the angular components and

$$M = \left\lceil \frac{Qk\zeta}{\pi} \right\rceil \tag{6.59}$$

for the axial components.

The following subsections introduce two elementary types of sources mounted on the baffle for both propagation spaces.



(A) Infinite Wedge (B) Finite Wedge

FIGURE 6.4: Pressure field of a point source located at $(0.2\,\mathrm{m}, \frac{\pi}{4}, 0)^T$, $Q = 1.5$, $\Delta k_z = 0.01\,\frac{\mathrm{rad}}{\mathrm{m}}$.

6.6.1 Point Source Model

The NGFs presented in Sections 6.3.3 and 6.3.4 can be applied to model the sound field of a point source on the rigid baffle V located at $(r_S, \phi', z')^T$ for the infinite and the finite wedge, respectively. The resulting pressure field can be calculated from Eq. (6.39) and Eq. (6.45), respectively, where

$$v_n(\mathbf{r}') = v_r(\mathbf{r}') = \frac{1}{r_S}\delta(r - r_S)\delta(\phi - \phi')\delta(z - z'). \tag{6.60}$$

Figure 6.4a shows the magnitude of the pressure field in the xy-plane ($z = 0$) for an infinite wedge with $\kappa = \frac{\pi}{2}$, $r_S = 0.2\,\mathrm{m}$, $\phi' = \frac{\pi}{4}$ and $z' = 0$. The interference between the radiated waves and the reflections from the rigid walls generate a clear radiation pattern. Particularly two angles at which destructive interference occurs point distinct nulls in the radiation pattern.

The same scenario for a finite wedge with $\kappa = \frac{\pi}{2}$ and $\zeta = 2\,\mathrm{m}$, where $z_1 = -1\,\mathrm{m}$ and $z_2 = 1\,\mathrm{m}$, is shown in Fig. 6.4b. In comparison to the field shown in Fig. 6.4a, the field in the finite wedge also suggests a weakened radiation in those same directions where Fig. 6.4a shows distinct destructive interference, the attenuation however being nowhere near as strong. This must be due to the now occurring reflections from the floor and ceiling that overlay their own interference pattern. The latter is responsible for increased maxima and minima observed when increasing the radius along the angle $\phi = \phi'$.

6.6.2 Piston Source Model

Another type of source that is simple to simulate with the given model is that of a piston source on the baffle. It can be modelled as a patch in the radial component of the velocity profile on the boundary V where $v_r(r_S, \phi', z') \neq 0$. For a patch located centrally on V, of angular width α and axial width B, the velocity profile is given by the product

$$v_r(r_S, \phi, z) = \Pi\left(\frac{\phi - \frac{\kappa}{2}}{\alpha}\right)\Pi\left(\frac{z}{B}\right) \tag{6.61}$$

where $\Pi(\cdot)$ denotes the rectangular function

$$\Pi(\phi) = \begin{cases} 1, & |\phi| \leq \frac{1}{2}, \\ 0, & |\phi| > \frac{1}{2}. \end{cases} \tag{6.62}$$

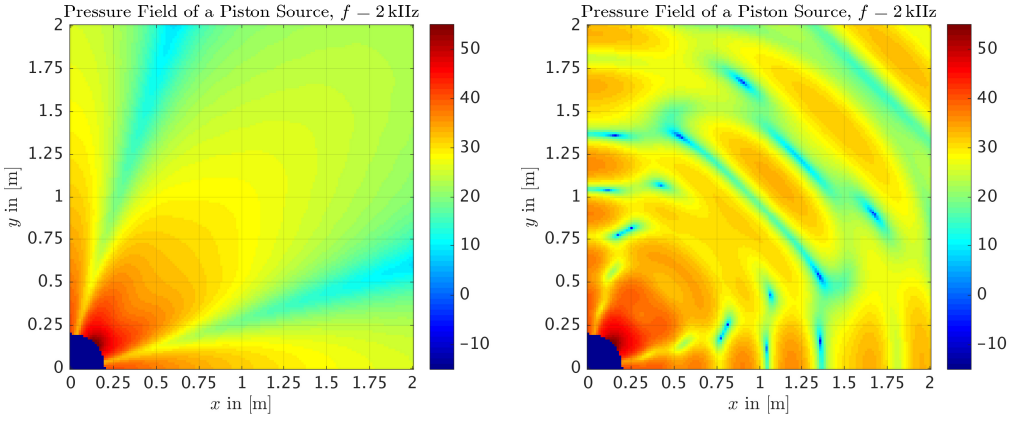
The corresponding radial component of the velocity spectra for the infinite and finite wedge, as calculated from (6.26) and (6.27), are

$$V_n(\omega, k_z, r_S) = \alpha B \mathrm{sinc}\left(k_z \frac{B}{2}\right) Z^*(k_z 0)\, \mathrm{sinc}\left(\frac{n\pi}{\kappa}\frac{\alpha}{2}\right) \Phi_n\left(\frac{\pi}{4}\right) \tag{6.63}$$

and

$$V_n^m(\omega, r_S) = \alpha B \mathrm{sinc}\left(\frac{m\pi}{\zeta}\frac{B}{2}\right) Z_m\left(\frac{\zeta}{2}\right) \mathrm{sinc}\left(\frac{n\pi}{\kappa}\frac{\alpha}{2}\right) \Phi_n\left(\frac{\pi}{4}\right), \tag{6.64}$$

respectively. In this case Eqs. (6.34) and (6.35), respectively, are more suitable to directly calculate the sound fields. The results for the infinite and finite wedge are shown in Figures 6.5a and 6.5b, respectively. With the piston size in a similar order of magnitude as the wave length,



(A) Infinite Wedge (B) Finite Wedge

FIGURE 6.5: Pressure field of a piston located at $(0.2\,\mathrm{m}, \frac{\pi}{4}, 0)^T$, $Q = 1.5$, $\Delta k_z = 0.01\,\frac{\mathrm{rad}}{\mathrm{m}}$, $\alpha = \frac{\pi}{6}$, $B = 0.1\,\mathrm{m}$.

the fields in Fig. 6.5 show similar characteristics as those in Fig. 6.4a, yet the attenuation with propagated distance is smaller than that of the point source. The overall level difference between the results for the point source and those for the piston source stems from the leading factor αB in Eqs. (6.63) and (6.64).

6.7 Measurements

In order to support the developed theory and model, holographic measurements of piston-like sources were conducted inside a quasi-infinite wedge with $\kappa = \pi$. The prototype of a hemi-cylindrical loudspeaker array [HFF16] with 15 Tang Band W1-1070SH drivers was used as a radiating source with $r_S = 0.1426\,\mathrm{m}$. The measured hologram data is compared to simulated data, assuming the sound field of rectangular pistons with an edge length equal to the diameter of

the drivers in the prototype, each located in the corresponding positions. Pressure spectra were calculated from the measured holograms in order to perform NAH. The reconstructed velocity profile of the radiating surface is then compared against the underlying actual source profile.

The next subsection introduces the measurement rig in detail.

6.7.1 Measurement Rig

The prototype was placed on a large planar baffle inside an anechoic chamber (see Fig. 6.6). Absorbing material was placed at the edges of the planar baffle to reduce the reflections caused

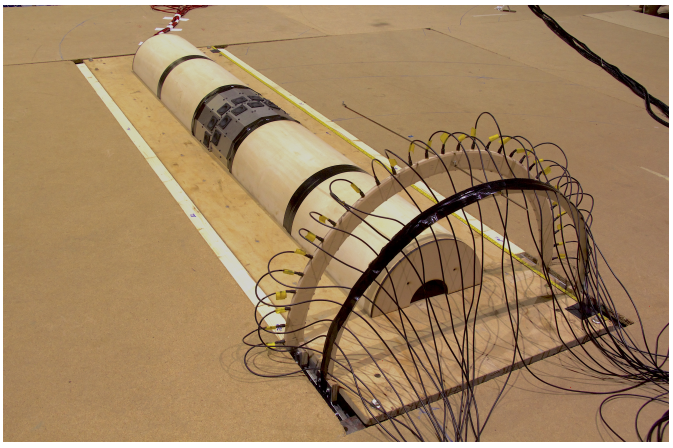


FIGURE 6.6: The measurement setup with a measurement array consisting of 30 *G.R.A.S. 40PL* array microphones fitted on an arch over the HCA with radius $r_H = 0.3\,\mathrm{m}$.

by the locally occurring change of acoustic impedance. The prototype was extended with wooden hemi-cylinders to locally provide conditions similar to those of an infinite hemi-cylinder.

Two sliding rails were fitted underneath the baffle, running parallel to the axis of the HCA on either side. Each held two carriers: one pair to support a wooden arc with radius $r_H = 0.3\,\mathrm{m}$, holding 30 *G.R.A.S. 40PL* array microphones (see Fig. 6.6) and one pair to hold a strain relief feeding the microphone cables to the array along the vertical axis. This arrangement avoids feeding a large number of cables along the microphone arc itself, thereby minimising the acoustic obstruction posed by the measurement arc. This setup allows for iterative scanning of the sound field radiated from the HCA along the vertical axis. The gaps in the baffle exposing the two rails were covered with pieces of timber to mitigate the acoustic impedance change occurring at the gap and thereby reducing reflections.

6.7.2 Angular Sampling and Resolution

The $S_\phi = 30$ microphones were arranged on the arc at the angles

$$\phi_u = \frac{\pi}{S_\phi}\left(u - \frac{1}{2}\right),\ u = 1 \ldots S_\phi, \tag{6.65}$$

yielding a uniform sampling pattern. Given the nature of the angular basis functions $\Phi_n(\phi)$ with the fixed phase relation, it follows from the sampling theorem that this microphone arrangement

allows for the recovery of orders up to

$$n = N_M = S_\phi - 1 = 29.$$

However, Figure 6.3 suggests that, up to a certain frequency, high angular orders will have already decayed below the noise floor once they have propagated to the measurement aperture given their evanescent nature. To investigate this further, one can look at the magnitude of the sum over the axial component of the radial functions

$$\Gamma_n^{(C)}(r, r_S, \kappa) = \int_{-K_z}^{K_z} \Gamma_n(k_z, r, r_S, \kappa) dk_z \tag{6.66}$$

as a function of both radius and frequency. Figure 6.7 shows the magnitude of $\Gamma_n^{(C)}$ normalised to the modal strength at $r = r_S$. This result suggests it is unlikely that orders higher than $n = 11$

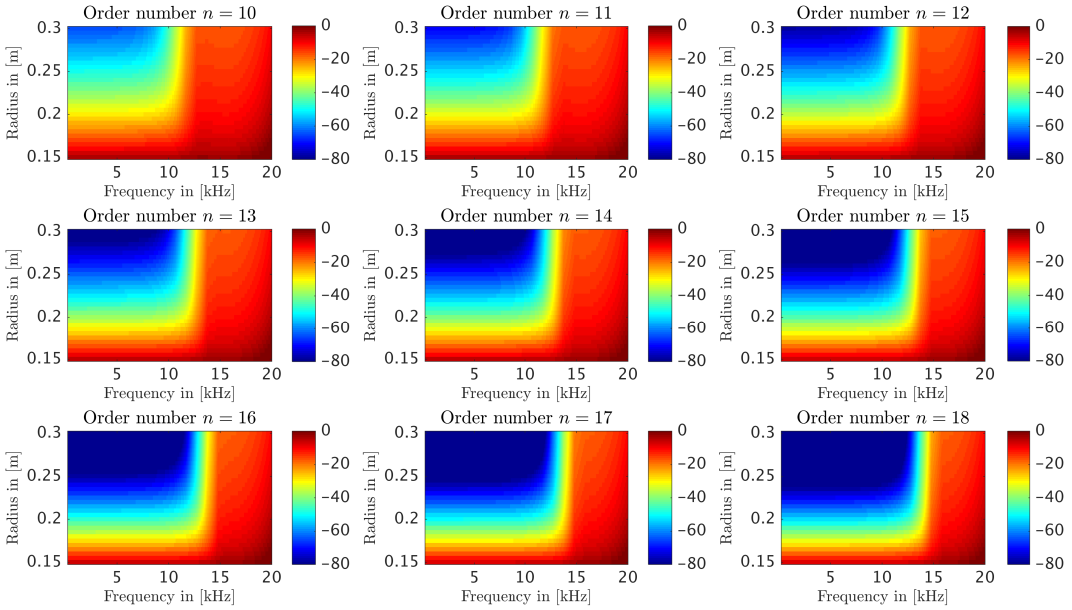


FIGURE 6.7: Modal decay from source surface ($r_S = 0.1426\,\text{m}$) to hologram surface ($r_H = 0.3\,\text{m}$) computed from $\Gamma_N^{(C)}$ within the audible frequency band. The data is normalised to the model strength on the surface of the loudspeaker array.

can be observed above the noise floor when they reach the measurement aperture at $r_H = 0.3\,\text{m}$ for frequencies below 10 kHz.

6.7.3 Vertical Sampling and Resolution

For the vertical extent of the measurement aperture, two different hologram lengths were used. Holograms with a single active driver were measured for $z \in [-0.5, 0.5]\,\text{m}$, covering $A_Z = 1\,\text{m}$ with $S_Z = 51$ samples. Holograms with multiple active drivers were measured from $z \in [-0.98, 0.98]\,\text{m}$ with $S_Z = 99$ samples and an aperture length $A_Z = 1.96\,\text{m}$. In either case, the sampling positions along the z-axis are

$$z_w = \Delta z \left(w - \frac{S_Z + 1}{2} \right), \; w = 1 \ldots S_Z. \tag{6.67}$$

It follows from the sampling theorem that the vertical sample spacing enables capture of axial components k_z up to

$$k_{z,\max} = \frac{\pi}{\Delta z} \approx 157.1 \, \frac{\text{rad}}{\text{m}}. \tag{6.68}$$

Due to the finite hologram aperture along the z-axis, a spatial 8-point Tukey window [Wil99] is applied to the sampled data to attenuate otherwise salient higher spatial frequency components, which would occur due to the discontinuity at the edges of the hologram data. Afterwards, all holograms were zero-padded to an aperture length ranging from $-3\,\text{m}$ to $3\,\text{m}$, so that $A_Z = 6\,\text{m}$ and $S_Z = 301$.

6.7.4 Analysis Equations

The analysis equation for the finite measurement aperture and set of samples is given by

$$\hat{P}_n(k_z, r_H) = \frac{\kappa}{S_\phi} \sum_{u=1}^{S_\phi} \Phi_n(\phi_u) \, \Delta z \sum_{w=1}^{S_Z} Z^*(k_z z_w) \, \hat{p}(\phi_u, z_w), \tag{6.69}$$

where $\hat{p}(\phi_m, z_w)$ is the pressure measured at the mth microphone at the height z_w at the hologram radius $r_H = 0.3\,\text{m}$, where the spatial window has already been applied. Similar to the frequency domain resolution of the Discrete Time Fourier Transform (DTFT), the pressure spectrum $\hat{P}_n(k_z, r_S)$ can be obtained at an arbitrary resolution in the k_z domain from the measurement data $\hat{p}(\phi_m, z_{m'})$. For the acquired holograms, the analysis was performed with a resolution of $\Delta k_z = \frac{2k_{z,\max}}{S_Z}$, matching that of a standard FFT operation.

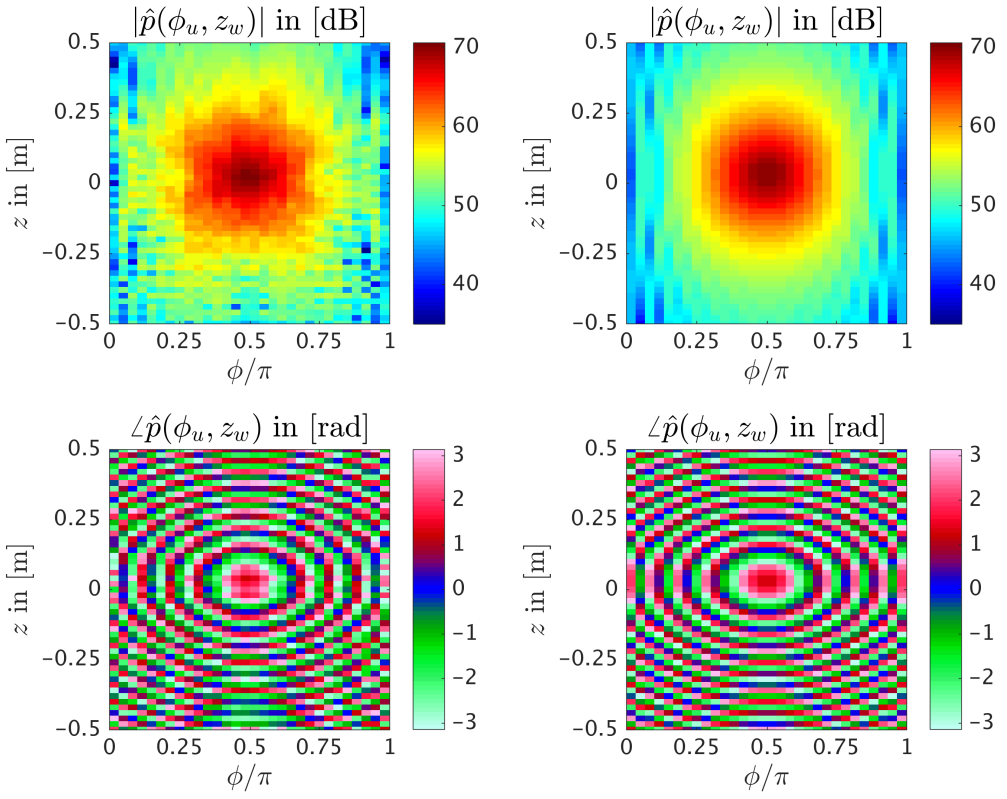
6.7.5 Result for a Single Driver

The measured holograms were not absolutely calibrated, so any given pressure data does NOT represent the actual SPL. To put the measured and the simulated holograms into perspective, the former were normalised to the latter, either at the measurement point closest to the centre of the active driver or at the measurement point closest to the strongest active driver.

Figure 6.8 shows the magnitude and phase of the holograms obtained for the field of a driver/piston located at $(r_S, \frac{\pi}{2}, +0.0315\,\text{m})^T$. Both magnitude and phase show the effects of reflections from the walls at the angles $\phi = 0$ and $\phi = \pi$ in both holograms.

The magnitude of the pressure spectra $\hat{P}_n(k_z, r_H)$ of both holograms are shown in Fig. 6.9. The most striking difference between the measured and the simulated data is the presence of measurement noise in the left plot. Due to the perfectly symmetric position of the driver, the simulated data does not show any contribution of odd angular orders n. This is also reflected in the spectrum from measured data with its prominent contributions for even orders up to approximately $n = 26$. However, it can also be observed that the entire spectrum has an intrinsic noise floor, next to other artefacts which might be attributed to further imperfections in the measurement rig.

Using the pressure spectrum $\hat{P}_n(k_z, r_H)$, the radial component of the velocity spectrum on the boundary of the prototype ($r_S = 0.1426\,\text{m}$) can be obtained from Eq. (6.53). The radial velocity



(A) Measured Hologram (B) Simulated Hologram

FIGURE 6.8: Magnitude and phase of both the measured data and the piston simulation for a driver at $(r_S, \frac{\pi}{2}, 0.0315\,\mathrm{m})^T$, $f = 8\,\mathrm{kHz}$.

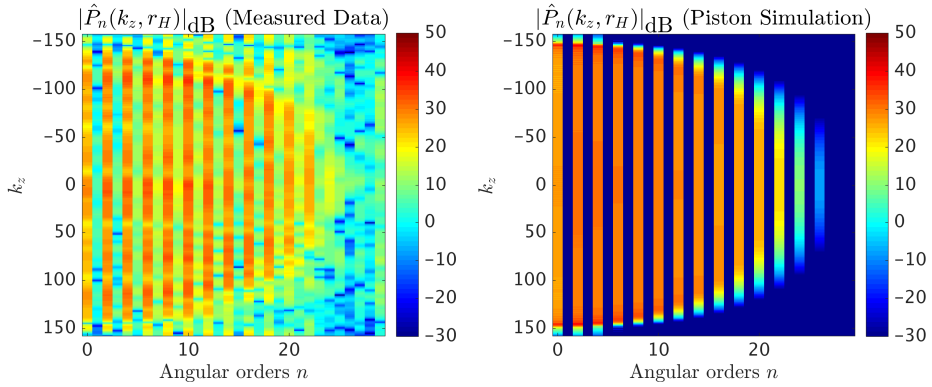
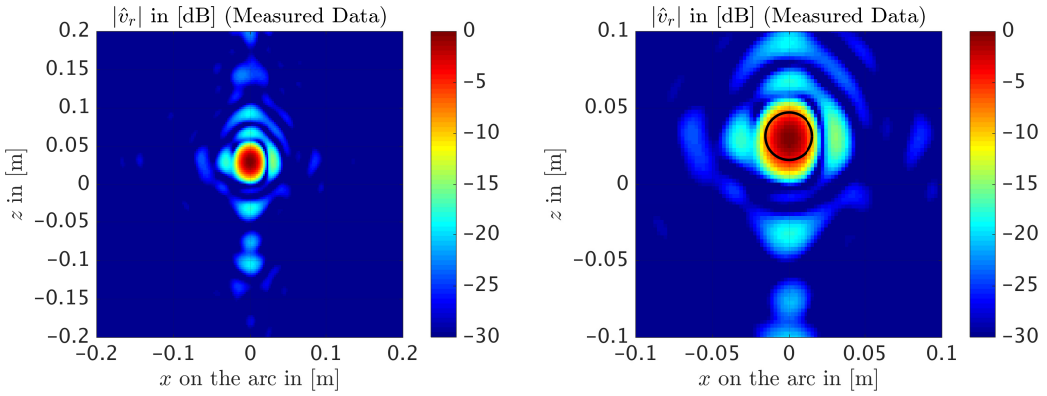


FIGURE 6.9: Spatial spectrum of the measured data vs. simulated measurement.

can be synthesised from numerical evaluation of Eq. (6.55). It was observed from the work with the measured data that the exact knowledge of the hologram radius r_H and the baffle radius r_S is critical when it comes to accurate reconstruction of the spreading of the radiation source along the z-axis.

The resulting radial velocity reconstruction, Fig. 6.10, shows a very clear maximum of the radial



(A) Velocity Reconstruction (B) Detail with Driver Position

FIGURE 6.10: (A) Velocity Reconstruction on the prototype surface, (B) Detailed excerpt with outline of driver, $f = 8\,\text{kHz}$.

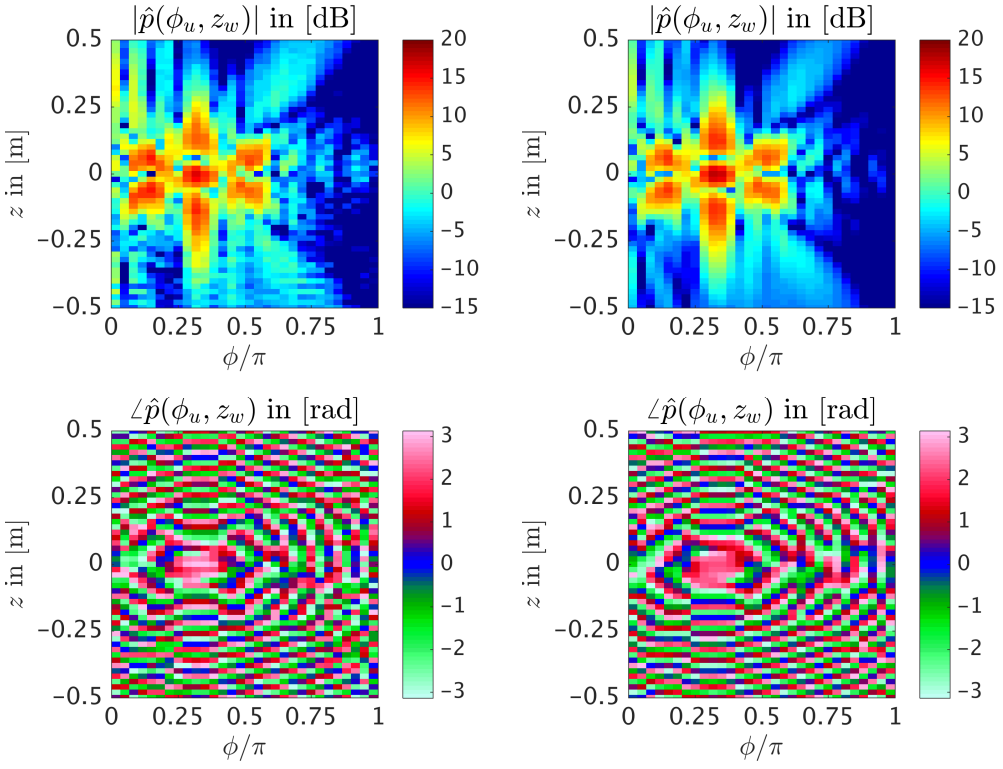
velocity component v_r. Note that the ϕ-axis was translated into a position x along the arc, where

$$x(\phi) = r_S\phi,\ \phi \in [0, \kappa].$$

In the more detailed plot on the right, the overlaid circle indicates the actual driver's position and dimension. It can be observed that the extent of the maximum along the arc (i.e. the ϕ-domain) matches the size of the driver with good approximation, however the spread along the z-axis is slightly larger. This can be explained by the rather large standoff distance $d_S = r_H - r_S \approx 15\,\text{cm}$. A consequence of the latter is that the necessary regularisation effectively reduces the spatial resolution of the reconstruction. The cause of the side lobes occurring next to the area associated to the driver is deemed to be the combination of a relatively short measurement aperture and the aforementioned reduced spatial resolution due to regularisation. If the regularisation could be reduced without introducing the artefacts of amplified measurement noise, then these sidelobes would be decreased and the area of the maximum associated with the driver would be even more distinctly shaped like a circle.

6.7.6 Result for 15 Active Drivers

Figure 6.11 shows both the measured and simulated hologram for all 15 drivers active. The individual source strength and phase was set to radiate a certain beam pattern (see [HFF16]). The interference between the fields radiated from the individual drivers is clearly seen in the beam patterns shown in the magnitude plots of both the measured and simulated holograms. The corresponding spectra are shown in Fig. 6.12. The spectrum from measured data shows the underlying noise floor, as did the spectrum for the single driver previously; However, its overall structure is substantially different, except of course for the distinction of the strength of modes within the radiation 'circle'. When looking at the velocity reconstruction results in Fig. 6.13, it can be seen that not all of the drivers appear active. This can be explained by the strong difference in the level of driving signals needed for the beamforming. Fig. 6.14 shows the source strength of the different drivers. Particularly, the driving signal at drivers 13 to 15 is more than



(A) Measured Hologram (B) Simulated Hologram

FIGURE 6.11: Measured pressure data vs. simulated data from piston model for all 15 drivers active, $f = 8\,\text{kHz}$.

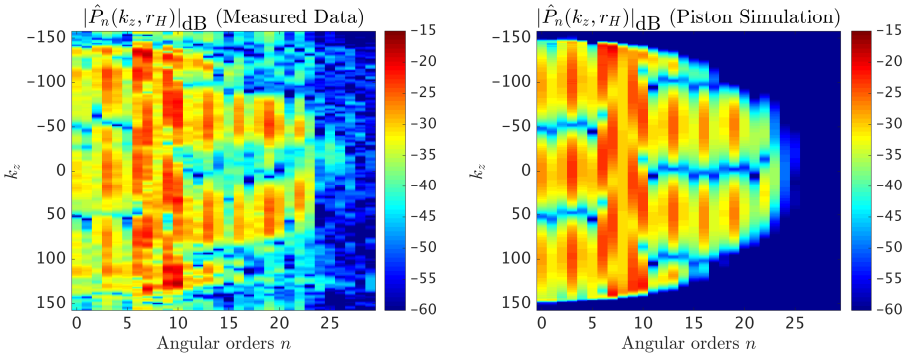


FIGURE 6.12: Spatial spectrum of the measured vs. simulated pressure, $f = 8\,\text{kHz}$.

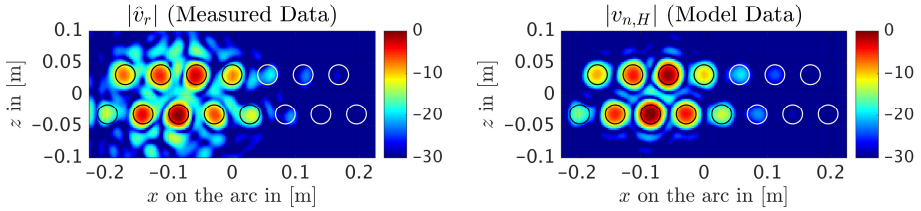


FIGURE 6.13: Reconstruction of v_r on the prototype surface with overlaid circles to indicate the driver positions from both measured (left) and simulated data (right), $f = 8\,\text{kHz}$. Note that the horizontal axis is given as $x(\phi) = \phi r_S$.

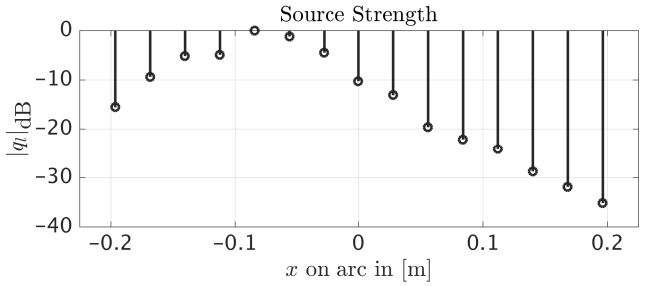


FIGURE 6.14: Normalised driving signal strength of the 15 drivers, $f = 8\,\text{kHz}$.

28 dB below that of the strongest driver. These drivers are thus likely not to be recovered from the hologram data due to noise.

6.7.7 Discussion

From the presented hologram data, it can be seen that there is generally a good match between the measurements and simulations, especially considering that the piston shape of measurement and simulation is slightly different. Physical effects such as the reflections from the rigid baffles at $\phi = 0$ and $\phi = \pi$ can be observed in both holograms, as expected from Section 6.6.

The results of the velocity reconstructions from NAH show an overall convincing match with the expectation. The occurring artefacts might be attributed to a relatively short scanning aperture and a relatively large standoff distance; a smaller hologram radius r_H would provide a better SNR and require less regularisation, which would increase the spatial resolution of the velocity reconstruction at the baffle radius r_S and thus reduce the level of the sidelobes. There is, however, a tradeoff to be made since a reduction of r_H shifts spatial aliasing effects down to lower frequencies as higher order wave components are stronger near the source. Some artefacts may also result from imperfections in the measurement setup, e.g. reflections from sudden impedance changes (transitions between elements of the hemi-cylindrical baffle, trenches housing the rails for the scanning arc and the latter itself, etc.).

The hypothesis made in Section 6.4.2 is supported by the velocity reconstruction results, since the spatial resolution (given the standoff distance) was sufficient to clearly distinguish the driver's diaphragm from the side lobes in the ϕ-domain yet not in the z-domain.

6.8 Summary & Outlook

Acoustic models for the propagation of waves and radiation from sources in infinite and finite wedge-shaped propagation spaces have been presented. From a basic expansion through orthonormal basis functions proposed for both types of propagation spaces, Rayleigh-like formulae as well as corresponding NGFs were derived from a basic expansion. These have been shown to provide simulated data that are well-matched with hologram data obtained from measurements. The identification of radial functions from the Rayleigh-like formulae enabled the formulation of a set of forward and inverse propagators that were then applied to perform NAH on the basis of the measured hologram data. Using the proposed NAH method, reconstructions of the normal

component of the velocity on a hemi-cylindrical baffle obtained from measured holograms were presented to confirm both the developed sound field model and its NAH suitability. A criterion for truncation of the model's series and integrals has been proposed.

It was found from the measurement results that a smaller standoff distance would produce more accurate results. The spatial resolution that remained after regularisation reduced the accuracy of the velocity sources along the vertical axis by comparison to the actual driver layout. A hypothesis suggesting that the angular resolution is less sensitive to an inadequate choice for the standoff distance than the axial resolution was proposed and confirmed from the measurement data.

Chapter 7

Beamforming with Wedge-Shaped Acoustic Arrays

Falk-Martin Hoffmann, Earl G. Williams, Filippo Maria Fazi and Simone Fontana

Abstract

This work derives a 3D sound field model for the pressure field radiated from loudspeaker arrays in acoustic wedge spaces. These wedge arrays are described through their normal velocity profile at a given boundary with constant radius around the origin. Using this model, a suitable beamforming technique with two different beam patterns is derived based on a mode-matching approach. The model is then used to study the beamforming performance as a function of frequency and distance from the array by analysing the newly introduced Arc Radial Functions (ARFs). The theory is tested on the basis of simulation results for different wedges and validated through holographic measurements with a hemi-cylindrical array. It is found that for greater distances, a far-field approximation of the model can be used to drive the array, whereas at short distances, the general model must be applied to yield accurate results.

7.1 Introduction

Loudspeaker arrays have grown to become established means for sound field control applications in various types of environments. Until recently, research was dominated by three principal array configurations: *linear arrays* (e.g. for line arrays in PA applications, 2D Wave Field Synthesis (WFS) etc.) [Ols57, BVV93], *spherical arrays* [DMN03, Pol05a, AS08] and *cylindrical/circular arrays* [KFV10, PB13, PBA15]. Linear arrays are probably the most visible/available to the public, ranging from PA applications to personal audio systems in the form of soundbars. Cylindrical/circular and spherical arrays are often found in the form of large installations for Wave Field Synthesis (WFS) or Higher Order Ambisonics (HOA) systems [DMN03, Pol05a, AS08,

SRA08, AWS10, Faz10, Ahr12], but also in compact form for sound sources with directivity control [ZSH07, MOA+10, KFV10, KFV11, PB13, SFNH14].

The term 'sound field control' can be interpreted in a relatively wide sense. The most classic form is probably the generation of a sound field in a dedicated reproduction area with holophony as the objective [BVV93, DMN03, PBA15]. Acoustic contrast control and pressure matching, as well as acoustic beamforming are also forms of sound field control [KFV10, CJO+13, FSO+14, FSOF15, FC16]. All those techniques can (but do not have to) be derived on the basis of wave propagation models, between which the physical accuracy may vary for a given environment. Two major assumptions are typically made to simplify such models: first, that the array is located in an infinite propagation space and, second, that the transducers are acoustically transparent (i.e. free field conditions). Both assumptions are typically not valid in practice, thus effecting a less than ideal performance of the real system. A good amount of work is published on the modelling of the diffraction as it occurs on spherical and cylindrical boundaries housing transducer arrays [Wil99, Mey01, ME02, Raf05, ZSH07, ZDG10], but also modelling of the reflections caused by the environment has been proposed to improve sound field control in reflective environments [BP14, BZPA15]. Pomberger et al. described the diffraction as it occurs on a rigid sphere in an infinite propagation space with limited range of directions [PZ13, PP14]. A special case of practical importance receiving so far only limited attention is a model for loudspeaker arrays radiating into the half-space.

An initial attempt to model a loudspeaker array arranged on an infinite, planar and rigid baffle was presented in [HFF16]. The proposed hemi-cylindrical arrays were designed to be mounted to an infinite planar baffle (e.g. a wall) and perform beamforming into the half-space. The proposed beamforming method is based on a far field wave propagation model and therefore may not model the near field of the array accurately. It is worth noting that the half-space is in fact a special case of a wedge-shaped propagation space, and thus the hemi-cylindrical array is a special case of a wedge-shaped array. The acoustic environment where such wedge arrays become relevant ranges from in-room applications (e.g. soundbars, PA systems, etc.) to cars or other means of transport, where the propagation space is inherently limited and can be described specifically or more abstractly as a wedge.

This work presents the derivation of a beamforming technique for loudspeaker arrays in infinite wedge propagation spaces [ABC+70]. The wave field models and basis functions for such propagation spaces have already found frequent use in the field of underwater acoustics [Buc84, LZ15] and a work on *Nearfield Acoustical Holography* (NAH) in wedge spaces was recently presented by the author in Chapter 6[1]. Using the existing models, two types of control filters for wedge arrays are proposed and tested in both simulations and experimental work.

The remainder of this work is organised as follows. The second section reviews the model for acoustic radiation from a boundary in wedge propagation spaces and develops a general expression for the velocity produced from an array of loudspeakers located on that boundary. The third section applies the developed array radiation model to solve the inverse problem and calculate the loudspeaker driving functions for a desired pressure field through a mode-matching approach. In that context, the low frequency behaviour and the occurrence of spatial aliasing caused by a finite number of loudspeakers are investigated. In Section four, the two types of beam patterns that are

[1]This manuscript draws on the findings of the (already submitted) manuscript presented in Chapter 6. The version that will be submitted for publication is going to refer to the other manuscript or paper, respectively.

used for the simulations and measurements are introduced. Section five presents and discusses the simulations of radiated beam patterns across frequency for two examples of wedge arrays, where one of them serves to introduce the prototype used for the experiments. The obtained experimental results are described and discussed in Section six, and the final section summarises the findings and gives an outlook on upcoming research.

7.2 Sound Field Model

This work uses the sound field model for the radiation from a velocity distribution on a curved baffle V around the origin of an *infinite wedge* propagation space, as presented in Chapter 6. Note that all considerations are made for the (quasi-)stationary case, where the factor $e^{-i\omega t}$ and the different quantities' dependency on ω was omitted for reasons of brevity.

The geometrical description of the infinite wedge is based on the cylindrical coordinate system, as given in Figure 7.1. A projection of the propagation space onto the xy-plane with the added

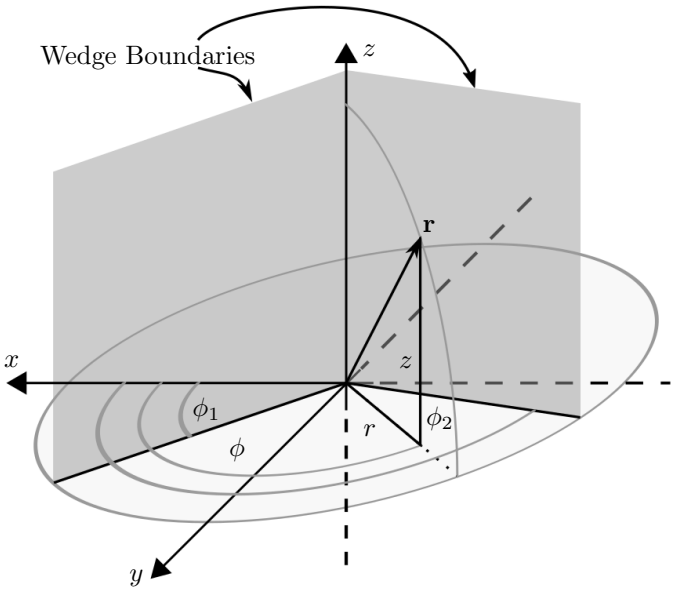


FIGURE 7.1: Propagation space with wedge boundaries described in cylindrical coordinates.

curved baffle V holding the loudspeaker array is shown in Figure 7.2.

The set of points *within* the propagation space W is uniquely defined through

$$W := \left\{ \mathbf{r} = (r, \phi, z)^T : r \geq r_S, \phi \in [\phi_1, \phi_2], z \in \mathbb{R} \right\}, \tag{7.1}$$

where r_S is the radius of the curved baffle around the origin. The wedge angle $\kappa = \phi_2 - \phi_1$ can take any value from the interval $[0, 2\pi]$. The boundaries of the propagation space comprise the curved baffle

$$V := \{ \mathbf{r} : r = r_S,\ \phi \in [\phi_1, \phi_2],\ z \in \mathbb{R} \}$$

and the wedge's flanks

$$S' := \{ \mathbf{r} : r > r_S,\ \phi = \phi_1, \phi_2,\ z \in \mathbb{R} \}.$$

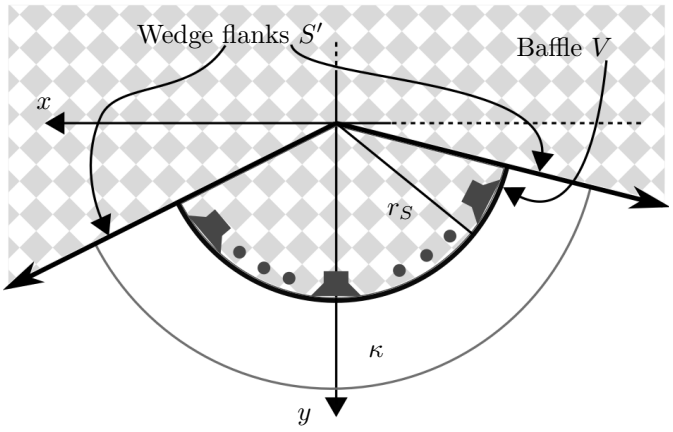


FIGURE 7.2: Wedge boundaries projected into the xy-plane with the loudspeaker array mounted on the curved baffle.

In the case of wedge arrays, all loudspeakers, i.e. all active acoustic sources, are located on V. Accordingly, the acoustic velocity on V is defined by the driving functions and the physical dimensions of the transducers. Without loss of generality but to simplify notation, it is assumed in the remainder of this work that $\phi_1 = 0$ and $\phi_2 = \kappa$.

The following subsection describes the mathematical model for the pressure radiated by a wedge array.

7.2.1 General Pressure Model

The general solution for the pressure field radiated from a given velocity distribution on V inside an infinite wedge with rigid boundaries was presented in Chapter 6 and is given by the expression

$$p(r,\phi,z) = \sum_{n=0}^{\infty} \Phi_n(\phi) \int_{-\infty}^{\infty} \Gamma_n(k_z,r) V_n(k_z,r_S) Z(k_z z)\, dk_z, \tag{7.2}$$

with

$$\Gamma_n(k_z,r) = i\rho_0 ck \frac{H_{\frac{n\pi}{\kappa}}(k_r r)}{k_r H'_{\frac{n\pi}{\kappa}}(k_r r_S)} \tag{7.3}$$

$$\Phi_n(\phi) = \nu_n \cos\left(\frac{n\pi}{\kappa}\phi\right), \tag{7.4}$$

$$Z(k_z z) = \frac{e^{ik_z z}}{\sqrt{2\pi}}, \tag{7.5}$$

where

$$k_r = \sqrt{k^2 - k_z^2} \quad \text{and} \quad \nu_n = \sqrt{\frac{2-\delta_n}{\kappa}},$$

κ is the wedge angle, k represents the acoustic wave number and δ_n denotes the Kronecker Delta. $H_{\frac{n\pi}{\kappa}}(x)$ refers to the Hankel function of the first kind[2] of order $\frac{n\pi}{\kappa}$, where the latter

[2]The superscript $^{(1)}$ commonly used to distinguish the Hankel functions of the first from that of the second kind is omitted for reasons of brevity.

is exclusively *integer* only for $\kappa = \pi$. The angular basis functions $\Phi_n(\phi)$ and the axial basis functions $Z(k_z z)$ are the elements of two Total Orthonormal Sets (TOS) [Kre78]. Their product creates a two-dimensional TOS. $\Gamma_n(k_z, r)$ denotes the corresponding radial functions for the radiation problem in the wedge space (see Chapter 6, Section 6.4), where the dependency on κ and r_S was omitted for the sake of brevity. The same applies for the angular basis functions' dependency on κ and ϕ_1.

The spatial velocity spectrum $V_n(k_z, r_S)$ in (7.2) is calculated as the spatial transform of the normal velocity (or vibration) profile $v(r_S, \phi, z)$ describing the particle displacement in the normal direction on the boundary V (see Chapter 6)

$$\begin{aligned} V_n(k_z, r_S) &= \mathcal{W}\left[v(r_S, \phi, z)\right] \\ &= \int_0^{\kappa} \int_{-\infty}^{\infty} v(r_S, \phi, z) \Phi_n(\phi)\, Z^*(k_z z)\, dz\, d\phi, \end{aligned} \tag{7.6}$$

where $(\cdot)^*$ denotes the complex conjugate and the operator $\mathcal{W}$ abbreviates the spatial transform. Inversely, $v(r_S, \phi, z)$ is obtained from $V_n(k_z, r_S)$ through the inverse operator

$$\begin{aligned} v(r_S, \phi, z) &= \mathcal{W}^{-1}[V_n(k_z, r_S)] \\ &= \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} V_n(k_z, r_S) \Phi_n(\phi)\, Z(k_z z)\, dk_z. \end{aligned} \tag{7.7}$$

Equations (7.6) and (7.7) define a transformation pair that may be regarded as generalised Fourier transform.

With the general sound field model at hand, the sound field control problem can now be theoretically solved by taking a mode-matching approach that yields a relation between the pressure spectrum

$$P_n(k_z, r) = \mathcal{W}[p(r, \phi, z)], \tag{7.8}$$

of a desired pressure field $p(r, \phi, z)$ at the radial coordinate r and the velocity spectrum $V_n(k_z, r_S)$ on V.

7.2.2 The General Solution to the Inverse Problem

Finding the required vibration profile v on V that yields a specific target pressure field p is an inverse problem that needs solving.

Applying the inverse operator $\mathcal{W}^{-1}[\cdot]$ defined in Eq. (7.7) to both sides of Eq. (7.2) yields

$$P_n(k_z, r) = \Gamma_n(k_z, r) V_n(k_z, r_S). \tag{7.9}$$

Equation (7.9) can be transformed into an expression for $V_n(k_z, r_S)$

$$V_n(k_z, r_S) = \frac{1}{\Gamma_n(k_z, r)} P_n(k_z, r), \tag{7.10}$$

if $k > 0$. The required normal velocity on the boundary V can then be obtained by replacing the velocity spectrum in (7.7) with the result in (7.10). From this it follows that a vibrating baffle

that is infinite along the z-axis is required to enable the reconstruction of *any* desired pressure field $p(r, \phi, z)$ that satisfies the homogeneous Helmholtz equation in W.

Controlling the vibration profile on the entire surface V is impractical. Therefore this work focusses on sound field control inside an arbitrary plane $z = z_B$ using an array with a finite number of loudspeakers arranged on an arc. These are henceforth referred to as wedge arrays. Before presenting the methods to identify the individual loudspeakers' driving functions, a general expression for the velocity spectrum of arbitrary wedge arrays is defined in the next subsection.

7.2.3 Velocity Spectrum of an Arbitrary Wedge Array

Let a wedge array be composed of L loudspeakers. Each speaker unit is described through the location of its geometrical centre $(r_S, \phi_l, z_l)^T \in V$ and its diaphragm's velocity profile $v_l(r_S, \phi, z)$. The index l is used to refer to the lth speaker in the array. In general, the corresponding velocity spectra, $V_n^l(k_z)$, are determined through the operator $\mathcal{W}$ defined in (7.6).

It shall be assumed that the diaphragm is a rectangular piston (see Chapter 6, Section 6.6.2) of angular width Θ and vertical length B. The corresponding velocity profile of the lth loudspeaker is then given by

$$v(r_S, \phi, z) = \Pi\left(\frac{\phi - \phi_l}{\Theta}\right) \Pi\left(\frac{z - z_l}{B}\right),$$

where $\Pi(\cdot)$ denotes the rectangular function. Applying $\mathcal{W}[\cdot]$ to the above equation yields the diaphragm's velocity spectrum

$$V_n^l(k_z) = V_n^D(k_z)\Phi_n(\phi_l)\, Z^*(k_z z_l) \tag{7.11}$$

with

$$V_n^D(k_z) = \Theta\, \mathrm{sinc}\left(\frac{n\pi}{\kappa}\frac{\Theta}{2}\right) B\, \mathrm{sinc}\left(k_z \frac{B}{2}\right). \tag{7.12}$$

The explicit dependency of $V_n^l(k_z)$ on r_S has been omitted for the sake of brevity. The expression in (7.11) was deliberately separated into a factor defining the shape of the diaphragm, $V_n^D(k_z)$, and two factors defining the position. Although this form may suggest otherwise, the general shift property is not valid for the transform $\mathcal{W}[\cdot]$.

For the velocity spectrum of the full wedge array it is convenient to assume that all loudspeakers have the same shape, i.e. $V_n^D(k_z)$ is equal for all L loudspeakers. This step is not strictly necessary in general, but it serves to simplify notation. Assuming linearity, the superposition principle enables the description of the overall velocity spectrum, $V_n(k_z, r_S)$, as a linear combination of the loudspeakers' velocity spectra $V_n^l(k_z)$. Weighting each of the latter with a driving function q_l yields

$$V_n(k_z, r_S) = V_n^D(k_z) \sum_{l=1}^{L} q_l \Phi_n(\phi_l)\, Z^*(k_z z_l). \tag{7.13}$$

The triple (q_l, ϕ_l, z_l) defines each loudspeaker's driving function, angular position and axial position, respectively. The specific form of (7.13) will proof useful in Section 7.3.

The result for $V_n(k_z, r_S)$ in (7.13) can now be used to formulate the array-specific forward problem in the modal domain.

7.2.4 The Array-Specific Forward Problem

Replacing $V_n(k_z, r_S)$ in (7.9) by the expression in (7.13) provides the relation of the array's driving functions, q_l, to the radiated pressure spectrum

$$P_n(k_z, r) = \Gamma_n(k_z, r) V_n^D(k_z) \sum_{l=1}^{L} q_l \Phi_n(\phi_l)\, Z^*(k_z z_l) \tag{7.14}$$

for $n \in \mathbb{N}$ and $k_z \in \mathbb{R}$. This poses a highly overdetermined equation system since the L driving functions q_l are to control an infinitely long vector of constants given by $P_n(k_z, r)$. Such a system may have a solution, but it cannot be solved exactly by a least-squares approach, unless one were to truncate the range of n and to sample and truncate the range of k_z; In that case one would obtain an approximation of the target pressure.

The modal combinations (n, k_z) falling into the region outside the "radiation circle", defined by $k \leq \sqrt{k_z^2 + (\frac{n\pi}{\kappa r_S})^2}$ [Wil99], are evanescent in their nature. Therefore their corresponding magnitudes $|P_n(k_z, r)|$ are subject to a steep decay as the radial coordinate r increases. This physical effect is mathematically described in Eq. (7.14) through to the magnitude of the radial functions $\Gamma_n(k_z, r)$. As a consequence, mostly coefficients $P_n(k_z, r)$ from within, or marginally outside the "radiation circle" need to be considered for a good approximation of the desired sound field. Therefore, limiting the range of n and k_z may not have a severely aggravating effect on the result. However, due to the *uncountable* nature of the elements within the interval along the k_z-axis, the vector formed by the $P_n(k_z)$ would still have an infinite number of entries, hence necessitating sampling of k_z for a least-squares solution.

The considerations made during the course of this section suggest that full control of the sound field inside the infinite wedge cannot be achieved with a finite set of transducers arranged within a finite area on V. This situation changes when the objective is simplified to controlling the sound field on an arc with radius r_B in the plane $z = z_B$ with loudspeakers arranged on the intersection of V with the plane $z = z_S$. In that case, the driving functions q_l can be identified through a mode-matching approach, which is shown in the following section.

7.3 Loudspeaker Driving Functions

7.3.1 Diaphragm Velocity as a Function of Driving Voltage

The driving functions derived in this section relate directly to the individual loudspeaker's *velocity*. When operating a classical electrodynamic transducer *above* its resonance frequency as a loudspeaker connected to an (ideal) voltage source, the applied voltage does not generate a proportional velocity v of the diaphragm but a proportional acceleration $\dot{v}$ instead (compare [BX08] p. 61, Figure 5.7). In the frequency domain, the relation between voltage u and velocity v is therefore given by

$$\frac{u(\omega)}{\omega} \sim v(\omega), \tag{7.15}$$

which corresponds to a integration in the time domain. It is therefore necessary to amend the driving signals $q_l(\omega)$ appropriately when operating an array built from electrodynamic

loudspeakers through a conventional power amplifier. Note that Equation (7.15) is only an approximation where it is assumed that the electrical impedance is predominantly resistive and that the mechanical impedance is dominated by the moving mass.

7.3.2 Model for the Mode-Matching Approach

From the completeness relation of TOSs [Kre78], it follows that an arbitrary target pressure field on the arc

$$\Psi := \{\mathbf{r} : r = r_B,\ \phi \in [0, \kappa],\ z = z_B\}$$

can be expressed as the infinite series expansion

$$p(r_B, \phi, z_B) = \sum_{n=0}^{\infty} P_n \Phi_n(\phi)\,. \tag{7.16}$$

To find a solution for the $L < \infty$ driving functions q_l dependent on the target coefficients P_n, the driving functions need first be expressed as a finite series expansion

$$q_l = \sum_{m=0}^{N} Q_m \Phi_n(\phi_l)\,,\ \forall l \in [1, L],\ N = L - 1. \tag{7.17}$$

The above expression poses a linear system of equations. Due to linear independence of $\Phi_n(\phi)$ for different orders, the system matrix has full rank and a unique solution for the coefficients Q_m exists, if all L angles ϕ_l are distinct.

Applying Eq. (7.17) in (7.13) and rearranging the double sum yields the expression

$$V_n(k_z, r_S) = V_n^D(k_z) \sum_{m=0}^{N} B_{n,m}(k_z) Q_m, \tag{7.18}$$

with

$$B_{n,m}(k_z) = \sum_{l=1}^{L} \Phi_m(\phi_l)\, \Phi_n(\phi_l)\, Z^*(k_z z_l)\,. \tag{7.19}$$

Equation (7.18) describes the modal relations between the velocity spectrum $V_n(k_z)$ and the coefficients Q_m for the most general case. The function $B_{n,m}(k_z)$ is determined by the positions of the loudspeakers, (ϕ_l, z_l). It is shown in the next section how a convenient choice of the latter is useful for the mode-matching approach.

7.3.3 Loudspeaker Positioning

The loudspeaker positions are now chosen as uniformly distributed along the intersection of V and $z = z_S$, i.e.

$$\phi_l = \frac{\kappa}{L}\left(l - \frac{1}{2}\right) \quad \text{and} \quad z_l = z_S,\ \forall l \in [1 \ldots L]. \tag{7.20}$$

This particular choice for ϕ_l simplifies the expression in Eq. (7.19) so that

$$B_{n,m}(k_z) = Z^*(k_z z_S)\,\frac{L}{\kappa} A_{n,m}, \tag{7.21}$$

$$\text{with} \quad A_{n,m} = \frac{\kappa}{L} \sum_{l=1}^{L} \Phi_m(\phi_l)\,\Phi_n(\phi_l)\,. \tag{7.22}$$

The orthogonality matrix, also referred to as the spatial aliasing matrix, is given through the coefficients

$$A_{n,m} = \sqrt{\sum_{u=0}^{\infty} \delta_{n-m-u2L} + \delta_{n+m-(u+1)2L}}. \tag{7.23}$$

Using the result of (7.21) in Eq. (7.18) and replacing the result for $V_n(k_z)$ in (7.2) yields the expression

$$p(r_B, \phi, z_B) = \frac{L}{\kappa} \sum_{n=0}^{\infty} \sum_{m=0}^{N} \Phi_n(\phi)\, A_{n,m} Q_m$$
$$\cdot \int_{-\infty}^{\infty} \Gamma_n(k_z, r_B) V_n^D(k_z) Z^*(k_z z_S)\, Z(k_z z_B)\, dk_z. \tag{7.24}$$

In order to further simplify this result and understand why it was convenient for all loudspeakers to have the same diaphragm and the chosen positions, it is necessary to take a closer look at the coefficients $A_{n,m}$.

7.3.4 Orthonormality and Aliasing

An example of the coefficients specified by (7.23) for $L = 15$, $n = 0 \ldots 33$ and $m = 0 \ldots N = 6$ is presented in Figure 7.3. It shows that the chosen positions of the L loudspeakers preserve a

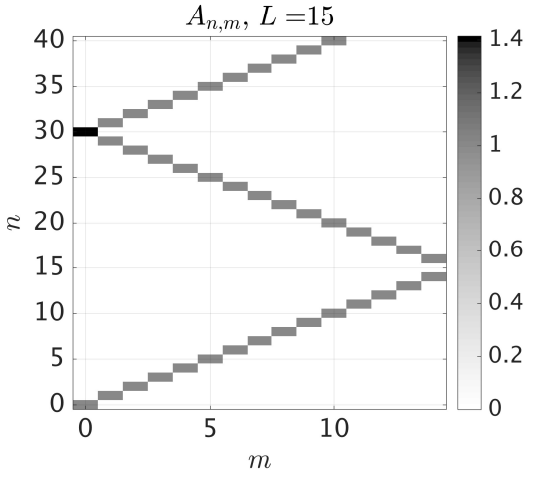


FIGURE 7.3: Visualisation of the orthogonality matrix, $L = 15$.

perfect orthonormality relation of the modes $\Phi_n(\phi)$ for $n = 0 \ldots N$. It also shows, however, that higher orders n are subject to aliasing from the orders $m = 0 \ldots N$, which can be predicted from (7.23). Note that both the pattern and thus also the expression on the right hand side of (7.23)

would likely have been much more complicated if not for the assumption that all loudspeakers are identical as specified in Section 7.2.3.

The analysis of the coefficients $A_{n,m}$ shows that the finite number of loudspeakers can still excite an infinite number of angular modes in the pressure field, even though they only seek to control the lowest L modes $\Phi_n(\phi)$. The excitation of modes of higher order then manifests itself in the form of spatial aliasing artefacts in the reproduced field.

As discussed in Section 7.2.3, not all angular modes actually propagate to the far field. Those that are evanescent decay rapidly away from the origin, making their contribution to the pressure on Ψ potentially insignificant. This is further investigated in the subsequent section.

7.3.5 The Arc Radial Functions

The integral and the leading factor of Equation (7.24) can be unified to form a new quantity that shall be denoted as Arc Radial Functions (ARFs)

$$\Gamma_n^{\Psi}(r_B, z_B) = \frac{L}{\kappa 2\pi} \int_{-\infty}^{\infty} \Gamma_n(k_z, r_B) V_n^D(k_z) e^{ik_z(z_B - z_S)} dk_z. \tag{7.25}$$

Together they describe how the modes excited by the specific set of driving functions q_l propagate to the arc Ψ. The expression for the pressure on the arc simplifies to the infinite expansion of the weighted modes $\Phi_n(\phi)$ and reads as

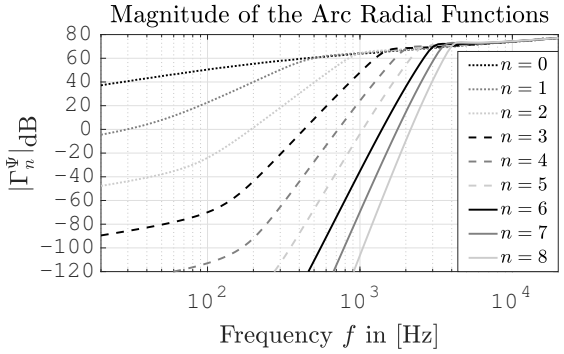
$$p(r_B, \phi, z_B) = \sum_{n=0}^{\infty} \sum_{m=0}^{N} \Phi_n(\phi)\, Q_m A_{n,m} \Gamma_n^{\Psi}(r_B, z_B). \tag{7.26}$$

The coefficients Q_m do not directly determine the modal strength of $\Phi_n(\phi)$ for $n = m$. Instead, the ARF describe how the nth angular mode excited on the source arc couples to the corresponding mode on the control arc Ψ. Which angular modes on the source arc get excited by the source distribution is determined by the coefficients $A_{n,m}$.

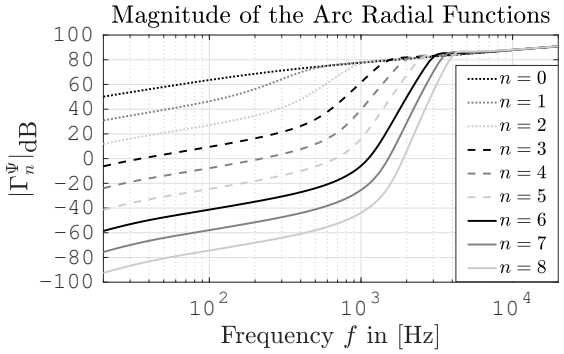
Figure 7.4 shows the development of the ARFs for different combinations of r_S and r_B with $\kappa = \frac{\pi}{2}$ for the velocity spectrum $V_n^D(k_z)$ of a point source. It can be observed that as frequency increases, eventually all depicted orders propagate equally well. Naturally the overall attenuation depends on the propagation distance r_B.

At low frequencies ($f < 3\,\mathrm{kHz}$) it can be seen that the higher its order the more a mode gets attenuated due to evanescence. The frequencies $f_n^C = \frac{n\pi c}{2\pi\kappa r_S}$ indicate where a mode of order n changes from propagating to evanescent behaviour. The expression for the edge frequency f_n^C was derived from the $n = kr$ rule that determines the transition from propagating to evanescent components [Wil99, WA01]. For the wedge, the rule reads $\frac{n\pi}{\kappa} = kr$. The dependency on r_S explains why the f_n^C shift to lower frequencies in Fig. 7.4c, in comparison to Fig. 7.4a.

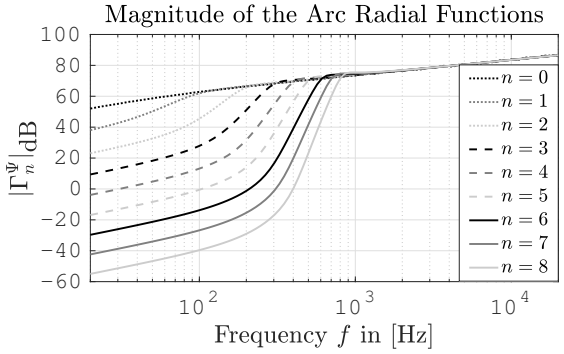
As frequency decreases further from f_n^C, there occurs yet another transition frequency $f_n^{C,\mathrm{NF}}$ where the slope becomes less steep. Where these 'edges' occur appears to depend on r_B, as the comparison of Fig. 7.4a and Fig. 7.4b suggests. These edges indicate a mode's transition from the far field to the near field of the radiating surface V. Comparison of all plots in Fig. 7.4 suggests that the transition frequency $f_n^{C,\mathrm{NF}}$ is inversely proportional to r_B. It follows implicitly



(A) $r_B = 2\,\mathrm{m}$, $r_S = 0.2\,\mathrm{m}$



(B) $r_B = 0.5\,\mathrm{m}$, $r_S = 0.2\,\mathrm{m}$



(C) $r_B = 2\,\mathrm{m}$, $r_S = 1\,\mathrm{m}$

FIGURE 7.4: Magnitude of Γ_n^Ψ, $\kappa = \frac{\pi}{2}$, $L = 15$, $z_S = z_B = 0\,\mathrm{m}$.

for the given orders $n = 0 \ldots N$ that the smaller the value of r_B the wider the frequency range where the corresponding modes demonstrate near field effects.

From the comparison of Fig. 7.4a and Fig. 7.4c, it can be deduced that a larger array can theoretically control the sound field it radiates with less effort at low frequency than a smaller array.

Figure 7.5 shows the ARF for $\kappa = \pi$. The transition frequencies f_n^C shift towards lower frequencies for the wider wedge angle, as can be expected. The discrepancy between the magnitudes of two consecutive evanescent modes is only half of that which can be seen for the case of $\kappa = \frac{\pi}{2}$. This had to be expected since the order $\frac{n\pi}{\kappa}$ of the Hankel functions contributing to Γ_n effectively increases twice as fast for $\kappa = \frac{\pi}{2}$ than for $\kappa = \pi$ (see Eq. (7.3)).

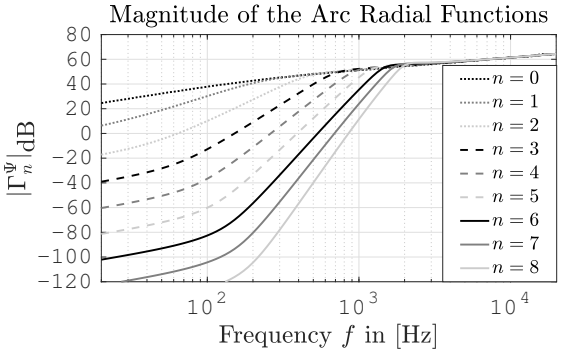


FIGURE 7.5: Magnitude Γ_n^Ψ with $\kappa = \pi$, $L = 15$, $r_B = 2\,\text{m}$, $r_S = 0.2\,\text{m}$, $z_S = z_B = 0\,\text{m}$.

The integral in (7.25) was calculated numerically using the procedure described in Chapter 6, Section 6.6. The bounds were chosen as $-K_z$ and K_z, where

$$K_z = \left\lceil -\frac{Qk}{\Delta k_z} \right\rceil \Delta k_z$$

with a step size $\Delta k_z = 0.01$. The parameter Q adjusts the range k_z and is chosen so that the bounds include all components for $n = 0$ that satisfy $20\log_{10}(|\Gamma_0(k_z, r_B)|) > -150\,\text{dB}$. Note that $\lceil\cdot\rceil$ represents the ceiling function.

While numerical integration may be one way to calculate the ARFs with very good approximation, it is computationally expensive. The SPA of the expression in (7.25) (see Appendix 7.8) yields the far field approximation for the ARF

$$\Gamma_{n,\text{FF}}^\Psi(R,\theta) \approx \frac{\rho_0 c L}{\kappa\pi} e^{i\frac{n\pi^2}{2\kappa}} \frac{e^{ikR}}{R} \frac{V_n^D(k\cos\theta)e^{-ikz_S\cos\theta}}{\sin\theta H'_{\frac{n\pi}{\kappa}}(k\sin\theta r_S)}. \tag{7.27}$$

The SPA of $\Gamma_n^\Psi(r_B, z_B)$ provides a result with good accuracy when r_B corresponds to the far field. Its expression has the advantage over Eq. (7.25) that it is given in analytic form and is thus fast to compute. It can be used alternatively in (7.26).

The difference between the ARFs and their far field approximations when evaluated in the near field becomes clear when comparing the results of Fig. 7.4a and Fig. 7.6. At high frequencies

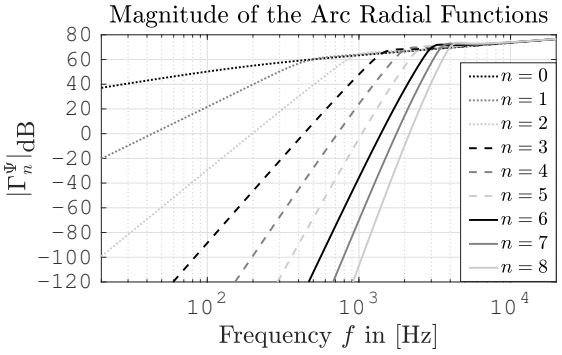


FIGURE 7.6: Magnitude of $\Gamma_{n,\text{FF}}^\Psi$, $\kappa = \frac{\pi}{2}$, $L = 15$, $R = r_B = 2\,\text{m}$, $\theta = \frac{\pi}{2}$, $z_S = 0\,\text{m}$.

there is hardly any difference, except for the small error of the numerical integration process. Below the edge frequencies f_n^C, the magnitudes of $\Gamma_{n,\text{FF}}^\Psi$ decrease monotonically with frequency,

unlike the values of Γ_n^{Ψ} that change decay once more at $f_n^{C,\mathrm{NF}}$. This difference has an impact on the array performance at low frequencies, as will be shown in Section 7.3.7.

Since the edge frequencies f_n^C do not change between the ARF and their far-field approximations, choosing one or the other for the calculation of the driving functions has no influence on the array's spatial aliasing performance.

7.3.6 Truncation of Higher Order Modes

For a given frequency, both types of arc radial functions, Γ_n^{Ψ} and $\Gamma_{n,\mathrm{FF}}^{\Psi}$, suggest that there is a mode of minimum order n_E that is the first to be evanescent. This mode and all subsequent thus significantly lose their contribution to the overall sound field, as they propagate. It can thus be assumed that there is an upper edge frequency where a truncation of the modelling order to $n = N$ has a negligible effect on the pressure $p(r_B, \phi, z_B)$. Such a truncation applied to the expression in (7.26) yields the approximation

$$p(r_B, \phi, z_B) \approx \sum_{n=0}^{N} \Gamma_n^{\Psi}(r_B, z_B) Q_n(\omega) \Phi_n(\phi) = \sum_{n=0}^{N} P_n \Phi_n(\phi)\,, \tag{7.28}$$

which itself can be described comprehensively through a set of coefficients $\{P_n\}_{n=0}^{N}$. Comparing factors between the two expressions in (7.28) and reorganising for Q_n yields

$$Q_n = \frac{1}{\Gamma_n^{\Psi}(r_B, z_B)} P_n. \tag{7.29}$$

Eq. (7.29) provides the driving function coefficients Q_n as a function of the coefficients P_n that define the desired pressure field on the arc Ψ. The values of $\Gamma_n^{\Psi}(r_B, z_B)$ describe how the different modes Q_n of the loudspeaker driving function expansion in (7.17) radiate into the pressure field. It was shown in Section 7.3.5 that the arc radial functions can take values of very small magnitude at low frequencies. Their reciprocals in (7.29) then take very large values, effecting very high power demands for the individual loudspeakers [HFF16]. In practice it is therefore necessary to apply regularisation to the inverse solution in (7.29) to protect the mechanical integrity of the loudspeakers [FSOF15] and, more generally, to ensure robustness and stability.

7.3.7 Regularised Driving Functions

Using a normalised Tikhonov regularisation, the inverse solution is reformulated to

$$\tilde{Q}_n = \frac{\Gamma_n^{\Psi *}(r_B, z_B)}{|\Gamma_n^{\Psi}(r_B, z_B)|^2 + \beta\sigma^2} P_n, \tag{7.30}$$

where $\sigma = \max_n |\Gamma_n^{\Psi}(r_B, z_B)|$. The parameter β adjusts the tradeoff between low-frequency directivity and maximum displacement of the loudspeakers' diaphragms, where a limitation of the latter improves the robustness of the system [FSOF15]. Using (7.30) in Eq. (7.17) provides the regularised individual loudspeaker driving functions

$$q_l = \sum_{n=0}^{N} \frac{\Gamma_n^{\Psi *}(r_B, z_B)}{|\Gamma_n^{\Psi}(r_B, z_B)|^2 + \beta\sigma^2} P_n \Phi_n(\phi_l) \tag{7.31}$$

for a desired pressure profile specified through the set $\{P_n\}_{n=0}^{N}$. The fact that r_B remains a parameter makes for the main difference to the driving functions presented in [HFF16], where the (arc) radial functions were derived from a plane wave scattering model through reciprocity and hence cannot be adjusted to a particular control radius. If r_B is large enough to be in the far field for all frequencies of interest, then there is no difference between using either Γ_n^{Ψ} or $\Gamma_{n,\mathrm{FF}}^{\Psi}$ for the calculation of q_l. The difference in performance of the driving functions when using the two types of ARFs is investigated in Section 7.6.

7.4 Beamforming

As with other geometric array arrangements, beamforming with wedge arrays aims to control the radiated pressure field in such a way that the pressure magnitude on the control arc Ψ resembles a designated beam pattern.

In this work, modal beamforming with two different types of beam patterns are considered. The first corresponds to a Dirac pulse in the ϕ-domain, and the second corresponds to a Rectangular pulse of variable width D. Both are approximated through finite expansions of the form

$$f(\phi) = \sum_{n=0}^{N} F_n \Phi_n(\phi)\,, \tag{7.32}$$

so that the two beam patterns differ with respect to their coefficients F_n. These can be calculated through

$$F_n = \mathcal{W}[f(\phi)\delta(z)]. \tag{7.33}$$

To ensure a unit magnitude in the steering direction ϕ_B, it is necessary to normalise these coefficients by the factor $\sum_{m=0} F_m \Phi_m(\phi_B)$. The exact expressions for the coefficients used in this work are presented below.

With the loudspeaker arrays under consideration, there is of course no limitation to modal beamforming. Other far-field beamforming methods such as maximum directivity or maximum White Noise Gain (WNG) [RPA+10, RK11, JHN17] are in principle applicable to the given array and propagation space geometry.

7.4.1 Beam Pattern I

The coefficients F_n^I of the beam pattern corresponding to a Dirac pulse are given by

$$F_n^I = \frac{\Phi_n(\phi_B)}{\sum_{m=0}^{N} \Phi_m(\phi_B)^2}. \tag{7.34}$$

Some examples of beam pattern I with $N = 10$ are shown in Figure 7.7 for values of ϕ_B.

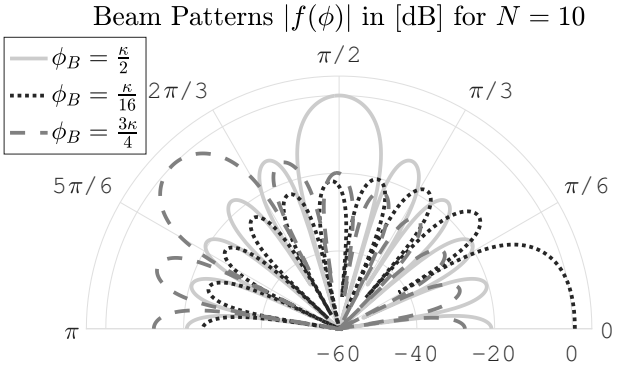


FIGURE 7.7: Examples for Beam Pattern I with different steering angles ϕ_B, $\kappa = \pi$.

7.4.2 Beam Pattern II

The coefficients F_n^{II} of the beam pattern corresponding to a Rectangular pulse of width D are given by

$$F_n^{II} = \frac{D\text{sinc}\left(\frac{n\pi}{\kappa}\frac{D}{2}\right)\Phi_n(\phi_B)}{\sum_{m=0}^{N} D\text{sinc}\left(\frac{m\pi}{\kappa}\frac{D}{2}\right)\Phi_m(\phi_B)^2}. \tag{7.35}$$

Beam pattern II is shown in Figure 7.8 for different steering angles. Unlike pattern I, pattern II

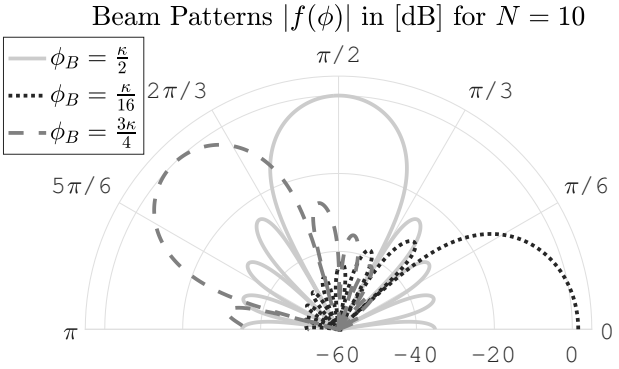


FIGURE 7.8: Examples for Beam Pattern II with different steering angles ϕ_B and $D = 0.2\pi$, $\kappa = \pi$.

enables control over the width of the main lobe.

Comparing the results in Figure 7.7 and Figure 7.8 shows that by allowing for an increased beam width, the side lobes' magnitudes decrease significantly. This can be interpreted as a trade-off between directivity and acoustic contrast. It can also be observed that the number of side lobes reduces slightly for Type II beams.

With a fully developed theory for sound field control with wedge arrays at hand, their performance shall be initially evaluated on the basis of simulations.

7.5 Simulations

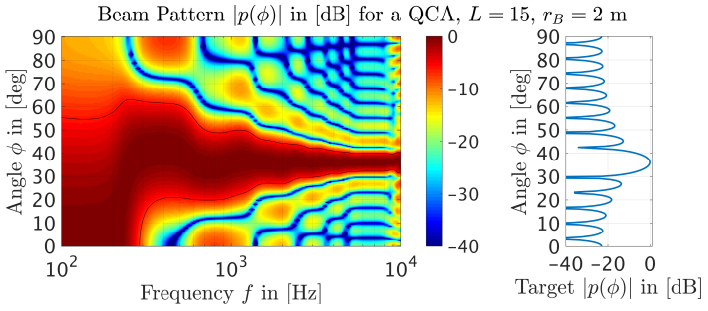
This section shows simulation results for the beamforming performance of wedge arrays with two wedge angles, a Quadrant Cylindrical Array (QCA), with $\kappa_1 = \frac{\pi}{2}$, and a Hemi-Cylindrical Array

(HCA), with $\kappa_2 = \pi$. The former corresponds to an array that is located in the corner of a room with rigid walls while the latter can be an array located on any large rigid wall. Note that all beam patterns presented in the remainder of this work include a black line that indicates the points where the main lobe has dropped by 6 dB from its maximum magnitude.

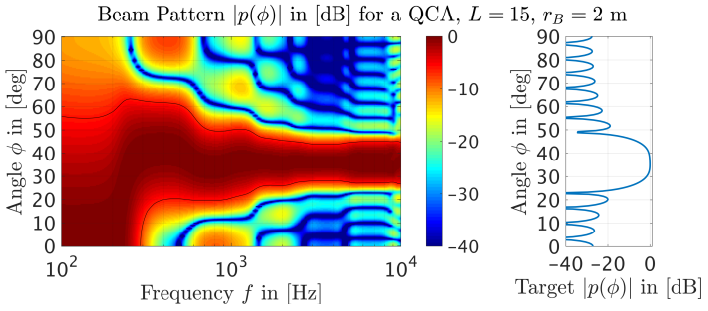
The radiated beam patterns were calculated on the basis of the sound field synthesis model in (7.2) for the plane $z = z_B = z_S$, where the infinite sum over n was truncated to $n = 30$ for the QCA and to $n = 60$ for the HCA. The speed of sound was assumed to be $c = 343\,\frac{\mathrm{m}}{\mathrm{s}}$ and the specific density of air $\rho_0 = 1.2041\,\frac{\mathrm{kg}}{\mathrm{m}^3}$. The velocity spectrum was calculated using Eq. (7.13), where different types of source models were used for the QCA and the HCA. The loudspeaker driving functions were obtained from Eq. (7.31), using the ARF in (7.25), unless stated otherwise. The simulations cover the two beam pattern types given in (7.34) and (7.35), and different control radii. All results are given within the frequency band from 100 Hz up to 10 kHz.

7.5.1 Case Study: Quadrant Cylindrical Array with $L = 15$

For the QCA with $L = 15$, all loudspeakers were modelled as point sources, i.e. $V_n^D(k_z) = 1$, with the velocity spectrum $V_n(k_z)$ as defined in Eq. (7.13). Figure 7.9 shows the normalised reproduced beam pattern of a QCA with $r_S = 0.2\,\mathrm{m}$ for the target patterns Type I and Type II. The lobe structures resemble that of the corresponding targets in the array's optimal frequency



(A) Beam Pattern Type I



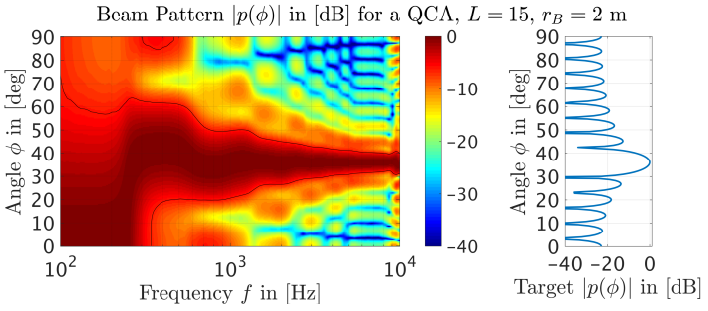
(B) Beam Pattern Type II, $D = 0.2\frac{\pi}{2}$

FIGURE 7.9: Simulated beam patterns for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\mathrm{m}$ and $r_S = 0.2\,\mathrm{m}$.

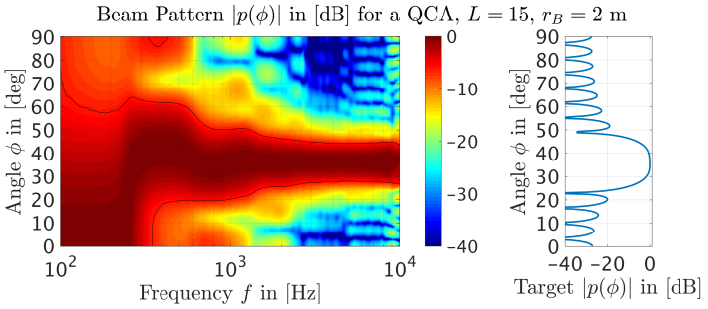
range between 5 and 8 kHz. Below that, as frequency decreases, the beam broadens while the number of side lobes drops. The side lobes remain distinct until directivity decreases so much that only a main lobe is left. This is founded in the required regularisation of the inverse solution in (7.31) with $\beta = 10^{-4}$ (see Section 7.3.7) [FSOF15]. The described effect is inherent to all

types of wedge arrays and will not be explicitly mentioned when discussing the simulation results presented below.

Figure 7.10 shows the beam pattern when the far field arc radial functions $\Gamma^{\Psi}_{n,\mathrm{FF}}$ were used to calculate the driving functions. The depicted beam patterns best approximate the target within



(A) Beam Pattern Type I



(B) Beam Pattern Type II, $D = 0.2\frac{\pi}{2}$

FIGURE 7.10: Simulated beam patterns for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\mathrm{m}$ and $r_S = 0.2\,\mathrm{m}$.

the aforementioned optimal frequency range, while, for most part of the displayed frequency range, the individual lobes are smeared and not as distinct as the target. This effect appears to increase towards low frequencies and can be traced back to the lack of control that the far-field ARFs provide at $r_B = 2\,\mathrm{m}$. This result suggests that the sound field still shows a near-field dominated behaviour for $r_B = 2\,\mathrm{m}$ and therefore requires the use of the ARF in (7.25) to calculate the driving functions for the best control of the beam pattern.

Spatial aliasing effects are visible in both Fig. 7.10 and Fig. 7.9 for frequencies higher than 8 kHz.

Figure 7.11 shows the beam pattern (Type I) radiated from a QCA with the same parameters as before apart from $r_S = 1\,\mathrm{m}$. Since the number of loudspeakers is the same, the effects of spatial aliasing are already visible at much lower frequencies. However, the directivity at low frequencies has improved significantly, as it was predicted for arrays with larger radius r_S in Section 7.3.5.

All simulations with the QCA show a certain variability of the angle at the main lobes' centre, particularly at low frequencies, where it deviates from the desired steering angle ϕ_B.

In principle, the observations described in this subsection for the QCA are transferable to any other wedge array. The next subsection shows the simulation of the performance of the HCA.

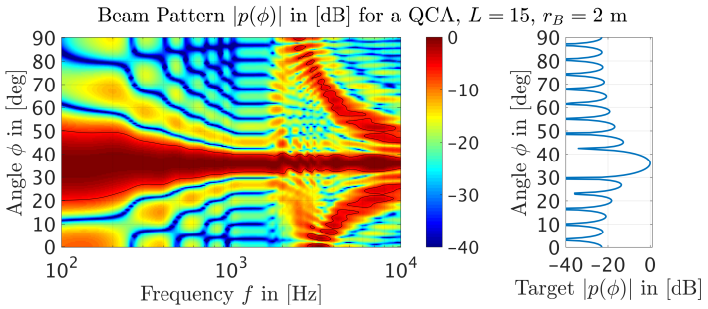


FIGURE 7.11: Simulated beam pattern for a QCA with $L = 15$, $N = 14$, $\beta = 10^{-4}$, $\phi_B = 36°$, $r_B = 2\,\text{m}$ and $r_S = 1\,\text{m}$.

7.5.2 Case Study: Hemi Cylindrical Array with $L = 15$

This case study provides reference results for the experimental data presented in Section 7.6. The prototype, which was previously presented in [HFF16] and Chapter 6, is depicted in Figure 7.12. It consists of $L = 15$ *Tang Band W1-1070SG* full-range drivers positioned at the angles

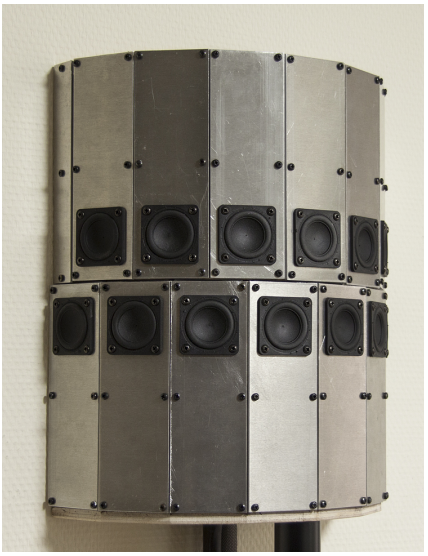


FIGURE 7.12: Prototype of a HCA with $r_S = 14.2\,\text{cm}$ and $L = 15$ *Tang Band W1-1070SG* full-range drivers.

$$\phi_l = \frac{\pi}{L+1} l,\ l = 1 \dots L \tag{7.36}$$

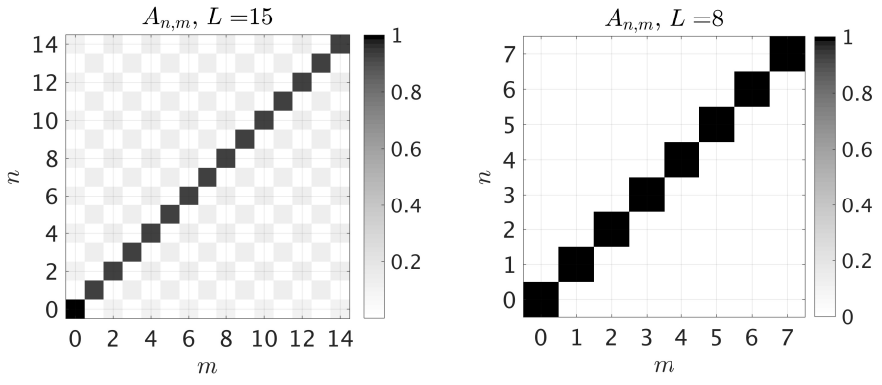
and at the corresponding z-axis values

$$z_l = z_B + [-z_S, +z_S, -z_S, \dots, -z_S], \tag{7.37}$$

where $z_S = 0.0275\,\text{m}$.

Unlike with the QCA, each loudspeaker in the HCA is modelled as a rectangular piston, using the coefficients $V_n^D(k_z)$ as defined in (7.12) with $B = 0.031\,\text{m}$ and $\Theta \approx 0.22$. Even though this seems an unusual choice to model a circular diaphragm, it must be said that the unavoidable order limitation effectively 'rounds the corners' of the modelled pistons, thus reducing the discrepancy.

The interleaved design aims to keep the array compact while increasing the number of drivers to gain a higher degree of spatial control. With this choice of ϕ_l and z_l, however, the evaluation of the sum in Eq. (7.22) does not yield the result in Eq. (7.23) for the interleaved arrangement. Instead, one obtains a more complicated relationship for $A_{n,m}$, which is shown in Figure 7.13a. That means when attempting to control the mth order, the system actually excites the mth order



(A) Full Prototype (B) Bottom Layer only

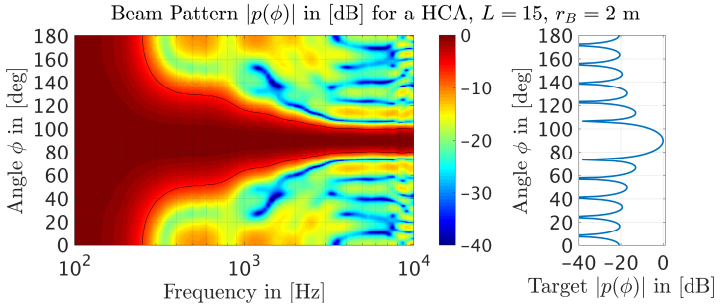
FIGURE 7.13: Visualisation of the orthogonality matrix $A_{n,m}$ for the full prototype and the bottom layer only.

but leaks energy into other orders as well. This is undesirable for a mode-matching approach, which assumes that only the modes it aims to control receive energy from the system. The prototype design is not orthogonal because it is missing 'half' a speaker at each end of the aperture, i.e. $\phi = 0, \pi$. If one uses only the bottom layer of the prototype (8 loudspeakers), the coefficients $A_{n,m}$ describe a perfect diagonal with all off-diagonal elements equal to zero, as depicted in Figure 7.13b. Despite the orthogonality flaw in the interleaved design, it is assumed that the loudspeaker driving functions can still be calculated from (7.31).

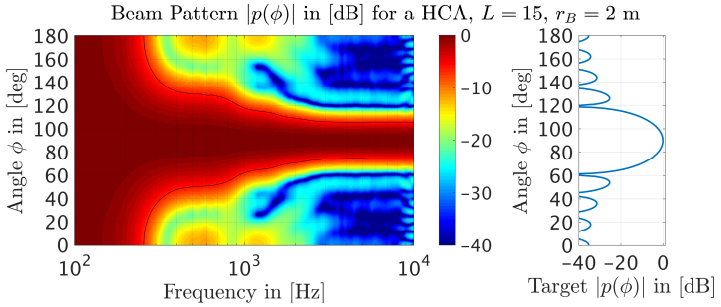
The performance of the full prototype for the target patterns Type I and II is shown in Figure 7.14. The maximum control order $N = L-1$ was reduced to 10 in order to push the lower aliasing edge frequency higher [FSOF15]. The result for the target Type I shows the highest directivity, while the array's performance with target Type II yields a very high side lobe suppression above 3 kHz. As it was observed for the QCA, the regularisation causes a reduction of directivity as frequency decreases. For both targets, the simulation shows a relatively stable beam width in the frequency range from 3 to 8 kHz; Spatial aliasing begins breaking up the beam pattern at frequencies beyond the latter.

Figure 7.15 shows the reproduced target beam patterns for the prototype's bottom layer only. When compared to Fig. 7.14a, as it can be expected, the frequency above which spatial aliasing occurs is much higher for interleaved configuration than for the bottom layer array. Aliasing effects start to show above 3 kHz for target Type I and only slightly further up for Type II. This effect is a result of the reduced magnitude of the coefficients F_n for higher values of n in the case of the Type II pattern.

It can also be seen that the full prototype does not reproduce a distinct target pattern like the bottom layer alone. Instead the side lobes in Figure 7.14 are smeared, independently of the type of target. This was observed before for the QCA, however, it was found to be caused by the

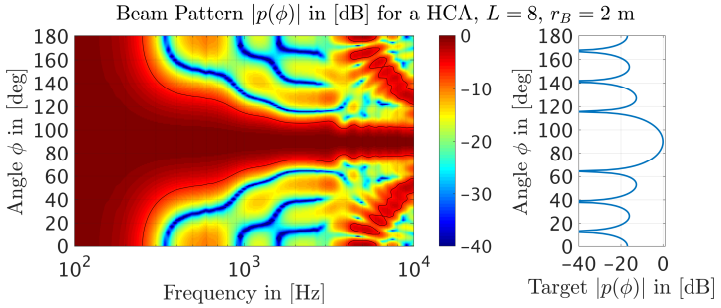


(A) Beam Pattern Type I

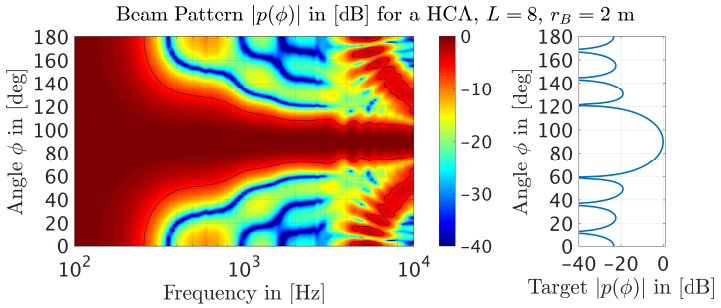


(B) Beam Pattern Type II, $D = 0.2\pi$

FIGURE 7.14: Simulated beam patterns for an HCA with $L = 15$, $N = 10$, $\beta = 0.1$, $\phi_B = \frac{\pi}{2}$ and $r_B = 2\,\mathrm{m}$.



(A) Beam Pattern Type I



(B) Beam Pattern Type II, $D = 0.2\pi$

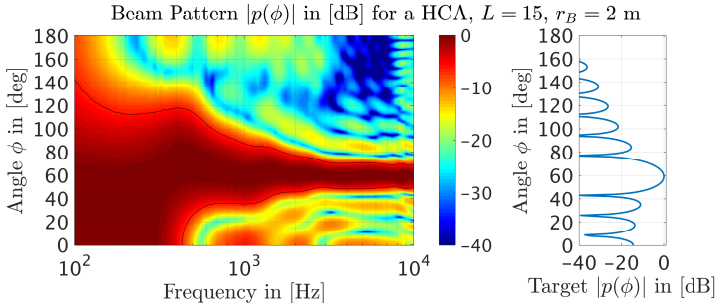
FIGURE 7.15: Simulated beam patterns for an HCA with $L = 8$, $N = 7$, $\beta = 0.1$, $\phi_B = \frac{\pi}{2}$ and $r_B = 2\,\mathrm{m}$.

imperfection of the far field ARF. In this case it appears that the lack of orthogonality inherent to the loudspeaker positions is responsible for this. However, even though the side lobes can no longer be clearly distinguished, they are lower than those for the bottom layer alone, showing that there is still a benefit from the interleaved layout. Furthermore, as it was already predicted

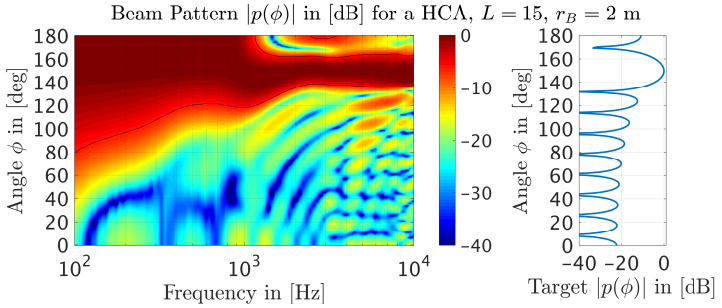
in Section 7.4, target Type II is shown to reduce the magnitude of the side lobes significantly over those achieved with target Type I.

It was found that decreasing the regularisation factor β does not improve the beam width at low frequencies for the interleaved configuration, yet it does so for the bottom layer. Therefore, in theory, the bottom layer by itself can yield a narrower beam at low frequencies than the interleaved configuration. This suggests that effects due to lack of orthogonality can also be mitigated by relatively strong regularisation, at the cost of reduced directivity at low frequencies. It is therefore suggested for future work to operate the prototype in the bottom layer configuration at low frequencies and make a transition to the interleaved configuration possibly around $f = 1\,\text{kHz}$.

The next two simulations consider only target Type I. The performance for the interleaved array with different steering angles ϕ_B in Figure 7.16 shows that the integrity of the main lobe remains relatively well preserved at frequencies beyond 3 kHz. The side lobes for the relatively extreme



(A) $\phi_B = \frac{\pi}{3}$



(B) $\phi_B = \frac{5\pi}{6}$

FIGURE 7.16: Theoretically achievable beam patterns for an array with $L = 15$ speakers, with $N = 10$, $\beta = 0.1$ and $r_B = 2\,\text{m}$.

steering angle of $\phi_B = \frac{5\pi}{6}$ rise up to just under 6 dB below the magnitude of the main lobe. It was found that this effect can be mitigated with a beam pattern Type II. At lower frequencies however, it appears that the main lobe no longer stays centred around the steering angle. This was already observed during the case study of the QCA and appears to occur only for steering angles $\phi_B \neq \frac{\kappa}{2}$. It can be assumed that this is caused by the combination of asymmetry and the break up of the lobe structure as a result of regularisation.

7.6 Measurements

7.6.1 Setup

Measurements to confirm the model for sound field control with wedge arrays were conducted for the case of the hemi-cylindrical loudspeaker arrays. For this purpose, the prototype was placed on a large planar baffle inside an anechoic chamber. Absorbing material (glass wool wedges) was placed at the edges of the planar baffle to reduce the reflections. The prototype was extended with wooden hemi-cylinders to better match the assumption of an infinite hemi-cylinder.

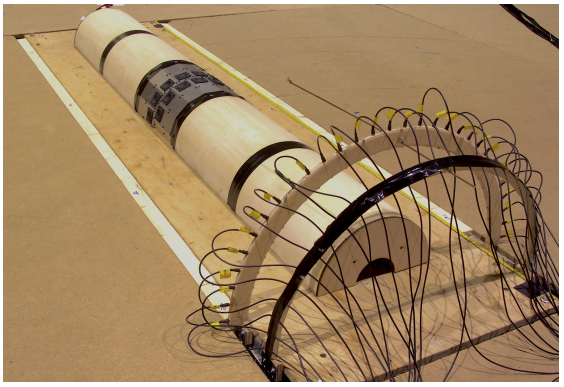


FIGURE 7.17: The measurement setup with a measurement array consisting of 30 *G.R.A.S. 40PL* array microphones fitted on an arch over the HCA with 30 cm radius.

The measurement arc with 30 *G.R.A.S. 40PL* array microphones is shown in Figure 7.17. This setup was used to acquire holograms from the radiated sound field at the radius $r_H = 0.3\,\mathrm{m}$. From this hologram data, it is possible to reconstruct the pressure field at a different points in the propagation space through acoustical holography, which is fully described for the wedge geometry in Chapter 6, Section 6.7, along with the measurement rig. The speed of sound and the specific density of air were chosen to match the in situ conditions, $c = 342.102\,\frac{\mathrm{m}}{\mathrm{s}}$ and $\rho_0 = 1.2130\,\frac{\mathrm{kg}}{\mathrm{m}^3}$.

Unless mentioned otherwise, all measurements results have been reconstructed at the control radius $r_C = 2\,\mathrm{m}$, where the holographic reconstruction of the pressure is denoted as $\hat{p}(r_C, \phi, z_S)$. The driving functions were calculated from (7.31). Unless stated otherwise, the steering angle was $\phi_B = 60°$ and the far field ARF $\Gamma^{\Psi}_{n,\mathrm{FF}}$ were used to calculate the driving functions.

7.6.2 Normalised Mean Squared Measurement Error

In order to assess the quality of the measurement data it was necessary to implement a measure that quantifies the error of the measured data $\hat{p}(r_C, \phi_w, z_S)$ with reference to the prediction $p(r_C, \phi_w, z_S)$ from the model. This work defines the normalised mean squared measurement error as

$$\bar{e}(\omega) = \frac{\sum_{w=1}^{W} |\hat{p}(r_C, \phi_w, z_S) - p(r_C, \phi_w, z_S)|^2}{\sum_{w=1}^{W} |p(r_C, \phi_w, z_S)|^2}, \tag{7.38}$$

where ϕ_w denotes all discrete angles where the sound field was reconstructed from hologram data. Note that this definition of $\bar{e}(\omega)$ also takes the phase error into account.

All subplots shown in Figures 7.18 and 7.19 present the corresponding mean measurement error along with the reconstruction data.

7.6.3 Results for Different HCA Beamforming Filters

The results of the HCA performance measurements are shown in Figures 7.18 and 7.19 for a selection of different parameters, along with the measurement error. Unless explicitly stated otherwise, all filters were calculated on the basis of the far field ARF, with $N = 10$ on the basis of beam pattern Type I.

Figures 7.18a, 7.18b and 7.18c show the performance of the HCA for different steering angles. Figures 7.18a and 7.18d show the performance of the HCA for Beam Pattern Type I and Type II at $\phi_B = 60°$.

Figures 7.18a, 7.19a, 7.19b and 7.19c show the performance of the HCA for Beam Pattern Type I at $\phi_B = 60°$ for four different filter settings, namely:

1. FF ARF & $r_C = 2\,\text{m}$ (see Fig. 7.18a),
2. NF ARF with $r_B = 2\,\text{m}$ & $r_C = 2\,\text{m}$ (see Fig. 7.19a),
3. FF ARF & $r_C = 0.25\,\text{m}$ (see Fig. 7.19b), and
4. NF ARF with $r_B = 0.25\,\text{m}$ & $r_C = 0.25\,\text{m}$ (see Fig. 7.19c).

The mean measurement error remains below 20 dB for the largest part of the frequency spectrum, which supports the model and the measurement technique.

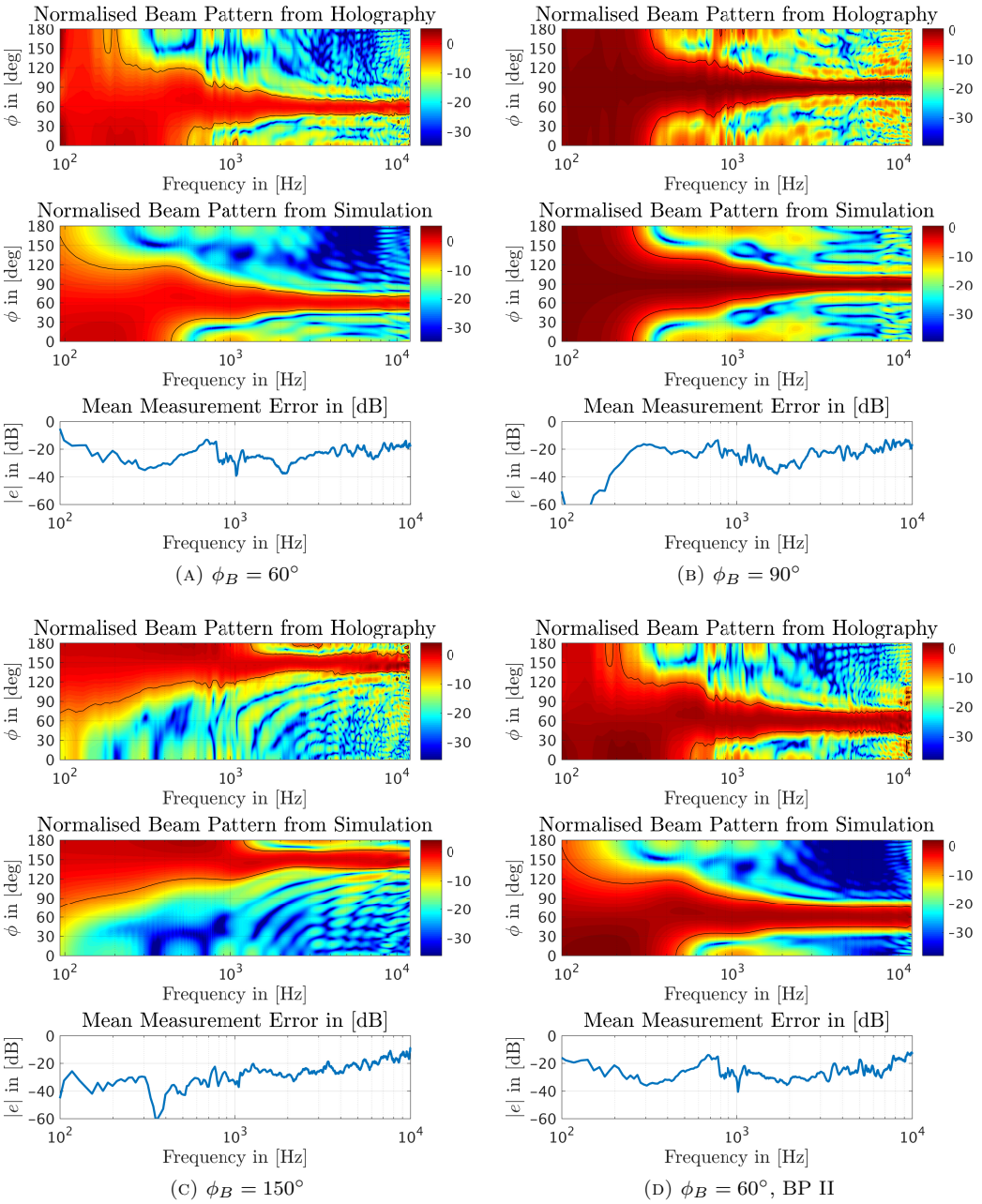
7.6.4 Discussion

The reconstruction from the measurement and the theoretical result show very good similarity, with all prominent features (i.e. side lobes, change of beam width, etc.) in relatively good match. The measurement mismatches are assumed to be due to on positioning errors of the microphones, imperfections in the acoustic setup due to reflections and finite structure and measurement aperture, and low SNR at low frequencies due to relatively small HCA output.

Of particular interest are probably those results shown in Figs. 7.19a, 7.19b and 7.19c since they confirm a good performance of the far-field ARF for reproduction radii of at least $R_C = 2\,\text{m}$. At this radius, no performance difference can be observed between the ARF and their far-field approximations. However, it can also be seen from the results in Figures 7.19b and 7.19c that, at small distances, filters based on the numerically calculated ARFs provide the desired performance in the near-field of the HCA, while far-field ARF-based filters fail to achieve the performance target.

The influence of the beam pattern type appears to be fairly small, yet, as it was predicted, the Type II beam pattern yields lower side lobes compared to the Type I beam pattern.

The results confirm the findings from the theoretical model, and hence indicate that the developed model for wedge loudspeaker arrays is valid and provides good results.



(A) $\phi_B = 60°$ (B) $\phi_B = 90°$

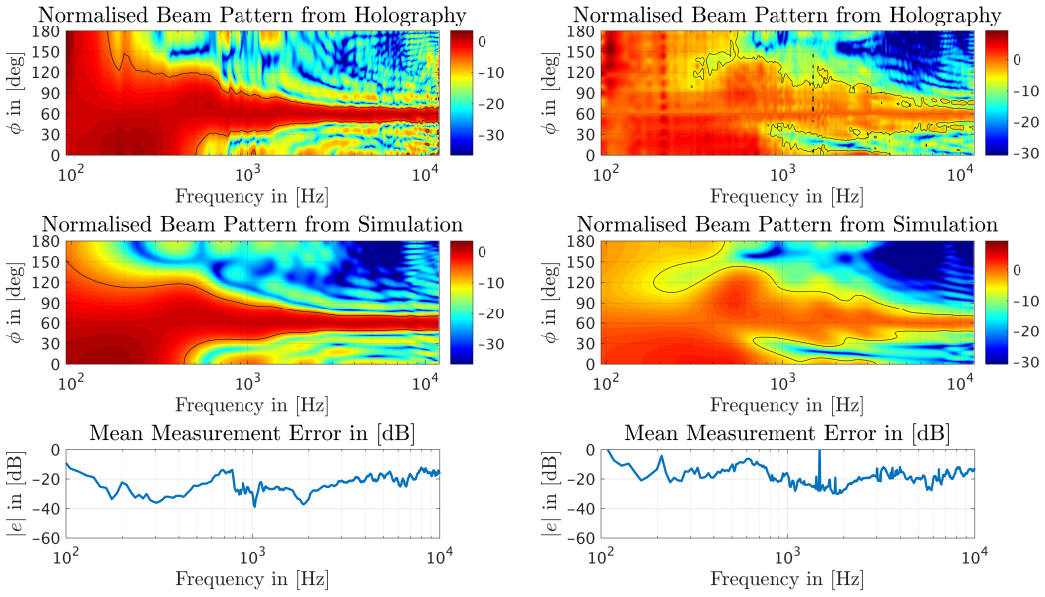
(C) $\phi_B = 150°$ (D) $\phi_B = 60°$, BP II

FIGURE 7.18: HCA performance for different steering angles ϕ_B and beam patterns.

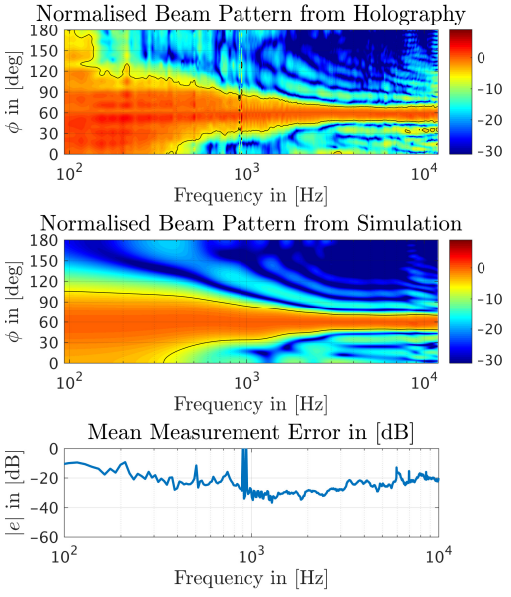
7.7 Summary & Outlook

An analytical model for the sound field of wedge loudspeaker arrays was developed.

The forward model maps the vibration profile of a loudspeaker array arranged in the shape of an arc on a curved baffle to the pressure in the propagation space. The radiation mechanism is described in the modal domain. Solving the inverse problem for a uniform loudspeaker distribution yields an expression for the control filters used to drive the array and control the radiated beam pattern. From the inverse solution, the aliasing behaviour of the system is predicted. To study the system's control performance at different radii and frequencies, the concept of Arc Radial



(A) NFs at $R_C = 2\,\mathrm{m}$ (B) FF at $R_C = 0.25\,\mathrm{m}$



(C) NF at $R_C = 0.25\,\mathrm{m}$

FIGURE 7.19: HCA performance for different steering angles ϕ_B, beam patterns and arc radial filters.

Functions (ARFs) is introduced. These were calculated numerically in their general form and through a far-field approximation.

Based on the developed model, the sound field for a QCA and an initial prototype of an HCA was simulated for different parameters and control filters to analyse the performance of wedge arrays. In order to confirm the data obtained through simulation, a series of measurements was conducted in a controlled environment. Using a special measurement rig, sound field holograms were acquired that enable acoustic holography to reconstruct the sound field at any given distance away from the prototype. The results obtained from this data confirm the simulation results and validate the developed model.

It was shown that for far-field control applications, the far-field filters yield results sufficiently similar to those obtained from the general filters. For near-field applications, however, it was shown that only the more accurate general filters provide adequate sound field control, as expected. The far-field filters presented in this work are effectively the same as those derived in [HFF16], and the corresponding findings therefore apply to either of them.

Future work will extend the model for the loudspeaker driving functions to allow for 3D sound field control and to derive driving functions and beamforming methods for loudspeaker arrays in finite wedges, for which the acoustic model was presented in Chapter 6.

7.8 Appendix I: Stationary Phase Approximation of the Arc Radial Functions

In [Wil99], Williams presented a generic formula for the SPA of integrals with the form

$$I(R) = \int_{-\infty}^{\infty} f(k_z) e^{iRg(k_z)} dk_z \tag{7.39}$$

that is given by

$$I(R) \approx f(k_{z0}) e^{iRg(k_{z0})} e^{-i\frac{\pi}{4}} \sqrt{\frac{2\pi}{R|g''(k_{z0})|}},\ g''(k_{z0}) < 0, \tag{7.40}$$

where k_{z0} denotes the *stationary phase point* that satisfies

$$\left.\frac{dg(k_z)}{dk_z}\right|_{k_z=k_{z0}} = 0 \tag{7.41}$$

and $g''(k_z)$ is the second derivative of g with respect to k_z. If $r \to \infty$, then the integral in (7.25) can be made of the form in (7.39) by replacing the Hankel function in $\Gamma_n(k_z, r)$ by its large argument asymptote [Wil99]

$$H_n(x) \sim \sqrt{\frac{2}{\pi x}} e^{i(x - n\pi/2 - \pi/4)} \tag{7.42}$$

and changing to spherical coordinates $r_B = R\sin\theta$ and $z_B = R\cos\theta$, yielding

$$\Gamma^{\Psi}_{n,\mathrm{FF}}(R,\theta) \approx$$

$$\frac{i\rho_0 ckLe^{-i\left(\frac{n\pi^2}{2\kappa}+\frac{\pi}{4}\right)}}{\kappa\pi^{\frac{3}{2}}\sqrt{2R\sin\theta}} \int_{-\infty}^{\infty} \underbrace{\frac{V_n^D(k_z)e^{-ik_z z_S}}{k_r^{\frac{3}{2}} H'_{\frac{n\pi}{\kappa}}(k_r r_S)}}_{f(k_z)} e^{iRg(k_z)} dk_z \tag{7.43}$$

where

$$g(k_z) = \sqrt{k^2 - k_z^2}\sin\theta + k_z\cos\theta.$$

Evaluating the integral in (7.43) using (7.40) yields the result of the SPA as the analytic expression for the far field approximation of the ARFs

$$\Gamma^{\Psi}_{n,\mathrm{FF}}(R, r_S) \approx \frac{\rho_0 cL}{\kappa\pi} e^{i\frac{n\pi^2}{2\kappa}} \frac{e^{ikR}}{R} \frac{V_n^D(k\cos\theta)e^{-ikz_S\cos\theta}}{\sin\theta H'_{\frac{n\pi}{\kappa}}(k\sin\theta\, r_S)}. \tag{7.44}$$

Chapter 8

Conclusions

A set of studies was presented in this work with the aim to further the state-of-the-art of sound field analysis and sound field control technology. The mutual ground for the different studies is the description of the pressure and velocity field, respectively, through an expansion of orthonormal basis functions in the cylindrical coordinate system. This facilitates solving the inverse problems associated with the considerations made for both the analysis and control of sound fields in a similar way.

The analysis of the components describing the forward and inverse operators (or forward and inverse propagators) is predominantly an analysis of the radial functions and their dependency on frequency. A strong attenuation for a mode in the forward operator/propagator potentially requires regularisation when solving the inverse problem due to ill-conditioning issues. This occurs for the sound field analysis as well as for the beamforming in the sound field control application. In the analysis cases, regularisation leads to a reduced spatial resolution, while with beamforming it leads to an increased beamwidth and thus a lower degree of local control.

This work has shown that the cylindrical coordinate system in combination with functional analysis of the sound field and specific boundary conditions provide a number of different ways to solve specific yet practically relevant problems. The presented findings contribute to the understanding of the investigated designs and techniques, and they provide an overview of their performance. Furthermore, they give a perspective on potential future ideas and investigations that are worth undertaking.

This chapter is organised in five sections. The first four provide general conclusions of the findings presented in the corresponding four main chapters of this thesis, while the last section gives an outlook on future developments in the field.

8.1 Tangential Pressure Gradient Array

For the considerations on circular microphone arrays composed of *Tangential Pressure Gradient* (TPG) sensors, the inverse problem was solved through an eigenvalue decomposition of the forward operator. The latter maps a given Herglotz Density to the pressure in the field around

the rigid-cylinder microphone array. The identified eigenvalues and corresponding eigenfunctions were used to invert the forward operator, yielding the inverse operator as the solution of the inverse problem.

The analysis of the eigenvalues revealed that, at low frequencies, the observation of the tangential pressure gradient yields a better-conditioned inverse problem than an observation of the pressure. As a consequence less regularisation is required, which then yields a higher spatial resolution than the design with pressure sensors only. At high frequencies, the eigenvalue analysis indicates a reduced robustness against spatial aliasing. Effectively, the optimum frequency band of the array is lower than that of a pressure sensor array. It follows that the choice of sensors determines the centre frequency of the optimum frequency band of the array, provided that the quality between sensors is comparable.

Furthermore, it was found that the observation of the TPG has the inherent flaw that the 0th order mode (the 'breathing cylinder') cannot be observed from the acquired signals. This is, in mathematical terms, due to a zero eigenvalue of that specific mode in the forward operator. As a consequence the inverse solution is non-unique, but it was shown that this issue can be resolved by adding a single pressure sensor to the array. The accuracy of the recovered information on the 0th order contribution was however shown to depend on the recovery of the remaining orders. This has a negative influence on the robustness of the 0th order information against spatial aliasing, as an error induced into the higher mode order coefficients propagates down to the 0th order.

Even though no distinct improvement of the array performance can be obtained by exclusively using TPG sensors, the presented study suggests that a combined array built from both sensor types can have a wider optimal frequency band. This assessment is based upon the assumption that both types of sensors are of the same quality. The problem of the reduced robustness against aliasing of the 0th order would be resolved. The application of a mixture of sensors was also shown to be beneficial in other applications of microphone arrays. For nearfield acoustical holography, different types of sensors were combined to increase the system's robustness against waves impinging on both sides of the measurement aperture (see Chapter 2, Section 2.2.4).

As an alternative to the pressure gradient sensor, it is worth conducting a similar study as the one presented in this work for tangential particle velocity sensors, e.g. the probe developed by *Microflown Technologies* or ribbon microphones. The relation between pressure gradient and particle velocity is given through Euler's equation. From the latter it may be predicted that the use of particle velocity instead of pressure gradient sensors can further improve the robustness against ill-conditioning at low-frequencies. As a consequence, the design would gain even more resolution at low frequencies than a design with TPG sensors.

The findings on the effect of TPG sensors in circular microphone arrays suggest that a combination of different types of microphone sensors (e.g. pressure, velocity, gradient) and different alignments of the directional sensors (e.g. radial, tangential, etc.) can be a promising way to increase the optimal frequency band of three-dimensional sound field measurement and recording systems. The variety of sensors would reduce the severity of ill-conditioning at low frequencies and thereby decrease the amount of regularisation needed for a stable solution of the inverse problem. Investigation through a prototype is required to determine if such a hybrid design can yield a wider optimal frequency range in practice than arrays with just one type of transducer.

8.2 DOA Estimation with a Finite Rate of Innovation Sound Field Model

It was shown that the *Herglotz Density* (HD) describing a field with a finite number of plane waves travelling in the same plane is periodic and has a *Finite Rate of Innovation* (FRI) following the definition by Vetterli et al. [VMB02]. The coefficients describing such an HD can be recovered from a circular microphone array within the latter's optimum frequency band. The total least-squares annihilating filter approach proposed by Blu et al. for time-domain signals with FRI was adopted for signals in the spatial domain to estimate the *Directions-of-Arrival* (DOAs) of plane waves impinging on the circular microphone array. The functionality of the method was confirmed through simulation studies and measured data acquired in controlled acoustical conditions.

The method requires $2L + 1$ microphones to extract the sound field coefficients that are needed to estimate the DOA of up to L plane waves. Based on acoustic measurements conducted using a circular microphone array with 15 sensors, the method produced DOA estimates with a mean *Direction Estimation Error* (DEE) of less than 4 % for two plane waves, and a mean DEE of less than 10 % for seven plane waves. The corresponding median DEE values were less than 2 % and 5 %, respectively. The discrepancy between the mean and median results suggests that the algorithm may produce outliers with significantly larger DEE values. These results for the accuracy that can be expected of the method are very similar to those obtained from a simulation study, which was conducted to assess the method's theoretical performance depending on the *Signal-to-Noise-Ratio* (SNR) and the number of plane waves.

The robustness of the proposed method against source types that are not compliant with the FRI data model (i.e. point sources and plane waves travelling out-of-plane) was investigated in the form of simulation studies. A basis function expansion in cylindrical coordinates was used for the analytical expressions for the pressure field of the different source types. The mechanisms compromising the annihilating filter estimation mechanism were investigated and discussed on the basis of the analytical expressions for the HD coefficients of non-compliant source types. It was shown that the DOA estimation accuracy of the proposed algorithm is still within respectable limits for a single non-compliant source. When isolated, the simulated results for point sources and plane waves travelling out of plane are comparable to sources compliant with the model. The experimental results confirmed that a single point source in the near field of the microphone array among a number of compliant sources in the overall sound field only mildly increases the mean DEE while the median DEE even improves, presumably due to a relatively high SNR.

It was found that the HD coefficients of a single (compliant) plane wave can be interpreted as a sequence of samples from taken a monochromatic analytic signal, where the frequency of the oscillation, the carrier, is determined by the DOA. The annihilating filter identified in the method attempts to suppress the carrier in the sequence of HD coefficients. This mechanism is compromised with non-compliant sources since the corresponding HD coefficients are an amplitude-modulated version of their compliant counterpart. It can be assumed that these additional spectral components influence the response of the identified annihilating filter, and in the final consequence also influence the DOA estimates.

It was also found that the proposed method does not require the signals associated with the impinging plane waves to be uncorrelated. This is a significant advantage over established

methods like MUSIC that explicitly require for the signals to be uncorrelated. In a reverberant environment, this may make the performance of the proposed method more robust to first order reflections that add to the direct sound impinging from a source.

It would be very interesting for future research to perform a thorough investigation on the influence of the amplitude-modulating components in the HD coefficients of non-compliant sources on the success of the annihilating filter method. Particularly the postulated masking of lower magnitude carriers from compliant sources by these additional spectral components needs to be first verified and then better understood.

It may be assumed that the problem of non-compliant plane waves can be solved by extending the estimation concept to three dimensions. That would require a cylindrical array with several layers of circular microphone arrays along the z-axis. From the multi-level data, a matrix instead of a vector of HD coefficients can be identified. It can be shown that, because of the form of the coefficients' dependency on k_z, the annihilating filter method can then be applied along the rows and columns separately. From these results, the azimuth and elevation component of the DOA can be identified. This was already attempted initially in the course of this project with promising results. For a proper verification further research is needed.

Two alternative approaches to enable three-dimensional DOA estimation with the annihilating filter method are to use a planar array or a spherical array. Where the planar case is very similar to the aforementioned multi-level cylindrical array, the spherical array appears to be a much greater challenge. One would need to first find a generic model for an annihilating filter in the spherical harmonics domain and then develop a method to identify the corresponding 'roots'. Another way to attempt this is by mapping the data measured on the sphere onto a different manifold where the annihilating filter method is more intuitive. This is a mathematically interesting problem that may require a more generic formulation of the the finite rate of innovation data model in spherical geometries.

The DOA estimation method proposed in this work would in principle serve to provide the required high accuracy estimates. However, a comparative study to determine its performance compared to other methods needs to be conducted. For such an investigation, aspects like estimation accuracy, robustness against corrupting factors and computational complexity need to be compared.

8.3 A General Radiation Model and Nearfield Acoustical Holography in Wedge Propagation Spaces

The model for sound radiation from a velocity distribution on the surface of a cylindrical segment inside the wedge was developed in the form of a basis function expansion. It was shown that the latter can be derived from the expansion for an external problem in cylindrical coordinates (see Chapter 2, Section 2.1.1) by introducing suitable boundary conditions to account for the reflections from the walls of the wedge. This yields a model for the infinite wedge. To model a finite wedge, a second pair of boundary conditions at specific values along the z-axis needs to be introduced.

Using the initial radiation model, Neumann-Green functions were formulated for the wedge propagation space. They were then taken as part of a Rayleigh-like integral equation that provides an alternative formulation of the radiation model.

Just like the general solution of the homogeneous Helmholtz equation in cylindrical coordinates, the solution for the radiation model of wedge propagation spaces relies on the separation of variables in the specific coordinate system. It describes a forward propagator that maps all modal combinations from the source's surface velocity to the pressure in the propagation space. It was shown that the inverse of the forward operator can be used to perform *Nearfield Acoustical Holography* (NAH) in the wave number domain. It was found that the resolution of the surface velocity reconstruction depends on the standoff distance, which is a common observation in the field of NAH. The transition from evanescent to propagating modal combinations of the forward propagator was investigated and a relation to the resolution of the NAH result along the circumferential and axial direction was drawn.

The application of NAH in wedge spaces may predominantly be of academic interest, since the velocity reconstruction would only work for a radiating surface with a constant radius around the origin. Nevertheless, the radiation model can be used to measure the far field response of sources that are located inside wedge-shaped propagation spaces (e.g. a loudspeaker in the corner of a room or on a flat wall). Klippel had already used the measurement of the modal components of the sound field observed in the near field of loudspeakers to reconstruct the far field radiation pattern of the unit [KB16, Kli17]. With the wedge model, the radiation pattern of loudspeakers can be predicted more accurately inside spaces with rigid boundary conditions (e.g. rooms, cinemas, etc.).

It would be interesting to investigate if the wedge model can be modified to describe the radiation inside a box-shaped room. One conceivable way to attempt this is to add further boundary conditions and find matching basis functions (i.e. a combination of the interior and the exterior domain, as done by Weinreich and Arnold for measurements of the radiation from violins [WA80]). Another approach could be to combine the wedge radiation model with a mirror source model.

The developed model could also measure the acoustic intensity radiated from sources that are typically standing on flat, rigid ground, e.g. large industrial machinery. The three dimensional analysis of the radiated sound field allows for the reconstruction of an object's radiation pattern and thus may help to identify components that are main sources of the operating noise. Under certain acoustic circumstances, the finite wedge model might further improve the accuracy of the measurement results and reduce the measurement effort by limiting the vertical expansion of the hologram through the introduction of a 'ceiling' and 'floor'.

8.4 Loudspeaker Arrays for Beamforming in Infinite Wedge Spaces

A beamforming method for compact arrays in infinite wedge spaces was presented. The model for the radiation from a source with constant radius around the origin of an infinite wedge propagation space was used to formulate the forward operator that maps the driving signals of the loudspeakers to the radiated beam pattern. It was shown that the sound field can be controlled

on an arc with specific radius and vertical position by solving the inverse problem in the modal domain to identify the required loudspeaker driving functions. The inverse problem is solved in the modal domain, where the coefficients defining the driving functions of the loudspeakers are put in relation with the desired beam pattern on the control arc through the introduced *Arc Radial Functions*.

The developed theory was confirmed through both numerical simulation and experimental data. Different beam patterns and arc radial functions were tested and it was shown that the theoretical model achieves a good agreement with the measured data. It was also shown that by using near field conformal arc radial functions, the sound field can even be controlled at very small distances from the array. At large distances, the numerically calculated arc radial functions can be replaced by their stationary phase approximation without affecting the array's performance.

The beamforming system requires the identification of an inverse solution that suffers from ill-conditioning at low frequencies and spatial aliasing at high frequencies. To preserve the integrity of the array, regularisation must be applied during the inversion to avoid driving signals with excessive amplitudes that might damage the loudspeaker units. The regularisation causes a drop in radiated pressure level, an increased beamwidth and the side lobes to move to different angles at low frequencies. To preserve a flat on-axis frequency response, the driving functions must be equalised accordingly.

It was observed that the side lobes occurring due to spatial aliasing can be reduced in level when one accepts a wider beamwidth. The wider beam corresponds to reduced strength of higher orders and it was found that the smaller the contributions of higher order modes the smaller the side lobes resulting from spatial aliasing. This is not surprising, as same is observed from the difference between the Fourier transform of a Dirac pulse and that of a rectangular function. It is also in accordance with the work by Fazi et al. [FSOF15], where the same was observed for circular arrays.

The loudspeaker array was modelled as a set of vibrating square patches, where each patch represents a single loudspeaker. A more suitable model for a circular diaphragm may be obtained through numerical calculation of coefficients corresponding to the individual drivers' diaphragm shapes.

For future work it would be worth to investigate the sound field control and beamforming performance of a multilevel loudspeaker array inside a finite wedge. To immediately address the question of practical relevance of such a system, it shall be suggested to design a system for the corner (with a right angle) of a room with rigid floor and ceiling. It can be assumed from the findings in this work that a multilevel loudspeaker arrays gains control of the beam pattern for both azimuth angle and elevation. Such a system would automatically include the reflections from the flanking walls as well as both floor and ceiling. This suggests that if turned into a loudspeaker system, this type of array has a significantly larger degree of control over the acoustics inside the listening room.

8.5 General Outlook

To this day it would seem that the problems of both spatial aliasing and ill-conditioning as they occur with microphone and loudspeaker arrays cannot easily be solved through design aspects while preserving practicality.

Microelectromechanical System (MEMS) microphones may one day help to push the edge frequency for spatial aliasing high enough so that it no longer falls within the audible frequency band. Their small size and improving audio quality would then allow to build arrays of very high order. Of course, the amount of data that would be acquired from such an array would increase rapidly along with the order, which poses a challenge for the required processing electronics and storage capacity.

While the MEMS technology would solve the problem through the number of sensors, the work by Alon and Rafaely [AR12, AR14, AR16] provides a signal processing approach to de-alias the beam patterns obtained with circular and spherical microphone arrays. Their beamforming method is based on a priori knowledge of the location of a source towards which to steer the beam. For a single source, this is already an improvement, but in more general sound fields the spatial aliasing problem has just transformed into another. Nevertheless, in many acoustical applications, the sources in the sound field may be assumed to radiate across a wide frequency spectrum. This assumption was previously proposed by Bernschütz in his work on bandwidth expansion for microphone arrays [Ber12]. A high resolution DOA estimation method can be used to first identify the directions of distinct sources in the sound field. With this a priori information about the direction of the sources, the de-aliasing beamformer proposed by Alon and Rafaely can be used to extract the isolated source signals. These can then be used to reconstruct an even higher resolution Ambisonics or circular harmonics signal describing the sound field at frequencies that would otherwise suffer from spatial aliasing artefacts.

Solving the spatial aliasing issue through signal processing can provide some improvement, but the success of the method is limited by the resolution of the DOA estimation method. Moreover, the approach still assumes that all sources at least radiate in the range from the optimum frequency band to higher frequencies, which may not always be the case. It may thus be a step forward, but does not provide a solution in general. The ideal case remains a continuous scanning of the sound field. An ultra-dense scanning of the sound field with MEMS microphones may provide the required quasi-continuous (within the audible band) spatial sampling. However it needs to be investigated if such a dense distribution of MEMS devices still satisfies the assumed boundary condition (rigid or acoustically transparent) on the array aperture. If the result is significantly different, the common models for the functional decomposition of the sound field around the array would no longer be valid. Of course, if transducers without any moving or obstructing parts can be invented, this problem would be solved immediately.

The ill-conditioning issue with both microphone arrays and compact loudspeaker arrays is determined through the radial functions and the noise in the measurement data. The only options to avoid ill-conditioning with microphone arrays are keeping a large array radius, use extremely linear and low noise transducers, and ensure high accuracy in the arrangement, as well as permanently monitoring the atmospheric conditions (i.e. the speed of sound and the specific density of air). However, it must be questioned if a high-order information at low frequencies

is practically relevant. With compact loudspeaker arrays, a large array radius can solve the problem as well, or alternatively, the use of loudspeakers with an extremely large dynamic range. The latter can, however, not be expected from conventional analogue transducers with moving diaphragms, as the mechanical suspension is a limiting factor. If a transducer can be invented where the sound is not radiated through the displacement of a diaphragm but in a different way, while it retains suitability for array applications, it would overcome the limitations on the system's directivity at low frequencies. This is however based on the assumption that such a transducer would have a significantly larger dynamic range.

With compact loudspeaker arrays, spatial aliasing can be tackled in a similar fashion as with microphone arrays. Increasing the number of transducers is certainly an option. However, with conventional moving diaphragm transducers, it cannot be expected that a design with a very small diaphragm yields a sufficiently high sound pressure level. However, a novel transducer design with the specifications discussed in the previous paragraph would be a potential solution to this approach. Alternatively, a diaphragm design with very high directivity at high frequencies would solve the problem as well. In that case, the sound fields from the individual drivers ideally no longer interfere with each other and can thus no longer create aliasing effects in the array's beam pattern.

The development of the theory and the practical validation of compact arrays in wedges already aims to move away from making free field assumptions for the surrounding propagation space and account for its room-like features instead. The next challenge is to not just model the wall or the corner that holds the array but also take further boundary conditions into account. The goal is to fully model the propagation space with all its (crude) boundaries (i.e. walls, floor and ceiling). Together with the compact loudspeaker arrays, this comprehensive spatial sound field model can yield a new degree of sound field control. Whether this is used to optimise the sound quality or the spaciousness of the sound field can be a decision that is taken when the system is designed to suit a purpose.

Moreover, such a comprehensive spatial sound field model is not just limited to sound field control applications but also poses a great opportunity for research on sound field capture, e.g. for virtual and augmented reality applications, but also for the normal loudspeaker development process. The holographic measurement in more realistic environments can be used to determine the radiation behaviour from a loudspeaker into a room. From the obtained information, the loudspeaker's design may be more easily optimised to yield the best sounding result for that specific environment.

Appendix A

Three Dimensional Sound Field Acquisition Based On The Tangential Component Of The Pressure Gradient

In this chapter, a theoretical analysis of a spherical microphone array is presented. The array is composed of sensors measuring the tangential component of the pressure gradient. The definition of the different boundaries and volumes involved in the formulation of the model for the sound field is the same as in Figure 4.1.

Like the circular array considerations in the previous chapter, the effects in the acoustic boundary layer of the spherical array will be neglected.

A.1 Model of the Sound Field

The three dimensional model for the pressure in the sound field around a spherical scatterer of radius r_s is given by

$$p(r,\theta,\phi) = 4\pi \sum_{n=0}^{\infty} r_n(kr) \sum_{m=-n}^{n} Y_n^m(\theta,\phi) \int_0^{2\pi} \int_0^{\pi} Y_n^m(\theta',\phi')^* q(\theta',\phi') \sin\theta' \theta' d\phi', \tag{A.1}$$

where the radial functions $r_n(kr)$ are defined as

$$r_n(kr_V) = i^n \begin{cases} j_n(kr_V) & \text{, FF} \\ j_n(kr_V) - \frac{j_n'(kr_s)}{h_n^{(1)\prime}(kr_s)} h_n^{(1)}(kr_V) & \text{, CS} \end{cases}. \tag{A.2}$$

for the open and rigid sphere, respectively.

The pressure gradient is a vector $\mathbf{g}$ defined as follows

$$\mathbf{g}(r,\theta,\phi) = \nabla \cdot p(r,\theta,\phi) = \frac{\partial p}{\partial r}\mathbf{e}_r + \underbrace{\frac{1}{r}\frac{\partial p}{\partial \theta}}_{g_\theta}\mathbf{e}_\theta + \underbrace{\frac{1}{r\sin\theta}\frac{\partial p}{\partial \phi}}_{g_\phi}\mathbf{e}_\phi, \tag{A.3}$$

where $\mathbf{e}_r, \mathbf{e}_\theta$ and $\mathbf{e}_\phi$ denote the three unit vectors in spherical coordinates. As will be seen later, for a full recovery of $q(\theta_y, \phi_y)$, it is necessary to make the actual output quantity a mixture of the two angular components g_θ and g_ϕ. One way to combine them is to generate a unit-norm linear combination

$$\mathbf{b}(\alpha) = \cos\alpha \cdot \mathbf{e}_\theta + \sin\alpha \cdot \mathbf{e}_\phi \tag{A.4}$$

of the two angular unit vectors and project the gradient vector onto it. Please note that α is the angle between $\mathbf{e}_\theta$ and an arbitrary vector $\mathbf{b}(\alpha)$ that lies in the plane spanned by $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$. The squared norm of the projected vector $\mathbf{g}_T = g_T \cdot \mathbf{b}(\alpha)$ will then be given by

$$g_T(\alpha, r, \theta, \phi) = \cos(\alpha) g_\theta + \sin(\alpha) g_\phi. \tag{A.5}$$

A.1.1 Angular Derivative in θ

The derivative of the pressure with respect to the polar angle θ is given by

$$\frac{1}{r}\frac{\partial p}{\partial \theta} = 4\pi \sum_{n=0}^{\infty} \frac{r_n(kr)}{r} \sum_{m=-n}^{n} \frac{\partial}{\partial \theta}\left[Y_n^m(\theta,\phi)\right] \langle Y_n^m | q \rangle_{\Omega_s}. \tag{A.6}$$

The derivative of Spherical Harmonics can be expressed through a combination of Spherical Harmonics of different orders:

$\underline{m \geq 0}$

$$\frac{\partial}{\partial \theta} Y_n^m(\theta,\phi) = \frac{1}{\sin\theta}\left[\sqrt{\frac{2n+1}{2n+3}}\sqrt{(n-m+1)(n+m+1)}Y_{n+1}^m(\theta,\phi) - (n+1)\cos\theta \cdot Y_n^m(\theta,\phi)\right] \tag{A.7}$$

$\underline{m < 0}$

$$\frac{\partial}{\partial \theta} Y_n^{-|m|}(\theta,\phi) = \frac{(-1)^{|m|}}{\sin\theta}\Bigg[\sqrt{\frac{2n+1}{2n+3}}\sqrt{(n-|m|+1)(n+|m|+1)}Y_{n+1}^{|m|}(\theta,\phi) - (n+1)\cos\theta \cdot Y_n^{|m|}(\theta,\phi)\Bigg]^*. \tag{A.8}$$

A.1.2 Angular Derivative in ϕ

The derivative of the pressure with respect to the azimuth angle ϕ is given by

$$\frac{1}{r\sin\theta}\frac{\partial p}{\partial \phi} = 4\pi \sum_{n=0}^{\infty} \frac{r_n(kr)}{r\sin\theta} \sum_{m=-n}^{n} \frac{\partial}{\partial \phi}\left[Y_n^m(\theta,\phi)\right] \langle Y_n^m | q \rangle_{\Omega_s}. \tag{A.9}$$

The derivative of Spherical Harmonics is considerably simpler than the one for the polar angle:

$\underline{m \geq 0}$

$$\frac{\partial}{\partial \phi} Y_n^m(\theta, \phi) = imY_n^m(\theta, \phi) \tag{A.10}$$

$\underline{m < 0}$

$$\frac{\partial}{\partial \phi} Y_n^{-|m|}(\theta, \phi) = (-1)^{|m|}(-i)mY_n^{|m|}(\theta, \phi)^*. \tag{A.11}$$

A.1.3 The Tangential Pressure Gradient

Using the formulae for the derivatives found in the previous two sections, the quantity describing the tangential pressure gradient in a scalar form as given in equation (A.5) can be written as follows.

$$g_T(\alpha, r, \theta, \phi) = 4\pi \sum_{n=0}^{\infty} \frac{r_n(kr)}{r \sin\theta} \sum_{m=-n}^{n} \left[D_n^m(\alpha) \cdot Y_{n+1}^m(\theta, \phi) - E_n^m(\theta, \alpha) \cdot Y_n^m(\theta, \phi) \right] \langle Y_n^m | q \rangle_{\Omega_s} \tag{A.12}$$

with

$$D_n^m(\alpha) = \cos\alpha \sqrt{\frac{2n+1}{2n+3}} \sqrt{(n-m+1)(n+m+1)} \tag{A.13}$$

$$E_n^m(\theta, \alpha) = (n+1)\cos\theta\cos\alpha - im\sin\alpha. \tag{A.14}$$

For negative m, it is actually sufficient to use the following relation.

$$Y_n^{-m}(\theta, \phi) = (-1)^m \left(Y_n^m(\theta, \phi) \right)^* \tag{A.15}$$

For the sake of convenience, the function

$$\Psi_n^m(\theta, \phi, \alpha) = \frac{1}{\sin\theta} \left[D_n^m(\alpha) \cdot Y_{n+1}^m(\theta, \phi) - E_n^m(\theta, \alpha) \cdot Y_n^m(\theta, \phi) \right] \tag{A.16}$$

is defined. This allows for the simplification of (A.12).

$$g_T(\alpha, r, \theta, \phi) = 4\pi \sum_{n=0}^{\infty} \frac{r_n(kr)}{r} \sum_{m=-n}^{n} \Psi_n^m(\theta, \phi, \alpha) \langle Y_n^m | q \rangle_{\Omega_s} \tag{A.17}$$

Assuming that the tangential gradient is observed on the boundary of a spherical microphone array of radius R_V, this leads to a new operator G_s.

$$g_T(\alpha, R_V, \theta, \phi) = (G_s q)(\alpha, R_V, \theta, \phi) := 4\pi \sum_{n=0}^{\infty} \frac{r_n(kR_V)}{R_V} \sum_{m=-n}^{n} \Psi_n^m(\theta, \phi, \alpha) \langle Y_n^m | q \rangle_{\Omega_s} \tag{A.18}$$

If $q(\theta_y, \phi_y)$ is square-integrable on the unit sphere, then it can be expressed through an infinite series

$$q(\theta_y, \phi_y) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} q_{n,m} \cdot Y_n^m(\theta_y, \phi_y), |q_{n,m}| < \infty, \forall m, n \tag{A.19}$$

Since spherical harmonics are orthonormal, it follows that

$$\begin{aligned}\langle Y_n^m|q\rangle_{\Omega_s} &= \int_0^{2\pi}\int_0^{\pi} Y_n^m(\theta_y,\phi_y)^* q(\theta_y,\phi_y)\sin\theta_y d\theta_y d\phi_y \\ &= \sum_{n'=0}^{\infty}\sum_{m'=-n'}^{n'} \underbrace{\int_0^{2\pi}\int_0^{\pi} Y_n^m(\theta_y,\phi_y)^* Y_{n'}^{m'}(\theta_y,\phi_y)\sin\theta_y d\theta_y d\phi_y}_{\delta_{n-n'}\delta_{m-m'}} \cdot q_{n',m'} \\ &= q_{n,m}. \end{aligned} \tag{A.20}$$

and hence that

$$g_T(\alpha,R_V,\theta,\phi) = (G_s q)(\alpha,R_V,\theta,\phi) := 4\pi\sum_{n=0}^{\infty}\frac{r_n(kR_V)}{R_V}\sum_{m=-n}^{n}\Psi_n^m(\theta,\phi,\alpha)q_{n,m} \tag{A.21}$$

A.2 Solution to the Inverse Problem

In Chapter 4 it has been shown that the solution to the inverse problem for a circular array can be obtained by using the components of the singular system in combination with equation (4.27). The solution found through the latter would be valid regardless of the underlying geometric model, provided that both conditions of Picard's theorem are satisfied.

A.2.1 Singular Value Decomposition of G_s

In order to find the solution to the inverse problem, it would now be desirable to perform a Singular Value Decomposition of the operator defined in equation (A.12). Considering this problem in the context of functional analysis, similar to the two dimensional case in Chapter 4, there is an operator G_s that transforms a sought after function $q(\theta_y,\phi_y)$, which is an element of an open Hilbert space A, into a function $g_T(\alpha,R_V,\theta,\phi)$, which is an element of an open Hilbert space B. The Hilbert space A is spanned by spherical harmonics Y_n^m as orthonormal basis functions. In conclusion, B is spanned by the basis functions $\Psi_n^m(\theta,\phi,\alpha)$, however, it is not guaranteed that they form an orthonormal basis. On the contrary, since they are composed of a weighted sum of two spherical harmonics of different order, it is extremely likely that they are generally not even orthogonal. The implications of this lead to the assumption that the inverse problem cannot be solved in a closed form, like it was possible in the two dimensional case.

For the special case of $\alpha = \frac{\pi}{2}$, however, the basis functions $\Psi_n^m(\theta,\phi,\alpha)$ are orthogonal and can also be reformulated to be orthonormal as well. This would then allow for a closed form solution for the orders $n = \mathbb{N}\backslash 0$ and the degrees $m \in [-n,\ldots,-1,1,\ldots,n]$. Unfortunately, the array would then be incapable of recovering the various coefficients of degree $m = 0$, which makes this design option rather impractical.

A.2.2 Solution Using Linear Algebra

If the Herglotz density is limited to an order N, it will be represented by the finite series

$$q(\theta_y, \phi_y) = \sum_{n=0}^{N} \sum_{m=-n}^{n} q_{n,m} \cdot Y_n^m(\theta_y, \phi_y). \tag{A.22}$$

Therefore, it can be represented in A by $(N+1)^2$ coefficients $q_{n,m}$. The tangential gradient g_T at the observation point (θ_1, ϕ_1) is then given by

$$g_T(\alpha, R_V, \theta_1, \phi_1) = \frac{4\pi}{R_V} \begin{pmatrix} i^0 r_0 \Psi_0^0(\theta_1, \phi_1, \alpha) \\ i^1 r_1 \Psi_1^{-1}(\theta_1, \phi_1, \alpha) \\ i^1 r_1 \Psi_1^0(\theta_1, \phi_1, \alpha) \\ i^1 r_1 \Psi_1^1(\theta_1, \phi_1, \alpha) \\ \vdots \\ i^N r_N \Psi_N^N(\theta_1, \phi_1, \alpha) \end{pmatrix}^T \begin{pmatrix} q_{0,0} \\ q_{1,-1} \\ q_{1,0} \\ q_{1,1} \\ \vdots \\ q_{N,N} \end{pmatrix} \tag{A.23}$$

For a number of L observation points (sensors), using matrix formulation, this can be extended to

$$\begin{pmatrix} g_T(\alpha, R_V, \theta_1, \phi_1) \\ \vdots \\ g_T(\alpha, R_V, \theta_L, \phi_L) \end{pmatrix} = \underbrace{\frac{4\pi}{R_V} \begin{pmatrix} i^0 r_0 \Psi_0^0(\theta_1, \phi_1, \alpha) & \dots & i^0 r_0 \Psi_0^0(\theta_L, \phi_L, \alpha) \\ i^1 r_1 \Psi_1^{-1}(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^{-1}(\theta_L, \phi_L, \alpha) \\ i^1 r_1 \Psi_1^0(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^0(\theta_L, \phi_L, \alpha) \\ i^1 r_1 \Psi_1^1(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^1(\theta_L, \phi_L, \alpha) \\ \vdots & \ddots & \vdots \\ i^N r_N \Psi_N^N(\theta_1, \phi_1, \alpha) & \dots & i^N r_N \Psi_N^N(\theta_L, \phi_L, \alpha) \end{pmatrix}^T}_{\mathbf{W}} \begin{pmatrix} q_{0,0} \\ q_{1,-1} \\ q_{1,0} \\ q_{1,1} \\ \vdots \\ q_{N,N} \end{pmatrix}. \tag{A.24}$$

If L exceeds the number of coefficients $q_{n,m}$ to be recovered, then (A.24) is overdetermined and can be solved in a least squares sense by applying the Moore-Penrose Pseudoinverse $\mathbf{W}^\dagger$ (compare Chapter 4, Section 4.3.2). Of course, (A.24) can be extended in a similar way as it has been done in chapter 4 by the information from one single pressure sensor to recover the 0th order $q_{0,0}$.

$$\begin{pmatrix} p(R_V, \theta_p, \phi_p) \\ g_T(\alpha, R_V, \theta_1, \phi_1) \\ \vdots \\ g_T(\alpha, R_V, \theta_L, \phi_L) \end{pmatrix} = \mathbf{W}' \begin{pmatrix} q_{0,0} \\ q_{1,-1} \\ q_{1,0} \\ q_{1,1} \\ \vdots \\ q_{N,N} \end{pmatrix}, \tag{A.25}$$

where

$$\mathbf{W}' = \frac{4\pi}{R_V} \begin{pmatrix} i^0 r_0 Y_0^0(\theta_p, \phi_p) & i^0 r_0 \Psi_0^0(\theta_1, \phi_1, \alpha) & \dots & i^0 r_0 \Psi_0^0(\theta_L, \phi_L, \alpha) \\ i^1 r_1 Y_1^{-1}(\theta_p, \phi_p) & i^1 r_1 \Psi_1^{-1}(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^{-1}(\theta_L, \phi_L, \alpha) \\ i^1 r_1 Y_1^0(\theta_p, \phi_p) & i^1 r_1 \Psi_1^0(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^0(\theta_L, \phi_L, \alpha) \\ i^1 r_1 Y_1^1(\theta_p, \phi_p) i^1 & r_1 \Psi_1^1(\theta_1, \phi_1, \alpha) & \dots & i^1 r_1 \Psi_1^1(\theta_L, \phi_L, \alpha) \\ \vdots & \vdots & \ddots & \vdots \\ i^N r_N Y_N^N(\theta_p, \phi_p) & i^N r_N \Psi_N^N(\theta_1, \phi_1, \alpha) & \dots & i^N r_N \Psi_N^N(\theta_L, \phi_L, \alpha) \end{pmatrix}^T .$$

The recovered coefficients are then given by

$$\begin{pmatrix} q_{0,0} \\ q_{1,-1} \\ q_{1,0} \\ q_{1,1} \\ \vdots \\ q_{N,N} \end{pmatrix} = (\mathbf{W}')^\dagger \begin{pmatrix} p(R_V, \theta_p, \phi_p) \\ g_T(\alpha, R_V, \theta_1, \phi_1) \\ \vdots \\ g_T(\alpha, R_V, \theta_L, \phi_L) \end{pmatrix} . \tag{A.26}$$

Appendix B

A De-Aliasing Strategy Based On Signal Processing

In the previous chapters it has been shown that the performance of circular microphone arrays not only depends on the type of applied transducers and their quality but also on their aliasing behaviour. Aliasing is an effect that occurs as a result of an inadequate sampling of a signal or any other form of continuously changing entity. For time domain signals, adequate sampling conditions are typically ensured with *anti-aliasing filters*, which restrict the frequency content of the signal to be sampled to a limited band. While in theory this can be achieved perfectly, in practice aliasing effects cannot be totally suppressed but are usually significantly reduced to an acceptable level. A spatially complex entity such as a sound field however cannot easily be 'smoothed'. As a consequence, only few options to implement actual anti-aliasing filters exist for spatial sampling and they usually come with either restrictions or certain drawbacks. One option is to use large diaphragm sensors, which bears the risk of potential distortion because of the sensor size; another option is to assume a stationary sound field to perform an integration over contiguous measurement points. Both methods offer only an intermediate solution as they do not actually solve the problem of aliasing and lead to a trade-off decision. It therefore appears to be necessary to approach the aliasing problem in another domain.

The requirements for the array design to overcome the problems of aliasing implied by the Nyquist-Shannon theorem evidently stand in the way of an important step forward in the field of sensor arrays. Even the latest transducer technology does not allow for a spatially continuous observation of the sound field, making it therefore necessary to seek for a way around the obligatory compliance of the sampling condition. Previously presented de-aliasing methods exploit the knowledge of the aliasing patterns for a specific sound field model [EJvS09, AR14]. This usually leads to a performance that depends on either a priori knowledge of the sound field or the potential of the applied dictionary. Both approaches have one aspect in common: they treat every frequency band of the signal independently.

This chapter presents an initial study of a novel de-aliasing strategy for circular sensor arrays. Its basic principle is to learn important and substantial information about the sound field from the optimal frequency band (see Chapter 4, Section 4.4) of an applied circular microphone array and

then apply that knowledge to recover original sound field information in frequency bands that are expected to be corrupted by aliasing. In other words, it is assumed that the sources in the field radiate over a wide frequency spectrum and their spatial distribution does not change with frequency. The de-aliasing strategy is divided into an *analysis stage* and a *de-aliasing stage*. The former analyses the sound field in the optimal frequency band to then generate a new sound field specific basis (also sometimes referred to as dictionary), while the latter applies the new basis in corrupted frequency bands to perform de-aliasing. Hence, the de-aliasing stage is completely independent from the analysis stage and both stages can be investigated and analysed separately.

As it has been mentioned above, the de-aliasing strategy is divided into two parts. The analysis stage is meant to identify a set of parameters that define precisely the basis of the current sound field. Three different design options are introduced: the CLEAN method, the Finite Rate of Innovation (FIR) method and the Compressive Sensing method. Only the first two have been further investigated in this work. In the subsequent de-aliasing stage, the acquired parameters can then be used to create a modified version of the basis. These modified basis functions predict the aliasing behaviour of the system and allow for the successful de-aliasing of corrupted data.

B.1 Sound Field Model

For the considerations in this chapter, the problem of sound field analysis, which is generally of a three-dimensional nature, is simplified to a two-dimensional problem. This means that, from the perspective of a listener, the sound field varies in the horizontal plane but not in the sagittal plane. Equivalently, the direction of a plane wave is only defined by the azimuth angle ϕ but not on the elevation angle θ (compare Section 3.9.3). That leads to the general assumption that the pressure field is invariant w.r.t. the x_3-axis.

$$p(x_1, x_2, x_3) = p(x_1, x_2) \tag{B.1}$$

Since the pressure is now no longer a function of x_3, it is more convenient for the interpretation of the integral in (4.1) to express the equation in cylindrical coordinates. With the parametrisation (compare Section 3.9.2)

$$x_1 = r\cos\phi, \text{ and} \tag{B.2}$$

$$x_2 = r\sin\phi, \tag{B.3}$$

the integration domain is given by the set

$$\Omega_c := \left\{ \hat{\mathbf{y}} = \begin{bmatrix} r'\cos\phi' \\ r'\sin\phi' \end{bmatrix} : \phi' \in [-\pi, \pi], r' = 1 \right\}. \tag{B.4}$$

The HWF can then be written as a function of r, ϕ, $\hat{\mathbf{y}}$ and ω.

$$p(r, \phi, \omega) = \int_{\Omega_c C} H(r, \phi, \hat{\mathbf{y}}, \omega) q(\hat{\mathbf{y}}, \omega) d\Omega(\hat{\mathbf{y}}) \tag{B.5}$$

Given the chosen parametrisation in the set definition of Ω_c in (B.4), the differential is given as

$$d\Omega_c = d\phi. \tag{B.6}$$

The Herglotz Kernel $H(r, \phi, \hat{\mathbf{y}}) = e^{ik\mathbf{x}\cdot\hat{\mathbf{y}}}$ can also be written as a function of r and ϕ through the Jacobi-Anger Expansion [CK98]

$$e^{ik\mathbf{x}\cdot\hat{\mathbf{y}}} = 2\pi \sum_{n=-\infty}^{\infty} i^n J_n(kr) \frac{e^{in\phi}}{\sqrt{2\pi}} \frac{e^{-in\phi'}}{\sqrt{2\pi}}, \tag{B.7}$$

where $J_n(x)$ is the Bessel function of order n. Replacing the Herglotz Kernel by the series expansion in (B.7) yields

$$p(r, \phi) = 2\pi \sum_{n=-\infty}^{\infty} i^n J_n(kr) \frac{e^{in\phi}}{\sqrt{2\pi}} \int_0^{2\pi} \frac{e^{-in\phi'}}{\sqrt{2\pi}} q(\phi') d\phi'. \tag{B.8}$$

This last equation is the basic pressure field model for the two-dimensional investigations in this chapter.

B.2 A New Set of Basis Functions For Sound Field Decomposition

From an information theory point of view, a sound field model like the one introduced in Equation (B.8), could contain an infinite amount of information. This can be easily proven by arguing that the model allows for an infinite number of plane waves, each of which is defined by two parameters: complex amplitude and direction of travel. In the model, both parameters are contained in the Herglotz Density $q(\phi)$. Another approach is that a perfect white noise signal (ranging from $t \to -\infty$ to $t \to \infty$), which has an infinite amount of information, can be mapped onto a one-dimensional circular manifold around the origin of the polar coordinate plane and would still not have changed its information content, despite having become periodic. Therefore, a HD of the form given in (4.10) can still carry an infinite amount of information. It is already known from the sampling theory of time-domain signals that such a signal cannot be sampled and successfully reconstructed since the bandwidth is infinite. The same applies for a Herglotz Density of infinite order.

As it has already been discussed in the introduction to this chapter, it is difficult to implement a type of spatial lowpass filter that limits the order of the Herglotz Density and thereby its information content to a finite amount. Nevertheless, it appears that the key to successful sampling of a continuous entity lies in the limitation of its information content. It is therefore not surprising that many successful approaches either require a-priori knowledge or impose a limit on the relevant amount of information in a particular use case [VMB02, Don06, CW08, BDV$^+$08, EJvS09, BB10c, BB10b, WEMJ11]. Both approaches seek to bypass the sampling theorem and overcome its limitations.

Within the optimum frequency band of a given circular microphone array, it is safe to assume that the HD coefficients $q_n, n \in [-N, \ldots, N]$ can be recovered with good accuracy. The spatial

lowpass effect of the operators in (4.16) or (4.18) suppresses contributions of orders $q_n, |n| > N$ on the measurement boundary ∂V and therefore serves as a system-inherent anti-aliasing filter. The recovered HD

$$\tilde{q}(\phi) = \sum_{n=-N}^{N} \tilde{q}_n \frac{e^{in\phi}}{\sqrt{2\pi}} \tag{B.9}$$

then provides a good representation of the original sound field.

The current form of the HD uses the TOS $\{a_n(\phi), n \in \mathbb{N}\}$. These basis functions have the valuable property of orthonormality, but they do not allow for a very concise way of representing the sound field by just a few parameters. Whether the number of plane waves in a sound field is very high or comes down to just one, the HD reconstruction in (B.9) always requires the knowledge of all $q_n, n \in [-N, \dots, N]$. For de-aliasing applications, this accurate but very general decomposition of the HD is not beneficial. It has already been shown by Alon and Rafaely [AR14] that the recording of a single plane wave can be de-aliased successfully. In sight of their results, it seems beneficial to analyse the recovered Herglotz Density w.r.t. the number of plane waves in the sound field and their individual direction of travel and thereby recovering a set of more specific sound field information. In order to do this, the first step is to take a closer look of the signature of a single plane wave in the HD.

B.2.1 Band Limited HD of a Single Plane Wave

From Chapter 4, Section 4.3.5, it is already known that the HD of a single plane wave with a complex amplitude b travelling in direction θ is given by

$$q(\phi) = b\delta(\phi - \theta). \tag{B.10}$$

Being a distribution and not a function, the Dirac pulse requires an infinite Fourier series for the best approximation. Moreover, the Dirac pulse is not square-integrable and therefore $\delta(\phi) \notin L^2(\Omega_c)$. In theory, this is not within a function space that is suitable for the operator H. However, Equation (B.10) can be rewritten as

$$q(\phi) = b \sum_{n=-\infty}^{\infty} \frac{e^{-in\theta}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}}, \tag{B.11}$$

using the completeness relation in Equation (3.24) for the TOS $\{a_n(\phi), n \in \mathbb{N}\}$ with $a_n(\phi) = \frac{e^{in\phi}}{\sqrt{2\pi}}$, where the expansion coefficients q_n are given by

$$q_n = b \cdot \frac{e^{-in\theta}}{\sqrt{2\pi}}. \tag{B.12}$$

If the series (B.11) is limited to $|n| \leq N$, the resulting $q(\phi)$ is an element of $L^2(\Omega_c)$ and can be mapped by the operator H. As it has already been shown in Subsections 4.3.6 and 4.4.5, the corresponding sound field provides an accurate approximation of a plane wave within a radius $r_R \approx \frac{N}{k}$. As a matter of fact, after the measurement of a perfect plane wave with an array built of L sensors, the Nyquist-Shannon sampling theorem suggests that only the coefficients $q_n, n \in [-N, \dots, N]$ with $N = \frac{L-1}{2}$ can be recovered perfectly, provided that an ideal anti-aliasing filter has been applied beforehand. Fortunately, a spatial anti-aliasing filter is not necessary

within the optimum frequency band due to the low pass behaviour of the propagation operator H.

A function similar to the one specified in (B.11) was defined in Chapter 3 as the circular sinc function. Equation (B.11) can then be rewritten as

$$\tilde{q}(\phi) = b \sum_{n=-N}^{N} \frac{e^{-in\theta}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}} = b \cdot \frac{1}{2\pi} \mathrm{csinc}_{(2N+1)}(\phi - \theta) \tag{B.13}$$

with

$$\mathrm{csinc}_{(2N+1)}(\phi) = \sum_{n=-N}^{N} e^{in\phi}. \tag{B.14}$$

This function can be seen as a basis function that corresponds to the approximation of the sound field of a plane wave travelling in the direction specified by θ within a circular area around the origin of radius $r_R \approx \frac{N}{k}$. Figure B.1 shows the circular sinc function for $L = 15$ and $\theta = \pi$. It

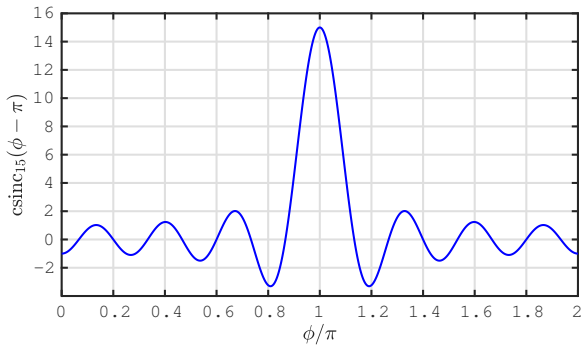


FIGURE B.1: Plot of the circular sinc function defined in (B.14) of order $L = 15$ with $\theta = \pi$.

very much resembles the sinc function $\mathrm{sinc}(x) = \frac{\sin(x)}{x}$, except for that the circular sinc function is periodic in 2π. In the continuous time domain the sinc function corresponds to a rectangular function in the frequency domain, which implies that the circular sinc function is strictly limited in spectral bandwidth. With a period of 2π, the circular sinc function's Fourier transform consists of a *finite* number of weighted Dirac pulses, which allows for its representation as a finite Fourier Series. Hence, analogous to the continuous time domain, the circular sinc function is *strictly limited in its spectral bandwidth* on the unit circle; this is obvious from its definition given in (3.34) when $N < \infty$.

This direct relation between the circular sinc function and a plane wave approximation offers the possibility to decompose a sound field into a finite number of M principal plane wave approximations, rather than into circular harmonics as given in (4.21). Each plane wave approximation would be specified by only two parameters, i.e. its amplitude and its incoming angle. A similar approach has already been successfully applied in [BB10c, BB10b], however with a different sound field analysis model and method, respectively, and a limitation to two principal plane waves. The following subsection introduces a new basis of the sound field that will be used in the de-aliasing stage.

B.2.2 Band Limited HD of M Plane Waves

Equation (B.13) states that the recovered HD of the sound field of a single plane wave approximation is equal to a weighted circular sinc function of order $L = 2N + 1$. In order for this to be true, it is required that the measurement was not corrupted by aliasing. As soon as this requirement is met, the HD of such a sound field no longer needs to be composed from a superposition of $2N + 1$ basis functions $a_n(\phi)$ but is simply equal to a weighted single circular sinc function that is shifted on the abscissa. This yields a comprehensive representation of the entire sound field through only two parameters.

Due to the linearity of the integral operator given in (4.16), the HD of a sound field with M plane waves is simply the sum of M corresponding circular sinc functions. Each of the M circular sinc functions is then specified by a direction of travel θ_m and an amplitude b_m. This then implies that a sound field consisting of M plane waves can be represented by an HD composed of M source-specific basis functions instead of L general basis functions from the TOS $\{a_n(\phi) : n \in [-N \dots N]\}$, yielding

$$\begin{aligned} q(\phi) &= \frac{1}{2\pi} \sum_{m=1}^{M} b_m \mathrm{csinc}_L(\phi - \theta_m) \\ &= \mathbf{c}(\phi)\mathbf{b} \end{aligned} \tag{B.15}$$

where

$$\mathbf{c}(\phi) = \frac{1}{2\pi} \begin{bmatrix} \mathrm{csinc}_L(\phi - \theta_1) & \dots & \mathrm{csinc}_L(\phi - \theta_M) \end{bmatrix}.$$

and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix}$$

The vector $\mathbf{c}(\phi)$ comprises the new basis of source-specific functions. An advantage of this decomposition into a source-specific basis is the reduction of an entire sound field representation to just $2M$ independent parameters.

B.2.3 Reconstructing the HD With The New Basis

Assuming that the directions of travel of the M plane waves are known, Equation (B.15) can be evaluated for a number angles $\phi_w = w\frac{2\pi}{W}, w = [1 \dots W]$ and $W \gg M$. The results for all W different angles can be merged into the following overdetermined equation system

$$\mathbf{q} = \begin{bmatrix} q(\phi_1) \\ \vdots \\ q(\phi_W) \end{bmatrix} = \begin{bmatrix} \mathbf{c}(\phi_1) \\ \vdots \\ \mathbf{c}(\phi_W) \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} = \mathbf{C}\mathbf{b}. \tag{B.16}$$

This compact formulation will be useful later on, when the amplitude of the identified plane waves needs to be recovered. Before that can be done, however, it is necessary to identify their directions of travel first.

One of the powerful properties of the original basis in form of the TOS $\{a_n(\phi), n \in [-N \dots N]\}$ was that any two elements satisfy the orthonormality relation in (3.22). The following subsection investigates the orthogonality of the (potential) elements of a basis such as the one specified by the vector $\mathbf{c}(\phi)$.

B.2.4 Autocorrelation Function and Orthogonality Property of the Circular Sinc Function

All the elements in the new basis are shifted versions of the circular sinc functions; the overall shape remains the same. The autocorrelation function $R(\theta)$ therefore provides a good measure to find out, whether or not two elements are orthogonal. The autocorrelation function of a function $f(\phi)$ in the domain Ω_c is defined by the integral

$$R_f(\theta) = \int_0^{2\pi} f(\phi) f(\phi - \theta) d\phi. \tag{B.17}$$

From Equation (3.20) it's been known that two elements of a function space $L^2(V)$ are orthogonal if their scalar product is equal to zero. The definition of the scalar product in (3.14) is relatively similar to that of the autocorrelation function in (B.17) except for the complex conjugation and the shifting of the function argument by θ. While the complex conjugation has no effect on a real-valued function, the function argument serves only as a parameter to shift one function w.r.t. the other. In other words, all zeros of the autocorrelation function $R_f(\theta)$ of a real-valued function f indicate those shifting parameters θ for which $f(\phi)$ is orthogonal to $f(\phi - \theta)$.

For $f(\phi) = \mathrm{csinc}_{2n+1}(\phi)$ one obtains

$$R_{\mathrm{csinc}}(\theta) = \int_0^{2\pi} \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi) \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta) d\phi.$$

Using the relation in (B.14) in the Equation above yields

$$\begin{aligned} R_{\mathrm{csinc}}(\theta) &= \int_0^{2\pi} \frac{1}{2\pi} \sum_{n=-N}^{N} e^{in\phi} \frac{1}{2\pi} \sum_{n'=-N}^{N} e^{in'(\phi-\theta)} d\phi \\ &= \int_0^{2\pi} \frac{1}{4\pi^2} \sum_{n=-N}^{N} \sum_{n'=-N}^{N} e^{i(n+n')\phi} e^{-in'\theta} d\phi \\ &= \frac{1}{2\pi} \sum_{n=-N}^{N} \sum_{n'=-N}^{N} e^{-in'\theta} \int_0^{2\pi} \frac{e^{i(n+n')\phi}}{2\pi} d\phi. \end{aligned}$$

Exploiting the orthonormality relation (3.22) then results in

$$R_{\mathrm{csinc}}(\theta) = \frac{1}{2\pi} \sum_{n=-N}^{N} e^{in\theta} = \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\theta).$$

It turns out that the autocorrelation function of the circular sinc function weighted by $\frac{1}{2\pi}$ is the function itself, as shown in Figure B.2. This result shows clearly that a function $\mathrm{csinc}_{2N+1}(\phi)$ is

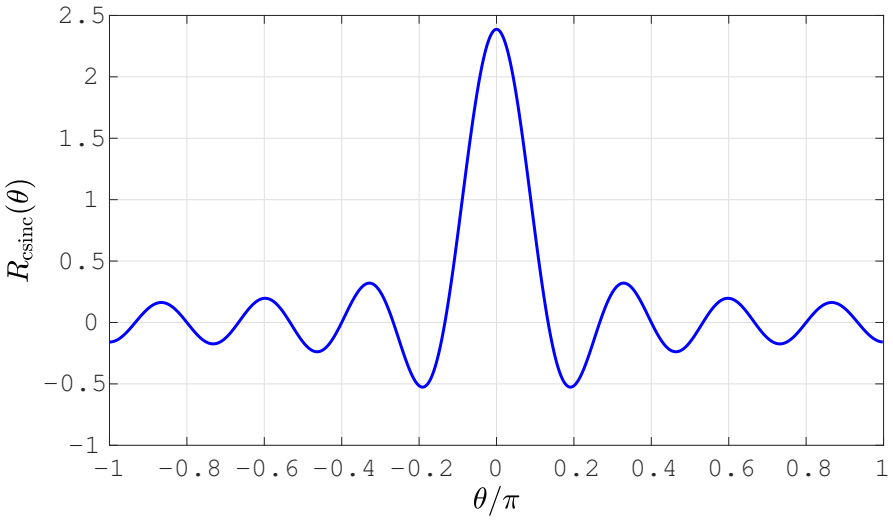


FIGURE B.2: Plot of the autocorrelation function of $\mathrm{csinc}_{2N+1}(\phi)$.

only orthogonal to a shifted version of itself $\mathrm{csinc}_{2N+1}(\phi-\theta)$ for $2N$ specific values of θ, namely

$$\theta_n = \frac{2\pi}{2N+1}n,\ n \in [1 \dots 2N].$$

In conclusion, the basis specified by $\mathbf{c}(\phi)$ is an orthogonal basis only if the elements are taken from the set

$$\left\{ \frac{1}{2\pi}\mathrm{csinc}_{2N+1}(\phi-\theta_m)\ :\ \theta_m = \frac{2\pi}{2N+1}m+\xi\ ,\ m \in [1 \dots M] \right\}, \tag{B.18}$$

where ξ is an arbitrary real-valued constant mutual to all θ_m. This result shall be of importance when it comes to the application of such a basis for the decomposition of a given HD or a sound field, respectively.

B.3 Methods to Identify the Parameters of the New Basis

There is more than one method to identify a new basis for the given data. The original basis specified by the set $\{a_n(\phi) : n \in [-N \dots N]\}$ is complete and orthonormal. A different basis can be found from an overcomplete dictionary by applying a *Basis Pursuit* approach [CDS98] that finds a new basis through convex optimisation of an l_1-norm constraint. This approach is frequently used in the field of compressed sensing and has been applied with microphone arrays before by Jin et al. [EJvS09, WEMJ11, WEJ12] to perform super-resolution. An alternative was proposed by Mallat and Zhang [MZ93]. Their *Matching Pursuit* method describes an iterative algorithm to obtain a new representation of given data from an overcomplete dictionary. If based on a TOS, the Matching Pursuit method leads to a perfect representation.

Going back to the original assumption that the sound field is actually composed of M plane waves instead of plane wave approximations, each of the plane waves can be assigned a direction of travel θ_m and a complex amplitude b_m. Hence, these $2M$ parameters define the sound field entirely. Since, however, a microphone array can in practice only recover enough HD coefficients

for a plane wave approximation, these parameters need to be extracted from the observed HD $\tilde{q}(\phi)$ or its finite number of coefficients $\tilde{q}_n$ instead, respectively. As explained later, the principal plane waves' amplitudes b_m can be easily recovered once the directions of travel θ_m have been identified.

In this work, the use of an overcomplete dictionary has been avoided, hence, in this report, compressed sensing approaches are only briefly discussed for the sake of completeness.

The knowledge that the individual components of the new basis are shifted versions of $\mathrm{csinc}_{2N+1}(\phi)$ allows to identify the individual basis components through either an iterative algorithm (Method I), a finite rate of innovation approach (Method II) or a compressed sensing approach (Method III).

B.3.1 Method I: Modified CLEAN Algorithm

The presented algorithm is basically a modified version of the iterative CLEAN algorithm by Högbom [Ḧ74]. It iteratively subtracts a finite number of principal plane wave components from the recovered HD $\tilde{q}(\phi)$ and thereby extracts their estimated directions of travel $\tilde{\theta}_m$. It is described as follows.

Algorithm

$q_0(\phi) = \tilde{q}(\phi)$
for $m = 1$ to M **do**
$\tilde{\theta}_m = \mathrm{argmax}_\phi |q_{m-1}(\phi)|$
$q_m(\phi) = q_{m-1}(\phi) - q_{m-1}(\theta_m) \cdot \frac{1}{2N+1}\mathrm{csinc}_{2N+1}(\phi - \tilde{\theta}_m)$
end for

The algorithm starts with a copy $q_0(\phi)$ of the recovered HD $\tilde{q}(\phi)$ and then finds the first angle $\phi = \tilde{\theta}_1$ for which $q_0(\phi)$ reaches a global maximum. It can be shown that this angle also corresponds to the maximum of the cross-correlation between $q_m(\phi)$ and $\frac{1}{2\pi}\mathrm{csinc}_{2N+1}(\phi)$ (see Appendix C). In that regard, the proposed algorithm is equivalent to the Matching Pursuit algorithm [MZ93, MCD14], only that this algorithm does not choose from a dictionary but identifies the matching basis function by itself. The angle $\tilde{\theta}_1$ shall be the first extracted parameter defining the direction of travel of the strongest principal plane wave. The next step is then to subtract the function $\frac{1}{2N+1}\mathrm{csinc}_{2N+1}(\phi - \tilde{\theta}_1)$, which is normalised w.r.t. its maximum amplitude, weighted by the value of $q_0(\tilde{\theta}_1)$ from the still unchanged copy of the recovered HD. The residual is then carried on to the next iteration of the algorithm. The algorithm terminates when M angles $\tilde{\theta}_m$ have been identified. The exact value for M can be either chosen empirically for a specific microphone array, to suit a particular application or even by using other a priori information on the environment (e.g. from visual sensors). Another option would be to leave the overall number of plane waves unspecified and let the algorithm terminate when the maximum value of $q_m(\phi)$ falls below a given threshold.

An example of the above algorithm applied to a HD composed of $M = 4$ randomly positioned and weighted plane waves is given in Figure B.3, indicating the state of the algorithm for each

iteration of the loop. For this figure, the weighting coefficients b_m had been chosen as real-valued for the sake of simplicity. It can be observed that a circular sinc function (CSF) is aligned

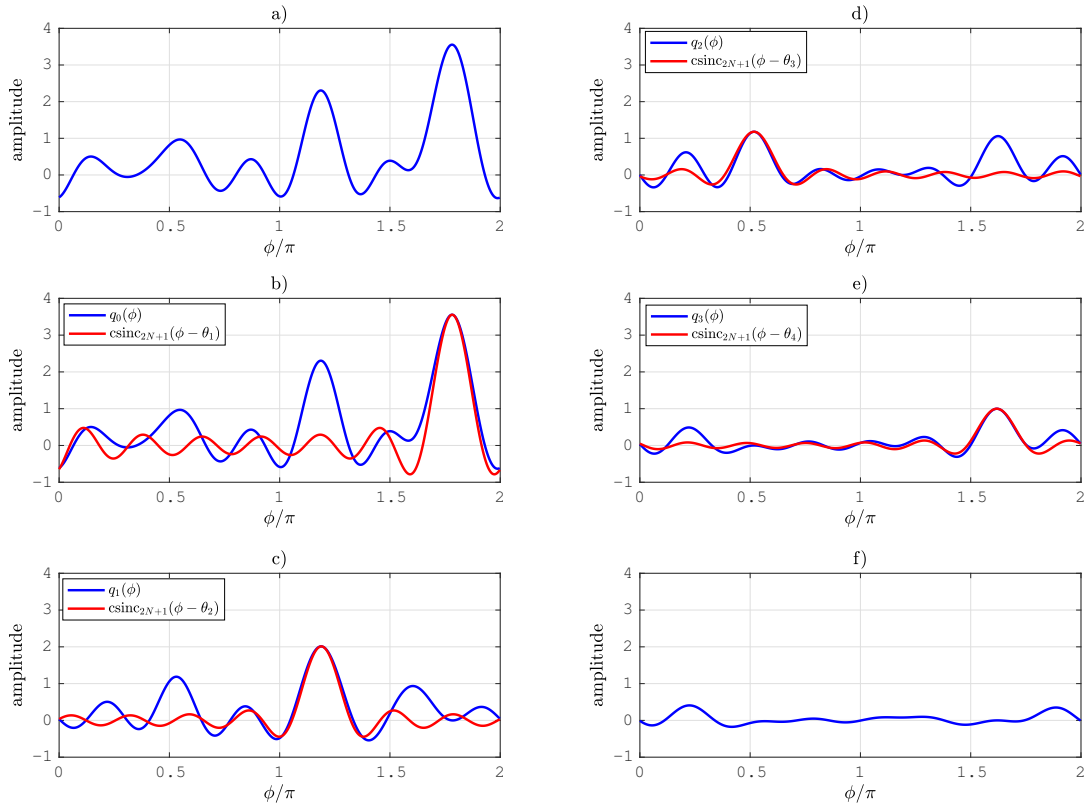


FIGURE B.3: Set of graphs describing the iterative basis acquisition of Method I: a) original HD, b) visualisation of iteration step $m = 1$, c) visualisation of iteration step $m = 2$, d) visualisation of iteration step $m = 3$, e) visualisation of iteration step $m = 4$, f) residual after the algorithm terminated.

alongside the current function $q_m(\phi)$ to match the maximum of the latter. To obtain the new residual, the exact same CSF is subtracted from $q_m(\phi)$. It is also obvious that the algorithm does not produce an ideal result, since the residual in the bottom right plot is not equal to zero. This can by easily explained by the fact that the basis functions of the form $\mathrm{csinc}_{2N+1}(\phi - \theta_m)$ are not orthogonal.

Even though this method appears to be rather unsophisticated, the results are surprisingly robust. The error induced by its imperfection (i.e. the lack of orthogonal basis functions) is investigated in a later section.

B.3.2 Method II: Finite Rate of Innovation (FRI) Approach

While Method I delivers only approximate results for the directions of travel of the principal plane waves that form the basis $\mathbf{c}(\phi)$, a method exists that allows to recover the exact values of θ_m from the coefficients q_n if the number of principal plane waves is smaller than or equal to N. This method requires a more elaborate approach that assumes the HD to have a *finite rate of innovation*. A comprehensive introduction to the theory behind signals with finite rate

of innovation is beyond the scope of this work and the interested reader is referred to pertinent literature [VMB02, DH09, BDV$^+$08]. This section briefly introduces the theory behind this particular application of this relatively new signal model.

The phrase 'finite rate of innovation' can also be paraphrased as 'finite number of degrees of freedom'. A Herglotz Density of the form

$$q(\phi) = \sum_{m=1}^{M} b_m \sum_{n=-\infty}^{\infty} \frac{e^{in(\phi-\theta_m)}}{2\pi}$$

has exactly $2M$ degrees of freedom, i.e. M directions of travel and M amplitudes. The Equation above can however be rearranged into a Fourier series

$$q(\phi) = \sum_{n=-\infty}^{\infty} \sum_{m=1}^{M} b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}}$$

with the Fourier coefficients

$$q_n = \sum_{m=1}^{M} b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}}.$$

The progression of the coefficients q_n can be seen as an infinite sequence, where n denotes the time index. Alternatively, it can be seen as the superposition of M infinite sequences defined by

$$q_n^{(m)} = b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}}$$

for $m \in [1 \dots M]$. For each signal $q_n^{(m)}$, the corresponding filter

$$a_n^{(m)} = \delta_n - e^{-i\theta_m}\delta_{n-1}$$

satisfies the equation

$$a_n^{(m)} * q_n^{(m)} = 0,$$

where the asterisk represents the discrete signal convolution. They are therefore also sometimes referred to as *annihilating filters*. The z-Transform of the filters $a_n^{(m)}$ is given by

$$A^{(m)}(z) = 1 - e^{-i\theta_m} z^{-1},$$

which is zero for $z = e^{i\theta_m}$. Each annihilating filter basically suppresses the component $e^{in\theta_m}$ in the sequence of the Fourier coefficients q_n.

If the individual progressions $q_n^{(m)}$ could actually be observed directly from the recovered coefficients q_n, it would be simple to identify the corresponding annihilating filters. Unfortunately this is not the case.

Going back to the originally observed coefficients q_n, it is now easy to see that

$$q_n = \sum_{m=1}^{M} q_n^{(m)}, \forall n \in [-N \dots N].$$

From the findings above, it can be assumed that one non-trivial solution for a_n to the equation

$$q_n * a_n = 0 \tag{B.19}$$

is the combination of all annihilating filters $a_n^{(m)}$ to one comprehensive annihilating filter a_n of the form

$$a_n = a_n^{(1)} * a_n^{(2)} * \ldots * a_n^{(M)} = \delta_n + \sum_{m=1}^{M} \alpha_m \delta_{m-n} \tag{B.20}$$

with its z-Transform given by

$$A(z) = \prod_{m=1}^{M} A^{(m)}(z) = \prod_{m=1}^{M} (1 - u_m z^{-1}),\ u_m = e^{-i\theta_m}.$$

The z-Transform can also be written using the coefficients α_m defined in Equation (B.20):

$$A(z) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \ldots + \alpha_M z^{-M} = (1 - u_1 z^{-1})(1 - u_2 z^{-1}) \ldots (1 - u_M z^{-1}).$$

From the above equation, it can be seen that the coefficients u_m are in fact the roots of the characteristic polynomial

$$p(z) = z^M + \alpha_1 z^{M-1} + \alpha_2 z^{M-2} + \ldots + \alpha_M.$$

In conclusion, once the annihilating filter a_n has been identified and the roots of its characteristic polynomial have been calculated, then the directions of travel of the M plane waves are recovered. The next step is therefore to find the annihilating filter a_n.

The convolution of discrete-time signals can be realised in a convolution matrix. With a limited length of the filter a_n, the equation system equivalent to (B.19) is given by

$$\begin{bmatrix} q_1 & q_0 & q_{-1} & \cdots & q_{-(M-1)} \\ q_2 & q_1 & q_0 & \cdots & q_{-(M-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_M & q_{M-1} & q_{M-2} & \cdots & q_0 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

It follows from Equation (B.20) that $a_0 = 1$. This can be exploited to avoid the trivial solution for the filter a_n. Rearranging the above equation system yields

$$\begin{bmatrix} q_0 & q_{-1} & \cdots & q_{-(M-1)} \\ q_1 & q_0 & \cdots & q_{-(M-2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_{M-1} & q_{M-2} & \cdots & q_0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = - \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_M \end{bmatrix}, \tag{B.21}$$

which is widely known in the field of auto-regressive filtering as a Yule-Walker equation system. This can ideally be solved by plain matrix inversion, since it can be expected that the matrix is full rank. However, this may not be the case in practice if the M plane waves are not distinct. Furthermore, Equation (B.21) implicitly gives a condition for the set of mode coefficients $\{q_n : n \in [-N_{\min} \ldots N_{\min}]\}$ that is needed to perform this type of analysis. From the system

matrix and the right hand side of Equation (B.21) it follows that

$$2N_{\min} + 1 = 2M \iff N_{\min} = M - \frac{1}{2}.$$

The last step is then to find the M distinct roots u_m of the characteristic polynomial $p(z)$, from which the estimated directions of travel $\tilde{\theta}_m$ can be calculated by the formula

$$\tilde{\theta}_m = i \ln(u_m), \forall c \in [1 \dots M]$$

where $\ln(x)$ denotes the principal value of the natural logarithm.

Figure B.4 gives an indication of the accuracy of the detection of the directions of travel when using the finite rate of innovation (FRI) method with an uncorrupted set of HD coefficients $\{q_n \ : \ n \in [-7 \dots 7]\}$ of a randomly generated HD composed of $M = 4$ plane wave approximations. The top graph shows the original HD $q(\phi)$. The bottom graph presents the four identified

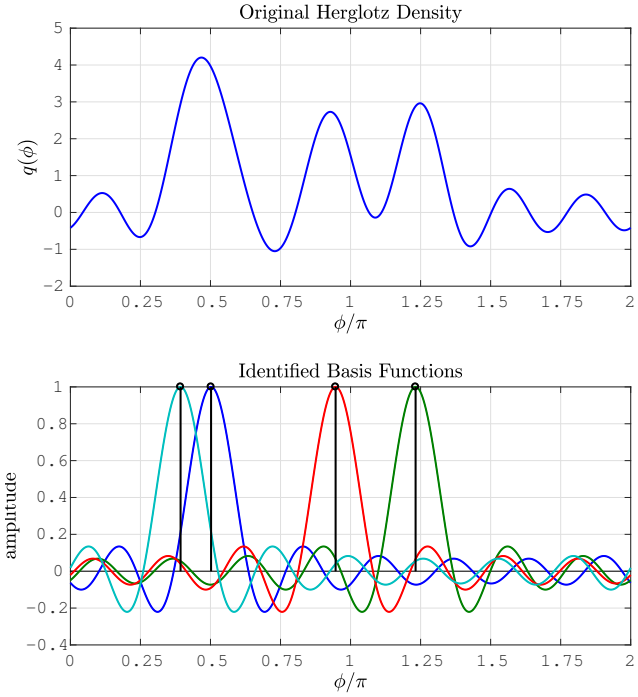


FIGURE B.4: Example of an original HD (top) and the basis functions obtained through Method II (bottom) for $M = 4$ and $N = 7$.

basis functions in different colours, while the black stems indicate the originally underlying directions of travel. It can be seen that the FRI method approximated the set of parameters $\{\theta_m : m \in [1 \dots M]\}$ with very high accuracy. The nearly negligible residual is credited to the limited numerical accuracy during the matrix inversion.

B.3.3 Method III: Compressive Sensing

Another method to find the directions of travel of the principal plane waves is based on compressed sensing. In this case, the matrix $\mathbf{C}_{\mathrm{CS}}$ serves as a huge dictionary of shifted circular sinc functions

of unity amplitude and the LS solution for $\mathbf{b}$ is obtained with an additional L_0-constraint, such that

$$\min(\|\mathbf{q} - \mathbf{C}_{\mathrm{CS}}\mathbf{b}_{\mathrm{LS}}\|_2) \text{ with } \|\mathbf{b}_{\mathrm{LS}}\|_0 = M.$$

This will lead to only those M circular sinc functions contributing to $q(\phi)$ which yield a minimised L_2-norm of the residual $\mathbf{q} - \mathbf{C}\mathbf{b}_{\mathrm{LS}}$.

Even though they will not be considered further in this work, it should be mentioned that compressed sensing approaches have shown to be more robust compared to plain LS methods in scenarios when the direction of travel lies between two atoms of the dictionary, i.e. none of the atoms perfectly matches the underlying data [CDS98, EJvS09]. In this case, the LS solution is likely to have a large number atoms in the dictionary contribute to the approximation, i.e. all those atoms that have a non-zero scalar product with the data. A compressed sensing solution should however ideally lead to an approximation involving only the two atoms closest to the ideal atom. Epain et al. investigated this in further detail for the problem of sound field reproduction [EJvS09].

B.3.4 Performance of the CLEAN and the FRI Method for $M > N$

The CLEAN method has no apparent limit to the number of identified plane waves due to its definition through an iterative algorithm. However, it must be doubted that the parameter estimation is going to be sufficiently accurate beyond a certain value for M.

The FRI method has a fixed limit $M = N$ up to which it works remarkably well (compare Sections B.3.2 and B.4.2), given that the coefficients q_n are not corrupted by either measurement noise or spatial aliasing. It is inherently limited to identify more than N plane waves, which can be seen from Equation (B.21), and it is yet to be investigated what the exact nature of the parameter estimates is in the case that the coefficients q_n define more than $M = N$ plane waves. So far it has been observed that the parameter estimates $\tilde{\theta}_m$ are complex valued in such a case. In order to allow for the method to function anyway, the parameters that are forwarded to the de-aliasing stage have been restricted to their real value.

B.4 Sound Field Decomposition with the New Basis

In the previous section, potential methods to obtain a new basis for the decomposition of a recovered Herglotz Density were introduced. This section first presents the technique to solve the equation system presented in (B.16) based on the new basis matrix $\mathbf{C}$. The two solutions, obtained through the inversion of the basis matrices found by Method I and Method II, are then compared. The quality of the acquired solutions is analysed for how accurately the original data can be reconstructed. To get a better idea of what factors affect the quality of the reconstruction, the accuracy of the parameter estimates delivered by Method I and Method II is investigated by looking at their average parameter identification error.

B.4.1 Solving the Inverse Problem with the New Basis

Comparing the problem of finding the complex amplitudes of the individual new basis components to the problem of the recovery of the HD from the observation of a microphone array, it can be seen that recovering the complex amplitudes b_m is also an inverse problem. The respective problem with the microphone array was ultimately solved through a matrix inversion (compare Chapter 4, Equation (4.52)). It will be shown that the current inverse problem can be solved in a similar manner. As it has already been mentioned in the beginning of subsection B.3, it is rather simple to find the amplitudes b_m of the M principal plane waves, once their directions of travel θ_m are known.

By simply inverting the equation system defined in (B.16), it can be solved for the vector $\mathbf{b}$, using the Moore-Penrose Pseudoinverse

$$\mathbf{C}^{\dagger} = (\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H. \tag{B.22}$$

This leads to the least squares (LS) solution for the vector containing the weights of the M plane waves

$$\mathbf{b}_{\mathrm{LS}} = \mathbf{C}^{\dagger}\mathbf{q}. \tag{B.23}$$

Note that this solution can also be found in the domain of the HD coefficients $\{q_n : n \in [-N \dots N]\}$, which has yet not been pursued further in this work for the sake of vividness of the process.

A comparison based on an example of the reconstruction performance of the bases acquired through Method I and II is given in Figure B.5. Both methods appear to find adequate matches of the underlying basis functions, however it can already be seen from Fig. B.5c that Method I is slightly inaccurate, while Method II (Figure B.5d) practically finds the exact basis functions. This is confirmed when looking at the reconstruction error for both methods in the two bottom graphs. While the reconstruction error is smaller than the original HD by a factor of approximately 10 for Method I, it falls down to a magnitude of around 10^{-15}. The reason why it does not drop to zero for Method II lies most probably in the finite numerical precision of MATLAB.

The impact of the limited accuracy of Method I on the sound field reconstruction based on the reconstructed HD can be seen from the graphs in Figure B.6. The original sound field for a frequency of $f = 1$ kHz is depicted in Figure B.6a. The sound field based on the HD representations with Methods I and II can be seen in Figures B.6b and B.6c, respectively. The apparent local error does not appear to be overwhelming, yet, when looking at the two bottom plots, it is evident that Method II achieves substantially better results than Method I.

B.4.2 Average Plane Wave Identification Error

The two methods designed to identify the individual directions of travel θ_m in a sound field composed of M plane waves were evaluated w.r.t. the average Parameter Identification Error (PIE). For each estimated parameter θ_m, the PIE in percent is given by

$$\mathrm{PIE} = \frac{|\theta_m - \tilde{\theta}_m|}{\pi} 100\%.$$

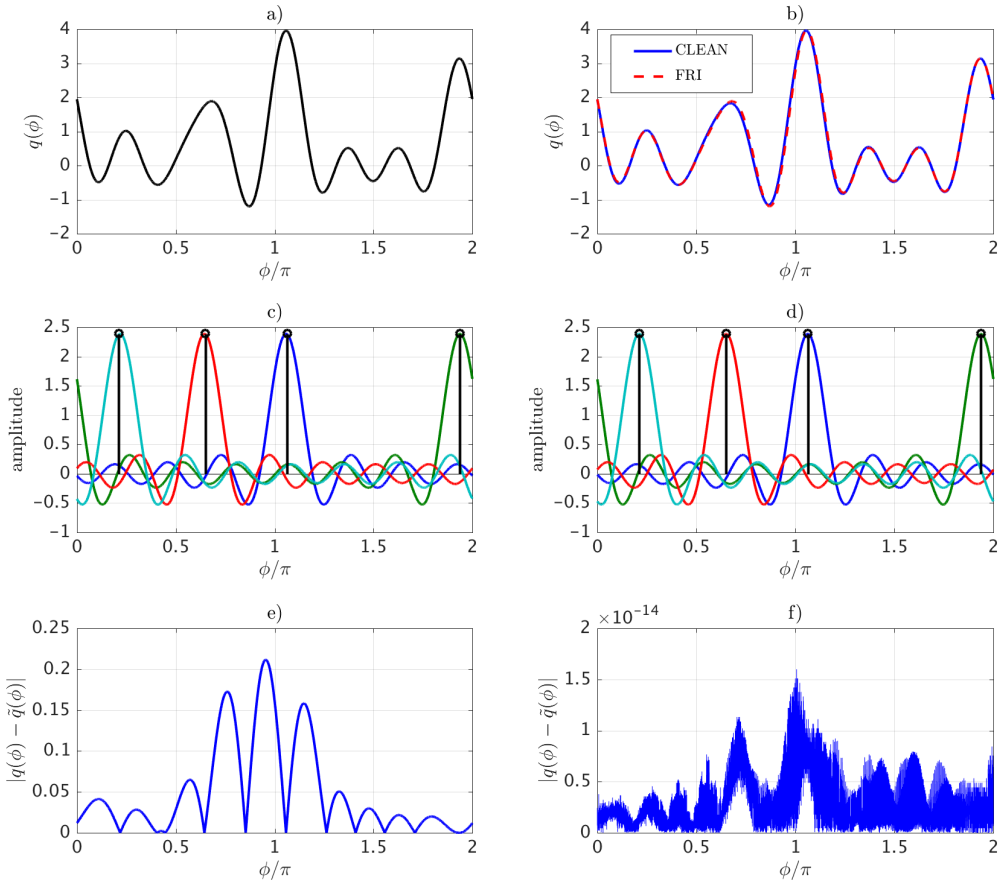


FIGURE B.5: a) Plot of the original HD, b) reconstructions of the HD based on the two bases gained through Method I and II, c) basis functions identified with Method I, d) basis functions identified with Method II, e) reconstruction error of Method I, f) reconstruction error of Method II.

This measure has been statistically investigated in a simulation study for both methods for different numbers of principal plane waves M. Every sample of a sound field was based on ideal Herglotz densities $q(\phi)$ specified by the form given in (B.15) with $L = 15$. Table B.1 shows the results for $M = 4, 5, 6, 7$, where $M = 7$ is the largest number of principal plane waves that can be identified with Method II (compare Sections B.3.2 and B.3.4). The presented results were

Method	$M = 4$	$M = 5$	$M = 6$	$M = 7$
I (CLEAN)	6.4504 %	8.3012 %	9.7109 %	10.5092 %
II (FRI)	$1.8098 \cdot 10^{-9}$ %	$5.0724 \cdot 10^{-8}$ %	$4.9562 \cdot 10^{-4}$ %	$2.1622 \cdot 10^{-4}$ %

TABLE B.1: PIE in percent for Methods I and II for $M = 4, 5, 6, 7$, calculated from 5000 different Herglotz Densities $q(\phi)$, each specified by a set of randomly chosen directions of travel θ_m and complex amplitudes b_m.

averaged over $5000 \cdot M$ individual parameter estimates, where M parameters have been recovered from every HD.

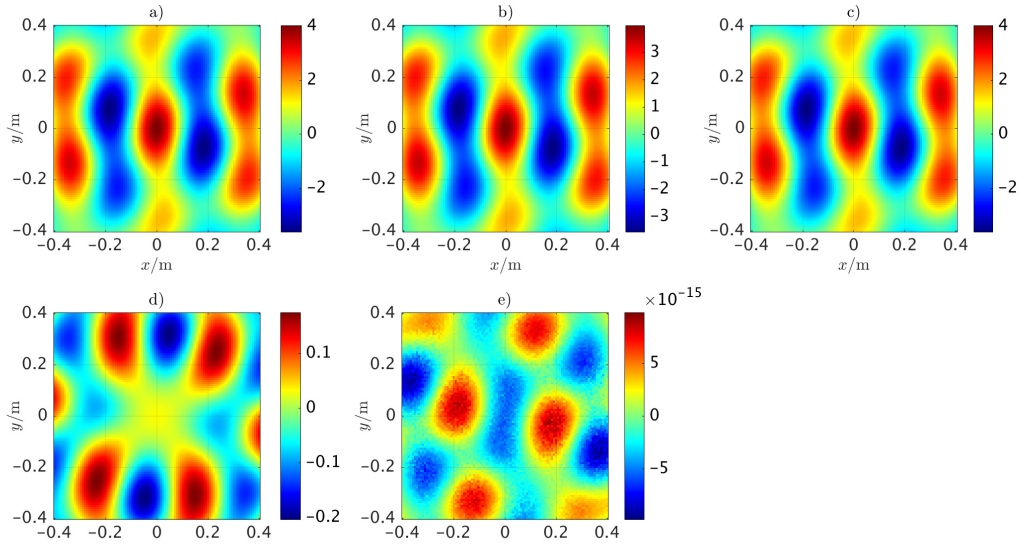


FIGURE B.6: a) Sound field corresponding to the original HD in Fig. B.5, b) sound field corresponding to the HD reconstructed with Method I, c) sound field corresponding to the HD reconstructed with Method II, d) sound field error with Method I, e) sound field error with Method II, $f = 1$ kHz.

B.5 De-Aliasing Stage

The methods developed for the sound field analysis stage to find a set of parameters defining the new basis provide information of different accuracy. It has already been shown by Alon and Rafaely [AR14] that the coefficients defining the beampattern of a beamformer can be optimised to cancel out aliasing effects (grating lobes etc.) that would normally occur in the baseline beam pattern. Like every beamformer, this approach is however set up for a specific look direction. The de-aliasing strategy proposed in this work seeks to de-alias a sound field whose principal components are already known, i.e. they have been acquired in a frequency band that is not corrupted by spatial aliasing.

The success of this strategy is based on the assumption that the sources in a sound field radiate over a wide frequency range. In other words, it is assumed that waves travelling in a direction θ_m in one frequency band are also travelling in the same direction in a different frequency band. In the previous section, these plane waves were denoted principal plane waves. If this assumption is correct and the principal plane waves in corrupted frequency bands have merely a different phase and amplitude from those identified in uncorrupted frequency bands, then aliasing effects can be reduced or even entirely removed. In order to do this, the occurring aliasing needs to be analytically predicted. A method to achieve this objective is discussed in the following subsection.

B.5.1 Predicting Aliasing of a Single Plane Wave

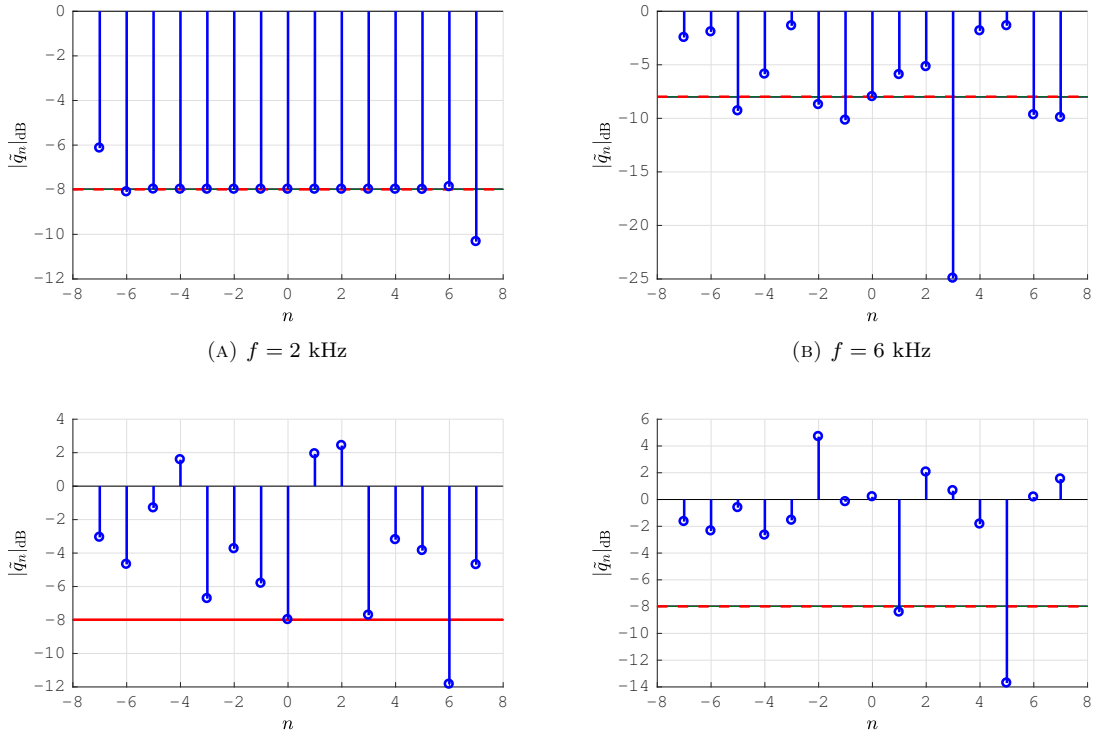
The aliasing behaviour of a circular microphone with TPG sensors has been investigated in Chapter 4, Section 4.3. The recovered coefficients of the HD $\tilde{q}_n$ are corrupted by orders $|n| > \frac{L-1}{2} = N$ as given by Equation (4.36). It has also been argued that for an array of pressure sensors, the result

is the same, only with different eigenvalues. The adequate equation in this case is

$$\tilde{q}_n = q_n + \sum_{\substack{u=-\infty \\ u\neq 0}}^{\infty} \frac{\nu_{n+uL}}{\nu_n} q_{n+uL}, \tag{B.24}$$

with $\nu_n = 2\pi R_n(kr_V)$.

The HD coefficients for a single plane wave can be calculated for any order $n \in \mathbb{N}$ by using Equation (B.12). Together with Equation (B.24), the corrupted coefficients $\tilde{q}_n$ that would be recovered from a simulated measurement of a circular array on a rigid structure with $L = 2N + 1$ sensors and a radius $r_V = 0.1$ m can be calculated. Figure B.7 shows the results for $\tilde{q}_n, n \in [-N \dots N]$ along with a red line indicating the ideal coefficient's magnitude (compare Chapter 4, Section 4.3.5). It can be observed that the coefficients for $f = 2$ kHz are only mildly affected, while the



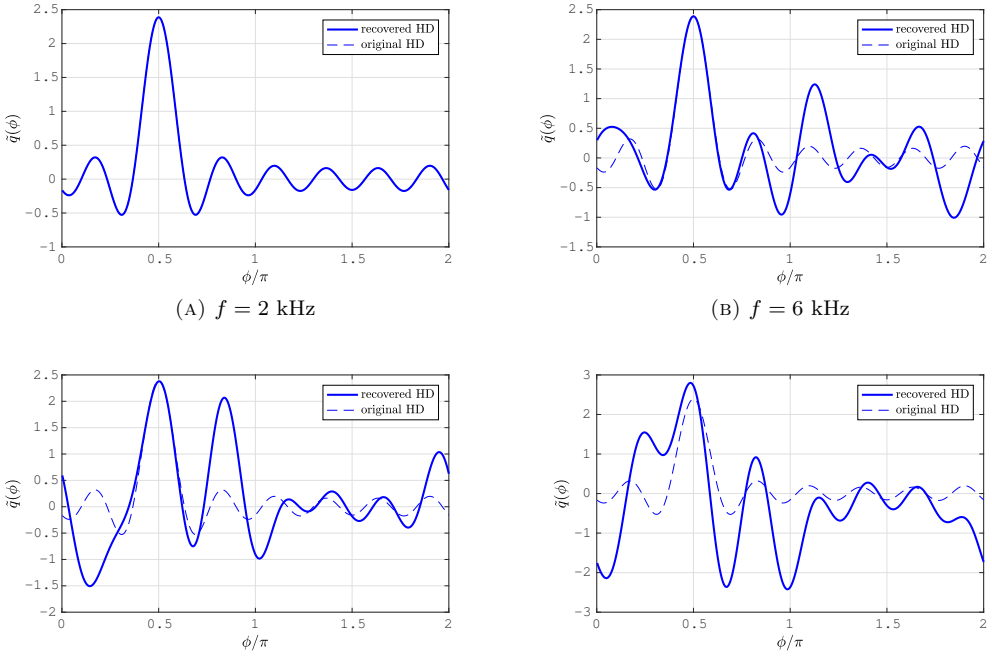
(A) $f = 2$ kHz (B) $f = 6$ kHz

(C) $f = 10$ kHz (D) $f = 16$ kHz

FIGURE B.7: Plots of the recovered Herglotz coefficients $\tilde{q}_n$ of a single plane wave with $\theta = \frac{\pi}{2}$ that were recovered in different frequency bands ($f = 2, 6, 10, 16$ kHz) from a simulated circular array with $L = 15$, $N = 7$, $r_V = 10$ cm and a maximum sound field order of $|n| = N_s = 40$.

plots for the other frequencies (Fig. B.7b, B.7c and B.7d) show severe degradation. The Herglotz densities that can be recovered from the corrupted coefficients $\tilde{q}_n$ are presented in Figure B.8. As it can be expected from the only mildly affected coefficients $\tilde{q}_n$ at $f = 2$ kHz, the corresponding HD in Fig. B.8a has a good resemblance to the ideal HD of a single plane wave (compare Fig. B.1). Equally to be expected is the severe degradation of the HD at even higher frequencies as it can be observed from Figures B.8b, B.8c and B.8d.

Another noticeable aspect of the increasing severity of the aliasing effects towards higher frequencies is the effect it has on the location on the main lobe that was originally centred at the angle



(A) $f = 2$ kHz (B) $f = 6$ kHz

(C) $f = 10$ kHz (D) $f = 16$ kHz

FIGURE B.8: Plots of the recovered Herglotz Densities $\tilde{q}(\phi)$ of a plane wave travelling in the direction $\theta_m = \frac{\pi}{2}$ from coefficients $\tilde{q}_n$ that were recovered in different frequency bands ($f = 2, 6, 10, 16$ kHz) from a simulated circular array with $L = 15$, $N = 7$, $r_V = 0.1$ m and a maximum sound field order of $|n| = N_s = 40$.

corresponding to the direction of travel (DOT) of the plane wave. While at $f = 6$ kHz there is still a distinct main lobe at $\phi = \theta = \frac{\pi}{2}$, at $f = 10$ kHz a side lobe of nearly similar height has emerged next to it. At 10 kHz the effects of aliasing are so dominant that the main lobe is alienated and shifted slightly off the original centre at $\phi = \frac{\pi}{2}$.

The findings above provide the option to predict the aliased shape of the HD of a single plane wave approximation up to the order $|n| = N_s$ travelling in the direction θ. The predicted aliased HD $q_{N_s}^{(a)}(\phi)$ is given by

$$q_{N_s}^{(a)}(\phi) = \frac{1}{2\pi}\mathrm{csinc}_{2N+1,N_s}^{(a)}(\phi-\theta) = \sum_{n=-N}^{N} q_{n,N_s}^{(a)}(\theta)\frac{e^{in\phi}}{\sqrt{2\pi}} \tag{B.25}$$

with

$$q_{n,N_s}^{(a)}(\theta_m) = \sum_{u=\left\lceil\frac{-(N_s+n)}{L}\right\rceil}^{\left\lfloor\frac{N_s-n}{L}\right\rfloor} \frac{\nu_{n+uL}}{\nu_n}\frac{e^{-i(n+uL)\theta}}{\sqrt{2\pi}}, \quad \forall n \in [-N \dots N]. \tag{B.26}$$

Once the aliasing can be anticipated, it is straightforward to show that the complex amplitude b of the original plane wave can be obtained from the corrupted HD $\tilde{q}(\phi)$ by solving the equation

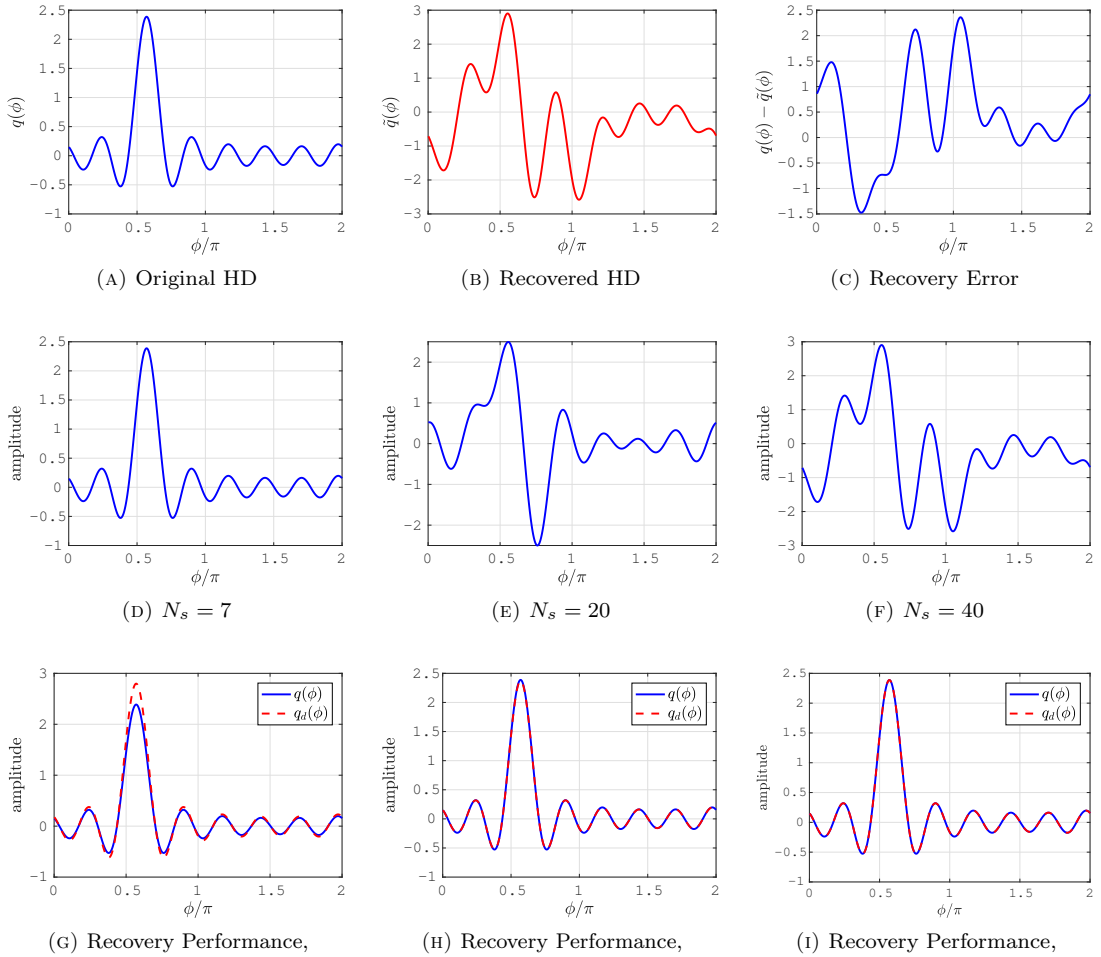
$$\tilde{q}(\phi) = b \cdot \frac{1}{2\pi}\mathrm{csinc}_{L,N_s}^{(a)}(\phi-\theta). \tag{B.27}$$

With the plane wave's complex amplitude b gained from the equation above, the de-aliased HD $q_d(\phi)$ can be reconstructed using Equation (B.11) up to an arbitrary order $|n| = N_R$:

$$q_d(\phi) = b \sum_{n=-N_R}^{N_R} \frac{e^{-in\theta}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}} = b \frac{1}{2\pi} \mathrm{csinc}_{2N_R+1}(\phi - \theta). \tag{B.28}$$

In conclusion, it is possible to de-alias the recovered HD of a single plane wave perfectly and even increase the spatial resolution beyond the maximum order N recovered by the array.

The importance of including all aliased orders when creating the circular sinc function in Equation (B.27) can be observed from the graphs in Figure B.9. The graphs in Figure B.9g, B.9h and



(A) Original HD (B) Recovered HD (C) Recovery Error

(D) $N_s = 7$ (E) $N_s = 20$ (F) $N_s = 40$

(G) Recovery Performance, $N_s = 7$ (H) Recovery Performance, $N_s = 20$ (I) Recovery Performance, $N_s = 40$

FIGURE B.9: a) HD of a single plane wave approximation up the 40th order with $\theta = \frac{4\pi}{7}$ at the frequency $f = 16$ kHz, b) HD as recovered from a simulated circular array with $L = 15$, $N = 7$ and $r_v = 0.1$ m, c) recovery error $q(\phi) - \tilde{q}(\phi)$, d) predicted aliased HD for $N_s = 7$, e) predicted aliased HD for $N_s = 20$, f) predicted aliased HD for $N_s = 40$, g) recovery performance for $N_s = 7$, h) recovery performance for $N_s = 20$, i) recovery performance for $N_s = 40$.

B.9i show the accuracy of $q_d(\phi)$ when the complex amplitude b has been recovered by solving Equation (B.28) for different values of N_s. It can be observed that even the basis function with $N_s = 7 = N$ does not perform too poorly, while the basis function with $N_s = 40$ allows to recover the original HD perfectly. It is shown below, that the choice of the parameter N_s can be used to

adjust the trade-off between robustness and potentially achievable de-aliasing performance when one attempts to de-alias the HD based on several plane wave approximations with imperfect DOT estimates θ.

The following subsection introduces the core method of the de-aliasing stage.

B.5.2 De-Aliasing of Sound Fields with M Plane Waves

In Section B.2 it has been described how the HD of a sound field with M principal plane wave approximations can be represented by a set of M circular sinc functions. The preceding section introduced how the aliasing for a single plane wave approximation can be anticipated and a definition for an 'aliased' circular sinc function has been derived. Due to the linearity of both the sound field model and the inverse operator that recovers the HD coefficients from the sensor signals, the findings above can be combined to anticipate the aliased HD of a sound field with M principal plane wave approximations up to the order $|n| = N_s$. Using Equation (B.25) to replace the circular sinc function in (B.15) yields

$$\begin{aligned}\tilde{q}(\phi) &= \frac{1}{2\pi}\sum_{m=1}^{M} b_m \mathrm{csinc}^{(a)}_{L,N_s}(\phi-\theta_m) \\ &= \mathbf{c}^{(a)}_{N_s}(\phi)\mathbf{b},\end{aligned} \tag{B.29}$$

where

$$\mathbf{c}^{(a)}_{N_s}(\phi) = \frac{1}{2\pi}\left[\mathrm{csinc}^{(a)}_{L,N_s}(\phi-\theta_1) \quad \dots \quad \mathrm{csinc}^{(a)}_{L,N_s}(\phi-\theta_M)\right].$$

and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix}.$$

To recover the amplitudes b_m from the corrupted HD $\tilde{q}_(\phi)$, Equation (B.29) needs to be expanded into an overdetermined linear equation system. Similarly to the system in (B.16), this can be obtained by evaluating (B.29) at $W \gg M$ different angles $\phi_w = w\frac{2\pi}{W}, w \in [1 \dots W]$.

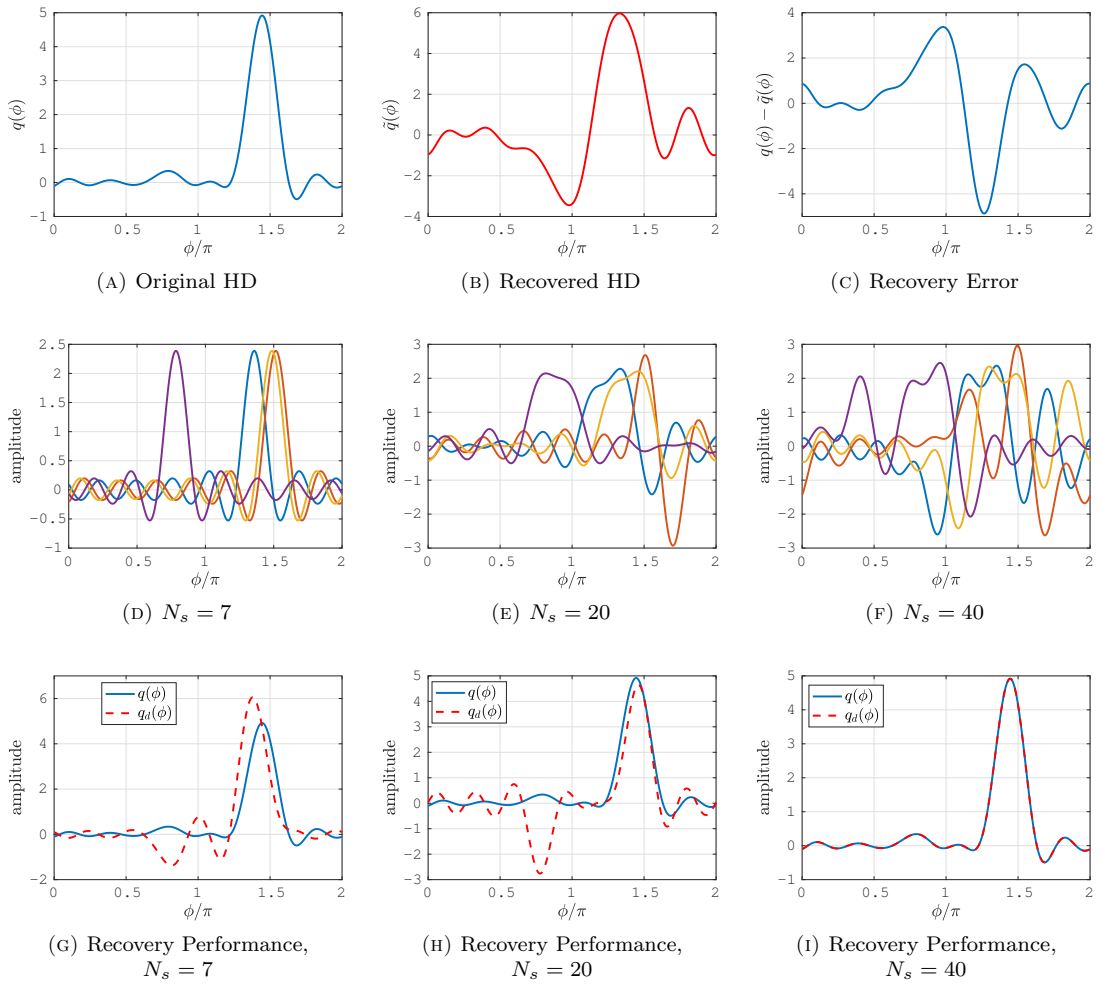
$$\mathbf{q} = \begin{bmatrix} \tilde{q}(\phi_1) \\ \vdots \\ \tilde{q}(\phi_W) \end{bmatrix} = \begin{bmatrix} \mathbf{c}^{(a)}_{N_s}(\phi_1) \\ \vdots \\ \mathbf{c}^{(a)}_{N_s}(\phi_W) \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} = \mathbf{C}\mathbf{b}. \tag{B.30}$$

By applying the Moore-Penrose Pseudoinverse $\mathbf{C}^{\dagger}$ to both sides of the above equation, the vector $\mathbf{b}$ containing the M complex amplitudes b_m can be estimated.

What has been done for the HD of a single plane wave can be equally accomplished for an HD of M plane waves. With the directions of travel θ_m and the corresponding complex amplitudes b_m known, the original HD can be reconstructed up to an arbitrary order $|n| = N_R$ by summing the results of Equation (B.28) for all M plane waves, yielding a solution for the de-aliased HD:

$$q_d(\phi) = \sum_{m=1}^{M} b_m \frac{1}{2\pi}\mathrm{csinc}_{2N_R+1}(\phi-\theta_m) = \sum_{m=1}^{M} b_m \sum_{n=-N_R}^{N_R} \frac{e^{-in\theta_m}}{\sqrt{2\pi}}\frac{e^{in\phi}}{\sqrt{2\pi}}. \tag{B.31}$$

Figure B.10 demonstrates that already a baseline de-aliasing performance with plain circular sinc functions shifted to the right DOTs θ_m reduces the difference between the original HD $q(\phi)$ and the de-aliased HD $q_d(\phi)$ (compare bottom left graph). With increasing value of the



(A) Original HD (B) Recovered HD (C) Recovery Error

(D) $N_s = 7$ (E) $N_s = 20$ (F) $N_s = 40$

(G) Recovery Performance, $N_s = 7$ (H) Recovery Performance, $N_s = 20$ (I) Recovery Performance, $N_s = 40$

FIGURE B.10: a) HD of a sound field with $M = 4$ plane wave approximations up the 40th order with $\theta = \frac{4\pi}{7}$ at the frequency $f = 16$ kHz, b) HD as recovered from a simulated circular array with $L = 15$, $N = 7$ and $r_v = 0.1$ m, c) recovery error $q(\phi) - \tilde{q}(\phi)$, d) predicted aliased circular sinc functions for $N_s = 7$, e) predicted aliased circular sinc functions for $N_s = 20$, f) predicted aliased circular sinc functions for $N_s = 40$, g) recovery performance for $N_s = 7$, h) recovery performance for $N_s = 20$, i) recovery performance for $N_s = 40$.

parameter N_s, which determines the accuracy of the anticipated aliasing, the de-aliased HD $q_d(\phi)$ becomes increasingly similar to the original HD $q(\phi)$ until the de-aliasing performance is perfect for $N_s = 40$.

The designed de-aliasing stage on its own allows for the perfect reconstruction of any original HD by solving the equation system given by (B.30) to find the complex amplitudes of the $M < W$ underlying plane wave approximations as long as the exact DOTs are known.

The next section investigates the combination of the sound field analysis (SFA) stage described in Section B.2 and the de-aliasing stage.

B.5.3 Position of the Side Lobes in the Aliased HD of a Single Plane Wave

One aspect that makes the de-aliasing performance very vulnerable to inaccurate estimates of the DOT parameters is the fact that the shape of an aliased circular sinc function depends on the DOT parameter itself. This can be demonstrated by looking at the aliased HD of a single plane wave, which is also a weighted aliased circular sinc function.

Splitting the right hand side of Equation (B.25) into the basic circular sinc function and the additional component $\epsilon_{L,N_s}(\phi,\theta)$ that occurs due to aliasing yields

$$q_{N_s}^{(a)}(\phi) = \underbrace{\sum_{n=-N}^{N} \frac{e^{-in\theta}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}}}_{\frac{1}{2\pi}\mathrm{csinc}_{2N+1}(\phi-\theta)} + \epsilon_{L,N,N_s}(\phi,\theta),$$

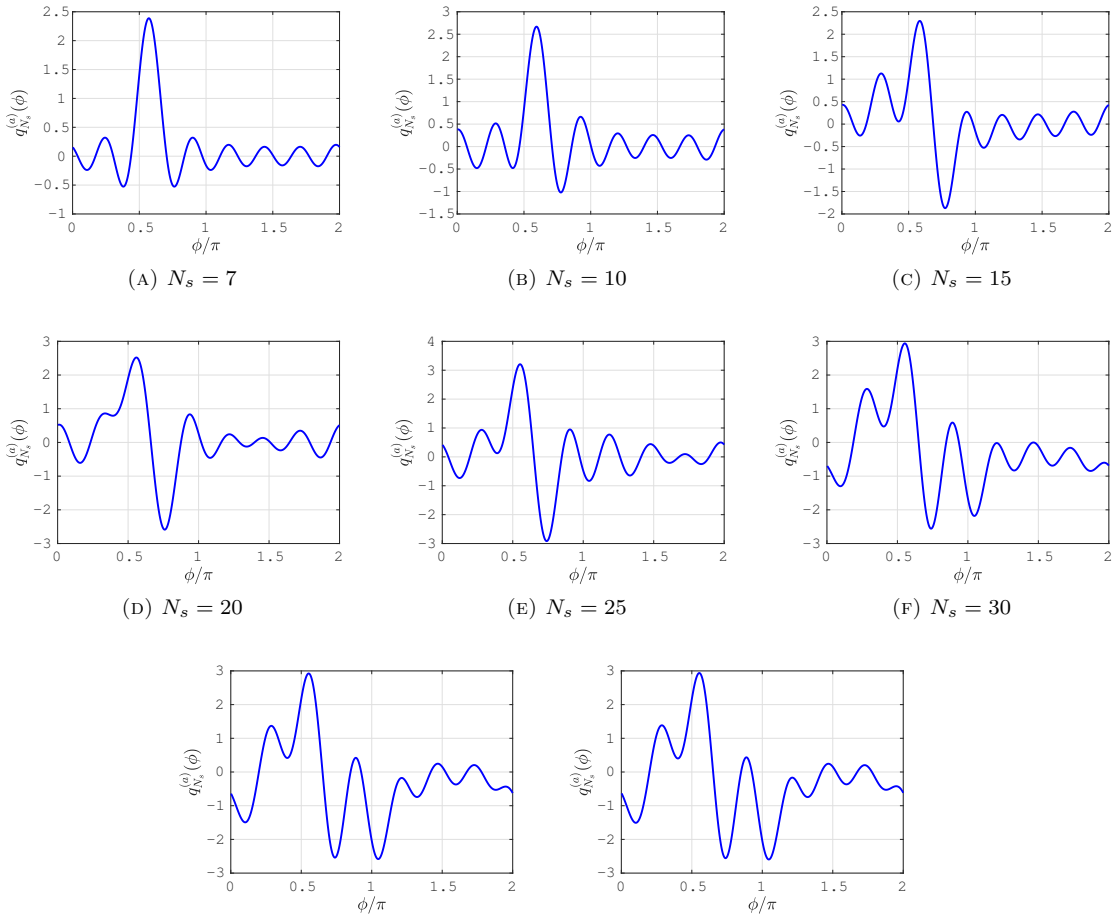
with

$$\epsilon_{L,N,N_s}(\phi,\theta) = \sum_{n=-N}^{N} \sum_{\substack{u=\lceil \frac{-(N_s+n)}{L} \rceil \\ u\neq 0}}^{\lfloor \frac{N_s-n}{L} \rfloor} \frac{\nu_{n+uL}}{\nu_n} \frac{e^{-iuL\theta}}{\sqrt{2\pi}} \frac{e^{in(\phi-\theta)}}{\sqrt{2\pi}}. \tag{B.32}$$

The definition above clearly indicates that the shape of $\epsilon_{L,N,N_s}(\phi,\theta)$ depends on θ (i.e. is not translation invariant).

This result is confirmed by the graphs in Figure B.11 that display the shape of the aliased HD of a single plane wave travelling in direction $\theta = \frac{4\pi}{7}$ for different maximum sound field orders N_s. For the case of $N_s = 7 = N$ there is no effect of aliasing as can be seen from Figure B.11a. As the value of N_s increases, the overall shape changes progressively up to $N_s = 35$. When comparing Figure B.11g and B.11h, it can be seen that there is only little difference despite the higher sound field order. This effect can be explained when looking at the fraction $\frac{\nu_{n+uL}}{\nu_n}$ in Equation (B.32), which is the weighting factor that defines how the energy contained in the $n+uL$th mode is transferred into the nth mode. From Figure 4.3 it can be anticipated that, for a given frequency f, the absolute value of this weighting factor is going to quickly fall towards zero for modes of orders $|n+uL| \gg |n|$. This effect is directly related to implications of the $N = kr_V$ rule [WA01], which says that for a plane wave the energy of modes $|n| > kr_V$ drops to a negligible level on the measurement boundary ∂V and therefore they do not contribute to aliasing effects. This behaviour is fully modelled by the factor $\frac{\nu_{n+uL}}{\nu_n}$ in Equation (B.32).

It has been shown that the position of the side lobes that occur as an effect of spatial aliasing is a function of the DOT. Since the de-aliasing stage attempts to recover the individual plane waves' complex amplitudes through reconstruction of the observed HD by using the aliased circular sinc functions of the respective DOT (compare Equation (B.30)), inaccurate parameters for the DOT must degrade the performance of the de-aliasing stage. This is confirmed by the findings presented in Section B.6.



(A) $N_s = 7$ (B) $N_s = 10$ (C) $N_s = 15$

(D) $N_s = 20$ (E) $N_s = 25$ (F) $N_s = 30$

(G) $N_s = 35$ (H) $N_s = 40$

FIGURE B.11: Plots of the aliased Herglotz Densities $q_{N_s}^{(a)}(\phi)$ of a plane wave travelling in the direction $\theta = \frac{4\pi}{7}$ for $f = 16$ kHz, $L = 15$, $N = 7$, $r_V = 0.1$ m and different maximum sound field orders $N_s \in [7, 10, 15, 20, 25, 30, 35, 40]$.

B.6 Performance Study of the Combination of the SFA Stage and the De-Aliasing Stage

It has been shown in the previous sections that the performance of the de-aliasing stage depends predominantly on the accuracy of the DOT estimates $\tilde{\theta}_m$. The nature of a method based on estimated parameters is that its performance cannot easily be anticipated by analytical calculations. Hence, the proposed de-aliasing strategy has been tested against unprocessed data in a simulation study over 10000 randomly generated sound fields (i.e. randomly chosen directions of travel and complex amplitudes, respectively,) consisting of different numbers of plane waves that were reproduced up to the 40th order. The evaluation of the performance was based on the error-to-signal-ratio (ESR) of the individual recovered HD. The SFA stage was generally set to identify all M plane waves in the sound field. However, for the FRI parameter estimation method, no more than $M = N = 7$ plane waves can be identified (see Section B.3.4), therefore for more than $N = 7$ plane waves and when based on the FRI parameter estimation, de-aliasing strategy is restricted to a basis with only $N = 7$ plane waves, even though the sound field actually contains more.

It is assumed that the sources in the sound field radiate across a wide frequency spectrum, so that a DOT identified at a frequency unaffected by spatial aliasing does also corresponds to the same source at higher frequencies where spatial aliasing is a problem. In this performance study, the parameters for the DOT were estimated from the HD coefficients obtained at $f = 1\,\text{kHz}$. These estimates were then used to perform de-aliasing at $f = 16\,\text{kHz}$. The complex amplitudes of the individual plane waves were randomly chosen such that their magnitudes fall into the interval $[0.1, 2]$.

B.6.1 Definition of the Error-To-Signal Ratio

The error-to-signal-ratio (ESR) of a recovered HD $q_R(\phi)$ is equivalent to the ESR defined in Equation (4.64) with the only difference that it is defined for the HD instead and is given as

$$\text{ESR} = 10 \cdot \log_{10} \left(\frac{\sum_{w=1}^{W} |q(w\frac{2\pi}{W}) - q_R(w\frac{2\pi}{W})|^2}{\sum_{w=1}^{W} |q(w\frac{2\pi}{W})|^2} \right). \tag{B.33}$$

For the case of the calculated ESR, the original HD $q(\phi)$ was calculated up to the order $N = 7$, which is exactly the order a circular array with $L = 15$ sensors is supposed to recover. The de-aliased Herglotz Density $q_R(\phi)$ is either equal to the output of the de-aliasing stage $q_d(\phi)$ or the unprocessed HD $\tilde{q}(\phi)$ recovered from the array observation, respectively. The latter will be referred to as the baseline performance. The reproduction order of the de-aliasing stage was set to $N_R = N = 7$, i.e. no super-resolution was attempted.

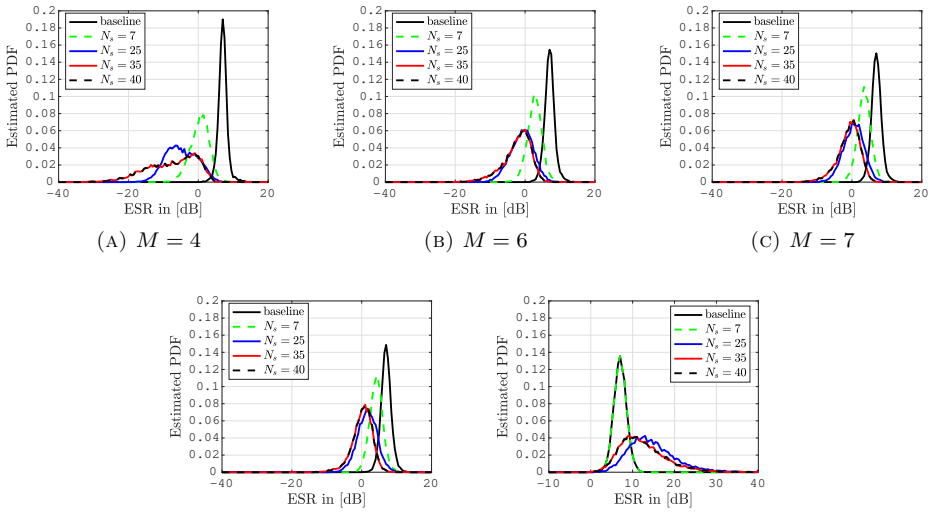
B.6.2 Statistical Performance Analysis

The average results for the achieved ESR are given in Table B.2 for the de-aliasing strategy (DS) and the baseline performance.

B.6.2.1 Results for the De-Aliasing Strategy With CLEAN Parameter Estimation

Figure B.12 shows the normalised histograms for the DS with the CLEAN based DOT parameter estimation. For the case of $M = 4$ plane waves, the ESR results for the baseline performance are concentrated between 0 and 10 dB, with an average ESR of 7.27 dB. The performance of the de-aliasing strategy (DS) with its various settings for N_s shows already an improvement over the baseline performance, with an average ESR decrease of $6.06 - 10.51\,\text{dB}$ compared to the baseline performance. However, the covered ESR ranges shown in the histograms are considerably wider than that of the baseline performance. The results for $N_s = 35$ and $N_s = 40$ show only little difference.

For $M = 6$ plane waves, the main body of the histogram for the baseline performance has changed only slightly. Again, the results for the DS clearly show an improvement over the baseline performance, yet is the span of ESR decrease smaller than that for $M = 4$. Once more the results for $N_s = 35$ and $N_s = 40$ show only little difference. However, the average gain over the baseline performance has reduced to just $4.05 - 8.18\,\text{dB}$ improvement.



(A) $M = 4$ (B) $M = 6$ (C) $M = 7$

(D) $M = 8$ (E) $M = 15$

FIGURE B.12: Histograms (or estimated Probability Density Functions (PDF)) of the Error-to-Signal-Ratio (ESR) for the De-Aliasing Strategy based on the CLEAN parameter estimation for various M and different aliasing model parameters N_s, for $f = 16\,\text{kHz}$ and the array parameters $L = 15$, $N = 7$, $r_V = 0.1\,\text{m}$.

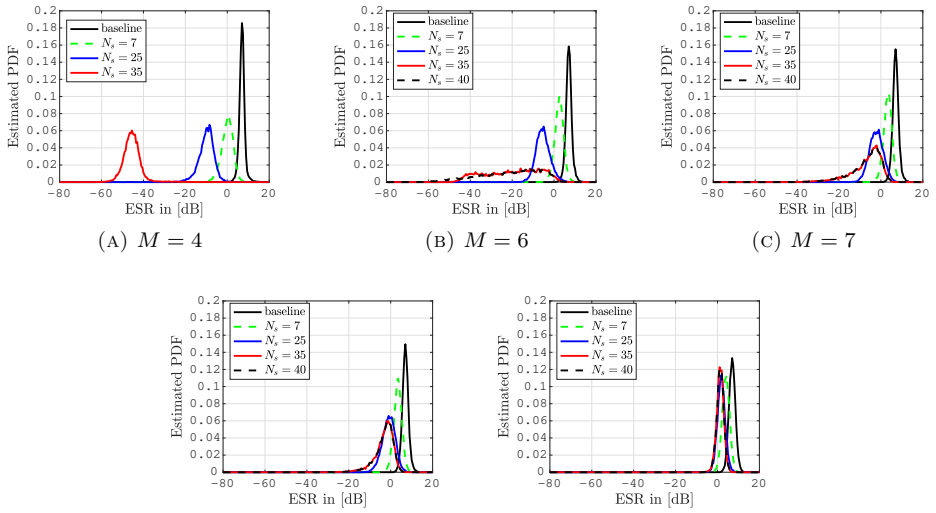
For $M = 7 = N$ and $M = 8$, the results show only little difference in the shape of the histograms compared to those for $M = 6$. From Table B.2 it can be observed that the average ESR of the DS shifts upwards with increasing number of plane waves.

For $M = 15 = L$, it can be seen that the baseline performance and that of the DS for $N_s = 7$ show almost exactly the same results. At the same time, the performance of the DS for $N_s = 25$, $N_s = 35$ and $N_s = 40$ has severely degraded, with the average ESR for all of them lying beyond $19.35\,\text{dB}$.

Both the histograms and the average ESR results of the DS with $N_s = 35$ and $N_s = 40$ show hardly any difference. This was to be anticipated, since it has already been shown in Section B.5.3 that the corruption of orders higher than 35 is negligible.

B.6.2.2 Results for the De-Aliasing Strategy With FRI Parameter Estimation

Figure B.13 shows the normalised histograms for the DS with the FRI based DOT parameter estimation. The histograms of the DS for the case of $N_s = 40$ have not been displayed for $M = 4$, as the average ESR lies below $-35\,\text{dB}$ (see Table B.2). The results of the baseline performance are inherently the same as before and only displayed for the sake of comparison. For the case of $M = 4$ and $M = 6$ plane waves, the performance of the de-aliasing strategy (DS) with its various settings for N_s shows significant improvement over the baseline performance. Unlike with the CLEAN parameter estimation, the results of the DS for different parameters N_s are much more distinct. This can be accredited to the precision of the parameter estimation of the FRI method. Its accurate parameter estimates allow for the potential of the aliasing model to clearly show in the histograms, as the (estimated) Probability Density Functions (PDF) for the ESR shift towards lower values on the abscissa with increasing value for N_s. The increase in number of plane waves from $M = 4$ to $M = 6$ causes the average ESR for the DS to increase differently for



(A) $M = 4$ (B) $M = 6$ (C) $M = 7$

(D) $M = 8$ (E) $M = 15$

FIGURE B.13: Histograms (or estimated Probability Density Functions (PDF)) of the Error-to-Signal-Ratio (ESR) for the De-Aliasing Strategy based on the FRI parameter estimation, where the simulations encompassed 10000 randomly generated Herglotz Densities, which have been taken out for various M and different aliasing model parameters N_s, for $f = 16\,\text{kHz}$ and the array parameters $L = 15$, $N = 7$, $r_V = 0.1\,\text{m}$.

the various values for N_s. In case of $N_s = 7$, the average ESR increases by about 2 dB, while for $N_s = 40$ the difference comes to almost 30 dB.

For $M = 7 = N$, the results of the DS for $N_s = 35$ and $N_s = 40$ show very similar results, with a vast spread from around -40 dB to just over 0 dB with an average ESR of approximately -3.37 dB. It appears that the accuracy of the FRI parameter estimation has mildly degraded and hence vanishes the difference between the performance for $N_s = 35$ and $N_s = 40$. Once more, the average ESR for $N_s = 7$ increased by little more than 0.8 dB due to the higher number of plane waves.

For $M = 8$, the results of the DS for $N_s = 35$ and $N_s = 40$ again show very similar results but would no longer spread down to values as low as before. It may be predicted that with further increase of M, eventually the performance of the DS would be the same for all four parameter values $N_s = 7$, $N_s = 25$, $N_s = 35$ and $N_s = 40$. This is confirmed by the histograms for $M = 15$ and by the results shown in Table B.2. However, it should be noted that the average ESR performance of the DS with the FRI parameter estimation for $M = 15$ outperforms that of the DS combined with the CLEAN parameter estimation, regardless of the chosen parameter N_s.

B.7 Conclusions

The developed de-aliasing strategy is divided into two stages. The first stage analyses a successfully recovered Herglotz Density as to its principal plane wave components. Each principal plane wave is defined by two parameters: an angle defining the direction of travel and a complex amplitude indicating amplitude and phase. It has been shown that the information gained from a microphone array measurement in the sound field with M plane waves can be perfectly de-aliased

Method	ESR_{dB}	ESR_{dB}, $\mathbf{N_s = 7}$	ESR_{dB}, $\mathbf{N_s = 25}$	ESR_{dB}, $\mathbf{N_s = 35}$	ESR_{dB}, $\mathbf{N_s = 40}$
M = 4					
CLEAN	-	1.21	-2.63	-3.51	-3.51
FRI	-	0.85	-8.67	-35.99	-36.56
Baseline	7.27	-	-	-	-
M = 6					
CLEAN	-	3.24	0.98	-0.89	-0.89
FRI	-	2.92	-3.80	-8.41	-8.41
Baseline	7.29	-	-	-	-
M = 7					
CLEAN	-	3.94	1.47	0.11	0.11
FRI	-	3.73	-1.05	-3.37	-3.37
Baseline	7.31	-	-	-	-
M = 8					
CLEAN	-	4.80	2.51	0.99	0.98
FRI	-	3.84	0.09	-1.53	-1.53
Baseline	7.29	-	-	-	-
M = 15					
CLEAN	-	7.32	22.33	19.35	20.90
FRI	-	4.32	2.22	1.61	1.61
Baseline	7.29	-	-	-	-

TABLE B.2: Average ESR at the output of the de-aliasing strategy and the baseline performance, where the simulations encompassed 10000 randomly generated Herglotz Densities, which have been taken out for various M and different aliasing model parameters N_s, for $f = 16\,\mathrm{kHz}$ and the array parameters $L = 15$, $N = 7$, $r_V = 0.1\,\mathrm{m}$.

when their directions of travel are known. This is possible because the knowledge of the aliasing pattern of a given array allows to anticipate the corruption inflicted by spatial aliasing on the observation of a plane wave travelling into a given direction. And since the corruption can be anticipated, the original information can also be restored by subtracting the erroneous information caused by the corruption.

It has also been shown that the success of perfectly de-aliasing an array observation depends highly on the exact knowledge of the directions of travel of the principal plane waves. Consequently, the quality of the de-aliasing strategy is mainly determined by the quality of the sound field analysis stage. However, the de-aliasing stage can be made more robust to inaccuracies of the input parameters, i.e. the estimated directions of travel of the principal plane wave components, by sacrificing the option of a perfect de-aliasing performance. It has been shown that, in the case of an inaccurate set of DOT estimates, it may lead to better de-aliasing results if the de-aliasing stage does not include the effects of the full aliasing pattern. In fact, this can be adjusted by just one parameter that limits the number of higher orders contributing to the aliasing pattern. That will of course lead to an inaccurate prediction of the aliasing effects. However the results suggest that the overall performance of the de-aliasing strategy with inaccurate DOT estimates may be better with an incomplete prediction of the aliasing, than with an entire prediction of the aliasing with the wrong DOT estimate. This can most probably be attributed to the shape of the aliased basis functions. The side lobes occurring in the representation of the principal plane waves due to spatial aliasing may become higher than the main lobe. This point is supported by the observation that the position of the side lobes also depends on the DOT. It can therefore be expected that a least-squares decomposition based on such a set of basis function would potentially lead to an even worse performance than the baseline offered by the array on its own.

Furthermore, it has been suggested that the number of higher orders that contribute to aliasing is limited by the well-known spatial lowpass behaviour of the operator defining the mapping of the HD to the pressure on the measurement boundary. That limits the higher contributions in the observed pressure domain.

Given the importance of the quality of the DOT estimates, two methods to recover the set of parameters from the array observation have been proposed and investigated. Method I is similar to the family of the CLEAN algorithms and theoretically allows for an infinite number of principal plane wave components. However, due to a lack of orthogonality of the different basis functions of which each corresponds to a principal plane wave component, it must be assumed that the de-aliasing performance will eventually degrade down to the baseline performance achieved by the array. Method II is based on a finite rate of innovation signal model and allows for the exact recovery of the principal plane waves' directions of travel. This is however limited by the fact that it can only detect a finite number of principal plane waves, where this number stands in a fixed relation to the number of sensors of the array. It must also be assumed that this method is vulnerable to the sound field being composed - even partially - of anything but perfect plane waves, which is typically the case. This method is thus so far only deemed of theoretical relevance.

Appendix C

Cross-Correlation of the HD and the Weighted Circular Sinc Function

Let the HD for a sound field composed of M plane waves be of the form

$$q(\phi) = \sum_{m=1}^{M} a_m \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta_m). \tag{C.1}$$

The cross-correlation function between the HD and the weighted circular sinc function $\frac{1}{2\pi}\mathrm{csinc}_{2N+1}(\phi)$ is then given by

$$\begin{aligned}
R_{\mathrm{HD,csinc}}(\theta) &= \int_0^{2\pi} \sum_{m=1}^{M} a_m \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta_m) \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta) d\phi \\
&= \sum_{m=1}^{M} a_m \int_0^{2\pi} \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta_m) \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\phi - \theta) d\phi \\
&= \sum_{m=1}^{M} a_m \frac{1}{4\pi^2} \sum_{n=-N}^{N} e^{-in\theta_m} \sum_{n'=-N}^{N} e^{-in'\theta} \underbrace{\int_0^{2\pi} e^{i(n+n')\phi} d\phi}_{2\pi\delta_{n+n'}} \\
&= \sum_{m=1}^{M} a_m \frac{1}{2\pi} \sum_{n=-N}^{N} e^{in\theta} e^{-in\theta_m} \\
&= \sum_{m=1}^{M} a_m \frac{1}{2\pi} \mathrm{csinc}_{2N+1}(\theta - \theta_m).
\end{aligned}$$

When substituting the variable θ by ϕ in the result for the cross-correlation function, then it can be seen that it is equal to the HD itself.

Appendix D

A Multichannel Low-Noise Signal Conditioning Amplifier for Microphone Arrays

In the context of the work on microphone arrays, a multichannel signal conditioning amplifier for the use with microphone arrays was developed. The design is a joint development of the fellow Ph.D. student, Fábio Casagrande Hirono, and the author. An overall number of 64 channels were built and arranged into multichannel units (16 channels each) by the electronic workshop of the ISVR.

The circuitry is based on the *THAT 1570 Low-Noise Differential Audio Preamplifier IC*. It was the design objective to create a flexible unit that can be combined with various types of microphones, while maintaining an low self noise and high linearity.

D.1 Circuit Design & Specifications

D.1.1 Full Electronic Circuit

The main circuit of the conditioning amplifier is shown in Figure D.1. The additional circuitry on the input side is needed to shield the conditioning amplifiers' input terminals from radio frequency interference and protect the instrumentation amplifier from phantom power supply failure. Figures D.2 and D.3 show the gain setting stage and the implementation of the switching mechanism through a dedicated integrated circuit, respectively.

D.1.2 Microphone Conditioning Amplifier: Specifications

- 16 channel microphone pre-amplifier
- Phantom Power: 48 Volt for studio/pro audio equipment or 5 Volt for electret capsules (see D.1.3), supplied directly from PSU

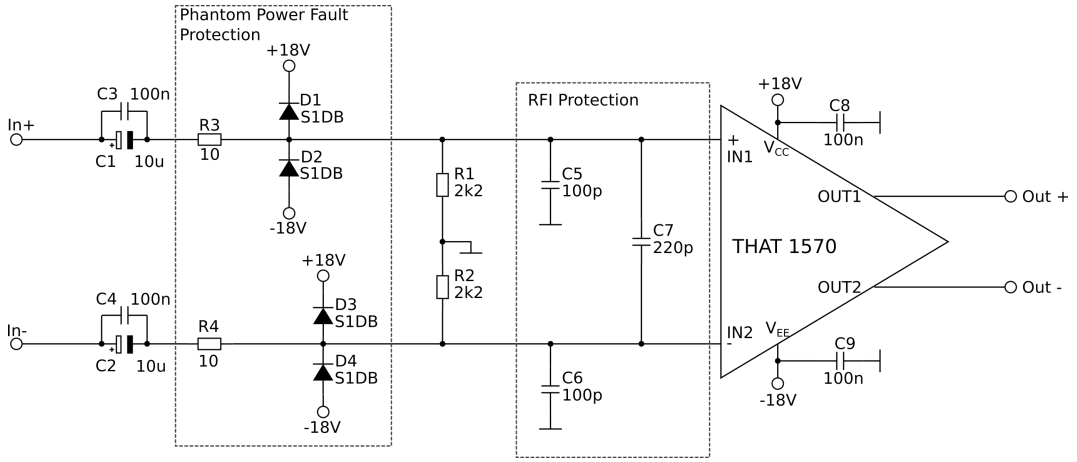


FIGURE D.1: Schematic of the conditioning amplifier.

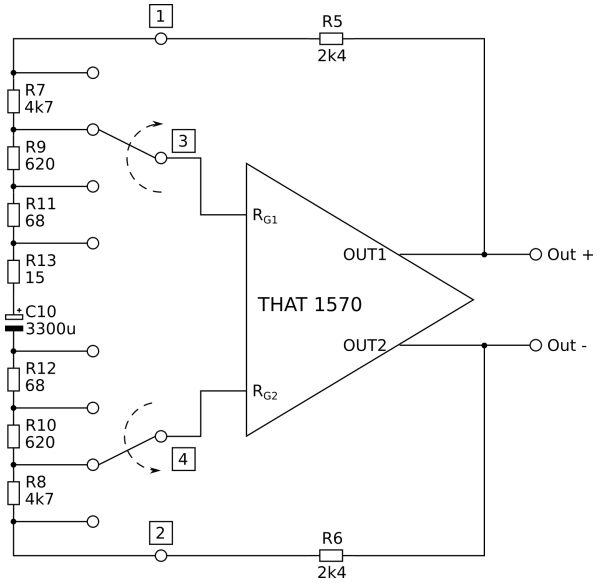


FIGURE D.2: Schematic of the gain control stage.

- switchable gain with either 3 dB, 20 dB, 40 dB or 60 dB gain for ALL 16 channels
- 1 x Sub-D-78 Connector multi-core balanced input
- 2 x Sub-D-25 Connectors multi-core balanced output (TASCAM's DB25 analogue standard) with 8 channels each
- 16 x balanced TRS (6.3 mm Jack) outputs

D.1.3 Power Supply Unit (PSU): Specfications

- General PSU for up to four 16-Channel microphone pre-amplifier units, ± 18 Volt
- 48 Volt phantom power for up to four 16-channel microphone pre-amplifier units
- 5 Volt phantom power for up to four 16-channel microphone pre-amplifier units

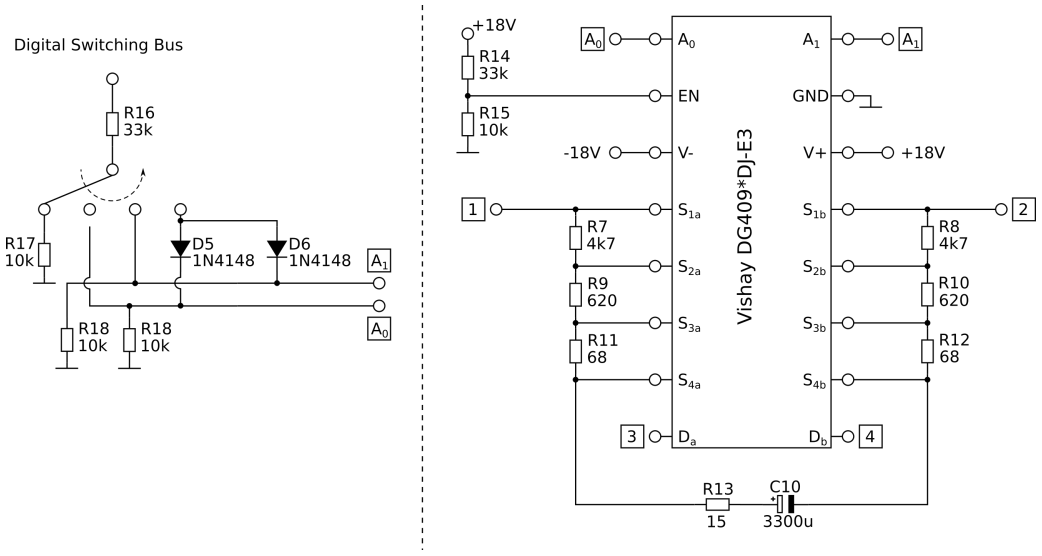


FIGURE D.3: Schematic of the gain switching circuit.

D.1.4 Sub-D-78 Connector: Pin-Layout

The chosen sub-d connector was used to enable for the connection of up to 16 microphones per amplifier unit. Using different internal wirings, the conditioning amplifier can be connected to either standard electret capsules, professional audio microphones (with or without P48 phantom power) and also *Brüel & Kjær* (B&K) *Type 4189-L-001* microphones.

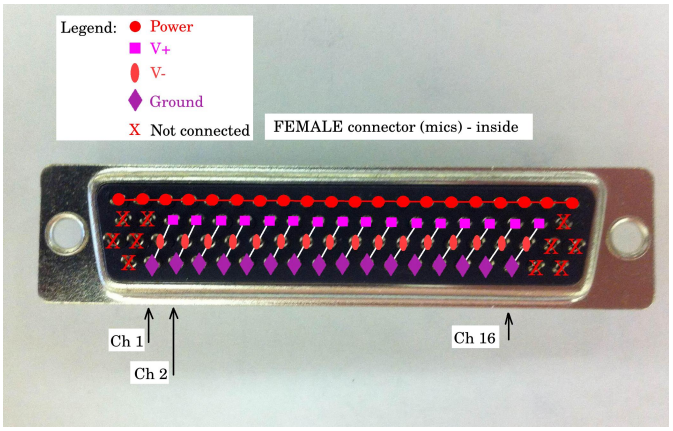
The exact pin layout for the male and female connector is shown in Figure D.4.

D.2 Performance Specifications

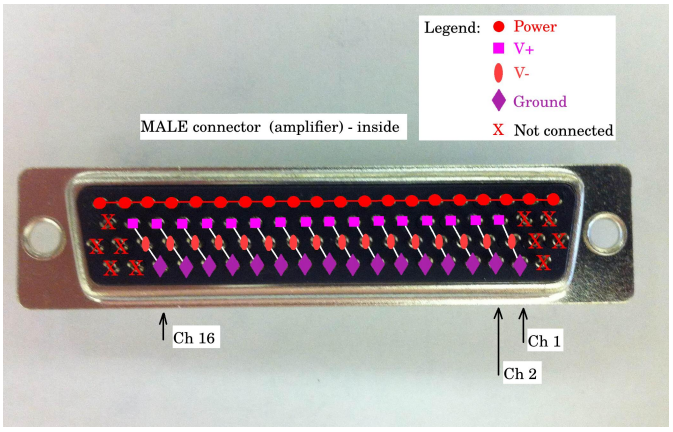
The performance specifications of the developed conditioning amplifier were determined through measurements with an *Audio Precision APx Series Audio Analyzer*, which was kindly provided by Tom Bell, *Bowers & Wilkins*.

D.2.1 Conditioning Amplifier

- input impedance: 2.2 kΩ
- Lower edge frequency: -3 dB @ 7.13 Hz
- maximum input signal level: +23.5 dBu
- maximum output signal level: +26.5 dBu
- Gain settings tolerance < 0.5 dB.
- Self Noise at the Output with the Input fixed on GND:
  - approx. -140 dBu @ 3 dB Gain



(A) Female Connector



(B) Male Connector

FIGURE D.4: Pin layout for the Sub-D-78 connector.

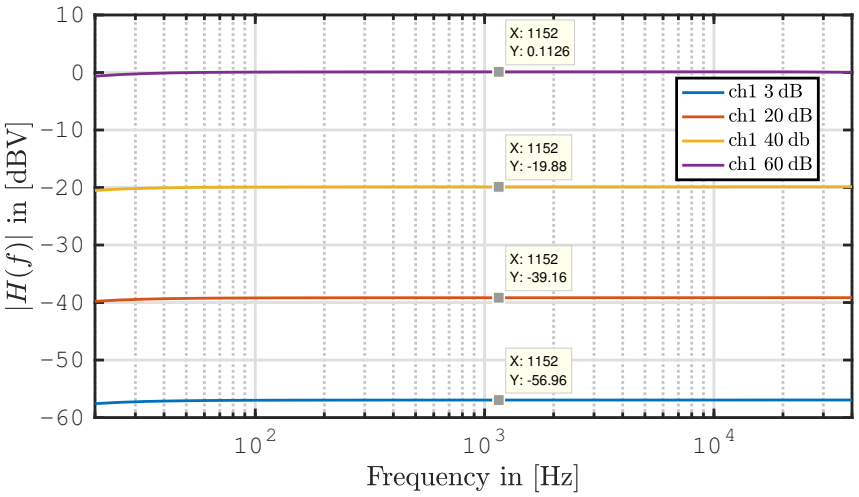
  - approx. −120 dBu @ 20 dB Gain
  - approx. −100 dBu @ 40 dB Gain
  - approx. −80 dBu @ 60 dB Gain

- Output Signal to THD+Noise Ratio (at 1 kHz for 1.25 mV RMS (≡ −55.85 dBu) at the input):
  - -111 dB @ 3 dB Gain
  - -117.8 dB @ 20 dB Gain
  - -125.7 dB @ 40 dB Gain
  - -127.8 dB @ 60 dB Gain

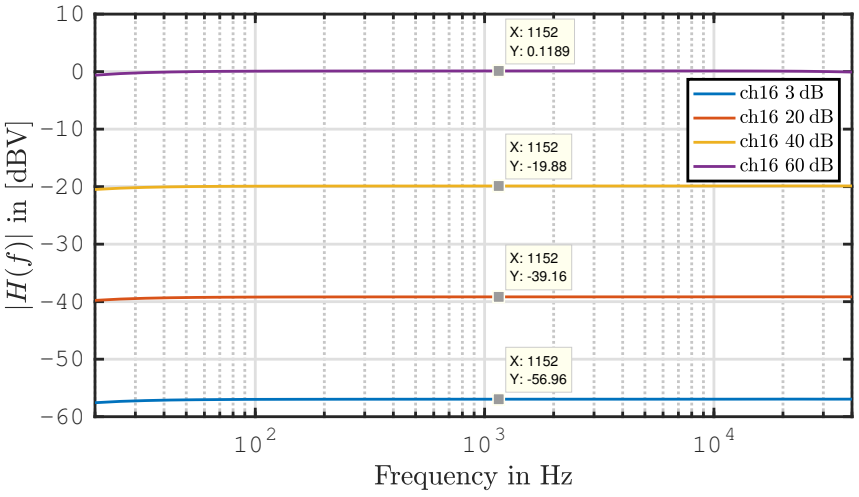
- Crosstalk into the output of Channel 2 for a signal level of −6 dBu at the output of Channel 1: less than −100 dBu within the audio band.

D.2.2 Frequency Response & THD+Noise Measurements

The frequency response was measured for Channels 1 and 16 of one of the four units. The results are shown in Figure D.5. As can be seen, the conditioning amplifier has a flat response within



(A) Frequency Response of Channel 1



(B) Frequency Response of Channel 16

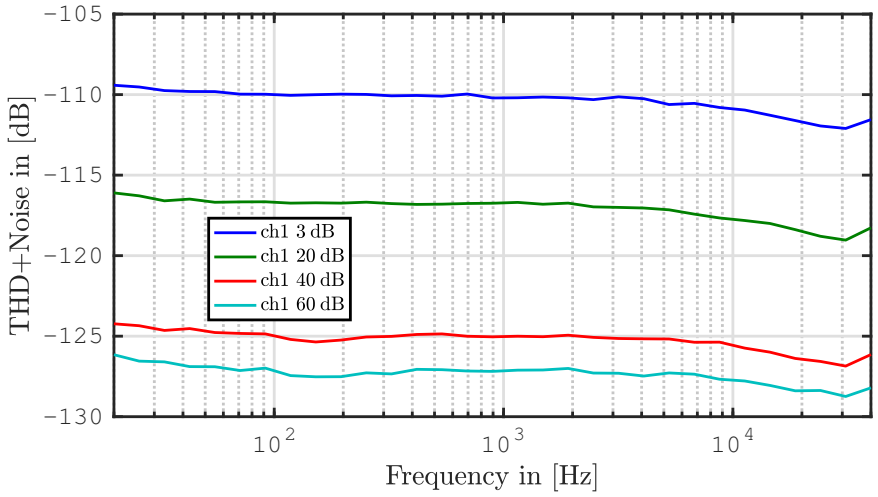
FIGURE D.5: Frequency responses of channels 1 and 16 of all four gain settings for an input signal with 1.25 mV RMS ($\equiv -55.85$ dBu) within the frequency band from 20 Hz up to 40 kHz.

the frequency range from 30 to 40 kHz.

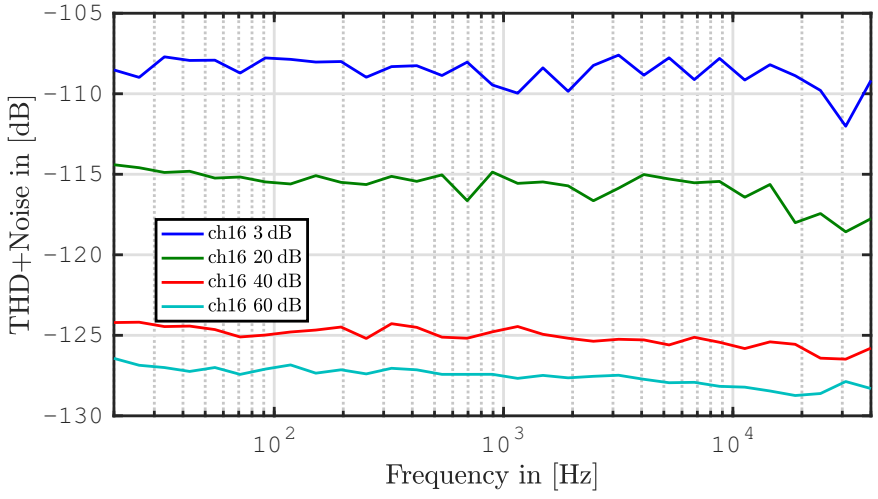
The results for the *Total Harmonic Distortion* (THD) + Noise measurements are shown in Figure D.6.

D.2.3 Power Supply Unit (PSU)

- maximum current from ±18 Volt supply: 1 A
- maximum current from 48 Volt phantom power: 1.4 A



(A) THD+Noise of Channel 1



(B) THD+Noise of Channel 16

FIGURE D.6: THD+Noise of channels 1 and 16 of all four gain settings for an input signal with 1.25 mV RMS (≡ −55.85 dBu) within the frequency band from 20 Hz up to 40 kHz.

Appendix E

Design & Realisation of a High-Quality Hemi-Cylindrical Array

A variable circular loudspeaker array with 32 loudspeakers in an interleaved arrangement has been designed. For drivers *Tang Band W1-1070-SG units* were chosen and the cabinets have been carefully designed to allow for the widest possible bandwidth. The interleaved design yields an effective increase in spatial resolution (reproduction of higher spatial orders) while the aperture radius remains the same. This newly designed array was constructed in the ISVR workshops.

E.1 Design Sketches

The design is specified in Figures E.1, E.2 and E.3. It allows for the array's use as either

- a full array with 32 interleaved speakers, or
- two half arrays with 15 interleaved speakers each.

The full structure has a size of just under 30 cm in diameter. As can be seen from Fig E.1, each individual cabinet is separated from the neighbouring cabinets by a wall involving 2 layers of 3 mm plywood that are connected by a layer of rubber adhesive applied near the edges to obtain a 1.5 mm air gap between them for improved acoustic decoupling. The front panels of each individual cabinet in the array are going to be made from aluminium to allow for a neater integration of the new drivers into the cabinet, preserving its intended acoustical and mechanical properties.

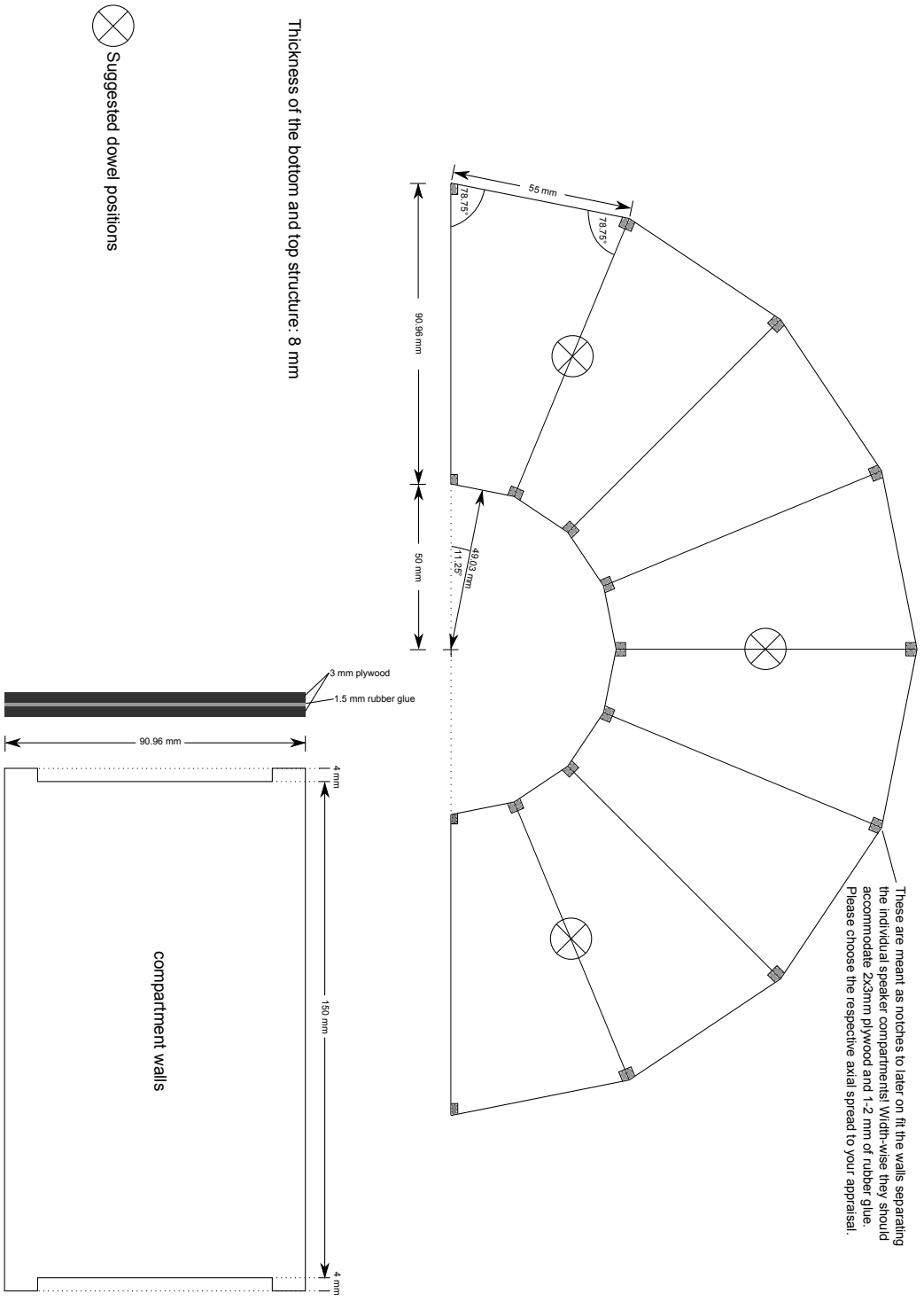


FIGURE E.1: Design of the lower layer of the Mirrorsys prototype.

FIGURE E.2: Design of the upper layer of the Mirrorsys prototype, with filler wedges for both, hemi-cylinder array and full array applications.

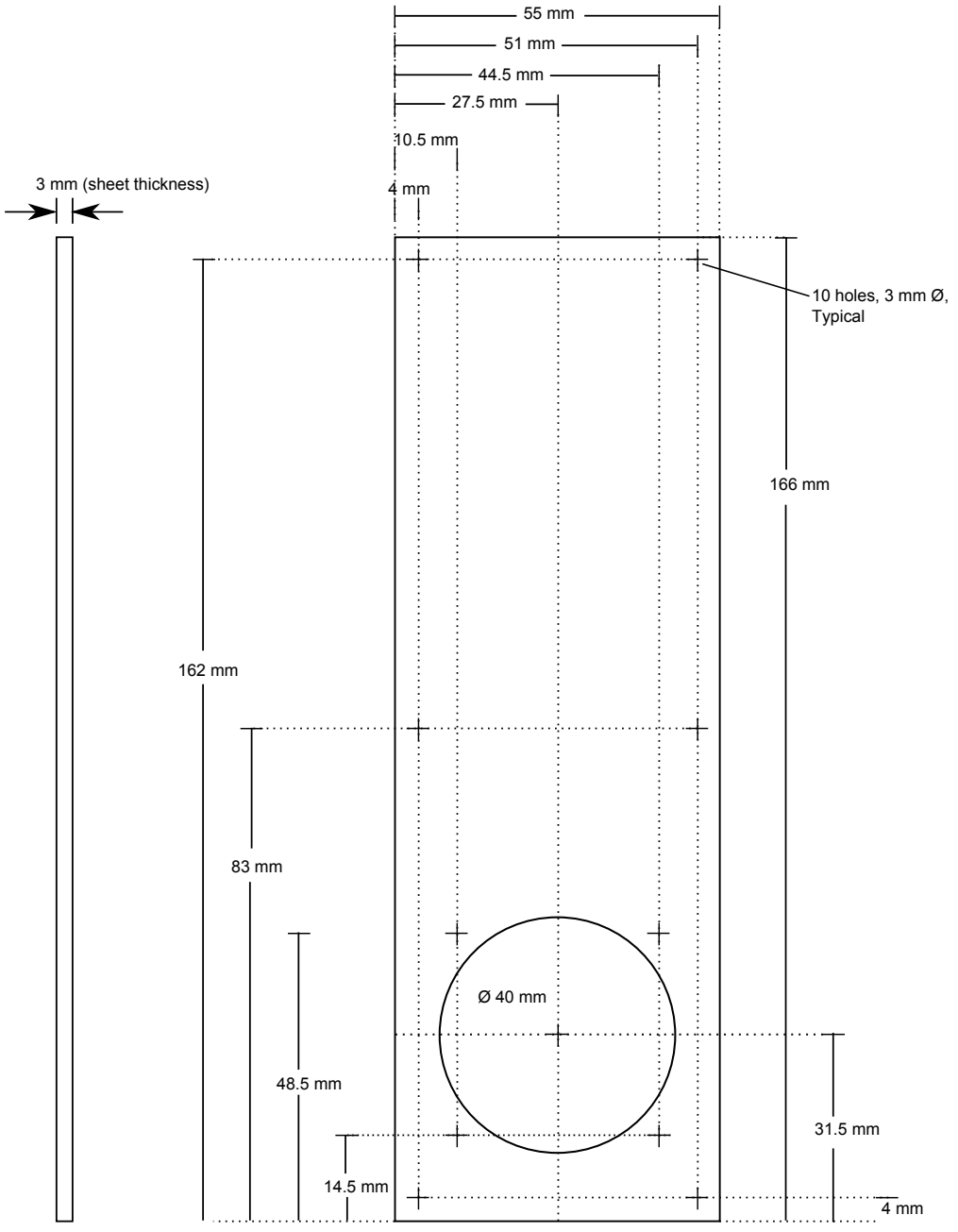


FIGURE E.3: Design of the front plate for the speaker compartments in the upper and lower layer.

E.2 Construction of the Array

E.2.1 Enclosure

As it has been discussed in the previous report, the enclosure of the HQ array was manufactured from plywood. It consists of two elements that, when combined, form an array of 15 independent drivers.

The top and bottom element consist of seven and eight cavities, respectively. The enclosure of the bottom element is depicted in Fig. E.4. Each cavity is separated from the adjacent cavities

FIGURE E.4: Enclosure of the HQ hemi-cylindrical array with damping material (upholstery cotton felt) in place.

by a wall made of two plywood boards that have been glued together, leaving a thin air gap between both boards. This design should reduce the mechanical/acoustical coupling between two adjacent cavities.

Each cavity is filled with a suitably cut piece of upholstery cotton felt to provide adequate acoustical damping inside the speaker cabinets. This particular material is a waste product of the textile industry and can be purchased for a relatively low price. A horizontal piece of plywood has been added halfway up the cavity front to ensure a solid structural integrity of the overall construction.

The cable carrying the loudspeaker signal accesses the cavity through a silicone sealed hole at the rear of the cavity.

E.2.2 Driver Units

All 15 drivers are Tang Band W1-1070SG full-range drivers (see Fig. E.5), where 'full-range' refers to the fact that the entire audio signal range is presented through these drivers. This does not mean, however, that the chosen drivers are capable of delivering a flat frequency response within the entire audible spectrum. Nevertheless, it can be assumed that the drivers present a

substantial improvement over the previous prototypes, especially with respect to its low-frequency radiation performance.

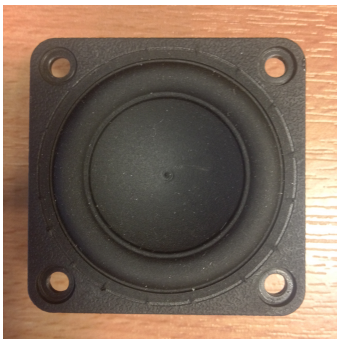


FIGURE E.5: Visible part of the Tang Band W1-1070SG driver.

E.2.3 Front Panels

The drivers are individually mounted onto an aluminium panel. To avoid acoustic leakage between these front panels and the front edges of the wooden cabinets, each panel was fitted with an additional silicone sealing along all four edges (see Fig. E.6). This ensures that, when the panel

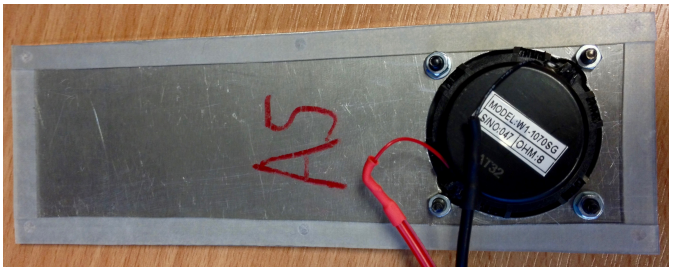


FIGURE E.6: Aluminium panel with driver in place and silicone sealing along the edges.

is mounted to the cabinets, the resulting cabinet is acoustically sealed, following the intended non-ported design.

E.2.4 The Assembled HQ Hemi-Cylindrical Array

The assembly of the array was concluded after fitting the front panels with their mounted drivers to the wooden structure of the array. The two segments were then connected to each other through designated dowels in the top and bottom structure of the upper and the lower segment, respectively. This yields an interleaved array design with 15 speakers as depicted in Figure E.7. The depicted array reaches a height of 327 mm and weighs approximately 2 kg.

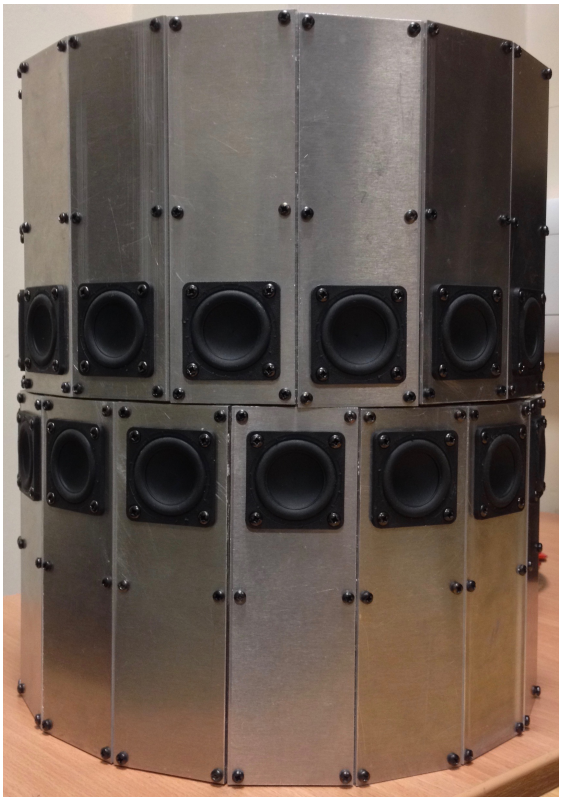


FIGURE E.7: Fully assembled HQ hemi-cylindrical array

Appendix F

Plane Wave Identification With Circular Arrays By Means Of A Finite Rate Of Innovation Approach

Falk-Martin Hoffmann, Filippo M. Fazi, Philip A. Nelson

Abstract

Many problems in the field of acoustic measurements depend on the direction of incoming wave fronts w.r.t. a measurement device or aperture. This knowledge can be useful for signal processing purposes such as noise reduction, source separation, de-aliasing and super-resolution strategies among others.

This paper presents a signal processing technique for the identification of the directions of travel for the principal plane wave components in a sound field measured with a circular microphone array. The technique is derived from a finite rate of innovation data model and the performance is evaluated by means of a simulation study for different numbers of plane waves in the sound field.

F.1 Introduction

There are many methods of decomposing a microphone array measurement of a sound field into an assessable number of parameters. Ignoring the differences between two-dimensional and three-dimensional technologies, a convenient approach to perform this analysis/decomposition is to decompose a sound field into a set of plane waves, e.g. as proposed in [Raf04, TK06, ZDG10, BB10c]. Plane waves have the advantage that they are elementary entities defined by only two

parameters: the direction of travel (DOT) and the complex amplitude. If the number of plane waves in a sound field is limited then it is also perfectly described by a small number of parameters. In that case, these parameters can be identified from the measurement made with a suitable microphone array.

One typically distinguishes between non-parametric (i.e. no underlying signal model) and parametric (e.g. based on a signal model) source detection or DOT estimation techniques. Examples of non-parametric techniques are adaptive beamforming based on maximum likelihood estimators or the MVDR beamformer [Cap69, VT02, ZFBZ08]. Prominent examples of parametric techniques are the MUSIC algorithm [Sch86, VT02], the EB-ESPRIT and the EB-DETECT algorithm [TK06], and Matching Pursuit variants such as CLEAN [Ḧ74]. Compressed sensing techniques could also be utilised to estimate parameters, yet are they typically used improve the performance of sensor arrays directly [EJvS09, WEMJ11]. Some of these techniques require a priori knowledge (e.g. the number of sources) or deploy statistical models to obtain estimates. This may potentially lead to usable, yet slightly inaccurate results. Many applications and techniques, however, require the exact knowledge of the DOT to ensure the best possible performance. Examples are source separation, de-noising and de-aliasing techniques based on beamforming for acoustical and sonar applications, as well as super-resolution techniques [HBC08, HL09, BIB13, AR14].

This work presents a novel application of a signal processing approach that theoretically allows to analytically recover the set of DOT of a finite number of principal plane waves from the measured data of a circular microphone array. Once recovered, this information can be used to further process the measured data, e.g. with simple beamforming techniques to suppress sources outside the main listening direction or even de-aliasing strategies [AR14]. This work is to be understood as an initial study based on idealised simulations and aims to introduce the proposed method. Therefore the impact of measurement noise, model mismatch and transducer imperfections is not within the scope of this work and will be addressed in future publications.

The next section introduces the sound field model, which is suitable to be combined with circular microphone arrays. The third section briefly describes how the data defining the sound field can be gained from the array observation. In the fourth section, the proposed method to identify the plane waves in the sound field based on a finite rate of innovation data model is presented. In the penultimate section, the proposed method is analysed with respect to its accuracy for different idealised sound field scenarios and the last section concludes the findings and gives an outlook on future work.

F.2 Sound Field Model

The general sound field model for this work is based on the Herglotz Wave Function (HWF) in the frequency domain [FNW12, HF14] given by

$$p(\mathbf{x},\omega) = \int_{\Omega} H(\mathbf{x},\hat{\mathbf{y}},\omega) q(\hat{\mathbf{y}}) d\Omega(\hat{\mathbf{y}}),\ \mathbf{x} \in \Lambda, \tag{F.1}$$

where Λ denotes a sphere of a given radius around the coordinate origin within which the homogeneous wave equation is satisfied at all points $\mathbf{x}$. The integration domain $\Omega = \{\hat{\mathbf{y}}$:

$\sqrt{y_1^2+y_2^2+y_3^2}=1\}$ encompasses all the points on the surface of the unit sphere. Equation (F.1) basically describes the sound field as the superposition of plane waves travelling in the directions $\hat{\mathbf{y}}$. The magnitude of each individual plane wave is determined by the *Herglotz Density* (HD) $q(\hat{\mathbf{y}})$ and its propagation characteristics are given through the *Herglotz Kernel* $H(\mathbf{x},\hat{\mathbf{y}},\omega)=e^{ik\mathbf{x}\cdot\hat{\mathbf{y}}}$.

While equation (F.1) is valid within $\Lambda\in\mathbb{R}^3$, this work considers only two-dimensional scenarios to reconcile the sound field model with the intended signal processing approach. Hence, the sound field is assumed as height-invariant, i.e. the pressure is constant along the x_3-axis, so that

$$p(x_1,x_2,x_3)=p(x_1,x_2).$$

It has been shown in [HF14] that, using the Jacobi-Anger expansion [CK98], the model of the pressure in equation (F.1) can be rewritten in polar coordinates as a function of the radius r and the polar angle ϕ

$$p(r,\phi)=2\pi\sum_{n=-\infty}^{\infty}i^n R_n(kr)\frac{e^{in\phi}}{\sqrt{2\pi}}\int_0^{2\pi}\frac{e^{-in\phi'}}{\sqrt{2\pi}}q(\phi')d\phi'. \tag{F.2}$$

Note that the angular frequency ω has been dropped from the notation for the sake of brevity. The radial function

$$R_n(kr)=\begin{cases}J_n(kr) & ,\text{FF}\\ J_n(kr)-\frac{J_n'(kr_s)}{H_n^{(1)\prime}(kr_s)}H_n^{(1)}(kr) & ,\text{CS}\end{cases} \tag{F.3}$$

allows for the modelling of free field conditions (FF) or the presence of a cylindrical scatterer at the origin (CS) with radius r_s [Wil99, TK06]. $J_n(\cdot)$ and $H_n^{(1)}(\cdot)$ denotes the Bessel-function and the Hankel-function of the first kind, respectively, where $J_n'(\cdot)$ and $H_n^{(1)\prime}(\cdot)$ denote their respective derivatives. The integration variable ϕ' in eq. (F.2) replaces the DOT $\hat{\mathbf{y}}$, so that it represents the angle between the positive x_1-axis and the vector $\hat{\mathbf{y}}$. The integration domain has become the unit circle $\Omega_c=\{\hat{\mathbf{y}}:\sqrt{x_1^2+x_2^2}=1\}$.

It is assumed here that the Herglotz density holds the basic information that defines the plane waves in the sound field, and that this needs to be recovered from a finite number of pressure observations made by a circular microphone array before the planned signal processing approach can be applied. This is typically referred to as the inverse problem, which is solved in the next section.

F.3 Recovering the Herglotz Density from Circular Array Measurements

The procedure for solving the inverse problem such as the one posed by equation (F.2) has been discussed in a number of works [Wil99, Raf05, TK06, Mey01]. For the sake of brevity, this derivation has been omitted. The analytical solution to the inverse problem to equation (F.2) is given by

$$q(\phi)=\sum_{n=-\infty}^{\infty}q_n\frac{e^{in\phi}}{\sqrt{2\pi}}, \tag{F.4}$$

with

$$q_n=\frac{1}{2\pi i^n R_n(kr_V)}\int_0^{2\pi}\frac{e^{-in\phi'}}{\sqrt{2\pi}}p(\phi')d\phi'. \tag{F.5}$$

For a circular microphone array with $L = 2N + 1$ uniformly distributed pressure sensors, the solution for the coefficients q_n in (F.5) can be approximated by

$$\tilde{q}_n = \frac{1}{i^n R_n(kr_V)L} \sum_{n=l}^{L} \frac{e^{-in\Delta\phi l}}{\sqrt{2\pi}} p(\Delta\phi l), \tag{F.6}$$

with $\Delta\phi = \frac{2\pi}{L}$. As a consequence of the discretisation, the HD can only be reconstructed up to the Nth order [Pol05b, HF14], so that

$$\tilde{q}(\phi) = \sum_{n=-N}^{N} \tilde{q}_n \frac{e^{in\phi}}{\sqrt{2\pi}}. \tag{F.7}$$

Furthermore, discretisation is also bound to lead to aliasing effects since the spatial complexity of sound fields can typically not be expressed through an order-limited HD, especially at high frequencies. However, the problem of aliasing is not within the scope of this work and considered a separate problem, but it has been described for circular arrays with pressure sensors in previous work by Poletti [Pol05b] and Alon and Rafaely [AR14]. In the following, it shall be assumed that the recovered coefficients $\tilde{q}_n$ are not corrupted by either aliasing, measurement noise or misalignment of the sensors.

The next section introduces a method of estimating the DOTs of up to N plane waves from the set of recovered coefficients $\{\tilde{q}_n : n \in [-N \dots N]\}$ on the basis of a finite rate of innovation signal model.

F.4 Finite Rate of Innovation Approach

This section presents how the parameters defining the plane waves in a sound field can be recovered based on a Finite Rate of Innovation (FRI) signal model. A comprehensive introduction to the theory behind signals with finite rate of innovation is beyond the scope of this work and the interested reader is referred to the pertinent literature [VMB02, DH09, BDV+08].

First, the new signal model for the Herglotz density is introduced. This model is based on a form for the HD specified by equation (F.7) but makes another assumption regarding the nature of the sound field. Then, a set of annihilating filters is calculated that is applied to the new signal model. These filters are specified by the DOT of the incoming plane waves. Finally, it is shown how these parameters can be extracted from the identified annihilating filters.

F.4.1 HD of a Single Plane Wave

From the definition of the HWF in equations (F.1) and (F.2), it can be seen that the HD of a single plane wave is theoretically a weighted Dirac delta shifted to the direction of travel θ and is given by

$$q(\phi) = b\delta(\phi - \theta), \tag{F.8}$$

where b is the complex amplitude of the plane wave. For a function with a period of 2π, the Dirac delta can be approximated by the Fourier Series

$$q(\phi) = \sum_{n=-\infty}^{\infty} b \frac{e^{-in\theta}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}}. \tag{F.9}$$

The above form of the HD can now be utilised in the FRI signal model, which is introduced in the following subsection.

F.4.2 New Model for the Herglotz Density

In the following, only Herglotz densities that define sound fields with up to $M = N$ plane waves are considered. Accordingly, these can be represented by the sum of the Herglotz densities of the M individual plane waves

$$q(\phi) = \sum_{m=1}^{M} b_m \sum_{n=-\infty}^{\infty} \frac{e^{in(\phi-\theta_m)}}{2\pi}. \tag{F.10}$$

A Herglotz Density of the above form has exactly $2M$ degrees of freedom, i.e. M directions of travel and M complex amplitudes. Alternatively, since the HD is periodic, one can say that it has a 'finite rate of innovation', which is an alternative way of expressing a 'finite number of degrees of freedom'.

Rearranging equation (F.10) yields

$$q(\phi) = \sum_{n=-\infty}^{\infty} \sum_{m=1}^{M} b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}} \frac{e^{in\phi}}{\sqrt{2\pi}} = \sum_{n=-\infty}^{\infty} q_n \frac{e^{in\phi}}{\sqrt{2\pi}} \tag{F.11}$$

with the Fourier coefficients

$$q_n = \sum_{m=1}^{M} b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}}. \tag{F.12}$$

The coefficients q_n can be seen as an infinite sequence. Alternatively, the coefficients q_n in (F.12) can be seen as the superposition of M infinite sequences defined by

$$q_n^{(m)} = b_m \frac{e^{-in\theta_m}}{\sqrt{2\pi}} \tag{F.13}$$

for $m \in [1 \dots M]$. Note that each of these sequences is uniquely related to *one* of the plane waves only. The following subsection describes how the individual parameters b_m and θ_m can be extracted from the q_n.

F.4.3 Definition of the Annihilating Filter

For each sequence $q_n^{(m)}$, there is a corresponding filter

$$a_n^{(m)} = \delta_n - e^{-i\theta_m} \delta_{n-1}$$

that satisfies the equation

$$a_n^{(m)} * q_n^{(m)} = 0,$$

where the asterisk represents the discrete signal convolution and δ_n denotes the Kronecker delta [Wil99, BSMM06], which is defined by

$$\delta_n = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

These filters are therefore also sometimes referred to as *annihilating filters* [VMB02]. The z-transform of the filters $a_n^{(m)}$ is given by

$$A^{(m)}(z) = 1 - e^{-i\theta_m} z^{-1},$$

which is zero for $z = e^{i\theta_m}$. Hence each annihilating filter suppresses the component $e^{in\theta_m}$ in the sequence of the Fourier coefficients q_n.

If the individual progressions $q_n^{(m)}$ could be observed directly from the recovered coefficients q_n, it would be simple to identify the corresponding annihilating filters. Unfortunately this is not the case. From equations (F.12) and (F.13) it follows however that

$$q_n = \sum_{m=1}^{M} q_n^{(m)}, \forall n \in \mathbb{Z}.$$

It can thus be assumed that one non-trivial solution for a_n that satisfies

$$q_n * a_n = 0 \tag{F.14}$$

is the combination of all annihilating filters $a_n^{(m)}$ to one comprehensive annihilating filter a_n of the form

$$a_n = a_n^{(1)} * a_n^{(2)} * \ldots * a_n^{(M)} = \delta_n + \sum_{m=1}^{M} \alpha_m \delta_{m-n} \tag{F.15}$$

with its z-transform given by

$$A(z) = \prod_{m=1}^{M} A^{(m)}(z) = \prod_{m=1}^{M} (1 - u_m z^{-1}),\ u_m = e^{-i\theta_m}. \tag{F.16}$$

The z-transform $A(z)$ can also be written using the coefficients α_m defined in equation (F.15):

$$A(z) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \ldots + \alpha_M z^{-M}. \tag{F.17}$$

From equations (F.16) and (F.17), it can be seen that the coefficients u_m are in fact the roots of the characteristic polynomial

$$\eta(z) = z^M + \alpha_1 z^{M-1} + \alpha_2 z^{M-2} + \ldots + \alpha_M. \tag{F.18}$$

In conclusion, once the annihilating filter a_n has been identified and the roots of its characteristic polynomial have been calculated, then the directions of travel $\{\theta_m : m \in [1 \ldots M]\}$ of the M

plane waves are successfully recovered. The next step is therefore to find the annihilating filter a_n.

F.4.4 Calculation of the Annihilating Filter

The convolution of discrete-time signals can be realised in a convolution matrix. With a limited length of the filter a_n, the linear equation system (LEQS) equivalent to (F.14) is given by

$$\begin{bmatrix} q_1 & q_0 & q_{-1} & \cdots & q_{-(M-1)} \\ q_2 & q_1 & q_0 & \cdots & q_{-(M-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_M & q_{M-1} & q_{M-2} & \cdots & q_0 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

It follows from equation (F.15) that $a_0 = 1$. This can be exploited to avoid the trivial solution for the filter a_n. Rearranging the above LEQS yields

$$\begin{bmatrix} q_0 & \cdots & q_{-(M-1)} \\ \vdots & \ddots & \vdots \\ q_{M-1} & \cdots & q_0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_M \end{bmatrix} = - \begin{bmatrix} q_1 \\ \vdots \\ q_M \end{bmatrix}, \tag{F.19}$$

which is widely known in the field of auto-regressive filtering as a Yule-Walker equation system. This can ideally be solved by plain matrix inversion, since it can be expected that the matrix is full rank. However, this may not be the case in practice if the M plane waves are not distinct [VMB02]. Furthermore, equation (F.19) implicitly gives a condition for the set of modal coefficients $\{q_n \,:\, n \in [-N_{\min} \dots N_{\min}]\}$ that is needed to perform this type of analysis. From the system matrix and the right hand side of equation (F.19) it follows that

$$2N_{\min} + 1 = 2M \iff N_{\min} = M - \frac{1}{2}. \tag{F.20}$$

The last step is then to find the M distinct roots u_m of the characteristic polynomial $\eta(z)$ in (F.18). These can be obtained through different algorithms. For this work, the built-in function *roots* of MATLAB has been utilised.

Once the roots have been found, the estimated directions of travel $\tilde{\theta}_m$ can be calculated by the formula

$$\tilde{\theta}_m = i \ln(u_m), \forall m \in [1 \dots M]$$

where $\ln(x)$ denotes the principal value of the natural logarithm.

F.4.5 Recovering the Complex Amplitudes of the Individual Plane Waves

With the DOT identified, equation (F.12) can be used to generate a linear equation system with the complex amplitudes $\{b_m \,:\, m \in [1 \dots M]\}$ as the unknown parameters and the Fourier

coefficients q_n form the vector of constants [VMB02].

$$\begin{bmatrix} q_0 \\ \vdots \\ q_{M-1} \end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix} 1 & \dots & 1 \\ e^{-i\theta_1} & \dots & e^{-i\theta_M} \\ \vdots & \ddots & \vdots \\ e^{-i(M-1)\theta_1} & \dots & e^{-i(M-1)\theta_M} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} \tag{F.21}$$

This so called Vandermonde equation system [VMB02] can be solved by matrix inversion to identify the complex amplitudes b_m.

F.5 Performance Evaluation

This section reports the results of numerical simulations to evaluate the performance of the proposed technique by means of one specific example and the results of a small simulation study.

F.5.1 An Example

Figure F.1 gives an indication of the accuracy of the estimated parameters $\{\theta_m \ : \ m \in [1 \dots M]\}$ when using the FRI based method with an artificially generated set of HD coefficients $\{q_n : |n| \leq 7\}$, which are given by equation (F.11) for a set of $M = 4$ randomly generated DOT θ_m and complex amplitudes b_m. The top graph shows the real- and imaginary part of the original HD $q(\phi)$. The

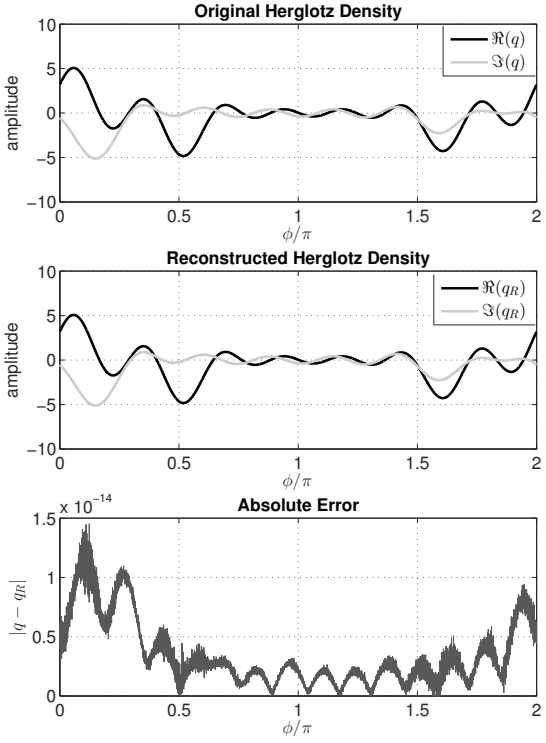


FIGURE F.1: Example of the results achieved with the proposed method for $M = 4$ and $N = 7$, (Top) real- and imaginary part of the original HD, (Middle) real- and imaginary part of the reconstructed HD, (Bottom) absolute reconstruction error as an indicator for the parameter estimation performance.

graph in the middle shows the real- and imaginary part of the HD $q_R(\phi)$ that was reconstructed on the basis of the identified parameters $\tilde{\theta}_m$ and $\tilde{b}_m$. The bottom graph shows the absolute error between the original HD and the reconstructed HD $|q - q_R|$. It can be seen that the FRI method approximates the set of parameters $\{\theta_m : m \in [1 \dots M]\}$ with very high accuracy. The nearly negligible error is accredited to the limited numerical accuracy during the matrix inversion.

F.5.2 Simulation Study

To assess both the average performance and the competitiveness of the proposed method, it was tested in a simulation study against the matching pursuit variant CLEAN [Ḧ74]. The latter is an iterative algorithm that can be described as follows:

1. Fit the signature of a single plane wave (see eq. (F.9)) in the Herglotz Density domain to the recovered data $\tilde{q}(\phi)$ so that their global maxima are aligned and even. The parameter that best aligns the two maxima is the DOT estimate $\tilde{\theta}_m$ for the current iteration's strongest plane wave. [1]
2. Subtract the identified signature from $\tilde{q}(\phi)$.
3. Repeated the first two steps until the designated number of M plane waves have been identified.

In this simulation study, 5001 uniformly spaced samples were used for the required reconstruction in the HD domain.

For the study, both methods were evaluated w.r.t. their average Parameter Identification Error (PIE) for the estimation of the DOTs. For an estimated parameter $\tilde{\theta}_m$, the PIE in percent is given by

$$\text{PIE} = \Big(|\theta_m - \tilde{\theta}_m| \ \mathbf{mod}\ \pi \Big)\, 100\,\%.$$

The above measure was calculated for all identified DOTs in 5000 different Herglotz Densities and the results were averaged afterwards. This was repeated for different numbers of plane waves M in the sound field. Every HD was of the form given in (F.11) with randomly chosen θ_m and b_m, respectively. That corresponds to data perfectly acquired through a circular array with $L = 15$ microphones (i.e. no spatial aliasing), allowing to reconstruct the sound field up to the order $N = 7$.

Table F.1 shows the study's results for $M = 4, 5, 6, 7$, where $M = 7$ is the largest number of plane waves that can be identified with the proposed method from the simulated sound field data (compare eq. (F.20)). It can be seen that, even for $M = 7$, the proposed method outperforms the CLEAN method significantly, providing sufficient accuracy for most applications.

Note that this simulation study has been conducted based on simulated, measurement noise-free and thus ideal conditions, where the number of plane waves identified always matched the number of plane waves in the field.

[1] It can be shown that $\tilde{\theta}_m$ also maximises the cross-correlation between the single plane wave signature and the current HD data $\tilde{q}(\phi)$.

Number of Plane Waves	Average PIE with FRI	Average PIE with CLEAN
$M = 4$	$4.0574 \cdot 10^{-9}$ %	6.3412 %
$M = 5$	$9.8317 \cdot 10^{-8}$ %	8.4108 %
$M = 6$	$5.1642 \cdot 10^{-6}$ %	9.7623%
$M = 7$	$6.3900 \cdot 10^{-3}$ %	10.4769%

TABLE F.1: PIE in percent for $M = 4, 5, 6, 7$, calculated for each value M from 5000 different Herglotz Densities $q(\phi)$, each specified by a set of randomly chosen directions of travel θ_m and complex amplitudes b_m.

F.6 Conclusion

A novel method to estimate the direction of travel of a given number of plane waves from the measurement of a circular microphone array has been presented. The method is based on FRI signal theory and has been evaluated by means of an initial simulation study w.r.t. the achieved average parameter identification error. It has been shown that the latter does not exceed 0.01 % for up to $M = 7$ plane waves with the proposed method. This assumes ideal measurement data acquired with 15 microphones in the absence of noise. The remaining inaccuracy can most presumably be blamed on the limited numerical precision of the computer system. The performance comparison of the proposed method and the CLEAN algorithm indicates that, for the conditions simulated, the FRI method clearly surpasses the CLEAN algorithm in terms of accuracy.

Future work is going to investigate the performance of the proposed method when applied to non-ideal data, covering the evaluation of its robustness against measurement noise, non-ideal plane waves (e.g. point sources in the near field of the measurement aperture) and transducer/aperture imperfections. Especially potential ill-conditioning of the matrices involved in the FRI method may have a crucial impact on its performance. Furthermore, the problem of when the number of plane waves to be identified does not match the number of plane waves in the sound field, i.e. it is greater or smaller, will be considered, as well as a comparison of the proposed method to further alternatives (e.g. MUSIC, EB-ESPRIT, etc.).

F.7 Acknowledgements

This work has been partially funded by the Royal Academy of Engineering and by the Engineering and Physical Sciences Research Council.

References

[AB79] Jont B. Allen and David A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

[ABC+70] J.S. Asvestas, J.J. Bowman, P.L. Christiansen, O. Einarsson, R.E. Kleinman, D.L. Sengupta, T.B.A. Senior, F.B. Sleator, P.L.E. Uslenghi, and N.R. Zitron. *Electromagnetic and Acoustic Scattering by Simple Shapes*. North-Holland Pub. Co., 1970.

[AFNR13] Paolo Annibale, Jason Filos, Patrick A. Naylor, and Rudolf Rabenstein. TDOA-based speed of sound estimation for air temperature and room geometry inference. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(2):234–246, feb 2013.

[Ahr12] Jens Ahrens. *Analytic Methods of Sound Field Synthesis*. T-Labs Series in Telecommunication Services. Springer, 2012.

[AR12] David Lou Alon and Boaz Rafaely. Spherical microphone array with optimal aliasing cancellation. In *Proceedings of the 27th Convention of Electrical and Electronics Engineering in Israel (IEEE)*, pages 1–5, November 2012.

[AR14] David Lou Alon and Boaz Rafaely. Spatial aliasing-cancellation for circular microphone arrays. In *4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014*, pages 137–141, May 2014.

[AR16] David Lou Alon and Boaz Rafaely. Beamforming with optimal aliasing cancellation in spherical microphone arrays. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 24(1):196–210, January 2016.

[AS08] J. Ahrens and S. Spors. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica united with Acustica*, 94(6):988–999, December 2008.

[AS12a] Jens Ahrens and Sascha Spors. A modal analysis of spatial discretization of spherical loudspeaker distributions used for sound field synthesis. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(9):2564–2574, November 2012.

[AS12b] Jens Ahrens and Sascha Spors. Wave field synthesis of a sound field described by spherical harmonics expansion coefficients. *The Journal of the Acoustical Society of America*, 131(3):2190–2199, March 2012.

[ASR15] David L. Alon, Jonathan Sheaffer, and Boaz Rafaely. Robust plane-wave decomposition of spherical microphone array recordings for binaural sound reproduction. *The Journal of the Acoustical Society of America*, 138(3):1925–1926, sep 2015.

[AW02] T.D. Abhayapala and D.B. Ward. Theory and design of high order sound field microphones using spherical microphone array. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 1949–1952, May 2002.

[AW05] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists*. New York, Academic Press, 6th edition, 2005.

[AWS10] Jens Ahrens, Hagen Wierstorf, and Sascha Spors. Comparison of higher order ambisonics and wave field synthesis with respect to spatial discretization artifacts in time domain. In *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, Oct 2010.

[BB05] Magella Bédard and Alain Berry. Development of a directivity controlled piezoelectric transducer for sound reproduction. *The Journal of the Acoustical Society of America*, 117(4):2455–2455, apr 2005.

[BB10a] Natasha Barrett and Svein Berge. High angular resolution planewave expansion (HARPEX). In *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, pages 4–9, 2010.

[BB10b] Svein Berge and Natasha Barrett. A new method for B-format to binaural transcoding. In *40th AES International Conference*, pages 8–10, Tokyo, Japan, 2010.

[BB10c] Sven Berge and Natasha Barrett. High angular resolution planewave expansion. In *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, 2010.

[BDV+08] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot. Sparse sampling of signal innovations. *IEEE Signal Processing Magazine*, 25(2):31–40, March 2008.

[Ber88] A.J. Berkhout. A holographic approach to acoustic control. *Journal of the Audio Engineering Society*, 36:977–995, December 1988.

[Ber12] Benjamin Bernschütz. Bandwidth extension for microphone arrays. In *133rd Convention of the Audio Engineering Society*, Oct 2012.

[BIB13] Mingsian R. Bai, Jeong-Guon Ih, and Jacob Benesty. *Acoustic Array Systems: Theory, Implementation, and Application*. Wiley-IEEE Press, 1st edition, 2013.

[Bla97] Jens Blauert. *Spatial Hearing : The Psychophysics of Human Sound Localization*. Cambridge, Mass. MIT Press, 1997.

[Blu31] Alan Dower Blumlein. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems, 1931.

[Blu35] Alan Dower Blumlein. Improvements in and relating to Electrical Sound Transmission Systems, 1935.

[BP14] Terence Betlehem and Mark A. Poletti. Two dimensional sound field reproduction using higher order sources to exploit room reflections. *Journal of the Acoustical Society of America*, 135(4):1820–1833, 2014.

[BSMM06] I.N. Bronstein, K.A. Semendjajew, G. Musiol, and H. Mühlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 2006.

[BT90] Michael J. Buckingham and Alexandra Tolstoy. An analytical solution for benchmark problem 1: The "ideal" wedge. *Journal of the Acoustical Society of America*, 87(4):1511–1513, 1990.

[Buc84] M.J. Buckingham. *Hybrid Formulation of Wave Propagation and Scattering*, chapter Acoustic Propagation in a Wedge-Shaped Ocean with Perfectly Reflecting Boundaries, pages 77–99. Springer Netherlands, Dordrecht, 1984.

[Buc89] Michael J. Buckingham. Theory of acoustic radiation in corners with homogeneous and mixed perfectly reflecting boundaries. *Journal of the Acoustical Society of America*, 86(6):2273–2291, 1989.

[BV95] Jeffery S. Bamford and John Vanderkooy. Ambisonic sound for us. In *99th Convention of the Audio Engineering Society*, Oct 1995.

[BVV93] A.J. Berkhout, D. De Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustical Society of America*, 93:2764–2778, May 1993.

[BX08] Jens Blauert and Ning Xiang. *Acoustics for Engineers*. Springer Berlin Heidelberg, 2008.

[BZPA15] Terence Betlehem, Wen Zhang, Mark A. Poletti, and Thushara D. Abhayapala. Personal sound zones: Delivering interface-free audio to multiple listeners. *IEEE Signal Processing Magazine*, 32(2):81–91, mar 2015.

[Can09] James V. Candy. *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2009.

[Cap69] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, August 1969.

[CAZ15] Hanchi Chen, Thushara D. Abhayapala, and Wen Zhang. Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis. *The Journal of the Acoustical Society of America*, 138(5):3081–3092, nov 2015.

[CDS98] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM JOURNAL ON SCIENTIFIC COMPUTING*, 20:33–61, 1998.

[CG74] Peter Graham Craven and Michael Anthony Gerzon. Coincident microphone simulation covering three dimensional space and yielding various directional outputs, 1974.

[Cha14] Peter John Chapman. Ambient atmospheric conditions and their influence on acoustic measurements. In *136th Convention of the Audio Engineering Society*, April 2014.

[CJO+13] Philip Coleman, Philip J. Jackson, Marek Olik, Martin Olsen, Martin Moller, and Jan Abildgaard Pedersen. The influence of regularization on anechoic performance and robustness of sound zone methods. In *Proceedings for the International Congress on Acoustics*, volume 19, page 055055. ASA, 2013.

[CK98] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, 1998.

[CK02] Joung-Woo Choi and Yang-Hann Kim. Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, 111(4):1695–1700, apr 2002.

[CLT09] Peter Graham Craven, Malcolm J. Law, and Chris Travis. Microphone arrays using tangential velocity sensors. In *Ambisonics Symposium*, June 2009.

[CS82] Andrés Cruz and Javier Sesma. Zeros of the Hankel function of real order and of its derivative. *Mathematics of Computation*, 39(160):639–645, October 1982.

[CSY07] Huawei Chen, Wee Ser, and Zhu Liang Yu. Optimal design of nearfield wideband beamformers robust against errors in microphone array characteristics. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(9):1950–1959, sep 2007.

[CW08] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, March 2008.

[Dan03] Jerome Daniel. Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*, May 2003.

[DBA09] Jacek Dmochowski, Jacob Benesty, and Sofiène Affès. On spatial aliasing in microphone arrays. *IEEE Transactions on Signal Processing*, 57(4):1383–1395, April 2009.

[DH09] P.L. Dragotti and F. Homann. Sampling signals with finite rate of innovation in the presence of noise. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2941–2944, April 2009.

[DMN03] Jérôme Daniel, Sébastien Moreau, and Rozenn Nicol. Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. In *114th Convention of the Audio Engineering Society*, 3 2003.

[Don06] David L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, April 2006.

[EDG03] Gary W. Elko, Eric Diethorn, and Tomas Gänsler. Room impulse response variation due to thermal fluctuation and its impact on acoustic echo cancellation. In *In International Workshop on Acoustic Echo and Noise Control*, 2003.

[EF07] N. Epain and E. Friot. Active control of sound inside a sphere via control of the acoustic pressure at the boundary surface. *Journal of Sound and Vibration*, 299(3):587 – 604, 2007.

[EJvS09] Nicolas Epain, Craig Jin, and André van Schaik. The application of compressive sampling to the analysis and synthesis of spatial sound fields. In *127th Convention of the Audio Engineering Society*, 2009.

[EK12] Yonina C. Eldar and Gitta Kutyniok, editors. *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[Ell01] Steven J. Elliott. *Signal Processing for Active Control*. Academic Press, 2001.

[Far00] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *108th Convention of the Audio Engineering Society*, Feb 2000.

[Far07] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *122nd Convention of the Audio Engineering Society, Vienna, Austria*, 2007.

[Faz10] Filippo M. Fazi. *Sound Field Reproduction*. PhD thesis, University of Southampton, Faculty of Engineering, Science and Mathematics, Institute of Sound and Vibration Research, 2010.

[FC16] Angelo Farina and Lorenzo Chiesi. Measuring spatial mimo impulse responses in rooms employing spherical transducer arrays. In *52nd AES International Conference on Sound Field Control - Engineering and Perception*, 2016.

[FLS63] R.P. Feynman, R.B. Leighton, and M. Sands. *The Feynman Lectures on Physics*, volume 1. Addison Wesley Longman, 1963.

[FM03] Leopold B. Felsen and Nathan Marcuvitz. *Radiation and Scattering of Waves*. John Wiley & Sons, 2003.

[FN07] Filippo M. Fazi and Philip A. Nelson. The ill-conditioning problem in sound field reconstruction. In *123rd Convention of the Audio Engineering Society*, Oct 2007.

[FN10] Filippo Fazi and Philip Nelson. Nonuniqueness of the solution of the sound field reproduction problem. In *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, May 2010.

[FN12] Filippo M. Fazi and Philip A. Nelson. Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control. *Acta Acustica united with Acustica*, 98(1):1–14, January 2012.

[FNW12] Filippo M. Fazi, M. Noisternig, and O. Warusfel. Representation of sound fields for audio recording and reproduction. In *Acoustics 2012: 11ème Congrès Français d'Acoustique. Annual meeting of the Institute of Acoustics, Nantes, FR, 23 - 27 Apr 2012*, 2012.

[FSO+14] Filippo Maria Fazi, Mincheol Shin, Ferdinando Olivieri, Simone Fontana, and Yue Lang. Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays. In *137th Convention of the Audio Engineering Society*. Audio Engineering Society, 2014.

[FSOF15] Filippo Maria Fazi, Mincheol Shin, Ferdinando Olivieri, and Simone Fontana. Low frequency performance of circular loudspeaker arrays. In *138th Convention of the Audio Engineering Society*. Audio Engineering Society, 2015.

[Ger73] Michael A. Gerzon. Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, 21(1):2–10, 1973.

[Ger75] Michael A. Gerzon. The design of precisely coincident microphone arrays for stereo and surround sound. *Proc. 50th AES Convention*, 1(1):50, 1975.

[H$\ddot{7}$4] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement*, 15:417, 1974.

[Han98] Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. Society for Industrial and Applied Mathematics, 1998.

[Han00] Per Christian Hansen. The L-curve and its use in the numerical treatment of inverse problems. In *in Computational Inverse Problems in Electrocardiology, ed. P. Johnston, Advances in Computational Bioengineering*, pages 119–142. WIT Press, 2000.

[Hay02] Simon O. Haykin. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002.

[HBC08] Yiteng Huang, Jacob Benesty, and Jingdong Chen. *Microphone Array Signal Processing*. Springer-Verlag Berlin Heidelberg, 2008.

[HBR95] G. Hayward, R.A. Banks, and L.B. Russell. A model for low noise design of ultrasonic transducers. In *Ultrasonics Symposium, 1995. Proceedings., 1995 IEEE*, volume 2, pages 971–974, November 1995.

[HF14] F.-M. Hoffmann and F.M. Fazi. Circular microphone array with tangential pressure gradient sensors. In *4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014*, pages 97–101, May 2014.

[HF15] F.-M. Hoffmann and F.M. Fazi. Theoretical study of acoustic circular arrays with tangential pressure gradient sensors. *IEEE/ACM Transactions on Audio, Speech, and Language Processing,*, 23(11):1762–1774, November 2015.

[HFF16] Falk-Martin Hoffmann, Filippo Maria Fazi, and Simone Fontana. Sound field control with hemi-cylindrical loudspeaker arrays. In *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*, 2016.

[HFN16] Falk-Martin Hoffmann, Filippo Maria Fazi, and Philip Nelson. Plane wave identification with circular arrays by means of a Finite Rate of Innovation Approach. In *140th Convention of the Audio Engineering Society*, May 2016. Student Technical Papers Award.

[HL09] Simon O. Haykin and K.J. Liu. *Handbook on Array processing and Sensor Networks*. John Wiley & Sons, 2009.

[HSdVB03] Edo Hulsebos, Thomas Schuurmans, Diemer de Vries, and Rinus Boone. Circular microphone array for discrete multichannel audio recording. In *Audio Engineering Society Convention 114*, Mar 2003.

[Idi01] Jérôme Idier, editor. *Bayesian Approach to Inverse Problems*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2001.

[JHN17] Daniel P. Jarrett, Emanuël A.P. Habets, and Patrick A. Naylor. *Theory and Applications of Spherical Microphone Array Processing*. Springer International Publishing, 2017.

[JJ07] Finn Jacobsen and Virginie Jaud. Statistically optimized near field acoustic holography using an array of pressure-velocity probes. *The Journal of the Acoustical Society of America*, 121(3):1550–1558, mar 2007.

[JKPS11] F.B. Jensen, W.A. Kuperman, M.B. Porter, and H. Schmidt. *Computational Ocean Acoustics*. Modern Acoustics and Signal Processing. Springer, 2011.

[KA08] J.E. Kirkebo and A. Austeng. Sparse cylindrical sonar arrays. *IEEE Journal of Oceanic Engineering*, 33(2):224–231, April 2008.

[KAH04] J.E. Kirkebo, A. Austeng, and S. Holm. Layout-optimized cylindrical sonar arrays. In *OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04*, volume 2, pages 598–602, Nov 2004.

[KB02] Mark Kahrs and Karlheinz Brandenburg, editors. *Applications of Digital Signal Processing to Audio and Acoustics*. Kluwer Academic Publishers, 2002.

[KB16] Wolfgang Klippel and Christian Bellmann. Holographic nearfield measurement of loudspeaker directivity. In *141st Convention of the Audio Engineering Society*, Sep 2016.

[KFCS00] L.E. Kinsler, A.R. Freya, A.B. Coppens, and J.V. Sanders. *Fundamentals of Acoustics*. Wiley, 4th edition edition, 2000.

[KFV10] Mihailo Kolundzija, Christof Faller, and Martin Vetterli. Baffled circular loudspeaker array with broadband high directivity. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010, 14-19 March 2010, Sheraton Dallas Hotel, Dallas, Texas, USA*, pages 73–76, 2010.

[KFV11] Mihailo Kolundzija, Christof Faller, and Martin Vetterli. Design of a compact cylindrical loudspeaker array for spatial sound reproduction. In *130th Convention of the Audio Engineering Society*. Audio Engineering Society, 2011.

[Kli17] Wolfgang Klippel. Sound quality of audio systems. Technical report, Dresden University of Technology, March 2017.

[Kre78] Erwin Kreyszig. *Introductory Functional Analysis with Application*. Wiley, 1978.

[KRWB10] Maor Kleider, Boaz Rafaely, Barak Weiss, and Eitan Bachmat. Golden-ratio sampling for scanning circular microphone arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, 18:2091–2098, 2010.

[KS07] A. Karbasi and A. Sugiyama. A new doa estimation method using a circular microphone array. In *2007 15th European Signal Processing Conference*, pages 778–782, Sept 2007.

[KSF89] Gary H. Koopmann, Limin Song, and John B. Fahnline. A method for computing acoustic fields based on the principle of wave superposition. *The Journal of the Acoustical Society of America*, 86(6):2433–2438, dec 1989.

[KV96] H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13(4):67–94, July 1996.

[LD07] Yhiyun Li and Ramani Duraiswami. Flexible and optimal design of spherical microphone arrays for beamforming. *IEEE Transactions on Speech and Audio Processing*, 15(2):702–714, 2007.

[LDGD04] Zhiyun Li, Ramani Duraiswami, Elena Grassi, and Larry S. Davis. Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1520–6149, 2004.

[lRD47] Jean le Rond D'Alembert. Recherches sur la courbe que forme une corde tenduë mise en vibration. *Histoire de l'académie royale des sciences et belles lettres de Berlin*, 3:214–219, 1747.

[LZ15] WenYu Luo and RenHe Zhang. Exact solution of three-dimensional acoustic field in a wedge with perfectly reflecting boundaries. *Science China Physics, Mechanics & Astronomy*, 58(9):1–10, 2015.

[Mal99] Dave G. Malham. Higher order Ambisonic systems for the spatialisation of sound. In *ICMC Proceedings*, 1999.

[Mal08] Stéphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition, 2008.

[MCD14] R. Mignot, Gilles Chardon, and Laurent Daudet. Low frequency interpolation of room impulse responses using compressed sensing. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 22:205–216, 2014.

[MDB06] Sébastien Moreau, Jérôme Daniel, and Stéphanie Bertet. 3D sound field recording with higher order Ambisonics – objective measurements and validation of a 4th order spherical microphone. In *Preprints of the 120th AES Convention*, 2006.

[ME02] Jens Meyer and Gary W. Elko. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1781–1784. IEEE, 2002.

[ME08] Jens Meyer and Gary W. Elko. Spherical harmonic modal beamforming for an augmented circular microphone array. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5280–5283. IEEE, 2008.

[Mey01] Jens Meyer. Beamforming for a circular microphone array mounted on spherically shaped objects. *Journal of the Acoustical Society of America*, 109(1):185–193, January 2001.

[MOA+10] Martin Møller, Martin Olsen, Finn T. Agerkvist, Jakob Dyreby, and Gert Kudahl Munch. Circular loudspeaker arrays with controllable directivity. In *Audio Engineering Society Convention 128*, May 2010.

[MTDG+09] Frank Melchior, Oliver Thiergart, Giovanni Del Galdo, Diemer de Vries, and Sandra Brix. Dual radius spherical cardioid microphone arrays for binaural auralization. In *Audio Engineering Society Convention 127*, Oct 2009.

[MWL85] J.D. Maynard, E.G. Williams, and Y. Lee. Nearfield acoustic holography: I. theory of generalized holography and the development of nah. *Journal of the Acoustical Society of America*, 78(4):1395–1413, 1985.

[MZ93] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, December 1993.

[Nyq28] H. Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32:110–113, July 1928.

[Oak97] C.G. Oakley. Calculation of ultrasonic transducer signal-to-noise ratios using the klm model. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 44(5):1018–1026, September 1997.

[OFC+13] Marek Olik, Jon Francombe, Philip Coleman, Philip J. B. Jackson, Martin Olsen, Martin Møller, Russell Mason, and Søren Bech. A comparative performance study of sound zoning methods in a reflective environment. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, Sep 2013.

[Off70] Franklin F. Offner. 1/f noise in semiconductors. *Journal of Applied Physics*, 41(12):5033–5034, nov 1970.

[OFNF16] Ferdinando Olivieri, Filippo Maria Fazi, Philip Nelson, and Simone Fontana. Comparison of strategies for accurate reproduction of a target signal with compact arrays of loudspeakers for the generation of zones of private sound and silence. *Journal of the Audio Engineering Society*, 64(11):905–917, dec 2016.

[OFSN15] Ferdinando Olivieri, Filippo Maria Fazi, Mincheol Shin, and Philip Nelson. Pressure-matching beamforming method for loudspeaker arrays with frequency dependent selection of control points. In *138th Convention of the Audio Engineering Society*, May 2015.

[Ols57] Harry F. Olson. *Elements of Acoustical Engineering.* D. Van Nostrand Company, Inc., 1957.

[OSF+13] Ferdinando Olivieri, Mincheol Shin, Filippo M Fazi, Philip A Nelson, and Peter Otto. Loudspeaker array processing for multi-zone audio reproduction based on analytical and measured electroacoustical transfer functions. In *Audio Engineering Society*

Conference: 52nd International Conference: Sound Field Control-Engineering and Perception. Audio Engineering Society, 2013.

[PAS12] M. A. Poletti, T. D. Abhayapala, and P. Samarasinghe. Interior and exterior sound field control using two dimensional higher-order variable-directivity sources. *The Journal of the Acoustical Society of America*, 131(5):3814–3823, may 2012.

[PB13] Mark Poletti and Terence Betlehem. Design of a prototype variable directivity loudspeaker for improved surround sound reproduction in rooms. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, Sep 2013.

[PBA15] Mark A. Poletti, Terence Betlehem, and Thushara D. Abhayapala. Higher-order loudspeakers and active compensation for improved 2d sound field reproduction in rooms. *Journal of the Audio Engineering Society*, 63(1/2):31–45, 2015.

[Pla10] Michel Plancherel. Contribution à létude de la représentation d'une fonction arbitraire par des intégrales définies. *Rendiconti del Circolo Matematico di Palermo*, 30(1):289–335, 1910.

[Pol00] Mark A. Poletti. A unified theory of horizontal holographic sound systems. *Journal of the Audio Engineering Society*, 48(12):1155–1182, 2000.

[Pol05a] M.A. Poletti. Three-dimensional surround sound systems based on spherical harmonics. *Journal of the Audio Engineering Society*, 53(11):1004–1025, November 2005.

[Pol05b] Mark A. Poletti. Effect of noise and transducer variability on the performance of circular microphone arrays. *Journal of the Audio Engineering Society*, 53(5):371–384, 2005.

[Pol07] Mark Poletti. Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts. *J. Audio Eng. Soc*, 55(7/8):598–610, 2007.

[Pol08] Mark Poletti. An investigation of 2-d multizone surround sound systems. In *125th Convention of the Audio Engineering Society*, Oct 2008.

[PP14] Hannes Pomberger and Florian Pausch. Design and evaluation of a spherical segment array with double cone. *Acta Acustica united with Acustica*, 100(5):921–927, September 2014.

[PPGM12] D. Pavlidi, M. Puigt, A. Griffin, and A. Mouchtaris. Real-time multiple sound source localization using a circular microphone array based on single-source confidence measures. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2625–2628, March 2012.

[PR05] Munhum Park and Boaz Rafaely. Sound-field analysis by plane-wave decomposition using spherical microphone array. *Journal of the Acoustical Society of America*, 118:3094–3103, November 2005.

[PZ13] H. Pomberger and F. Zotter. Modal sound field decomposition applicable for a limited range of directions. In A. Peretti and J. Scheuren, editors, *Fortschritte der Akustik, AIA-DAGA2013*, Berlin (Deutschland), 03 2013. procedure: without peer reviewing.

[Raf04] Boaz Rafaely. Plane-wave decomposition of the sound field on a sphere by spherical convolution. *Journal of the Acoustical Society of America*, 4:2149–2157, October 2004.

[Raf05] Boaz Rafaely. Analysis and design of spherical microphone arrays. *IEEE Transactions on Speech and Audio Processing*, 13(1):135–143, January 2005.

[Raf09] Boaz Rafaely. Spherical loudspeaker array for local active control of sound. *The Journal of the Acoustical Society of America*, 125(5):3006, 2009.

[RK89] R. Roy and T. Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, July 1989.

[RK11] B. Rafaely and D. Khaykin. Optimal model-based beamforming and independent steering for spherical loudspeaker arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(7):2234–2238, Sept 2011.

[RPA+10] Boaz Rafaely, Yotam Peled, Morag Agmon, Dima Khaykin, and Etan Fisher. *Spherical Microphone Array Beamforming*, pages 281–305. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[RWB07] Boaz Rafaely, Barak Weiss, and Eitan Bachmat. Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, 55:1003–1010, 2007.

[SAP14] P.N. Samarasinghe, T.D. Abhayapala, and M.A. Poletti. Room reflections assisted spatial sound field reproduction. In *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO), 2014*, pages 1352–1356, Sept 2014.

[Sch18] W. Schottky. Über spontane stromschwankungen in verschiedenen elektrizitätsleitern. *Annalen der Physik*, 362(23):541–567, 1918.

[Sch26] W. Schottky. Small-shot effect and flicker effect. *Phys. Rev.*, 28:74–103, July 1926.

[Sch86] R.O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation,*, 34(3):276–280, March 1986.

[SEA02] G.-B. Stan, J.J. Embrechts, and D. Archambeau. Comparison of different impulse response measurement techniques. *Journal of the Audio Engineering Society*, 50:249–262, 2002.

[SF74] G. M. Sessler and J. C. French. Noise due to brownian motion of the diaphragm, of electret microphones. *The Journal of the Acoustical Society of America*, 55(2):444–444, feb 1974.

[SFNH14] Mincheol Shin, Filippo M. Fazi, Philip A. Nelson, and Fabio C. Hirono. Controlled sound field with a dual layer loudspeaker array. *Journal of Sound and Vibration*, 333(16):3794–3817, aug 2014.

[SM16] N. Stefanakis and A. Mouchtaris. Direction of arrival estimation in front of a reflective plane using a circular microphone array. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 622–626, Aug 2016.

[SR06] Sascha Spors and Rudolf Rabenstein. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. In *120th Convention of the Audio Engineering Society*, May 2006.

[SRA08] Sascha Spors, Rudolf Rabenstein, and Jens Ahrens. The theory of wave field synthesis revisited. In *Proc. 124th Convention of the Audio Engineering Society, Amsterdam, May 17-20*, 2008.

[Tar04] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.

[Teu05] Heinz Teutsch. *Wavefield Decomposition Using Microphone Arrays and Its Application to Acoustic Scene Analysis*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005.

[Teu07] Heinz Teutsch. *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*. Springer Berlin Heidelberg, 2007.

[Tit25] E.C. Titchmarsh. A contribution to the theory of Fourier Transforms. *Proceedings of The London Mathematical Society*, s2-23:279–289, 1925.

[TK06] Heinz Teutsch and Walter Kellermann. Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays. *Journal of the Acoustical Society of America*, 120(5):2724–2736, November 2006.

[VB88] Barry D. Van Veen and Kevin M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24, April 1988.

[VMB02] Martin Vetterli, Pina Marziliano, and Thierry Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002. IEEE Signal Processing Society's 2006 Best Paper Award.

[VT02] H.L. Van Trees. *Detection, Estimation, and Modulation Theory: Optimum array processing*. Detection, Estimation, and Modulation Theory. Wiley, 2002.

[WA80] Gabriel Weinreich and Eric B. Arnold. Method for measuring acoustic radiation fields. *The Journal of the Acoustical Society of America*, 68(2):404–411, August 1980.

[WA01] Darren B. Ward and Thushara D. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697–707, September 2001.

[Web19] A.-G. Webster. Acoustical impedance, and the theory of horns and of the phonograph. *Proceedings of the National Academy of Science*, 5:275–282, July 1919.

[WEJ12] P.K.T. Wu, N. Epain, and C. Jin. A dereverberation algorithm for spherical microphone arrays using compressed sensing techniques. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 4053–4056, March 2012.

[WEMJ11] A. Wabnitz, N. Epain, A. McEwan, and C. Jin. Upscaling ambisonic sound scenes using compressed sensing techniques. In *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–4, Oct 2011.

[WHH03] Earl G. Williams, Brian H. Houston, and Peter C. Herdic. Fast fourier transform and singular value decomposition formulations for patch nearfield acoustical holography. *Journal of the Acoustical Society of America*, 114(3):1322–1333, 2003.

[Wil99] Earl G. Williams. *Fourier Acoustic: Sound Radiation and Nearfield Acoustical Holography.* Academic Press, 1999.

[Wil01] Earl G. Williams. Regularization methods for near-field acoustical holography. *Journal of the Acoustical Society of America*, 110(4):1976–1988, 2001.

[WVHK06] Earl G. Williams, Nicolas Valdivia, Peter C. Herdic, and Jacob Klos. Volumetric acoustic vector intensity imager. *Journal of the Acoustical Society of America*, 120(4):1887–1897, 2006.

[ZDG10] Dmitry N. Zotkin, Ramani Duraiswami, and Nail A. Gumerov. Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(1):2–16, January 2010.

[ZFBZ08] Cha Zhang, D. Florencio, D.E. Ba, and Zhengyou Zhang. Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings. *IEEE Transactions on Multimedia*, 10(3):538–548, April 2008.

[Zie53] A. Van Der Ziel. Shot noise in semiconductors. *Journal of Applied Physics*, 24(2):222–223, feb 1953.

[ZJBC09] Yong-Bin Zhang, Finn Jacobsen, Chuan-Xing Bi, and Xin-Zhao Chen. Near field acoustic holography based on the equivalent source method and pressure-velocity transducers. *The Journal of the Acoustical Society of America*, 126(3):1257–1263, sep 2009.

[ZS13] Franz Zotter and Sascha Spors. Is sound field control determined at all frequencies? how is it related to numerical acoustics? In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, Sep 2013.

[ZSH07] Franz Zotter, Alois Sontacchi, and Robert Höldrich. Modeling a spherical loudspeaker system as multipole source. *Fortschritte der Akustik*, 33(1):221, 2007.

[ZZ14] Markus Zaunschirm and Franz Zotter. Measurement-based modal beamforming using planar circular microphone arrays. In *Proc. of the EAA Joint Symposium on Auralization and Ambisonics, Berlin*, 2014.