

The Single-layer Potential Approach Applied to Sound Field Synthesis Including Cases of Non-enclosing Distributions of Secondary Sources

vorgelegt von
Diplom-Ingenieur
Jens Ahrens
aus Friedberg

Von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften
Dr.-Ing.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Alexander Raake
Berichter: Prof. Dr.-Ing. Sebastian Möller
Berichter: Dr.-Ing. Dr. Tech. h. c. Jens Blauert

Tag der wissenschaftlichen Aussprache: 04.10.2010

Berlin 2010

D 83

Abstract

The present dissertation treats the topic of sound field synthesis. The focus lies thereby on serving human listeners although the results can be also exploited in other applications such as underwater acoustics or ultrasonics. A fundamental formulation of the problem is derived based on standard integral equations and the *single-layer potential approach* is identified as a useful tool in order to derive a general solution. An explicit solution is derived exemplarily for inward-radiating spherical distributions of secondary sources.

The drawback of the single-layer potential approach is the fact that it requires secondary source distributions which enclose the receiver volume. Extensions to the single-layer potential approach are proposed which allow for a derivation of explicit solutions for circular, planar, and linear distributions of secondary sources. Based on above described formulation it is shown that the two established analytical approaches of *Wave Field Synthesis* and *Near-field Compensated Higher Order Ambisonics* constitute specific solutions to the general problem which are covered by the single-layer potential solution and its extensions.

The physical theory of the single-layer potential approach requires that the employed distributions of secondary sources are continuous. Such continuous distributions can not be implemented in practice with today's available loudspeaker technology but discrete distributions of loudspeakers have to be used. The consequences of this spatial discretization of the secondary source distribution are analyzed in detail for all above mentioned geometries in different spatial frequency domains, in temporal frequency domain, and in time domain. Two fundamental results are derived: Firstly, the discretization leads to repetitions of the secondary source driving function in a spatial frequency domain which is determined by the geometry of the secondary source distribution under consideration. And secondly, the bandwidth of the driving function with respect to the according spatial frequency domain has essential influence on the properties of the synthesized sound field. As a consequence, the concept of categorizing sound field synthesis approaches according to the bandwidth of the driving function into *narrowband*, *wideband*, and *fullband* approaches is proposed.

It is finally shown how different types of spatial bandwidth limitation can be employed in order to locally increase the accuracy of the synthesized sound field. This concept is termed *local sound field synthesis*.

This thesis presents an instrumentalized analysis of the fundamental physical properties of the problem. Although the presented work aims at audio presentation to human listeners, perception can only be marginally be considered. However, care was taken that the results are presented such that they can be directly used as a basis for experimental perceptual evaluation.

Zusammenfassung

Die vorliegende Dissertation behandelt das Thema der Schallfeldsynthese. Im Fokus steht dabei die Darbietung von Audiosignalen. Die vorgestellten Ergebnisse lassen sich jedoch in anderen Gebieten wie z.B. der Unterwasserakustik oder Ultraschalltechnik anwenden. Eine grundlegende Formulierung des Problems auf Basis von etablierten Integralgleichungen wird erarbeitet. Die Verwendung der *Methode des Einschichtpotentials* als zweckdienliche allgemeine Lösungsmethode wird vorgeschlagen. Eine explizite Lösung wird beispielhaft für einwärts strahlende kugelförmige Sekundärquellenverteilungen erarbeitet.

Den Nachteil der Einschichtpotentialmethode stellt der Umstand dar, dass die behandelten Sekundärquellenverteilungen das Zielvolumen einschließen müssen. Es werden Erweiterungen der Einschichtpotentialmethode vorgeschlagen, welche Lösungen für kreisförmige, ebene und zeilenförmige Sekundärquellenverteilungen ermöglichen. Anhand dieser Ergebnisse wird gezeigt, dass die beiden etablierten analytischen Methoden der *Wellenfeldsynthese* und des *Near-Field-Compensated-Higher-Order-Ambisonics* Spezialfälle der allgemeinen Lösung darstellen und in der Lösung über die Einschichtpotentialmethode enthalten sind.

Die physikalischen Grundlagen der Einschichtpotentialmethode erfordern, dass die betrachteten Sekundärquellenverteilungen kontinuierlich sind. Solche kontinuierlichen Verteilungen lassen sich mit der zur Verfügung stehenden Lautsprechertechnologie praktisch nicht umsetzen. Es müssen diskrete Lautsprecherverteilungen verwendet werden. Die Auswirkungen dieser räumlichen Abtastung der Sekundärquellenverteilung wird detailliert für alle oben genannten Geometrien im Raumfrequenzbereich, im Zeitfrequenzbereich sowie im Zeitbereich untersucht. Zwei grundlegende Ergebnisse werden erarbeitet: Erstens, die Abtastung führt zu Wiederholungen der Sekundärquellenansteuerungsfunktion in einem Raumfrequenzbereich, der durch die Geometrie der betrachteten Sekundärquellenverteilung bestimmt ist. Und zweitens, die räumliche Bandbreite der Sekundärquellenansteuerungsfunktion hat grundlegenden Einfluss auf die Eigenschaften des synthetisierten Schallfeldes. Deshalb wird vorgeschlagen, die verschiedenen Methoden der Schallfeldsynthese anhand der räumlichen Bandbreite der Ansteuerungsfunktion in *schmalbandige*, *breitbandige* und *vollbandige* Methoden einzuordnen.

Letztlich wird gezeigt, wie verschiedene Arten der räumlichen Bandbreiteneinschränkung angewendet werden können, um lokal die Genauigkeit des synthetisierten Schallfeldes zu erhöhen. Dieses Konzept wird *lokale Schallfeldsynthese* getauft.

Die vorliegende Dissertation stellt eine instrumentalisierte Analyse der grundlegenden Eigenschaften des Problems dar. Obwohl die behandelten Methoden primär auf den Menschen als Empfänger abzielen, kann die menschliche Wahrnehmung der synthetisierten Schallfelder nur marginal berücksichtigt werden. Die Ergebnisse wurden jedoch derart aufgearbeitet, dass sie direkt als Basis für eine weiterführende experimentelle Evaluierung der Wahrnehmung dienen können.

Acknowledgments

My special thanks go to Sebastian Möller for putting immeasurable efforts in providing perfect working conditions and for giving me the freedom to work on the topic of sound field synthesis. And, of course, I thank him for reviewing the present dissertation. Irene Hube-Achter's efforts have also contributed to a considerable extent to the pleasantness of my working conditions which I am also very thankful for.

Jens Blauert deserves general acknowledgements for exciting and inspiring discussions over the years; and he deserves special acknowledgements for reviewing the present dissertation and for giving valuable comments and suggestions.

I wish to thank all of my colleagues at Quality and Usability Lab, most notably Matthias Geier, Karim Helwani and Hagen Wierstorf of the audio technology group, Warcel Wältermann and Alexander Raake, and I wish to thank the management of Deutsche Telekom Laboratories for their support and enthusiasm for spatial audio.

The last and thus most important paragraph is dedicated to Sascha Spors who deserves most pronounced acknowledgments for various efforts including introducing me to the topic of sound field synthesis, guiding me through all these years that I have spent at Quality and Usability Lab and Deutsche Telekom Laboratories, and also for organizing my employment after a single phone call. And finally, I am especially thankful for the fact that we have shared and do still share so many of our interests and for the coincidence that brought us together.

Contents

1	Introduction	1
1.1	The History of Audio Presentation	1
1.2	Motivation of the Presented Research	2
1.3	Nomenclature	4
2	Physical Fundamentals of Sound Fields	7
2.1	The Wave Equation	7
2.1.1	General	7
2.1.2	Solutions in Cartesian Coordinates	8
2.1.3	Solutions in Spherical Coordinates	9
2.2	Representations of Sound Fields	12
2.2.1	Representation of Sound Fields as Series of Spherical Harmonics	12
2.2.2	Selected Properties of Bandlimited Spherical Harmonics Series	15
2.2.3	Multipoles	18
2.2.4	Far-Field Radiation	18
2.2.5	The Wavenumber Domain	20
2.2.6	The Angular Spectrum Representation	21
2.3	Boundary Conditions	23
2.3.1	Dirichlet Boundary Condition	23
2.3.2	Neumann Boundary Condition	23
2.3.3	Sommerfeld Radiation Condition	24
2.4	Green's Functions	25
2.5	The Kirchhoff-Helmholtz Integral	25
2.6	Problem Formulation	26
2.6.1	Wave Field Synthesis	27
2.6.2	Simple Source Formulation and Equivalent Scattering Problem	33
2.6.3	Potential Theory	33
2.6.4	Numerical Approaches	37
2.7	Summary	37
3	Continuous Secondary Source Distributions	39
3.1	Introduction	39
3.2	Spherical Secondary Source Distributions	39
3.2.1	Derivation of the Driving Function	40
3.2.2	Synthesized Sound Field	42
3.2.3	Incorporation of Secondary Sources With Complex Radiation Properties	43

3.2.4	Near-Field Compensated Higher Order Ambisonics	47
3.3	Circular Secondary Source Distributions	48
3.3.1	Derivation of the Driving Function	49
3.3.2	Synthesized Sound Field	51
3.3.3	Incorporation of Secondary Sources With Complex Radiation Properties	52
3.4	Planar Secondary Source Distributions	55
3.4.1	Derivation of the Driving Function	55
3.4.2	Physical Interpretation	57
3.4.3	Synthesized Sound Field And Example Driving Function . . .	58
3.5	Linear Secondary Source Distributions	59
3.5.1	Derivation of the Driving Function	59
3.5.2	Synthesized Sound Field And Example Driving Function . . .	60
3.5.3	Incorporation of Secondary Sources With Complex Radiation Properties	62
3.5.4	A Note on Wave Field Synthesis Employing Linear Secondary Source Distributions	64
3.5.5	Truncated Linear Secondary Source Distributions	65
3.6	Approximate Solution for Arbitrary Convex Secondary Source Dis- tributions	67
3.7	Summary	68
4	Discrete Secondary Source Distributions	69
4.1	Introduction	69
4.2	Excursion: Discretization of Time-Domain Signals	69
4.3	Spherical Secondary Source Distributions	73
4.3.1	Discretization of the Sphere	73
4.3.2	Discretization of the Driving Function	74
4.3.3	Properties of the Synthesized Sound Field in Time-Frequency Domain	77
4.4	Circular Secondary Source Distributions	81
4.4.1	Discretization of the Driving Function	81
4.4.2	On the Spatial Bandwidth of Wave Field Synthesis With Cir- cular Secondary Source Distributions	84
4.4.3	Properties of the Synthesized Sound Field in Time-Frequency Domain	86
4.4.4	Properties of the Synthesized Sound Field in Time Domain . .	89
4.4.5	Optimizing the Synthesis with Respect to a Given Receiver Location	95
4.5	Planar Secondary Source Distributions	99
4.6	Linear Secondary Source Distributions	102
4.6.1	Discretization of the Driving Function	102
4.6.2	Properties of the Synthesized Sound Field in Time-Frequency Domain	105
4.6.3	Properties of the Synthesized Sound Field in Time Domain . .	107
4.6.4	Spatial Discretization in Wave Field Synthesis Employing Lin- ear Secondary Source Distributions	109

4.6.5	Optimizing the Synthesis with Respect to a Given Receiver Location	109
4.7	Further Aspects of Discretization and Spatial Truncation With Planar and Linear Secondary Source Distributions	112
4.8	On the Spatial Bandwidth of Numerical Solutions	113
4.9	Summary	114
5	Conclusions and Outlook	115
A	Coordinate Systems	117
B	Definition of the Fourier Transform	119
C	Fourier Transforms of Selected Quantities	121
C.1	Fourier Transforms of a Plane Wave	121
C.2	Fourier Transforms of the Free-Field Green's Function	121
D	Convolution Theorems	123
D.1	Fourier Series Domain	123
D.2	Spherical Harmonics Domain	123
E	Miscellaneous Mathematical Considerations	127
E.1	Translation of Spherical Harmonics Expansions	127
E.2	Recursion Formulae for Exterior-to-Interior Sectorial Translation . . .	129
E.3	The Stationary Phase Approximation Applied to the Rayleigh I Integral	129
	References	131

Chapter 1

Introduction

1.1 The History of Audio Presentation

Since the invention of the telephone, the first electro-acoustic communication device patented by Alexander Graham Bell in 1876 (Bell, 1876), a great variety of audio presentation methods have evolved.

Due to the single loudspeaker which is employed in the telephone only *monaural* auditory information can be provided which limits the presentable spatial information (Blauert, 1997). As early as in 1881, two parallel telephone channels were used in order to transmit performances from the Paris Opera House (du Moncel, 1881; Torick, 1998). The service was commercialized a few years later and termed *Théâtrophone*. The enabled provision of *binaural* auditory information essentially extends the transmittable spatial information.

Later on, two and more loudspeakers were used in *stereophony* which bases on the work of Alan Blumlein carried out in the 1930s (Alexander, 1999). Up to now, stereophony is still the most wide-spread audio presentation method which provides binaural cues. An extension of Stereophony was used in 5.1 systems (employing 5 regular loudspeakers plus one subwoofer) and larger systems (e.g. 22.2 with 22 regular loudspeakers plus 2 subwoofers (Rumsey, 2001; Hamasaki *et al.*, 2005)) installed mainly in cinemas. A further milestone was set by Snow's *acoustic curtain* (Steinberg & Snow, 1934) which formed the basis for advanced sound field synthesis techniques like Wave Field Synthesis (see below). The initial idea comprised a high number of transducers but practical implementations employed only a few and have not received wide attention. From the 1970s on, *Quadraphony* (Torick, 1998) and *Ambisonics* (Gerzon, 1973) were developed in order to provide a domestic surround experience which stereo was not capable of delivering at that time.

Although all above mentioned approaches were initially motivated from a physical perspective, later research showed that their success can be largely attributed to psycho-acoustical properties of the human auditory system, e.g. (Theile, 1980). In order to achieve balanced presentation over a larger listening area than stereophony and alike permit, methods targeting the physical synthesis of sound fields over an extended area have evolved in the recent decades. The best known representatives of the latter are *Near-field Compensated Higher Order Ambisonics* (NFC-HOA) (Daniel, 2001) and *Wave Field Synthesis* (WFS) (Berkhout *et al.*, 1993). Due to the high number of loudspeakers employed which can reach several hundred channels

or even more (de Vries, 2009), the latter approaches are also referred to as *massive multichannel audio presentation methods*, or *holophony*, or *sound field synthesis*. The latter term is used in this thesis for convenience. When frequencies in the audible range are considered sound field synthesis generally indeed addresses human listeners. Ultrasonic methods find application in medical imaging and underwater acoustics, e.g. (Jones, 2001). This thesis focuses on the audible frequency range and thus on applications serving human listeners.

All of above mentioned loudspeaker-based presentation methods can be interpreted as employing a loudspeaker setup which – partly or fully – encloses the listening area. Recently, outward radiating loudspeaker setups become more and more popular. Typically, spherical arrangements are used and the primary target is the synthesis of the radiation properties of a given sound source (Pollow & Behler, 2009; Zotter, 2009a). Although the work presented in this thesis can be straightforwardly extended to such radiation synthesis, it is not the primary focus.

Alternatively to the loudspeaker-based methods, headphone-based methods were proposed which elaborate the ideas behind the Théâtrophone. The acoustical properties of the human body, most notably the head and the outer ears, are imitated in order to create auditory events with specific spatial attributes. These acoustical properties are described by *head-related transfer functions* (HRTFs) and are individual for one person (Blauert, 1997). Typically, a given scene is recorded with a mannequin or a person with ear-mounted microphones, or HRTFs obtained from measurements are imposed on the signals (Hammershøi & Møller, 2002). This approach is also referred to as *binaural presentation*.

Headphones are particularly suited for such presentation since the signals at both ears can be controlled individually. When loudspeakers are used, appropriate cross-talk has to be applied which exhibits fundamental limitations (Gardner, 1997; Nelson & Rose, 2005; Kim *et al.*, 2006).

Different ways of categorizing above mentioned approaches are possible considering e.g. the number of listeners addressed, the size of the preferred listening area, whether the method itself employs HRTFs or addresses the listeners' HRTFs, or whether a physical synthesis of a sound field or rather the evocation of a specific perception is targeted.

The choice of categories depends on the considered situation and purpose of categorization and is not further discussed in this thesis.

1.2 Motivation of the Presented Research

The ultimate goal of research in the field of audio presentation in the recent decades has been the creation of an authentic and plausible auditory perception both in terms of timbre and spatial attributes. *Authenticity* refers to the degree to which the perception is consistent with the perception of a given original. *Plausibility* refers to the degree to which the perception of a given artificial scene is consistent with the experience and expectations a listener.

An undisputable prerequisite for spatial presentation is the capability of a method of producing binaural auditory cues. This prerequisite is fulfilled by all of the audio presentation methods presented in Sec. 1.1 which employ more than

one loudspeaker (either in space or in a headphone). The perceptual evaluation especially with respect to spatial perception is still at an early stage but has received more attention during the last few years (Gabrielsson & Sjgren, 1979; Rumsey, 2002; Rumsey *et al.*, 2005; Lindau *et al.*, 2007; Wittek, 2007; Bertet, 2009). An extensive characterization of the different methods in terms of their perceptual properties has not been available.

What is common to all methods – potentially apart from sound field synthesis methods – is the fact that the size of the optimal listening area can not be arbitrarily extended. Stereophony, Quadraphony, and Ambisonics exhibit a pronounced *sweet spot* in the center of the loudspeaker setup outside of which presentation quality is deteriorated (Dutton, 1962; Bamford & Vanderkooy, 1995). A similar limitation for headphone-based and crosstalk-cancellation based methods is obvious.

The psycho-acoustical mechanisms in the perception of stereophony, Quadraphony, Ambisonics, and similar methods have not been ultimately revealed and open questions persist (Theile, 1980; Blauert, 1997). However, the available results suggest that it is not possible to arbitrarily extend the preferred listening area and even balanced and fully predictable presentation for two or more listeners seems questionable.

The perceptual evaluation of sound field synthesis methods like NFC-HOA and WFS has not received much attention yet. The literature is restricted to a number of localization experiments such as (Start, 1997; de Bruijn, 2004; Sanson *et al.*, 2008) and a limited number of more sophisticated investigations like (Wittek, 2007; Bertet, 2009). Although sound field synthesis methods can potentially satisfy arbitrarily large listening areas, this capability has not been proven due to unavoidable artifacts in practice. As will be discussed in detail in Chap. 4, the properties of the arising artifacts can be influenced so that e.g. regions with only weak artifacts can be created by the cost of stronger artifacts elsewhere. A more or less even distribution of artifacts over the entire listening area can also be achieved.

Neither the detailed properties nor the perception of such artifacts have been investigated so far. A common conceptual framework for methods like NFC-HOA and WFS has not been available and the methods are treated as distinct concepts. Since the fundamental relations between different methods have not been revealed a transfer of results obtained for a specific method to other methods can not be performed and analyses have to be performed distinctly such as e.g. in (Daniel *et al.*, 2003).

This thesis presents a fundamental physical concept which clearly reveals the relationships between the different methods and allows for the transfer of results. Furthermore, a detailed – yet instrumentalized – analysis of the arising artifacts is performed. The motivation is to lay the basis for an experimental perceptual investigation. Such an investigation may in turn lead to the provision of criterions for optimization of the presentation quality by an appropriate *shaping* of the unavoidable artifacts. It might thereby be possible to create a preferred listening area which is significantly larger than for other methods so that multiple listener can be satisfied.

Note that numerical sound field synthesis methods such as (Kirkeby & Nelson, 1993; Ward & Abhayapala, 2001; Daniel, 2001; Poletti, 2005; Hannemann & Donohue, 2008; Kolundžija *et al.*, 2009) are not treated in this thesis. The properties of

such methods especially with respect to artifacts in practice can not be predicted but specific scenarios have to be considered individually. Deduction of fundamental properties is thus difficult.

1.3 Nomenclature

This thesis employs the following notational conventions.

For scalar variables, lower case denotes the time domain, upper case the time-frequency domain, e.g. $s(\mathbf{x}, t)$ vs. $S(\mathbf{x}, \omega)$. Vectors are denoted by lower case bold-face, e.g. \mathbf{k} . The three-dimensional position vector in Cartesian coordinates is given as $\mathbf{x} = [x \ y \ z]^T$; the coordinate systems employed are presented in App. A. The definition of the Fourier transform is outlined in App. B. In order to emphasize which frequency domain is considered in a given situation it is explicitly referred to time-frequency domain and space-frequency domain.

When it is referred to a *sound field* $s(\mathbf{x}, t)$ in this thesis, it is referred to the *sound pressure*, i.e. the *local pressure deviation* from the ambient pressure (in the present case the atmospheric pressure) caused by a sound wave. The SI (*Système international d'unités*) unit of sound pressure is the *pascal* ($1\text{Pa} = 1\frac{\text{N}}{\text{m}^2}$) (Bureau International des Poids et Mesures, 2006). The time-frequency spectrum of a sound field $S(\mathbf{x}, \omega)$, i.e. the spectral amplitude density of $s(\mathbf{x}, t)$, is thus given in $\text{Pa} \cdot \text{s}$ or $\frac{\text{Pa}}{\text{Hz}}$ respectively (Girod *et al.*, 2001). For convenience, $S(\mathbf{x}, \omega)$ is referred to as a “sound field in time-frequency domain” or “sound pressure in time-frequency domain” in this thesis.

Angles are given in radians if not indicated as different.

The following two examples of a plane wave and a spherical wave sound field illustrate further notational conventions. The sound pressure deviation $S_{\text{pw}}(\mathbf{x}, \omega)$ in time-frequency domain caused by a plane wave sound field propagating in direction \mathbf{k}_{pw} is given by

$$S_{\text{pw}}(\mathbf{x}, \omega) = \hat{S}_{\text{pw}}(\omega) e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}}, \quad (1.1)$$

with

$$\mathbf{k}_{\text{pw}}^T = [k_{\text{pw},x} \ k_{\text{pw},y} \ k_{\text{pw},z}] \quad (1.2)$$

$$= k_{\text{pw}} \cdot [\cos \theta_{\text{pw}} \sin \phi_{\text{pw}} \ \sin \theta_{\text{pw}} \sin \phi_{\text{pw}} \ \cos \phi_{\text{pw}}] \quad (1.3)$$

and $(\theta_{\text{pw}}, \phi_{\text{pw}})$ being the propagation direction of the plane wave in spherical coordinates. i denotes the imaginary unit ($i^2 = -1$).

The right hand side of (1.1) is composed of two components:

1. A time-frequency component $\hat{S}_{\text{pw}}(\omega)$ which represents the information with respect to time such as a sine wave or a music signal.
2. A spatial transfer function $e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}}$ representing the spatial information.

The spatial transfer function $e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}}$ as used in (1.1) is of dimension 1 so that $\hat{S}_{\text{pw}}(\omega)$ has to be of the unit $\frac{\text{Pa}}{\text{Hz}}$ in order that (1.1) is correct.

Now consider an outgoing spherical wave sound field $S_{\text{sw}}(\mathbf{x}, \omega)$ originating from the coordinate origin given by

$$S_{\text{sw}}(\mathbf{x}, \omega) = \hat{S}_{\text{sw}}(\omega) \frac{e^{-i\frac{\omega}{c}r}}{r} . \quad (1.4)$$

The spatial transfer function in (1.4) is of unit $\frac{1}{\text{m}}$ so that $\hat{S}_{\text{sw}}(\omega)$ has to be of unit $\frac{\text{Ns}}{\text{m}}$.

This inconsistency is a consequence of the simplifying notational conventions. In order to explicitly account for the physical meaning of the involved functions, quantities like the density of the medium in which the sound wave propagates and alike have to be considered explicitly (Williams, 1999). For notational simplicity, this thesis employs the convention applied widely in the scientific literature of exclusively considering the spatial transfer function of a given sound field or similar quantity under consideration neglecting the time information as well as any constant factors. The explicit composition of the involved time components such as $\hat{S}_{\text{pw}}(\omega)$ and $\hat{S}_{\text{sw}}(\omega)$ is not relevant in the presented investigation and is therefore not treated. The reader is referred to (Williams, 1999).

The results derived e.g. in Chap. 3 are thus not *driving signals* for the individual secondary sources of a given secondary source distribution but *driving functions*. Exclusively, scenarios which are spatially static are considered in this thesis so that in order to derive the driving signal $\hat{D}(\mathbf{x}_0, \omega)$ for a secondary source at position \mathbf{x}_0 from the driving function $D(\mathbf{x}_0, \omega)$, the driving function has to be applied to the input signal $\hat{S}_{\text{in}}(\omega)$ (which represents the information with respect to time) as

$$\hat{D}(\mathbf{x}_0, \omega) = \hat{S}_{\text{in}}(\omega) \cdot D(\mathbf{x}_0, \omega) \quad (1.5)$$

in time-frequency domain or as

$$\hat{d}(\mathbf{x}_0, t) = \hat{s}_{\text{in}}(t) *_t d(\mathbf{x}_0, t) \quad (1.6)$$

in time domain whereby the asterisk $*_t$ denotes convolution with respect to time.

As is common in electrical engineering, complex notation is used for purely real harmonic time-domain signals (Girod *et al.*, 2001). I.e. a unit amplitude cosine wave $\hat{s}_{\text{cos}}(t)$ of radian frequency $\omega_0 = 2\pi f_0$ is notated as

$$\hat{s}_{\text{cos}}(t) := e^{i\omega_0 t} . \quad (1.7)$$

The actual time-domain signal is then obtained by considering exclusively the real part of $\hat{s}_{\text{cos}}(t)$ as

$$\Re\{\hat{s}_{\text{cos}}(t)\} = \cos(\omega_0 t) . \quad (1.8)$$

A list of the most frequently used symbols can be found below.

c	speed of sound in air
i	imaginary unit, $i = \sqrt{-1}$
∇	gradient defined in (2.4) and (2.10)
$\Re\{\cdot\}$	real part
$\Im\{\cdot\}$	imaginary part
$G(\mathbf{x} - \mathbf{x}_0, \omega)$	free-field Green's function for excitation at \mathbf{x}_0
$Y_n^m(\beta, \alpha)$	spherical harmonic of n -th degree and m -th order, defined in (2.15)
$\hat{S}_n^m(r, \omega)$	spherical harmonics expansion coefficients of sound field $S(\mathbf{x}, \omega)$, defined in (2.20)
$\check{S}_{n,i}^m(\omega)$ or $\check{S}_n^m(\omega)$	interior expansion coefficients of sound field $S(\mathbf{x}, \omega)$, defined in (2.21a)
$\check{S}_{n,e}^m(\omega)$	exterior expansion coefficients of sound field $S(\mathbf{x}, \omega)$, defined in (2.21b)
$\tilde{S}(k_x, k_y, z, \omega)$	sound field $S(\mathbf{x}, \omega)$ considered in wavenumber domain with respect to k_x and k_y
$\check{S}(\cdot)$	angular spectrum representation of $S(\cdot)$, defined in (2.38a)
$P_n^m(\cdot)$	associated Legendre function of n -th degree and m -th order (Gumerov & Duraiswami, 2004)
$(I I)_{nn'}^{mm'}(\cdot)$	translation coefficient for interior-to-interior translation
$(E I)_{nn'}^{mm'}(\cdot)$	translation coefficient for exterior-to-interior translation
(α, β)	direction given by azimuth α and colatitude β
$\text{sinc } x$	Sinc function, $\text{sinc } x = \frac{\sin(\pi x)}{\pi x}$
$\langle \cdot \rangle$	inner product (Weisstein, 2002)
$\mathbf{x} = [x \ y \ z]^T$	position vector in Cartesian coordinates (App. A)
\mathbf{x}^T	transposition of vector \mathbf{x}
$ x $	absolute value (Weisstein, 2002)
$ \mathbf{x} $	vector norm (Weisstein, 2002)
$\vec{\mathbf{e}}_x$	unit vector pointing in x -direction
$\delta(\cdot)$	Dirac delta function
$\delta_{nn'}$	Kronecker Delta, defined in (2.18)
$\partial\Omega$	boundary enclosing volume Ω_i
Ω_i	volume enclosed by boundary $\partial\Omega$
Ω_e	domain exterior to boundary $\partial\Omega$
$\gamma_{n_1, n_2, n}^{m_1, m_2, m}$	Gaunt coefficient, defined in (D.6)
$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$	Wigner 3j-Symbol as defined in (Weisstein, 2002)
$\mathcal{E} \begin{pmatrix} m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{pmatrix}$	E-symbol, defined in (D.7)
$(\cdot)!$	factorial (Gumerov & Duraiswami, 2004)
$\frac{\partial}{\partial \mathbf{n}}$	gradient in direction \mathbf{n} , refer to (2.44)
$dA(\mathbf{x}_0)$	infinitesimal surface element

Chapter 2

Physical Fundamentals of Sound Fields

2.1 The Wave Equation

2.1.1 General

In order for a sound field $s(\mathbf{x}, t)$ to be physically possible it has to satisfy the *scalar wave equation* in the domain (i.e. the volume) of interest. When *source-free* domains are considered, the wave equation is termed being *homogeneous* and is given by (Morse & Feshbach, 1953)

$$\nabla^2 s(\mathbf{x}, t) + \frac{1}{c^2} \frac{\partial^2 s(\mathbf{x}, t)}{\partial t^2} = 0 . \quad (2.1)$$

c denotes the speed of sound in air which is assumed to be $343 \frac{\text{m}}{\text{s}}$ throughout this thesis. The zero on the right hand side of (2.1) indicates the absence of sources. Consistently, when (2.1) exhibits a source term on the right hand side, it is termed *inhomogeneous*.

The *Laplacian* ∇^2 is a scalar differential operator yielded by applying twice the *gradient* ∇ . Explicit expressions for ∇ will be introduced in Sec. 2.1.2 and 2.1.3 in conjunction with the solutions to the wave equation with respect to different coordinate systems.

Assuming *steady-state* conditions and harmonic time dependence and applying a temporal Fourier transform as defined in App. B on the time domain wave equation (2.1) yields the *scalar Helmholtz equation* which is given by (Morse & Feshbach, 1953)

$$\nabla^2 S(\mathbf{x}, \omega) + k^2 S(\mathbf{x}, \omega) = 0 . \quad (2.2)$$

Eq. (2.2) will play a central role in this thesis. k is termed wavenumber (although it is rather a coefficient than a number) and is of unit $\frac{\text{rad}}{\text{m}}$. It is related to the radian frequency ω via

$$k^2 = \left(\frac{\omega}{c} \right)^2 . \quad (2.3)$$

The radian frequency ω is related to the time frequency f via $\omega = 2\pi f$ and is of unit $\frac{\text{rad}}{\text{s}}$. The *wavelength* λ , measured in m, is given by $\lambda = \frac{c}{f} = \frac{2\pi}{k}$.

Note that exclusively sound propagation in *homogeneous* and *non-dissipative* (i.e. lossless) media is considered throughout this thesis. Finally, the wave equation (2.1) and the Helmholtz equation (2.2) assume that the medium which is air in the present case is perfectly linear. This assumption is only met for infinitesimal sound pressures so that (2.1) and (2.2) are essentially approximations. However, it has been shown that (2.1) and (2.2) provide useful results when sound pressures are considered which are below the threshold of pain of the human auditory system (Gumerov & Duraiswami, 2004, p. 2-3).

2.1.2 Solutions in Cartesian Coordinates

The gradient ∇ in Cartesian coordinates is given by (Weisstein, 2002)

$$\nabla = \frac{\partial}{\partial x} \vec{e}_x + \frac{\partial}{\partial y} \vec{e}_y + \frac{\partial}{\partial z} \vec{e}_z, \quad (2.4)$$

whereby \vec{e}_i denotes the unit vector in indexed direction. Refer to App. A for an illustration of the coordinate system.

Solutions to the Helmholtz equation (2.2) in Cartesian coordinates are given by (Williams, 1999, p. 21)

$$S(\mathbf{x}, \omega) = \hat{S}(\omega) e^{-i(k_x x + k_y y + k_z z)} = \hat{S}(\omega) e^{-i\mathbf{k}^T \mathbf{x}}, \quad (2.5)$$

and thus constitute plane waves.

Eq. (2.5) is satisfied as long as the *dispersion relation*

$$k^2 = k_x^2 + k_y^2 + k_z^2 \quad (2.6)$$

is fulfilled. Thus, the wavenumber k represents the length of the propagation vector $\mathbf{k} = [k_x \ k_y \ k_z]^T$. Eq. (2.6) can be rearranged to read

$$k_y^2 = k^2 - k_x^2 - k_z^2. \quad (2.7)$$

Note that there is no restriction on the values of k_x^2 and k_z^2 in (2.7) provided that they are real (Williams, 1999, p. 21). Taking the square root of (2.7) yields

$$k_y = \begin{cases} \pm \sqrt{k^2 - k_x^2 - k_z^2} & \text{for } k^2 \geq k_x^2 + k_z^2 \\ \pm i \sqrt{k_x^2 + k_z^2 - k^2} & \text{for } k_x^2 + k_z^2 \geq k^2 \end{cases} \quad (2.8)$$

since k is non-negative.

The first case in (2.8) represents a *propagating* or *homogeneous* plane wave. The vector \mathbf{k} points into the direction of propagation. Refer to Fig. 2.1(a) for a simulation.

The second case in (2.8) (with complex k_y) represents an *evanescent* or *inhomogeneous* wave. Inserting k_y into (2.5) yields

$$S_{\text{pw}}(\mathbf{x}, \omega) = \hat{S}_{\text{pw}}(\omega) e^{\pm \sqrt{k_x^2 + k_z^2 - k^2} y} e^{-i(k_{\text{pw},x} x + k_{\text{pw},z} z)}. \quad (2.9)$$

Note that the first exponential in (2.9) is purely real. For $y > 0$, the positive exponent in the first exponential in (2.9) is non-physical since it blows up for $y \rightarrow +\infty$ so that the solution is restricted to the decaying term (the negative exponent) for this case (Williams, 1999). Refer to Fig. 2.1(b) for a simulation of an evanescent wave decaying in y direction.

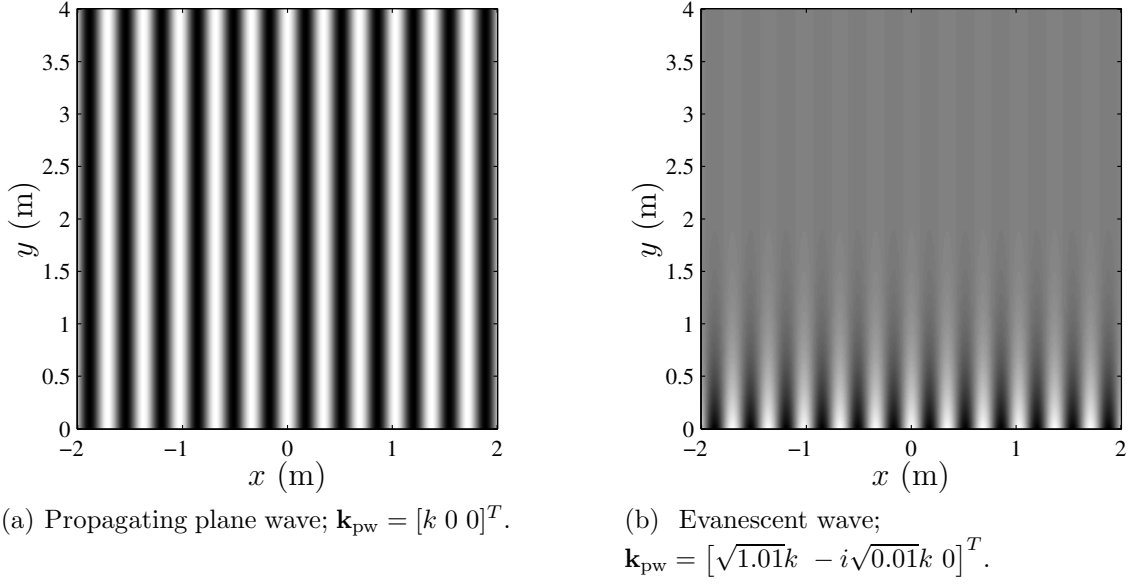


Figure 2.1: Propagating and evanescent waves of frequency $f_{\text{pw}} = 1000$ Hz. A cross-section through the horizontal plane is shown.

2.1.3 Solutions in Spherical Coordinates

The gradient ∇ in spherical coordinates is given by (Weisstein, 2002)

$$\nabla = \frac{\partial}{\partial r} \vec{\mathbf{e}}_r + \frac{1}{r} \frac{\partial}{\partial \beta} \vec{\mathbf{e}}_\beta + \frac{1}{r \sin \beta} \frac{\partial}{\partial \alpha} \vec{\mathbf{e}}_\alpha . \quad (2.10)$$

Refer to App. A for an illustration of the coordinate system.

Solutions to the Helmholtz equation (2.2) in spherical coordinates are obtained by *separation of variables* (Gumerov & Duraiswami, 2004, p. 41) and are of the form

$$S(\mathbf{x}, \omega) = \Pi(r) \Theta(\alpha) \Phi(\beta) . \quad (2.11)$$

The radial solutions $\Pi(r)$ are given by the *spherical Bessel functions* $j_n\left(\frac{\omega}{c}r\right)$ and the *spherical Neumann functions* $y_n\left(\frac{\omega}{c}r\right)$ of order $n \in \mathbb{N}_0$. Another set of solutions is given by the *spherical Hankel functions*

$$h_n^{(1,2)}\left(\frac{\omega}{c}r\right) = j_n\left(\frac{\omega}{c}r\right) \pm i y_n\left(\frac{\omega}{c}r\right) . \quad (2.12)$$

Refer to Fig. 2.2 for illustrations.

It can be shown that $h_n^{(2)}\left(\frac{\omega}{c}r\right)$ represents outgoing waves and $h_n^{(1)}\left(\frac{\omega}{c}r\right)$ represents incoming waves for the definition of the Fourier transform used in this thesis (refer to App. B). $j_n\left(\frac{\omega}{c}r\right)$ represents standing wave solutions and

$$j_n\left(\frac{\omega}{c}r\right) = \frac{1}{2} \left(h_n^{(1)}\left(\frac{\omega}{c}r\right) + h_n^{(2)}\left(\frac{\omega}{c}r\right) \right) \quad (2.13)$$

holds.

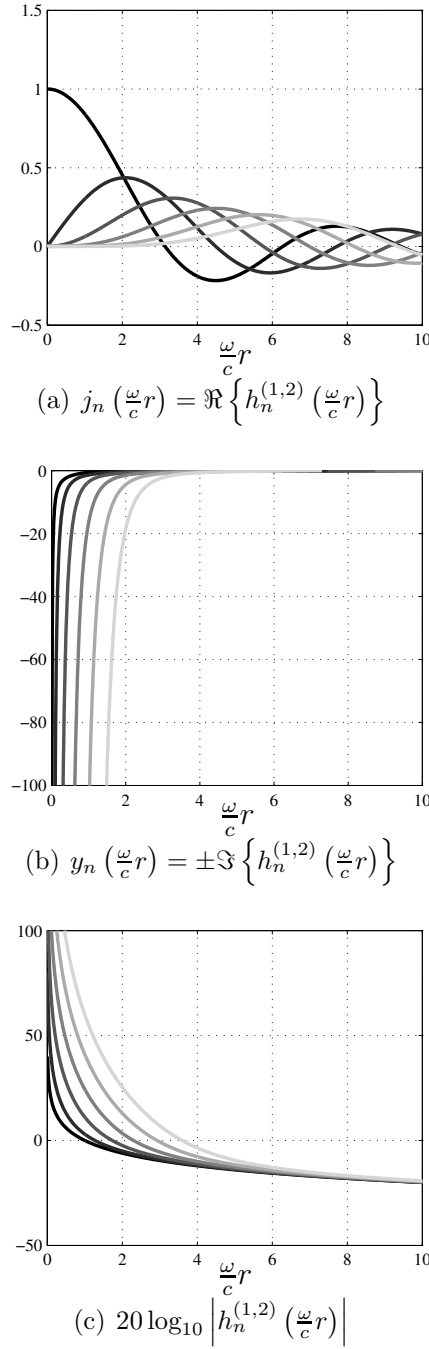


Figure 2.2: Bessel, Neumann, and Hankel functions for $0 \leq n \leq 5$. Brighter color indicates a higher order n .

In certain situations the large-argument approximation of the spherical Hankel functions given by (Gumerov & Duraiswami, 2004, p. 59)

$$h_n^{(1,2)}\left(\frac{\omega}{c}r\right) \approx (\mp i)^{(n+1)} \frac{e^{\pm i \frac{\omega}{c}r}}{\frac{\omega}{c}r} = (\mp i)^n h_0^{(1,2)}\left(\frac{\omega}{c}r\right) \quad \forall \quad \frac{\omega}{c}r \rightarrow +\infty \quad (2.14)$$

will be employed in order to simplify problems. In the present context (2.14) constitutes a far-field/high-frequency approximation.

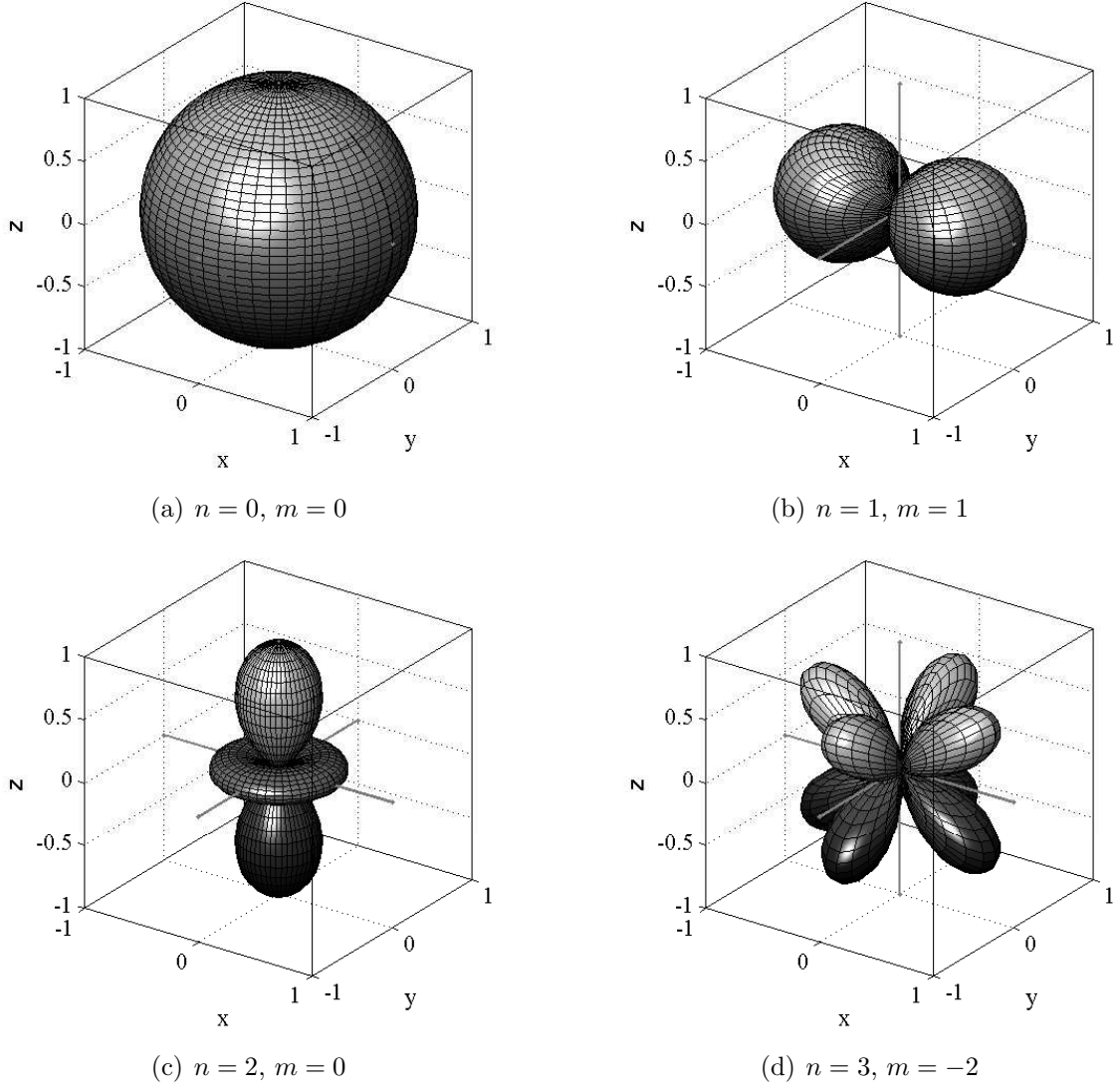


Figure 2.3: $|\Re\{Y_n^m(\beta, \alpha)\}|$ for a selection of n and m .

The azimuthal solutions $\Theta(\alpha)$ in (2.11) are given by the exponentials $e^{im\alpha}$ with $m \in \mathbb{Z}$ and the colatitudinal solutions $\Phi(\beta)$ are given by the *associated Legendre functions* $P_n^m(\cos \beta)$. Both the exponentials $e^{im\alpha}$ and the associated Legendre functions are orthogonal for a given order m .

The only fundamental property of associated Legendre functions mentioned here is the fact that they vanish for $|m| > n$. General properties of $e^{im\alpha}$ and the associated Legendre functions can be deduced from the illustrations of spherical harmonics in Fig. 2.3 (see explanations below).

The solutions of the Helmholtz equation for the angular variables α and β are typically combined together with normalization factors into the *spherical surface harmonics* or *spherical harmonics* $Y_n^m(\beta, \alpha)$. In this thesis, the definition of the spherical harmonics from (Gumerov & Duraiswami, 2004) is employed which is

given by

$$Y_n^m(\beta, \alpha) = (-1)^m \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_n^{|m|}(\cos \beta) e^{im\alpha} . \quad (2.15)$$

Like the associated Legendre functions, spherical harmonics $Y_n^m(\beta, \alpha)$ vanish for $|m| > n$. Refer to Fig. 2.3 for an illustration of selected spherical harmonics.

Note that other variants of the definition (2.15) exist which differ mainly with respect to the factor $(-1)^m$, e.g. (Condon & Shortley, 1935; Arfken & Weber, 2005; Williams, 1999). The choice of this factor is not essential but is rather made upon practical considerations.

The advantage of definition (2.15) is the fact that it inherently handles negative m and avoids the case differentiation which is required in alternative definitions. Furthermore, the complex conjugate $Y_n^m(\beta, \alpha)^*$ can be expressed by negating the degree m as (Gumerov & Duraiswami, 2004)

$$Y_n^m(\beta, \alpha)^* = Y_n^{-m}(\beta, \alpha) . \quad (2.16)$$

Spherical harmonics are orthonormal so that the relation

$$\int_0^{2\pi} \int_0^\pi Y_n^m(\beta, \alpha) Y_{n'}^{-m'}(\beta, \alpha) \sin \beta \, d\beta \, d\alpha = \delta_{nn'} \delta_{mm'} \quad (2.17)$$

holds (Williams, 1999), whereby $\delta_{nn', mm'}$ denotes the Kronecker Delta defined as (Weisstein, 2002)

$$\delta_{nn'} = \begin{cases} 1 & \text{for } n = n' \\ 0 & \text{for } n \neq n' \end{cases} . \quad (2.18)$$

Furthermore, spherical harmonics satisfy the completeness relation (Williams, 1999)

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\beta, \alpha) Y_n^{-m}(\beta', \alpha') = \delta(\alpha - \alpha') \delta(\beta - \beta') . \quad (2.19)$$

2.2 Representations of Sound Fields

2.2.1 Representation of Sound Fields as Series of Spherical Harmonics

As mentioned above, spherical harmonics constitute an orthonormal and complete set of solutions to the Helmholtz equation (2.2). Any solution $S(\mathbf{x}, \omega)$ (i.e. any sound field) can thus be expressed by its according expansion coefficients $\hat{S}_n^m(r, \omega)$ as (Arfken & Weber, 2005, p. 790)

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \hat{S}_n^m(r, \omega) Y_n^m(\beta, \alpha) . \quad (2.20)$$

The representation of a function $S(\mathbf{x}, \omega)$ as such a double series is a generalized Fourier series known as a *Laplace series* (Arfken & Weber, 2005, p. 790).

It can be shown that *interior* and *exterior* problems have to be considered separately (Williams, 1999, p. 207, 217). Interior problems are problems which consider domains which are free of sound sources and obstacles, i.e. all sound sources and obstacles are located outside the considered domain. Exterior problems on the other hand consider domains which are exterior to a distribution of sound sources and obstacles. Exterior problems do not necessarily extend to infinity. They can thus as well be interior with respect to a second sound source distribution. In the latter case, this interjacent problem is then described as a superposition of an interior and an exterior problem.

When considering series of spherical surface harmonics the boundaries to interior and exterior problems are spherical and are centered around the origin of the coordinate system employed. The interior domain is thus a sphere centered around the origin of the coordinate system which is tangent to the closest sound source of a source distribution and which does not cut through the source distribution at any point. The precise definition of the exterior domain is accordingly. Refer to Fig. 2.4 for an illustration.

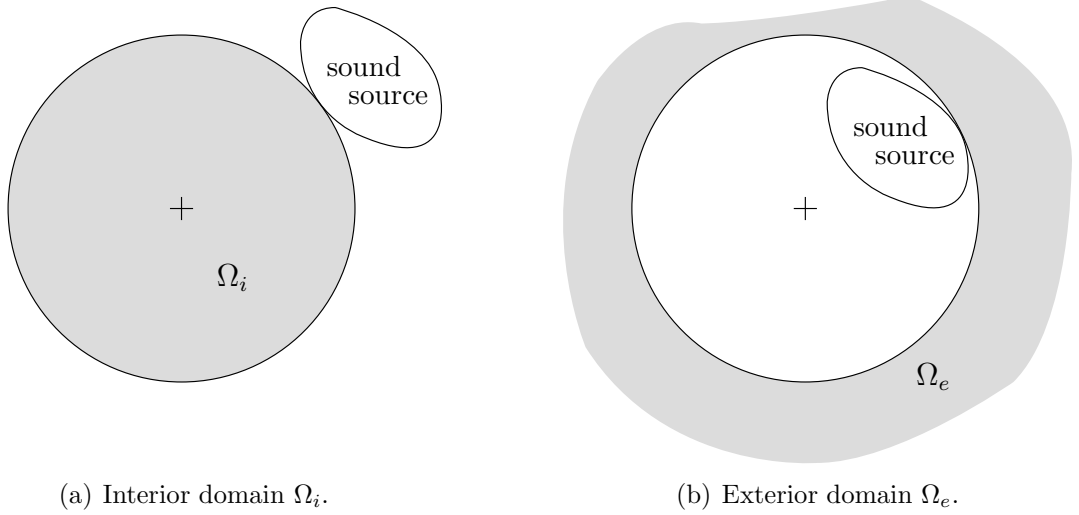


Figure 2.4: Examples of interior and exterior problems. Shaded areas denote the domains of interest. The cross indicates the origin of the coordinate system.

Any sound field $S(\mathbf{x}, \omega)$ can be described in the interior domain Ω_i by

$$S_i(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{n,i}^m(\omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha), \quad (2.21a)$$

and in the exterior domain Ω_e by

$$S_e(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{n,e}^m(\omega) h_n^{(2)}\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha). \quad (2.21b)$$

Note that the existence of an exterior domain suggests that the sound source or the sound source distribution which evokes the sound field under consideration has finite spatial extent.

The coefficients $\check{S}_{n,i}^m(\omega)$ and $\check{S}_{n,e}^m(\omega)$ respectively can be obtained by exploiting the orthogonality of the spherical harmonics as

$$\check{S}_{n,i}^m(\omega) = \frac{1}{j_n\left(\frac{\omega}{c}r\right)} \int_0^{2\pi} \int_0^\pi S(\mathbf{x}, \omega) Y_n^{-m}(\beta, \alpha) \sin \beta \, d\beta \, d\alpha \quad (2.22)$$

for the interior problem and accordingly for the exterior problem. This thesis considers mainly interior problems and the index “ i ” is generally dropped for notational convenience except for specific situations.

Since expansions (2.21) converge uniquely and uniformly, the order of summation may be exchanged (Gumerov & Duraiswami, 2004, p. 75). If the spherical harmonics $Y_n^m(\beta, \alpha)$ are then expressed by their explicit formulation (2.15), the Fourier series which is inherent to (2.21) is revealed. It is given by

$$S(\mathbf{x}, \omega) = \underbrace{\sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} \check{S}_n^m(\omega) j_n\left(\frac{\omega}{c}r\right) (-1)^m \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_n^{|m|}(\cos \beta) e^{im\alpha}}_{= \mathring{S}_m(r, \beta, \omega)} \quad (2.23)$$

exemplarily for the interior expansion. The Fourier series expansion coefficients of $S(\mathbf{x}, \omega)$ are denoted by $\mathring{S}_m(r, \beta, \omega)$. As mentioned in Sec. 2.1.3, the basis functions $e^{im\alpha}$ of the Fourier series are orthogonal for $m \in \mathbb{Z}$. Furthermore, they constitute a complete set and the orthogonality relation (Williams, 1999)

$$\frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{im\alpha} e^{-im\alpha'} = \delta(\alpha - \alpha') \quad (2.24)$$

holds. The inverse operation to (2.23) is given by

$$\mathring{S}_m(r, \beta, \omega) = \frac{1}{2\pi} \int_0^{2\pi} S(\mathbf{x}, \omega) e^{-im\alpha} \, d\alpha \quad (2.25)$$

The expansions of the most basic sound fields in free-field, namely spherical and plane waves, are (Williams, 1999; Gumerov & Duraiswami, 2004)

$$\frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_s|}}{|\mathbf{x}-\mathbf{x}_s|} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\left(-i\right)\frac{\omega}{c}h_n^{(2)}\left(\frac{\omega}{c}r_s\right)Y_n^{-m}(\beta_s, \alpha_s)j_n\left(\frac{\omega}{c}r\right)Y_n^m(\beta, \alpha)}_{= \check{S}_{n,sw,i}^m} \quad \forall r < r_s \quad (2.26a)$$

$$\frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_s|}}{|\mathbf{x}-\mathbf{x}_s|} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\left(-i\right)\frac{\omega}{c}j_n\left(\frac{\omega}{c}r_s\right)Y_n^{-m}(\beta_s, \alpha_s)h_n^{(2)}\left(\frac{\omega}{c}r\right)Y_n^m(\beta, \alpha)}_{= \check{S}_{n,sw,e}^m} \quad \forall r > r_s \quad (2.26b)$$

for a spherical wave originating from (r_s, α_s, β_s) and

$$e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{4\pi i^{-n} Y_n^{-m}(\phi_{\text{pw}}, \theta_{\text{pw}})}_{= \check{S}_{n,\text{pw}}^m} j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad (2.27)$$

for a plane wave with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}})$. For plane waves no exterior expansion exists since the source is assumed to be at infinite distance, thus making the interior domain infinite.

Occasionally in this thesis, a given sound field will be considered with respect to two different coordinate systems. The spherical harmonics expansions of the given sound field with respect to the two coordinate systems are related by a translation operation. This translation of coordinate systems is not straightforward. App. E.1 summarizes one compact representation thereof. Selected alternative representations are outlined in Sec. 3.2.3 and 3.3.3. An extensive treatment can be found in (Gumerov & Duraiswami, 2004).

2.2.2 Selected Properties of Bandlimited Spherical Harmonics Series

Consider a bandlimited series

$$S(\mathbf{x}, \omega) \approx \sum_{n=0}^{N-1} \sum_{m=-n}^n \check{S}_n^m(r, \omega) Y_n^m(\beta, \alpha). \quad (2.28)$$

Above a certain threshold N_{min} , (2.28) converges uniformly for given r and ω (Kennedy *et al.*, 2007; Gumerov & Duraiswami, 2004) so that any such bandlimited series constitutes an approximation of $S(\mathbf{x}, \omega)$ the error of which decreases with increasing $N > N_{\text{min}}$. In the case of (2.28), i.e. $\check{S}_n^m(r, \omega) = 0 \ \forall \ n > N-1$, one speaks of an *N-truncated sum* (Gumerov & Duraiswami, 2004, p. 75), an expansion with spatial bandwidth $N-1$, or an $(N-1)$ -th order expansion. When simulations are presented in this thesis which depict quantities of infinite order, the order of the simulations is chosen such that the result becomes indistinguishable from the exact representation.

A thorough analysis of accuracy and properties of bandlimited expansions like (2.28) is cumbersome since the properties strongly depend on a number of factors including the propagation direction of the sound field $S(\mathbf{x}, \omega)$ under consideration in the domain of interest. The reader is referred to (Gumerov & Duraiswami, 2004, chap. 9) for an extensive mathematical treatment. An explicit review of this treatment is waived here since the perceptual consequences of such a spatial bandwidth limitation can not be deduced from mathematical treatments.

In the following the most basic properties of spatially bandlimited expansion which are important in the context of this thesis are summarized. Note that the properties presented below can not be seen as general. They are valid only if the stated assumptions are met.

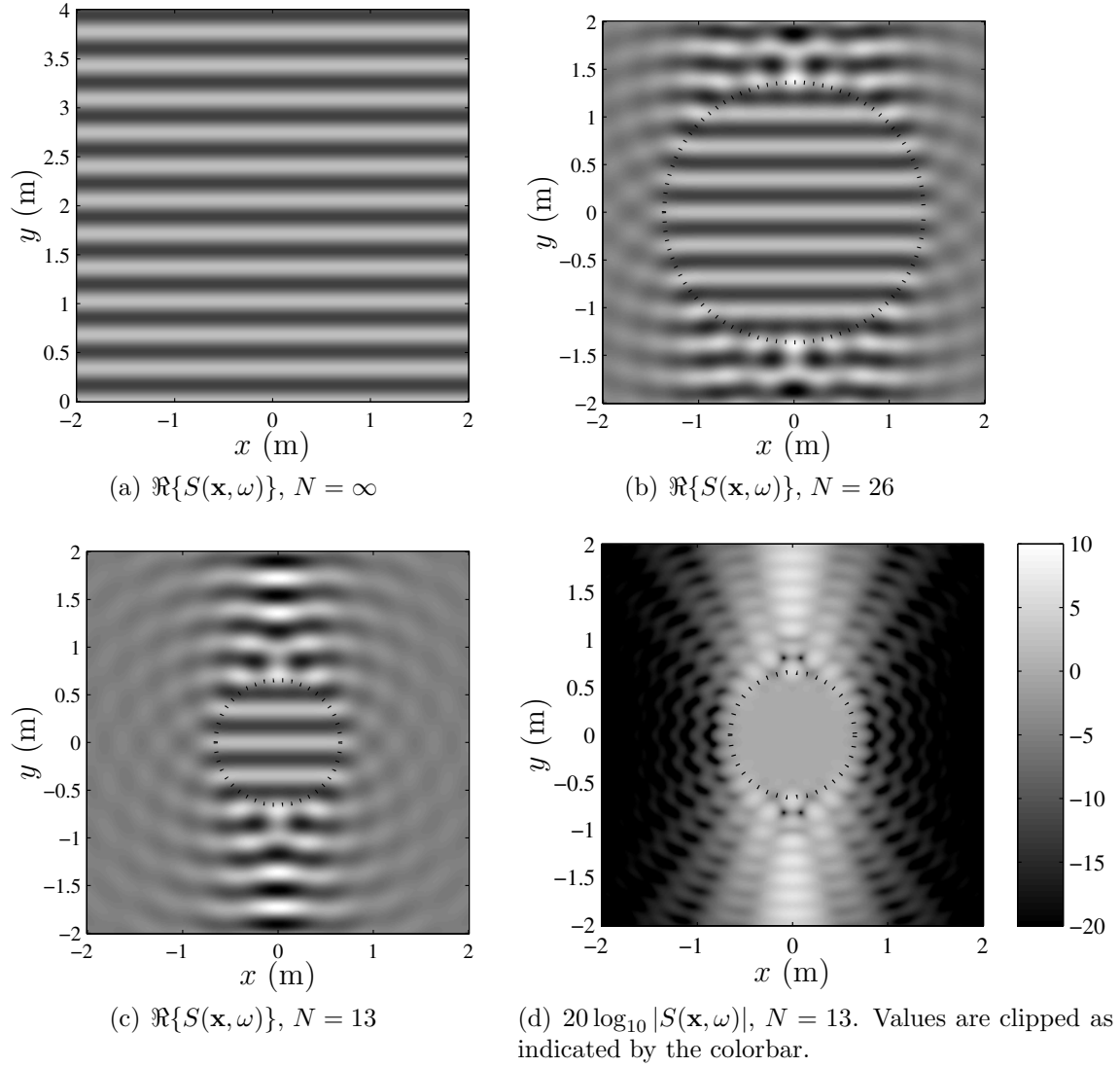


Figure 2.5: A monochromatic plane wave sound field $S(\mathbf{x}, \omega)$ with propagation direction $(-\frac{\pi}{2}, \frac{\pi}{2})$ of frequency $f = 1000$ Hz (Fig. 2.5(a)) and bandlimited approximations thereof with different bandwidths (Fig. 2.5(b)–(d)). The dotted circles bound the r_{N-1} -region. A cross-section through the horizontal plane is shown.

Interior Expansions

The properties of interior spherical harmonics expansions can be summarized as follows: Low orders generally describe the represented sound field close to the expansion center (i.e. the origin of the coordinate system), and higher orders describe the represented sound field at locations at far distances of the expansion center.

This circumstance is directly reflected by the properties of the spherical Bessel functions $j_n(\cdot)$ (refer to Fig. 2.2(a) in Sec. 2.1.3): The higher the order n of the Bessel function, the higher is the argument $\frac{\omega}{c}r$ at which the maximum value is reached (Abramowitz & Stegun, 1968).

Typically, the domain inside which a bandlimited sound field description is considered to be comparable to its full-band analogon is assumed to be inside the radius

r_{N-1} at which $(N-1) = \frac{\omega}{c} r_{N-1}$, whereby $(N-1)$ is the highest order contained in the expansion (Gumerov & Duraiswami, 2004, p. 427). In the remainder of this thesis, the domain bounded by a sphere of radius r_{N-1} will be referred to as r_{N-1} -region. Note that r_{N-1} is directly proportional to the time frequency f .

Note furthermore that a bandlimited approximation is exact at the expansion center – i.e. the origin of the coordinate system – since the only mode which contributes there is the zero-th order mode. At the origin all higher modes are equal to zero.

Fig. 2.5 depicts a monochromatic plane wave with propagation direction $(-\frac{\pi}{2}, \frac{\pi}{2})$ (Fig. 2.5(a)), a 25th-order approximation of the plane wave (Fig. 2.5(b)), a 12th-order approximation of the plane wave (Fig. 2.5(c)), and the magnitude of the latter (Fig. 2.5(d)). The circles bound the corresponding r_{N-1} -region. It can be seen that the approximation with larger bandwidth describes the original sound field over a larger volume. As apparent especially in Fig. 2.5(d), outside the r_{N-1} -region the amplitude of the bandlimited approximation can be higher than that of the exact representation. This circumstance constitutes *Gibbs phenomenon* (Weisstein, 2002).

From Fig. 2.5(c) and 2.5(d) it is evident that the approximation can exhibit very low amplitude in those locations which are outside of the r_{N-1} -region and which are not along the channel of propagation of the sound field which crosses the r_{N-1} -region.

Consider now a sound field carrying a signal which is broadband with respect to the time frequency. When the spatial bandwidth of the sound field is constant over the entire time-frequency range, then the sound field has a larger “extent” at low frequencies than at high frequencies. At positions closer to the expansion center more energy is apparent at higher time frequencies than at farther positions.

Exterior Expansions

In the following it is assumed for convenience that the sound source under consideration is located in the origin of the coordinate system.

An elementary type of sound source is a point source the spatio-temporal transfer function of which is given by (Williams, 1999; Gumerov & Duraiswami, 2004)

$$\frac{e^{-i\frac{\omega}{c}r}}{r} = -i\frac{\omega}{c}h_0^{(2)}\left(\frac{\omega}{c}r\right) = -\sqrt{4\pi}i\frac{\omega}{c}h_0^{(2)}\left(\frac{\omega}{c}r\right)Y_0^0(\beta, \alpha), \quad (2.29)$$

and thus employs only 0-th order.

For illustration of the properties of sound sources with more complex radiation properties consider a sound source whose spatio-temporal transfer function is given by (Ahrens & Spors, 2010b)

$$\check{G}_{n,e}^m(\omega) = \begin{cases} (-1)^{(m+n)}i^{-n}\frac{(N-1)!N!}{(N+n)!(N-n-1)!}Y_n^{-m}(\beta_{\text{or}}, \alpha_{\text{or}}) & \forall n \leq N-1 \\ 0 & \text{elsewhere} \end{cases} \quad (2.30)$$

Note that (2.30) was derived from the in-phase driving function in Ambisonics amplitude panning (Daniel, 2001). $(\alpha_{\text{or}}, \beta_{\text{or}})$ denotes the main radiation direction of the source, i.e. its nominal orientation. Eq. (2.30) represents a purely real spatio-temporal transfer function.

Refer to Fig. 2.6 which depicts the sound field radiated by sources whose spatio-temporal transfer functions are given by (2.30) for $(\alpha_{\text{or}}, \beta_{\text{or}}) = (0, \frac{\pi}{2})$ and $N = 4$ and $N = 21$ respectively. The far-field directivities of the two sound sources are also depicted (refer to Sec. 2.2.4 for a treatment of far-field radiation). It can be seen especially in Fig. 2.6(c) that the emitted sound field exhibits very high values in the vicinity of the sound source which is located in the origin of the coordinate system. This circumstance is also represented in Fig. 2.2(c) by the fact that the higher the order n of a Hankel function the larger is its magnitude especially for low arguments.

The high pressure values apparent in Fig. 2.6(a) and Fig. 2.6(c) are caused by the evanescent components of the sound field. Note that the sound pressure in Fig. 2.6(c) clips over a larger area than in Fig. 2.6(a). Considerable evanescent field components indicate the vicinity of vibrating surface (i.e a sound source) (Williams, 1999). Larger bandwidths thus suggest a larger spatial extent of a source. Note however that this is not a general rule.

Finally, it can be seen from Fig. 2.6 that the directivity of the source with bandwidth $N = 21$ exhibits a stronger focus in the main radiation direction.

A strong frequency dependency like with the properties of interior expansion treated above is not present here.

2.2.3 Multipoles

Radiating solutions to the Helmholtz equation can be represented by *multipole expansions*, i.e. by combinations of monopoles located at infinitesimal distance from each other (Gumerov & Duraiswami, 2004). Lower order multipoles are also referred to as *monopoles, dipoles, quadrupoles, octopoles, etc.*

Multipole expansions are closely related to spherical harmonics expansion with the fundamental difference that the former are not unique and thus do not form a basis in the strict sense. Multipole expansions will only play a marginal role in the context of this thesis and are therefore not treated in detail but only their existence is mentioned. The reader is referred to (Gumerov & Duraiswami, 2004) for a more extensive treatment.

2.2.4 Far-Field Radiation

As outlined in Sec. 2.2.1, the spatio-temporal transfer function of any stationary sound source of finite spatial extent can be represented in the exterior domain by a series of spherical harmonics $Y_n^m(\beta, \alpha)$ and appropriate coefficients, eq. (2.21b). In the remainder of this section it is assumed that the sound source under consideration is located in the origin of the coordinate system.

When the spatio-temporal transfer function is evaluated in the far-field, i.e. for $\frac{\omega}{c}r \rightarrow +\infty$, then the large-argument approximation of the spherical Hankel functions (2.14) can be applied on (2.21b) (Williams, 1999). This results in

$$G(\mathbf{x}, \omega) \approx h_0^{(2)}\left(\frac{\omega}{c}r\right) \sum_{n=0}^{\infty} \sum_{m=-n}^n i^n \check{G}_{n,e}^m(\omega) Y_n^m(\beta, \alpha). \quad (2.31)$$

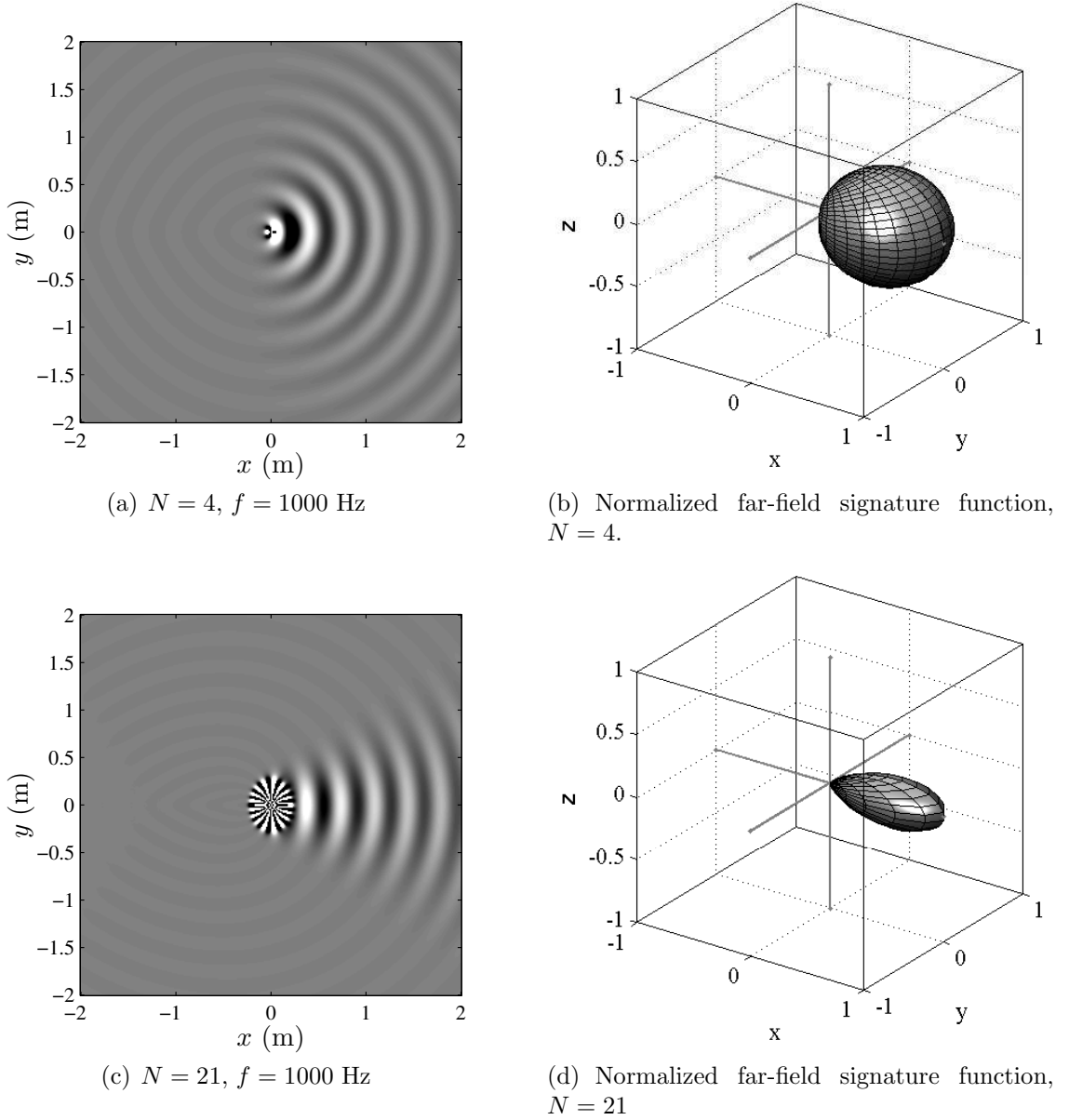


Figure 2.6: Sound fields the horizontal plane and far-field signature functions of monochromatic sound sources with a spatio-temporal transfer function given by (2.30).

Thus, at sufficient distance any stationary sound source of finite spatial extent radiates like a point source ($h_0^{(2)}(\cdot) \sim \frac{1}{r} e^{-i\frac{\omega}{c}r}$, see (2.29)) whereby the angular dependency of the transfer function is given by an appropriate summation of the coefficients $\check{G}_{n,e}^m(\omega)$.

Note that it is actually not rigorous to apply the large-argument approximation on (2.21b) since the former does not hold uniformly in n . Rigorous treatments can be found in (Colton & Kress, 1998; Gumerov & Duraiswami, 2004) which also lead to (2.31). The detailed derivation of (2.31) is not performed here since it is not relevant for the remainder of this thesis.

It is a remarkable property of the coefficients $\check{G}_{n,e}^m(\omega)$ that they serve both as an

exact representation of a sound field (as in (2.21b)) as well as a far-field representation (as in (2.31)). The double summation as used in (2.31) is also referred to as *far-field signature function* or simply *signature function* (Gumerov & Duraiswami, 2004). Furthermore, (2.31) proves that the signature function corresponds to what is commonly referred to as *directivity* or *directivity function* (Williams, 1999; Blackstock, 2000) and constitutes the two-dimensional equivalent of polar diagrams.

Examples of far-field signature functions are depicted in Fig. 2.6.

2.2.5 The Wavenumber Domain

The *spatial Fourier transform* $\tilde{S}(\cdot)$ of a sound field $S(\mathbf{x}, \omega)$ is defined in (B.3) and is stated here again for convenience as

$$\tilde{S}(k_x, y, z, \omega) = \int_{-\infty}^{\infty} S(\mathbf{x}, \omega) e^{ik_x x} dx \quad (2.32)$$

exemplarily for the x -dimension. The inverse operation to (2.32) is given by (B.4) in App. B. The spatial Fourier domain is also referred to as *wavenumber domain* or *k-space* (Williams, 1999).

Note that the existence of the Fourier transform of a given function $S(\mathbf{x}, \omega)$ is not explicitly proven in this thesis. A strict formalism requires showing that $S(\mathbf{x}, \omega)$ fulfills specific prerequisites (Girod *et al.*, 2001). It is implicitly assumed throughout this thesis that the latter is the case.

Due to the separability of the Cartesian coordinate system (Morse & Feshbach, 1953), the spatial Fourier transform can be applied independently along all three dimensions of space. The dependent variables of a given quantity in the space-frequency domain indicate with respect to which dimension the space-frequency domain is considered. E.g. $\tilde{S}(k_x, y, z, \omega)$ means that $S(\mathbf{x}, \omega)$ is considered in the wavenumber domain only with respect to k_x ; $\tilde{S}(k_x, k_y, z, \omega)$ means that $S(\mathbf{x}, \omega)$ is considered in the wavenumber domain with respect to k_x and k_y .

Note that the forward and inverse spatial Fourier transforms as used in this thesis ((2.32) and (B.4)) use signs in the exponent which are reversed with respect to the forward and inverse temporal Fourier transforms defined in (B.1) and (B.2) respectively. The motivation to do so is related to the propagation direction of plane waves as explained in the following.

The inverse spatial Fourier transform over a function $\tilde{S}(\mathbf{k}, \omega)$ with respect to all three spatial dimensions is given by

$$S(\mathbf{x}, \omega) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \tilde{S}(\mathbf{k}, \omega) e^{-i\mathbf{k}^T \mathbf{x}} dk_x dk_y dk_z. \quad (2.33)$$

The exponential function in (2.33) can be interpreted as the spatial component of a plane wave propagating in direction \mathbf{k} (refer to (2.5) and App. C.1). Thus, the spatial Fourier domain constitutes a plane wave representation of a sound field with respect to a three-dimensional space. The wave vector $\mathbf{k} = k \cdot [\cos \theta \sin \phi \quad \sin \theta \sin \phi \quad \cos \phi]^T$ then points into the direction of propagation of the plane wave component under

consideration. The latter is also represented by the colatitude ϕ and the azimuth θ . Using signs in the exponent of the spatial Fourier transform similar to the temporal one as e.g. in (Rabenstein *et al.*, 2006) results in the angles ϕ and θ describing the direction the plane wave is “coming from”, which is considered less elegant.

In order to illustrate the physical meaning of the wavenumber k_i , the analogies of the spatial Fourier transform (2.32) and the temporal Fourier transform defined in (B.1) are outlined below exemplarily for the x and k_x -dimensions respectively.

The frequency variable in the time Fourier transform is the *radian time frequency* ω which is related to the *time frequency* f via $\omega = 2\pi f$. In practice, the time-frequency scale (not the radian frequency scale) is used in order to refer to specific values.

The frequency variable in the spatial Fourier transform is the *wavenumber in x -direction* k_x . k_x can thus be interpreted as the *spatial radian frequency* and is of unit $\frac{\text{rad}}{\text{m}}$. Via the relation $k_x = 2\pi f_x$, a *space frequency* f_x can be established.

Note that $\lambda_x = \frac{2\pi}{k_x} = \frac{1}{f_x}$ is termed *trace wavelength* in x direction and k_x is termed *trace wavenumber* in x direction (Williams, 1999).

2.2.6 The Angular Spectrum Representation

Consider a sound field $S(\mathbf{x}, \omega)$ which is given by its spatial spectrum $\tilde{S}(k_x, y, k_z, \omega)$ at any plane $y = \text{const.}$ as

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{S}(k_x, y, k_z, \omega) e^{-i(k_x x + k_z z)} dk_x dk_z . \quad (2.34)$$

Due to the separability of the Cartesian coordinate system (Arfken & Weber, 2005), the Helmholtz equation (2.2) may be considered independently for each dimension of the Cartesian coordinate system. Inserting $\tilde{S}(k_x, y, k_z, \omega)$ into the Helmholtz equation (2.2) reformulated exclusively for the y -coordinate yields

$$\frac{\partial^2}{\partial y^2} \tilde{S}(k_x, y, k_z, \omega) + k_y^2 \tilde{S}(k_x, y, k_z, \omega) = 0 , \quad (2.35)$$

whereby

$$k_y = \sqrt{k^2 - k_x^2 - k_z^2}, \quad \forall \quad k_x^2 + k_z^2 \leq k^2 \quad (2.36a)$$

$$k_y = i\sqrt{k_x^2 + k_z^2 - k^2}, \quad \forall \quad k_x^2 + k_z^2 > k^2 . \quad (2.36b)$$

A propagating sound field is described when (2.36a) is satisfied and an evanescent sound field is described when (2.36b) is satisfied.

There are two solutions to (2.35) which are given by

$$\tilde{S}_1(k_x, y, k_z, \omega) = \check{S}_1(k_x, k_z, \omega) e^{ik_y y} \quad (2.37a)$$

$$\tilde{S}_2(k_x, y, k_z, \omega) = \check{S}_2(k_x, k_z, \omega) e^{-ik_y y} . \quad (2.37b)$$

Introducing (2.37) into (2.34) yields two expressions for $S(\mathbf{x}, \omega)$ which are given by

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \check{S}_1(k_x, k_z, \omega) e^{-i(k_x x + k_z y + k_z z)} dk_x dk_z \quad (2.38a)$$

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \check{S}_2(k_x, k_z, \omega) e^{-i(k_x x - k_z y + k_z z)} dk_x dk_z. \quad (2.38b)$$

$\check{S}_1(k_x, k_z, \omega)$ and $\check{S}_2(k_x, k_z, \omega)$ are termed the *angular spectrum representation* or *plane wave spectrum* of $S(\mathbf{x}, \omega)$ in a source-free half-space (Nieto-Vesperinas, 2006). The integral (2.38a) is convergent for $y \geq 0$ and represents $S(\mathbf{x}, \omega)$ in the case that all sound sources are located at $y < 0$. Eq. (2.38b) is convergent for $y \leq 0$ and represents $S(\mathbf{x}, \omega)$ in the case that all sound sources are located at $y > 0$.

Substituting k_x , k_z , and k_z by $k \cos \theta_{pw} \sin \phi_{pw}$, $k \sin \theta_{pw} \sin \phi_{pw}$, and $k \cos \phi_{pw}$ respectively clearly reveals the motivation for terming it *angular* representation. The angular spectrum represents the decomposition of a sound field which is specified over a given plane into a continuum of plane waves with given (complex) amplitudes and directions of propagation. For simplicity the reference plane is typically assumed to be one of the planes containing two of the coordinate axes.

In the remainder of this thesis exclusively the case that all sound sources are located at $y < 0$ will be considered. The index in the angular spectra is therefore omitted so that $\check{S}(\cdot) = \check{S}_1(\cdot)$.

Eq. (2.38) takes the form of a two-dimensional inverse Fourier transform and can thus be inverted by the forward transform as indicated in App. B. Setting then $y = 0$ yields

$$\check{S}(k_x, k_z, \omega) = \iint_{-\infty}^{\infty} S(x, 0, z, \omega) e^{i(k_x x + k_z z)} dx dz, \quad (2.39)$$

which represents the relation between the boundary value $S(x, 0, z, \omega)$ of the sound field $S(\mathbf{x}, \omega)$ at the reference plane (in this case the x - z -plane) and its angular spectrum representation $\check{S}(k_x, k_z, \omega)$.

Introducing (2.39) into (2.38a) yields

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} S(\mathbf{x}_0, \omega) \underbrace{\iint_{-\infty}^{\infty} e^{-i(k_x(x-x_0) + k_z y + k_z(z-z_0))} dk_x dk_z}_{= \mathcal{P}(\mathbf{x} - \mathbf{x}_0, \omega)} dx_0 dz_0, \quad (2.40)$$

with $\mathbf{x}_0 = [x_0 \ 0 \ z_0]^T$. $\mathcal{P}(\mathbf{x} - \mathbf{x}_0, \omega)$ is termed *wavefield propagator* (Nieto-Vesperinas, 2006).

Eq. (2.40) describes the relationship between the sound field $S(\mathbf{x}, \omega)$ at an arbitrary point \mathbf{x} in the half-space $y \geq 0$ and its boundary value $S(x, 0, z, \omega)$ at the reference plane.

An extensive literature exists regarding the theoretical limits on applicability of, and analytical solutions to, the angular spectrum decomposition. Refer to the standard literature on *Fourier optics* such as (Nieto-Vesperinas, 2006) for references.

Note finally that spherical harmonics expansion coefficients, the wavenumber domain, and the angular spectrum representation amongst others are termed *spatial spectra* or *space-frequency domains* since they describe the spatial composition of a given sound field.

2.3 Boundary Conditions

Boundary conditions are imposed on solutions to the wave equation (2.1) in order to consider the physical properties of the boundary of the domain under consideration. In *internal* or *interior* problems this domain is finite (refer to Sec. 2.3.1 and 2.3.2), in *external* or *exterior* problems it is infinite (Sec. 2.3.3).

The possible range of boundary conditions can be classified into two fundamental categories:

1. *homogeneous boundary conditions*
2. *inhomogeneous boundary conditions*

Homogeneous boundary conditions describe stationary boundaries; inhomogeneous boundary conditions describe vibrating boundaries. Problems involving mixtures of the two categories can be solved by a superposition of the two corresponding solutions and are also referred to as *mixed problems*.

The following sections give a brief overview of those boundary conditions which are important in the context of this thesis. Only the most fundamental types of boundary conditions are stated. Refer to (Gumerov & Duraiswami, 2004; Morse & Feshbach, 1953) for a detailed treatment.

2.3.1 Dirichlet Boundary Condition

Dirichlet boundary conditions affect the sound pressure. The *homogeneous Dirichlet boundary condition* is given by

$$S(\mathbf{x}, \omega) = 0 \quad \forall \quad \mathbf{x} \in \partial\Omega \quad (2.41)$$

and describes *sound-soft* (i.e. pressure-release) boundaries. It states that the sound pressure $S(\mathbf{x}, \omega)$ vanishes at the boundary $\partial\Omega$.

The *inhomogeneous Dirichlet boundary condition*

$$S(\mathbf{x}, \omega) = f_D(\mathbf{x}, \omega) \quad \forall \quad \mathbf{x} \in \partial\Omega \quad (2.42)$$

states that the sound pressure $S(\mathbf{x}, \omega)$ equals an arbitrary square integrable function $f_D(\mathbf{x}, \omega)$ at boundary $\partial\Omega$.

2.3.2 Neumann Boundary Condition

The *homogeneous Neumann boundary condition* is given by

$$\left. \frac{\partial S(\mathbf{x}, \omega)}{\partial \mathbf{n}(\mathbf{x})} \right|_{\partial\Omega} = 0 \quad (2.43)$$

and describes *sound-hard* (thus rigid) boundaries. For interior problems $\mathbf{n}(\mathbf{x})$ denotes the inward pointing surface normal on the boundary $\partial\Omega$. The operator $\frac{\partial}{\partial \mathbf{n}(\mathbf{x})}$ is termed *directional gradient* or *directional derivative* and is given by (Morse & Feshbach, 1953; Weisstein, 2002)

$$\frac{\partial}{\partial \mathbf{n}(\mathbf{x})} S(\mathbf{x}, \omega) = \langle \nabla S(\mathbf{x}, \omega), \mathbf{n} \rangle, \quad (2.44)$$

whereby the brackets $\langle \cdot \rangle$ indicate inner product (Weisstein, 2002). In the present case, the latter can also be interpreted as scalar (dot) product. The inner product of $\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right]^T$ and $\mathbf{n}(\mathbf{x}) = [n_x, n_y, n_z]^T = [\cos \alpha_n \sin \beta_n, \sin \alpha_n \sin \beta_n, \cos \beta_n]^T$ is given by

$$\langle \nabla, \mathbf{n}(\mathbf{x}) \rangle = \cos \alpha_n \sin \beta_n \frac{\partial}{\partial x} + \sin \alpha_n \sin \beta_n \frac{\partial}{\partial y} + \cos \beta_n \frac{\partial}{\partial z}. \quad (2.45)$$

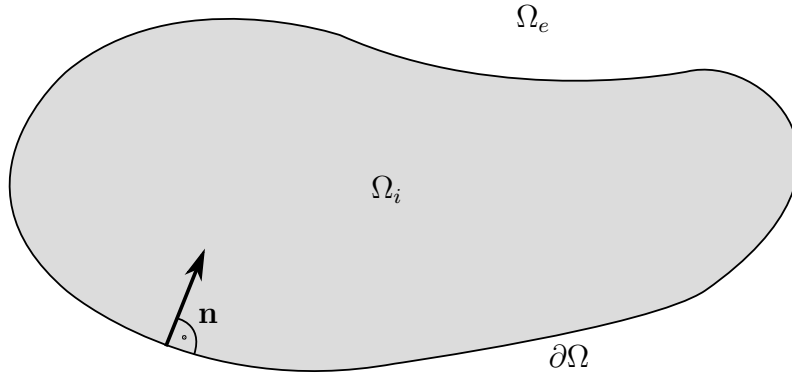


Figure 2.7: Illustration of interior domain Ω_i which is enclosed by boundary $\partial\Omega$. Ω_e is the domain exterior with respect to $\partial\Omega$. \mathbf{n} denotes the inward pointing surface normal on $\partial\Omega$.

Eq. (2.43) states that the gradient of the sound pressure in direction of the normal $\mathbf{n}(\mathbf{x})$ on the boundary pointing into the domain of interest vanishes at the boundary $\partial\Omega$. Note that the directional gradient of a pressure field is directly proportional to the particle velocity (Williams, 1999). A vanishing directional gradient of the sound pressure means also a vanishing particle velocity and thus a rigid boundary. Refer to Fig. 2.7 for an illustration of the interior example.

Finally, the *inhomogeneous Neumann boundary condition* is given by

$$\left. \frac{\partial S(\mathbf{x}, \omega)}{\partial \mathbf{n}(\mathbf{x})} \right|_{\partial\Omega} = f_N(\mathbf{x}) \Big|_{\partial\Omega} \quad (2.46)$$

and imposes an arbitrary square integrable function $f_N(\mathbf{x})$ on the directional gradient of the sound pressure $S(\mathbf{x}, \omega)$ at the boundary $\partial\Omega$.

2.3.3 Sommerfeld Radiation Condition

The *Sommerfeld radiation condition* is given by (Gumerov & Duraiswami, 2004)

$$\lim_{r \rightarrow +\infty} r \left(\frac{\partial}{\partial r} S(\mathbf{x}, \omega) + i \frac{\omega}{c} S(\mathbf{x}, \omega) \right) = 0 \quad (2.47)$$

for the definitions of the Fourier transform used in this thesis. It is employed in exterior problems and provides a boundary condition at infinity. A sound field $S(\mathbf{x}, \omega)$ satisfying (2.47) is composed of outgoing waves only. In simple words, the Sommerfeld radiation condition takes care that no energy contributions to the sound field under consideration stem from infinity.

2.4 Green's Functions

In the context of this thesis, solutions $G(\mathbf{x}|\mathbf{x}_0, \omega)$ to the inhomogeneous Helmholtz equation

$$\nabla^2 G(\mathbf{x}|\mathbf{x}_0, \omega) + k^2 G(\mathbf{x}|\mathbf{x}_0, \omega) = -\delta(\mathbf{x} - \mathbf{x}_0) \quad (2.48)$$

are termed *Green's functions* (Gumerov & Duraiswami, 2004). $\delta(\mathbf{x} - \mathbf{x}_0)$ denotes a three-dimensional Dirac delta function at position \mathbf{x}_0 with represents excitation of space at \mathbf{x}_0 . Green's functions thus describe the response of the domain of interest to a spatial Dirac excitation and thus the way sound propagates. When considered in time domain (i.e. $g(\mathbf{x}|\mathbf{x}_0, t)$), they can be interpreted as the spatial impulse response of the domain.

Under free-field conditions $G(\cdot)$ is typically denoted by $G_0(\cdot)$ and given by

$$G_0(\mathbf{x} - \mathbf{x}_0, \omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\mathbf{x} - \mathbf{x}_0|}}{|\mathbf{x} - \mathbf{x}_0|} . \quad (2.49)$$

Note that $G_0(\mathbf{x} - \mathbf{x}_0, \omega)$ is shift-invariant ($G_0(\mathbf{x} - \mathbf{x}_0, \omega)$ vs. $G_0(\mathbf{x}|\mathbf{x}_0, \omega)$) (Williams, 1999). $G_0(\mathbf{x} - \mathbf{x}_0, \omega)$ can be interpreted as the spatio-temporal transfer function of a monopole sound source located at \mathbf{x}_0 (Williams, 1999).

When $G(\mathbf{x}|\mathbf{x}_0, \omega)$ satisfies given Neumann boundary conditions, one speaks of a *Neumann Green's function* and accordingly for Dirichlet conditions.

The directional gradient $\frac{\partial G_0(\mathbf{x}, \omega)}{\partial \mathbf{e}_i}$ of $G_0(\mathbf{x}, \omega)$ in a given direction \mathbf{e}_i will also occasionally be of importance in this thesis. Exemplarily, the gradient of $G_0(\mathbf{x}, \omega)$ in x -direction is given by

$$\frac{\partial G_0(\mathbf{x}, \omega)}{\partial x} = \frac{1}{4\pi} \left(i\frac{\omega}{c} - \frac{x}{r} \right) \frac{e^{-i\frac{\omega}{c}r}}{r^2} . \quad (2.50)$$

Eq. (2.50) can be interpreted as the spatio-temporal transfer function of a dipole source whose main axis is along the x -axis (Williams, 1999). The far-field signature function of $\frac{\partial G_0(\mathbf{x}, \omega)}{\partial x}$ is similar to Fig. 2.3(b).

Since exclusively the free-field Green's function is employed in this thesis, the index 0 is omitted in the remainder.

2.5 The Kirchhoff-Helmholtz Integral

The *Kirchhoff-Helmholtz Integral* (or *Kirchhoff Integral* or *Helmholtz Integral*) is one of the essential theorems in acoustics. For interior problems it is given by (Williams,

1999)

$$a(\mathbf{x})P(\mathbf{x}, \omega) = - \oint_{\partial\Omega} \left(G(\mathbf{x}|\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}(\mathbf{x}_0)} S(\mathbf{x}, \omega) \Big|_{\mathbf{x}=\mathbf{x}_0} - S(\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}(\mathbf{x}_0)} G(\mathbf{x}|\mathbf{x}_0, \omega) \right) dA(\mathbf{x}_0) , \quad (2.51)$$

with

$$a(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega_i \\ \frac{1}{2} & \text{if } \mathbf{x} \in \partial\Omega \\ 0 & \text{if } \mathbf{x} \in \Omega_e \end{cases} .$$

$\partial\Omega$ denotes a surface enclosing the source-free volume Ω_i , $A(\mathbf{x}_0)$ an infinitesimal surface element of $\partial\Omega$, \mathbf{x}_0 a point on $\partial\Omega$; Ω_e denotes the domain outside $\partial\Omega$, $G(\mathbf{x}|\mathbf{x}_0, \omega)$ a Green's function fulfilling the given boundary conditions, and $\frac{\partial}{\partial \mathbf{n}(\mathbf{x}_0)}$ the gradient in direction of the inward pointing surface normal $\mathbf{n}(\mathbf{x}_0)$. Refer to Fig. 2.7. An according formulation of (2.51) for exterior problems exists (Williams, 1999).

The Kirchhoff-Helmholtz Integral (2.51) represents solutions to the homogeneous Helmholtz equation (2.2) with inhomogeneous boundary conditions. The sound field $P(\mathbf{x}, \omega)$ described by (2.51) equals $S(\mathbf{x}, \omega) \forall \mathbf{x} \in \Omega_i$ provided that $S(\mathbf{x}, \omega)$ is source-free in Ω_i .

The Kirchhoff-Helmholtz Integral thus states that the sound pressure $S(\mathbf{x}, \omega)$ evoked by a sound source distribution located outside an enclosing surface $\partial\Omega$ is uniquely determined inside $\partial\Omega$ by the sound pressure $S(\mathbf{x}, \omega)$ on $\partial\Omega$ and the gradient of the sound pressure in direction of the inward pointing surface normal on $\partial\Omega$. The sound field in the exterior domain Ω_e is not described by the Kirchhoff-Helmholtz Integral ($a(\mathbf{x}) = 0$ if $\mathbf{x} \in \Omega_e$). The latter can therefore not be employed for backward problems (Williams, 1999).

Under free-field conditions, i.e. when the boundary $\partial\Omega$ is *acoustically transparent*, then $G(\mathbf{x}|\mathbf{x}_0, \omega)$ is given by the free-field Green's function (2.49).

2.6 Problem Formulation

The Kirchhoff-Helmholtz Integral (2.51) presented in Sec. 2.5 provides a direct method for sound field synthesis. As mentioned in Sec. 2.4, under free-field conditions the Green's function $G(\mathbf{x} - \mathbf{x}_0, \omega)$ employed in the Kirchhoff-Helmholtz Integral can be interpreted as the spatio-temporal transfer function of a monopole sound source and its directional gradient $\frac{\partial}{\partial \mathbf{n}} G(\cdot)$ as the spatio-temporal transfer function of a dipole sound source whose main axis lies parallel to \mathbf{n} (Williams, 1999). Reinterpreted in terms of sound field synthesis, by means of an enclosing acoustically transparent continuous layer of *secondary monopole sources* and an according layer of *secondary dipole sources*, any source-free sound field can be synthesized inside this enclosing boundary. The term *secondary source* represents the fact that such a sound source is not the primary source of the auditory event which is desired to be evoked in the listener. It is rather such that an ensemble of such secondary sources

synthesizes a sound field which exhibits specific properties which in turn lead to the desired auditory event.

However, this approach to sound field synthesis requires two layers of secondary sources which is considered inconvenient. Typically, it is desired to avoid the dipole layer since it is more difficult to implement in practice. The fact that the sound field synthesized via the Kirchhoff-Helmholtz Integral is zero outside the secondary source distribution and thus that the acoustical properties of the listening room are negligible is only a theoretical benefit (Fazi & Nelson, 2007).

It is thus rather desired to employ a monopole-only formulation. In the following sections it is shown that methods exist which may be employed in order to solve the problem of sound field synthesis which avoid the necessity of secondary dipole sources. An overview over the fundamental formulations is provided and the properties of the different approaches are discussed and compared and finally the choice of the single-layer potential approach presented in Sec. 2.6.3 is justified.

2.6.1 Wave Field Synthesis

Wave Field Synthesis (WFS) (Berkhout *et al.*, 1993) is an established approach for sound field synthesis. The initial formulation considered infinitely extended planar distributions of secondary sources and was then extended to the employment of linear distributions and finally to more complex one dimensional distributions like circles and alike (Start, 1996). For didactical purposes, the following review of the fundamentals of WFS is not chronological.

Planar Secondary Source Distributions

The initial formulation of WFS is derived from Rayleigh's integral formulas, most notably Rayleigh's first integral formula (Berkhout *et al.*, 1993). The latter states that the sound field $S(\mathbf{x}, \omega)$ evoked by any arbitrary sound source distribution which is located outside a given half-space is uniquely determined by the gradient of the sound field under consideration evaluated on the planar boundary of that half-space. The gradient has to be taken in direction of the normal vector pointing into the source-free half-space.

Mathematically, Rayleigh's first integral formula is given in time-frequency domain by (Williams, 1999; Berkhout *et al.*, 1993)

$$P(\mathbf{x}, \omega) = - \int_{\partial\Omega} \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}, \omega) \Big|_{\mathbf{x}=\mathbf{x}_0} \cdot G(\mathbf{x} - \mathbf{x}_0, \omega) dA(\mathbf{x}_0), \quad (2.52)$$

\mathbf{x}_0 denotes a position on the plane $\partial\Omega$; $S(\mathbf{x}, \omega)$ denotes an arbitrary sound field which is source-free in one of the half-spaces bounded by $\partial\Omega$. The latter is referred to as *target half-space*. Refer to Fig. 2.8 for an illustration. Due to the close relationship between (2.52) and the angular spectrum representation presented in Sec. 2.2.6, the properties of both representations with respect to convergence are similar (Nieto-Vesperinas, 2006).

$\frac{\partial}{\partial \mathbf{n}}$ denotes the gradient in direction of \mathbf{n} , the unit length normal vector on the plane $\partial\Omega$ pointing into the target half-space. And finally, $P(\mathbf{x}, \omega)$ is the sound

pressure evoked by the planar monopole distribution. $P(\mathbf{x}, \omega)$ is perfectly symmetric with respect to $\partial\Omega$ and is identical to $S(\mathbf{x}, \omega)$ for all positions inside the target half-space.

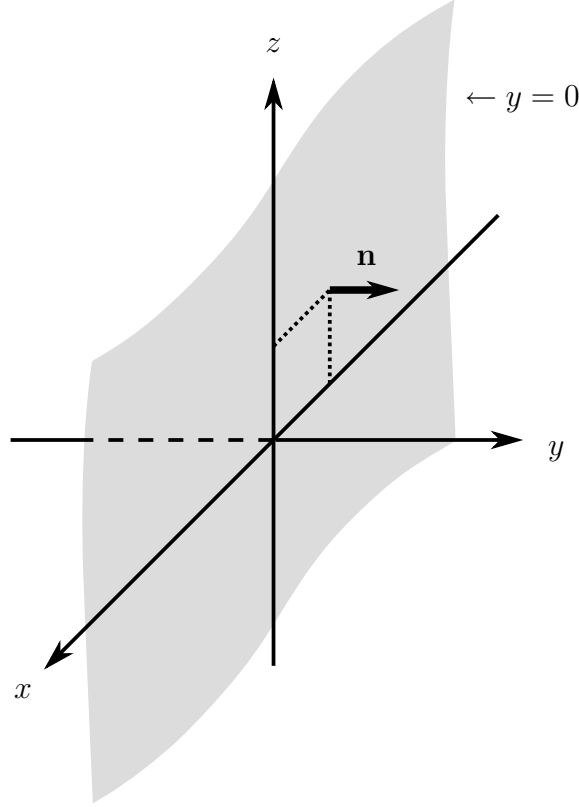


Figure 2.8: Illustration of Rayleigh's first integral formula. For convenience it is assumed that the boundary $\partial\Omega$ of the target half-space is situated along the x - z -plane. It is indicated by the grey shading and has infinite extent. The target half-space contains the positive y -axis.

Reinterpreted in terms of sound field synthesis, Rayleigh's first integral formula states that the sound field of any arbitrary virtual source distribution which is located outside of the target half-space can be perfectly synthesized by a continuous planar distribution of secondary monopole sources which are driven with the driving function (Berkhout *et al.*, 1993)

$$D(\mathbf{x}_0, \omega) = -\frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}, \omega) \Big|_{\mathbf{x}=\mathbf{x}_0}. \quad (2.53)$$

In practical implementations, loudspeakers with closed cabinets are employed which behave approximately like monopole sources when lower frequencies are considered. An analytical method compensating for deviations of the loudspeaker radiation characteristics from omnidirectionality was proposed in (de Vries, 1996). However, the latter approach constitutes an approximation due to the involved application of the stationary phase approximation (Williams, 1999).

The secondary source driving function (2.53) is only valid for planar secondary source distributions. This constitutes an essential drawback.

Arbitrary Convex Secondary Source Distributions

The extension of WFS to non-planar secondary source distributions is typically derived via the Kirchhoff-Helmholtz integral presented in Sec. 2.5 (Start, 1996; Spors *et al.*, 2008). In order to avoid the necessity of the employment of secondary dipole sources as it is required by the Kirchhoff-Helmholtz integral, it is assumed that the secondary source distribution is acoustically rigid (i.e. Neumann boundary conditions apply). The drawback is the fact that this assumption does not permit the employment of secondary monopole sources but requires that the secondary source exhibit a spatio-temporal transfer function equal to the according Neumann Green's function. The latter is dependent on the geometry under consideration. In the present context this means that the required radiation properties of the employed secondary sources are dependent on the shape of the secondary source contour. A closed form solution for this Neumann Green's function can only be found for simple geometries like cylinders and spheres (Williams, 1999). For convenience, an alternative interpretation is presented below which yields the same result.

As pointed out in (Fazi *et al.*, 2009), it is helpful to approach WFS by considering the equivalent problem of scattering of sound waves at a sound-soft object whose geometry is identical to that of the secondary source distribution. Sound-soft objects exhibit ideal pressure release boundaries, i.e. a homogeneous Dirichlet boundary condition is assumed.

When the wavelength λ of the wave field under consideration is much smaller than the dimensions of the scattering object and when the object is convex the so-called *Kirchhoff approximation* or *physical optics approximation* can be applied (Colton & Kress, 1998). The surface of the scattering object is divided into a region which is *illuminated* by the incident wave, and a *shadowed* area. The problem under consideration is then reduced to far-field scattering off the illuminated region whereby the surface of the scattering object is assumed to be locally plane. The shadowed area has to be discarded in order to avoid an unwanted *secondary diffraction* (Colton & Kress, 1998). The convexity is required in order to avoid scattering of the scattered sound field.

For such small wave lengths any arbitrary convex enclosing secondary monopole distribution may also be assumed to be locally plane. Consequently, a high-frequency approximation of the driving function for the synthesis of a given desired sound field may be derived from (2.53) when only those secondary sources are employed which are located in that region which is illuminated by the virtual sound field.

The better the assumptions of the physical optics approximation are fulfilled, most notably the wave length under consideration being significantly smaller than the dimensions of the secondary source distribution, the smaller is the resulting inaccuracy.

The illuminated area can be straightforwardly determined via geometrical considerations as indicated in Fig. 2.9 for a virtual plane wave. The area illuminated by a virtual plane wave is bounded by two lines parallel to the propagation vector \mathbf{k}_{pw} of the plane wave passing the secondary source distribution in a tangent-like manner.

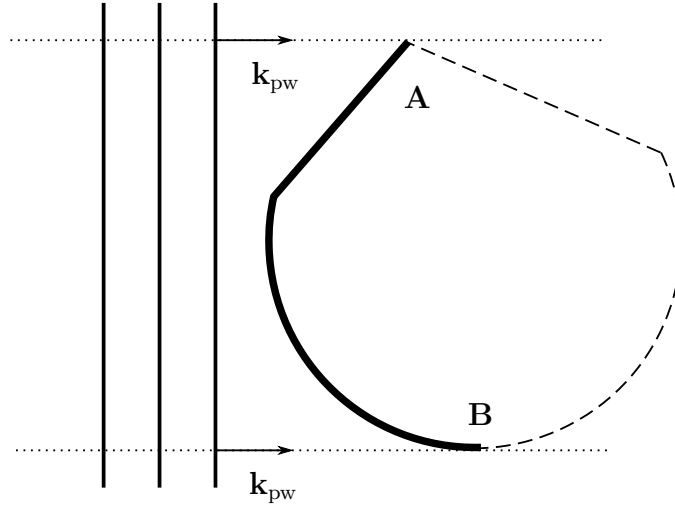


Figure 2.9: Secondary source selection for a virtual plane wave with propagation direction \mathbf{k}_{pw} . Thick solid lines indicate the area which is illuminated by the virtual plane wave. The illuminated area corresponds to the active secondary sources. The dashed line indicates the shadowed part of the secondary source distribution. The two dotted lines are parallel to \mathbf{k}_{pw} and pass the secondary source distribution in a tangent-like manner. In case **A** tapering has to be applied, in case **B** not.

The driving signal is thus approximately given by

$$D(\mathbf{x}_0, \omega) \approx -w(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}, \omega) \Big|_{\mathbf{x}=\mathbf{x}_0}, \quad (2.54)$$

whereby the window function $w(\mathbf{x}_0) = 1$ if \mathbf{x}_0 belongs to the illuminated area or $w(\mathbf{x}_0) = 0$ if \mathbf{x}_0 belongs to the shadowed area. Explicitly, $w(\mathbf{x}_0)$ for a virtual plane wave with propagation vector \mathbf{k}_{pw} is given by (Spors *et al.*, 2008)

$$w(\mathbf{x}_0) = \begin{cases} 1 & \text{if } \langle \mathbf{k}_{pw}, \mathbf{n}(\mathbf{x}_0) \rangle > 0 \\ 0 & \text{elsewhere} \end{cases}. \quad (2.55)$$

If the proper tangent on the boundary of the illuminated area is not parallel to \mathbf{k}_{pw} or is not defined (like the boundary of a planar distribution of finite size) a degenerated problem is considered (case **A** in Fig. 2.9). That means, the illuminated area is incomplete and artifacts have to be expected. The perceptual prominence of such spatial truncation artifacts can be reduced by the application of *tapering*, i.e. an attenuation of the secondary sources towards the edges of the illuminated area (Start, 1997).

It has been shown that the illuminated area does not need to be smooth. Corners are also possible with only little additional error introduced (Verheijen, 1997).

Accuracy

As mentioned above, the better the assumptions of the physical optics approximation are fulfilled, most notably the wave length under consideration being significantly

smaller than the dimensions of the secondary source distribution, the smaller is the resulting inaccuracy. This circumstance is illustrated in the following.

The sound field synthesized by a theoretical continuous spherical secondary source distribution of radius $R = 1.5$ m driven in order to synthesize a monochromatic virtual plane wave of unit amplitude and with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ is depicted in Fig. 2.10. This radius of the secondary source distribution corresponds to the wavelength of a sound wave of around 230 Hz. That means that for frequencies much higher than 230 Hz, the physical optics approximation is justified and it is expected that the error is negligible. Consequently, for frequencies of around 230 Hz and below a considerable inaccuracy has to be expected (Ahrens & Spors, 2009c).

The derivation of the sound fields depicted in Fig. 2.10 is only briefly outlined since the details are not relevant for the remainder of this thesis. The synthesized sound fields were derived in the spherical harmonics domain via (3.4) and then composed using (2.21). The spherical harmonics representation was obtained via analytical spherical harmonics transforms of the driving function and the window $w(\mathbf{x}_0)$ and using (D.9).

Fig. 2.10(a) shows the synthesized sound field for a plane wave of $f_{\text{pw}} = 200$ Hz. Indeed, some distortion of the wave front occurs especially for $-1 < y < 0$ m. For a plane wave of $f_{\text{pw}} = 1000$ Hz on the other hand no considerable error is apparent as can be seen in Fig. 2.10(b) (Ahrens & Spors, 2009c).

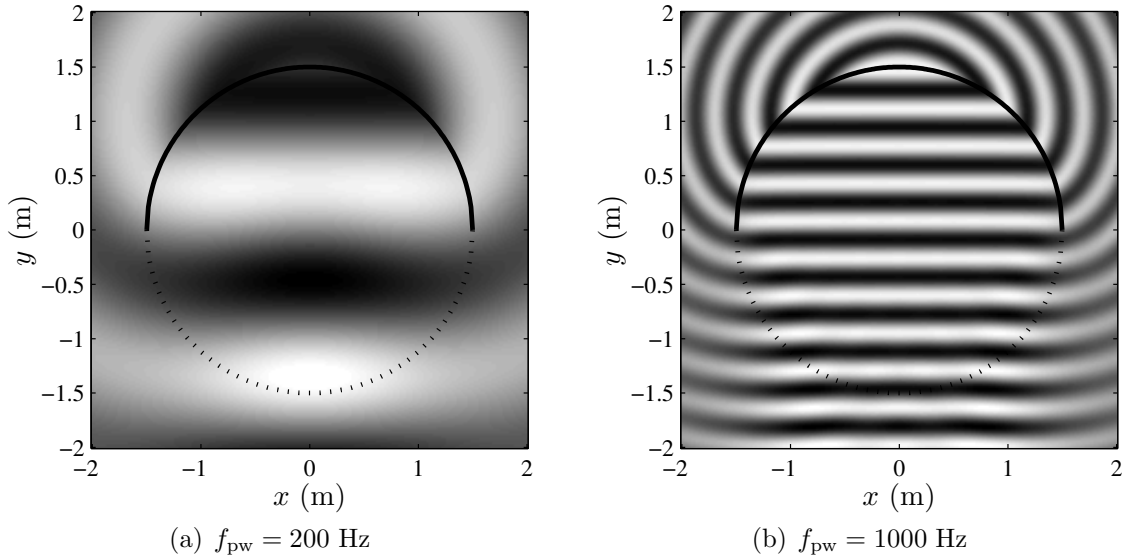


Figure 2.10: WFS of a virtual plane wave for different frequencies via a continuous spherical secondary source distribution. A cross-section through the horizontal plane. The solid line indicates the area with active secondary sources, i.e. the area which is illuminated by the virtual sound field; the dotted line indicates the shadowed area.

2^{1/2}-dimensional Synthesis

For many practical applications of WFS it is sufficient to restrict the synthesis to the horizontal plane and employing a linear distribution of secondary sources. This situation is then referred to as *2^{1/2}-dimensional synthesis* (Start, 1997). The term “2^{1/2}-dimensional” reflects the fact that the synthesis is neither purely two-dimensional nor purely three-dimensional but rather something in between. In the remainder of this subsection the 2^{1/2}-dimensional driving function will be derived from the three-dimensional one.

The WFS synthesis equation is given by (2.52) which is reformulated here as

$$P(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} D(\mathbf{x}_0, \omega) \cdot \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|} dz_0 dx_0 . \quad (2.56)$$

In order to simplify the notation it is assumed that the secondary source distribution is located in the x - z -plane, i.e. $\mathbf{x}_0 = [x_0, 0, z_0]^T$ and that synthesis in that part of the horizontal plane is targeted which contains the positive y -axis, i.e. $z = 0, y > 0$. Refer also to Fig. 2.8.

Eq. (2.56) can be approximated in the horizontal plane via the stationary phase approximation as (Berkhout *et al.*, 1993)

$$P(\mathbf{x}|_{z=0}, \omega) \approx \int_{-\infty}^{\infty} \underbrace{D(\mathbf{x}_0|_{z_0=0}, \omega) \sqrt{\frac{2\pi}{i\frac{\omega}{c}}} \sqrt{(x-x_0)^2 + y^2}}_{= D_{2.5D}(x, y, \omega)} \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|} \Big|_{z_0=0} dx_0 . \quad (2.57)$$

as outlined in detail in App. E.3. The planar secondary source distribution has thus degenerated to a linear one which is located along the x -axis. Note that (2.57) constitutes a high-frequency approximation (Williams, 1999).

Assigning all factors which arose due to the stationary phase approximation in (2.57) to the driving function $D(\mathbf{x}_0|_{z_0=0}, \omega)$ yields the 2^{1/2}-dimensional driving function $D_{2.5D}(x, y, \omega)$. However, $D_{2.5D}(x, y, \omega)$ is dependent on the listening position (x, y) . Typically, it is desired that the synthesis satisfies an extended receiver area. The driving function is therefore referenced to a given distance $d_{\text{ref}} > 0$ by setting the square root in (2.57) to $\sqrt{d_{\text{ref}}}$. Refer to Sec. 3.5.4 for an interpretation of this referencing.

The 2^{1/2}-dimensional driving function $D_{2.5D}(x, y, \omega)$ is finally given by (Berkhout *et al.*, 1993)

$$D_{2.5D}(x, y, \omega) = \sqrt{\frac{2\pi d_{\text{ref}}}{i\frac{\omega}{c}}} D(\mathbf{x}_0|_{z_0=0}, \omega) . \quad (2.58)$$

The 2^{1/2}-dimensional correction of the driving function $D(\mathbf{x}_0|_{z_0=0}, \omega)$ is equal for all secondary sources and can therefore be applied on the input signal before the latter is distributed to the individual secondary sources. Refer also to Sec. 3.3.2 and 3.5.2 for a more detailed treatment of 2^{1/2}-dimensional synthesis.

Note that (2.58) only holds for linear secondary source distributions. In order to allow for the employment of convex one-dimensional contours, the physical optics approximation presented above can be applied.

2.6.2 Simple Source Formulation and Equivalent Scattering Problem

The *simple source formulation* of the problem is obtained by constructing two equivalent but spatially disjunct problems (Williams, 1999). Besides the interior Kirchhoff-Helmholtz Integral (2.51), an equivalent exterior Kirchhoff-Helmholtz Integral is formulated with the same boundary $\partial\Omega$ but with outward pointing normal vector (Williams, 1999). It is further assumed that the sound pressure is continuous and the directional gradient is discontinuous when approaching the boundary $\partial\Omega$ from both sides. The latter assumptions represent the distribution of secondary sources on $\partial\Omega$. Additionally, the exterior sound field caused by the source distribution has to satisfy the Sommerfeld radiation condition (2.47).

Subtracting the resulting interior from the exterior problem formulation under free-field assumptions results in

$$P(\mathbf{x}, \omega) = \oint_{\partial\Omega} D(\mathbf{x}_0, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) dA(\mathbf{x}_0) , \quad (2.59)$$

whereby $D(\mathbf{x}_0, \omega)$ denotes the driving function of the secondary sources. Note that only the monopole layer is apparent in (2.59).

The continuity conditions for the pressure and its gradient on the boundary $\partial\Omega$ can be interpreted in terms of an *equivalent scattering problem* (Fazi *et al.*, 2009). Here, the secondary source distribution is replaced by a sound-soft object (i.e. Dirichlet boundaries are assumed) that scatters the impinging sound field $S(\mathbf{x}, \omega)$. Inside the boundary $\partial\Omega$, the scattered sound field $P(\mathbf{x}, \omega)$ corresponds to the impinging virtual sound field $S(\mathbf{x}, \omega)$.

The driving signal $D(\mathbf{x}_0, \omega)$ (or source strength (Williams, 1999)) is then given by

$$D(\mathbf{x}_0, \omega) = \left(\frac{\partial S_e(\mathbf{x}, \omega)}{\partial \mathbf{n}(\mathbf{x}_0)} - \frac{\partial S(\mathbf{x}, \omega)}{\partial \mathbf{n}(\mathbf{x}_0)} \right) \Big|_{\mathbf{x}=\mathbf{x}_0} . \quad (2.60)$$

$\mathbf{n}(\mathbf{x}_0)$ denotes the inward pointing surface normal and $S_e(\mathbf{x}, \omega)$ the scattered field in the exterior domain. Inside $\partial\Omega$, the synthesized sound field $P(\mathbf{x}, \omega)$ coincides with the desired sound field $S(\mathbf{x}, \omega)$.

Although the simple source approach has not received considerable attention in sound field synthesis so far - except for (Poletti, 2005) - it is of special interest since it links the well documented results from scattering theory to sound field synthesis and therefore provides interesting insights into the general problem. The drawback is the fact that an exterior field $S_e(\mathbf{x}, \omega)$ has to be constructed from the desired interior field $S(\mathbf{x}, \omega)$ in order to find the driving function $D(\mathbf{x}_0, \omega)$.

2.6.3 Potential Theory

The Kirchhoff-Helmholtz Integral (2.51) can be split into two integrals which are given by (Colton & Kress, 1998)

$$S_{\text{monopole}}(\mathbf{x}, \omega) = \oint_{\partial\Omega} D_{\text{monopole}}(\mathbf{x}_0, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) dA(\mathbf{x}_0) . \quad (2.61)$$

and

$$S_{\text{dipole}}(\mathbf{x}, \omega) = \oint_{\partial\Omega} D_{\text{dipole}}(\mathbf{x}_0, \omega) \frac{\partial G(\mathbf{x} - \mathbf{x}_0, \omega)}{\partial \mathbf{n}(\mathbf{x}_0)} dA(\mathbf{x}_0) . \quad (2.62)$$

when free-field conditions are assumed; $\mathbf{x}_0 \in \partial\Omega$. $\mathbf{n}(\mathbf{x}_0)$ denotes the inward pointing surface normal at \mathbf{x}_0 . $S_{\text{monopole}}(\mathbf{x}, \omega)$ and $S_{\text{dipole}}(\mathbf{x}, \omega)$ are termed *acoustic single-layer* and *double-layer potential* respectively and are widely used in a number of disciplines especially in the solution to scattering problems (Colton & Kress, 1998). $D_{\text{monopole}}(\mathbf{x}_0, \omega)$ and $D_{\text{dipole}}(\mathbf{x}_0, \omega)$ are termed *density* of the potentials.

The relation between a *vector field* $\mathbf{V}(\mathbf{x}, \omega)$ and its *scalar potential* $S(\mathbf{x}, \omega)$ is given by (Gumerov & Duraiswami, 2004, p. 3)

$$\mathbf{V}(\mathbf{x}, \omega) = -\nabla S(\mathbf{x}, \omega) . \quad (2.63)$$

As stated by *Euler's equation* (Williams, 1999, p. 15), $\mathbf{V}(\mathbf{x}, \omega)$, i.e. the negative sound pressure gradient in time-frequency domain, is directly proportional to the particle velocity in time-frequency domain. Sloppily speaking, the sound pressure $S(\mathbf{x}, \omega)$ is the potential of the particle velocity, thus a *harmonic velocity potential*.

The term *single-layer* used above reflects the fact that exclusively one layer of secondary monopoles is considered for the given free-field conditions. The term *double-layer* reflects the fact that the directional gradient $\frac{\partial G(\mathbf{x} - \mathbf{x}_0, \omega)}{\partial \mathbf{n}}$ of the free-field Green's function can be interpreted as a secondary dipole source which in turn can be represented by a combination of two monopoles. I.e. the double layer can be described as two single layers of monopoles (refer to Sec. 2.2.3 and 2.4).

Again, $D(\mathbf{x}_0, \omega)$ denotes the driving function of the secondary source distribution.

The double-layer potential (2.62) is inconvenient for the problem of sound field synthesis since it requires secondary dipoles (refer also to Sec. 2.6). As mentioned above, a double layer may be interpreted as a combination of single layers.

Therefore, the remainder of this section concentrates on the single-layer formulation (2.61). The index “monopole” in (2.61) is omitted for convenience.

In order to find the solution to (2.61), i.e. in order to find the appropriate driving function $D(\mathbf{x}, \omega)$ which synthesizes the desired sound field $S(\mathbf{x}, \omega)$, it is assumed at first stage that $S(\mathbf{x}, \omega)$ is considered exclusively on the boundary $\partial\Omega$, i.e. $\mathbf{x} \in \partial\Omega$ (Morse & Feshbach, 1953).

Eq. (2.61) can be interpreted as an operator \mathcal{A} acting on $D(\mathbf{x}, \omega)$ as (Morse & Feshbach, 1953)

$$(\mathcal{A}D)(\mathbf{x}, \omega) = \oint_{\partial\Omega} D(\mathbf{x}_0, \omega) G(\mathbf{x}|\mathbf{x}_0, \omega) dA(\mathbf{x}_0) . \quad (2.64)$$

\mathcal{A} is a *Fredholm operator* which is acting on a *Sobolev space* if

- its range is closed;
- its kernel is of finite dimensions;
- its cokernel is of finite dimensions.

From the equivalent scattering problem it is known that \mathcal{A} (Giroire, 1982)

- is a Fredholm operator of zero index;
- is an isomorphism if and only if ω is not an eigenvalue of the interior Dirichlet problem,

so that it can be concluded that \mathcal{A} constitutes a *compact operator*.

Such a compact operator can be expanded into a series of basis functions $\psi_n(\mathbf{x})$ as (Morse & Feshbach, 1953)

$$(\mathcal{A}D)(\mathbf{x}, \omega) = \sum_{n=1}^N \underbrace{\langle \bar{\psi}_n(\mathbf{x}), D(\mathbf{x}, \omega) \rangle}_{= \tilde{D}_n(\omega)} \tilde{G}_n(\omega) \psi_n(\mathbf{x}) \quad \forall \quad 1 \leq N \leq \infty, \quad (2.65)$$

whereby $\langle \cdot \rangle$ denotes the scalar product and $\bar{\psi}_n(\mathbf{x})$ the adjoint of $\psi_n(\mathbf{x})$. For the Green's functions considered in this thesis, $\bar{\psi}_n(\mathbf{x}) = \psi_n(\mathbf{x})^*$, whereby the asterisk $*$ denotes complex conjugation. $\tilde{G}_n(\omega)$ are the eigenvalues of \mathcal{A} and $\psi_n(\mathbf{x})$ constitutes a complete set of solutions to the wave equation which is orthogonal on $\partial\Omega$. The orthogonality relation

$$\oint_{\partial\Omega} \bar{\psi}_n(\mathbf{x}_0) \psi_m(\mathbf{x}_0) dA(\mathbf{x}_0) = a_n \delta_{nm} \quad (2.66)$$

and the completeness relation

$$\sum_{n=1}^N a_n \bar{\psi}_n(\mathbf{x}) \psi_n(\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0) \quad (2.67)$$

thus hold, whereby a_n is a normalization constant; δ_{nm} denotes the *Kronecker delta* and $\delta(\mathbf{x} - \mathbf{x}_0)$ a multidimensional *Dirac pulse*.

The projection $\tilde{D}_n(\omega)$ of the driving function $D(\mathbf{x}, \omega)$ onto the basis functions $\psi_n(\mathbf{x})$ is obtained via (Morse & Feshbach, 1953)

$$\tilde{D}_n(\omega) = \langle \bar{\psi}_n(\mathbf{x}), D(\mathbf{x}, \omega) \rangle = \oint_{\partial\Omega} D(\mathbf{x}_0, \omega) \bar{\psi}_n(\mathbf{x}_0) dA(\mathbf{x}_0), \quad (2.68)$$

so that $D(\mathbf{x}, \omega)$ can be represented by $\tilde{D}_n(\omega)$ as

$$D(\mathbf{x}, \omega) = \sum_{n=1}^N \tilde{D}_n(\omega) \psi_n(\mathbf{x}). \quad (2.69)$$

Similarly, it can be shown that the Fredholm kernel $G(\mathbf{x}, \omega)$ can be represented as (Morse & Feshbach, 1953)

$$G(\mathbf{x}, \omega) = \sum_{n=1}^N \tilde{G}_n(\omega) \psi_n(\mathbf{x}) \bar{\psi}_n(\mathbf{x}_0). \quad (2.70)$$

The solution to (2.61) is obtained by expanding all involved quantities – the desired sound field $S(\mathbf{x}, \omega)$, the driving function $D(\mathbf{x}, \omega)$, and the Green's function $G(\mathbf{x}, \omega)$ – into series of the basis functions $\psi_n(\mathbf{x})$ as

$$\begin{aligned} \sum_{n=1}^N \tilde{S}_n(\omega) \psi_n(\mathbf{x}) &= \oint_{\partial\Omega} \sum_{n=1}^N \tilde{D}_n(\omega) \psi_n(\mathbf{x}_0) \sum_{n'=1}^N \tilde{G}_{n'}(\omega) \psi_{n'}(\mathbf{x}) \bar{\psi}_{n'}(\mathbf{x}_0) dA(\mathbf{x}_0) \\ &= \sum_{n=1}^N \tilde{D}_n(\omega) \sum_{n'=1}^N \tilde{G}_{n'}(\omega) \psi_{n'}(\mathbf{x}) \oint_{\partial\Omega} \psi_n(\mathbf{x}_0) \bar{\psi}_{n'}(\mathbf{x}_0) dA(\mathbf{x}_0) \quad (2.71) \end{aligned}$$

Due to orthogonality (2.66), the last integral in (2.71) vanishes unless $n' = n$ so that

$$\sum_{n=1}^N \tilde{S}_n(\omega) \psi_n(\mathbf{x}) = \sum_{n=1}^N a_n \tilde{D}_n(\omega) \tilde{G}_n(\omega) \psi_n(\mathbf{x}) . \quad (2.72)$$

In order that (2.72) holds, all coefficients have to be equal, thus

$$\tilde{S}_n(\omega) = a_n \tilde{D}_n(\omega) \tilde{G}_n(\omega) . \quad (2.73)$$

The comparison of coefficients in (2.73) is also termed *mode-matching* since $\psi_n(\mathbf{x})$ are referred to as *modes*. Eq. (2.73) can be rearranged to be (Spors & Ahrens, 2008b)

$$\tilde{D}_n(\omega) = \frac{\tilde{S}_n(\omega)}{a_n \tilde{G}_n(\omega)} , \quad (2.74)$$

provided that $\tilde{G}_n(\omega)$ does not vanish. The driving function $D(\mathbf{x}, \omega)$ is finally obtained from (2.74) via (2.69).

Note that above reviewed procedure can also be interpreted to be a *singular value decomposition* (Fazi *et al.*, 2008b).

As stated above, (2.74) only holds on the contour $\partial\Omega$, i.e. on the secondary source contour. Since the Fredholm operator \mathcal{A} is an isomorphism as stated above, it can be concluded that the solution (2.74) holds in the entire interior domain Ω , i.e. for $\mathbf{x} \in \Omega$ (Morse & Feshbach, 1953; Giroire, 1982).

The solution (2.74) has the following fundamental properties.

- *Non-uniqueness*: At the eigenfrequencies of the interior Dirichlet problem (2.74) is non-unique. These eigenfrequencies represents resonances of the cavity under consideration. The solutions in this case are given by the null-space of operator \mathcal{A} . It is reported the non-uniqueness is not a severe problem (Copley, 1968; Giroire, 1982). Actually, it has not been reported that consequences of the non-uniqueness have been observed in practice.
- *Ill-conditioning*: Small eigenvalues $\tilde{G}_n(\omega)$ can give rise to ill-conditioning. Modes with vanishing eigenvalues can not be controlled at all. A counter-measure is regularization or discarding of problematic modes (Fazi & Nelson, 2010). As with the non-uniqueness, the practical consequences of this ill-conditioning are not clear.

2.6.4 Numerical Approaches

A number of numerical approaches such as (Kirkeby & Nelson, 1993; Ward & Abhayapala, 2001; Daniel, 2001; Poletti, 2005; Hannemann & Donohue, 2008; Kolundžija *et al.*, 2009) have been proposed for the problem of sound field synthesis. Typically, an equation system is derived either directly or in a transformed domain. The former is then solved using a given optimization criterion. Such methods are typically very flexible in terms of the secondary source layout. Typical optimization criterions are the minimization of a given error signal.

The fundamental drawbacks are firstly the fact that so far, the proposed optimization criterions are restricted to measures the relation of which to perception is unclear. Secondly, the optimization criterions are not aware of fundamental physical restrictions of the secondary source setup under consideration such as $2^{1/2}$ -dimensionality or consequences of the spatial discrete property of real-world setups (refer to Chap. 4). This circumstance leads in practice to an increased amount of regularization which has to be applied in order that the energy of the loudspeaker driving signals stays at moderate levels.

At moderately low frequencies all proposed methods have been shown to provide comparable results, see e.g. (Fazi & Nelson, 2007). The synthesis at high time frequencies where an accurate physical synthesis of the desired sound field is not possible with discrete secondary source setups has hardly been investigated apart e.g. from (Kolundžija *et al.*, 2009). The properties of different optimization criterions in such critical situations are not known and can not be predicted.

2.7 Summary

Based on the physical fundamentals outlined in Sec. 2.1–2.5 four analytical concepts for sound field synthesis were presented:

1. Direct implementation of the Kirchhoff-Helmholtz integral (Sec. 2.5),
2. Wave Field Synthesis (Sec. 2.6.1),
3. The simple source formulation (Sec. 2.6.2), and
4. The single-layer potential approach (Sec. 2.6.3).

Ad 1.: As discussed in the introduction of Sec. 2.6, direct implementation of the Kirchhoff-Helmholtz integral is inconvenient in practice.

Ad 2.: WFS generally represents a high-frequency approximation.

Ad 3.: The simple source formulation requires the construction of an exterior problem according to the interior problem under consideration.

The single-layer potential approach can therefore be identified as a promising candidate for the solution to the problem of sound field synthesis and will be elaborated with a concrete example in Sec. 3.2. One benefit of WFS over the single-layer

potential approach is the fact that WFS does not require secondary source distributions which enclose the receiver area. As will be shown in Sec. 3.3–3.5, modifications of the single-layer potential approach enable the employment of non-enclosing secondary source distributions for both three-dimensional as well as $2^{1/2}$ -dimensional synthesis.

What is common to approaches 2.–4. in theory and to all above mentioned approaches in practice is the fact that interaction of the synthesized sound field with the listening room has to be expected. This circumstance can have essential impact on perception. It is not useful to consider the acoustical environment via application of the according boundary conditions on the employed Green's function. This is due to the fact that these boundary conditions exhibit considerable time variance, e.g. persons can move inside the room, windows and doors can be opened, and temperature changes affect the speed of sound in air (Petrausch *et al.*, 2005).

For simplicity, free-field conditions are typically assumed for the synthesis and methods which actively compensate for the influence of the listening room are additionally applied. Such methods work preferably adaptively and examples are (Kirkeby *et al.*, 1998; Betlehem & Abhayapala, 2005; Lopez *et al.*, 2005; Corteel, 2006; Gauthier & Berry, 2006; Spors *et al.*, 2007). This thesis does not consider the problem of listening room compensation but focuses on the fundamental physical properties of sound field synthesis systems. Free-field conditions are assumed in the remainder for convenience.

Chapter 3

Continuous Secondary Source Distributions

3.1 Introduction

The single-layer potential solution presented in Sec. 2.6.3 requires continuous distributions of secondary sources. Such continuous distributions can not be implemented in practice with today's available loudspeaker technology but discrete setups have to be used. However, the investigation of continuous secondary source distributions gives valuable insights into the fundamental physical properties of the problem and is therefore treated in this chapter. The spatial discretization of the secondary source distribution as performed in practice is treated in Chap. 4.

The theory presented in Sec. 2.6.3 is very flexible in terms of the geometry of the secondary source contour under consideration provided that the latter simply encloses the receiver area. Although the solutions to such potentially complicated contours are mathematically well understood, the required basis functions are only available for simple geometries like spheroids and similar.

Actually, the complexity of practical implementation restricts the useful geometries to spherical secondary source distributions. These are treated in Sec. 3.2. However, geometries like circles, planes, and lines of secondary sources have also been proven to be useful in practice (de Vries, 2009). These latter geometries do not fulfill the assumptions under which the single-layer potential solution is valid as explained further below. Modifications of the single-layer potential solution provide solutions for such *imperfect* geometries whereby certain restrictions apply as investigated in detail in Sec. 3.3–3.5.

3.2 Spherical Secondary Source Distributions

When continuous spherical secondary source contours are considered all prerequisites for the application of the single-layer potential solution presented in Sec. 2.6.3 are fulfilled.

The sound field $S(\mathbf{x}, \omega)$ synthesized by a spherical secondary source contour of radius R centered around the coordinate origin is given by (Ahrens & Spors, 2008a;

Fazi *et al.*, 2009; Zotter *et al.*, 2009)

$$S(\mathbf{x}, \omega) = \int_0^{2\pi} \int_0^\pi D(\mathbf{x}_0, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) \sin \beta \, d\beta \, d\alpha, \quad (3.1)$$

with $\mathbf{x}_0 = R [\cos \alpha_0 \sin \beta_0 \quad \sin \alpha_0 \sin \beta_0 \quad \cos \beta_0]^T$. Refer to Fig. 3.1 for an illustration of the setup. Eq. (3.1) is generally also referred to as *synthesis equation*.

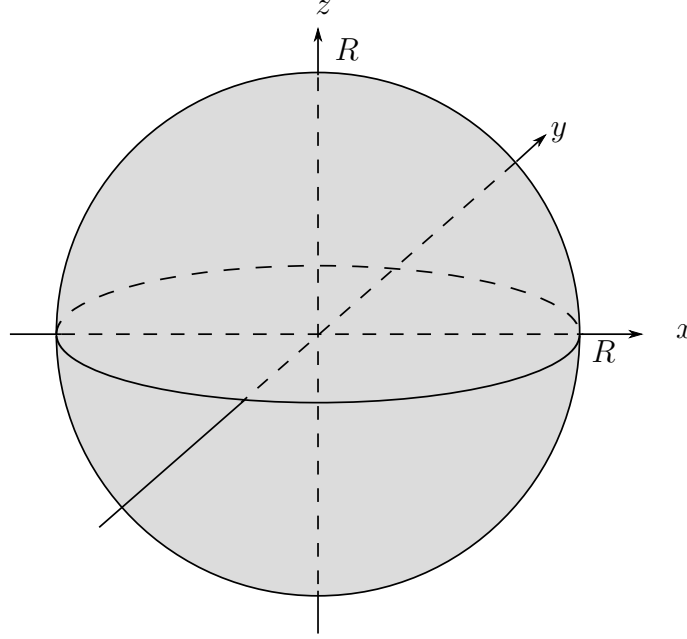


Figure 3.1: Spherical secondary source distribution of radius R centered around the coordinate origin.

Note that, for convenience, an additional restriction is introduced in (3.1) compared to (2.61). Eq. (3.1) implies that $G(\cdot)$ is shift-invariant ($G(\mathbf{x} - \mathbf{x}_0, \omega)$ vs. $G(\mathbf{x}|\mathbf{x}_0, \omega)$ in (2.61)). For the considered free-field conditions, this does not constitute a constraint. It will be of relevance for the treatment presented in Sec. 3.2.3.

Reformulating (3.1) explicitly in spherical coordinates as

$$S(\mathbf{x}, \omega) = \int_0^{2\pi} \int_0^\pi D(\alpha_0, \beta_0, \omega) G(r, R, \alpha - \alpha_0, \beta - \beta_0, \omega) \sin \beta \, d\beta \, d\alpha \quad (3.2)$$

reveals that for this particular secondary source contour, the invariance has to be apparent with respect to a rotation around the coordinate origin. Note that (3.2) may also be formulated as a rotation as it is done in (Driscoll & Healy, 1994; Rafaely, 2004).

3.2.1 Derivation of the Driving Function

Following the procedure outlined in Sec. 2.6.3 requires that $S(\mathbf{x}, \omega)$, $D(\mathbf{x}_0, \omega)$, and $G(\mathbf{x} - \mathbf{x}_0, \omega)$ are expanded into appropriate orthogonal basis functions in order to

derive a mode-matching equation similar to (2.73). For the geometry under consideration these orthogonal basis functions are given by the spherical surface harmonics presented in Sec. 2.1.3. This procedure can indeed be straightforwardly applied yielding the desired result. As will be shown in the treatment of non-enclosing secondary source contours such as circular, planar, and linear ones it is useful to derive the mode-matching equation via an alternative yet equivalent procedure as presented below (Ahrens & Spors, 2008a).

Eq. (3.2) and (3.1) can be interpreted as a convolution along the surface of a sphere as

$$S(\mathbf{x}, \omega) = D(\mathbf{x}|_{r=R}, \omega) *_{\text{sph}} G(\mathbf{x} - \mathbf{x}_0|_{\alpha_0=0, \beta_0=0}, \omega) . \quad (3.3)$$

In that case, the convolution theorem

$$\dot{S}_n^m(r, \omega) = 2\pi R \sqrt{\frac{4\pi}{2n+1}} \dot{D}_n^m(\omega) \cdot \dot{G}_n^0(r, \omega) , \quad (3.4)$$

applies (Driscoll & Healy, 1994). The convolution theorem (3.4) directly corresponds to the mode-matching equation (2.73) whereby the former facilitates the interpretation of the involved quantities.

Comparing (3.4) and (3.3) reveals the meaning of the individual quantities apparent in (3.4):

- $\dot{S}_n^m(r, \omega)$: Spherical harmonics expansion coefficients of the synthesized sound field.
- $\dot{D}_n^m(\omega)$: Spherical harmonics expansion coefficients of the driving function.
- $\dot{G}_n^0(r, \omega)$: Spherical harmonics expansion coefficients of the spatio-temporal transfer function of the secondary source positioned at $(\alpha_0 = 0, \beta_0 = 0)$, i.e. at the north pole of the secondary source distribution, expanded around the origin of the coordinate system.

The asymmetry of the convolution theorem (3.4), $\dot{S}_n^m(r, \omega)$ vs. $\dot{G}_n^0(r, \omega)$ is a consequence of the definition of (3.1) as left convolution. An according convolution theorem for right convolutions exists (Driscoll & Healy, 1994).

Rearranging (3.4) yields

$$\dot{D}_n^m(\omega) = \frac{1}{2\pi R} \sqrt{\frac{2n+1}{4\pi}} \frac{\dot{S}_n^m(r, \omega)}{\dot{G}_n^0(r, \omega)} . \quad (3.5)$$

When introducing the explicit expressions for the coefficients $\dot{S}_n^m(r, \omega)$ and $\dot{G}_n^0(r, \omega)$ given by (2.21a) into (3.5),

$$\dot{D}_n^m(\omega) = \frac{1}{2\pi R} \sqrt{\frac{2n+1}{4\pi}} \frac{\check{S}_n^m(\omega) \cdot j_n\left(\frac{\omega}{c}r\right)}{\check{G}_n^0(\omega) \cdot j_n\left(\frac{\omega}{c}r\right)} , \quad (3.6)$$

it can be seen that the parameter r appears both in the numerator as well as in the denominator in (3.6) in the Bessel function $j_n\left(\frac{\omega}{c}r\right)$. For $j_n\left(\frac{\omega}{c}r\right) \neq 0$, $j_n\left(\frac{\omega}{c}r\right)$ and thus r cancel out directly. For $\frac{\omega}{c}r = 0$, de l'Hôpital's rule (Weisstein, 2002)

can be applied to proof that $j_n(0)$ also cancels out. The driving function is thus independent from the receiver position in these cases.

However, in particular situations, i.e. when $j_n\left(\frac{\omega}{c}r\right) = 0$ and $\frac{\omega}{c}r \neq 0$, (3.6) can be undefined. In this case *forbidden frequencies* arise (Williams, 1999; Fazi & Nelson, 2010) which represent resonances of the spherical cavity. A mathematical workaround to get rid of forbidden frequencies and therefore to avoid computational instabilities in practical implementations is to reference the synthesized sound field to the center of the secondary source distribution (Williams, 1999). Then, all Bessel functions in (3.5) cancel out yielding

$$\dot{D}_n^m(\omega) = \frac{1}{2\pi R} \sqrt{\frac{2n+1}{4\pi}} \frac{\check{S}_n^m(\omega)}{\check{G}_n^0(\omega)}. \quad (3.7)$$

In order that (3.7) holds, $\check{G}_n^0(\omega)$ may not exhibit zeros which is fulfilled under free-field conditions.

The secondary source driving function $D(\alpha, \beta, \omega)$ for the synthesis of a desired sound field with expansion coefficients $\check{S}_n^m(\omega)$ is then (Ahrens & Spors, 2008a; Fazi *et al.*, 2009; Zotter *et al.*, 2009)

$$D(\alpha, \beta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\frac{1}{2\pi R} \sqrt{\frac{2n+1}{4\pi}} \frac{\check{S}_n^m(\omega)}{\check{G}_n^0(\omega)}}_{= \dot{D}_n^m(\omega)} Y_n^m(\beta, \alpha). \quad (3.8)$$

Note that, in practical applications, the summation in (3.8) can not be performed over an infinite number of addends but has to be truncated. Further discussion of a suitable choice of summation bounds is carried out in Sec. 4.3.

3.2.2 Synthesized Sound Field

Eq. (3.8) can be verified by inserting it into (3.1). After interchanging the order of integration and summation and exploitation of the orthogonality of the spherical harmonics, one arrives at

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{n,i}^m(\omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad \forall r < R, \quad (3.9)$$

which proves perfect synthesis in the interior domain. In the exterior domain, the synthesized sound field can be determined to be

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{n,i}^m(\omega) \frac{\check{G}_{n,e}^0(\omega)}{\check{G}_{n,i}^0(\omega)} h_n^{(2)}\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad \forall R < r. \quad (3.10)$$

Fig. 3.2 depicts the sound field synthesized by a continuous spherical secondary source distribution of radius $R = 1.5$ m driven in order to synthesize a virtual plane wave of unit amplitude. The according coefficients are given in explicit form in (2.26) and (2.27). Both the interior and exterior domain are shown in Fig. 3.2. Note that the sound field in the exterior domain corresponds to the scattered field in the equivalent scattering problem Sec. 2.6.2.

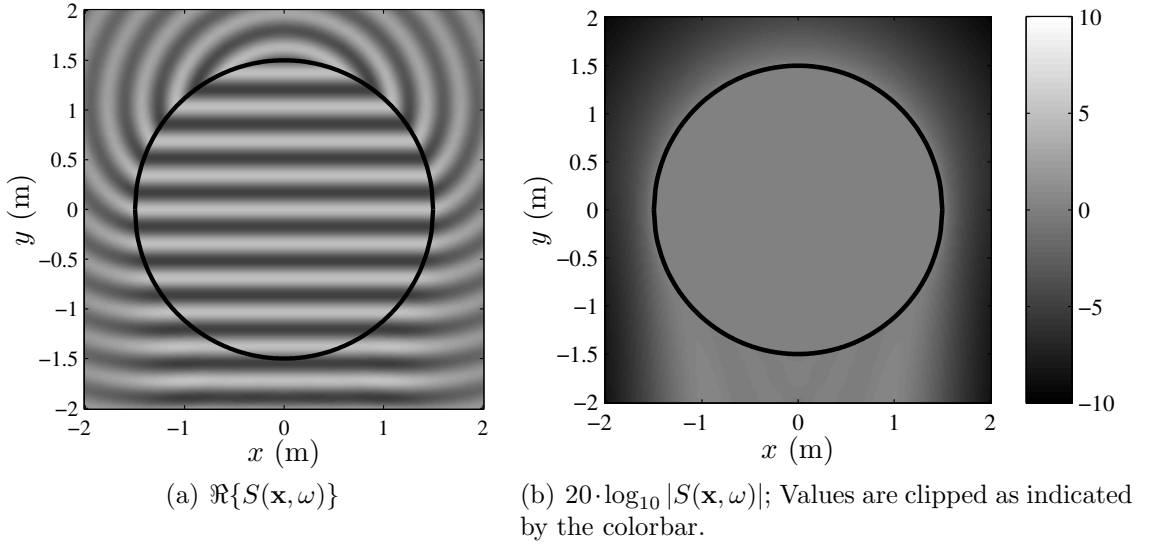


Figure 3.2: A virtual plane wave of unit amplitude and of frequency $f_{\text{pw}} = 1000$ Hz propagating into direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ synthesized by a continuous distribution of secondary monopole sources. A cross-section through the horizontal plane is shown. The black line indicates the secondary source distribution.

3.2.3 Incorporation of Secondary Sources With Complex Radiation Properties

The solutions derived in Sec. 2.6.3 and 3.2.1 assume a single layer of a harmonic sound pressure potential which can be interpreted as a layer of monopole sound sources. However, the latter are generally not available in practice when the entire audible frequency range is considered. Practical implementations rather employ loudspeakers with closed cabinets. These can indeed be assumed to be omnidirectional as long as the considered wavelength is significantly larger than the dimensions of the loudspeaker, thus at low frequencies. At higher frequencies, complex radiation patterns evolve (Fazi *et al.*, 2008a).

As mentioned in Sec. 2.2.3, sound sources of finite spatial extent can also be represented by multipoles which are combinations of monopoles located at infinitesimal distance from each other. If an appropriate combination of acoustically transparent single-layer potentials – thus a *multi-layer potential* – is assumed secondary sources with complex radiation properties can be handled as shown below. Recall that $G(\mathbf{x}, \omega)$ has to be shift-invariant in order for the derivation outlined in Sec. 3.2.1 to hold. This means that all employed secondary sources have to exhibit equal radiation characteristics and have to be orientated appropriately. Note that in a strict sense, it is not appropriate to term $G(\mathbf{x}, \omega)$ a *Green's function* when it is represented by a multipole since a multipole does not satisfy (2.48). For convenience, the symbol $G(\mathbf{x}, \omega)$ is used nevertheless.

It was mentioned in Sec. 3.2.1 that the coefficients $\check{G}_{n,i}^0(\omega)$ apparent in the driving function (3.8) represent the spatio-temporal transfer function of a secondary source which is positioned at the north pole of the sphere, thus at $\mathbf{x}_0 = [0 \ 0 \ R]^T$. The

expansion center is the origin of the coordinate system. This follows directly from the convolution theorem (3.3) or (3.4) respectively.

However, typical loudspeaker directivity measurements such as (Fazi *et al.*, 2008a) or similar yield the coefficients $\check{G}_{n',e}^{m'}(\omega)$ (see below) of an expansion of the loudspeaker's spatio-temporal transfer function around the acoustical center of the loudspeaker. The acoustical center of a loudspeaker is referred to as the position of the latter in the remainder. For convenience, it is assumed in the following that the loudspeaker under consideration is positioned at $\mathbf{x}_0 = [0 \ 0 \ R]^T$ and is orientated towards the origin of the global coordinate system.

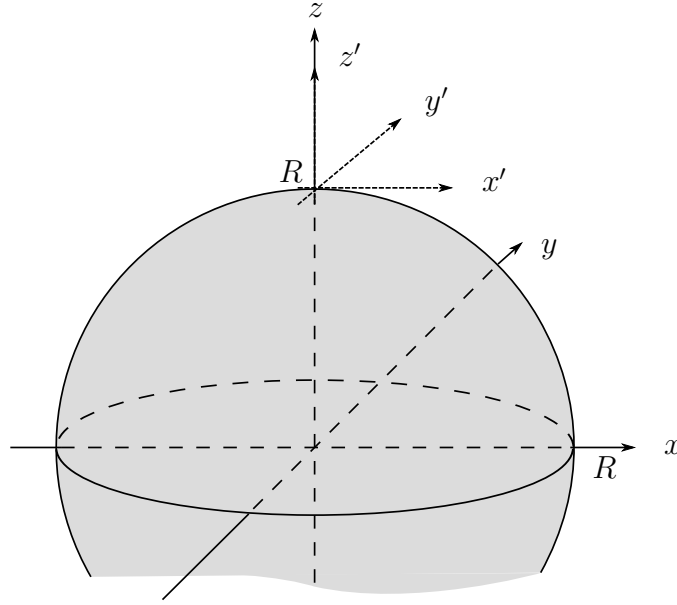


Figure 3.3: Local coordinate system with origin at position $\mathbf{x}_0 = [0 \ 0 \ R]^T$. The sphere indicates the secondary source distribution.

A local coordinate system is established with origin at \mathbf{x}_0 which can be transformed into the global coordinate system via a simple translation (refer to Fig. 3.3) (Ahrens & Spors, 2010b). Then, the spatio-temporal transfer function $G_e(\mathbf{x}', \omega)$ of the considered loudspeaker with respect to the local coordinate system can be described as

$$G_e(\mathbf{x}', \omega) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \check{G}_{n',e}^{m'}(\omega) h_{n'}^{(2)}\left(\frac{\omega}{c}r'\right) Y_{n'}^{m'}(\beta', \alpha') \quad (3.11)$$

with respect to the local coordinate system. Note that

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}) = \mathbf{x} + \Delta\mathbf{x} = \mathbf{x} - R\vec{\mathbf{e}}_z, \quad (3.12)$$

with $\Delta\mathbf{x} = [0 \ 0 \ R]$. $\vec{\mathbf{e}}_z$ denotes the unit vector pointing into positive z -direction.

The translation of the coordinate system required in order to obtain the coefficients $\check{G}_{n,i}^0(\omega)$ required by the secondary source driving function (3.8) from the directivity coefficients $\check{G}_{n',e}^{m'}(\omega)$ can be performed using the translation theorem described in (E.3) and (E.4). Doing so results in a presentation of $\check{G}_{n,i}^0(\omega)$ which is

dependent on all coefficients $\check{G}'_{n',e}{}^{m'}(\omega)$. As will be shown below, an alternative formulation provided by (Gumerov & Duraiswami, 2004) leads to a representation of $\check{G}_{n,i}^0(\omega)$ which requires only a subset of $\check{G}'_{n',e}{}^{m'}(\omega)$.

The required translation from the local coordinate system to the global one takes place coaxially in negative z -direction. This can be expressed in terms of a translation in positive z -direction as (Gumerov & Duraiswami, 2004, (3.2.54), p. 103; (3.2.86), p. 113)

$$G_i(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\sum_{n'=|m|}^{\infty} \check{G}'_{n',e}{}^m(\omega) (-1)^{n+n'} (E|I)_{nn'}^m(\Delta r, \omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha)}_{= \check{G}_{n,i}^m(\omega)} , \quad (3.13)$$

whereby $(E|I)_{nn'}^m(\Delta r, \omega)$ are termed *coaxial translation coefficients*. The notation $(E|I)$ indicates that the translation represents a change from an exterior expansion to an internal expansion. Note that m' is replaced with m in (3.13) for convenience.

From the driving function (3.8) it can be deduced that not all coefficients $\check{G}_{n,i}^m(\omega)$ are needed but only $\check{G}_{n,i}^0(\omega)$

$$\check{G}_{n,i}^0(\omega) = \sum_{n'=|m|}^{\infty} \check{G}'_{n',e}{}^0(\omega) (-1)^{n+n'} (E|I)_{nn'}^0(\Delta r, \omega) . \quad (3.14)$$

This reveals that only the subset $\check{G}'_{n',e}{}^0(\omega)$ of the secondary source directivity coefficients $\check{G}'_{n',e}{}^{m'}(\omega)$ need to be known. The former represent those modes of $G(\mathbf{x}', \omega)$ which are symmetric with respect to rotation around the vertical axis through the expansion center.

This fact further facilitates the translation significantly. The required *zonal* translation coefficients can be computed from combinations of the initial values (Gumerov & Duraiswami, 2004, (3.2.103), p. 116; (3.2.96), p. 115)

$$(E|I)_{n0}^0(\Delta r, \omega) = (-1)^n \sqrt{2n+1} h_n^{(2)}\left(\frac{\omega}{c}\Delta r\right) \quad (3.15)$$

$$(E|I)_{0n'}^0(\Delta r, \omega) = \sqrt{2n'+1} h_{n'}^{(2)}\left(\frac{\omega}{c}\Delta r\right) \quad (3.16)$$

via the recursion formula (Gumerov & Duraiswami, 2004, (3.2.90), p. 113)

$$\begin{aligned} a_{n'-1}(E|I)_{nn'-1}^0(\Delta r, \omega) - a_{n'}(E|I)_{nn'+1}^0(\Delta r, \omega) \\ = a_n(E|I)_{n+1n'}^0(\Delta r, \omega) - a_{n-1}(E|I)_{n-1n'}^0(\Delta r, \omega) , \end{aligned} \quad (3.17)$$

with (Gumerov & Duraiswami, 2004, (2.2.8), p. 67)

$$a_n = \frac{n+1}{\sqrt{(2n+1)(2n+3)}} . \quad (3.18)$$

Note that $a_{-1} = 0$.

It can be shown that the zonal translation coefficients are of the form (Ahrens & Spors, 2010b)

$$(E|I)_{nn'}^0(\Delta r, \omega) = \sum_{l'=0}^{n'} c^{l',n,n'} h_{n+2l'-n'}^{(2)}\left(\frac{\omega}{c}\Delta r\right), \quad (3.19)$$

whereby $c^{l',n,n'}$ is a real number derived from (3.15)–(3.18).

In order that the driving function (3.8) is defined neither mode $\check{G}_{n,i}^0(\omega)$ may exhibit zeros. From (3.14) it can be seen that each mode of $\check{G}_{n,i}^0(\omega)$ is given by a summation over all coefficients $\check{G}_{n',e}^0(\omega)$ multiplied by the respective translation coefficient $(E|I)_{nn'}^0(R, \omega)$. The translation coefficients $(E|I)_{nn'}^0(R, \omega)$ are linear combinations of spherical Hankel functions of the same argument but of different orders (refer to (3.19)). Spherical Hankel functions of different orders are linearly independent (Williams, 1999). Thus, since spherical Hankel functions do not exhibit zeros, a linear combination of spherical Hankel functions and therefore the translation coefficients do not exhibit zeros either. The fact of whether $\check{G}_{n,i}^0(\omega)$ vanishes or not is thus essentially dependent on the properties of the secondary source directivity coefficients $\check{G}_{n',e}^0(\omega)$.

Secondary source directivity coefficients $\check{G}_{n',e}^0(\omega)$ yielded from measurements of real loudspeakers do not per se result in a well-behaved driving function. Therefore (preferably frequency dependent) regularization has to be applied in order to yield a realizable solution. Contrary to conventional multichannel regularization, the presented approach allows for independent regularization of each mode n of the driving function. Thereby, stable modes need not be regularized while the regularization of individual unstable modes can be assumed to be favorable compared to conventional regularization of the entire filter (Ahrens & Spors, 2010b).

The fact that only the coefficients $\check{G}_{n',e}^0(\omega)$ need to be measured or modeled provides potential to facilitate the implementation of the presented approach in practice. Conventional methods of measuring $\check{G}_{n',e}^{m'}(\omega)$ such as (Fazi *et al.*, 2008a) employ spherical microphone arrays which require an enormous amount of microphones even for a moderate spatial resolution. It might be possible to exploit the rotational symmetry of the basis functions for the coefficients $\check{G}_{n',e}^0(\omega)$ in order to significantly reduce the number of microphones employed in the measurement so that e.g. an arch of microphones is sufficient instead of a full sphere.

Recall that a method enabling the employment of secondary sources with complex radiation characteristics in WFS was proposed in (de Vries, 1996). Contrary to this method the approach outlined above does not constitute an approximation but provides exact results.

Example

In order to illustrate the general properties of the presented approach a spherical distribution of highly directional secondary sources whose spatio-temporal transfer function is given by the coefficients $\check{G}_{n',e}^{m'}(\omega)$ given by (2.30) with $(\alpha'_{\text{or}}, \beta'_{\text{or}}) = (0, \pi)$ and $N' = 13$ is considered in the following. The normalized far-field signature function of $G(\cdot)$ is depicted in Fig. 3.4(a).

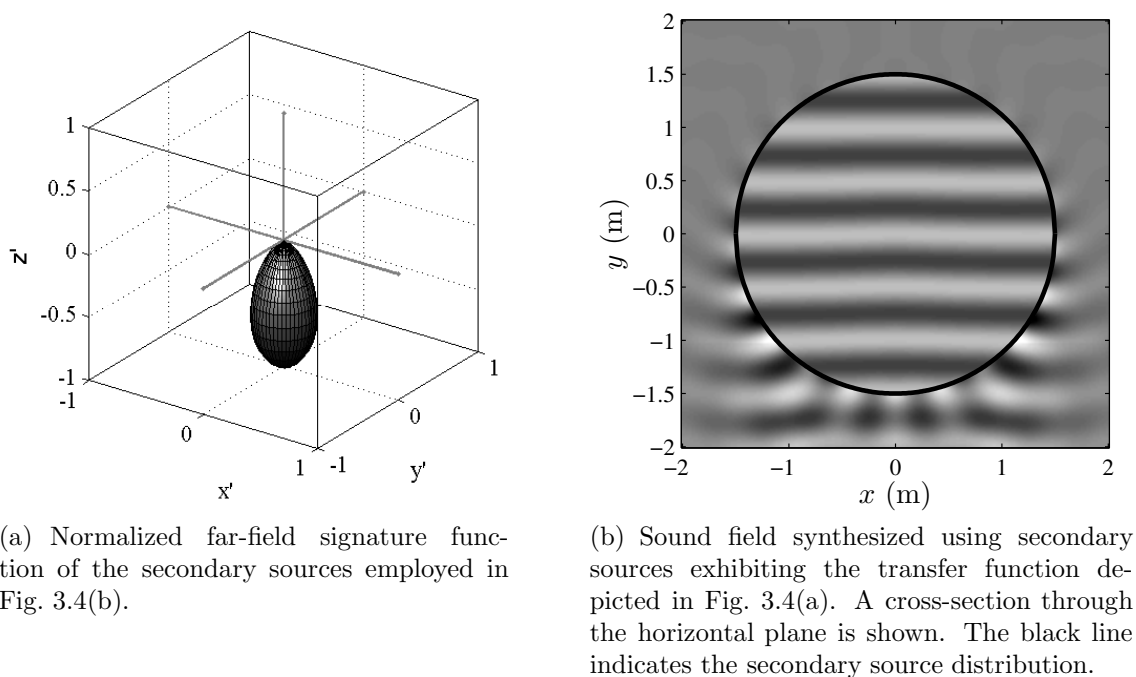


Figure 3.4: Synthesis of a virtual plane wave of unit amplitude and of frequency $f_{pw} = 700$ Hz propagating into direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ using secondary sources with complex radiation properties.

Fig. 3.4(b) depicts a continuous spherical distribution of secondary sources with a directivity as explained above synthesizing a virtual plane wave of $f_{pw} = 700$ Hz. As theoretically predicted, the virtual sound field is indeed perfectly synthesized inside the secondary source distribution. Outside the secondary source distribution the synthesized sound field is considerably different to that sound field synthesized by secondary monopoles in Fig. 3.2(a).

3.2.4 Near-Field Compensated Higher Order Ambisonics

The *Ambisonics* approach was proposed in the early 1970s in order to provide a three-dimensional listening experience, e.g. (Gerzon, 1973). Although initially motivated from a physical perspective – though without explicit physical justification – it employed a number of simplifications in order to yield simple practical solutions. The most essential simplification applied is the assumption that the secondary sources radiate plane waves which results in pure amplitude panning driving functions for simple secondary source setups (Daniel, 2001).

Practical implementations were typically realized with only a handful of loudspeakers, so that one can not speak of a physical synthesis of a sound field. The success of Ambisonics has later been attributed to psycho-acoustical properties of the human auditory system (Gerzon, 1992) and psycho-acoustic criteria have been included into the approach (Gerzon, 1980).

In (Bamford, 1995; Daniel, 2001), the traditional Ambisonics approach was then extended by systematically undoing the simplifications which were initially carried

out and increasing the resolution to *Higher Order Ambisonics* (HOA). The term *higher order* reflects the fact that the latter approach employs higher orders than first order as in the traditional approach.

Finally, the secondary sources were modeled as monopole sources which lead to the approach which is referred to as *Near-field Compensated Higher Order Ambisonics* (NFC-HOA) (Daniel, 2003) and the term *sound field synthesis* is justified here. The term *near-field* in this particular context represents the fact that the secondary sources are not assumed to be at infinite distance.

In the NFC-HOA approach, the secondary sources are typically located on the surface of a sphere. Mathematically, the involved quantities are expanded into series of spherical harmonics which allows for a mode-matching procedure which leads to an equation system that is solved for the optimal loudspeaker driving functions. These drive the loudspeakers such that their superposed sound fields best approximate the desired one in a given sense:

$$S(\mathbf{x}, \omega) = \sum_{l=0}^{L-1} D(\mathbf{x}_l, R, \omega) \cdot G(\mathbf{x} - \mathbf{x}_l, \omega) , \quad (3.20)$$

where $S(\mathbf{x}, \omega)$ denotes the desired sound field, $D(\mathbf{x}_l, R, \omega)$ the driving function of the loudspeaker located at position

$$\mathbf{x}_l = R \cdot [\cos \alpha_l \sin \beta_l \quad \sin \alpha_l \sin \beta_l \quad \cos \beta_l]^T ,$$

and $G(\mathbf{x} - \mathbf{x}_l, \omega)$ its spatio-temporal transfer function. Typically, numerical algorithms are employed to find the appropriate loudspeaker driving functions.

Modern formulations of NFC-HOA, e.g. (Fazi *et al.*, 2009; Zotter *et al.*, 2009), assume a continuous secondary source distribution and the mode-matching is solved analytically. Reformulating (3.20) in an analytical manner leads directly to (3.1). The convolution theorem (3.4) is the analog to the mode-matching which is performed in the NFC-HOA approach.

It can therefore be concluded that modern formulations of NFC-HOA and comparable constitute the single-layer potential solution to the problem of sound field synthesis employing spherical secondary source contours thus retroactively physically justifying these approaches.

From a modern perspective, the terms *Lower-resolution Ambisonics* for the conventional approach including HOA and *Higher-resolution Ambisonics* for NFC-HOA seem more appropriate since the application of the near-field compensation is a more fundamental property than the number of orders considered.

3.3 Circular Secondary Source Distributions

Sound field synthesis systems are frequently restricted to synthesis in the horizontal plane and secondary sources are arranged on a circle. In this case, the propagation direction of the synthesized sound field can only be controlled in the horizontal plane. For such a setup the free-field Green's function required by the single-layer approach presented in Sec. 2.6.3 can be interpreted as the spatial transfer function of a line source perpendicular to the target plane (Williams, 1999). This case is

treated e.g. in (Poletti, 2000; Wu & Abhayapala, 2009). A variety of such purely two-dimensional problems are treated in (Spors, 2005; Rabenstein *et al.*, 2006).

However, horizontal implementations of sound field synthesis systems usually employ loudspeakers with closed cabinets whose spatial transfer function is three-dimensional. This secondary source type mismatch prevents perfect synthesis of arbitrary source-free sound fields inside the receiver plane since the assumption of an enclosing distribution on which the single-layer potential approach bases is violated and a $2^{1/2}$ -dimensional scenario is considered. Refer also to Sec. 2.6.1.

As will be shown below, the procedure of finding the single-layer potential solution to the problem as presented in Sec. 3.2.1 leads to a useful solution which is yet imperfect as a consequence of the underlying fundamental physical limitations (Ahrens & Spors, 2008a).

For convenience, it is assumed in the following that only horizontally propagating sound fields are desired to be synthesized. The question of how such a perceptually adequate horizontal projection of a three-dimensional sound field can be obtained has not been investigated in detail so far.

3.3.1 Derivation of the Driving Function

For a circular secondary source distribution of radius R which is located inside the horizontal plane and centered around the origin of the coordinate system, the synthesis equation is given by (Ahrens & Spors, 2008a)

$$S(\mathbf{x}, \omega) = \int_0^{2\pi} D(\alpha_0, \omega) \cdot G(\mathbf{x} - \mathbf{x}_0, \omega) R d\alpha_0 \quad (3.21)$$

$$= \int_0^{2\pi} D(\alpha_0, \omega) \cdot G\left(r, R, \alpha - \alpha_0, \beta - \frac{\pi}{2}, \omega\right) R d\alpha_0, \quad (3.22)$$

with $\mathbf{x}_0 = R [\cos \alpha_0 \ \sin \alpha_0 \ 0]^T$. Refer to Fig. 3.5.

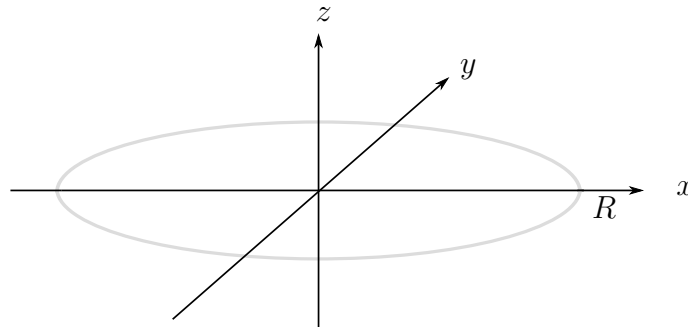


Figure 3.5: Circular secondary source distribution of radius R in the horizontal plane and centered around the coordinate origin.

Eq. (3.22) can be interpreted as a circular convolution and thus the convolution theorem (Girod *et al.*, 2001)

$$\mathring{S}_m(r, \omega) = 2\pi R \mathring{D}_m(\omega) \mathring{G}_m(r, \omega) \quad (3.23)$$

and therefore

$$\mathring{D}_m(\omega) = \frac{1}{2\pi R} \frac{\mathring{S}_m(r, \omega)}{\mathring{G}_m(r, \omega)} \quad (3.24)$$

relating the Fourier series expansion coefficients $\mathring{D}_m(\omega)$, $\mathring{S}_m(r, \omega)$, and $\mathring{G}_m(r, \omega)$ applies.

Comparing (3.23) and (3.22) reveals the meaning of the individual quantities apparent in (3.23):

- $\mathring{S}_m(r, \omega)$: Fourier-series expansion coefficients of the synthesized sound field.
- $\mathring{D}_m(\omega)$: Fourier-series expansion coefficients of the driving function.
- $\mathring{G}_m(r, \omega)$: Fourier-series expansion coefficients of the spatio-temporal transfer function of the secondary source positioned at $(\alpha_0 = 0, \beta_0 = \frac{\pi}{2})$, i.e. at $\mathbf{x}_0 = [R \ 0 \ 0]^T$.

With (3.24) and (2.24), $D(\alpha, \omega)$ can be determined to be

$$D(\alpha, \omega) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi R} \frac{\mathring{S}_m(r, \omega)}{\mathring{G}_m(r, \omega)} e^{im\alpha} . \quad (3.25)$$

Introducing the explicit formulation of the Fourier series expansion coefficients $\mathring{S}_m(r, \omega)$ and $\mathring{G}_m(r, \omega)$ given by (2.23) into (3.25) yields the explicit driving function $D(\alpha, \omega)$. Analysis of the latter reveals that unlike the case of spherical secondary source distributions treated in section Sec. 3.2.1, the radius r does not cancel out. r appears both in the numerator as well as in the denominator in the summation over n in the argument of the spherical Bessel function $j_n(\frac{\omega}{c}r)$. The driving function is therefore dependent on the receiver position. This finding has already been derived in (Ward & Abhayapala, 2001). It is thus required to reference the synthesized sound field to a specific radius which is then the only location where the synthesis is correct. For convenience, the center of the secondary source distribution ($r = 0$) is chosen.

At a first stage, setting $r = 0$ in (3.25) leads to an undefined expression of the form $\frac{0}{0}$ for $n \neq 0$ since spherical Bessel functions of argument 0 equal 0 $\forall n \neq 0$. Application of de l'Hôpital's rule (Weisstein, 2002) proves that the expression is defined for $r = 0$ and finally yields the driving function $D_{2.5D}(\alpha, \omega)$ for 2^{1/2}-dimensional synthesis as (Ahrens & Spors, 2008a)

$$D_{2.5D}(\alpha, \omega) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi R} \frac{\mathring{S}_{|m|}^m(\omega)}{\mathring{G}_{|m|}^m(\omega)} e^{im\alpha} . \quad (3.26)$$

Note that the summation over n in (2.23) reduces to a single addend with $n = |m|$ (Gumerov & Duraishwami, 2004, p. 49).

As with the driving function for spherical secondary source distributions treated in Sec. 3.2.1, the summation in (3.8) can not be performed over an infinite number of addends in practical applications. Further discussion of a suitable choice of summation bounds is carried out in Sec. 4.4.

3.3.2 Synthesized Sound Field

The sound field $S_{2.5D}(\mathbf{x}, \omega)$ synthesized by the circular secondary source distribution can be deduced from (3.23), (3.26) and (2.23) as (Ahrens & Spors, 2008a)

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{|m|,i}^m(\omega) \frac{\check{G}_{n,i}^m(\omega)}{\check{G}_{|m|,i}^m(\omega)} j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad \forall r < R, \quad (3.27)$$

and

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{S}_{|m|,i}^m(\omega) \frac{\check{G}_{n,e}^m(\omega)}{\check{G}_{|m|,i}^m(\omega)} h_n^{(2)}\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad \forall R < r. \quad (3.28)$$

Fig. 3.6 depicts the sound field synthesized by a continuous circular secondary monopole distribution with $R = 1.5$ m driven in order to synthesize a virtual plane wave of $f_{pw} = 1000$ Hz with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$.

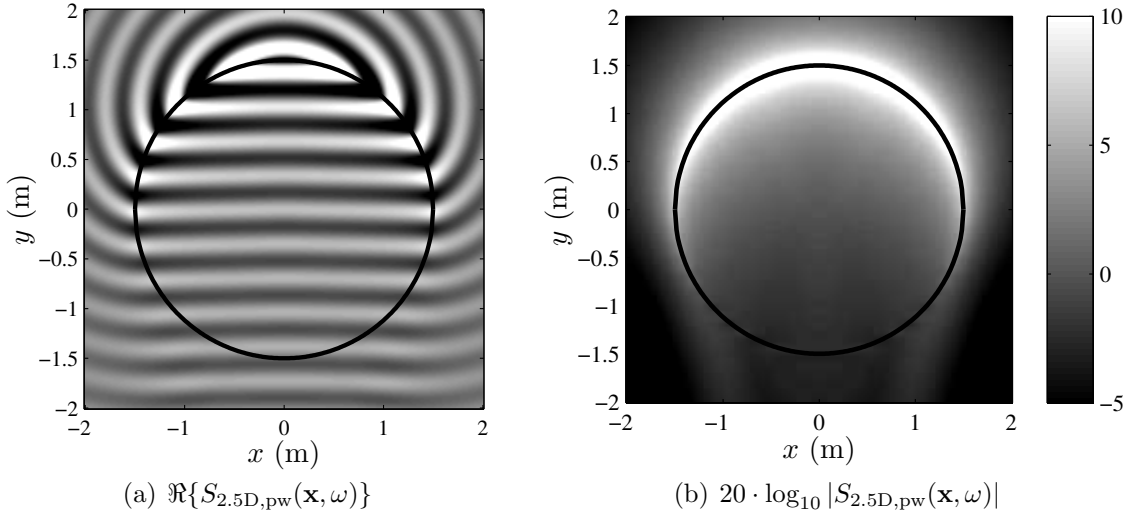


Figure 3.6: Sound pressure $S_{2.5D,pw}(\mathbf{x}, \omega)$ of a continuous circular distribution with radius $R = 1.5$ m of secondary monopole sources synthesizing a virtual plane wave of $f_{pw} = 1000$ Hz and unit amplitude with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ referenced to the coordinate origin. The secondary source distribution is indicated by the black line.

From Fig. 3.6(a) it can be seen that the wave fronts of $S(\mathbf{x}, \omega)$ in the interior domain are not perfectly straight inside the horizontal plane but slightly concave. An amplitude decay of approximately 3 dB per doubling of the distance is apparent when following the propagation path of the plane wave (Ahrens & Spors, 2008a). Fig. 3.6(b) further illustrates this amplitude decay by depicting the magnitude of the sound pressure in logarithmic scale. This inherent amplitude error is typical for $2^{1/2}$ -dimensional synthesis and is also known from WFS (Sonke *et al.*, 1998). Further investigation of the synthesized sound field reveals that subtle spectral alterations are present in the temporal broadband case. Refer also to Sec. 4.4.3 for a further analysis.

3.3.3 Incorporation of Secondary Sources With Complex Radiation Properties

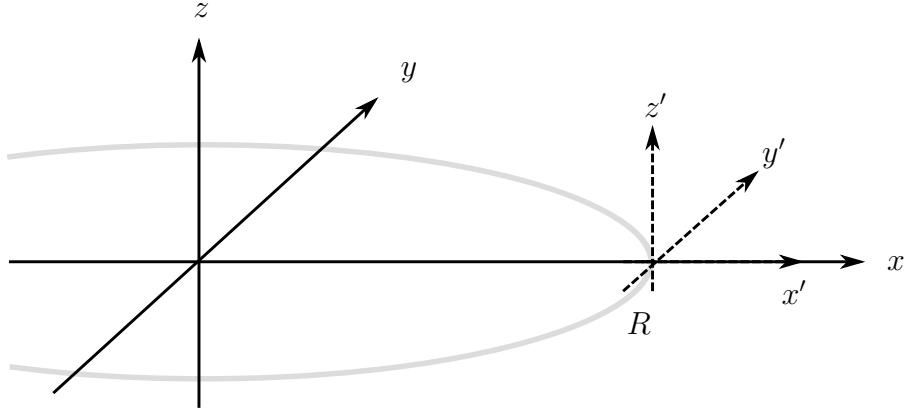


Figure 3.7: Local coordinate system with origin at $\mathbf{x}_0 = [R \ 0 \ 0]^T$. The gray line indicates the secondary source distribution.

Similarly to the case of spherical secondary source distributions treated in Sec. 3.2.3, a multi-layer potential has to be assumed if secondary sources with complex radiation characteristics have to be considered.

As outlined in Sec. 3.3.1, the coefficients $\check{G}_{|m|,i}^m(\omega)$ apparent in the driving function (3.26) describe the spatio-temporal transfer function of a secondary source which is positioned at $\mathbf{x}_0 = [R \ 0 \ 0]^T$. The expansion center is the origin of the coordinate system. This follows directly from the convolution theorem (3.23).

In order to derive the coefficients $\check{G}_{|m|,i}^m(\omega)$ apparent in the driving function in terms of the coefficients $\check{G}_{n',e}^{m'}(\omega)$ (Sec. 3.2.3) a local coordinate system with origin at \mathbf{x}_0 is established which can be transformed into the global coordinate system by a simple translation (Ahrens & Spors, 2009a). Refer to Fig. 3.7.

Then, the spatio-temporal transfer function $G(\mathbf{x}', \omega)$ of the considered loudspeaker can be described by (3.11) with respect to the local coordinate system. In this case,

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}) = \mathbf{x} + \Delta\mathbf{x} , \quad (3.29)$$

with $\Delta\mathbf{x} = [R \ 0 \ 0]$, $\Delta r = R$, $\Delta\alpha = 0$, and $\Delta\beta = \frac{\pi}{2}$.

As with spherical secondary source distributions in Sec. 3.2.3, it is beneficial to employ the formulation presented in (Gumerov & Duraiswami, 2004) for the translation of the coordinate system instead of using (E.4). In the present case, the translation from the local coordinate system to the global one takes place coaxially in negative x -direction.

As shown in App. E.1, $G(\mathbf{x}, \omega)$ can be expressed in the interior domain with

respect to the global coordinate system as

$$G(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \check{G}'_{n',e}{}^{m'}(\omega) (-1)^{n+n'} (E|I)_{nn'}^{mm'}(\Delta\mathbf{x}, \omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha)}_{= \check{G}_{n,i}^m(\omega)} . \quad (3.30)$$

From the driving function (3.26) it can be deduced that not all coefficients $\check{G}_{n,i}^m(\omega)$ are required but only $\check{G}_{|m|,i}^m(\omega)$ so that

$$\check{G}_{|m|,i}^m(\omega) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \check{G}'_{n',e}{}^{m'}(\omega) (-1)^{|m|+n'} (E|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega) . \quad (3.31)$$

This facilitates the translation because the *sectorial* translation coefficients $(E|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega)$ are easier to calculate than the *tesseral* coefficients $(E|I)_{nn'}^{mm'}(\Delta\mathbf{x}, \omega)$ (Gumerov & Duraiswami, 2004, p. 108). The symmetry relation (Gumerov & Duraiswami, 2004, eq. (3.2.49), p. 103)

$$(E|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega) = (-1)^{|m|+n'} (E|I)_{n'|m|}^{-m'-m}(\Delta\mathbf{x}, \omega) \quad (3.32)$$

can be exploited.

The sectorial translation coefficients on the right hand side of (3.32) can be computed recursively from combinations of the initial value (Gumerov & Duraiswami, 2004, eq. (3.2.5), p. 95)

$$(E|I)_{n'0}^{m'0}(\Delta\mathbf{x}, \omega) = \sqrt{4\pi} (-1)^{n'} h_{n'}^{(2)}\left(\frac{\omega}{c}\Delta r\right) Y_{n'}^{-m'}(\Delta\beta, \Delta\alpha) \quad (3.33)$$

via the recursion formulae (E.6) and (E.7) given in App. E.2.

Also required is the initial value (Gumerov & Duraiswami, 2004, eq. (3.2.51), p. 103)

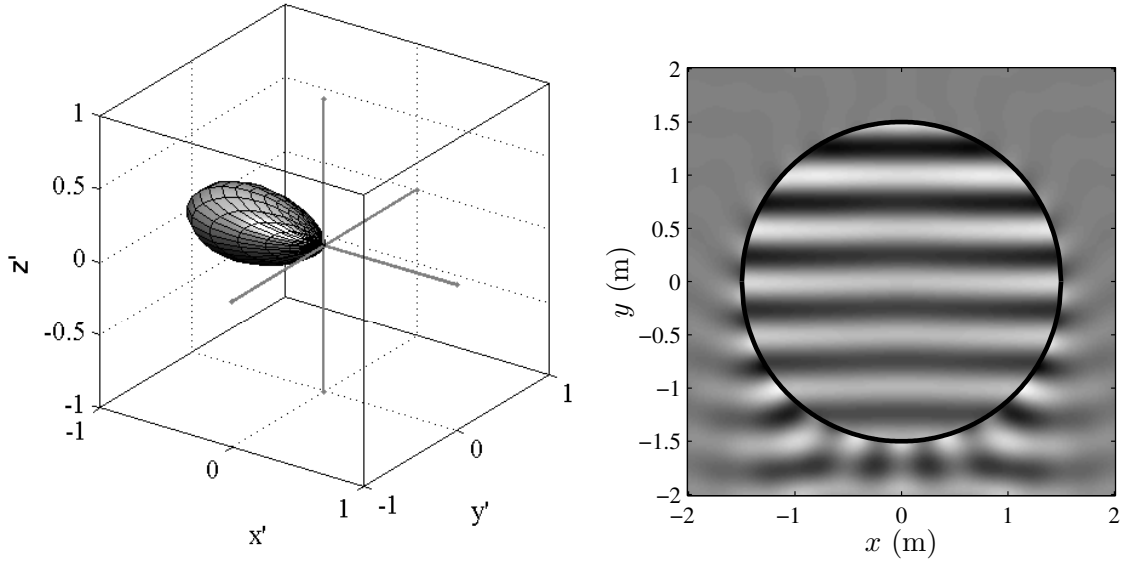
$$(E|I)_{0|m|}^{0m}(\Delta\mathbf{x}, \omega) = \sqrt{4\pi} h_{|m|}^{(2)}\left(\frac{\omega}{c}\Delta r\right) Y_{|m|}^m(\Delta\beta, \Delta\alpha) . \quad (3.34)$$

It can be shown that the sectorial translation coefficients are of the form (Ahrens & Spors, 2009a)

$$(E|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega) = \sum_{l'=0}^{|m|} c^{l',m',n',m} h_{n'-|m|+2l'}^{(2)}\left(\frac{\omega}{c}R\right) P_{n'-|m|+2l'}^{m'-m}(0) , \quad (3.35)$$

whereby $c^{l',m',n',m}$ is a real number derived from (3.33), (3.34), (E.6), (E.7), and (E.8). All factors in (3.35) are always different from zero except for $P_{n'-|m|+2l'}^{m'-m}(0)$ which exhibits zeros wherever $n' - |m| + 2l' + m' - m$ is odd (Abramowitz & Stegun, 1968). The latter is equivalent to the case of $n' + m'$ being odd. To take account for this, the summations in (3.31) are modified as

$$\check{G}_{|m|,i}^m(\omega) = \sum_{n'=0}^{\infty} \sum_{k'=0}^{n'} \check{G}'_{n',e}{}^{2k'-n'}(\omega) (-1)^{|m|+n'} (E|I)_{|m|,n'}^{m,2k'-n'}(\Delta\mathbf{x}, \omega) . \quad (3.36)$$



(a) Normalized far-field signature function of the secondary sources employed in Fig. 3.8(b). (b) Sound field synthesized using secondary sources exhibiting the transfer function depicted in Fig. 3.8(a). A cross-section through the horizontal plane is shown. The black line indicates the secondary source distribution.

Figure 3.8: Synthesis of a virtual plane wave of unit amplitude and of frequency $f_{\text{pw}} = 700$ Hz propagating into direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ using secondary sources with complex radiation properties.

This reveals that only the coefficients $\check{G}_{n',e}^{2k'-n'}(\omega)$ have to be known in order to compute the directivity filter which potentially facilitates practical measurement or modeling.

The properties of $\check{G}_{n',e}^{2k'-n'}(\omega)$ in the present case are similar to those of the coefficients apparent with spherical secondary source distributions so that the reader is referred to Sec. 3.2.3 for details (Ahrens & Spors, 2009a).

Example

In order to illustrate the presented approach a circular distribution of highly directional secondary sources whose spatio-temporal transfer function is given by the coefficients $\check{G}_{n',e}^{m'}(\omega)$ (eq. (2.30)) with $(\alpha'_{\text{or}}, \beta'_{\text{or}}) = (\pi, \frac{\pi}{2})$ and $N' = 13$. The normalized far-field signature function of $G(\cdot)$ is depicted in Fig. 3.8(a).

The translation theorems presented in App. E.1 were employed in the simulation in Fig. 3.8(b) in order to determine the coefficients $\check{S}_n^m(\omega)$ for $n \neq |m|$.

Fig. 3.8(b) depicts a continuous circular distribution of secondary sources with a directivity as explained above synthesizing a virtual plane wave of $f_{\text{pw}} = 700$ Hz. Inside the secondary source distribution, the synthesized sound field is similar to that sound field synthesized by secondary monopoles (Fig. 3.6(a)). Outside of the secondary source distribution the two sound fields differ considerably.

3.4 Planar Secondary Source Distributions

In order to find the driving function for planar secondary source distributions, the single-layer potential formulation from Sec. 2.6.3 is modified. Assume a volume enclosed by a uniform single layer. The boundary consists of a disc Ω_0 and a hemisphere Ω_{hemi} of radius r_{hemi} as depicted in Fig. 3.9 (Williams, 1999, p. 275). As $r_{\text{hemi}} \rightarrow \infty$, the disc Ω_0 turns into an infinite plane and the volume under consideration turns into a half-space. The latter is referred to as *target half-space*. Additionally, the Sommerfeld radiation condition (2.47) is invoked, i.e. it is assumed that there are no contributions to the desired sound field to be synthesized which originate from infinity so that only the planar part of the boundary needs to be considered (Ahrens & Spors, 2010c). Note that similarity of the considered scenario to the one represented by the Rayleigh integral (2.52).

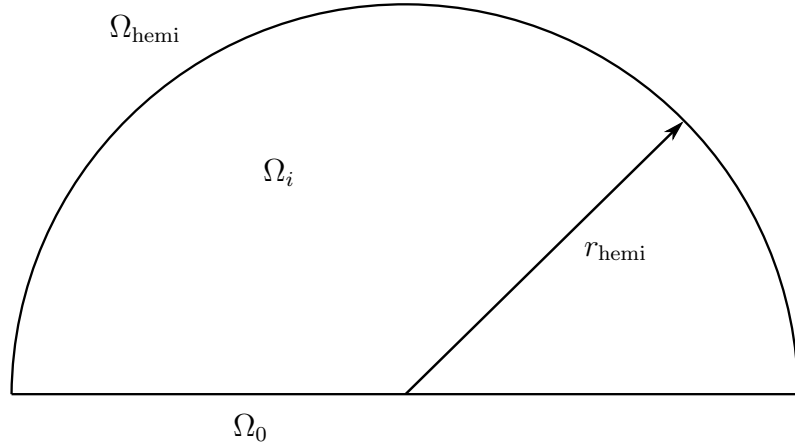


Figure 3.9: Cross-section through a boundary consisting of a hemisphere and a disc.

As a consequence, arbitrary sound fields which are source-free in the target half-space and which satisfy the Sommerfeld radiation condition may now be described by an integration over the infinite plane Ω_0 . For convenience, it is assumed in the following that the boundary of the target half-space (i.e. the secondary source distribution) is located in the x - z -plane, and the target half-space is assumed to include the positive y -axis as depicted in Fig. 3.10.

3.4.1 Derivation of the Driving Function

The synthesis equation for an infinite uniform planar secondary source distribution is given by (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$S(\mathbf{x}, \omega) = \iint_{-\infty}^{\infty} D(\mathbf{x}_0, \omega) \cdot G(\mathbf{x} - \mathbf{x}_0, \omega) dx_0 dz_0 . \quad (3.37)$$

with $\mathbf{x}_0 = [x_0 \ 0 \ z_0]^T$. Here, (3.37) implies that $G(\cdot)$ is invariant with respect to translation along the secondary source contour (Williams, 1999). Note the resemblance of (3.37) to the Rayleigh integral (2.52) (Berkhout, 1987; Williams, 1999).

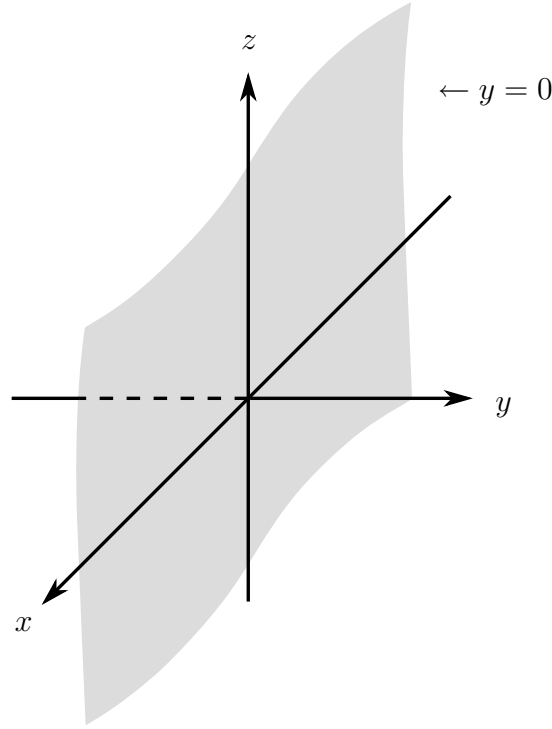


Figure 3.10: Illustration of the setup of a planar secondary source situated along the x - z -plane. The secondary source distribution is indicated by the grey shading and has infinite extent. The target half-space is the half-space bounded by the secondary source distribution and containing the positive y -axis.

Eq. (3.37) essentially constitutes a two-dimensional convolution along the spatial dimensions x and z respectively. This fact is revealed when (3.37) is rewritten as (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$\begin{aligned}
 S(\mathbf{x}, \omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D([x_0 \ 0 \ z_0]^T, \omega) \ G([x \ y \ z]^T - [x_0 \ 0 \ z_0]^T, \omega) \ dx_0 dz_0 \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(x_0, 0, z_0, \omega) \ G(x - x_0, y, z - z_0, \omega) \ dx_0 dz_0 \\
 &= D(\mathbf{x}|_{y=0}, \omega) *_{x,z} G(\mathbf{x}, \omega), \tag{3.38}
 \end{aligned}$$

where the asterisk $*_i$ denotes convolution with respect to the indexed spatial dimension (Girod *et al.*, 2001). Thus, the convolution theorem

$$\tilde{S}(k_x, y, k_z, \omega) = \tilde{D}(k_x, k_z, \omega) \cdot \tilde{G}(k_x, y, k_z, \omega) \tag{3.39}$$

holds (Girod *et al.*, 2001).

The secondary source driving function $\tilde{D}(k_x, k_z, \omega)$ in wavenumber domain is given by

$$\tilde{D}(k_x, k_z, \omega) = \frac{\tilde{S}(k_x, y, k_z, \omega)}{\tilde{G}(k_x, y, k_z, \omega)}, \tag{3.40}$$

and in time-frequency domain by (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$D(x, z, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{\tilde{S}(k_x, y, k_z, \omega)}{\tilde{G}(k_x, y, k_z, \omega)} e^{-i(k_x x + k_z z)} dk_x dk_z. \quad (3.41)$$

In order that $\tilde{D}(k_x, k_z, \omega)$ and $D(x, z, \omega)$ are defined $\tilde{G}(k_x, y, k_z, \omega)$ may not exhibit zeros.

Note that $\tilde{G}(k_x, y, k_z, \omega)$ is the spatial spectrum of the secondary source located at the coordinate origin. This follows directly from (3.38) and (3.39). The incorporation of measured or modeled complex secondary source transfer functions is straightforward and does not require a translation of the coordinate system as it was the case for spherical and circular secondary source contours (Sec. 3.2.3 and 3.3.3, respectively).

Eq. (3.41) suggests that $D(x, z, \omega)$ is dependent on the distance y of the receiver to the secondary source distribution since y is apparent on the right hand side of (3.41). It will be shown in Sec. 3.4.3 that under certain circumstances, y does indeed cancel out making $D(x, z, \omega)$ independent from the location of the receiver.

3.4.2 Physical Interpretation

Integrals like (3.37) are termed *Fredholm integrals* (Morse & Feshbach, 1953). One remarkable fundamental property of Fredholm integrals is the fact that – contrary to the *Fredholm operators* which represent enclosing secondary source distributions as discussed in Sec. 2.6.3 – the solution always exists (Morse & Feshbach, 1953). Practically speaking, resonances in the target volume can not evolve.

Since the physical fundamentals given by the Fredholm integral theory suggest that any sound field satisfying the Sommerfeld radiation condition (Sec. 2.3.3) can be perfectly synthesized by a continuous planar distribution of appropriate secondary sources, the driving function (3.41) must be independent of y . It is therefore justified to set $y = 0$ in (3.41) which yields

$$D(x, z, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \underbrace{\frac{\tilde{S}(k_x, 0, k_z, \omega)}{\tilde{G}(k_x, 0, k_z, \omega)}}_{= \tilde{D}(k_x, k_z, \omega)} e^{-i(k_x x + k_z z)} dk_x dk_z. \quad (3.42)$$

Comparison of (3.42) and (2.38a) reveals that the ratio of $\tilde{S}(k_x, 0, k_z, \omega)$ and $\tilde{G}(k_x, 0, k_z, \omega)$ constitutes the angular spectrum representation $\tilde{D}(k_x, k_z, \omega)$ of the driving function $D(x, z, \omega)$.

Comparison of (3.37) and (2.40) suggests that the driving function $D(x, z, \omega)$ acts as a representation of $S(\mathbf{x}|_{x=0, z=0}, \omega)$ on the secondary source contour. The relation of $S(\mathbf{x}|_{x=0, z=0}, \omega)$ to the sound field $S(\mathbf{x}, \omega)$ in the target half-space is represented by the Green's function $G(\mathbf{x} - \mathbf{x}_0, \omega)$ which serves as propagator.

Finally, the angular spectrum representation outlined in Sec. 2.2.6 can significantly facilitate the derivation of driving functions. As mentioned in Sec. 2.2.6, the angular spectrum representation constitutes a representation of a given sound in

terms of a continuum of plane waves with respect to a reference plane. That means if the reference plane coincides with the secondary source distribution and the according angular representation of the sound field under consideration is known, then the driving function for synthesis of this sound field can be obtained from the driving function of a plane wave. Plane waves have to be combined according to the sound field's angular representation. An explicit solution to (3.37) for the sound field under consideration is therefore not required.

Refer to Sec. 3.4.3 for the calculation of the plane wave driving function.

3.4.3 Synthesized Sound Field And Example Driving Function

The sound field synthesized by a continuous planar secondary monopole distribution driven according to (3.41) is yielded by inserting (3.41) into (3.37). To solve the integrals one has to substitute $u = x_0 - x$ and $v = z_0 - z$ and follow the procedure outlined in appendix C.2. One arrives then at (C.2) proofing perfect synthesis in the target half-space (Ahrens & Spors, 2010d).

In the remainder of this section, the derivation of the driving function for a sample plane wave of given propagation direction to be synthesized by a continuous planar distribution of secondary point sources is demonstrated.

The explicit expressions for $\tilde{S}(k_x, y, k_z, \omega)$ and $\tilde{G}(k_x, y, k_z, \omega)$ are derived in the appendices and are given by (C.4) and (C.9). Due to the constrained validity of the involved transformations, the following equations are only valid for $k_{pw,y} > 0$ (refer to App. C), i.e. for plane waves propagating into the target half-space.

Inserting (C.4) and (C.9) into (3.40) and exploiting the sifting property of the delta function (Girod *et al.*, 2001) yields

$$\tilde{D}(k_x, k_z, \omega) = 8\pi^2 j k_{pw,y} \cdot \delta(k_x - k_{pw,x}) \delta(k_z - k_{pw,z}) 2\pi \delta(\omega - \omega_{pw}) . \quad (3.43)$$

Note that $\tilde{D}(k_x, k_z, \omega)$ is indeed independent from y under the given assumptions.

Finally, the driving function is given by

$$D(x, z, \omega) = 2j k_{pw,y} e^{-jk_{pw,x}x} e^{-jk_{pw,z}z} 2\pi \delta(\omega - \omega_{pw}) . \quad (3.44)$$

Transferred to the time domain and formulated for broadband signals using (Girod *et al.*, 2001), (3.44) reads (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$d(x, z, t) = \frac{2}{c} \sin \theta_{pw} \sin \phi_{pw} \frac{\partial}{\partial t} \hat{s} \left(t - \frac{x}{c} \cos \theta_{pw} \sin \phi_{pw} - \frac{z}{c} \cos \phi_{pw} \right) , \quad (3.45)$$

where $\hat{s}(t)$ denotes the time domain signal that the plane wave carries. Thus, the driving signal for a secondary source at a given location is yielded by differentiating the time domain input signal with respect to time and weighting and delaying it. The differentiation and the weight are independent from the position of the secondary sources and can therefore be performed on the input signal. The delay is dependent both on the propagation direction of the desired plane wave as well as on the position of the secondary source. It therefore has to be performed individually for each secondary source. This constitutes a computationally efficient implementation

scheme compared to the numerical approaches in (Ward & Abhayapala, 2001; Hanemann & Donohue, 2008; Kirkeby & Nelson, 1993). The implementation scheme of the presented approach is similar to that of WFS (Spors *et al.*, 2008) (refer also to Sec. 2.6.1).

Finally, note that the temporal differentiation in (3.45) compensates for the spatial integration taking place in (3.37).

3.5 Linear Secondary Source Distributions

Despite the simple driving function for the planar secondary source array, this setup will be rarely implemented due to the enormous amount of loudspeakers necessary. Typically, audio presentation systems employ linear arrays or a combination thereof. For convenience, the secondary source array is assumed to be along the x -axis (thus $\mathbf{x}_0 = [x_0 \ 0 \ 0]^T$, refer to Fig. 3.11).

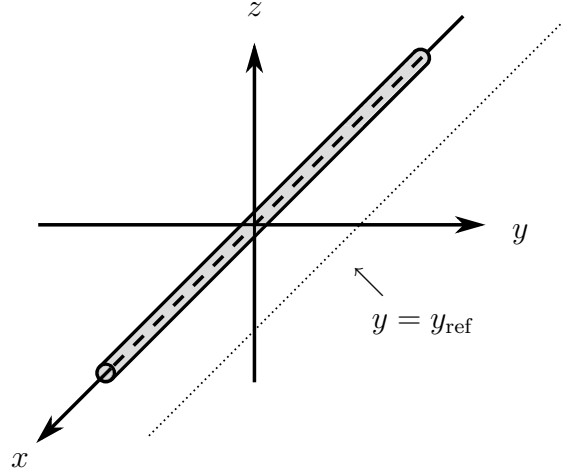


Figure 3.11: Illustration of the setup of a linear secondary source situated along the x -axis. The secondary source distribution is indicated by the grey shading and has infinite extent. The target half-plane is the half-plane bounded by the secondary source distribution and containing the positive y -axis. Thin dotted lines indicate the reference line (see text).

3.5.1 Derivation of the Driving Function

For above described setup the synthesis equation is given by (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$S(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} D(\mathbf{x}_0, \omega) \cdot G(\mathbf{x} - \mathbf{x}_0, \omega) dx_0 . \quad (3.46)$$

Similarly to (3.37), (3.46) can be interpreted as a convolution along the x -axis (Berkhout, 1987; Verheijen, 1997; Girod *et al.*, 2001) and the convolution

theorem

$$\tilde{S}(k_x, y, z, \omega) = \tilde{D}(k_x, \omega) \cdot \tilde{G}(k_x, y, z, \omega) \quad (3.47)$$

holds. The secondary source driving function in wavenumber domain is thus given by

$$\tilde{D}(k_x, \omega) = \frac{\tilde{S}(k_x, y, z, \omega)}{\tilde{G}(k_x, y, z, \omega)}, \quad (3.48)$$

and in temporal spectrum domain by (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$D(x, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{S}(k_x, y, z, \omega)}{\tilde{G}(k_x, y, z, \omega)} e^{-jk_x x} dk_x. \quad (3.49)$$

Again, $\tilde{G}(k_x, y, z, \omega)$ may not exhibit zeros.

3.5.2 Synthesized Sound Field And Example Driving Function

In the following, the synthesis of a virtual plane wave of unit amplitude and given propagation direction is considered. $\tilde{S}(k_x, y, z, \omega)$ and $\tilde{G}(k_x, y, z, \omega)$ for a plane wave and secondary monopole sources are given by (C.3) and (C.8).

Inserting (C.3) and (C.8) into (3.49) and applying the sifting property of the Dirac delta function yields (Ahrens & Spors, 2010d)

$$\tilde{D}(k_x, \omega) = \frac{2\pi\delta(k_x - k_{pw,x})e^{-jk_{pw,y}|y|}e^{-jk_{pw,z}z}}{-\frac{j}{4}H_0^{(2)}\left(\sqrt{\left(\frac{\omega_{pw}}{c}\right)^2 - k_{pw,x}^2}\sqrt{y^2 + z^2}\right)} 2\pi\delta(\omega - \omega_{pw}). \quad (3.50)$$

$H_0^{(2)}(\cdot)$ denotes the cylindrical Hankel function of second kind (Williams, 1999). Note that y and z are apparent in the expression for the driving function (3.50) suggesting that (3.46) can only be satisfied for positions on the surface of a cylinder determined by $d = \sqrt{y^2 + z^2}$.

However, with such a linear secondary source distribution, the k_x , k_y and k_z components of the synthesized sound field can not be controlled individually (Williams, 1999). The secondary source distribution radiates conical wave fronts which have only one degree of freedom. The term $\left(\frac{\omega_{pw}}{c}\right)^2 - k_{pw,x}^2$ in (3.50) is constant for a given radian frequency ω_{pw} and given $k_{pw,x}$ and the relations

$$\left(\frac{\omega_{pw}}{c}\right)^2 - k_{pw,x}^2 = k_{pw,y}^2 + k_{pw,z}^2 \quad (3.51)$$

$$= \underbrace{k_{pw}^2(\sin^2 \theta_{pw} \sin^2 \phi_{pw} + \cos^2 \phi_{pw})}_{= k_{pw,\rho}^2} \quad (3.52)$$

$$= \text{const}$$

hold due to the dispersion relation (2.6). In order to illustrate (3.51) and (3.52) the problem is reformulated in cylindrical coordinates. It is assumed that the linear

axis of the coordinate system coincides with the secondary source distribution. $k_{\text{pw},\rho}$ denotes the radial wavenumber.

Relation (3.52) states that the radial wavenumber $k_{\text{pw},\rho}$ is solely dependent on the time frequency and the $k_{\text{pw},x}$ component of the virtual plane wave. For a given azimuth θ_{pw} of the propagation direction of the desired virtual plane wave, the zenith angle ϕ_{pw} is determined by relations (3.51) and (3.52) and vice versa.

In other words, when a correct propagation direction of the synthesized virtual plane wave is desired, (3.46) can only be satisfied for receiver positions on a straight line parallel to the secondary source distribution (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d). In spherical coordinates, this receiver line is determined by $d = \sqrt{y^2 + z^2}$ and $(\alpha = \theta_{\text{pw}}, \beta = \phi_{\text{pw}})$. This finding is in analogy to the synthesis of a plane wave by a circular arrangement of secondary point sources where the synthesized sound field has to be referenced to a point (refer to Sec. 3.3.1). As a consequence, a correct propagation direction of the synthesized sound field can only be achieved inside a target half-plane containing the secondary source distribution and the reference line.

The horizontal half-plane containing the positive y -axis is chosen as target half-plane, thus $y > 0, z = 0$. Consequently, also the propagation directions of the desired plane wave have to be restricted to the horizontal plane ($\phi_{\text{pw}} = \frac{\pi}{2}$ or $k_{\text{pw},z} = 0$). Furthermore, y in (3.50) is set to the reference distance $y_{\text{ref}} > 0$ (Fig. 3.11). As mentioned in Sec. 3.3, this type of synthesis is typically referred to as 2^{1/2}-dimensional synthesis.

With above mentioned referencing, (3.50) simplifies to

$$\tilde{D}(k_x, \omega) = \frac{4j \cdot e^{-jk_{\text{pw},y}y_{\text{ref}}}}{H_0^{(2)}(k_{\text{pw},y}y_{\text{ref}})} \cdot 2\pi\delta(k_x - k_{\text{pw},x}) 2\pi\delta(\omega - \omega_{\text{pw}}), \quad (3.53)$$

and finally

$$D(x, \omega) = \frac{4j \cdot e^{-jk_{\text{pw},y}y_{\text{ref}}}}{H_0^{(2)}(k_{\text{pw},y}y_{\text{ref}})} \cdot e^{-jk_{\text{pw},x}x} 2\pi\delta(\omega - \omega_{\text{pw}}). \quad (3.54)$$

Transferred to the time domain and formulated for broadband signals, (3.54) reads (Ahrens & Spors, 2010c; Ahrens & Spors, 2010d)

$$d(x, t) = f(t) *_t \hat{s} \left(t - \frac{x}{c} \cos \theta_{\text{pw}} \sin \phi_{\text{pw}} - \frac{y_{\text{ref}}}{c} \sin \theta_{\text{pw}} \sin \phi_{\text{pw}} \right). \quad (3.55)$$

$f(t)$ denotes a filter with frequency response

$$F(\omega) = \frac{4j}{H_0^{(2)}(k_{\text{pw},y}y_{\text{ref}})}, \quad (3.56)$$

the asterisk $*_t$ denotes convolution with respect to time, and $\hat{s}(t)$ the time domain signal that the plane wave carries. Thus, the time domain driving signal for a secondary source at a given location is yielded by applying a delay and a filter on the time domain input signal. The transfer function $F(\omega)$ of the filter has high pass characteristics with a slope of approximately 3 dB per octave.

$F(\omega)$ is exclusively dependent on the propagation direction of the desired plane wave and on the amplitude reference distance y_{ref} . It is therefore equal for all

secondary sources and it is sufficient to perform the filtering only once on the input signal before distributing the signal to the secondary sources. The delay is dependent both on the propagation direction of the desired plane wave and on the position of the secondary source. It therefore has to be performed individually for each secondary source.

As with planar secondary source distributions, this constitutes a computationally efficient implementation scheme compared to the numerical approaches in (Ward & Abhayapala, 2001; Hannemann & Donohue, 2008; Kirkeby & Nelson, 1993).

Inserting (3.54) into (3.46) yields the sound field synthesized by a continuous linear secondary monopole source distribution driven to synthesize the sample plane wave. Solving the integral as indicated in Sec. 3.4.3 yields

$$S(\mathbf{x}, \omega) = \frac{e^{-jk_{\text{pw},y}y_{\text{ref}}}}{H_0^{(2)}(k_{\text{pw},y}y_{\text{ref}})} e^{-jk_{\text{pw},x}x} H_0^{(2)}\left(k_{\text{pw},y}\sqrt{y^2 + z^2}\right). \quad (3.57)$$

For $y = y_{\text{ref}}$ and $z = 0$ equation (3.57) exactly corresponds to the desired sound field. However, for $y \neq y_{\text{ref}}$ or $z \neq 0$ the synthesized sound field differs from the desired one. The arising artifacts are easily identified when the far-field/high-frequency region is considered ($k_{\text{pw},y}y_{\text{ref}} \gg 1$, $k_{\text{pw},y}\sqrt{|y|^2 + z^2} \gg 1$).

There, the Hankel functions apparent in (3.57) can be replaced by their large argument approximation $H_n^{(2)}(z) = \sqrt{\frac{2}{\pi z}} e^{-j(z - n\frac{\pi}{2} - \frac{\pi}{4})}$ (Williams, 1999). The approximated synthesized sound field reads then (Ahrens & Spors, 2010d)

$$S_{\text{appr}}(\mathbf{x}, \omega) = \sqrt{\frac{y_{\text{ref}}}{\sqrt{y^2 + z^2}}} e^{-jk_{\text{pw},x}x} e^{-jk_{\text{pw},y}\sqrt{y^2 + z^2}}. \quad (3.58)$$

In the horizontal plane (the target plane, $z = 0$) in the far-field/high-frequency region, the amplitude of the synthesized sound field $S(\mathbf{x}, \omega)$ shows a decay proportional to $\sqrt{|y|}^{-1}$, i.e. of approximately 3 dB with each doubling of the distance to the secondary source array – the classical amplitude decay for 2^{1/2}-dimensional plane wave synthesis (Sec. 3.3.2). In the near-field/low-frequency region the amplitude decay is slightly different and additionally, some subtle spectral deviations are apparent. The latter circumstance is further discussed in Sec. 4.6.3.

Refer to Fig. 3.12. It depicts the real part and the magnitude of the sound pressure of a continuous linear distribution of secondary point sources synthesizing a virtual plane wave of $f_{\text{pw}} = 1000$ Hz and unit amplitude with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$ referenced to the distance $y_{\text{ref}} = 1.0$ m.

3.5.3 Incorporation of Secondary Sources With Complex Radiation Properties

The incorporation of secondary sources with complex radiation characteristics into the driving functions for planar and linear secondary source distributions (3.41) and (3.49) is less cumbersome than for spherical and circular arrays (Sec. 3.2.3 and 3.3.3). The driving functions in the former cases incorporate $\tilde{G}(\cdot)$ which is the spatio-temporal transfer function of the secondary source located at the origin

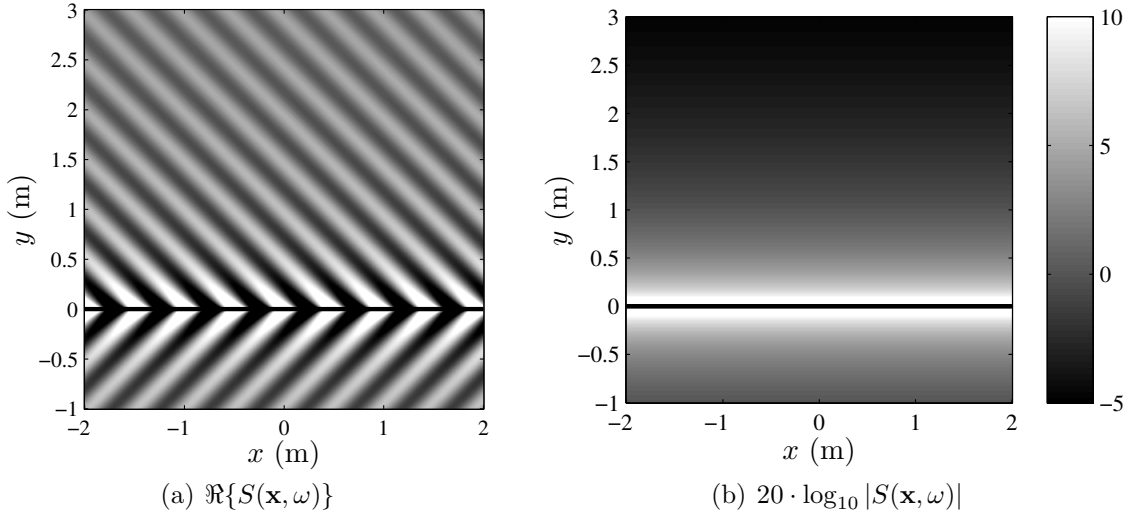


Figure 3.12: Sound pressure $S(\mathbf{x}, \omega)$ of a continuous linear distribution of secondary point sources synthesizing a virtual plane wave of $f_{\text{pw}} = 1000$ Hz and unit amplitude with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$ referenced to the distance $y_{\text{ref}} = 1.0$ m. The secondary source distribution is indicated by the black line. Only the horizontal plane is shown. The values are clipped as indicated by the colorbars.

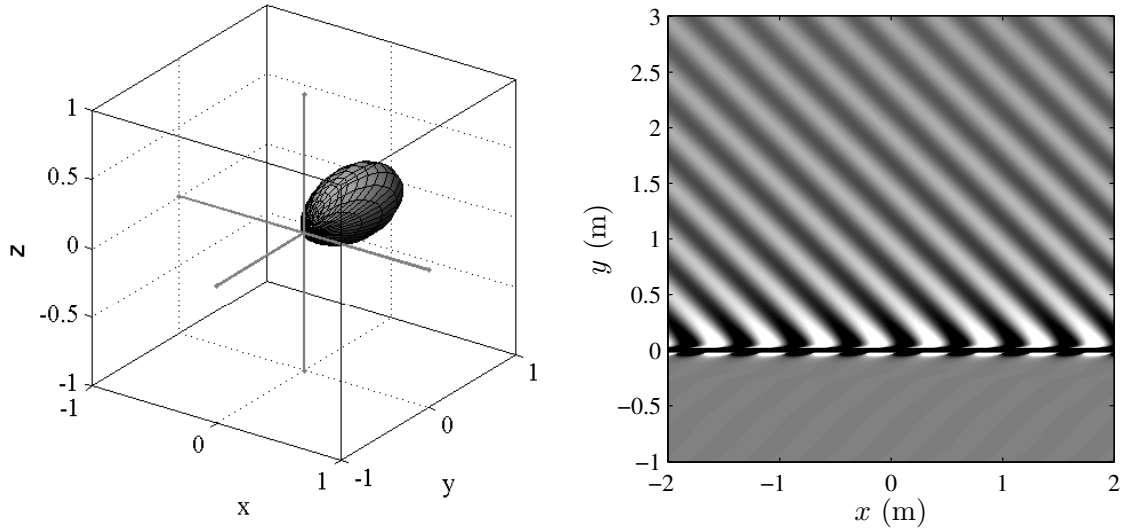
of the coordinate system. This transfer function can be directly obtained from measurements, e.g. employing a linear array of microphones in the horizontal plane and parallel to the x -axis at distance y_{ref} .

In the following example a linear distribution of secondary sources is assumed whose spatio-temporal transfer function $G(\mathbf{x}, \omega)$ is given by (2.30) with $(\alpha_{\text{or}}, \beta_{\text{or}}) = (\frac{\pi}{2}, \frac{\pi}{2})$ and $N = 13$. The far-field signature function is depicted in Fig. 3.13(a).

The secondary source distribution is driven in order to synthesize a virtual plane wave with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$. $\tilde{G}(k_x, y, z, \omega)$ has been calculated numerically in the simulation since an analytical treatment is not straightforward. A simulation of the synthesized sound field is shown in Fig. 3.13(b).

The synthesized sound fields are very similar inside the target half-plane for secondary monopoles (Fig. 3.12(a)) and complex secondary sources (Fig. 3.13(b)) (Ahrens & Spors, 2010a). The latter exhibits slight irregularities close to the secondary source distribution.

For the distribution of monopoles, the sound field synthesized in the half-space other than the target half-space is a perfect mirrored copy of the sound field in target half-space. For the distribution of complex sources, the sound field synthesized in the other half-space differs from the perfect mirrored copy with respect to amplitude and phase (Ahrens & Spors, 2010a). The wave fronts are perfectly straight inside the horizontal plane at sufficient distance from the secondary source distribution.



(a) Normalized far-field signature function of the secondary sources employed in Fig. 3.13(b).

(b) Sound field synthesized by secondary sources exhibiting the transfer function depicted in Fig. 3.13(a).

Figure 3.13: Sound fields in the horizontal plane synthesized by continuous distributions of secondary sources. Desired sound field is a monochromatic plane wave of frequency $f_{\text{pw}} = 1000$ Hz with unit amplitude and propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$.

3.5.4 A Note on Wave Field Synthesis Employing Linear Secondary Source Distributions

Deriving the WFS driving function for a monochromatic plane wave with radian frequency ω_{pw} propagating into the direction $(\theta_{\text{pw}}, \frac{\pi}{2})$ for a linear secondary source distribution yields (Ahrens & Spors, 2010d)

$$D_{\text{WFS}, 2.5\text{D}}(x, \omega) = \sqrt{8\pi y_{\text{ref}}} \sqrt{ik_{\text{pw}}} \sin \theta_{\text{pw}} e^{-jk_{\text{pw},x}x} \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) . \quad (3.59)$$

Recall that 2^{1/2}-dimensional WFS constitutes a high-frequency approximation of the underlying problem (Sec. 2.6.1). In order to compare the WFS solution with the solution presented in Sec. 3.5.1 the high-frequency approximation of the latter is considered, which is given by

$$D_{\text{appr}, 2.5\text{D}}(x, \omega) = \sqrt{8\pi y_{\text{ref}}} \sqrt{jk_{\text{pw}}} \sqrt{\sin \theta_{\text{pw}}} e^{-jk_{\text{pw},x}x} \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) . \quad (3.60)$$

As a consequence of the fact that the driving functions of the two approaches differ by an amplitude factor, the synthesized sound fields differ as well by the same factor.

The synthesized sound fields can only be compared in the high-frequency region because the WFS driving function only holds there. It can indeed be shown that (Ahrens & Spors, 2010d)

$$S_{\text{WFS}, \text{pw}}(\mathbf{x}, \omega) = \sqrt{\sin \theta_{\text{pw}}} \cdot S_{\text{appr}, \text{pw}}(\mathbf{x}, \omega) . \quad (3.61)$$

where $S_{\text{appr}, \text{pw}}(\mathbf{x}, \omega)$ is given by (3.58). From (3.57) and (3.58) it can be seen that the presented approach provides the desired result: A sound field which coincides

with the desired one on the receiver line. It can therefore be concluded that the standard WFS driving function for virtual plane waves (3.59) has to be corrected by a factor of $\sqrt{\sin \theta_{\text{pw}}}$ in order to perform comparably to the presented approach in the high-frequency region (Ahrens & Spors, 2010d).

The source of deviation in WFS seems to lie in the stationary phase approximation in (2.57). In the traditional WFS formulation like (de Vries, 1996; Verheijen, 1997; Start, 1997) the result of this stationary phase approximation is interpreted as a referencing of the synthesized sound field to a line which is parallel to the secondary source distribution. From (2.57) it becomes clear that the synthesized sound field in WFS is actually not referenced to a line but to a circle around the individual secondary sources. The apparent consequence is the incorrect amplitude when $\sqrt{\sin \theta_{\text{pw}}} \neq 1$. This amplitude deviation is low for $\sqrt{\sin \theta_{\text{pw}}} \approx 1$ but can reach several dB for $\sqrt{\sin \theta_{\text{pw}}}$ deviating strongly from 1, i.e. for virtual plane wave fronts which are not approximately parallel to the secondary source distribution.

This type of systematic amplitude error has not been investigated for virtual sound fields other than plane waves. However, the property of the present approach being exact on the reference line unlike WFS has been exploited in various ways, e.g. (Spors & Ahrens, 2010b; Spors & Ahrens, 2010a).

3.5.5 Truncated Linear Secondary Source Distributions

Unlike the secondary source distributions treated in Sec. 3.5.2, practical implementations of sound field synthesis systems can not be of infinite length. The consequences of this spatial truncation are treated in this section. For convenience, a continuous linear secondary source distribution which is truncated in x -dimension is explicitly considered.

The spatial truncation is modeled by multiplying the secondary source driving function $D(\mathbf{x}_0, \omega)$ with a suitable window function $w(x_0)$ (Start, 1997). Incorporating $w(x_0)$ into equation (3.46) yields the sound field $S_{tr}(\mathbf{x}, \omega)$ of a truncated planar source distribution as

$$S_{tr}(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} w(x_0) D(\mathbf{x}_0, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) dx_0 . \quad (3.62)$$

The convolution theorem (3.47) then reads (Girod *et al.*, 2001)

$$\tilde{S}_{tr}(k_x, y, z, \omega) = \frac{1}{2\pi} \underbrace{\left(\tilde{w}(k_x) *_{k_x} \tilde{D}(k_x, \omega) \right)}_{= \tilde{D}_{tr}(k_x, \omega)} \tilde{G}(k_x, y, z, \omega) , \quad (3.63)$$

whereby the asterisk $*_{k_x}$ denotes convolution with respect to the space frequency variable k_x .

The finite extent of a secondary source distribution of length L centered around $x = 0$ can be modeled by a rectangular window $w_R(x)$ as

$$w_R(x) = \text{rect}\left(\frac{x}{L}\right) = \begin{cases} 1 & \text{for } |x| \leq \frac{L}{2} \\ 0 & \text{elsewhere} . \end{cases} \quad (3.64)$$

The Fourier transformation of $w_R(x)$ with respect to x is (Williams, 1999)

$$\tilde{w}_R(k_x) = L \cdot \frac{\sin \frac{k_x L}{2}}{\frac{k_x L}{2}} = L \cdot \text{sinc} \left(\frac{k_x L}{2\pi} \right). \quad (3.65)$$

For the interpretation of (3.63) again the synthesis of a virtual plane wave is considered. Recall $\tilde{D}(k_x, \omega)$ given by (3.53). The convolution of $\tilde{D}(k_x, \omega)$ with $\tilde{w}_R(k_x)$ is essentially a spatial low pass filtering operation smearing $\tilde{D}(k_x, \omega)$ along the k_x -axis. The Dirac $\delta(k_x - k_{\text{pw},x})$ apparent in (3.53) turns into a $\text{sinc}(\cdot)$. The truncated secondary source distribution therefore exhibits distinctive complex radiation properties. Due to the wavenumber domain representation of the synthesized sound field in (3.63) the properties of the synthesized sound field $\tilde{S}_{tr}(k_x, y, z, \omega)$ can be directly obtained from the properties of the truncated driving function $\tilde{D}_{tr}(k_x, \omega)$ as discussed below.

The main lobe of the $\text{sinc}(\cdot)$ function points into the propagation direction of the desired virtual plane wave. However, the synthesized sound field will not exhibit perfectly plane wave fronts but a certain curvature due to the smearing of the energy of the spatial spectrum (Ahrens & Spors, 2010d). The side lobes of the $\text{sinc}(\cdot)$ function result in components in the synthesized sound field propagating into other directions than the desired virtual plane wave. Note that the side lobes exhibit alternating algebraic sign (i.e. the lobes are not in phase) and that there are zeros between the lobes.

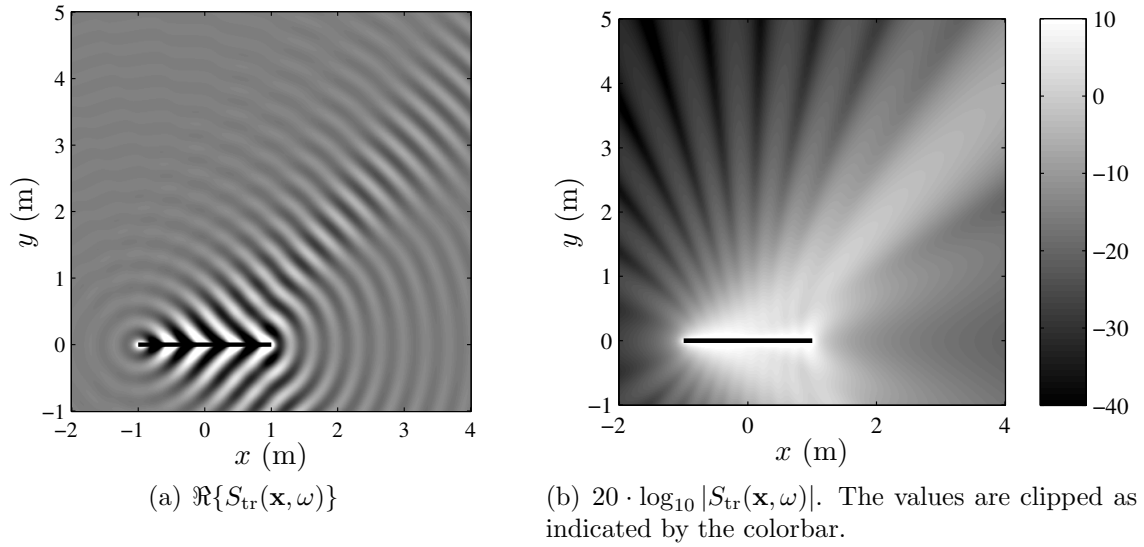


Figure 3.14: Sound pressure $S_{tr}(\mathbf{x}, \omega)$ of a continuous linear distribution of secondary point sources and of length $L = 2$ m synthesizing a virtual plane wave of $f_{\text{pw}} = 1000$ Hz and unit amplitude with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$ referenced to the distance $y_{\text{ref}} = 1.0$ m. The secondary source distribution is indicated by the black line.

Refer to Fig. 3.14. It depicts the sound field synthesized by a continuous truncated linear secondary monopole source distribution. In Fig. 3.14(b), the directivity

lobes due to truncation are clearly apparent. It is also evident from Fig. 3.14 that the local propagation direction of the synthesized sound field strongly depends on the position of the receiver.

Real-world implementations of planar sound field synthesis systems are of course also truncated in z -dimension. Due to the separability of the Cartesian coordinate system (Morse & Feshbach, 1953), the truncation in the two dimensions can be treated independently. The procedure outlined above has to be applied also on the z -dimension.

Further analysis reveals that truncation artifacts can be interpreted as additional sound sources located at the ends of the secondary source distribution (Verheijen, 1997). Of course, other window functions than the rectangular one can be applied some of which provide potential to shape truncation artifacts in order to make them perceptually less disturbing. This process is an established technique in WFS and is referred to as *tapering* (Verheijen, 1997). Typically, windows with cosine-shaped shoulders are applied.

3.6 Approximate Solution for Arbitrary Convex Secondary Source Distributions

The argumentation via the physical optics approximation presented in Sec. 2.6.1 can be employed in order to find an approximate solution for arbitrary convex two-dimensional secondary source distributions based on the solution for planar contours presented in Sec. 3.4 and to find an approximate solution for arbitrary convex one-dimensional secondary source distributions based on the solution for linear contours presented in Sec. 3.5.

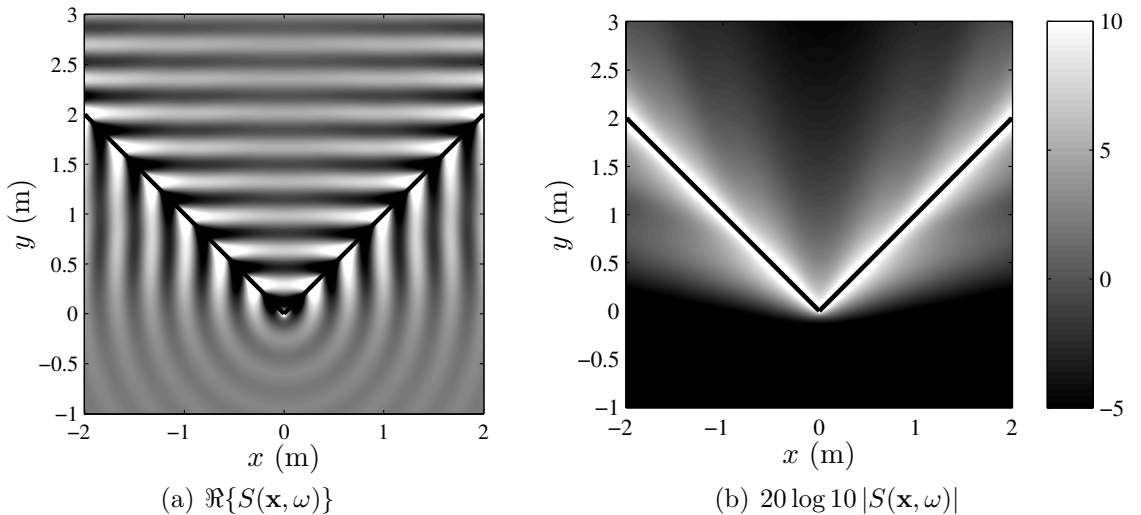


Figure 3.15: Sound fields in the horizontal plane synthesized by two continuous linear distributions of secondary sources which make up an angle of $\frac{\pi}{2}$. Desired sound field is a monochromatic plane wave of frequency $f_{\text{pw}} = 1000$ Hz with unit amplitude and propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{2}, \frac{\pi}{2})$.

For illustration, Fig. 3.15 depicts the sound field synthesized by a combination of two adjoined infinite linear secondary source distributions. It has already been shown in the context of WFS in (Verheijen, 1997) that non-smooth illuminated areas do not introduce a considerable additional error. Similar situations have been investigated in the field of Fourier optics in the context of *Kirchhoff diffraction* (Arfken & Weber, 2005; Nieto-Vesperinas, 2006).

3.7 Summary

The single-layer potential solution to the problem of sound field synthesis was demonstrated for continuous spherical secondary source contours. It was concluded that this solution corresponds to the well established method of *Near-field Compensated Higher Order Ambisonics*.

The single-layer potential approach was then extended to infinite planar secondary source distributions whereby certain restrictions with respect to the propagation direction of the synthesized sound field had to be accepted.

Furthermore, the concept of $2^{1/2}$ -dimensional synthesis was presented for circular and linear distributions of secondary sources. $2^{1/2}$ -dimensional synthesis employs one-dimensional distributions of secondary sources in order to perform two-dimensional synthesis employing secondary sources with three-dimensional spatio-temporal transfer functions. The typical artifacts for $2^{1/2}$ -dimensional synthesis well-known from the WFS context have been confirmed. The most prominent artifact is a generally incorrect amplitude decay of the synthesized sound field.

The solution for planar secondary source distributions was found to be closely related to the well-established concept of angular spectrum representation.

Consequences of spatial truncation of infinite planar and linear secondary source distributions were discussed.

For all geometries discussed, methods for considering the radiation properties of the employed secondary sources were presented. It is emphasized that unlike previously published approaches, the presented methods are not a compensation for deviations of the loudspeaker radiation characteristics from certain assumptions (e.g. omnidirectionality). It is rather such that the formulation of the approach allows for an explicit consideration thereof.

Chapter 4

Discrete Secondary Source Distributions

4.1 Introduction

The continuous secondary source distributions treated in Chap. 3 can not be implemented with today's available technology. Continuous distributions have to be approximated by a finite number of discrete loudspeakers. The consequences of this circumstance are the topic of this chapter.

Commonly, loudspeakers with closed cabinets are employed in practice which are assumed to be omnidirectional, i.e. monopole pressure sources. This assumption is indeed fulfilled for low frequencies of a few hundred Hertz but at higher frequencies, complex radiation patterns evolve (Fazi *et al.*, 2008a). For simplicity, the present chapter investigates the consequences of discretization of the secondary source distribution under the assumption that ideal secondary monopole sources are employed. Spherical secondary source distributions are treated in Sec. 4.3, circular ones in Sec. 4.4, planar ones in Sec. 4.5, and linear ones in Sec. 4.6.

In order to avoid unnecessary redundancies, time-domain properties of the synthesized sound fields with a focus on human auditory perception (Sec. 4.4.4 and 4.6.3) as well as an advanced technique termed *local sound field synthesis* are exclusively treated for circular and linear contours (Sec. 4.4.5 and Sec. 4.6.5 respectively). The general properties of spherical contours in this context can be deduced from the results for circular contours; the properties of planar contours can be deduced from the results for linear contours.

4.2 Excursion: Discretization of Time-Domain Signals

As outlined in (Verheijen, 1997), it is useful to emphasize analogies between spatial discretization and discretization of time-domain signals. Therefore, the latter is briefly reviewed in this section.

Assume a continuous time-domain signal $s_0(t)$ whose time-frequency spectrum

is given by

$$S_0(\omega) = \int_{-\infty}^{\infty} s_0(t) e^{-i\omega t} dt . \quad (4.1)$$

In order that $s_0(t)$ can be stored in a digital system it is discretized in time at *sampling frequency* f_s , i.e. with the constant sampling interval $T_s = \frac{1}{f_s}$, as (Girod *et al.*, 2001; Zayed, 1993)

$$s_{0,s}(t) = s_0(t) \underbrace{\sum_{\mu=-\infty}^{\infty} \delta(t - T_s \mu)}_{= \xi(t)} . \quad (4.2)$$

The time-frequency spectrum $S_{0,s}(\omega)$ of the sampled signal is given by

$$S_{0,s}(\omega) = \int_{-\infty}^{\infty} s_0(t) \xi(t) e^{-i\omega t} dt . \quad (4.3)$$

Eq. (4.3) constitutes a Fourier transform of the product of two functions. The *multiplication theorem* of the Fourier transform states that the result can be expressed as a convolution of the time-frequency spectra $S_0(\omega)$ and $\xi(\omega)$ of the two functions with respect to the frequency (Girod *et al.*, 2001). Explicitly,

$$S_{0,s}(\omega) = \frac{1}{2\pi} S_0(\omega) *_\omega \xi(\omega) . \quad (4.4)$$

The Fourier transform $\xi(\omega)$ of the sampling pulse train $\xi(t)$ is given by (Girod *et al.*, 2001)

$$\xi(\omega) = \int_{-\infty}^{\infty} \sum_{\mu=-\infty}^{\infty} \delta(t - T_s \mu) e^{-i\omega t} dt \quad (4.5)$$

$$= \frac{2\pi}{T_s} \sum_{\mu=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{T_s} \mu\right) , \quad (4.6)$$

so that $S_{0,s}(\omega)$ is finally given by

$$\begin{aligned} S_{0,s}(\omega) &= S_0(\omega) *_\omega \frac{1}{T_s} \sum_{\mu=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{T_s} \mu\right) \\ &= \frac{1}{T_s} \sum_{\mu=-\infty}^{\infty} S_0(\omega - 2\pi f_s \mu) . \end{aligned} \quad (4.7)$$

Eq. (4.7) states that the time-frequency spectrum $S_{0,s}(\omega)$ of a time-discrete signal is given by repetitions of period $\omega_s = 2\pi f_s$ of the time-frequency spectrum $S_0(\omega)$ of the initial continuous signal. For $\mu = 0$, (4.7) corresponds to $S_0(\omega)$, the spectrum of the continuous signal $s_0(t)$, scaled by $\frac{1}{T_s}$.

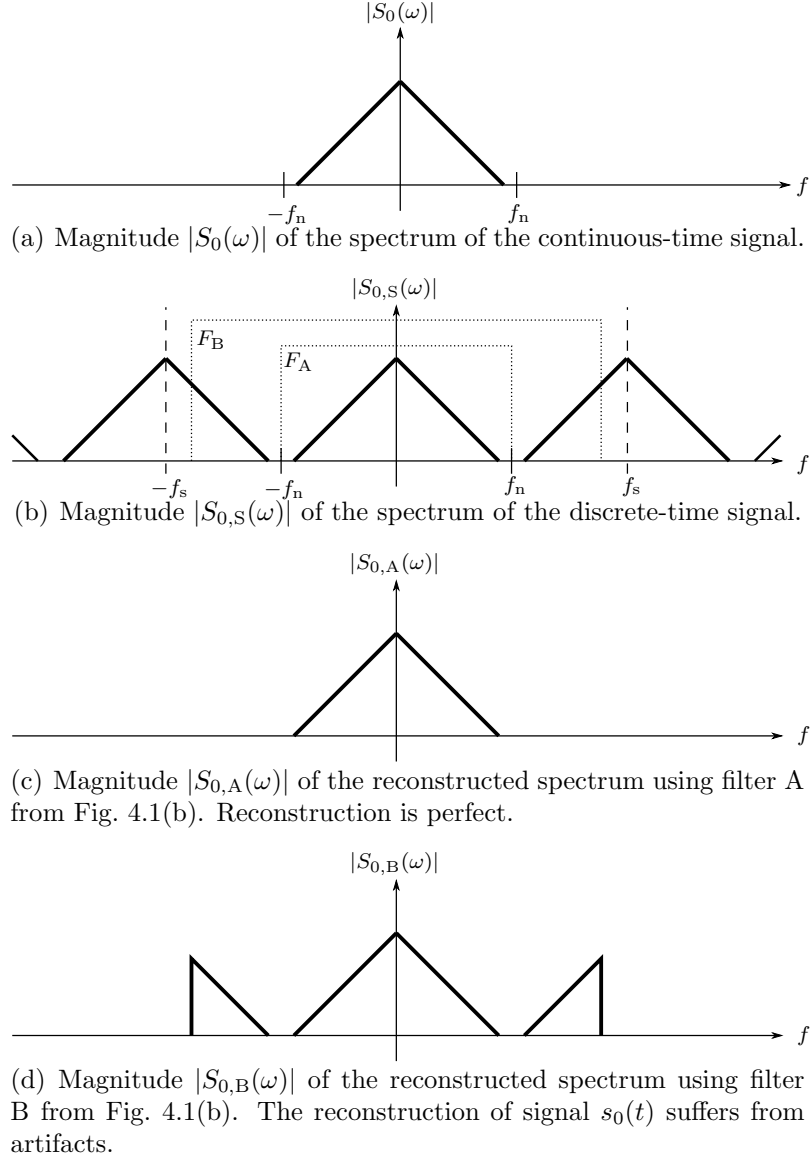


Figure 4.1: Sampling of a purely real bandlimited time-domain signal.

It is possible to perfectly reconstruct the initial time-domain signal $s_0(t)$ from the discretized signal $s_{0,S}(t)$ if certain assumptions are met. The procedure is indicated in Fig. 4.1(a). Fig. 4.1(a) sketches the time-frequency spectrum $S_0(\omega)$ of continuous time-domain signal $s_0(t)$. The according time-frequency spectrum $S_{0,S}(\omega)$ of the discretized signal $s_{0,S}(t)$ is indicated in Fig. 4.1(b). Note that it is assumed that $s_0(t)$ is bandlimited such that its energy is exclusively contained at frequencies at or below $f_n = \frac{f_s}{2}$. f_n is termed *Nyquist frequency* (Girod *et al.*, 2001).

Due to the bandlimitedness of $s_0(t)$, the spectral repetitions of the discretized signal do not overlap. By applying an appropriate lowpass filter (the transfer function of which is indicated by the dotted line marked F_A in Fig. 4.1(b)), the continuous time domain signal $s_0(t)$ can be perfectly reconstructed as indicated in Fig. 4.1(c). The filter F_A is also termed *interpolation filter* or *reconstruction filter*.

Two circumstances lead to a corrupted reconstruction of $s_0(t)$: (Girod *et al.*,

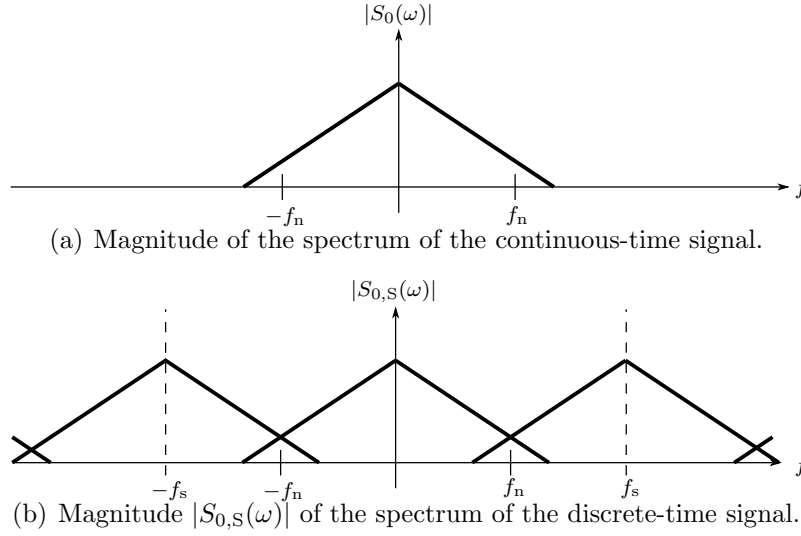


Figure 4.2: Sampling of a signal exhibiting energy above f_n .

2001)

1. If the passband of the reconstruction filter is wider than $2f_n = f_s$ like the filter whose transfer function is marked F_B in Fig. 4.1(b), then the spectral repetitions are not perfectly suppressed in the reconstruction. This type of error is generally referred to as *reconstruction error*.
2. If $s_0(t)$ exhibits energy above f_n the spectral repetitions overlap and interfere. Refer to Fig. 4.2 for a sketch. It is not possible to separate the baseband from the discretized signal and the reconstruction is corrupted by *aliasing*.

The reconstruction $S_{0,S,\text{rec}}(\omega)$ from the time-discrete representation $S_{0,S}(\omega)$ can be represented in time-frequency domain by as (Girod *et al.*, 2001)

$$S_{0,S,\text{rec}}(\omega) = S_{0,S}(\omega) \cdot F_A(\omega) , \quad (4.8)$$

whereby $F_A(\omega)$ denotes the transfer function of the reconstruction filter. If the bandwidth of $S_0(\omega)$ and the properties of the reconstruction filter $F_A(\omega)$ are according then $S_{0,S,\text{rec}}(\omega) = S_0(\omega)$ and the reconstruction is perfect.

Fig. 4.3 summarizes the process of sampling a continuous time-domain signal $S_0(\omega)$ and reconstructing the signal $S_{0,S,\text{rec}}(\omega)$ from the time-discrete representation $S_{0,S}(\omega)$ via a reconstruction filter with transfer function $F_A(\omega)$.

In the remainder of this chapter, the investigation of discretization of a time-domain signal is adapted to the spatial discretization of the secondary source distributions investigated in Chap. 3. For convenience, spatial discretization is modeled by a discretization of the corresponding driving function. Thus, a continuous distribution of secondary sources is assumed which is driven at discrete points. The essential benefit of this approach is the fact that all integral and convolution theorems exploited in Chap. 3 are still valid. The consequences of spatial discretization can therefore be deduced from an investigation of the properties of the discretized driving functions.

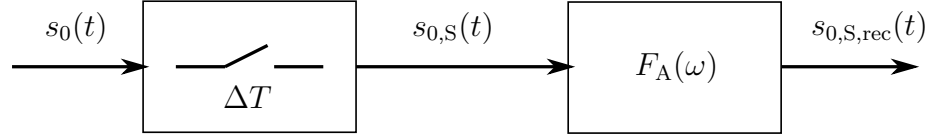


Figure 4.3: Schematic of the process of discretization and reconstruction of the continuous time-domain signal $S_0(\omega)$. $F_A(\omega)$ denotes the transfer function of the reconstruction filter.

4.3 Spherical Secondary Source Distributions

4.3.1 Discretization of the Sphere

In contrast to the sampling of time-domain signals outlined in Sec. 4.2, it is not obvious how sampling of a spherical secondary source distribution can be performed in a favorable way. Generally, the discretization grid shall be such that the orthogonality relation of the spherical harmonics (2.17) holds (Driscoll & Healy, 1994). Eq. (2.17) reformulated using a discretized integral is given by

$$\sum_l w_l Y_n^m(\beta_l, \alpha_l) Y_{n'}^{-m'}(\beta_l, \alpha_l) = \delta_{nn'} \delta_{mm'} . \quad (4.9)$$

The weights w_l compensate for a potentially uneven distribution of the sampling points.

It can be shown that sampling schemes can be found for which (4.9) does indeed hold when spatially bandlimited functions are considered (Driscoll & Healy, 1994). An exact uniform sampling is exclusively provided by layouts based on one of the five *platonic solids* tetrahedron, cube, octahedron, dodecahedron, and icosahedron (Weisstein, 2002).

In general, the available sampling strategies can be categorized into (quasi) uniform and non-uniform approaches. The most popular approaches are *hyperinterpolation*, *quadrature*, and the *(weighted) least-squares solution*; all of which exhibit benefits and drawbacks (Zotter, 2009b). For convenience, a non-uniform layout given by the *Gauß sampling scheme* is chosen here due to its relatively simple mathematical description (Mohlenkamp, 1999; Driscoll & Healy, 1994).

When a Gauß sampling scheme with $2L^2$ sampling points is assumed, the azimuth angle α_0 is sampled equiangularly at $2L$ locations and the zenith angle β_0 is sampled at L locations. This results in a sampling grid which is symmetric with respect to the horizontal plane.

Mathematically, the Gaussian sampling grid $\Phi(\alpha, \beta, L)$ is given by (Driscoll & Healy, 1994)

$$\Phi(\alpha, \beta, L) = \frac{\pi}{2L^2} \sum_{l_1=0}^{2L-1} \delta(\alpha - \alpha_{l_1}) \sum_{l_2=0}^{L-1} w_{l_2} \delta(\beta - \beta_{l_2}) \quad (4.10)$$

with

$$\alpha_{l_1} = \frac{2\pi l_1}{2L} . \quad (4.11)$$

The angles β_{l_2} are computed as the zeros of the L -th degree Legendre polynomial $P_L(\cos \beta_{l_2}) \stackrel{!}{=} 0$. Refer to Fig. 4.4 for an example grid.

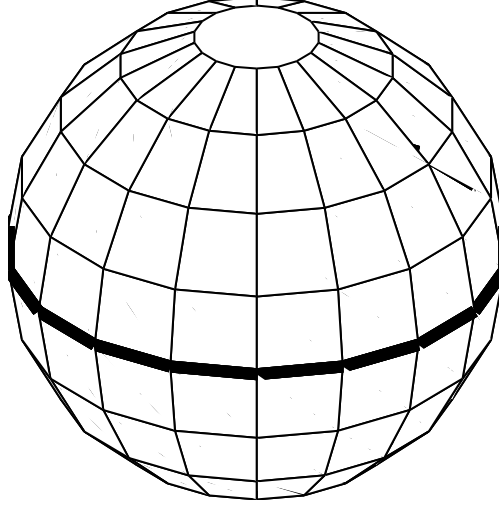


Figure 4.4: Gauß sampling grid for $L = 9$. The sampling points are represented by the intersections of the lines. The thick line indicates the equator.

The process of calculating the weights w_{l_2} is outlined in (Driscoll & Healy, 1994). Note that in the simulations in this chapter, the MATLAB scripts provided by (The Chebfun Team, 2009) are employed.

4.3.2 Discretization of the Driving Function

The analysis of the consequences of spatial discretization of a representation of a sound field on the surface of a sphere has been performed in (Rafaely *et al.*, 2007). An alternative approach will be presented below which allows for a frequency-dependent modal decomposition of the synthesized sound field.

It can be shown via (3.4) from Sec. 3.2.1 that the expansion coefficients $\check{S}_n^m(\omega)$ of the synthesized sound field are given by a multiplication of the spherical harmonics expansion coefficients $\check{D}_n^m(\omega)$ of the driving function and the expansion coefficients $\check{G}_n^m(\omega)$ of the spatio-temporal transfer function of the secondary sources. Thus, if it is possible to determine the coefficients $\check{D}_{n,s}^m(\omega)$ of the sampled driving function $D_s(\mathbf{x}, \omega)$, the synthesized sound field $S(\mathbf{x}, \omega)$ can be determined via its expansion coefficients $\check{S}_n^m(\omega)$.

The spherical harmonics transform of the sampled driving function $D(\mathbf{x}, \omega)$ is given by

$$\check{D}_{n,s}^m(R, \omega) = \int_0^{2\pi} \int_0^\pi \Phi(\alpha, \beta, L) D(\mathbf{x}, \omega) Y_n^{-m}(\beta, \alpha) \sin \beta \, d\beta d\alpha . \quad (4.12)$$

Eq. (4.12) constitutes the spherical harmonics transform of a product of the functions $\Phi(\alpha, \beta, L)$ and $D(\mathbf{x}, \omega)$. As derived in App. D.2, this spherical harmonics transform

can be formulated in terms of the spherical harmonics expansion coefficients $\mathring{\Phi}_n^m(L)$ and $\mathring{D}_n^m(\omega)$ as

$$\mathring{D}_{n,S}^m(R, \omega) = \sum_{n_1=0}^{\infty} \sum_{m_1=-n_1}^{n_1} \sum_{n_2=0}^{\infty} \mathring{\Phi}_{n_1}^{m_1}(L) \mathring{D}_{n_2}^{m-m_1}(R, \omega) \gamma_{n_1, n_2, n}^{m_1, m-m_1, m}, \quad (4.13)$$

whereby $\gamma_{n_1, n_2, n}^{m_1, m-m_1, m}$ denotes the *Gaunt coefficient* and is given by (D.6).

The spherical harmonics transform $\mathring{\Phi}_{n_1}^{m_1}(L)$ of the Gauß sampling grid can be determined via

$$\mathring{\Phi}_{n_1}^{m_1}(L) = \int_0^{2\pi} \int_0^{\pi} \Phi(\alpha, \beta, L) Y_{n_1}^{-m_1}(\beta, \alpha) \sin \beta \, d\beta d\alpha. \quad (4.14)$$

The integrals in (4.14) can be solved independently as

$$\begin{aligned} \int_0^{2\pi} \sum_{l_1=0}^{2L-1} \delta\left(\alpha - \frac{2\pi l_1}{2L}\right) e^{-im_1\alpha} \, d\alpha &= \sum_{l_1=0}^{2L-1} e^{-im_1 2\pi \frac{l_1}{2L}} \\ &= \begin{cases} 2L & \forall \, m_1 = \mu 2L, \, \mu \in \mathbb{Z} \\ 0 & \text{elsewhere} \end{cases}, \end{aligned}$$

and

$$\int_0^{\pi} \sum_{l_2=0}^{L-1} w_{l_2} \delta(\beta - \beta_{l_2}) P_{n_1}^{|m_1|}(\cos \beta) \sin \beta \, d\beta = \sum_{l_2=0}^{L-1} w_{l_2} P_{n_1}^{|m_1|}(\cos \beta_{l_2}) \sin \beta_{l_2}. \quad (4.15)$$

From the parity properties of the sampling locations β_{l_2} , the associated Legendre functions, and the sine function in (4.15), it can be deduced that the result equals zero for $m_1 + n_1$ being odd.

The spherical harmonics expansion coefficients $\mathring{\Phi}_{n_1}^{m_1}(L)$ of the sampling grid are finally given by

$$\mathring{\Phi}_{n_1}^{m_1}(L) = \begin{cases} \frac{\pi(-1)^{m_1}}{L+1} \sqrt{\frac{2n_1+1}{4\pi} \frac{(n_1-|m_1|)!}{(n_1+|m_1|)!}} \sum_{l_2=0}^{L-1} w_{l_2} P_{n_1}^{|m_1|}(\cos \beta_{l_2}) \sin \beta_{l_2} & \forall \, m_1 = \mu 2L \\ 0 & \text{elsewhere} \end{cases}. \quad (4.16)$$

Introducing (4.16) into (4.13), changing the order of summations, and considering selection rule 2 from Sec. D.2 yields

$$\mathring{D}_{n,S}^m(R, \omega) = \sum_{\mu=-\infty}^{\infty} \sum_{n_2=|m-\mu 2L|}^{\infty} \mathring{D}_{n_2}^{m-\mu 2L}(R, \omega) \Upsilon_{n_2, n}^{\mu, m}(L), \quad (4.17)$$

with

$$\Upsilon_{n_2, n}^{\mu, m}(L) = \sum_{n_1=|n-n_2|}^{n+n_2} \mathring{\Phi}_{n_1}^{\mu 2L}(L) \gamma_{n_1, n_2, n}^{\mu 2L, m-\mu 2L, m}. \quad (4.18)$$

Via the selection rules of the Gaunt coefficient outlined in App. D.2 and the symmetry relations of the involved Wigner3j-Symbols (Weisstein, 2002) it can be shown that $\mathring{D}_{n,S}^m(R, \omega)$ given by (4.17) is composed of the coefficients $\mathring{D}_n^m(R, \omega)$ of the continuous driving function plus repetitions with respect to n and m of $\mathring{D}_n^m(R, \omega)$. The period of the repetitions both in n and m is $2L$. It can furthermore be shown that $\mathring{D}_{n,S}^m(R, \omega) = \mathring{D}_n^m(R, \omega)$ for the case of $\mu = 0$. For convenience, simulations of a sample scenario are presented below for illustration of above outlined properties of $\mathring{D}_{n,S}^m(R, \omega)$.

The sound field $\mathring{S}_{n,S}^m(r, \omega)$ synthesized by the discrete secondary source distribution is given in spherical harmonics domain by (3.4), which is stated here again for convenience as

$$\mathring{S}_{n,S}^m(r, \omega) = 2\pi R \sqrt{\frac{4\pi}{2n+1}} \mathring{D}_{n,S}^m(\omega) \cdot \mathring{G}_n^0(r, \omega) . \quad (4.19)$$

Note the similarity between (4.19) and (4.8): A discretized function being composed of repetitions of the underlying continuous function is weighted in a transformed domain in order to yield the desired quantity.

This analogy greatly facilitates the interpretation of (4.19). The discretization of the secondary source distribution leads to repetitions in the spatial spectrum of the driving function. When these repetitions do not overlap – and thus do not corrupt the base band – and the properties of the spatio-temporal transfer function of the employed secondary sources are such that these repetitions are suppressed, then the synthesized sound field is unaffected by the discretization. The spatio-temporal transfer function $\mathring{G}_n^0(r, \omega)$ of the employed secondary sources can thus be interpreted as the analogon of the reconstruction filter denoted $F_A(\omega)$ in the time discretization example in Fig. 4.3. Fig. 4.5 depicts an adaptation of Fig. 4.3 to the present situation.

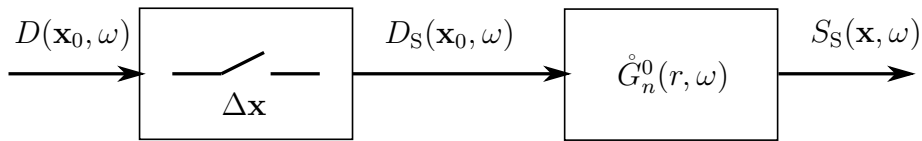


Figure 4.5: Schematic of the spatial discretization process for spherical secondary source distributions.

In order to illustrate the properties of the repetitions which occur in the angular domain due to the discretization of the driving function, the scenario of a discrete spherical distribution of radius $R = 1.5$ m composed of 1568 secondary monopole sources on a Gauß grid ($L = 28$) synthesizing a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ is considered. In this case, the spherical harmonics coefficients $\mathring{D}_n^m(\omega)$ of the continuous driving function $D(\mathbf{x}, \omega)$ are given by

$$\mathring{D}_n^m(\omega) = \begin{cases} \frac{2i}{R} \sqrt{\frac{2n+1}{4\pi}} \frac{i^{-n} Y_n^{-m}(\frac{\pi}{2}, -\frac{\pi}{2})}{\frac{\omega}{c} h_n^{(2)}(\frac{\omega}{c} R) Y_n^0(0,0)} & \forall \ n, |m| \leq M \\ 0 & \text{elsewhere} \end{cases} , \quad (4.20)$$

whereby (2.26a) and (2.27) were used. The choice of the bandlimit M is yet to be determined. It is introduced since an infinite bandwidth as suggested by (3.8) can generally not be implemented in practice. The driving function (3.8) for spherical secondary source distributions is thus finally given by

$$D(\alpha, \beta, \omega) = \sum_{n=0}^M \sum_{m=-n}^n \mathring{D}_n^m(\omega) Y_n^m(\beta, \alpha) . \quad (4.21)$$

Fig. 4.6(a) depicts the magnitude of $\mathring{D}_n^m(\omega)$ given by (4.20) for $M \rightarrow +\infty$. Note that in Fig. 4.6 (as well as in Fig. 4.7) the magnitude is indicated both via brightness as well as via transparency. Values below the lower limit indicated by the errorbars are fully transparent; opacity increases proportionally to the magnitude and reaches full opacity for values above the upper limit indicated by the errorbars.

When sampling of the driving function is considered, it can be seen from Fig. 4.6(b) that parts of the spectral repetitions with considerable energy overlap and interfere. The period of the repetitions of $2L = 56$ with respect to both n and m is also apparent. Since the repetitions also leak into the baseband of $\mathring{D}_{n,S}^m(\omega)$ *spatial aliasing* occurs.

Choosing a spatial bandlimit of the driving function as $M \leq L$ prevent the spectral repetitions from corrupting the baseband of $\mathring{D}_{n,S}^m(\omega)$ and thus suppresses spatial aliasing as depicted in Fig. 4.7. Since $\mathring{G}_n^0(r, \omega)$ is not bandlimited – as can be deduced from (2.26a) – the spectral repetitions are not suppressed and the synthesized sound field suffers from a *reconstruction error*.

Note that it is common in sound field synthesis to refer to this reconstruction error as spatial aliasing, e.g. (Verheijen, 1997; Pueo *et al.*, 2007; Zotter *et al.*, 2009; Wu & Abhayapala, 2009). This thesis does not follow this convention and employs a strict segregation of aliasing and reconstruction errors. Strictly speaking, aliasing constitutes a corruption of the baseband due to overlapping spectral repetitions (Girod *et al.*, 2001). Artifacts which are a consequence of the circumstance that the reconstruction filter does not perfectly suppress spectral repetitions are termed reconstruction error. Therefore, the notion of a *spatial aliasing frequency* as commonly used (Verheijen, 1997; Theile, 2004; Pueo *et al.*, 2007), i.e. the frequency below which on considerable artifacts arise, is not appropriate here. Refer to Sec. 4.2 for a discussion of the terminology in the time discretization example.

4.3.3 Properties of the Synthesized Sound Field in Time-Frequency Domain

The discussion of the properties of the discretized driving function in Sec. 4.3.2 suggests that the spatial bandwidth of the employed continuous driving function has essential impact on the properties of the sound field synthesized by a discrete distribution of secondary sources. This circumstance is indeed evident from Fig. 4.8 which depicts the synthesized sound field for different bandwidths and time frequencies. The synthesis of a plane wave with parameters outlined in Sec. 4.3.2 is considered. The synthesized sound field was derived using (4.17) and (3.4).

At rather low frequencies, no considerable difference between the case of $M =$

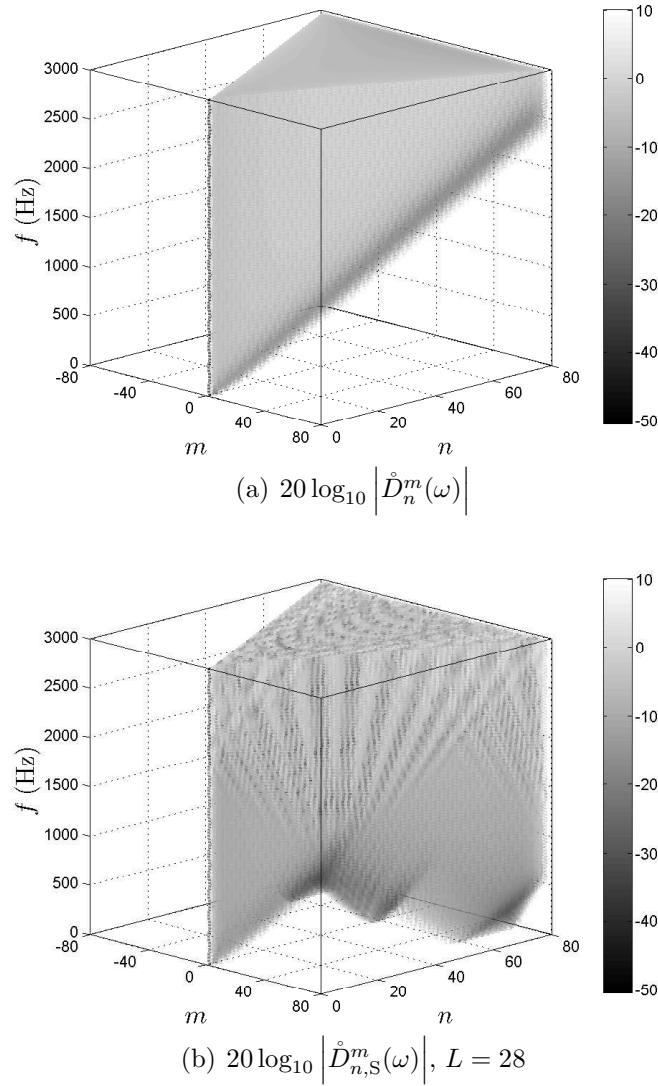


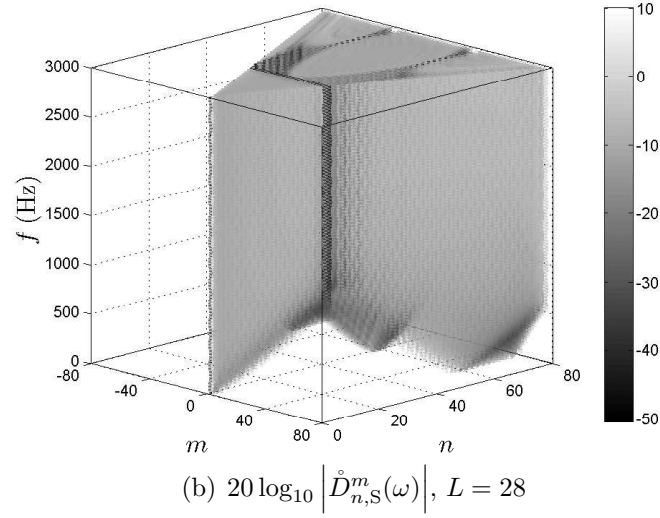
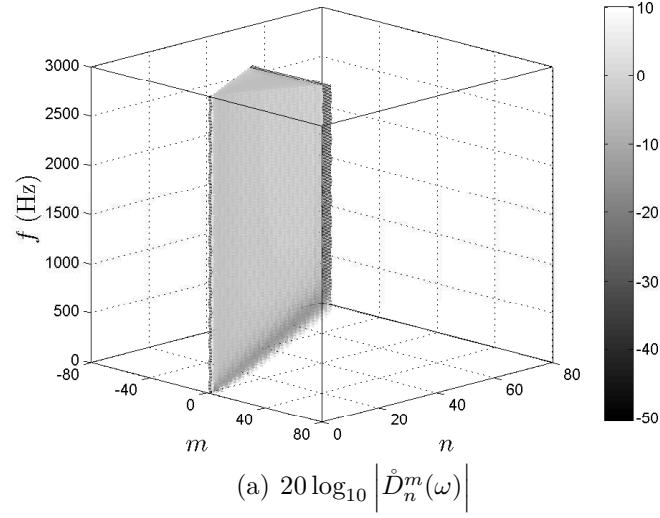
Figure 4.6: Illustration of the properties of the driving function; $M \rightarrow +\infty$

L in Fig. 4.8(a) and the case of $M \rightarrow +\infty$ in Fig. 4.8(b) is apparent. Obvious differences arise at higher frequencies as discussed in the following.

As mentioned in Sec. 4.3.2, a spatial bandlimit of $M \leq L$ leaves the lower orders of the driving function and thus of the synthesized sound field uncorrupted. A region of nearly artifact-free synthesis arises around the center of the secondary source distribution as evident in particular from Fig. 4.8(c) and 4.8(e). Recall that the lower orders typically describe the sound field around the center of the expansion (refer to Sec. 2.2.2). In the present case, the latter coincides with the center of the secondary source distribution.

This region of nearly artifact-free synthesis is bounded by a sphere of radius r_M . For a bandwidth of the driving function of $M \rightarrow +\infty$ artifacts arise over the entire interior domain at higher frequencies (Fig. 4.8(f)).

A further detailed analysis of the synthesized sound fields is not performed here.

Figure 4.7: Illustration of the properties of the driving function; $M = 28$

As will be shown in Sec. 4.4, the properties of circular secondary source distributions with respect to spatial discretization are very similar to those of spherical ones. For convenience, further detailed discussion is performed *ibidem*.

Due to the fundamental impact of the spatial bandwidth of the driving function on the properties of the synthesized sound field, it is proposed to categorize the synthesis with respect to the spatial bandwidth of the continuous driving function into spatially *narrowband*, *wideband*, and *fullband* synthesis. The term narrowband is applied when the bandwidth of the continuous driving function is so low that the spectral repetitions due to spatial discretization do not overlap, i.e. $M \leq L$ (as in Fig. 4.8(a), 4.8(c), or 4.8(e)).

The term fullband ($M \rightarrow +\infty$) reflects the fact that the spatial bandwidth of the driving function is so large that a further increase of the bandwidth does not lead to considerable changes in the domain of interest (as in Fig. 4.8(b), 4.8(d), or

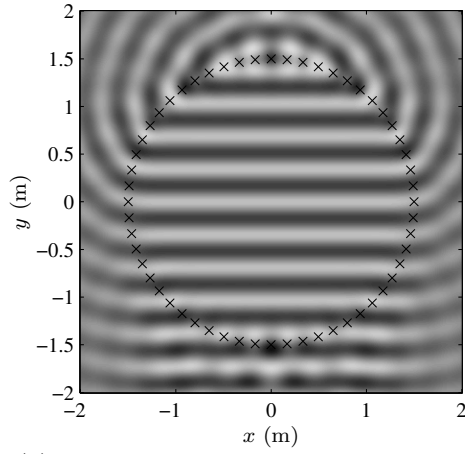
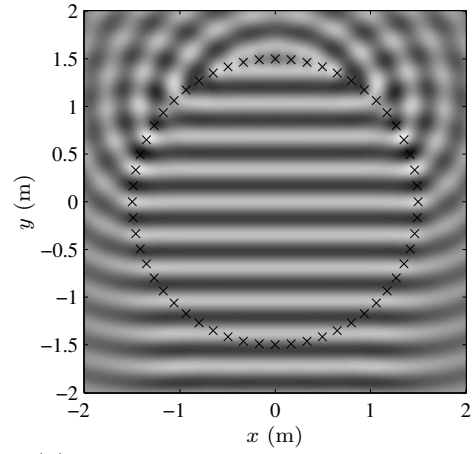
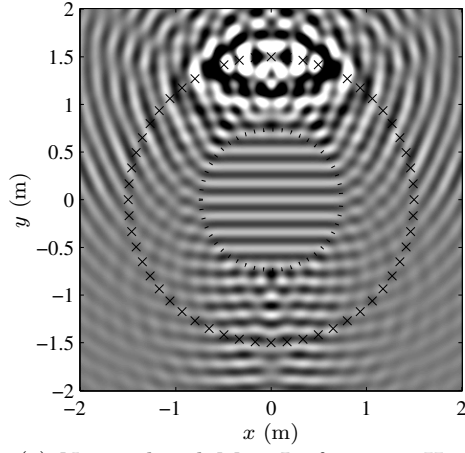
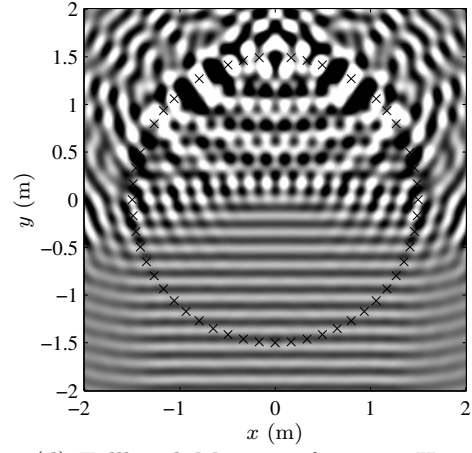
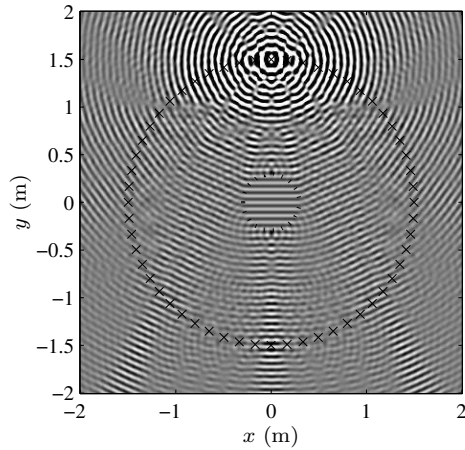
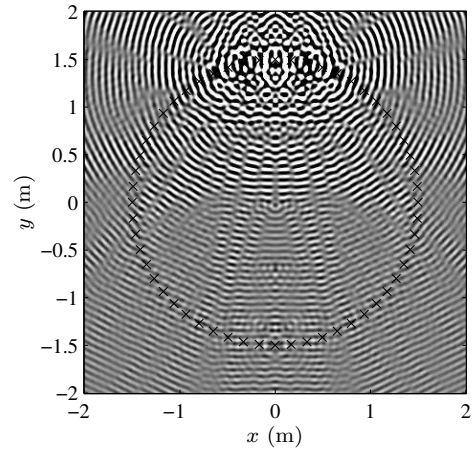
(a) Narrowband $M = L$, $f = 1000$ Hz.(b) Fullband $M \rightarrow \infty$, $f = 1000$ Hz.(c) Narrowband $M = L$, $f = 2000$ Hz.(d) Fullband $M \rightarrow \infty$, $f = 2000$ Hz.(e) Narrowband $M = L$, $f = 5000$ Hz.(f) Fullband $M \rightarrow \infty$, $f = 5000$ Hz.

Figure 4.8: Synthesized sound field in the horizontal plane for the synthesis of a plane wave for different bandwidths of the driving function. The marks indicate the positions of those secondary sources which are contained in the horizontal plane. The dotted lines bound the r_{27} -region in the narrowband case. $L = 28$, thus 1568 secondary sources are employed.

4.8(f)).

Driving functions with $L \ll M \ll +\infty$ may be termed spatially wideband. Wideband driving functions exhibit thus a significantly larger spatial bandwidth than narrowband driving functions (so that overlaps of the spectral repetitions occur); but a further increase of the bandwidth does considerably change the properties of the synthesized sound field.

The spatial bandwidth limitation does not need to be a sharp truncation as performed above but a smooth fade-out towards higher orders may also be applied. This latter approach is especially promising for wideband driving functions. The properties of the latter can not be investigated in this thesis and are subject to future work.

Note that the near-field compensated higher order Ambisonics approach as it is typically applied (e.g. in (Daniel, 2001; Zotter *et al.*, 2009; Fazi *et al.*, 2009)) constitutes narrowband synthesis.

4.4 Circular Secondary Source Distributions

In this section, the procedure outlined in Sec. 4.3 is adapted to circular secondary source contours. Again, the employment of a discrete secondary source distribution is modeled by a discretization of the driving function. For circular contours, uniform sampling can straightforwardly be achieved via equiangular sampling with a sampling interval equal to an integer fraction of 2π .

4.4.1 Discretization of the Driving Function

In the following, it is assumed that the circular secondary source contour under consideration is sampled equiangularly at L points. The sampling interval is thus $\Delta\alpha = \frac{2\pi}{L}$. The discretized driving function $D_S(\alpha, \omega)$ is given by (Girod *et al.*, 2001)

$$D_S(\alpha, \omega) = \underbrace{\frac{1}{L} \sum_{l=0}^{L-1} \delta\left(\alpha - \frac{l}{L} 2\pi\right)}_{= \Psi(\alpha, L)} D(\alpha, \omega) . \quad (4.22)$$

The Fourier series expansion coefficients $\mathring{D}_{m,S}(\omega)$ of the discretized driving function $D_S(\alpha, \omega)$ are given by (Williams, 1999)

$$\mathring{D}_{m,S}(\omega) = \frac{1}{2\pi} \int_0^{2\pi} \Psi(\alpha, L) D(\alpha, \omega) e^{-im\alpha} d\alpha . \quad (4.23)$$

Eq. (4.23) constitutes the Fourier series transform of a product of the functions $\Psi(\alpha, L)$ and $D(\alpha, \omega)$. As derived in App. D.1, this Fourier series transform can be formulated in terms of the Fourier series expansion coefficients $\mathring{\Psi}_m(L)$ and $\mathring{D}_m(\omega)$ as

$$\mathring{D}_{m,S}(\omega) = \sum_{m_1=-\infty}^{\infty} \mathring{\Psi}_{m_1}(L) \mathring{D}_{m-m_1}(\omega) . \quad (4.24)$$

The Fourier series transform of the equiangular sampling grid $\Psi(\alpha, L)$ is given by (Weisstein, 2002)

$$\begin{aligned}\mathring{\Psi}_{m_1}(L) &= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{L} \sum_{l=0}^{L-1} \delta\left(\alpha - \frac{2\pi l}{L}\right) e^{-im_1\alpha} d\alpha \\ &= \frac{1}{L} \sum_{l=0}^{L-1} e^{-im_1 2\pi \frac{l}{L}} \\ &= \begin{cases} 1 & \forall m_1 = \mu L, \mu \in \mathbb{Z} \\ 0 & \text{elsewhere} \end{cases},\end{aligned}$$

so that $\mathring{D}_{m,S}(\omega)$ is finally given by (Girod *et al.*, 2001; Spors & Rabenstein, 2006; Ahrens & Spors, 2008a)

$$\mathring{D}_{m,S}(\omega) = \sum_{\mu=-\infty}^{\infty} \mathring{D}_{m-\mu L}(\omega). \quad (4.25)$$

The spatial spectrum $\mathring{D}_{m,S}(\omega)$ of the sampled driving function is thus composed of repetitions of the spatial spectrum $\mathring{D}_m(\omega)$ of the continuous driving function (eq. (4.25)) with a period of L .

According to (3.23), the synthesized sound field $\mathring{S}_{m,S}(r, \omega)$ in Fourier series domain is given by $\mathring{D}_{m,S}(\omega)$ weighted by the spatio-temporal transfer function $\mathring{G}_m(r, \omega)$ of the secondary sources as

$$\mathring{S}_{m,S}(r, \omega) = 2\pi R \mathring{D}_{m,S}(\omega) \mathring{G}_m(r, \omega). \quad (4.26)$$

Eq. (4.26) constitutes the analogon to (4.8) and (4.19). The adaptation of Fig. 4.3 and 4.5 to the present situation is depicted in Fig. 4.9.

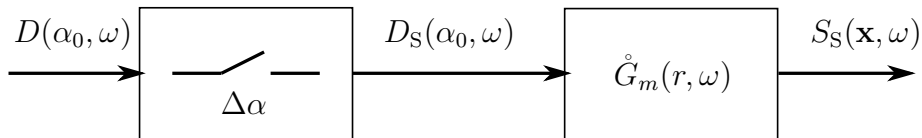


Figure 4.9: Schematic of the spatial discretization process for circular secondary source distributions.

In order to illustrate the consequences of the repetitions which occur in the Fourier series domain due to the discretization of the driving function, the scenario of a discrete circular distribution of radius $R = 1.5$ m composed of $L = 56$ equiangularly spaced secondary monopole sources synthesizing a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ is considered.

It was noted in Sec. 3.3.1 that the summation in the driving function (3.26) can not be performed over an infinite amount of coefficients in practice. The Fourier

coefficients $\mathring{D}_m(\omega)$ of the continuous driving function therefore have to be chosen to be

$$\mathring{D}_m(\omega) = \begin{cases} \frac{2i}{R} \frac{i^{-|m|} Y_{|m|}^{-m}\left(\frac{\pi}{2}, -\frac{\pi}{2}\right)}{\frac{\omega}{c} h_{|m|}^{(2)}\left(\frac{\omega}{c} R\right) Y_{|m|}^{-m}\left(\frac{\pi}{2}, 0\right)} & \forall |m| \leq M \\ 0 & \text{elsewhere} \end{cases}, \quad (4.27)$$

whereby (2.26a) and (2.27) were used. The driving function (3.26) for circular secondary source contours is thus finally given by

$$D_{2.5D}(\alpha, \omega) = \sum_{m=-M}^M \mathring{D}_m(\omega) e^{im\alpha}. \quad (4.28)$$

The choice of the bandlimit M is discussed below.

The Fourier coefficients $\mathring{D}_m(\omega)$ of the continuous driving function are illustrated in Fig. 4.10(a) for different frequencies for $M \rightarrow +\infty$. Fig. 4.10(b) depicts the Fourier coefficients of the discretized driving function and for $M \rightarrow +\infty$. It can be seen that for this infinite angular bandwidth, the spectral repetitions overlap and interfere and thus spatial aliasing in the strict sense occurs (Spors & Rabenstein, 2006; Spors & Ahrens, 2008a; Ahrens & Spors, 2008a).

Such an overlapping of the spectral repetitions can be avoided by limiting the angular bandwidth M (i.e. the order) of the driving function (4.27) as (Spors & Rabenstein, 2006; Spors & Ahrens, 2008a; Ahrens & Spors, 2008a)

$$M \leq \begin{cases} \frac{L}{2} - 1 & \text{for even } L \\ \frac{L-1}{2} & \text{for odd } L \end{cases}. \quad (4.29)$$

For the current setup of $L = 56$ discrete sampling points (i.e. loudspeakers), a choice of $M \leq 27$ is thus suitable. The Fourier coefficients of the continuous bandlimited driving function are depicted in Fig. 4.10(c), and those of the discretized bandlimited driving function in Fig. 4.10(d). Note that a spatial bandwidth limitation of the driving function can also be achieved by a bandwidth limitation of the desired sound field.

The properties of the spatio-temporal transfer function $\mathring{G}_m(r, \beta, \omega)$ of the secondary source have essential influence on the synthesized sound field (eq. (4.26)). When $\mathring{G}_m(r, \beta, \omega)$ suppresses the spectral repetitions of the driving function in the case of (4.29), the synthesized sound field is unaffected by the discretization. $\mathring{G}_m(r, \beta, \omega)$ is illustrated in Fig. 4.11 for $r = \frac{R}{2}$ and $r = R$ in the horizontal plane ($\beta = \frac{\pi}{2}$). It can be seen that $\mathring{G}_m(r, \beta, \omega)$ is not bandlimited so that the spectral repetitions in $\mathring{D}_{m,S}(\omega)$ are not suppressed and the synthesized sound field suffers from a reconstruction error.

Similarly to the case of spherical secondary source distributions presented in Sec. 4.3.2, a driving function with a spatial bandwidth M which satisfies (4.29) is termed *spatially narrowband* driving function (refer to Fig. 4.10(c)); a driving function with a spatial bandwidth $M \rightarrow +\infty$ is termed *spatially fullband* driving function (refer to Fig. 4.10(b)). Driving functions with $\frac{L}{2} \ll M \ll +\infty$ may be termed *spatially wideband*.

As with spherical secondary source contours, the spatial bandwidth limitation does not need to be a sharp truncation but a smooth fade-out towards higher orders may also be applied.

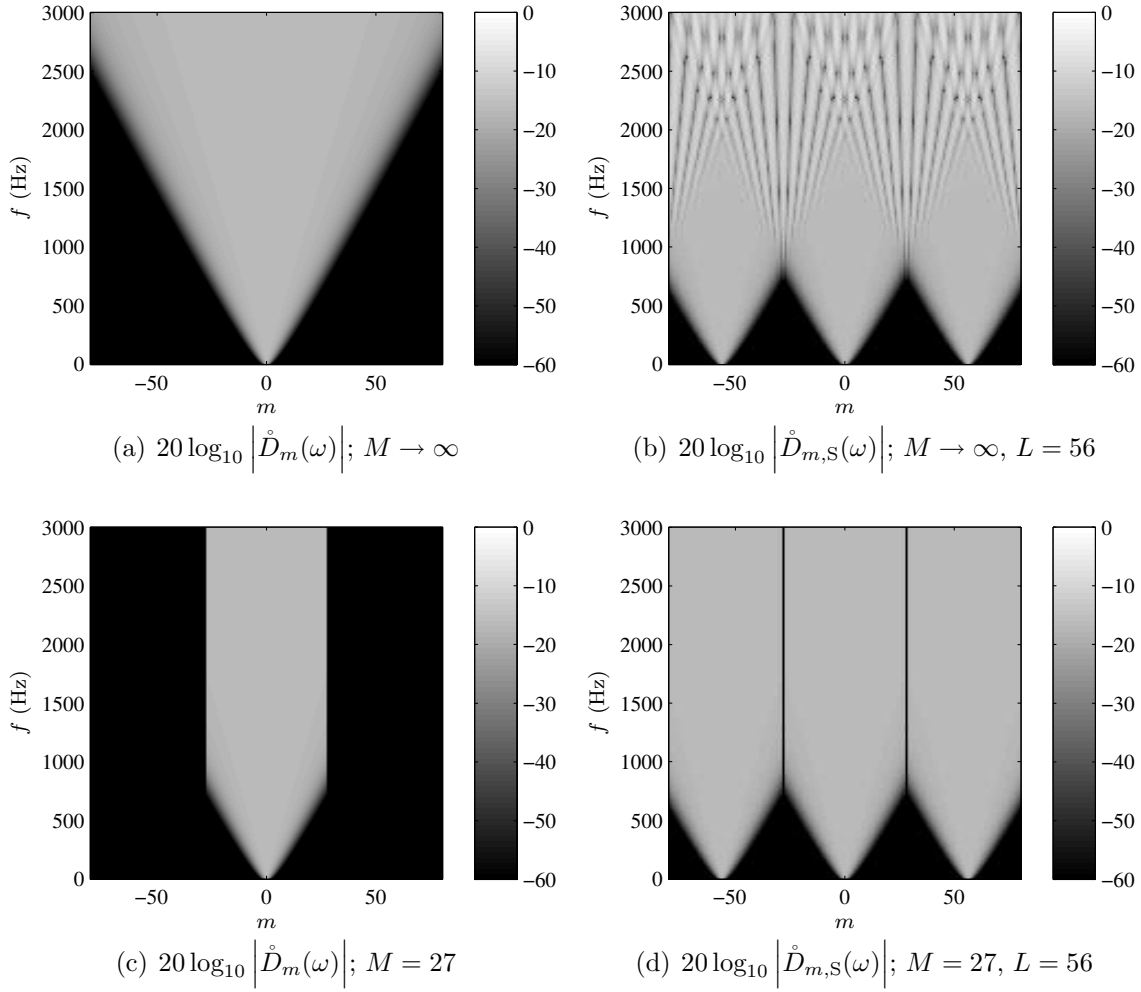


Figure 4.10: Illustration of the properties of the driving function.

4.4.2 On the Spatial Bandwidth of Wave Field Synthesis With Circular Secondary Source Distributions

Before the detailed analysis of the properties of the sound field synthesized by a discrete secondary source distribution is performed, the spatial bandwidth of WFS is investigated in order to facilitate the integration of the obtained results into previously published results on WFS such as (Start, 1997; de Bruijn, 2004; Sanson *et al.*, 2008; Wittek, 2007).

As discussed in Sec. 2.6.1, WFS constitutes a high-frequency approximation of the problem under consideration when non-planar distributions of secondary sources are considered and minor deviations from the desired sound field occur for continuous distributions. $2^{1/2}$ -dimensional WFS constitutes a further high-frequency approximation which only holds at distances to the secondary source distribution which are significantly larger than the wavelength under consideration.

As has been discussed in Sec. 4.3.2 and 4.4.1, the spatial bandwidth of the driving function is expected to have essential influence on the evolving discretization

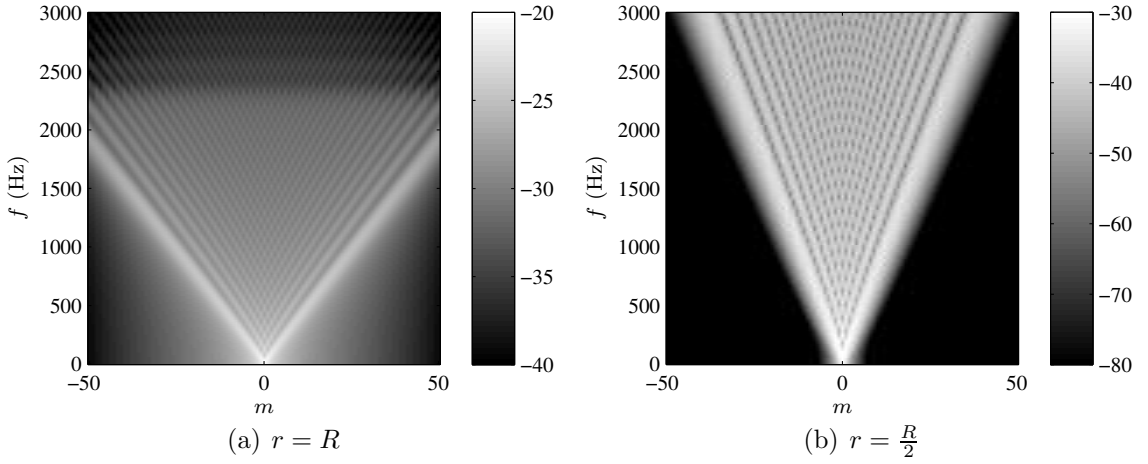


Figure 4.11: $20 \log_{10} |\dot{G}_m(r, \omega)|$. Note the different scalings of the colorbar.

artifacts. This section investigates the spatial bandwidth of WFS with enclosing secondary source distributions on the example of a circular contour. The properties of WFS with respect to spatial discretization can then be directly deduced based on the results from the analysis presented in Sec. 4.3.3, 4.4.3, and 4.4.4.

Assume a continuous circular secondary source distribution of radius R centered around the coordinate origin as depicted in Fig. 3.5. Combining (2.54) and (2.58) yields an approximation for the $2^{1/2}$ -dimensional driving function $D(\mathbf{x}_0, \omega)$ as

$$D(\mathbf{x}_0, \omega) \approx w(\mathbf{x}_0) \sqrt{\frac{2\pi y_{\text{ref}}}{i\frac{\omega}{c}}} D_{3D}(\mathbf{x}_0, \omega) . \quad (4.30)$$

In order to get an indication of the spatial bandwidth of $D(\mathbf{x}_0, \omega)$ the latter has to be transformed to the according space-frequency domain as

$$\dot{D}_m(\omega) = \frac{1}{2\pi} \int_0^{2\pi} D(\mathbf{x}_0, \omega) e^{-im\alpha} d\alpha . \quad (4.31)$$

As evident from (4.30), (4.31) constitutes a Fourier series transform over the product of two functions. As derived in App. D.1, the result of such a transform is given by a convolution of the Fourier series expansion coefficients of the two functions (eq. (D.3)). For the present case, this means

$$\dot{D}_m(\omega) = \sum_{m_1=-\infty}^{\infty} \dot{w}_{m_1} \dot{D}_{m-m_1}(\omega) . \quad (4.32)$$

Assuming a virtual plane wave with propagation direction $(\theta_{\text{pw}}, \frac{\pi}{2})$ to be synthesized, $w(\alpha)$ is given by (Spors *et al.*, 2008)

$$w(\alpha, \theta_{\text{pw}}) = \begin{cases} 1 & \text{for } \alpha - \frac{\pi}{2} \leq \theta_{\text{pw}} \leq \alpha + \frac{\pi}{2} \\ 0 & \text{elsewhere} \end{cases} . \quad (4.33)$$

The Fourier expansion coefficients $\dot{w}_m(\theta_{\text{pw}})$ of $w(\alpha, \theta_{\text{pw}})$ can be determined to be (Ahrens & Spors, 2008b)

$$\dot{w}_m(\theta_{\text{pw}}) = \begin{cases} \frac{1}{2} & \text{for } m = 0 \\ i \frac{e^{-im\theta_{\text{pw}}}}{2\pi m} (i^{-m} - i^m) & \text{for } m \neq 0 \end{cases} \quad (4.34)$$

Since $\dot{w}_m(\theta_{\text{pw}})$ never vanishes the result of the convolution in (4.32) and thus the WFS driving function has always infinite spatial bandwidth. This proves that WFS constitutes fullband synthesis. Similar results can be obtained for spherical secondary source distributions and other non-planar and non-linear geometries.

4.4.3 Properties of the Synthesized Sound Field in Time-Frequency Domain

For convenience, only the interior domain is considered in the following. The sound field $S_S(\mathbf{x}, \omega)$ which is synthesized by a discrete circular secondary source distribution as described in Sec. 4.4.1, can be calculated by inserting (4.25) into (3.23) and composing $S_S(\mathbf{x}, \omega)$ from its Fourier coefficients $\dot{S}_S(r, \beta, \omega)$ as indicated in (2.23). Exchanging then the order of summation yields

$$S_S(\mathbf{x}, \omega) = 2\pi R \sum_{n=0}^{\infty} \sum_{m=-n}^n \dot{D}_{m,S}(\omega) \check{G}_n^m(\omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha) \quad (4.35)$$

From (4.25) and (4.35) and the simulations depicted in Fig. 4.12 it can be deduced that

- As outlined in Sec. 4.4.1, $\dot{D}_{m,S}(\omega)$ is never bandlimited. Thus, $S_S(\mathbf{x}, \omega)$ always exhibits infinite bandwidth.
- If a narrowband driving function is chosen, $\dot{D}_{m,S}(\omega) = \dot{D}_m(\omega)$ holds for all $|m| \leq M$, so that the lower orders $n \leq M$ stay uncorrupted (the summation over m is bounded to the interval $[-n, n]$) (Spors & Ahrens, 2008a; Ahrens & Spors, 2008a). A region of nearly artifact-free synthesis arises around the center of the secondary source distribution. Recall that the lower orders typically describe the sound field around the center of the expansion (refer to Sec. 2.2.2). This region is bounded by a circle of radius r_M . Fig. 4.12(a), 4.12(c), and 4.12(e) depict this case for different time frequencies. For low frequencies, the r_M -limit fills the entire interior domain (Fig. 4.12(a)) and gets smaller proportional to the frequency (Fig. 4.12(c)), 4.12(e)). The higher orders of the synthesized sound field – and thus locations beyond r_M – are corrupted since the properties of the secondary sources do not perfectly suppress the spectral repetitions (Sec. 4.4.1). The energy of the artifacts outside r_M is not evenly distributed and regions arise with an amplitude with several dB below that of the desired component (e.g. around position $\mathbf{x} = [0.5 \ -1 \ 0]^T$ m in Fig. 4.12(e) and Fig. 4.13(a)). The location of these regions of significantly lower amplitude is dependent on the frequency. The arising artifacts can be locally interpreted as plane wave fronts with different propagation direction than the desired virtual plane wave (e.g. around position $\mathbf{x} = [1 \ -0.5 \ 0]^T$ m in Fig. 4.12(e)).

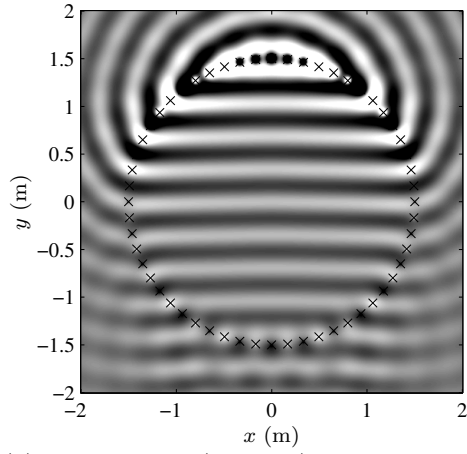
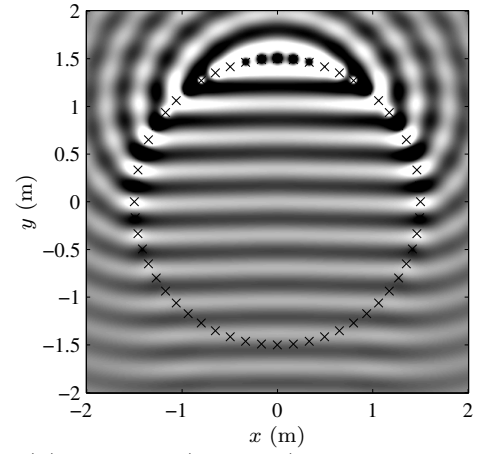
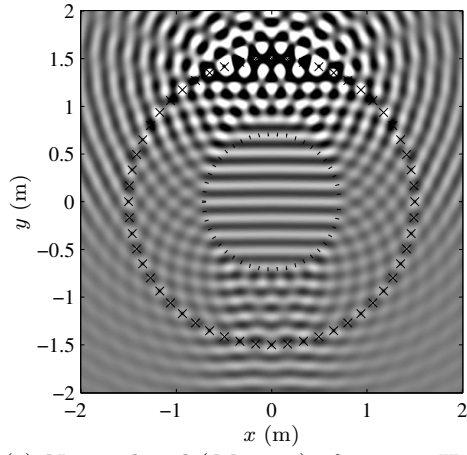
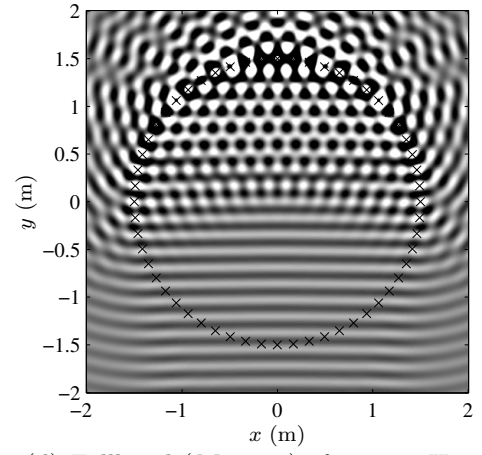
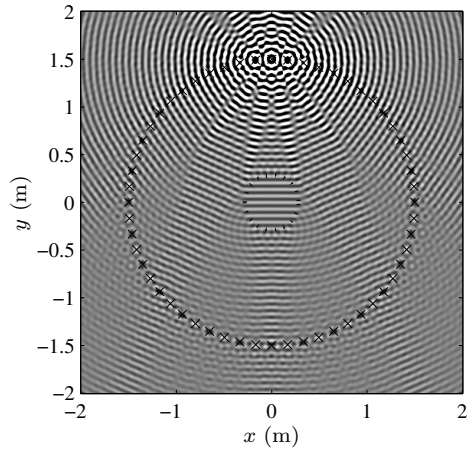
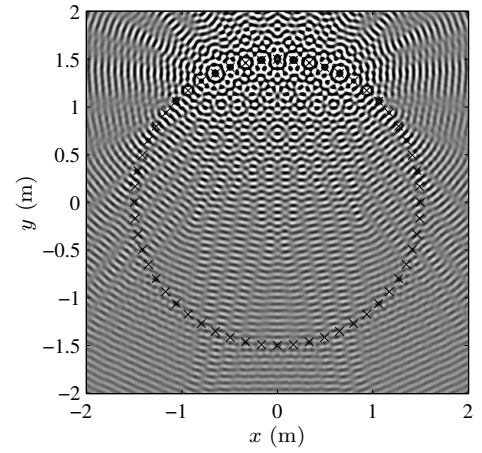
(a) Narrowband ($M = 27$), $f = 1000$ Hz(b) Fullband ($M \rightarrow \infty$), $f = 1000$ Hz(c) Narrowband ($M = 27$), $f = 2000$ Hz(d) Fullband ($M \rightarrow \infty$), $f = 2000$ Hz(e) Narrowband ($M = 27$), $f = 5000$ Hz(f) Fullband ($M \rightarrow \infty$), $f = 5000$ Hz

Figure 4.12: Synthesized sound field in the horizontal plane for the synthesis of a virtual plane wave for different bandwidths of the driving function. The marks indicate the positions of the secondary sources. The dotted circle bounds the r_M region in the narrowband case. $L = 56$ secondary sources are employed.

- A fullband driving function also allows for a synthesis which is free of considerable artifacts at lower frequencies as shown in Fig. 4.12(b). This is due to the fact that no considerable energy from the spectral repetitions leaks into the lower orders at lower frequencies (refer also to Fig. 4.10(b)).
At higher frequencies also the lower orders are corrupted and artifacts are distributed over the entire receiver area. The spatial structure of the arising artifacts can not be interpreted. The overall amplitude of the resulting sound field is more balanced over the entire receiver area than with narrowband synthesis (compare Fig. 4.13(a) to Fig. 4.13(b)).
- Evaluating (4.35) exclusively for $\mu = 0$ represents the desired component of the synthesized sound field. All cases of $\mu \neq 0$ represent discretization artifacts. Thorough inspection of Fig. 4.13(a) suggests that spatial discretization artifacts can be beneficial in narrowband synthesis since such artifacts provide energy in regions which would exhibit very low amplitude if discretization artifacts were absent. The latter circumstance is also referred to as *friendly aliasing* (Zotter & Pomberger, 2010).
Note that in Fig. 4.13(a) only that “ray” of the synthesized sound field which passes the center belongs to the desired sound field. All other components are due to spatial discretization.
- Finally, note that if only the horizontal plane is considered, circular secondary source distributions are capable of achieving results which are comparable to those of spherical secondary source distributions (Fig. 4.8) with a fraction of the number of loudspeakers (1568 vs. 56).

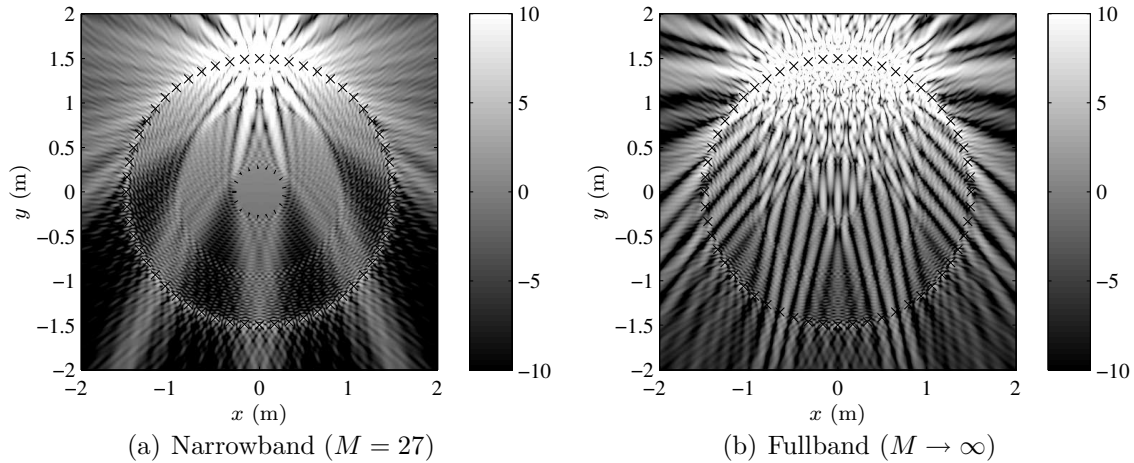


Figure 4.13: Magnitude $20 \log_{20} |S_S(\mathbf{x}, \omega)|$ of the sound fields depicted in Fig. 4.12(e) and Fig. 4.12(f); The dotted circle bounds the r_M region in the narrowband case. $f = 5000$ Hz.

Fig. 4.14 depicts the magnitude of the transfer function of the discrete secondary source distribution for different receiver positions. The scenario considered is the same as in Fig. 4.12. Fig. 4.14(a) and 4.14(c) show the transfer function for receiver

points which are distributed over the entire interior domain. Fig. 4.14(b) and 4.14(d) show positions which are in the vicinity of each other. It can be seen that:

- For the narrowband driving function, the transfer function is indeed perfectly flat at the center of the secondary source distribution (see the black line in Fig. 4.14(a)).
- Other positions in the narrowband scenario show some minor deviations of the transfer function from a perfectly flat response at low frequencies (Fig. 4.14(a) and 4.14(b)). Above a frequency of approximately 2000 Hz strong deviations from the flat response with wide gaps and peaks arise (Fig. 4.14(a)). The transfer function exhibits very little local variation (Fig. 4.14(b)).
- In the fullband examples (Fig. 4.14(c) and 4.14(d)), the transfer functions exhibits as well minor deviations from the perfectly flat response below 1000 Hz for all receiver positions.
- Above approximately 1500 Hz, narrow peaks and gaps arise with the fullband driving function with large global variation (Fig. 4.14(c)). Strong local variation is also apparent in Fig. 4.14(d) above a few kHz. This large local variation has already been detected in WFS (Wittek, 2007). Since the perceived sound coloration is significantly less than the simulations suggest, it is assumed that some kind of averaging takes place in the human auditory system which evens out the transfer function (Wittek, 2007). Note that this interpretation does not hold for narrowband synthesis as shown in Fig. 4.14(b). It has therefore to be expected that the latter case leads to significantly stronger coloration than fullband synthesis.
- For all receiver positions in fullband synthesis, the transfer function exhibits a highpass behavior with a slope of approximately 3 dB per octave above 1500 Hz. Since this slope is similar for all receiver positions it can be compensated for by an appropriate pre-filtering of the input signal. This general compensation for the highpass slope is a standard method in WFS (Spors & Ahrens, 2010a).

4.4.4 Properties of the Synthesized Sound Field in Time Domain

The analyses presented in Sec. 4.4.3 revealed the spectral characteristics of spatial discretization artifacts. It has recently been shown in (Geier *et al.*, 2010) that the time-domain characteristics of spatial discretization artifacts in synthetic sound fields can have essential influence on perception. In the time domain, such artifacts can occur as correlated signals arriving before (pre-echoes) or after (echoes) the desired wave front. So far, pre-echoes have only be observed in the synthesis of focused virtual sound sources in WFS (Spors *et al.*, 2009). Echoes have been observed in the synthesis of virtual point sources in WFS (Vogel, 1993).

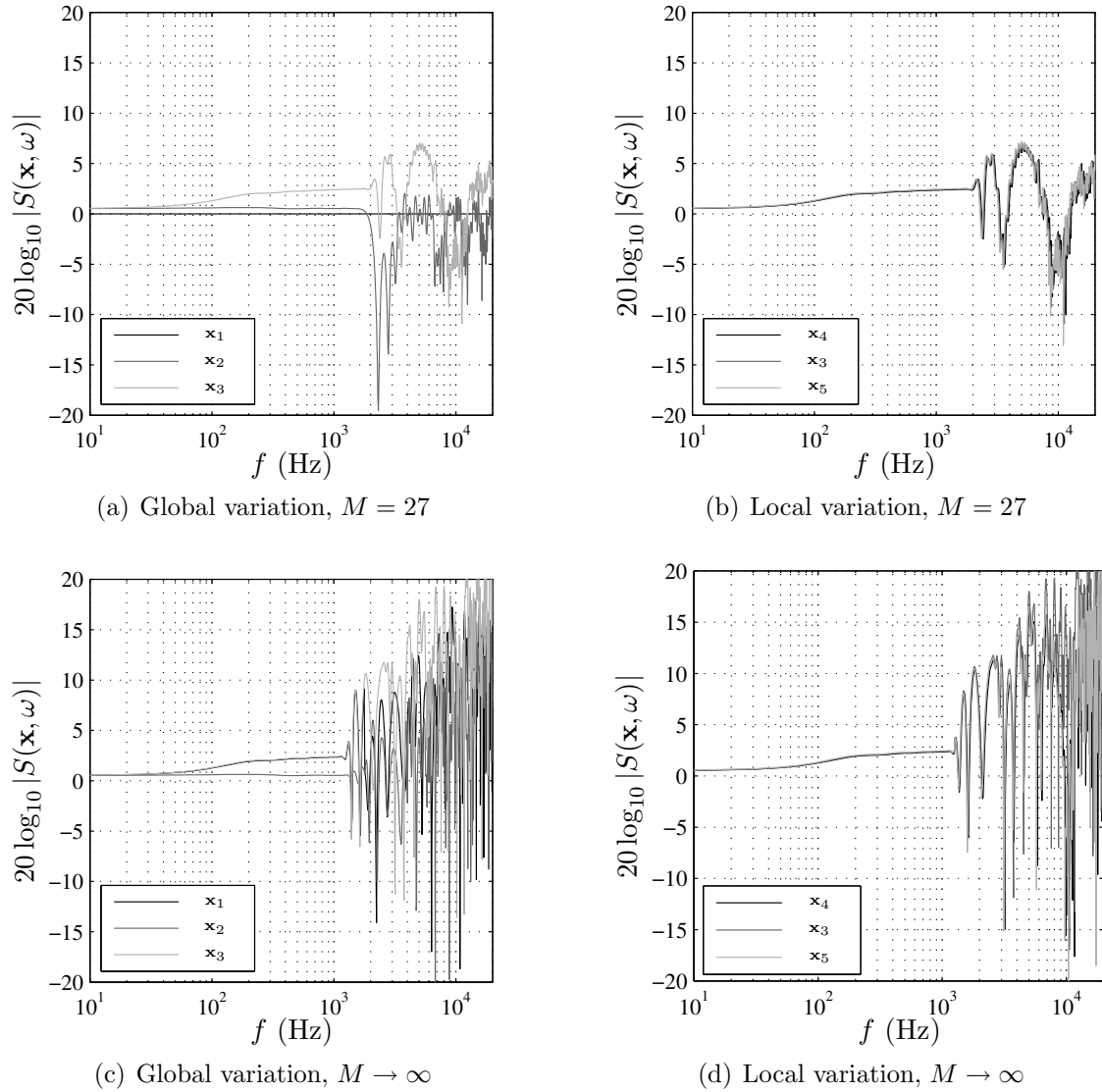


Figure 4.14: Transfer function of a circular distribution of 56 monopoles driven in order to synthesize a virtual plane wave for different listening positions. $\mathbf{x}_1 = [0 \ 0 \ 0]^T$, $\mathbf{x}_2 = [0.7 \ 0 \ 0]^T$ m, $\mathbf{x}_3 = [0 \ 0.7 \ 0]^T$ m, $\mathbf{x}_4 = [0 \ 0.69 \ 0]^T$ m, $\mathbf{x}_5 = [0 \ 0.71 \ 0]^T$ m

Since an analytical treatment in time domain is not straightforward, the sample scenario considered in Sec. 4.4.3 is numerically transferred to time domain and the result is analyzed below.

Note that time-domain simulations of WFS (and thus of fullband synthesis, Sec. 4.4.2) have also been presented in the classical WFS literature such as (Vogel, 1993) and simulations of NFC-HOA have been presented in (Daniel, 2003). However, detailed analysis and comparison have not been performed.

The critical property of the human auditory system to mention at this point is the *precedence effect* which is a fundamental mechanism in spatial hearing (Wallach *et al.*, 1949; Haas, 1951; Blauert, 1997). The precedence effect describes the phenomenon that the direction of a perceived sound is not altered by echoes of this

sound which may arrive from different directions in a time window of 1–40 ms after the leading wave front. Furthermore, the echoes are not perceived as such but as a room impression, so that in the time window of 1–40 ms fusion to one auditory percept occurs. In the case of sound field synthesis the possibility hence exists that the spatial discretization artifacts have no influence on the perceived direction of the auditory event and are not perceived as echoes. This means also that pre-echoes are more critical than echoes, because they arrive before the desired wave front and can influence the perceived direction due to triggering of the precedence effect.

On the other hand, the precedence effect only occurs if the relative level of the echoes occurring after the leading wave front is not higher than 10–15 dB. Thus, if the amplitude of the desired wave front is much higher than the amplitudes of the pre-echoes, the pre-echoes will be perceived as an additional auditory event.

Fig. 4.15 shows still images of the spatio-temporal impulse response of the loudspeaker system under consideration when driven in order to synthesize a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ for different time instances. A cross-section through the horizontal plane is shown. Fig. 4.15(a), 4.15(c), and 4.15(e) show narrowband synthesis, Fig. 4.15(b), 4.15(d), and 4.15(f) show fullband synthesis.

Fig. 4.16 shows impulse responses of the loudspeaker system for a specific listening position in narrowband synthesis (Fig. 4.16(a) and 4.16(c)) and fullband synthesis (Fig. 4.16(b) and 4.16(d)). Fig. 4.16(c) and 4.16(d) show the impulse responses from Fig. 4.16(a) and 4.16(b) respectively but lowpass and highpass filtered with cutoff frequencies f_{cutoff} as indicated. In all figures the absolute value of the sound pressure is shown in dB, i.e.

$$20 \log_{10} |\Re\{s_S(\mathbf{x}, t)\}| . \quad (4.36)$$

The time t is chosen such that the virtual plane wave front passes the center of the loudspeaker array at $t = 0$ ms.

As described above, the major findings which can be deduced from time domain simulations are the properties of the first arriving wave fronts and the occurrence of additional and correlated wave fronts (echoes). The latter are a consequence of the chosen spatial bandwidth of the driving function in combination with the fact that a finite number of spatially discrete loudspeakers is employed.

As outlined in Sec. 4.4.3, considerable artifacts have to be expected above a given time frequency f_a . In fullband synthesis, f_a is approximately constant over the entire listener area. For the present setup it lies between $f_a = 1400$ Hz and $f_a = 2500$ Hz depending on the receiver position.

This situation is more complicated in narrowband synthesis. There, it is such that an almost artifact-free region evolves around the center of the secondary source distribution which gets smaller with frequency. For frequencies below 1400 Hz, this artifact-free region fills the entire receiver area and reaches the size of a human head at approximately 10 kHz for the present loudspeaker array.

In the following, the observations deduced from the illustrations in Fig. 4.15 and Fig. 4.16 are summarized and interpreted in terms of perception.

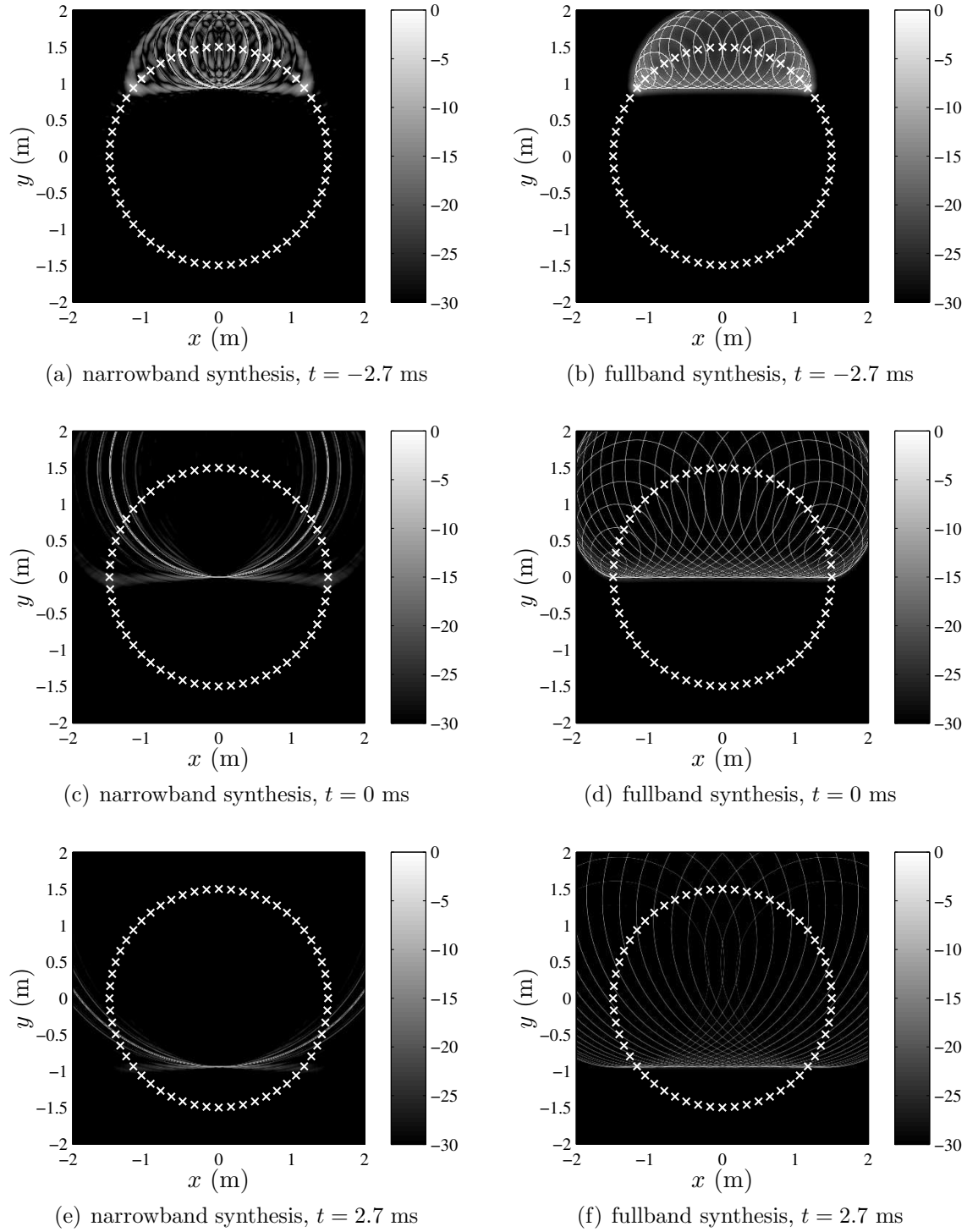


Figure 4.15: Impulse responses of the loudspeaker system in the horizontal plane when driven in order to synthesize a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$. The absolute value of the time domain sound pressure is shown in dB for different instances of time. The left column shows narrowband synthesis, the right column shows fullband synthesis. The marks indicate the positions of the loudspeakers.

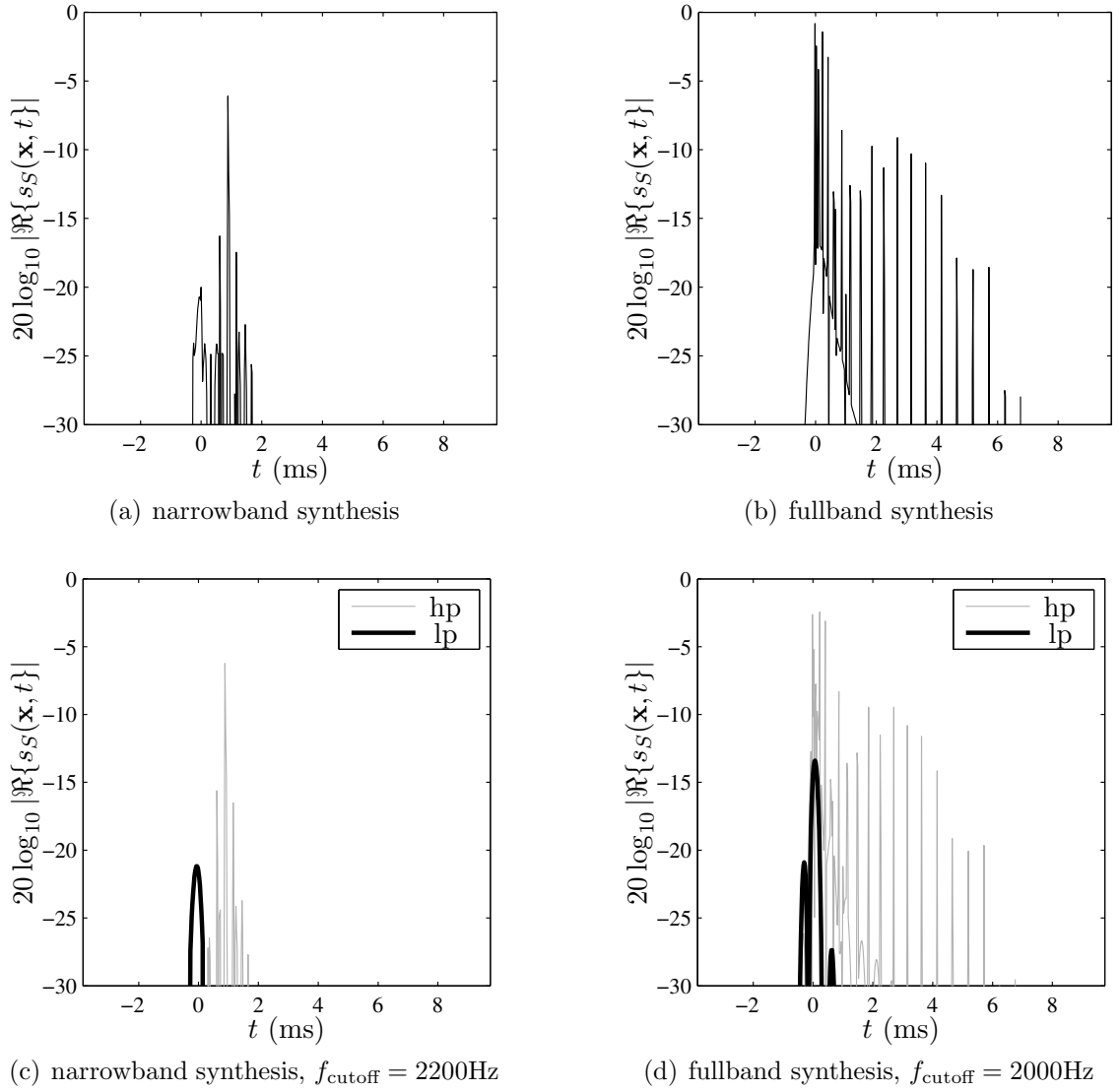


Figure 4.16: Impulse responses of the loudspeaker system at position $\mathbf{x} = [1 \ 0 \ 0]^T$ m when driven in order to synthesize a virtual plane wave with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (-\frac{\pi}{2}, \frac{\pi}{2})$. Fig. 4.16(c) and 4.16(d) show the impulse responses from Fig. 4.16(a) and 4.16(b) but highpass ('hp') and lowpass ('lp') filtered with a cutoff frequency of f_{cutoff} . The absolute value of the sound pressure is shown in dB. The plane wave passes the center of the array at $t = 0$ ms with amplitude 0 dB.

First wave front and echoes

Fullband synthesis exhibits a pronounced first wave front at all listening positions. Above f_a , this first wave front is slightly distorted but keeps its straight shape. Spatial discretization artifacts in the form of high-frequency echoes follow the first wave front for all listening positions above f_a . As pointed out in (Berkhout *et al.*, 1993), WFS (which is fullband synthesis, Sec. 4.4.2) can be seen as *wave front synthesis*.

The broadband first wave front is followed by a dense sequence of echoes of approximately similar amplitude for $0 \text{ ms} < t < 0.2 \text{ ms}$ (refer to Fig. 4.15(d)). This dense sequence is followed by a slightly sparser sequence of high-frequency echoes for $0.2 \text{ ms} < t < 6 \text{ ms}$ with decreasing amplitude. The time interval between successive echoes in the sparser part of the impulse response is in the order of some hundred μs . These high-frequency echoes arrive from various directions and are rather homogeneously distributed over the entire receiver area. It can be shown that each of the active loudspeakers produces one of these echoes (Vogel, 1993). Consequently, larger loudspeaker setups lead to longer impulse responses and a larger loudspeaker spacing leads to longer intervals between the echoes.

In narrowband synthesis the plane wave front is accurately synthesized around the central listening position (refer to Fig. 4.15(c)). At other listening positions, especially at positions lateral to the center, the synthesized sound field consists of a number of echoes which impinge at different times and from different directions on the listener.

Comparison of Fig. 4.15(c) with the simulations in Fig. 4.12(a), 4.12(c), and 4.12(e) reveals that the first wave front arriving carries the low time-frequency content. This is also confirmed by the impulse response of the narrowband scenario, as depicted in Fig. 4.16(c). The thick black curve represents energy below $f_{\text{cutoff}} = 2200 \text{ Hz}$, the thin gray curve represents energy above f_{cutoff} . The virtual plane wave is accurately synthesized at these low time frequencies whereby it exhibits a slightly concave shape containing some distortion for positions lateral to the center.

After the first wave front, a number of echoes arrive successively from a direction which approximately coincides with the direction of that loudspeaker at which the virtual plane wave first “touches” the loudspeaker contour. Comparison of Fig. 4.15(c) with monochromatic simulations in Fig. 4.12(a), 4.12(c), and 4.12(e) reveals that these echoes contain high time frequencies. Again, this is confirmed by Fig. 4.16(c). Note that the amplitude of the loudest echo is at almost 15 dB above the first wave front (Fig. 4.16(a)). The distance in time between the adjacent wave fronts is significantly lower than 1 ms for the loudspeaker system under consideration. A wider loudspeaker spacing leads to a larger distance between the wave fronts.

It is evident from inspection of Fig. 4.15 (and Fig. 4.16(a) and 4.16(b)) that the impulse response of the system is significantly shorter for narrowband synthesis than for fullband synthesis for a given listener position. While no considerable energy is present at all positions for $y > 0 \text{ m}$ in narrowband synthesis for $t = 2.7 \text{ ms}$ in Fig. 4.15(c) the discretization artifacts in fullband synthesis are still obvious (Fig. 4.15(d)).

Recall finally that, as explained in Sec. 4.4.3, the energy distribution over the entire receiver area is very inhomogeneous for frequencies above f_a . At certain locations dependent on the considered frequency, the synthesized sound field exhibits a significantly lower amplitude than desired.

Perception

The accurate synthesis of the first wave front in fullband synthesis leads to very good auditory localization for non-focused virtual sources over the entire listening

area (Start, 1997; de Bruijn, 2004; Sanson *et al.*, 2008). This is in accordance with the conclusion that was drawn above based on the precedence effect. The high-frequency echoes due to spatial discretization are not perceivable as echoes nor do they change the perceived direction of the virtual plane wave. Recall that the echoes arrive in a time window smaller than 6 ms, are lower in amplitude, and contain fewer spectral components than the first wave front. Informal listening confirms absence of perceivable echoes, but the echoes do add some sense of spaciousness. This is another well-known phenomenon of the precedence effect and enables humans to properly localize auditory events in non-anechoic environments (Blauert, 1997). Due to the unnatural pattern of echoes and the corresponding comb filtering of the transfer function also slight coloration is perceivable.

For narrowband synthesis a separation in time between the wave fronts for low and high frequencies takes place. Therefore, there exists no spectral overlap between the first wave front and the later echoes. This leads to a weaker precedence effect (Litovsky *et al.*, 1999). Also the high-frequency echoes are 15 dB higher in amplitude than the first wave front. This suggests that the high time-frequency content of the synthesized sound field is localized in direction of the loudspeakers producing these echoes (see above). This is in contrast to the low frequency content which impinges from the desired direction.

Informal listening shows that high and low time-frequency contents are indeed localized at different directions for listening positions lateral to the center. The auditory event is thus split into two. One event is composed exclusively of the high time-frequency content, the other event is composed of the low time-frequency content.

In general, it is expected that narrowband synthesis provides a less homogeneous perception than fullband synthesis when the entire listening area is considered. On the other hand, at the center of the loudspeaker array, narrowband synthesis is expected to cause less coloration than fullband synthesis due to the absence of any echoes at this location in narrowband synthesis.

4.4.5 Optimizing the Synthesis with Respect to a Given Receiver Location

It was shown in Sec. 4.4.3 that the synthesis of a sound field which is bandlimited according to (4.29) leads to a region around the center of the secondary source distribution which is free of considerable discretization artifacts. It will be shown in this section that a bandlimitation with respect to an expansion around any given location inside the area surrounded by the secondary source distribution does indeed also lead to such a region of high physical accuracy at the according location (Ahrens & Spors, 2009b).

This approach is only presented for circular secondary source distributions and not for spherical ones since the situation is similar for both geometries.

Limiting the Spatial Bandwidth with Respect to a Local Coordinate System

Limiting the spatial bandwidth of a sound field $S(\mathbf{x}, \omega)$ with respect to an expansion in a local coordinate system with origin at the global coordinate \mathbf{x}_c yields (Ahrens & Spors, 2009b)

$$S_{N'}(\mathbf{x}', \omega) = \sum_{n'=0}^{N'-1} \sum_{m'=-n'}^{n'} \check{S}'_{n'}{}^{m'}(\omega) j_{n'}\left(\frac{\omega}{c} r'\right) Y_{n'}^{m'}(\beta', \alpha'), \quad (4.37)$$

whereby $N' - 1$ denotes the local angular bandwidth. Again, the spatial bandwidth limitation does not need to be a sharp truncation but a smooth fade-out towards higher orders may also be applied. For simplicity, sharp truncation is applied.

r' and α' denote the position coordinates with respect to a local coordinate system whose origin is at $\mathbf{x}_c = [x_c \ y_c \ 0]^T$ and which is obtained by a translation of the global coordinate system. A similar situation is depicted in Fig. 3.7 whereby in the present case, \mathbf{x}_c is not necessarily on the x -axis. Note that $r' = r'(\mathbf{x})$ and $\alpha' = \alpha'(\mathbf{x})$.

For the calculation of the driving function (4.28) the coefficients $\check{S}_{|m|}^m(\omega)$ with respect to expansion in the global coordinate system are required. The expansion (4.37) has therefore to be expressed in the global coordinate system. Similar to (E.3), this translation is given by

$$S_{N'}(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\sum_{n'=0}^{N'-1} \sum_{m'=-n'}^{n'} \check{S}'_{n'}{}^{m'}(\omega) (-1)^{n+n'} (I|I)_{nn'}^{mm'}(\Delta\mathbf{x}, \omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m(\beta, \alpha)}_{= \check{S}_n^m(\omega)}, \quad (4.38)$$

so that the coefficients $\check{S}_{|m|}^m(\omega)$ required by the driving function are given by

$$\check{S}_{|m|, N'}^m(\omega) = \sum_{n'=0}^{N'-1} \sum_{m'=-n'}^{n'} \check{S}'_{n'}{}^{m'}(\omega) (-1)^{n+n'} (I|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega). \quad (4.39)$$

Two spatial bandlimitations are apparent in the driving function: (Ahrens & Spors, 2009b)

1. $S_{N'}(\mathbf{x}, \omega)$ is bandlimited with respect to an expansion around \mathbf{x}_c . The bandlimit is denoted by N' . From (4.38) it can be deduced that $S_{N'}(\mathbf{x}, \omega)$ nevertheless exhibits infinite spatial bandwidth with respect to expansion around the global coordinate origin.
2. The driving function (4.28) on the other hand is bandlimited with respect to expansion around the coordinate origin. This bandlimit is denoted by M . The desired component of the synthesized sound field is bandlimited in both senses.

Spatial discretization properties

The spatial bandwidth limitation introduced in (4.37) leads to favorable spatial discretization properties as described in this section. For convenience, the synthesis of

a virtual plane wave with propagation direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (-\frac{\pi}{2}, \frac{\pi}{2})$ is considered. The coefficients $\check{S}'_{n'}(\omega)$ in this case correspond to the coefficients $\check{S}_{n,\text{pw}}^m(\omega)$ given by (2.27).

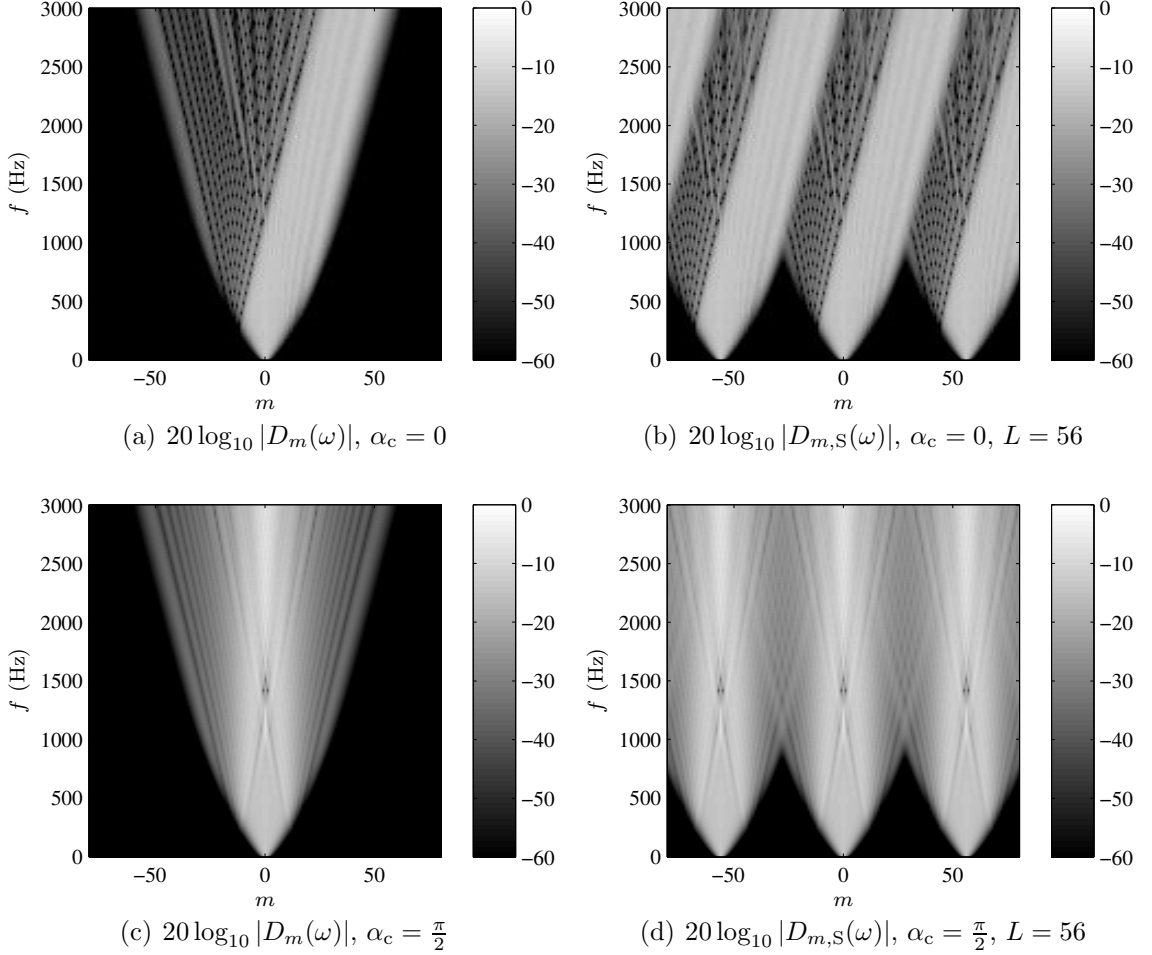


Figure 4.17: Magnitude of the Fourier coefficients with respect to the expansion around the origin of the global coordinate system. $r_c = 0.75$ m, $N' = 15$

In Fig. 4.17 it can be seen that the energy of the angular spectrum $\mathring{D}_m(\omega)$ of the continuous proposed driving function is distributed such that the spectral repetitions due to spatial sampling overlap only in regions of low energy. This enables the application of a driving function (4.28) with a bandlimit M significantly higher than M in the narrowband case still avoiding considerable overlap (Ahrens & Spors, 2009b). Generally, a choice $M \rightarrow +\infty$ will be made which leads to a *locally bandlimited fullband driving function*. Since the spectral repetitions do nevertheless introduce considerable energy into the lower orders of the driving function, the synthesized sound field will suffer from considerable spatial aliasing and other reconstruction errors. Since no interference of the high-energy regions occurs, spatial aliasing and the reconstruction errors evolve in spatial locations at significant distance from the local expansion center.

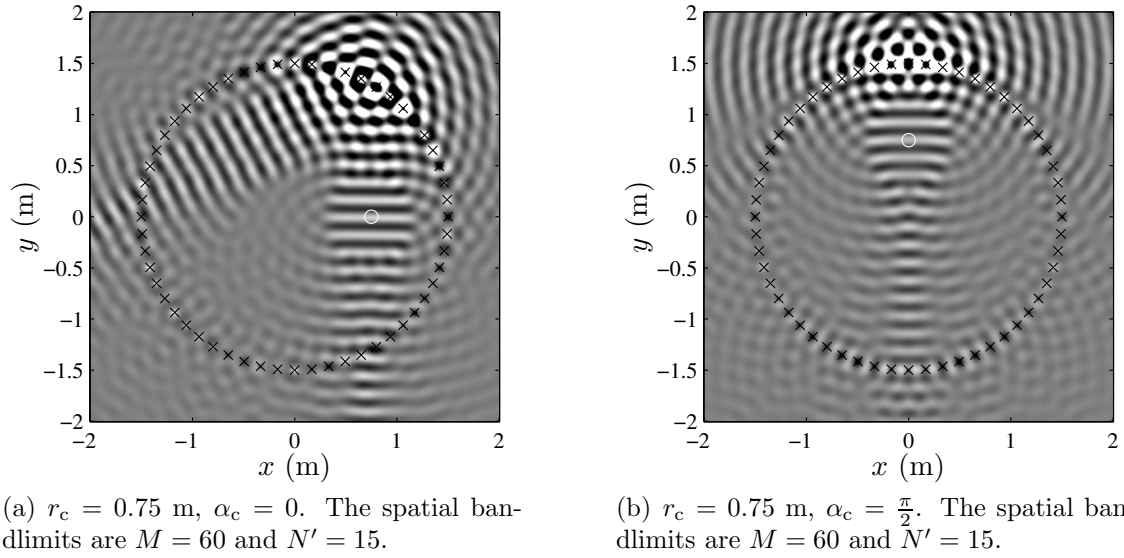


Figure 4.18: Sound fields synthesized by a circular distribution of $L = 56$ discrete loudspeakers and with radius $R = 1.5$ m synthesizing a plane wave of frequency $f = 2000$ Hz with propagation direction $(\theta_{pw}, \phi_{pw}) = (-\frac{\pi}{2}, \frac{\pi}{2})$. The black marks indicate the positions of the secondary sources; white circles indicate \mathbf{x}_c , the centers of the local coordinate systems.

Two examples of the application of the proposed driving function are shown in Fig. 4.18. It can be seen that regions of high accuracy do indeed evolve around the expansion centers \mathbf{x}_c marked by the white circles. These regions have a radius of $r_{N'}$.

Outside these regions, strong deviations from the desired sound field arise. Similar like with the conventional driving function, the regions of increased accuracy become smaller with increasing time frequency of the synthesized sound field. When comparing Fig. 4.18(b) to the application of the conventional narrowband and full-band driving functions illustrated in Fig. 4.12(c) and 4.12(d), it can be seen that the locally bandlimited approach indeed enables the accurate synthesis of the desired sound field in locations where the conventional approach fails to do so. The synthesis can thus be optimized with respect to a given – potentially dynamic – target area.

This approach is referred to as *local sound field synthesis*.

Efficient Implementation

For the efficient implementation of (potentially dynamic) local sound field synthesis, a reformulation of the coordinate translation from Sec. 3.3.3 for interior-to-interior ($(I|I)$) translation can be employed (Gumerov & Duraiswami, 2004). All relations for the coefficients $(E|I)$ given in Sec. 3.3.3 hold on a similar manner for $(I|I)$ whereby initial values (Gumerov & Duraiswami, 2004, eq. (3.2.9), p. 96)

$$(I|I)_{n'0}^{m'0}(\Delta\mathbf{x}, \omega) = \sqrt{4\pi} (-1)^{n'} j_{n'}\left(\frac{\omega}{c}\Delta R\right) Y_{n'}^{-m'}(\Delta\beta, \Delta\alpha) \quad (4.40)$$

and (Gumerov & Duraiswami, 2004, eq. (3.2.52), p. 103)

$$(I|I)_{0|m|}^{0m}(\Delta\mathbf{x}, \omega) = \sqrt{4\pi} j_{|m|} \left(\frac{\omega}{c} \Delta r \right) Y_{|m|}^m(\Delta\beta, \Delta\alpha) . \quad (4.41)$$

have to be employed instead of (3.33) and (3.34).

4.5 Planar Secondary Source Distributions

An infinite discrete planar secondary source array of constant spacing between adjacent secondary sources of Δx and Δz in x - and z -direction respectively is considered in this section. The spatial discretization is modeled by a sampling of the driving function as (Spors, 2006)

$$D_S(x, z, \omega) = \sum_{\eta=-\infty}^{\infty} \delta(x - \Delta x \eta) \sum_{\nu=-\infty}^{\infty} \delta(z - \Delta z \nu) \cdot D(x, z, \omega) . \quad (4.42)$$

Similarly like in (4.7), it can be shown that $\tilde{D}_S(k_x, k_z, \omega)$ is then (Ahrens & Spors, 2010d)

$$\tilde{D}_S(k_x, k_z, \omega) = \sum_{\eta=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} \tilde{D} \left(k_x - \frac{2\pi}{\Delta x} \eta, k_z - \frac{2\pi}{\Delta z} \nu, \omega \right) , \quad (4.43)$$

and spectral repetitions in k_x - k_z -domain become apparent. According to (3.39), the synthesized sound field \tilde{S}_S is given by

$$\tilde{S}_S(k_x, y, k_z, \omega) = \tilde{D}_S(k_x, k_z, \omega) \cdot \tilde{G}(k_x, y, k_z, \omega) \quad (4.44)$$

Eq. (4.44) constitutes the analogon to (4.8), (4.19), and (4.26). The adaptation of Fig. 4.3, 4.5, and 4.9 to the present situation is depicted in Fig. 4.19.

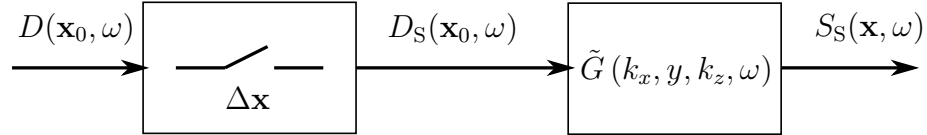


Figure 4.19: Schematic of the spatial discretization process for planar secondary source distributions.

For convenience, the example of the synthesis of a unit amplitude plane wave with propagation vector \mathbf{k}_{pw} (refer to (1.2)) is considered in the following. The sound field $S_S(\mathbf{x}, \omega)$ synthesized by a discrete secondary source distribution as described above is yielded by inserting (3.43) into (4.43), and the result and (C.9) into (3.37). Applying then an inverse Fourier transform along k_x and k_y finally yields (Ahrens & Spors, 2010d)

$$S_S(\mathbf{x}, \omega) = 2ik_{pw,y} \sum_{\eta=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} \tilde{G} \left(\frac{2\pi}{\Delta x} \eta + k_{pw,x}, y, \frac{2\pi}{\Delta z} \nu + k_{pw,z}, \omega \right) \\ \times e^{-i \left(\frac{2\pi}{\Delta x} \eta + k_{pw,x} \right) x} e^{-i \left(\frac{2\pi}{\Delta z} \nu + k_{pw,z} \right) z} \cdot 2\pi \delta(\omega - \omega_{pw}) . \quad (4.45)$$

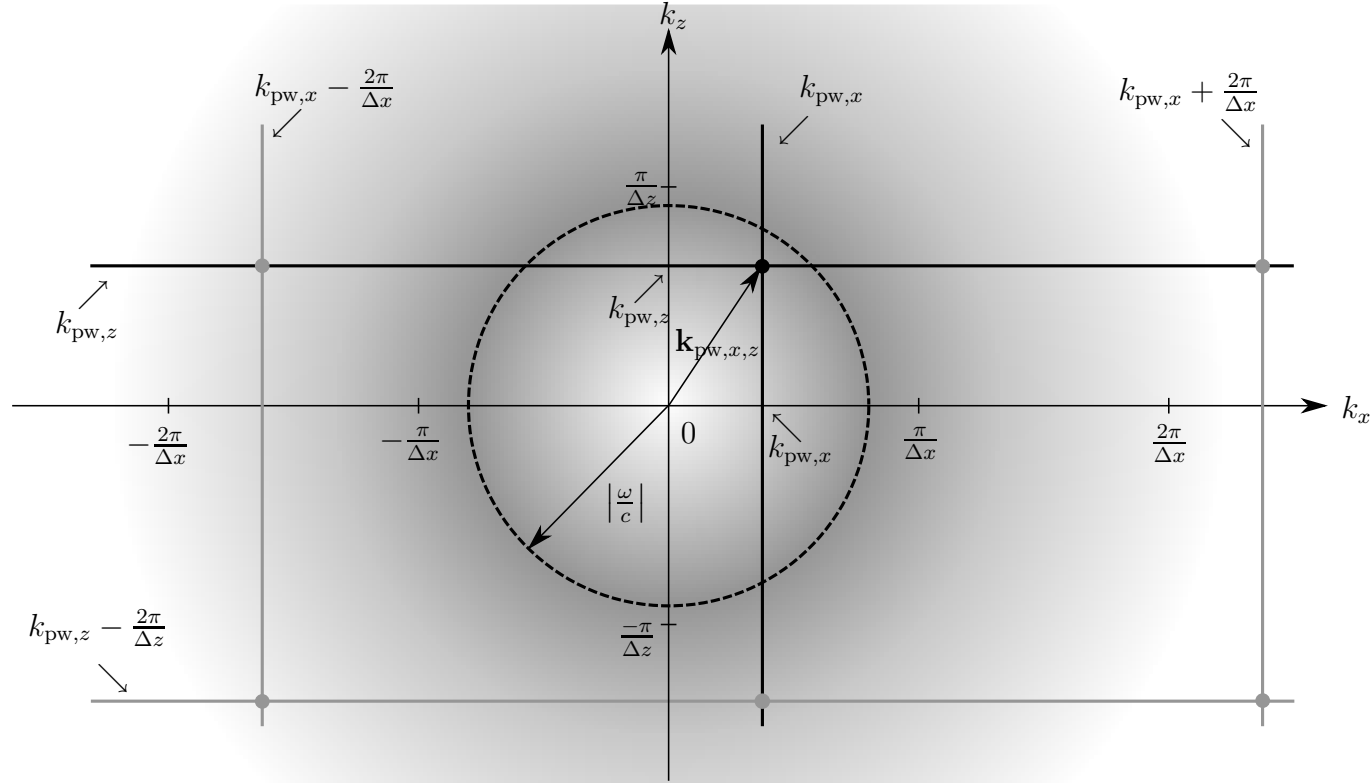


Figure 4.20: Illustration of $\tilde{G}(k_x, y, k_z, \omega)$ reflecting the properties of discrete planar secondary source distributions (eq. (4.45)). The vector

$\mathbf{k}_{pw,x,z} = [k_{pw,x} \ k_{pw,z}]^T$ represents the propagation direction of the synthesized plane wave projected onto the k_x - k_z -plane. The dots \bullet indicate synthesized components. Black solid lines and black dots represent quantities occurring with continuous secondary source distributions. Grey lines and dots represent quantities occurring additionally due to the spatial discretization. The grey shading indicates the amplitude of $\tilde{G}(\cdot)$. Locations outside the circle represent evanescent sound fields, locations inside the circle represent propagating sound fields.

$S_S(\mathbf{x}, \omega)$ is thus given by a summation over a multiplication of three factors. The latter describe the synthesized sound field along each individual dimension of space. Eq. (4.45) evaluated for $(\eta = 0, \nu = 0)$ constitutes the desired plane wave. The other terms in the sum for $\eta \neq 0$ and $\nu \neq 0$ are a consequence of spatial discretization.

For each individual order η and ν , the synthesized sound field in x - and z -direction is given by complex exponential functions. The amplitude is therefore constant along the respective dimension and the phase changes harmonically. The synthesized sound field along the y -dimension is determined by the secondary source transfer function $\tilde{G}(k_x, y, k_z, \omega)$ given by (C.9). Since $\tilde{G}(k_x, y, k_z, \omega)$ essentially determines the properties of $S_S(\mathbf{x}, \omega)$, the investigation is limited to the properties of the former (Ahrens & Spors, 2010d).

Fig. 4.20 illustrates $\tilde{G}(k_x, y, k_z, \omega)$ in the k_x - k_z -plane. For ease of illustration, a schematic is used here instead of a simulation. Note that the properties of the involved quantities are investigated more in detail in conjunction with linear secondary source distributions in Sec. 4.6.1. The essential mechanisms are similar with linear distributions but detailed illustration of the latter is more convenient due to the lesser degrees of freedom. A basic analysis is given in the following.

For a fixed time frequency ω , $k_x = \frac{2\pi}{\Delta x}\eta + k_{pw,x}$ is represented by straight lines perpendicular to the k_z -axis in Fig. 4.20. $k_z = \frac{2\pi}{\Delta z}\nu + k_{pw,z}$ is represented by straight lines perpendicular to the k_x -axis. $\tilde{G}(k_x, y, k_z, \omega)$ has a pole on a circular region of radius $\frac{\omega}{c}$ centered around the origin of the coordinate system.

The different components of $S_S(\mathbf{x}, \omega)$ are given by the intersections of the above described lines in the k_x - k_z -plane. The desired plane wave is indicated in Fig. 4.20 by the intersection of the two lines inside the circle of radius $\frac{\omega}{c}$.

Two categories of discretization artifacts can be identified: a) Evanescent components and b) propagating plane wave components which are additional to the desired one.

Artifacts belonging to category a) are illustrated in Fig. 4.20. They are represented by intersections of lines occurring at locations where $\sqrt{k_x^2 + k_z^2} > \left|\frac{\omega}{c}\right|$. It can be seen from (C.9) that $S_S(\mathbf{x}, \omega)$ is evanescent for exactly these locations. Note that the exponent in (C.9) is purely real for $\sqrt{k_x^2 + k_z^2} > \left|\frac{\omega}{c}\right|$. The existence of evanescent components in the synthesized sound field has already been indicated in (Pueo *et al.*, 2007).

Since neither η nor ν is bounded, these evanescent discretization artifacts can not be avoided. Due to the monotonically decreasing amplitude of $\tilde{G}(k_x, y, k_z, \omega)$ (indicated by the grey shading in Fig. 4.20) for $\sqrt{k_x^2 + k_z^2} > \left|\frac{\omega}{c}\right|$, the higher the orders η and ν of the discretization contributions are, the lower are their amplitudes.

The discretization artifacts of category b) occur only in special situations: When the distance Δx or Δz between adjacent loudspeakers is so large, respectively if the time frequency ω is so high that lines other than those for $(\eta = 0, \nu = 0)$ intersect inside the circular region bounded by the pole of $\tilde{G}(k_x, y, k_z, \omega)$. In this case, the discretization artifacts are additional plane wave contributions whose propagation direction is determined by the location of the points of intersection and is therefore dependent on the radian frequency ω .

Note that this situation is not apparent in Fig. 4.20. For ease of clarity $\Delta x, \Delta z$,

and ω in Fig. 4.20 where chosen such that the lines for $\eta \neq 0$ and $\nu \neq 0$ only intersect outside the circular boundary between the regions of propagating and evanescent components.

A segregation of spatial aliasing in the strict sense (as explained in Sec. 4.3.3) and other reconstruction errors is not useful in the present case of planar secondary source distributions synthesizing plane waves. This is due to the fact that the spatial spectrum of the continuous driving function is given by a single delta function and thus an overlap for repetitions can not occur. Since it is practically significantly more relevant whether the arising artifacts are propagating or evanescent, the term *spatial aliasing* may be employed when propagating artifacts are considered.

It is not straightforward to derive a revealing analytical *anti-aliasing condition* for planar secondary source distributions which prevents the synthesis of unwanted propagating components. This is due to the fact that the sampling in x -dimension and the sampling in z -dimension interact and can not be treated independently. The conditions (Ahrens & Spors, 2010d)

$$\left(\frac{\omega}{c}\right)^2 < \left(\frac{2\pi}{\Delta x} - |k_{pw,x}|\right)^2 + k_{pw,z}^2 \quad (4.46a)$$

$$\left(\frac{\omega}{c}\right)^2 < k_{pw,x}^2 + \left(\frac{2\pi}{\Delta z} - |k_{pw,z}|\right)^2 \quad (4.46b)$$

both have to be met.

The concept of narrowband and wide-/fullband driving functions as it was proposed for spherical and circular secondary source distributions (Sec. 4.3.2 and 4.4.1) is not useful here since a bandwidth limitation restricts the possible propagation directions of the synthesized sound field. Recall that the driving function for the synthesis of a plane wave (3.43) consists of Dirac delta functions in k_x and k_y . A limitation of the spatial bandwidth can generally only be applied by a transposition of the delta function to lower space frequencies which results in a change of the propagation direction.

As will be shown in Sec. 4.6, the spatial sampling properties of planar and linear secondary source distributions are essentially similar. In order to avoid redundancies, detailed analyses are only presented for linear distributions in Sec. 4.6.

4.6 Linear Secondary Source Distributions

4.6.1 Discretization of the Driving Function

Applying the procedure outlined in Sec. 4.5 to linear secondary source distributions leads to a discretized driving function $\tilde{D}_S(k_x, \omega)$ given by (Spors, 2006)

$$\tilde{D}_S(k_x, \omega) = \sum_{\eta=-\infty}^{\infty} \tilde{D}\left(k_x - \frac{2\pi}{\Delta x}\eta, \omega\right), \quad (4.47)$$

and spectral repetitions in k_x -domain become apparent. According to (3.47), the synthesized sound field \tilde{S}_S is given by

$$\tilde{S}_S(k_x, y, z, \omega) = \tilde{D}_S(k_x, \omega) \cdot \tilde{G}(k_x, y, z, \omega) \quad (4.48)$$

Eq. (4.48) constitutes the analogon to (4.8), (4.19), (4.26), and (4.44). The adaptation of Fig. 4.3, 4.5, 4.9, and 4.19 to the present situation is depicted in Fig. 4.21.

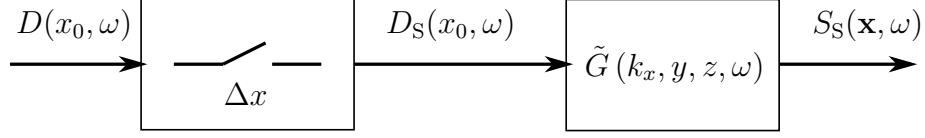


Figure 4.21: Schematic of the spatial discretization process for linear secondary source distributions.

As with planar secondary source distributions, the synthesis of a virtual plane wave propagating inside the horizontal plane is considered in the following. Inserting (3.53) into (4.47) and the result and (C.8) into (3.46) yields the synthesized sound field $S_S(\mathbf{x}, \omega)$ given by (Ahrens & Spors, 2010d)

$$S_S(\mathbf{x}, \omega) = \frac{4ie^{-ik_{pw,y}y_{ref}}}{H_0^{(2)}(k_{pw,y}y_{ref})} \cdot 2\pi\delta(\omega - \omega_{pw}) \sum_{\eta=-\infty}^{\infty} e^{-i\left(\frac{2\pi}{\Delta x}\eta + k_{pw,x}\right)x} \times \tilde{G}\left(\frac{2\pi}{\Delta x}\eta + k_{pw,x}, y, z, \omega\right). \quad (4.49)$$

Again, $S_S(\mathbf{x}, \omega)$ is given by a complex exponential function along the x -dimension. The properties of the secondary sources reflected by $\tilde{G}(k_x, y, z, \omega)$ given by (C.8) determine $S_S(\mathbf{x}, \omega)$ in radial direction, i.e. along $\sqrt{y^2 + z^2}$.

The situation for discrete linear secondary source distributions is very similar to that of discrete planar distributions discussed in Sec. 4.5: The considered region of the wavenumber space, in this case the k_x -axis, is divided into regions implying different properties of the synthesized sound field. i) Locations where $|k_x| < \left|\frac{\omega}{c}\right|$ represent a combination of propagating and evanescent sound fields, ii) locations where $|k_x| > \left|\frac{\omega}{c}\right|$ represent purely evanescent sound fields.

This finding is deduced from the properties of the secondary source transfer function $\tilde{G}(k_x, y, z, \omega)$. For $|k_x| < \left|\frac{\omega}{c}\right|$, $\tilde{G}(k_x, y, z, \omega)$ is given by the zero-th order Hankel function of second kind $H_0^{(2)}(\cdot)$ (refer to (C.8)). This indicates a combination of a propagating and an evanescent sound field (Williams, 1999). For $|k_x| > \left|\frac{\omega}{c}\right|$, $\tilde{G}(k_x, y, z, \omega)$ is given by the zero-th order modified Bessel function of second kind $K_0(\cdot)$. $K_0(\cdot)$ is purely real and decreases strictly monotonically with increasing argument, i.e. with increasing distance $\sqrt{y^2 + z^2}$ to the secondary source distribution.

Fig. 4.22 illustrates $\tilde{G}(k_x, y, z, \omega)$ in the horizontal plane, i.e. for $z = 0$ for two different distances y . The edges for the triangular structure in Fig. 4.22 correspond to $|k_x| = \left|\frac{\omega}{c}\right|$.

It can be deduced that the magnitude of $\tilde{G}(k_x, y, z, \omega)$ drops quickly when the evanescent region is entered whereby the slope is less steep closer to the secondary source, i.e. for smaller y . Obviously, evanescent components are more pronounced in the vicinity of the source.

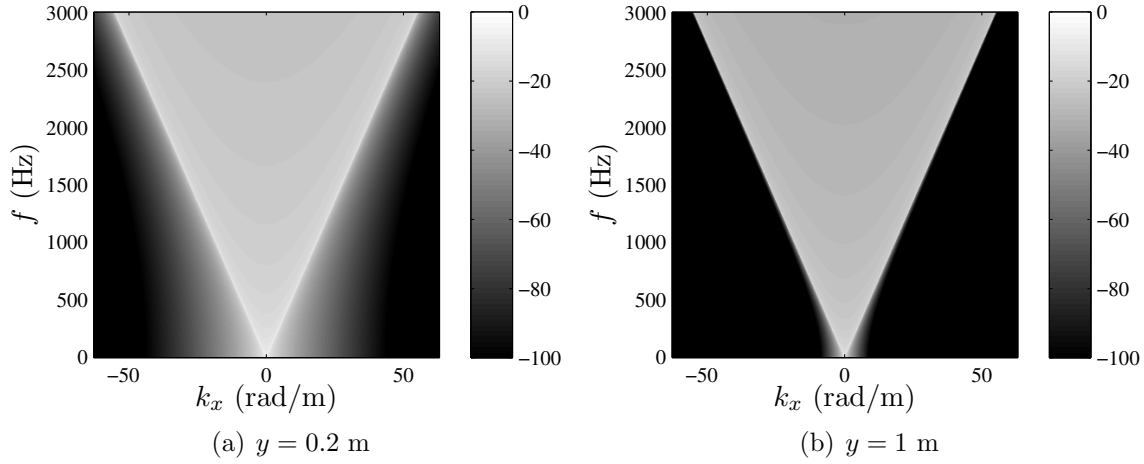


Figure 4.22: $20 \log_{10} |\tilde{G}(k_x, y, z, \omega)|$ for $z = 0$ and varying y .

Furthermore, it can be deduced from (4.49) and Fig. 4.22 that all propagating components of the synthesized sound field (i.e. components triggering the region of $\tilde{G}(k_x, y, z, \omega)$ inside the triangular structure) have comparable amplitude.

The locations $k_x = \frac{2\pi}{\Delta x} \eta + k_{\text{pw},x}$ in (4.49) are represented by black dots in Fig. 4.23. Locations where $|k_x| < \left| \frac{\omega}{c} \right|$ represent the synthesis of the combination of a propagating and an evanescent sound field as described by the Hankel function. Locations where $|k_x| > \left| \frac{\omega}{c} \right|$ indicate the synthesis of a purely evanescent component. As with planar secondary source distributions, the purely evanescent components can not be avoided since η is not bounded. Again, higher orders η lead to lower amplitudes of the contributions in the purely evanescent region $|k_x| > \left| \frac{\omega}{c} \right|$.

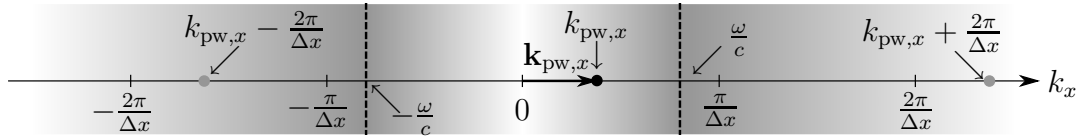


Figure 4.23: Illustration of the consequences of the discretization of the secondary source distributions for linear distributions by means of illustrating $\tilde{G}(k_x, y, z, \omega)$. The dots \bullet indicate synthesized components. Black solid lines and black dots represent quantities occurring with continuous secondary source distributions. Grey lines and dots represent quantities occurring additionally due to the spatial discretization. The grey shading indicates the amplitude of $\tilde{G}(\cdot)$. The vector $\mathbf{k}_{\text{pw},x} = [k_{\text{pw},x}]$ represents the propagation direction of the virtual plane wave projected onto the k_x -axis. Locations outside the interval $[-\frac{\omega}{c}; \frac{\omega}{c}]$ represent evanescent sound fields, locations inside represent propagating sound fields.

If only $\eta = 0$ falls into the region where $|k_x| < \left| \frac{\omega}{c} \right|$, the synthesized propagating sound field consists exclusively of the desired sound field plus an according evanescent component. This situation is illustrated in Fig. 4.23. Note that all synthesized

propagating components are accompanied by an additional evanescent component as described by the Hankel function in (4.49).

However, if the spacing Δx between adjacent secondary sources is large enough, respectively if the radian frequency ω is chosen high enough, then also synthesized components for $\eta \neq 0$ fall into the region where $|k_x| < \left|\frac{\omega}{c}\right|$. In this case, propagating discretization artifacts arise which are accompanied by an according evanescent component as discussed above. This situation is not illustrated in Fig. 4.23. These propagating discretization artifacts constitute additional wave fronts which are straight inside the horizontal plane. Informally, one speaks of additional plane waves.

The according location inside the region where $|k_x| < \left|\frac{\omega}{c}\right|$ determines the k_x -component of the propagation direction of the additional wave fronts. Note that the propagation directions of the additional wave fronts are dependent on the radian frequency ω . This finding has been derived in (Spors, 2008) for purely two-dimensional synthesis.

For reasons similar to those discussed in Sec. 4.5 for planar secondary source distributions, segregation of spatial aliasing in the strict sense and other reconstruction errors is not useful either for linear secondary source distributions synthesizing virtual plane waves.

The anti-aliasing condition preventing undesired propagating aliasing contributions can be graphically deduced from Fig. 4.23. It is given by

$$\omega < \frac{2\pi c}{\Delta x (1 + |\cos \theta_{pw}|)} . \quad (4.50)$$

Eq. (4.50) has already been derived in (Spors, 2006) for purely two-dimensional synthesis and in (Verheijen, 1997; Start, 1997; Pueo *et al.*, 2007; Ahrens & Spors, 2010d) for 2^{1/2}-dimensional synthesis.

As with planar secondary source distributions outlined in Sec. 4.5, the concept of narrowband and wide-/fullband driving functions is not useful with linear distributions.

4.6.2 Properties of the Synthesized Sound Field in Time-Frequency Domain

Refer to Fig. 4.24 for simulations of the sound field synthesized by a discrete linear secondary source distribution when driven in order to synthesize a virtual plane wave.

For a loudspeaker spacing of $\Delta x = 0.2$ m and a frequency of $f = 1000$ Hz as depicted in Fig. 4.24(a), 4.24(c), and 4.24(e), exclusively evanescent spatial discretization artifacts are apparent.

A higher frequency of $f = 1500$ Hz evokes an additional propagating wave which propagates in direction $\theta \approx 2$ rad $\approx 115^\circ$ with an amplitude similar to the desired sound field. Refer to Fig. 4.24(b), 4.24(d), and 4.24(f). The evanescent discretization artifacts in this situation exhibit very low amplitude and are not visible in the figures.

Choosing an even higher frequency or a larger secondary source spacing results in more propagating artifacts each of which with an individual propagation direction as discussed in Sec. 4.6.1.

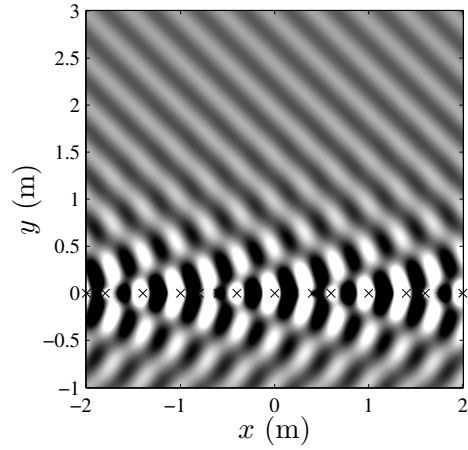
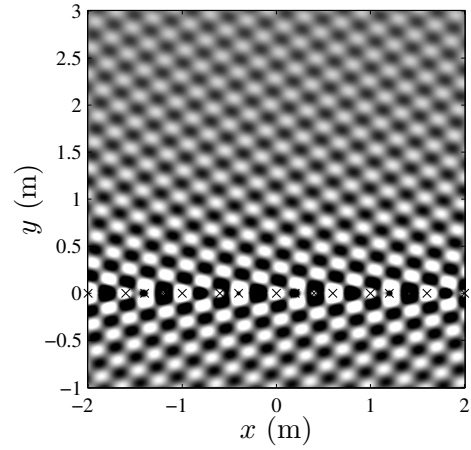
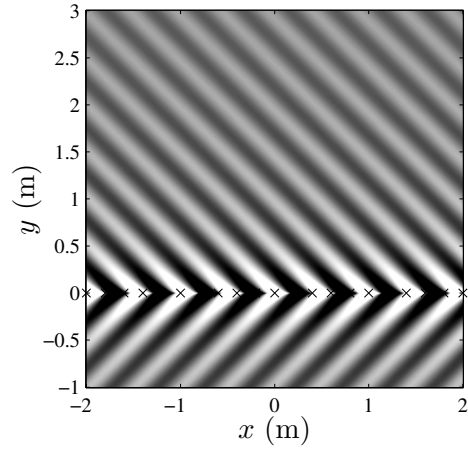
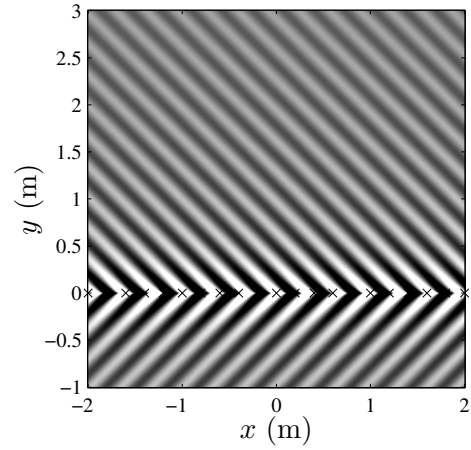
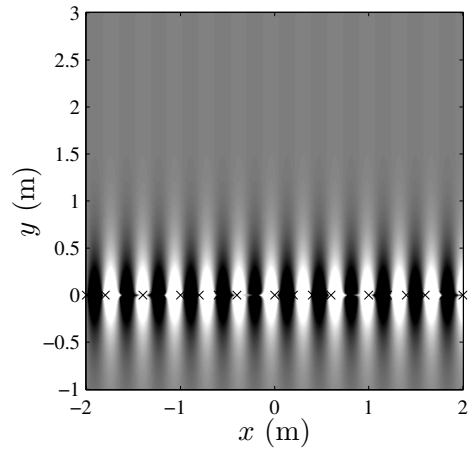
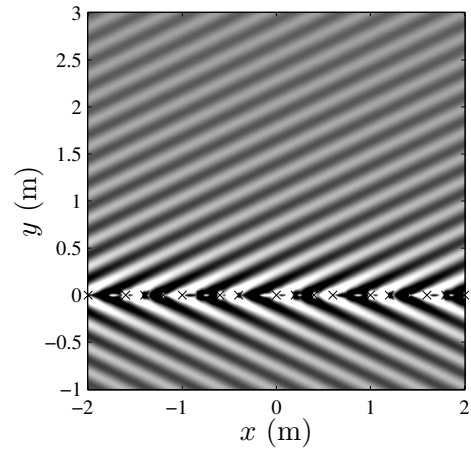
(a) synthesized sound field, $f = 1000$ Hz(b) synthesized sound field, $f = 1500$ Hz(c) desired component, $f = 1000$ Hz(d) desired component, $f = 1500$ Hz(e) discretization artifacts, $f = 1000$ Hz(f) discretization artifacts, $f = 1500$ Hz

Figure 4.24: Sound field synthesized by a discrete linear secondary source distribution; $\Delta x = 0.2$ m, $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\frac{\pi}{4}, \frac{\pi}{2})$, $y_{\text{ref}} = 1$ m; the marks indicate the secondary sources.

4.6.3 Properties of the Synthesized Sound Field in Time Domain

Fig. 4.25(a) shows a still image of the impulse response of a discrete linear secondary source distribution with a loudspeaker spacing of $\Delta x = 0.2$ m. Fig. 4.26 shows the impulse response for a specific receiver position. The secondary source distribution is driven in order to synthesize a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (\frac{\pi}{4}, \frac{\pi}{2})$. The absolute value of the time-domain sound pressure is shown in dB, i.e.

$$20 \log_{10} |\Re\{s_s(\mathbf{x}, t)\}| . \quad (4.51)$$

The representation of the driving function in time domain was obtained using (3.55) and applying a numerical Fourier transform on (3.56).

The observations are similar to those found in fullband synthesis using circular secondary source distributions discussed in Sec. 4.4.4. The discussion is kept brief and the reader is referred to Sec. 4.4.4 for details.

From Fig. 4.25(a) and 4.26 it can be deduced that:

- The synthesized wave front is perfectly straight. As with fullband synthesis with circular secondary source distributions discussed in Sec. 4.4.4, this suggests good localization due to the precedence effect.
- After the initial wave front high frequency echoes arise the strongest of which generally arrive from similar directions. The echoes are likely to produce coloration since they arrive at intervals below 1 ms.
- Note that contrary to fullband synthesis with circular secondary source distributions, the impulse response of the linear secondary source contour has infinite length (Fig. 4.26).
- As pointed out in Sec. 3.5.4, WFS exhibits similar properties like the solution presented in Sec. 3.5 with slightly less accuracy in the lower frequencies. The analysis above confirms once more that WFS indeed constitutes a method for synthesis of *wave fronts* (Berkhout *et al.*, 1993).

From the transfer function of the above described system to three different receiver positions along the y -axis depicted in Fig. 4.25(b) it can be deduced that:

- The transfer function is perfectly flat below a given frequency f_a at y_{ref} . For other positions, slight deviations arise. These deviations are individual for each position (actually for each distance to the secondary source distribution) and can therefore not be compensated for.
- The amplitude decay with distance y is apparent in the transfer function.
- Above f_a , densely spaced prominent notches and peaks of 10 dB or more occur.
- Above f_a , the transfer function exhibits a highpass character with a slope of approximately 3 dB per octave for all listening positions. As with full spatial bandwidth synthesis using circular secondary source distributions, this highpass character can be compensated for (Spors & Ahrens, 2010a).

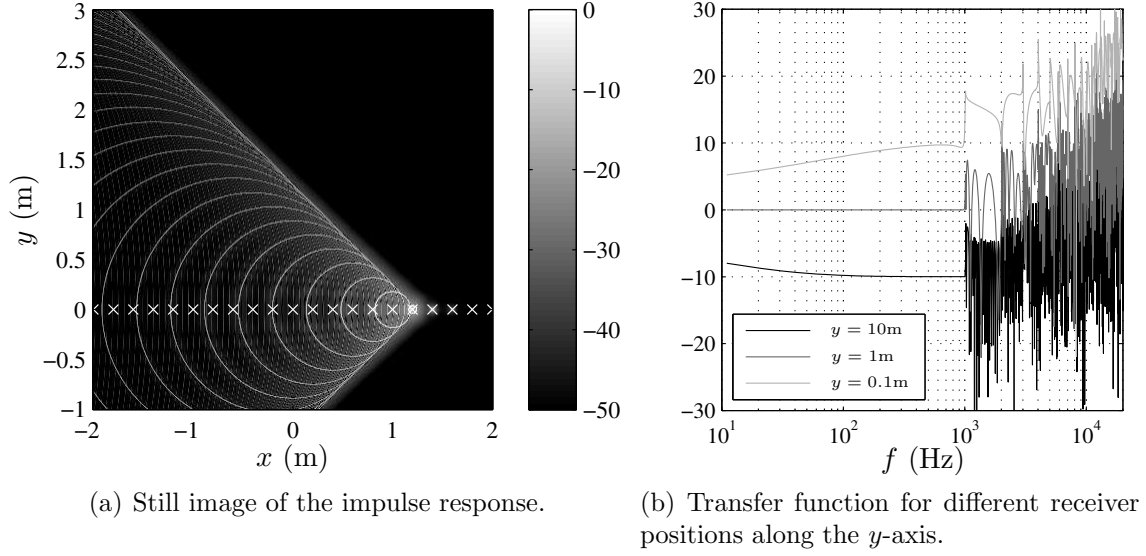


Figure 4.25: Impulse response and transfer function of a discrete secondary source distribution driven in order to synthesized a virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (\frac{\pi}{4}, \frac{\pi}{2})$. The secondary source spacing is $\Delta x = 0.2$ m; the white marks indicate the secondary sources.

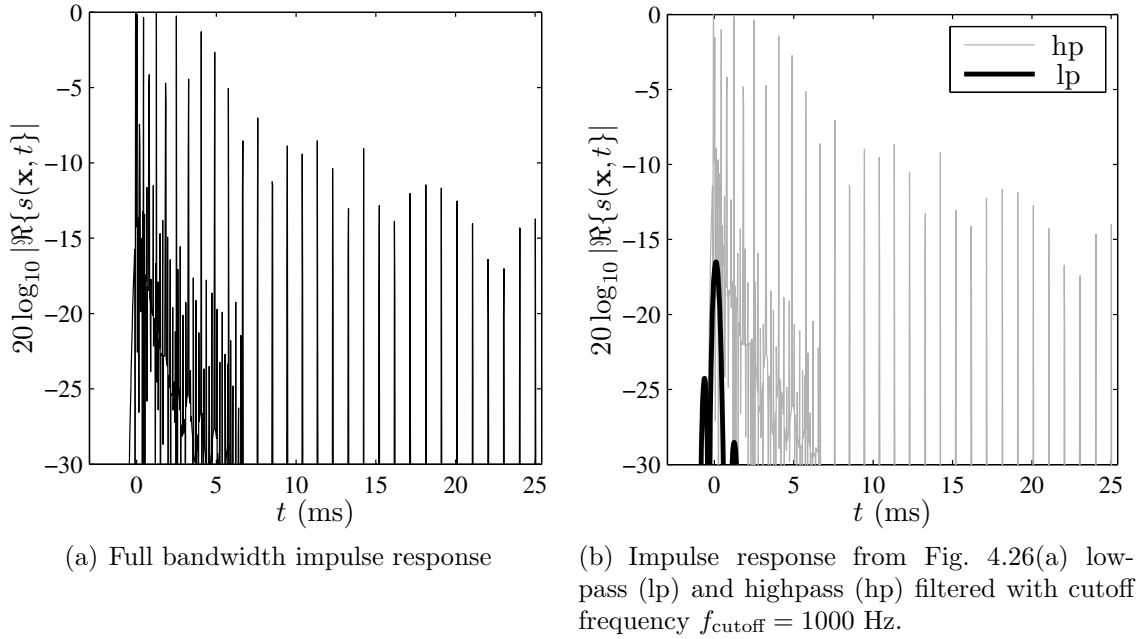


Figure 4.26: Impulse response of a discrete infinitely long linear secondary source distribution with a spacing of $\Delta x = 0.2$ m driven in order to synthesize virtual plane wave with propagation direction $(\theta_{pw}, \phi_{pw}) = (\frac{\pi}{4}, \frac{\pi}{2})$. The considered location is $\mathbf{x} = [0 \ 1 \ 0]^T$ m.

- Although not apparent from the simulations, it can be shown that the transfer function exhibits strong local variation especially at frequencies significantly above f_a . This variation is similar to that arising in fullband synthesis with circular secondary source distribution depicted in Fig. 4.14(d).

4.6.4 Spatial Discretization in Wave Field Synthesis Employing Linear Secondary Source Distributions

Sec. 3.5.4 has shown that – apart from a systematic amplitude deviation – the WFS driving function for the synthesis of a virtual plane wave using a linear distribution of secondary sources is essentially similar to the driving function investigated in Sec. 4.6.1 and 4.6.3. Consequently, the properties of WFS with linear secondary source distributions with respect to spatial discretization are essentially similar and can therefore be deduced from Sec. 4.6.1 and 4.6.3.

4.6.5 Optimizing the Synthesis with Respect to a Given Receiver Location

In Sec. 4.4.5 local sound field synthesis employing discrete circular secondary source distributions was shown. The local increase of physical accuracy was achieved by concentrating the energy of the continuous driving function (or correspondingly the energy of the desired sound field) at a small region in the space-frequency domain in order to avoid overlaps of the inevitable spectral repetitions. Since also spectral repetitions occur with discrete planar and linear secondary source distributions, spatial bandwidth limitation of the driving function can avoid the overlap of regions containing considerable energy. For convenience, this technique is only demonstrated for linear secondary source distributions but not for planar ones.

The driving function for the synthesis of a virtual plane wave by a linear distribution of secondary monopoles is given by (3.53). It is composed of a weighted Dirac delta function in the k_x -domain which makes a bandlimitation impossible without changing the propagation direction of the synthesized sound field. Therefore, the synthesis of the sound field of a virtual monopole source is considered in the following.

The generic driving function $\tilde{D}(k_x, \omega)$ in wavenumber domain for linear secondary source distributions is given by (3.48). The spatial spectrum $\tilde{S}(k_x, y, z, \omega)$ of the sound field of a monopole sound source located at $\mathbf{x}_s = [x_s \ y_s \ 0]^T$ can be deduced from $\tilde{G}_0(k_x, y, z, \omega)$ given by (C.8) via the shift theorem of the Fourier transform as (Girod *et al.*, 2001; Spors & Ahrens, 2010b)

$$\tilde{S}(k_x, y, z, \omega) = e^{ik_x x_s} \tilde{G}_0(k_x, y - y_s, z, \omega), \quad (4.52)$$

so that the driving function $\tilde{D}(k_x, \omega)$ explicitly reads

$$\tilde{D}(k_x, \omega) = e^{ik_x x_s} \times \begin{cases} \frac{H_0^{(2)}\left(\sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2} (y_{\text{ref}} - y_s)\right)}{H_0^{(2)}\left(\sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2} y_{\text{ref}}\right)} & \text{for } 0 \leq |k_x| < \left|\frac{\omega}{c}\right| \\ \frac{K_0\left(\sqrt{k_x^2 - \left(\frac{\omega}{c}\right)^2} (y_{\text{ref}} - y_s)\right)}{K_0\left(\sqrt{k_x^2 - \left(\frac{\omega}{c}\right)^2} y_{\text{ref}}\right)} & \text{for } 0 < \left|\frac{\omega}{c}\right| < |k_x| \end{cases}. \quad (4.53)$$

In the following, a virtual point source at $\mathbf{x}_s = [0 \ -1 \ 0]^T$ m and $y_{\text{ref}} = 1$ m is considered. Eq. (4.53) for these parameters is depicted in Fig. 4.27(a) and the corresponding synthesized sound field in Fig. 4.28(a). Note that the latter was derived via a numerical Fourier transform since an analytical expression is not available.

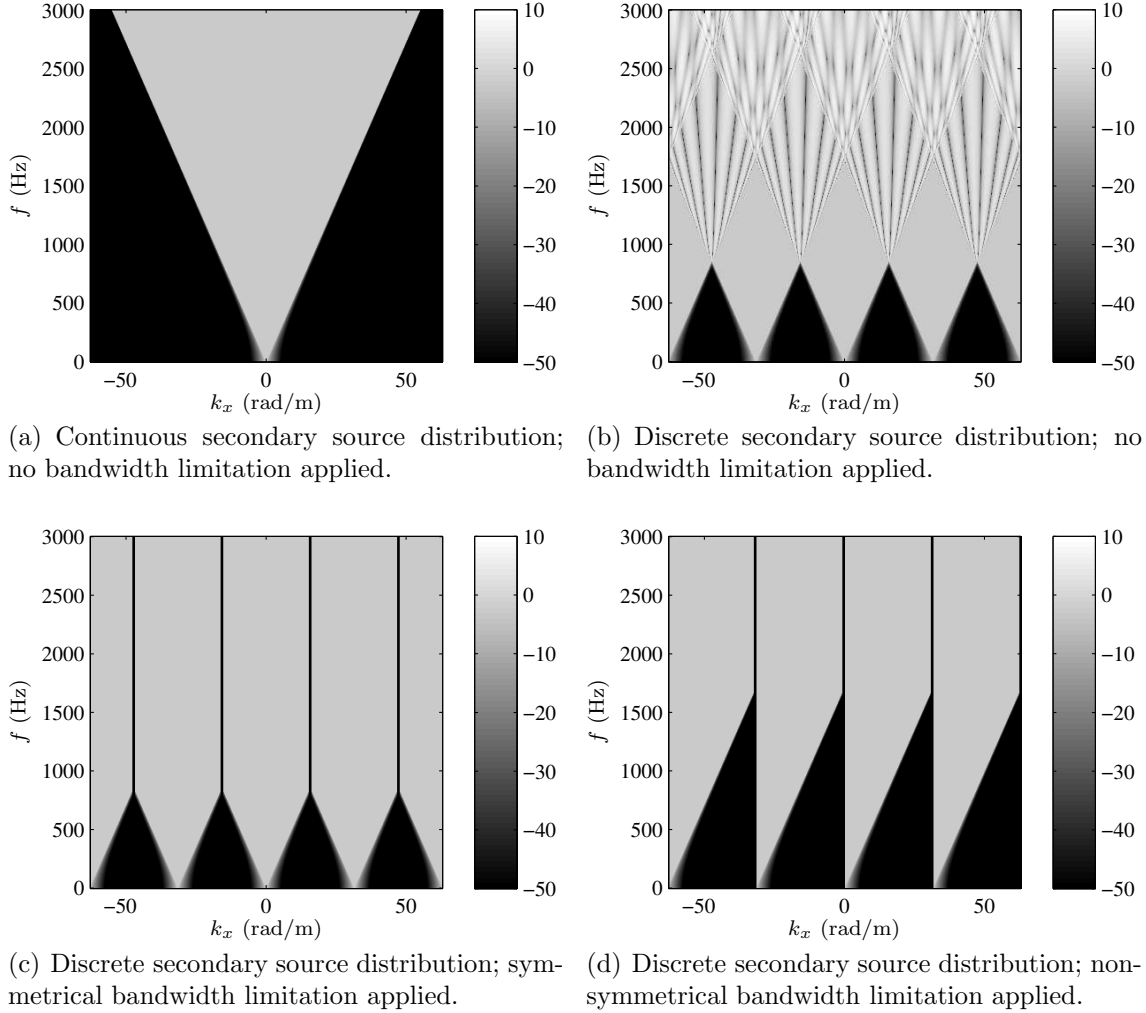
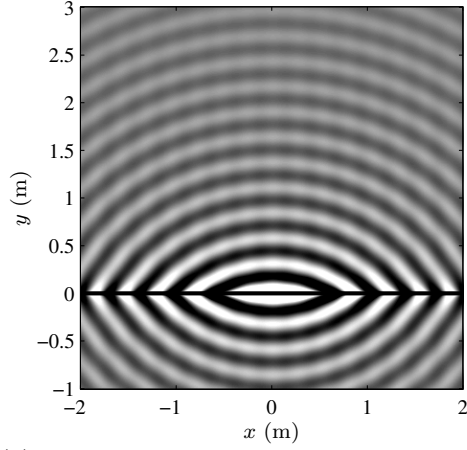
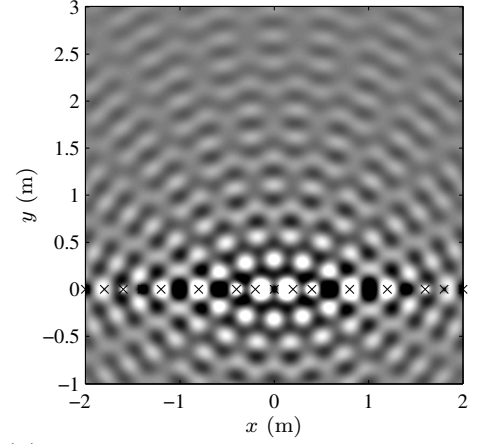


Figure 4.27: $20 \log_{10} |\tilde{D}(k_x, \omega)|$ for continuous (Fig. 4.27(a)) and discrete linear secondary source distributions (Fig. 4.27(b)–(d)); $\Delta x = 0.2$ m; $y_{\text{ref}} = 1$ m.

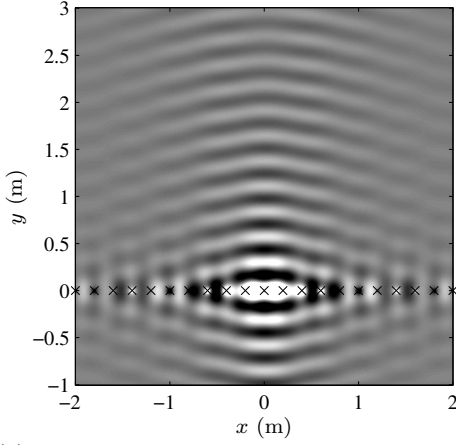
Since $\tilde{D}(k_x, \omega)$ is not bandlimited with respect to k_x , discretization of the driving function leads to an interference of the spectral repetitions above approximately 800 Hz for a secondary source spacing of $\Delta x = 0.2$ m (Fig. 4.27(b)) and thus a corruption of the synthesized sound field (Fig. 4.28(b)). A bandlimitation with respect to k_x can be straightforwardly performed by setting selected components of $\tilde{D}(k_x, \omega)$ to zero. Of course, more advanced weighting may also be applied. For simplicity, only the former approach is treated here. Note that such a spatial bandwidth limitation in order to reduce discretization artifacts has been proposed in (Verheijen, 1997) though detailed properties of the synthesized sound field have not been investigated.



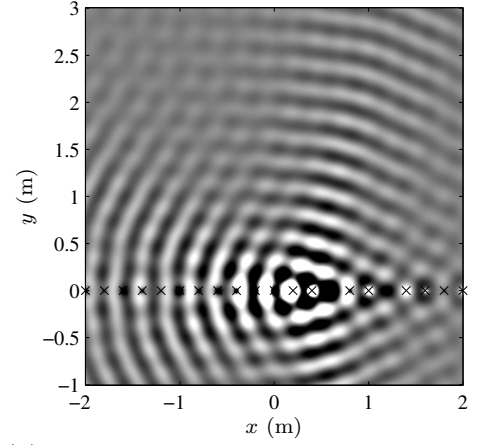
(a) Continuous secondary source distribution; no bandwidth limitation applied.



(b) Discrete secondary source distribution; no bandwidth limitation applied.



(c) Discrete secondary source distribution; symmetrical bandwidth limitation similar to Fig. 4.27(c) applied.



(d) Discrete secondary source distribution; non-symmetrical bandwidth limitation similar to Fig. 4.27(d).

Figure 4.28: Illustration of the influence of the bandwidth of the driving function for the synthesis of a virtual point source at position $\mathbf{x}_s = [0 \ -1 \ 0]^T$ m emitting a monochromatic signal of $f = 1300$ Hz; $\Re\{S(\mathbf{x}, \omega)\}$ is shown. In the continuous case Fig. 4.28(a), the secondary source distribution is indicated by the black line; In the discrete cases in Fig. 4.28(b)–(d), the marks indicate the secondary sources. The secondary source spacing is $\Delta x = 0.2$ m.

Narrowband synthesis (avoiding overlaps of the spectral repetitions) is achieved with a passband of the continuous driving function with a width of smaller than $\frac{2\pi}{\Delta x}$. For a secondary source spacing of $\Delta x = 0.2$ m as employed in Fig. 4.28 this means that the passband has to be smaller or equal to approximately $31 \frac{\text{rad}}{\text{m}}$.

Limiting the spatial bandwidth of $\tilde{D}(k_x, \omega)$ in a manner symmetrical to $k_x = 0$ (Fig. 4.27(c)) results in a synthesized sound field which is less corrupted by spatial aliasing artifacts but the energy of which propagates primarily in direction perpendicular to the secondary source distribution. As a consequence, the amplitude of the synthesized sound field is significantly too low a certain locations in the target

half-plane.

Limiting the spatial bandwidth of $\tilde{D}(k_x, \omega)$ in a manner which is not symmetrical to $k_x = 0$ (Fig. 4.27(d)) allows for a steering of the primary propagation direction of the synthesized sound field into a given direction. The synthesis can therefore be optimized with respect to a given location of the receiver (e.g. the listener).

Local sound field synthesis is thus also possible using linear distributions of secondary sources.

4.7 Further Aspects of Discretization and Spatial Truncation With Planar and Linear Secondary Source Distributions

In order to assess the properties of spatially truncated discrete secondary source distributions (which is in fact what is found in real-life), the findings derived in Sec. 3.5.5 and 4.5 (or Sec. 4.6 respectively) have to be combined (Pueo *et al.*, 2007; Ahrens & Spors, 2010d). For convenience, a discrete linear secondary source distribution which is truncated in x -dimension is explicitly considered.

From (3.63) and (4.43) it can be deduced that the synthesized sound field $\tilde{S}_{S,tr}(k_x, y, z, \omega)$ of a truncated discrete linear secondary source distribution is given in wavenumber domain by

$$\tilde{S}_{S,tr}(k_x, y, z, \omega) = \frac{1}{2\pi} \underbrace{\left(\sum_{\eta=-\infty}^{\infty} \tilde{w}(k_x) *_{k_x} \tilde{D} \left(k_x - \frac{2\pi}{\Delta x} \eta, \omega \right) \right)}_{= \tilde{D}_{S,tr}(k_x, \omega)} \tilde{G}(k_x, y, z, \omega) . \quad (4.54)$$

For the interpretation of (4.54) again the synthesis of a plane wave is considered. Recall the plane-wave driving function given by (3.54). The spatial truncation does not only smear the energy of the desired components along the k_x but also the repetitions due to discretization. It can thus happen that such a contribution due to discretization which is propagating for an infinite discrete secondary source distribution is partly smeared into the evanescent region $0 < |\frac{\omega}{c}| < |k_x|$ (Pueo *et al.*, 2007; Ahrens & Spors, 2010d). Vice versa, a contribution due to discretization which is evanescent for an infinite discrete secondary source distribution can partly be smeared into the propagating region where $0 < |k_x| < |\frac{\omega}{c}|$. As a consequence, the interaction of spatial sampling and truncation results in a reduced spatial fine structure of the synthesized sound field.

It has to be noted that the undesired evanescent components in the synthesized sound field exhibit an amplitude which is decaying rapidly with the distance to the secondary source array. They become negligible already at moderate distances (Williams, 1999; Spors & Ahrens, 2007).

Above derived findings are supported by results from (Kennedy *et al.*, 2007) where it is shown that a bandlimited sound field has a limited complexity in a given spherical region. Thus, it can be resynthesized by a limited number of secondary sources. Inversely, a limited number of secondary sources – e.g. a truncated sampled array – is then only capable of synthesizing a sound field with limited complexity.

Due to the complex structure of the sound field of a truncated secondary source distribution as discussed in Sec. 3.5.5, the amplitude of the individual propagating aliasing components is strongly dependent on the location of the receiver as is the amplitude of the desired component (Spors, 2006; Pueo *et al.*, 2007).

Fig. 4.29 shows a combination of the conditions depicted in Fig. 3.14 and 4.24(b), i.e. the sound field synthesized by a truncated discrete secondary source distribution at a frequency where propagating discretization artifacts arise. The desired virtual plane wave propagates mainly in direction $(\theta_{pw}, \phi_{pw}) = (\frac{\pi}{4}, \frac{\pi}{2})$. The propagating discretization artifacts propagate into an essentially different direction. Only at locations close to the secondary source distribution do the two components of the synthesized sound field overlap.

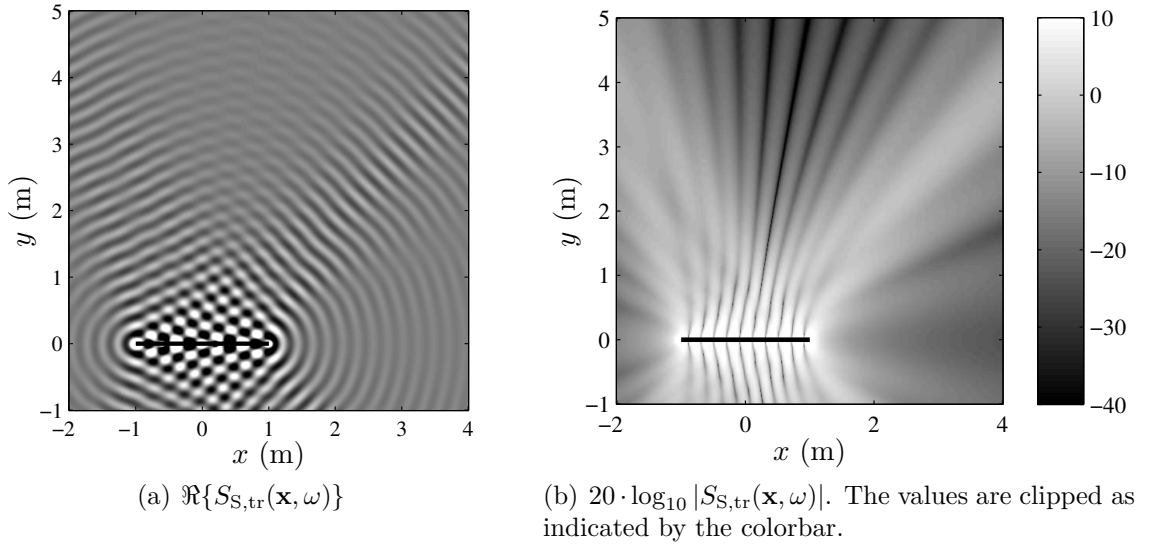


Figure 4.29: Sound pressure $S_{S,tr,pw}(\mathbf{x}, \omega)$ of a discrete linear distribution of secondary point sources synthesizing a virtual plane wave of $f_{pw} = 1000$ Hz and unit amplitude with propagation direction $(\theta_{pw}, \phi_{pw}) = (\frac{\pi}{4}, \frac{\pi}{2})$ referenced to the distance $y_{ref} = 1.0$ m. The secondary source distribution is located along the black line. $L = 2$ m, $\Delta x = 0.2$ m.

4.8 On the Spatial Bandwidth of Numerical Solutions

The numerical approaches for sound field synthesis mentioned in Sec. 2.6.4, i.e. (Kirkeby & Nelson, 1993; Ward & Abhayapala, 2001; Daniel, 2001; Poletti, 2005; Hannemann & Donohue, 2008; Kolundžija *et al.*, 2009) all employ local optimization criteria. As shown above, such a local optimization is achieved via a limitation of the spatial bandwidth of the secondary source driving function. Depending on the location and shape of the region for which the synthesis is optimized either classical narrowband synthesis takes place which is similar to the one treated

in Sec. 4.3.2 and Sec. 4.4.1 or local sound field synthesis as treated in Sec. 4.4.5 and Sec. 4.6.5 takes place.

4.9 Summary

In this chapter, the consequences of spatial discretization of the continuous secondary source distributions as treated in Chap. 3 on the synthesized sound field were investigated. It was found common for all geometries of secondary source contours that the discretization leads to repetitions in the spatial spectra of the driving function. With spherical contours, the repetitions occur in the spherical harmonics domain, with circular contours in the Fourier series domain, with planar contours in the wavenumber domain with respect to two dimensions, and similarly in the wavenumber domain with respect to one dimension with linear secondary source contours.

Typical practical implementations of sound field synthesis methods use loudspeaker spacings of several centimeters. This results in considerable discretization artifacts above a few thousand Hertz. Since the audible frequency range can be assumed to significantly exceed 15 kHz the synthesized sound field will always be corrupted when the entire potential receiver area is considered.

The most fundamental conclusion which can be drawn from the presented results is the fact that the spatial bandwidth of the desired sound field – and thus of the driving function – has essential influence of the synthesized sound field. The concept of categorizing the methods with respect to their spatial bandwidth into narrowband, wideband, and fullband methods was proposed and elaborated.

Narrowband methods avoid overlaps of the spectral repetitions and typically lead to regions in the receiver area in which the accuracy of the synthesis is significantly higher than at other locations. Fullband methods create artifacts which are rather evenly distributed over the receiver area. The category of wideband methods (methods with a bandwidth in between narrowband and fullband) was not investigated in detail and is subject to future work.

The established method of near-field higher order Ambisonics was found to be a narrowband method; Wave Field Synthesis was found to be a fullband method. It can not be decided at this stage if a high or low spatial bandwidth of the driving function is preferable in a specific situation.

The representation of the synthesized sound field in wavenumber domain when planar and linear secondary source distributions are considered allowed for a segregation of artifacts in terms of propagating and evanescent components. Such an analysis is not straightforward for spherical and circular distributions.

Further considerations on the spatial bandwidth led to the concept of *local sound field synthesis* which locally increases the accuracy by the cost of stronger artifacts elsewhere.

Chapter 5

Conclusions and Outlook

The physical fundamentals of sound field synthesis were treated in this thesis. Explicit solutions for the driving functions for continuous spherical, circular, planar, and linear distributions of secondary sources were derived based on the single-layer potential solution and modifications thereof. It was shown that continuous spherical and planar secondary source distributions allow for a perfect synthesis of a given desired source-free sound field with minor systematic limitations.

Circular and linear distributions on the other hand exhibit essential limitations most notably in terms of the possible propagation directions of the synthesized sound field and in terms of the amplitude decay of the latter. However, for such a $2^{1/2}$ -dimensional synthesis no general underlying theory is available which allows for a detailed prediction of the properties of the synthesized sound field. Each situation has to be investigated individually.

Near-field compensated higher-order Ambisonics was identified to be the single-layer potential solution for spherical secondary source distributions. The Wave Field Synthesis solution to the problem constitutes a high-frequency approximation of the presented approach.

Continuous secondary source distributions as mentioned above can not be implemented in practice but discrete distributions of loudspeakers have to be used. The analysis of the properties of such discrete secondary source distributions revealed that the spatial discretization leads to repetitions of the spatial spectrum of the driving function. The spatial bandwidth of the secondary source driving function has thus essential impact on the properties of the synthesized sound field. It was therefore proposed to categorize sound field synthesis approaches with respect to their spatial bandwidth into narrowband, wideband, and fullband approaches. Modifications of the spatial bandwidth of the driving function allow for a local increase of the accuracy of the synthesis.

The investigation of the influence of the radiation properties of the employed loudspeakers on the synthesized sound field was beyond the scope of the present thesis and is subject to future work.

The presented work is restricted to a mainly instrumentalized analysis of the properties of the investigated synthetic sound fields. Some prediction of human perception of simple scenarios could be given based on an analysis of the evolving wave fronts. However, in order to better understand the perception extensive experimental investigations have to be performed.

Appendix A

Coordinate Systems

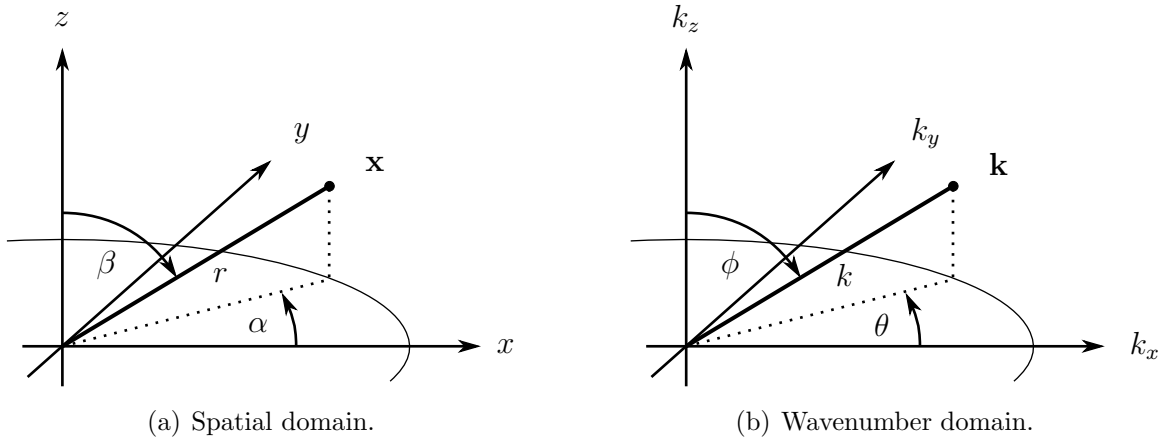


Figure A.1: The coordinate systems used in this paper.

The coordinate systems used in this thesis are depicted in Fig. A.1. The spherical coordinates (r, α, β) are related to the Cartesian coordinates $[x, y, z]^T$ by (Weisstein, 2002)

$$r = \sqrt{x^2 + y^2 + z^2} \quad (\text{A.1a})$$

$$\alpha = \arctan\left(\frac{y}{x}\right) \quad (\text{A.1b})$$

$$\beta = \arccos\left(\frac{z}{r}\right), \quad (\text{A.1c})$$

where $r \in [0, \infty)$, $\alpha \in [0, 2\pi)$, and $\phi \in [0, \pi]$, and the inverse tangent must be suitably defined to take the correct quadrant of (x, y) into account (Weisstein, 2002).

The Cartesian coordinates $[x, y, z]^T$ are related to the spherical coordinates (r, α, β) by

$$x = r \cos \alpha \sin \beta \quad (\text{A.2a})$$

$$y = r \sin \alpha \sin \beta \quad (\text{A.2b})$$

$$z = r \cos \beta. \quad (\text{A.2c})$$

The angles α and θ are termed *azimuth*, β and ϕ are termed *spherical polar angle*, or *zenith angle*, or *colatitude*.

Appendix B

Definition of the Fourier Transform

The temporal Fourier transform used in this work is defined as (Bracewell, 2000)

$$S(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} s(\mathbf{x}, t) e^{-i\omega t} dt . \quad (\text{B.1})$$

The inverse temporal Fourier transform is therefore

$$s(\mathbf{x}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\mathbf{x}, \omega) e^{i\omega t} d\omega . \quad (\text{B.2})$$

The spatial Fourier transform is defined as

$$\tilde{S}(k_x, y, z, \omega) = \int_{-\infty}^{\infty} S(\mathbf{x}, \omega) e^{ik_x x} dx \quad (\text{B.3})$$

exemplarily for the x -dimension. The corresponding inverse spatial Fourier transform is

$$S(\mathbf{x}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{S}(k_x, y, z, \omega) e^{-ik_x x} dk_x . \quad (\text{B.4})$$

Note that reversed exponents are used in the spatial Fourier transform compared to the temporal one. The motivation for this choice is outlined in Sec. 2.2.5.

Appendix C

Fourier Transforms of Selected Quantities

C.1 Fourier Transforms of a Plane Wave

A monochromatic plane wave with radian frequency ω_{pw} and wave vector \mathbf{k}_{pw} is given by (Williams, 1999)

$$s(\mathbf{x}, t) = e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}} \cdot e^{i\omega_{\text{pw}} t} \quad (\text{C.1a})$$

$$= e^{i\frac{\omega_{\text{pw}}}{c}(ct - \mathbf{n}_{\text{pw}}^T \mathbf{x})} , \quad (\text{C.1b})$$

with \mathbf{n}_{pw} denoting the unit length vector pointing in the same direction like \mathbf{k}_{pw} , i.e. in propagation direction of the plane wave. The term in brackets in (C.1b) is termed *Hesse normal form* of a plane propagating in direction \mathbf{n}_{pw} with speed c (Weisstein, 2002).

The Fourier transform of $s(\mathbf{x}, t)$ with respect to t yields (Girod *et al.*, 2001)

$$S(\mathbf{x}, \omega) = e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}} \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) . \quad (\text{C.2})$$

A further Fourier transform with respect to x yields

$$\tilde{S}(k_x, y, z, \omega) = 2\pi\delta(k_x - k_{\text{pw},x}) e^{-ik_{\text{pw},y}y} e^{-ik_{\text{pw},z}z} \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) , \quad (\text{C.3})$$

a further Fourier transform with respect to z yields

$$\tilde{S}(k_x, y, k_z, \omega) = 4\pi^2\delta(k_x - k_{\text{pw},x}) e^{-ik_{\text{pw},y}y} \delta(k_z - k_{\text{pw},z}) \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) , \quad (\text{C.4})$$

and finally a further Fourier transform with respect to y yields

$$\tilde{S}(\mathbf{k}, \omega) = 8\pi^3\delta(\mathbf{k} - \mathbf{k}_{\text{pw}}) \cdot 2\pi\delta(\omega - \omega_{\text{pw}}) . \quad (\text{C.5})$$

C.2 Fourier Transforms of the Free-Field Green's Function

The three-dimensional free-field Green's function for excitation at the coordinate origin is given in time domain by (Williams, 1999)

$$g(\mathbf{x}, t) = \frac{1}{4\pi} \frac{\delta\left(t - \frac{r}{c}\right)}{r} . \quad (\text{C.6})$$

The temporal Fourier transform of (C.6) is then

$$G(\mathbf{x}, \omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}r}}{r} . \quad (\text{C.7})$$

The Fourier transform with respect to x is calculated by applying Euler's formula (Weisstein, 2002) and using (Gradshteyn & Ryzhik, 2000, (3.876-1) and (3.876-2)) and (Morse & Feshbach, 1953, p. 1323). It is given by

$$\tilde{G}(k_x, y, z, \omega) = \begin{cases} -\frac{i}{4} H_0^{(2)} \left(\sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2} \sqrt{y^2 + z^2} \right) & \text{for } 0 \leq |k_x| < \left|\frac{\omega}{c}\right| \\ \frac{1}{2\pi} K_0 \left(\sqrt{k_x^2 - \left(\frac{\omega}{c}\right)^2} \sqrt{y^2 + z^2} \right) & \text{for } 0 < \left|\frac{\omega}{c}\right| < |k_x| . \end{cases} \quad (\text{C.8})$$

$H_0^{(2)}(\cdot)$ denotes the zero-th order Hankel function of second kind, $K_0(\cdot)$ the zero-th order modified Bessel function of second kind (Williams, 1999). A further Fourier transform with respect to z is yielded using (Gradshteyn & Ryzhik, 2000, (6.677-3), (6.677-4), and (6.677-5)). It is given by

$$\tilde{G}(k_x, y, k_z, \omega) = \begin{cases} -\frac{i}{2} \frac{e^{-i\sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2 - k_z^2} y}}{\sqrt{\left(\frac{\omega}{c}\right)^2 - k_x^2 - k_z^2}} & \text{for } 0 \leq \sqrt{k_x^2 + k_z^2} < \left|\frac{\omega}{c}\right| \\ \frac{1}{2} \frac{e^{-\sqrt{k_x^2 + k_z^2 - \left(\frac{\omega}{c}\right)^2} y}}{\sqrt{k_x^2 + k_z^2 - \left(\frac{\omega}{c}\right)^2}} & \text{for } 0 < \left|\frac{\omega}{c}\right| < \sqrt{k_x^2 + k_z^2} . \end{cases} \quad (\text{C.9})$$

Note that (C.9) is only valid for $y > 0$ (Gradshteyn & Ryzhik, 2000).

Finally, $\tilde{G}(\mathbf{k}, \omega)$ is yielded using (Gradshteyn & Ryzhik, 2000, (3.893-2)). It is given by

$$\tilde{G}(\mathbf{k}, \omega) = \tilde{G}(k, \omega) = \frac{1}{k^2 - \left(\frac{\omega}{c}\right)^2} . \quad (\text{C.10})$$

Appendix D

Convolution Theorems

D.1 Fourier Series Domain

A representation of the Fourier series expansion coefficients $\mathring{H}_m(r, \beta, \omega)$ of a function $H(\mathbf{x}, \omega)$ which is given by a multiplication of two functions $F(\mathbf{x}, \omega)$ and $G(\mathbf{x}, \omega)$ as

$$H(\mathbf{x}, \omega) = F(\mathbf{x}, \omega) \cdot G(\mathbf{x}, \omega) \quad (\text{D.1})$$

in terms of the Fourier series expansion coefficients $\mathring{F}_m(r, \beta, \omega)$ and $\mathring{G}_m(r, \beta, \omega)$ of $F(\mathbf{x}, \omega)$ and $G(\mathbf{x}, \omega)$ respectively is derived in this section. Applying (2.25) yields

$$\begin{aligned} \mathring{H}_m(r, \beta, \omega) &= \frac{1}{2\pi} \int_0^{2\pi} F(\mathbf{x}, \omega) G(\mathbf{x}, \omega) e^{-im\alpha} d\alpha \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{m_1=-\infty}^{\infty} \mathring{F}_{m_1}(r, \beta, \omega) e^{im_1\alpha} \sum_{m_2=-\infty}^{\infty} \mathring{G}_{m_2}(r, \beta, \omega) e^{im_2\alpha} e^{-im\alpha} d\alpha \\ &= \frac{1}{2\pi} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \mathring{F}_{m_1}(r, \beta, \omega) \mathring{G}_{m_2}(r, \beta, \omega) \int_0^{2\pi} e^{i(m_1+m_2-m)\alpha} d\alpha . \end{aligned} \quad (\text{D.2})$$

The integral in (D.2) vanishes unless $m_1 + m_2 - m = 0$ or $m_2 = m - m_1$ respectively. In these cases it equals 2π so that finally (Girod *et al.*, 2001)

$$\mathring{H}_m(r, \beta, \omega) = \sum_{m_1=-\infty}^{\infty} \mathring{F}_{m_1}(r, \beta, \omega) \mathring{G}_{m-m_1}(r, \beta, \omega) = \mathring{F}_m(r, \beta, \omega) *_m \mathring{G}_m(r, \beta, \omega) , \quad (\text{D.3})$$

which represents a *convolution theorem* for the Fourier series expansion.

D.2 Spherical Harmonics Domain

The procedure outlined in Sec. D.1 is adapted here in order to obtain a representation of the coefficients $\mathring{H}_n^m(r, \omega)$ of a function $H(\mathbf{x}, \omega)$ which is given by a multiplication

of two functions $F(\mathbf{x}, \omega)$ and $G(\mathbf{x}, \omega)$ as

$$H(\mathbf{x}, \omega) = F(\mathbf{x}, \omega) \cdot G(\mathbf{x}, \omega) \quad (\text{D.4})$$

in terms of the coefficients $\mathring{F}_n^m(r, \omega)$ and $\mathring{G}_n^m(r, \omega)$ of $F(\mathbf{x}, \omega)$ and $G(\mathbf{x}, \omega)$ respectively. Applying (2.22) yields

$$\begin{aligned} \mathring{H}_n^m(r, \omega) &= \int_0^{2\pi} \int_0^\pi F(\mathbf{x}, \omega) G(\mathbf{x}, \omega) Y_n^{-m}(\alpha, \beta) \sin \beta \, d\beta \, d\alpha \\ &= \int_0^{2\pi} \int_0^\pi \sum_{n_1=0}^\infty \sum_{m_1=-n_1}^{n_1} \mathring{F}_{n_1}^{m_1}(r, \omega) Y_{n_1}^{m_1}(\alpha, \beta) \sum_{n_2=0}^\infty \sum_{m_2=-n_2}^{n_2} \mathring{G}_{n_2}^{m_2}(r, \omega) \\ &\quad \times Y_{n_2}^{m_2}(\alpha, \beta) Y_n^{-m}(\alpha, \beta) \sin \beta \, d\beta \, d\alpha \\ &= \sum_{n_1=0}^\infty \sum_{m_1=-n_1}^{n_1} \sum_{n_2=0}^\infty \sum_{m_2=-n_2}^{n_2} \mathring{F}_{n_1}^{m_1}(r, \omega) \mathring{G}_{n_2}^{m_2}(r, \omega) \\ &\quad \times \underbrace{\int_0^{2\pi} \int_0^\pi Y_{n_1}^{m_1}(\alpha, \beta) Y_{n_2}^{m_2}(\alpha, \beta) Y_n^{-m}(\alpha, \beta) \sin \beta \, d\beta \, d\alpha}_{= \gamma_{n_1, n_2, n}^{m_1, m_2, m}}. \quad (\text{D.5}) \end{aligned}$$

Integrals like the one in (D.5) often appear in problems in quantum mechanics and their properties are well investigated (Arfken & Weber, 2005). The result is a real number and these integrals are also referred to as *Gaunt coefficients* $\gamma_{n_1, n_2, n}^{m_1, m_2, m}$ (Sébilleau, 1998). The integral form of $\gamma_{n_1, n_2, n}^{m_1, m_2, m}$ as given in (D.5) is inconvenient for evaluation since it can not be solved analytically. More convenient is the representation (Gumerov & Duraiswami, 2004, eq. (3.2.28), p. 99)

$$\gamma_{n_1, n_2, n}^{m_1, m_2, m} = \frac{1}{4\pi} \sqrt{\frac{(2n_1+1)(2n_2+1)(2n+1)}{4\pi}} \mathcal{E} \begin{pmatrix} m_1 & m_2 & -m \\ n_1 & n_2 & n \end{pmatrix}. \quad (\text{D.6})$$

The *E-symbol* $\mathcal{E}(\cdot)$ is defined as (Gumerov & Duraiswami, 2004, eq. (3.2.27), p. 99)

$$\mathcal{E} \begin{pmatrix} m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{pmatrix} = 4\pi \varepsilon_{m_1} \varepsilon_{m_2} \varepsilon_{m_3} \begin{pmatrix} n_1 & n_2 & n_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} n_1 & n_2 & n_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \quad (\text{D.7})$$

with

$$\varepsilon_m = i^{m+|m|} = \begin{cases} (-1)^m & \forall \, m \geq 0 \\ 1 & \forall \, m \leq 0 \end{cases} \quad (\text{D.8})$$

and $\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$ denoting the *Wigner 3j-Symbol*. The Wigner 3j-Symbol is defined in (Weisstein, 2002). The MATLAB simulations presented in this thesis employ the script provided by (Kraus, 2008).

The E-symbol and thus the Gaunt coefficients $\gamma_{n_1, n_2, n}^{m_1, m_2, m}$ satisfy the following selection rules:

1. $m_2 = m - m_1$.

2. $|n - n_2| \leq n_1 \leq n + n_2$ (*triangle inequalities* or *triangle rule* (Weisstein, 2002)).
3. $n + n_1 + n_2$ is even or zero.

If these rules are not satisfied then $\gamma_{n_1, n_2, n}^{m_1, m_2, m} = 0$. Actually, it can be shown that $\gamma_{n_1, n_2, n}^{m_1, m_2, m}$ vanishes in more cases than stated above (Gjellestad, 1955; Gumerov & Duraiswami, 2004). In order to retain notational clarity the selection rules are only occasionally explicitly considered.

Reformulating (D.5) by explicitly considering rule 1 reads then (Arfken & Weber, 2005; Shirdhonkar & Jacobs, 2005)

$$\mathring{H}_n^m(r, \omega) = \sum_{n_1=0}^{\infty} \sum_{m_1=-n_1}^{n_1} \sum_{n_2=0}^{\infty} \mathring{F}_{n_1}^{m_1}(r, \omega) \mathring{G}_{n_2}^{m-m_1}(r, \omega) \gamma_{n_1, n_2, n}^{m_1, m-m_1, m} \quad (\text{D.9})$$

$$= \mathring{F}_n^m(r, \omega) *_n^m \mathring{G}_n^m(r, \omega) . \quad (\text{D.10})$$

Eq. (D.9) constitutes a *convolution theorem* for the spherical harmonics expansion.

Appendix E

Miscellaneous Mathematical Considerations

E.1 Translation of Spherical Harmonics Expansions

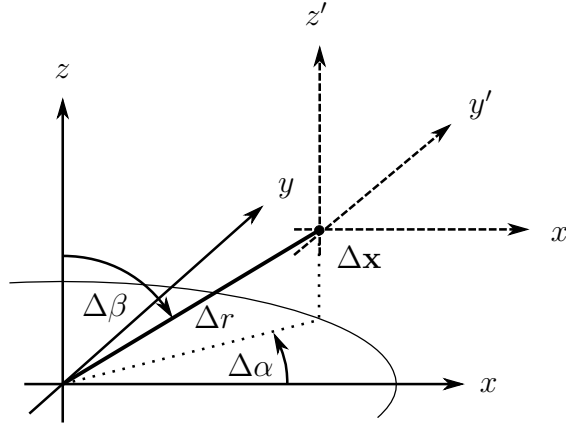


Figure E.1: Illustration of the local coordinate system employed in (E.1).

Assume the coefficients $\check{S}'_{n',e}{}^{m'}(\omega)$ represent an exterior sound field $S(\mathbf{x}, \omega)$ with respect to a local coordinate system with origin at $\Delta\mathbf{x}$ which can be transformed into the global coordinate system by a simple translation as depicted in Fig. E.1. Then $S(\mathbf{x}', \omega)$ can be described as (refer to (2.21b))

$$S(\mathbf{x}', \omega) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \check{S}'_{n',e}{}^{m'}(\omega) h_{n'}^{(2)}\left(\frac{\omega}{c}r'\right) Y_{n'}^{m'}(\beta', \alpha') \quad (\text{E.1})$$

with respect to the local coordinate system. Note that $\mathbf{x}' = \mathbf{x}'(\mathbf{x}) = \mathbf{x} + \Delta\mathbf{x}$.

It is now desired to describe $S(\mathbf{x}, \omega)$ by means of a spherical harmonics expansion around the origin of the global coordinate system. This *translation* of the coordinate system is described below.

Assuming that the origin of the global coordinate system is located in the exterior domain with respect to the local coordinate system, then it must be possible to expand the term $h_{n'}^{(2)} \left(\frac{\omega}{c} r' \right) Y_{n'}^{m'}(\beta', \alpha')$ with respect to the global coordinate system as (Gumerov & Duraiswami, 2004, ch. 3.2)

$$h_{n'}^{(2)} \left(\frac{\omega}{c} r' \right) Y_{n'}^{m'}(\beta', \alpha') = \sum_{n=0}^{\infty} \sum_{m=-n}^n (-1)^{n+n'} (E|I)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega) j_n \left(\frac{\omega}{c} r \right) Y_n^m(\beta, \alpha), \quad (\text{E.2})$$

since this term constitutes a solution to the wave equation. The notation $(E|I)$ indicates that the translation represents a change from an exterior expansion to an interior expansion (Williams, 1999; Gumerov & Duraiswami, 2004). The factor $(-1)^{n+n'}$ arises since the coefficients $(E|I)$ are defined in (Gumerov & Duraiswami, 2004) for translation in opposite direction. Refer also to (ibidem, eq. (3.2.54), p. 103).

Inserting (E.2) in (E.1) and re-ordering of the sums reveals the general form of $\check{S}_n^m(\omega)$ as

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \check{S}_{n',e}^{m'}(\omega) (-1)^{n+n'} (E|I)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega) j_n \left(\frac{\omega}{c} r \right) Y_n^m(\beta, \alpha)}_{= \check{S}_n^m(\omega)}. \quad (\text{E.3})$$

From (Gumerov & Duraiswami, 2004, eq. (3.2.30), (3.2.36); p. 100, 101) it can be deduced that the translation coefficients $(E|I)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega)$ are given by

$$(E|I)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega) = \sum_{n''=|n'-n|}^{n'+n} i^{n+n''-n'} \sqrt{\frac{(2n+1)(2n'+1)(2n''+1)}{4\pi}} \times \mathcal{E} \begin{pmatrix} m' & -m & m-m' \\ n' & n & n'' \end{pmatrix} h_{n''}^{(2)} \left(\frac{\omega}{c} \Delta r \right) Y_{n''}^{m'-m}(\Delta \beta, \Delta \alpha). \quad (\text{E.4})$$

$\mathcal{E}(\cdot)$ is defined in (D.7).

Similar considerations like above yield the translation coefficients $(E|E)$ and $(I|I)$ for exterior-to-exterior and interior-to-interior translation respectively as (Gumerov & Duraiswami, 2004, eq. (3.2.18), (3.2.46); p. 97, 102)

$$(E|E)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega) = (I|I)_{nn'}^{mm'}(\Delta \mathbf{x}, \omega) = \sum_{n''=|n'-n|}^{n'+n} i^{n+n''-n'} \sqrt{\frac{(2n+1)(2n'+1)(2n''+1)}{4\pi}} \times \mathcal{E} \begin{pmatrix} m' & -m & m-m' \\ n' & n & n'' \end{pmatrix} j_{n''} \left(\frac{\omega}{c} \Delta r \right) Y_{n''}^{m'-m}(\Delta \beta, \Delta \alpha). \quad (\text{E.5})$$

Note that every second addend in the summations in (E.4) and (E.5) is zero. This is not explicitly indicated to retain notational clarity.

Eq. (E.4) and (E.5) do not represent the most efficient translation operators. However, they are employed in this thesis since they are the most compact expressions. Refer to (Gumerov & Duraiswami, 2004) for alternatives.

E.2 Recursion Formulae for Exterior-to-Interior Sectorial Translation

As outlined in Sec. 3.3.3, the sectorial translation coefficients $(E|I)_{|m|n'}^{mm'}(\Delta\mathbf{x}, \omega)$ can be computed using (Gumerov & Duraiswami, 2004, eq. (3.2.79), p. 109)

$$\begin{aligned} b_{-m}^m (E|I)_{n', |m|}^{m', m}(\Delta\mathbf{x}, \omega) \\ = b_{n'}^{m'} (E|I)_{n'-1, |m+1|}^{m'+1, m+1}(\Delta\mathbf{x}, \omega) - b_{n'+1}^{-m'-1} (E|I)_{n'+1, |m+1|}^{m'+1, m+1}(\Delta\mathbf{x}, \omega) , \end{aligned} \quad (\text{E.6})$$

for $m \leq 0$ and (Gumerov & Duraiswami, 2004, eq. (3.2.78), p. 108)

$$\begin{aligned} b_m^{-m} (E|I)_{n', m}^{m', m}(\Delta\mathbf{x}, \omega) \\ = b_{n'}^{-m'} (E|I)_{n'-1, m-1}^{m'-1, m-1}(\Delta\mathbf{x}, \omega) - b_{n'+1}^{m'-1} (E|I)_{n'+1, m-1}^{m'-1, m-1}(\Delta\mathbf{x}, \omega) , \end{aligned} \quad (\text{E.7})$$

for $m \geq 0$ with (Gumerov & Duraiswami, 2004, eq. (2.2.10), p. 68)

$$b_n^m = \begin{cases} \sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}} & \text{for } 0 \leq m \leq n \\ -\sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}} & \text{for } -n \leq m < 0 \\ 0 & \text{for } |m| > n . \end{cases} \quad (\text{E.8})$$

E.3 The Stationary Phase Approximation Applied to the Rayleigh I Integral

The objective of this section is approximating the Rayleigh I integral (2.56) in the horizontal plane. Consider the integral over z_0 in (2.56) assuming that the driving function $D(\mathbf{x}_0, \omega)$ is independent from z_0 , thus (Berkhout *et al.*, 1993)

$$\int_{-\infty}^{\infty} \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|} \Big|_{z=0} dz_0 . \quad (\text{E.9})$$

Such an integral can be approximated by the stationary phase approximation (Williams, 1999). The latter provides an approximative solution to integrals of the form

$$I = \int_{-\infty}^{\infty} f(z_0) e^{i\zeta(z_0)} dz_0 \quad (\text{E.10})$$

which is given by

$$I \approx \sqrt{\frac{2\pi i}{\zeta''(z_p)}} f(z_p) e^{i\zeta(z_p)} . \quad (\text{E.11})$$

$\zeta''(z_0)$ denotes the second derivative of $\zeta(z_0)$ with respect to z_0 . z_p denotes the *stationary phase point* which corresponds to the zero of $\zeta'(z_0)$.

In the present case ($z = 0$),

$$f(z_0) = \frac{1}{4\pi} \cdot \frac{1}{\sqrt{(x-x_0)^2 + y^2 + z_0^2}}, \quad (\text{E.12})$$

$$\zeta(z_0) = -\frac{\omega}{c} \sqrt{(x-x_0)^2 + y^2 + z_0^2}, \quad (\text{E.13})$$

$$\zeta'(z_0) = -\frac{\omega}{c} \frac{1}{\sqrt{(x-x_0)^2 + y^2 + z_0^2}} z_0. \quad (\text{E.14})$$

Thus $z_p = 0$.

$$\zeta''(z_0) = -\frac{\omega}{c} \frac{1}{\sqrt{(x-x_0)^2 + y^2 + z_0^2}} + \frac{\omega}{c} \frac{z_0^2}{((x-x_0)^2 + y^2 + z_0^2)^{\frac{3}{2}}} \quad (\text{E.15})$$

so that (Berkhout *et al.*, 1993)

$$\zeta''(z_p) = -\frac{\omega}{c} \frac{1}{\sqrt{(x-x_0)^2 + y^2}}. \quad (\text{E.16})$$

Inserting above results in (E.11) and the result in (2.56) yields the 2^{1/2}-dimensional approximation of the Rayleigh I integral (2.57).

References

- Abramowitz, M., & Stegun, I. A. (eds). 1968. *Handbook of Mathematical Functions*. New York: Dover Publications Inc.
- Ahrens, J., & Spors, S. 2008a. An Analytical Approach to Sound Field Reproduction Using Circular and Spherical Loudspeaker Distributions. *Acta Acustica utd. with Acustica*, **94**(6), 988–999.
- Ahrens, J., & Spors, S. 2008b (May). Focusing of virtual sound sources in higher order Ambisonics. *Paper 7378 of: 124th Convention of the AES*.
- Ahrens, J., & Spors, S. 2009a (Aug.). An Analytical Approach to 2.5D Sound Field Reproduction Employing Circular Distributions of Non-Omnidirectional Loudspeakers. *Pages 814–818 of: 17th European Signal Processing Conference (EUSIPCO)*.
- Ahrens, J., & Spors, S. 2009b (Apr.). An Analytical Approach to Sound Field Reproduction with a Movable Sweet Spot Using Circular Distributions of Loudspeakers. *Pages 273–276 of: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- Ahrens, J., & Spors, S. 2009c (Oct.). On the Secondary Source Type Mismatch in Wave Field Synthesis Employing Circular Distributions of Loudspeakers. *Paper 7952 of: 127th Convention of the AES*.
- Ahrens, J., & Spors, S. 2010a (Mar.). An Analytical Approach To 2.5D Sound Field Reproduction Employing Linear Distributions Of Non-Omnidirectional Loudspeakers. *Pages 105–108 of: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- Ahrens, J., & Spors, S. 2010b (Mar.). An Analytical Approach to 3D Sound Field Reproduction Employing Spherical Distributions of Non-Omnidirectional Loudspeakers. *Pages 1–5 of: IEEE Int. Symp. on Comm., Control and Sig. Proc. (ISCCSP)*.
- Ahrens, J., & Spors, S. 2010c (May). Applying The Ambisonics Approach On Planar And Linear Arrays Of Loudspeakers. *Paper 3 of: 2nd Int. Symp. on Ambisonics and Spherical Acoustics*.
- Ahrens, J., & Spors, S. 2010d. Sound Field Reproduction Using Planar and Linear Arrays of Loudspeakers. *IEEE Trans. on Sp. and Audio Proc.*, **18**(8), 2038–2050.

- Alexander, R. J. 1999. *The Inventor of Stereo: The Life and Works of Alan Dower Blumlein*. Oxford: Focal Press.
- Arfken, G., & Weber, H. 2005. *Mathematical Methods for Physicists*. 6th edition edn. San Diego: Elsevier Academic Press.
- Bamford, J. S. 1995. *An Analysis of Ambisonics Sound Systems of First and Second Order*. M.Sc. thesis, University of Waterloo, Ont. Canada.
- Bamford, J. S., & Vanderkooy, J. 1995 (Oct.). Ambisonic sound for us. *Paper 4138 of: 99th Convention of the AES*.
- Bell, A. G. 1876. *Improvement in Telegraphy*. US patent 174465.
- Berkhout, A. J. 1987. *Applied Seismic Wave Theory*. Amsterdam: Elsevier Publishing Company.
- Berkhout, A. J., de Vries, D., & Vogel, P. 1993. Acoustic control by wave field synthesis. *JASA*, **93**(5), 2764–2778.
- Bertet, S. 2009. *Formats audio 3D hiérarchiques: Caractérisation objective et perceptive des systèmes ambisonics d'ordres supérieurs*. PhD thesis, INSA Lyon. text in French.
- Betlehem, T., & Abhayapala, T. D. 2005. Theory and design of sound field reproduction in reverberant rooms. *JASA*, **117**(4), 2100–2111.
- Blackstock, D. T. 2000. *Fundamentals of Physical Acoustics*. New York: Wiley and Sons, Inc.
- Blauert, J. 1997. *Spatial Hearing*. New York: Springer.
- Bracewell, R. N. 2000. *The Fourier Transform and Its Applications*. Singapore: McGraw-Hill.
- Bureau International des Poids et Mesures. 2006. *The International System of Units (SI)*.
- Colton, D., & Kress, R. 1998. *Inverse Acoustic and Electromagnetic Scattering Theory*. Berlin: Springer.
- Condon, E. U., & Shortley, G. H. 1935. *The Theory of Atomic Spectra*. Cambridge: Cambridge University Press.
- Copley, L. G. 1968. Fundamental Results Concerning Integral Representations in Acoustic Radiation. *JASA*, **44**, 28–32.
- Corteel, E. 2006. Equalization in an Extended Area Using Multichannel Inversion and Wave Field Synthesis. *JAES*, **54**(12), 1140–1161.

- Daniel, J. 2001. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia* [Representations of Sound Fields, Application to the Transmission and Reproduction of Complex Sound Scenes in a Multimedia Context]. PhD thesis, Université Paris 6. text in French.
- Daniel, J. 2003 (May). Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. *Paper 16 of: 23rd International Conference of the AES*.
- Daniel, J., Nicol, R., & Moreau, S. 2003 (Mar.). Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging. *Paper 5788 of: 114th Convention of the AES*.
- de Bruijn, W. 2004. *Application of Wave Field Synthesis in Videoconferencing*. PhD thesis, Delft University of Technology.
- de Vries, D. 1996. Sound Reinforcement by Wavefield Synthesis: Adaptation of the Synthesis Operator to the Loudspeaker Directivity Characteristics. *JAES*, **44**(12), 1120–1131.
- de Vries, D. 2009. *Wave Field Synthesis*. AES Monograph. New York: AES.
- Driscoll, J. R., & Healy, D. M. 1994. Computing Fourier Transforms and Convolutions on the 2-Sphere. *Advances in Applied Mathematics*, **15**(2), 202–250.
- du Moncel, T. 1881. The Telephone at the Paris Opera. *Scientific American*, Dec., 422–423.
- Dutton, G. F. 1962. The Assessment of Two-Channel Stereophonic Reproduction Performance in Studio Monitor Rooms, Living rooms and Small Theatres. *JAES*, **10**(2), 98–105.
- Fazi, F., & Nelson, P. 2007 (Sept.). A theoretical study of sound field reconstruction techniques. *Pages paper ELE-04-002 of: 19th International Congress on Acoustics*.
- Fazi, F., & Nelson, P. 2010 (May). Nonuniqueness of the Solution of the Sound Field Reproduction Problem. *Paper 2 of: 2nd Int. Symp. on Ambisonics and Spherical Acoustics*.
- Fazi, F., Brunel, V., Nelson, P., Hörchens, L., & Seo, J. 2008a (May). Measurement and Fourier-Bessel Analysis of Loudspeaker Radiation Patterns Using a Spherical Array of Microphones. *Paper 7354 of: 124th Convention of the AES*.
- Fazi, F., Nelson, P., & Potthast, R. 2009 (June). Analogies and Differences between 3 Methods for Sound Field Reproduction. *Pages 23–28 of: Ambisonics Symposium*.
- Fazi, F. M., Nelson, P. A., Christensen, J. E. N., & Seo, J. 2008b (Oct.). Surround system based on three dimensional sound field reconstruction. *Paper 7555 of: 125th Convention of the AES*.

- Gabrielsson, A., & Sjøgren, H. 1979. Perceived sound quality of sound-reproducing systems. *JASA*, **65**(4), 1019–1033.
- Gardner, W. G. 1997. *3-D Audio Using Loudspeakers*. PhD thesis, Massachusetts Institute of Technology.
- Gauthier, P.-A., & Berry, A. 2006. Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory. *JASA*, **119**(5), 2721–2737.
- Geier, M., Wierstorf, H., Ahrens, J., Wechsung, I., Raake, A., & Spors, S. 2010 (May). Perceptual Evaluation of Focused Sources in Wave Field Synthesis. *Paper 8069 of: 128th Convention of the AES*.
- Gerzon, M. A. 1973. With-height sound reproduction. *JAES*, **21**, 2–10.
- Gerzon, M. A. 1980 (Feb.). Practical Periphony: The Reproduction of Full-Sphere Sound. *Paper 1571 of: 65th Convention of the AES*.
- Gerzon, M. A. 1992 (Mar.). General Metatheory of Auditory localization. *Paper 3306 of: 92th Convention of the AES*.
- Girod, B., Rabenstein, R., & Stenger, A. 2001. *Signals and Systems*. New York: J.Wiley & Sons.
- Giroire, J. 1982. Integral equation methods for the Helmholtz equation. *Integral Equations and Operator Theory*, **5**(1), 506–517.
- Gjellestad, G. 1955. Note On The Definite Integral Over Products Of Three Legendre Functions. *PNAS*, **41**(Nov.), 954–956.
- Gradshteyn, I. S., & Ryzhik, I. M. 2000. *Table of Integrals, Series, and Products*. San Diego: Academic Press.
- Gumerov, N. A., & Duraiswami, R. 2004. *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Amsterdam: Elsevier.
- Haas, W. 1951. The Influence of a Single Echo on the Audibility of Speech. *Acustica*, **1**, 49–58.
- Hamasaki, K., Hiyama, K., & Okumura, R. 2005 (May). The 22.2 Multichannel Sound System and Its Application. *Paper 6406 of: 118th Convention of the AES*.
- Hammershøi, D., & Møller, H. 2002. Methods for Binaural Recording and Reproduction. *Acustica*, **88**(3), 303–311.
- Hannemann, J., & Donohue, K. D. 2008. Virtual Sound Source Rendering Using a Multipole-Expansion and Method-of-Moments Approach. *JAES*, **56**(6), 473–481.

- Jones, J. P. 2001. Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media. *JASA*, **110**(1), 48–58.
- Kennedy, R. A., Sadeghi, P., Abhayapala, T. D., & Jones, H. M. 2007. Intrinsic Limits of Dimensionality and Richness in Random Multipath Fields. *IEEE Trans. on Sig. Proc.*, **55**(6), 2542–2556.
- Kim, Y., Deille, O., & Nelson, P. A. 2006. Crosstalk cancellation in virtual acoustic imaging systems for multiple listeners. *Journal of Sound and Vibration*, **297**(1-2), 251–266.
- Kirkeby, O., & Nelson, P. A. 1993. Reproduction of plane wave sound fields. *JASA*, **94**(5), 2992–3000.
- Kirkeby, O., Nelson, P. A., Hamada, H., & Orduna-Bustamante, F. 1998. Fast Deconvolution of Multichannel Systems using Regularization. *IEEE Trans. on Sp. and Audio Proc.*, **6**(2), 189–195.
- Kolundžija, M., Faller, C., & Vetterli, M. 2009 (May). Sound Field Reconstruction: An Improved Approach For Wave Field Synthesis. *Paper 7754 of: 126th Convention of the AES*.
- Kraus, K. 2008. *Wigner3j Symbol*. www.mathworks.com/matlabcentral/fileexchange. [Online; accessed 10-Jan-2010].
- Lindau, A., Hohn, T., & Weinzierl, S. 2007 (May). Binaural resynthesis for comparative studies of acoustical environments. *Paper 7032 of: 122nd Convention of the AES*.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. 1999. The precedence effect. *JASA*, **106**(4), 1633–1654.
- Lopez, J. J., Gonzalez, A., & Fuster, L. 2005 (Oct.). Room compensation in wave field synthesis by means of multichannel inversion. *Pages 146–149 of: IEEE Workshop on Appl. of Sig. Proc. to Audio and Acoustics (WASPAA)*.
- Mohlenkamp, M. J. 1999. Fast transform for spherical harmonics. *J. Fourier Analy. Applicat.*, **5**(2/3), 159–184.
- Morse, P. M., & Feshbach, H. 1953. *Methods of Theoretical Physics*. Minneapolis: Feshbach Publishing, LLC.
- Nelson, P. A., & Rose, J. F. W. 2005. Errors in two-point sound reproduction. *JASA*, **118**(1), 193–204.
- Nieto-Vesperinas, M. 2006. *Scattering and Diffraction in Physical Optics*. Singapore: World Scientific Publishing.
- Petrausch, S., Spors, S., & Rabenstein, R. 2005. Simulation and Visualization of Room Compensation for Wave Field Synthesis with the Functional Transformation Method. *Paper 6547 of: 119th Convention of the AES*.

- Poletti, M. A. 2000. A Unified Theory of Horizontal Holographic Sound Systems. *JAES*, **48**(12), 1155–1182.
- Poletti, M. A. 2005. Three-dimensional surround sound systems based on spherical harmonics. *JAES*, **53**(11), 1004–1025.
- Pollow, M., & Behler, G. 2009. Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization. *Acta Acustica utd. with Acustica*, **6**, 1082–1092.
- Pueo, B., Lopez, J. J., Escolano, J., & Bleda, S. 2007. Analysis of Multiactuator Panels in Space-Time Wavenumber Domain. *JAES*, **55**(12), 1092–1106.
- Rabenstein, R., Steffen, P., & Spors, S. 2006. Representation of two-dimensional wave fields by multidimensional signals. *EURASIP Signal Proc. Magazine*, **86**(6), 1341–1351.
- Rafaely, B. 2004. Plane-wave decomposition of the sound field on a sphere by spherical convolution. *JASA*, **116**(4), 2149–2157.
- Rafaely, B., Weiss, B., & Bachmat, E. 2007. Spatial Aliasing in Spherical Microphone Arrays. *IEEE Trans. on Signal Proc.*, **55**(3), 1003–1010.
- Rumsey, F. 2001. *Spatial Audio*. Oxford: Focal Press.
- Rumsey, F. 2002. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *JAES*, **50**(9), 651–666.
- Rumsey, F., Kassier, S., Zieliski R., & Bech, S. 2005. On the relative importance of spatial and timbral fidelities in judgements of degraded multichannel audio quality. *JASA*, **118**(2), 968–976.
- Sanson, J., Corteel, E., & Warusfel, O. 2008 (May). Objective and subjective analysis of localization accuracy in Wave Field Synthesis. *Paper 7361 of: 124th Convention of the AES*.
- Sébilleau, D. 1998. On the computation of the integrated products of three spherical harmonics. *J. Phys. A: Math. Gen.*, **31**, 7157–7168.
- Shirdhonkar, S., & Jacobs, D. 2005 (Oct.). Non-negative Lighting and Specular Object Recognition. *Pages 1323–1330, Vol. 2 of: IEEE Int. Conference on Computer Vision*.
- Sonke, J.-J., Labeeuw, J., & de Vries, D. 1998 (May). Variable Acoustics by Wave-field Synthesis: A Closer Look at Amplitude Effects. *Paper 4712 of: 104th Convention of the AES*.
- Spors, S. 2005. *Active Listening Room Compensation for Spatial Sound Reproduction Systems*. PhD thesis, University of Erlangen-Nuremberg.

- Spors, S. 2006 (Mar.). Spatial aliasing artifacts produced by linear loudspeaker arrays used for wave field synthesis. *Pages 1–4 of: IEEE Int. Symp. on Comm., Control and Sig. Proc.*
- Spors, S. 2008 (Mar.). Investigation of spatial aliasing artifacts of wave field synthesis in the temporal domain. *Pages 223–224 of: 34rd German Annual Conference on Acoustics (DAGA).*
- Spors, S., & Ahrens, J. 2007 (Mar.). Analysis of near-field effects of wave field synthesis using linear loudspeaker arrays. *Paper 29 of: 30th Intern. Conference of the AES.*
- Spors, S., & Ahrens, J. 2008a (Oct.). A Comparison of Wave Field Synthesis and Higher-Order Ambisonics with Respect to Physical Properties and Spatial Sampling. *Paper 7556 of: 125th Convention of the AES.*
- Spors, S., & Ahrens, J. 2008b. Towards a theory for arbitrarily shaped sound field reproduction systems. *JASA*, **123**(5), 3930.
- Spors, S., & Ahrens, J. 2010a (May). Analysis and improvement of pre-equalization in 2.5-dimensional wave field synthesis. *Paper 8121 of: 128th Convention of the AES.*
- Spors, S., & Ahrens, J. 2010b (Mar.). Reproduction of Focused Sources by the Spectral Division Method. *Paper 17.4 of: IEEE Int. Symp. on Comm., Control and Sig. Proc. (ISCCSP).*
- Spors, S., & Rabenstein, R. 2006 (May). Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Arrays used for Wave Field Synthesis. *Paper 6711 of: 120th Convention of the AES.*
- Spors, S., Buchner, H., Rabenstein, R., & Herbordt, W. 2007. Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering. *JASA*, **122**(1), 354–369.
- Spors, S., Rabenstein, R., & Ahrens, J. 2008 (May). The Theory of Wave Field Synthesis Revisited. *Paper 7358 of: 124th Convention of the AES.*
- Spors, S., Wierstorf, H., Geier, M., & Ahrens, J. 2009 (Oct.). Physical and perceptual properties of focused sources in wave field synthesis. *Paper 7914 of: 127th Convention of the AES.*
- Start, E. W. 1996 (May). Application of curved arrays in wave field synthesis. *Paper 4143 of: 100th Convention of the AES.*
- Start, E. W. 1997. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, Delft University of Technology.
- Steinberg, J. C., & Snow, W. B. 1934. Symposium on Wire Transmission of Symphonic Music and Its Reproduction in Auditory Perspective: Physical Factors. *Bell Systems Technical Journal*, **XIII**(2).

- The Chebfun Team. 2009. *The Chebfun Project*. <http://www2.maths.ox.ac.uk/chebfun>. [Online; accessed 09-Dec-2009].
- Theile, G. 1980. *On the localisation in the superimposed soundfield*. PhD thesis, Technische Universität Berlin.
- Theile, G. 2004 (Mar.). Spatial perception in WFS rendered sound fields. *Pages 27–30 of: Proceedings of the Joint Congress CFA/DAGA*.
- Torick, E. 1998. Highlights in the history of multichannel sound. *JAES*, **46**(1/2), 27–31.
- Verheijen, E. N. G. 1997. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, Delft University of Technology.
- Vogel, P. 1993. *Application of Wave Field Synthesis in Room Acoustics*. PhD thesis, Delft University of Technology.
- Wallach, H., Newman, E. B., & Rosenzweig, M. R. 1949. The Precedence Effect in Sound Localization. *American Journal of Psychology*, **57**, 315–336.
- Ward, D. B., & Abhayapala, T. D. 2001. Reproduction of a Plane-Wave Sound Field using an Array of Loudspeakers. *IEEE Trans. on Sp. and Audio Proc.*, **9**(6), 697–707.
- Weisstein, E. W. 2002. *CRC Concise Encyclopedia of Mathematics*. London: Chapman & Hall/CRC.
- Williams, E. G. 1999. *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London: Academic Press.
- Wittek, H. 2007. *Perceptual differences between wavefield synthesis and stereophony*. PhD thesis, University of Surrey.
- Wu, Y. J., & Abhayapala, T. D. 2009. Theory and Design of Soundfield Reproduction using Continuous Loudspeaker Concept. *IEEE Trans. on Audio, Speech and Language Proc.*, **17**(1), 107–116.
- Zayed, A. I. 1993. *Advances in Shannon's Sampling Theory*. New York: CRC Press.
- Zotter, F. 2009a. *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*. Doctoral Thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz.
- Zotter, F. 2009b (Mar.). Sampling Strategies for Acoustic Holography/Holophony on the Sphere. *Pages 1107–1110 of: NAG-DAGA*.
- Zotter, F., & Pomberger, H. 2010 (May). Ambisonic Decoding with and without Mode-Matching: Case Study Using the Hemisphere. *Paper 3 of: 2nd Int. Symp. on Ambisonics and Spherical Acoustics*.

-
- Zotter, F., Pomberger, H., & Frank, M. 2009 (May). An Alternative Ambisonics Formulation: Modal Source Strength Matching and the Effect of Spatial Aliasing. *Paper 7740 of: 126th Convention of the AES.*