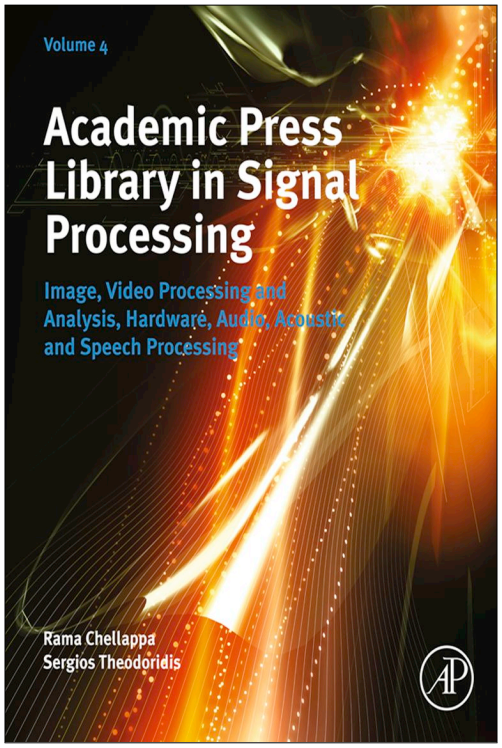


From Rudolf Rabenstein et al., Sound Field Synthesis. In: Rama Chellappa and Sergios Theodoridis, editors, Academic Press Library in Signal Processing. Vol 4, Image, Video Processing and Analysis, Hardware, Audio, Acoustic and Speech Processing, Chennai: Academic Press, 2014, p. 915-979.
ISBN: 978-0-12-396501-1

CHAPTER

Sound Field Synthesis

32

Rudolf Rabenstein[*], **Sascha Spors**[†], **and Jens Ahrens**[‡]

[*] *University Erlangen-Nürnberg, Erlangen, Germany*
[†] *University Rostock, Rostock, Germany*
[‡] *Microsoft Research, Redmond, USA*

4.32.1 Introduction

This chapter treats the topic of sound field synthesis in the context of creation of virtual acoustic environments. The latter as well as the corresponding rendering methods are introduced in this section. Then the organization of this chapter is presented.

4.32.1.1 Virtual acoustic environments

The term *virtual environment* is not well defined and used in a variety of ways. Nevertheless, two aspects appear to be common to all approaches: From the technical viewpoint, virtual environments are created by the convergence of different multimedia technologies. From a user perspective, virtual environments create a sense of immersion, where users perceive no more single displays or loudspeakers but the complete audiovisual scene as an entity.

The prevalent multimedia technologies in virtual environments are video and audio reproduction, where mass-produced hardware and media standards are available. These are sometimes complemented by more experimental devices like data gloves, head-mounted displays, haptic interaction, or alike.

Virtual environments are applied in the entertainment industry for movie reproduction in different formats and for games, in communications for teleconferencing, and in simulation for the evaluation of the design of buildings and machines, for operator training, etc.

The term virtual environment is often used for advanced video technologies only, while sound is considered as an add-on to support the visual content. In contrast, the focus lies here explicitly on the creation of *acoustic virtual environments*. They may be used for audio reproduction in its own right or in conjunction with video reproduction.

The technologies for creating acoustic virtual environments can be roughly classified into *head related methods* and *room related methods* [1]. Head related methods attempt to create the proper acoustic signals at both ears of one listener. The set of methods employed for this purpose is called binaural technology. They are addressed briefly in the beginning of Section 4.32.1.2.

Room related methods establish a sound field within a room where one or more listeners can sit or walk around. A classical approach is to pan a sound source between one or more pairs of loudspeakers like in two-channel stereo and extensions thereof. More recent approaches to sound field synthesis apply

Academic Press Library in Signal Processing. http://dx.doi.org/10.1016/B978-0-12-396501-1.00032-7

a high number of reproduction channels which are treated as spatial samples of a continuous source distribution around the listening area [2]. Finally, the theory of multiple-input, multiple-output systems gives rise to the family of multipoint methods. Section 4.32.1.2 reviews aspects of these room related approaches.

From all room related methods, the approaches to sound field synthesis are closest to the reproduction of the physically correct sound field. Their presentation constitutes the main focus of this chapter.

4.32.1.2 Rendering of virtual acoustic environments

The technical process of creating an acoustic virtual environment is also called rendering in parallel to the rendering of visual scenes on video displays. Some aspects of head related and room related rendering methods are briefly reviewed here. The presentation starts with a glimpse on head-related transfer functions. Then various panning methods from classical stereo reproduction to more recent approaches are discussed. Sound field synthesis is put in perspective to the other room related methods here, but the resulting rendering methods are discussed in more depth in the remainder of the chapter. The overview on room related methods is concluded by a review of multipoint methods.

4.32.1.2.1 Reproduction based on head-related transfer functions

The human auditory system exploits the acoustic characteristics of the outer ear in order to perform spatial scene analysis [3]. The outer ear includes the pinnae, head and upper torso. The acoustic properties from the outer ear can be captured by measuring the transfer function from an acoustic source to a defined position in the ear canal of both ears. These functions are known as head-related transfer functions (HRTFs) and are individual for one person. HRTFs are measured in anechoic, reverberant or simulated environments. Virtual sound sources are created by filtering the signal of the virtual source by the its left and right ear HRTF, and reproducing the resulting signals via headphones in order to have independent control over the signals at the two eardrums. For proper auralization also the transfer function of the headphone has to be compensated for. This reproduction method is sometimes also referred to as *binaural reproduction*.

The HRTFs depend on the position of the listener's ears, the position of the acoustic source, the acoustic properties of the source and the environment. Hence, a database of HRTFs is required that covers all potential parameters. In order to limit the effort, often only the head orientation in the horizontal plane for one listener and source position is taken into account. For auralization, the head orientation has to be tracked and the corresponding set of HRTFs has to be applied. A head tracked system is typically termed as *dynamic binaural reproduction* system. The benefit of binaural reproduction is its low complexity in conjunction with the high spatial quality that can be achieved by a well designed system. Drawbacks are that HRTFs are listener dependent and that binaural reproduction gets complex for multiple listeners since head-tracking and appropriate HRTF databases are required for each listener (position). Furthermore, the reproduction via headphones decouples the listener from their environment so that e.g., the communication with other people in the same room becomes unpleasant.

Binaural reproduction can also be performed via loudspeakers. Here, an appropriate crosstalk cancellation [4] has to be employed in order to control the signals at both ears of the listener independently. Crosstalk cancelation typically exhibits a very pronounced sweet spot and is likely to produce strong artifacts at positions some few centimeters from the sweet spot.

4.32.1.2.2 Panning approaches

Panning approaches apply amplitude differences (and/or time delays) to a low number of loudspeakers (typically a pair, triple, or quadruple) in order to create the impression of phantom sources. Stereophony is the most widespread variant of these approaches which is based on the use of a pair of loudspeakers. However, generalizations of the basic concept have been developed like e.g., vector base amplitude panning (VBAP) [5] and Ambisonics amplitude panning (AAP) [6]. Although all these techniques are partly physically motivated, it is agreed nowadays that their success can be exclusively attributed to psychoacoustic properties of the human auditory system.

A number of studies have been conducted in order to clarify the perception of Stereophony, see e.g., [7] for references. The assumed underlying psycho-acoustical mechanism is termed *summing localization*, e.g., [7, p. 9] and [3, p. 204]. Summing localization refers to the superposition of (typically a low number of) sound fields carrying sufficiently coherent signals impinging at a time interval smaller than approximately 1 ms. It is assumed that the superposition of the sound fields at the listener's ears leads to summed signals, the components of which can not be discriminated by the human hearing system. An extension of the concept of summing location is Theile's *association theory* published ibidem.

Panning approaches are typically realized by applying weights (and/or delays) to each loudspeaker according to a given panning law. These panning laws are derived by either considering the physics of the problem (e.g., the sine law) or by psychoacoustic experiments. A major benefit of panning approaches is their low complexity. However it is well known, that they exhibit a very pronounced sweet spot and that the impression of a phantom source at lateral and rear positions is unreliable for typical setups [5]. Outside of the sweet spot, the spatial impression is heavily distorted and also some impairment in terms of sound color may occur [7]. Hence, these techniques are only suitable for small audiences if spatial sound reproduction with high resolution is desired.

4.32.1.2.3 Sound field synthesis

In order to provide the potential of satisfying a significantly larger receiver area that above mentioned approaches allow for, methods have been proposed that aim at the physical synthesis of a given sound field. These methods are termed *sound field synthesis* [2].

The problem of sound field synthesis may be formulated in words as follows:

A given ensemble of elementary sound sources shall be driven such that the superposition of the sound fields emitted by the individual elementary sound sources best approximates a sound field with given desired properties over an extended area.

Such an extended area may be a volume or a surface. The employed elementary sound sources will be termed *secondary sources* in the remainder of this chapter [8, e.g., p. 106]. In practical implementations, loudspeakers will be used as secondary sources. The term "secondary source" has been established in the context of scattering problems where the influence of a given object on an incident field is described by a distribution of secondary sources that are located along the surface of the object and that replace the latter, e.g., [9].

In order to facilitate the mathematical treatment and in order to facilitate the exploitation of results that have been achieved in closely related problems such as acoustical scattering [9], the ensemble of secondary sources under consideration will be assumed to be continuous and will therefore be referred to as a *distribution of secondary sources*.

Three basic analytic approaches have been proposed in the literature based on the framework outlined above. They are termed Near-field Compensated Higher-Order Ambisonics (NFC-HOA), Spectral Division Method (SDM), and Wave Field Synthesis (WFS), respectively, and they will all be presented later in this chapter.

4.32.1.2.4 *Multipoint approaches*

The approaches to sound field synthesis introduced so far are based upon a spatially continuous formulation of the underlying physical problem. The spatial sampling is typically introduced at a later stage into the driving function. Besides these approaches a number of approaches exist that are based on a spatially discrete formulation of the problem. The published approaches differ in terms of their discretization scheme and signal domain in which the problem is formulated. Most of the approaches assume a spatially discrete secondary source and receiver distribution. Typically a number of receiver points are defined which are located around or within the desired listening area. The pressure field of the virtual source should be synthesized as accurately as possible at these points by a weighted combination of the secondary sources. This is often formulated in terms of a matrix equation that can be understood as a discretized version of the synthesis equation (32.134) that constitutes the starting point for all analytic approaches. The desired pressure at the defined receiver points and the driving function for the secondary sources are combined into vectors. The acoustic paths from all secondary sources to all receiver points are characterized by a matrix of transfer functions. The resulting set of linear equations in form of a matrix equation is then solved with respect to the vector of driving functions. Typically least-square error (LSE) approaches are used. The basic scheme or variants are formulated either in the time domain, in the frequency domain [10], or in terms of spherical harmonics [11]. For the latter methods often only the secondary source distribution is explicitly spatially sampled. The traditional formulation of Higher Order Ambisonics [12] is a prominent example.

The properties of the multipoint approaches are typically somewhere between NFC-HOA, SDM, and WFS and depend heavily on the number of secondary sources, receiver points and their spatial configuration.

The major benefit of the multipoint methods are the potentially very flexible loudspeaker layouts for which an approximate (least-squares) solution can be found. As for the solution of the integral equation (32.134), also the solution of the matrix equation in the multipoint approach poses an inverse problem in acoustics. It is well known that in practice these are often ill-conditioned at high frequencies and the result therefore becomes unpredictable [13]. Besides this problem, the drawbacks of the multipoint approaches are numerical complexity and that they provide only little insight into fundamental properties of the reproduced sound field.

4.32.1.3 Organization of this chapter

The remainder of this chapter discusses methods for sound field synthesis. It starts with the foundations from physics that lead to the acoustic wave equation in Section 4.32.2. Then the focus shifts from sound waves to signals that carry space-time information and their representations in different coordinate systems and in the frequency domain in Section 4.32.3. The response to sound sources and its description by the Green's function is presented in Section 4.32.4. Section 4.32.5 continues with the Kirchhoff-Helmholtz integral equation as the physical foundation of sound field synthesis.

The following sections present three distinct methods for sound field synthesis: Near-field Compensated Higher Order Ambisonics in Section 4.32.6, the Spectral Division Method in Section 4.32.7, and Wave Field Synthesis in Section 4.32.8. These sections share a common structure: a description of the rendering method on the basis of a continuous source distribution, the introduction of discrete loudspeaker positions by spatial sampling and application examples, and extensions of the basic principle.

4.32.2 Acoustic wave equation

The acoustic wave equation is introduced here at first in a coordinate free representation. Then those coordinate systems are presented which are most frequently used for the description of sound fields.

4.32.2.1 Coordinate free representation

Detailed derivations of the acoustic wave equation can be found in many classical and modern books on acoustics [14–20]. A concise derivation with a clear indication of the involved assumptions and simplifications is found e.g., in [19]. Therefore only a few remarks on the underlying physical principles are given here.

The description of wave propagation in a fluid involves three field quantities, the sound pressure $p(\boldsymbol{x}, t)$, the particle velocity $\boldsymbol{v}(\boldsymbol{x}, t)$, and the mass density $\varrho(\boldsymbol{x}, t)$. They depend on the space coordinate vector $\boldsymbol{x}$ and on time t. Three different relations for each pair of variables can be established from the first principles of physics:

- the conservation of mass which links the mass density $\varrho(\boldsymbol{x}, t)$ and the particle velocity $\boldsymbol{v}(\boldsymbol{x}, t)$,
- the equations for the thermodynamical state of the fluid involving the mass density $\varrho(\boldsymbol{x}, t)$ and the pressure $p(\boldsymbol{x}, t)$,
- the equation of motion (Newton's second law) for the sound pressure $p(\boldsymbol{x}, t)$ and the particle velocity $\boldsymbol{v}(\boldsymbol{x}, t)$.

Figure 32.1 shows the interrelations between these three field quantities and the governing physical laws in graphical form.

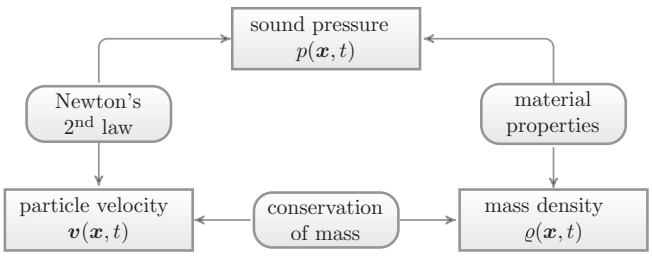


FIGURE 32.1

Fundamental physical principles for the description of sound propagation.

The propagation medium for sound waves is assumed to be air at standard conditions. It can be regarded as an ideal gas for the purpose of room acoustics and for audible frequencies. This assumption simplifies the relations for the thermodynamical state to the extent that variations in the mass density can be expressed by variations of the sound pressure. Then the mass density ϱ can be eliminated such that two expressions for the sound pressure p and the particle velocity $\boldsymbol{v}$ remain.

One relation results from the conservation of mass and the assumption of an ideal gas. It establishes the relation between the compression of the air and the velocity gradient

$$\frac{1}{c^2}\frac{\partial p(\boldsymbol{x},t)}{\partial t} + \varrho_0 \nabla \boldsymbol{v}(\boldsymbol{x},t) = 0. \tag{32.1}$$

Here ϱ_0 denotes the static density of the air and c the sound speed.

The other relation is Newton's second law in differential form (Euler's law). It links the gradient of the sound pressure $p(\boldsymbol{x},t)$ and the acceleration of the air particles

$$\nabla p(\boldsymbol{x},t) + \varrho_0 \frac{\partial \boldsymbol{v}(\boldsymbol{x},t)}{\partial t} = 0. \tag{32.2}$$

Finally the scalar homogeneous wave equation is obtained by eliminating the vector of the particle velocity $\boldsymbol{v}$ from (32.1) and (32.2). Taking the partial time derivative of (32.1) and the gradient of (32.2) results in

$$\nabla^2 p(\boldsymbol{x},t) - \frac{1}{c^2}\frac{\partial^2 p(\boldsymbol{x},t)}{\partial t^2} = 0. \tag{32.3}$$

The square of the Nabla-Operator for the spatial derivative $\nabla^2 = \Delta$ is the Laplace operator. It takes different forms depending on the chosen coordinate system.

4.32.2.2 Coordinate systems

4.32.2.2.1 *Introduction*

The ability to convert between different systems of coordinates allows to find the most suitable one for a given problem. In many cases it is possible to exploit certain spatial symmetries of typical wave fields and to use a coordinate system which reflects these symmetries in the most simple way. An example is the use of spherical coordinates for the sound field of a single point source (acoustic monopole). The radial symmetry of the problem leads to a sound field which is constant with respect to the azimuth and zenith angle of spherical coordinates. In other cases, the choice of the coordinate system is induced by the shape of the enclosure. The wave propagation in a cylindrical pipe like the bore of a wind instrument is an example for the favorable use of cylindrical coordinates. Aligning the longitudinal axis with the center of the bore brings the walls of the enclosure to a constant value of the radial axis. The following section gives a short overview on frequently used coordinate systems.

4.32.2.2.2 *Overview on frequently used spatial coordinate systems*

A schedular overview on the most frequently used spatial coordinate systems is given in Table 32.1.

Cartesian coordinates are most frequently used when no inherent spatial structure of a spatial region of interest is given. An example are world coordinates for general audiovisual scenes in virtual reality.

Table 32.1 Frequently Used Coordinate Systems in Two and Three Spatial Dimensions

Two-dimensional	Three-dimensional
Cartesian	Cartesian
Polar	Cylindrical
	Spherical

Cartesian coordinates allow object and viewpoint manipulations like translation or rotation with simple linear transformations. Also the transition between spatially two-dimensional and three-dimensional representations is easily accomplished by projections from 3D to 2D.

Cartesian coordinates are also the most suitable choice for spatial structures with linear or planar shape. Examples are shoebox models of acoustic spaces for simplified analytical modal investigations or for the mirror image source method in room acoustics. Another natural choice for Cartesian coordinates are line or planar arrays of microphones or loudspeakers, where the alignment of the spatial structure with the coordinate axes is obvious.

For user centric systems, *polar* or *spherical* coordinate systems are more suitable. They define acoustic events relative to a listener who resides naturally in the center of the coordinate systems. Sound sources are then defined by their distance and by their azimuth and zenith angles in the 3D case. For sound events in a horizontal plane around the listener's ears, a 2D polar coordinate system is often sufficient.

Polar and spherical coordinates are frequently used to describe source localization and for the definition of head related transfer functions (HRTFs). The Ambisonics sound reproduction method is based on a sound field description with spherical harmonics.

Cylindrical coordinates are suitable for the descriptions of cylindrical waveguides. They are used in musical acoustics to describe woodwind instruments, organ pipes and alike. In sound reproduction, cylindrical coordinates are of theoretical value to describe the transition between 3D and 2D descriptions of sound fields.

4.32.2.2.3 Cartesian coordinates

In Cartesian coordinates the three-dimensional space coordinates are denoted by

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{32.4}$$

The Cartesian coordinates for two dimensions are given in the same way for $z = 0$.

The Laplace operator in Cartesian coordinates consists of the second order derivatives calculated along each orthogonal spatial dimension

$$\nabla^2 = \Delta = \text{div}\,\text{grad} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{32.5}$$

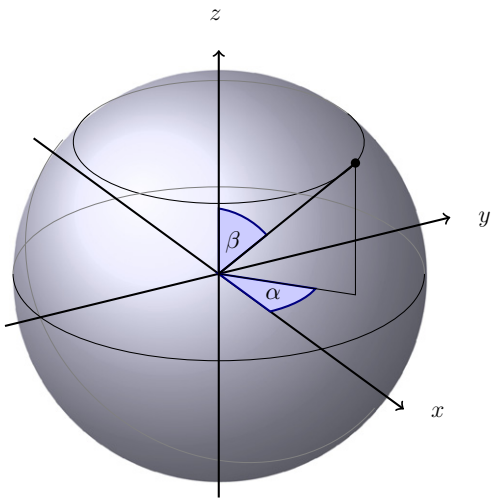


FIGURE 32.2

Spherical coordinates.

4.32.2.2.4 *Spherical coordinates*

Spherical coordinates specify the distance r from the origin, an azimuth angle α in the horizontal plane and a zenith angle β for the elevation, see Figure 32.2. Spherical coordinates are expressed in terms of Cartesian coordinates as

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} \cos\alpha \sin\beta \\ \sin\alpha \sin\beta \\ \cos\beta \end{bmatrix}, \tag{32.6}$$

with the inverse relations

$$r^2 = |\boldsymbol{x}|^2 = x^2 + y^2 + z^2, \tag{32.7}$$

$$\tan\alpha = \frac{y}{x}, \tag{32.8}$$

$$\cos\beta = \frac{z}{r}, \tag{32.9}$$

and the vector notation

$$\boldsymbol{r} = \begin{bmatrix} r \\ \alpha \\ \beta \end{bmatrix}. \tag{32.10}$$

The Laplace operator in spherical coordinates is given by

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\beta}\frac{\partial}{\partial \beta}\left(\sin\beta \frac{\partial}{\partial \beta}\right) + \frac{1}{r^2 \sin^2\beta}\frac{\partial^2}{\partial \alpha^2}. \tag{32.11}$$

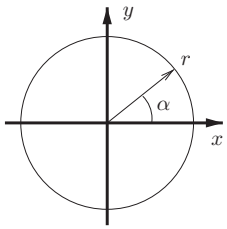


FIGURE 32.3

Polar coordinates.

The second order partial differentials can be written in different ways. For example the first term may appear as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}p(r)\right) = \frac{\partial^2}{\partial r^2}p(r) + \frac{2}{r}\frac{\partial}{\partial r}p(r) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rp(r)). \tag{32.12}$$

The equivalence of these three expressions is shown with the chain rule and the product rule of differentiation.

4.32.2.2.5 *Polar coordinates*

Polar coordinates in space are a two-dimensional representation for the distance r from the origin and a rotation by the angle α (see Figure 32.3). They correspond to the spherical coordinates with $\beta = \frac{\pi}{2}$. The relations to the two-dimensional Cartesian coordinates are given by

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = r\begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \tag{32.13}$$

with the inverse relations

$$r^2 = |\boldsymbol{x}|^2 = x^2 + y^2, \tag{32.14}$$

$$\tan\alpha = \frac{y}{x}. \tag{32.15}$$

For a concise notation, the polar coordinates are also arranged in vector fashion

$$\boldsymbol{r} = \begin{bmatrix} r \\ \alpha \end{bmatrix}. \tag{32.16}$$

The Laplace operator in polar coordinates is given by

$$\Delta = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\alpha^2}. \tag{32.17}$$

4.32.3 Signal representations

This section provides links between the physical quantities of the acoustic wave equation and between multidimensional signals and their numerous representations. These representations distinguish between signals in the time and space domain and between signals in the associated frequency and wavenumber domains. Furthermore, space- or wavenumber-dependent signals in two or three spatial dimensions can be represented in various spatial coordinate systems.

The resulting multitude of different signal representations is not easy to handle with regards to the mathematical notation. This chapter follows the good practice in signal processing to distinguish between signals in time, frequency, and possibly other domains. The corresponding notational details are introduced in due course. They include small and capital letters, subscripts and superscripts, the tilde and other graphemes, although such signal "decorations" are sometimes deemed superfluous in the literature on acoustics (see e.g., [21, Section A.3]).

4.32.3.1 Introduction

The discussion of the acoustic wave equation in Section 4.32.2 relied on the sound pressure and the particle velocity as physical quantities. They served to establish fundamental relations like conservation of mass or Newton's law of motion. However, for the purpose of sound rendering, also another aspect of these quantities is of importance. The temporal and spatial variations of the sound pressure carry information from a sound source to the listeners. This information may be explicit such as speech or musical notes or it may be implicit such as perceived genre or timbre. In any way, the sound pressure governed by the acoustical wave equation is not only a physical quantity but also a signal in the sense of communications. While Section 4.32.2 considered the sound pressure as a physical quantity, this section and the following ones emphasize the signal character.

Signals in communications are mostly one-dimensional and time dependent, e.g., the varying voltage received by an antenna or picked up by a microphone. Signals of this kind are not only represented by their temporal variations but also—equivalently—by their frequency content. The connection between the time and the frequency domain is either provided by the phasor approach or by integral transformations like the Fourier or the Laplace transformation.

- The phasor approach considers monofrequent signals of the form

$$u_1(t) = U e^{i\omega_1 t} \tag{32.18}$$

  with the imaginary unit i, the angular frequency ω_1, and the complex amplitude U. The signal $u_1(t)$ depends on the time variable t while the angular frequency ω_1 is an arbitrary but fixed parameter. Different values of ω_1 define different signals $u_1(t)$. To indicate clearly the different nature of time and frequency, the fixed angular frequency carries the index 1 and the function $u_1(t)$ shares the same index.
- The transformation approach works with the equivalence of a signal and its spectrum

$$u(t) \quad \circ\!\!-\!\!\bullet \quad U(\omega). \tag{32.19}$$

The transformation symbol ○—● is a shorthand notation for the Fourier transformation $U(\omega) = \mathcal{F}\{u(t)\}$. A formal definition is given in Section 4.32.3.3. Here the Fourier transform $U(\omega)$ depends on the frequency variable ω. As a free variable, ω is not indexed.

For signals which depend only on time the dual character of a signal and its spectrum in the form of (32.19) is quite simple and easy to handle. This situation is different for the sound pressure $p(\boldsymbol{x}, t)$ which depends on time and space and thus constitutes a multidimensional signal. Therefore two different transformations are required, one for the time and one for the space variable. The latter has to consider the number of spatial dimensions. In addition the nature of the spatial transformation depends on the chosen spatial coordinate system (Cartesian, polar, spherical).

These different representations are introduced now as extensions of the well-known phasor concept. It turns out that its generalization to time- and space-dependent signals leads to another basic concept, the so-called plane wave. This extension of the phasor approach to the sound pressure is presented in Section 4.32.3.2. Section 4.32.3.3 introduces the set of transformations required for space variables in Cartesian coordinates. Extensions to two-dimensional spatial problems in polar coordinates are given in Section 4.32.3.4 and to three-dimensional spatial problems in spherical coordinates in Section 4.32.3.5.

4.32.3.2 The phasor approach for the wave equation

4.32.3.2.1 *One-dimensional phasors*

Phasors are eigenfunctions of one-dimensional linear and time-invariant (LTI) systems. They are signals of the form

$$u_1(t) = U(\omega_1)e^{i\omega_1 t} \tag{32.20}$$

with the angular frequency ω_1 and the complex amplitude U as in (32.18). The response of an LTI-system with the complex frequency response $H(\omega)$ to a phasor with the amplitude $U = 1$ is also a phasor with a complex amplitude $H(\omega_1)$ and the same angular frequency, see Figure 32.4.

The frequency response $H(\omega)$ is a function of the angular frequency ω. When the corresponding system is excited by a monofrequent signal then only its value $H(\omega_1)$ at the excitation frequency ω_1 determines the output signal.

4.32.3.2.2 *Multidimensional phasors*

The analysis of space-time systems requires to extend the phasor concept to multiple dimensions. Similar to (32.20), a phasor for a system depending on time t and space $\boldsymbol{x}$ has the form

$$u_{01}(\boldsymbol{x}, t) = \tilde{U}(\boldsymbol{k}_0, \omega_1)e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})}. \tag{32.21}$$

Again, ω_1 is the angular frequency with respect to time. Its counterpart with respect to space is the *wave vector* $\boldsymbol{k}_0$ which is of the same dimension as the vector of space coordinates $\boldsymbol{x}$. Similar to the

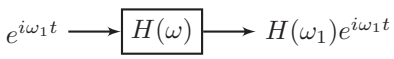


FIGURE 32.4

Response of an LTI-system to a phasor.

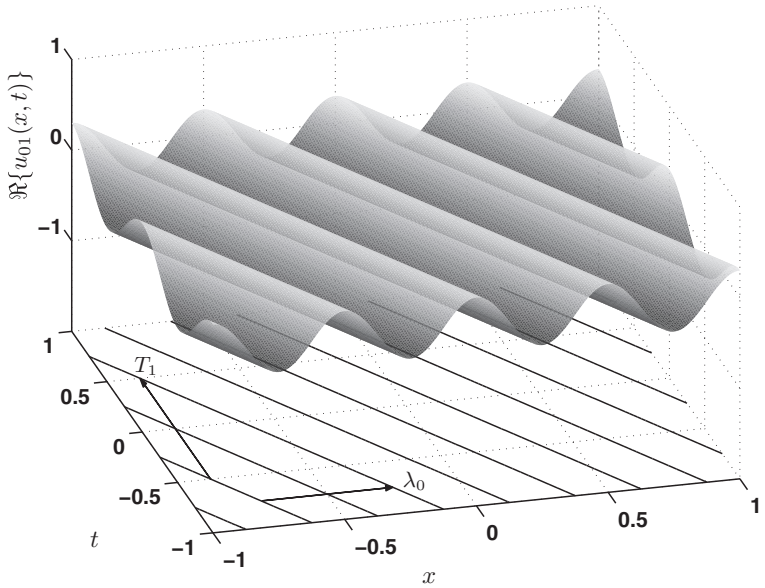


FIGURE 32.5

Real part of the (1+1)D multidimensional phasor $u_{01}(x, t)$ from (32.21). The figure clearly shows the periodicity both in time and space.

one-dimensional phasor, the angular frequency ω_1 and the wave number $\boldsymbol{k}_0$ are fixed values and denoted by an index. The tilde in $\tilde{U}(\boldsymbol{k}_0, \omega_1)$ indicates that $\tilde{U}$ is a function of both spatial and temporal frequencies.

The wave vector $\boldsymbol{k}_0$ can be expressed by its magnitude, the *wave number* k_0, and the unit length vector $\boldsymbol{n}_0$ as

$$\boldsymbol{k}_0 = k_0 \boldsymbol{n}_0 \quad \text{with} \quad k_0 = |\boldsymbol{k}_0| \quad \text{and} \quad \boldsymbol{n}_0 = \frac{1}{k_0}\boldsymbol{k}_0. \tag{32.22}$$

Due to the periodicity of the complex exponential function in (32.21), also $u_{01}(\boldsymbol{x}, t)$ is periodic both in time and space

$$u_{01}(\boldsymbol{x}, t) = u_{01}(\boldsymbol{x} + \lambda_0 \boldsymbol{n}_0, t + T_1), \tag{32.23}$$

with the period T_1 and the wavelength λ_0

$$T_1 = \frac{2\pi}{\omega_1}, \quad \lambda_0 = \frac{2\pi}{k_0}. \tag{32.24}$$

This periodicity is obvious from Figure 32.5 for one spatial dimension, where only the real part $\cos(\omega_1 t + k_0 x)$ is shown.

4.32.3.2.3 Multidimensional phasors and the acoustic wave equation

For sound rendering, the relation of multidimensional phasors to the acoustic wave equation is of special interest. It turns out that the multidimensional phasors of the form of (32.21) are solutions of the wave equation, if certain relations between angular frequency and wave number hold. To see these relations,

insert (32.21) into the acoustic wave Eq. (32.3). Carrying out the derivations with respect to time and space gives

$$\left[k_0^2 - \left(\frac{\omega_1}{c}\right)^2\right] u_{01}(\boldsymbol{x}, t) = 0. \tag{32.25}$$

There exists a nontrivial solution $u_{01}(\boldsymbol{x}, t)$ only iff

$$\omega_1 = \pm c\, k_0. \tag{32.26}$$

This close tie between the angular frequencies ω_1 and k_0 of the multidimensional phasor is imposed by the acoustic wave equation. It is called the *dispersion relation*.

4.32.3.2.4 Physical interpretation of multidimensional phasors

A closer look at $u_{01}(\boldsymbol{x}, t)$ under the restriction (32.26) shows that $u_{01}(\boldsymbol{x}, t) = \text{const}$ for

$$\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x} \pm ct = \text{const}. \tag{32.27}$$

For each of the two signs, (32.27) describes a plane in space which propagates with the speed c in or against the direction of the normal vector $\boldsymbol{n}_0$ as shown in Figure 32.6. Therefore a solution of the wave equation with the arbitrary complex amplitude $\tilde{P}$

$$p_{01}(\boldsymbol{x}, t) = \tilde{P}(\boldsymbol{k}_0, \omega_1) e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})} \tag{32.28}$$

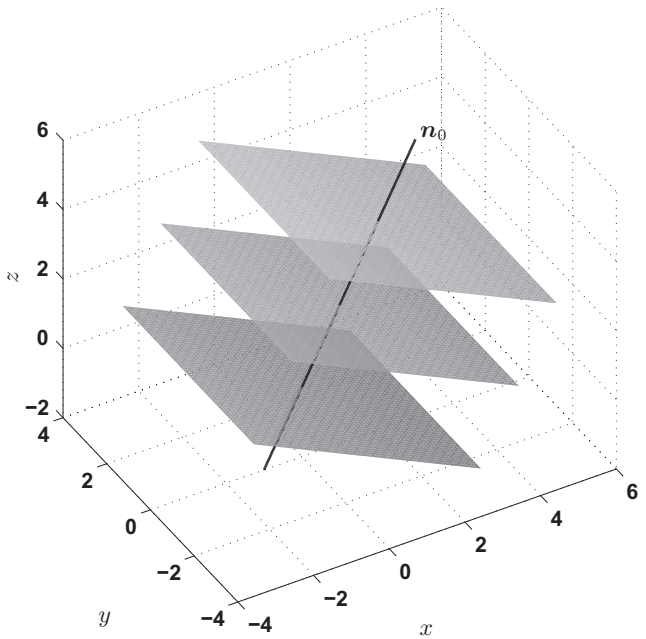


FIGURE 32.6

Plane in space which propagates with the speed c in the direction of the normal vector $\boldsymbol{n}_0$.

is called a *plane wave* or due to its complex exponential nature also a *harmonic plane wave*. Physics-oriented texts sometimes emphasize the analogy to monofrequent light by the designation *monochromatic plane wave*. With the dispersion relation (32.26), the plane wave can be written in various forms ($\tilde{P} = 1$)

$$p_{01}(\boldsymbol{x}, t) = e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})} = e^{ik_0(ct + \boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x})} = e^{i\omega_1(t + \frac{1}{c}\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x})}. \tag{32.29}$$

The periodicity of the plane wave in time and space is reflected by the periodicity of the complex exponential either in dimensionless variables $\omega_1 t$ and $\boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x}_0$, with respect to time, or with respect to space, as shown in (32.29)

$$\begin{aligned} p_{01}(\boldsymbol{x}, t) &= p_{01}(\boldsymbol{x}, t + T_1) = p_{01}(\boldsymbol{x} + \lambda_0 \boldsymbol{n}_0, t) \\ &= e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})} = e^{i\omega_1(t + T_1 + \frac{1}{c}\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x})} = e^{ik_0(ct + \boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x} + \lambda_0)}. \end{aligned} \tag{32.30}$$

4.32.3.3 Fourier transformations in time and space

The phasor concept introduced in Section 4.32.3.2 does not only describe monofrequent signals and plane waves in an elegant way. It provides also the link to general kinds of signals and their spectral representations with respect to time and space. These relations are first established for the Fourier transformation with respect to time and then with respect to space.

4.32.3.3.1 Fourier transformation with respect to time

Consider the one-dimensional phasor $u_1(t)$ from (32.20)

$$u_1(t; \omega_1) = U(\omega_1) e^{i\omega_1 t} \tag{32.31}$$

which describes a monofrequent signal with angular frequency ω_1. Since the complex amplitude U may vary with different angular frequencies, it is written here as a function of ω_1. Furthermore, ω_1 is added as parameter to the list of variables in $u_1(t; \omega_1)$. Variables and parameters are separated by a semicolon.

A general time dependent signal $u(t)$ can be regarded as a superposition of monofrequent signals $u_1(t; \omega_1)$ with different angular frequencies ω_1 from the range $-\infty < \omega_1 < \infty$ with different corresponding complex amplitudes $U(\omega_1)$

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u_1(t; \omega_1) d\omega_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(\omega_1) e^{i\omega_1 t} d\omega_1. \tag{32.32}$$

The special property of the complex exponential function

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(\omega_1 - \omega)t} dt = \delta(\omega_1 - \omega) \tag{32.33}$$

allows to invert the relation (32.32) as

$$\int_{-\infty}^{\infty} u(t) e^{-i\omega t} dt = \int_{-\infty}^{\infty} U(\omega_1) \delta(\omega_1 - \omega) d\omega_1 = U(\omega). \tag{32.34}$$

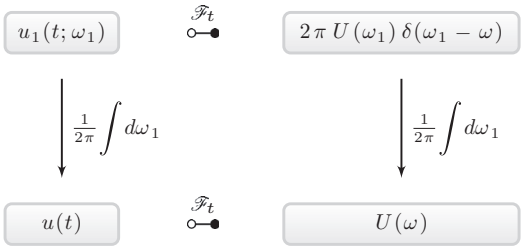


FIGURE 32.7

Representation of the Fourier transformation in time as integration of a phasor with respect to angular frequency ω_1.

Equations (32.32) and (32.34) constitute the Fourier transform pair

$$U(\omega) = \mathscr{F}_t\{u(t)\} = \int_{-\infty}^{\infty} u(t)e^{-i\omega t}dt, \tag{32.35}$$

$$u(t) = \mathscr{F}_t^{-1}\{U(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} U(\omega)e^{i\omega t}d\omega. \tag{32.36}$$

The subscript t in $\mathscr{F}_t$ denotes the Fourier transformation with respect to time. Its relation to the one-dimensional phasor by an integration with respect to frequency is shown in Figure 32.7. The Fourier transform of the phasor in the upper right corner follows from the orthogonality (32.33). The derivation of this well-known relation is presented here as an introduction to the more involved Fourier transformation with respect to space.

4.32.3.3.2 *Fourier transformation with respect to space*

Similar as in the preceding section, the derivation of the Fourier transformation with respect to space starts with the multidimensional phasor from (32.21)

$$u_{01}(\boldsymbol{x}, t; \boldsymbol{k}_0, \omega_1) = \tilde{U}(\boldsymbol{k}_0, \omega_1)e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})}. \tag{32.37}$$

As in (32.31), the dependence of the complex amplitude $\tilde{U}(\boldsymbol{k}_0, \omega_1)$ on the wave vector $\boldsymbol{k}_0$ and the angular frequency ω_1 has been made explicit in the notation.

In a first step similar to (32.32) all phasors with the same wave vector $\boldsymbol{k}_0$ but with varying angular frequency ω_1 are used to describe the superposition

$$u_0(\boldsymbol{x}, t; \boldsymbol{k}_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{U}(\boldsymbol{k}_0, \omega_1)e^{i(\omega_1 t + \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x})}d\omega_1 = \mathscr{F}_t^{-1}\{\tilde{U}(\boldsymbol{k}_0, \omega_1)e^{i\boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x}}\}. \tag{32.38}$$

Inverting (32.38) gives

$$\tilde{U}(\boldsymbol{k}_0, \omega_0)e^{i\boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x}} = \mathscr{F}_t\{u_0(\boldsymbol{x}, t; \boldsymbol{k}_0)\}. \tag{32.39}$$

In a second step the superposition is extended to all possible values of the elements of the wave vector $\boldsymbol{k}_0$

$$U(\boldsymbol{x},\omega) = \frac{1}{(2\pi)^d}\int_{-\infty}^{\infty} \mathscr{F}_t\{u_0(\boldsymbol{x},t;\boldsymbol{k}_0)\}d\boldsymbol{k}_0 = \frac{1}{(2\pi)^d}\int_{-\infty}^{\infty} \tilde{U}(\boldsymbol{k}_0,\omega_1)e^{i\boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x}}d\boldsymbol{k}_0. \tag{32.40}$$

The number of spatial dimensions is equal to d, $d = 1, 2, 3$; the integrals are understood as multiple integrals over all components of the wave vector $\boldsymbol{k}_0$.

Using the relation of the complex exponential function in multiple dimensions (compare (32.33))

$$\frac{1}{(2\pi)^d}\int_{-\infty}^{\infty} e^{i(\boldsymbol{k}_0-\boldsymbol{k})^{\mathrm{T}}\boldsymbol{x}}d\boldsymbol{x} = \delta(\boldsymbol{k}_0-\boldsymbol{k}) \tag{32.41}$$

allows to turn (32.40) into

$$\int_{-\infty}^{\infty} U(\boldsymbol{x},\omega)e^{-i\boldsymbol{k}^{\mathrm{T}}\boldsymbol{x}}d\boldsymbol{x} = \tilde{U}(\boldsymbol{k},\omega). \tag{32.42}$$

The results (32.40) and (32.42) constitute the spatial Fourier transformation in multiple dimensions

$$\tilde{U}(\boldsymbol{k},\omega) = \mathscr{F}_{\boldsymbol{x}}\{U(\boldsymbol{x},\omega)\} = \int_{-\infty}^{\infty} U(\boldsymbol{x},\omega)e^{-i\boldsymbol{k}^{\mathrm{T}}\boldsymbol{x}}d\boldsymbol{x}, \tag{32.43}$$

$$U(\boldsymbol{x},\omega) = \mathscr{F}_{\boldsymbol{x}}^{-1}\{\tilde{U}(\boldsymbol{k},\omega)\} = \frac{1}{(2\pi)^d}\int_{-\infty}^{\infty} \tilde{U}(\boldsymbol{k},\omega)e^{i\boldsymbol{k}^{\mathrm{T}}\boldsymbol{x}}d\boldsymbol{k}. \tag{32.44}$$

4.32.3.3.3 Fourier transformation with respect to space and time

The Fourier transformations with respect to time and space can be compiled as in Figure 32.8 below. The transformation symbols $\circ\!\!-\!\!\bullet$ denote the transformation pairs $\mathscr{F}_t$ from (32.35, 32.36) and $\mathscr{F}_{\boldsymbol{x}}$ from (32.43, 32.44).

The sequence of the transformations has been shown above in the order

$$u(\boldsymbol{x},t) \quad \overset{\mathscr{F}_t}{\circ\!\!-\!\!\bullet} \quad U(\boldsymbol{x},\omega) \quad \overset{\mathscr{F}_{\boldsymbol{x}}}{\circ\!\!-\!\!\bullet} \quad \tilde{U}(\boldsymbol{k},\omega)$$

(via lower left corner of Figure 32.8), but the order of the transformations can also be reversed (via upper right corner).

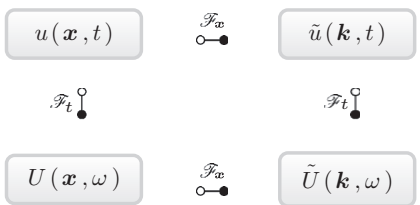


FIGURE 32.8

Fourier transformation in time and space.

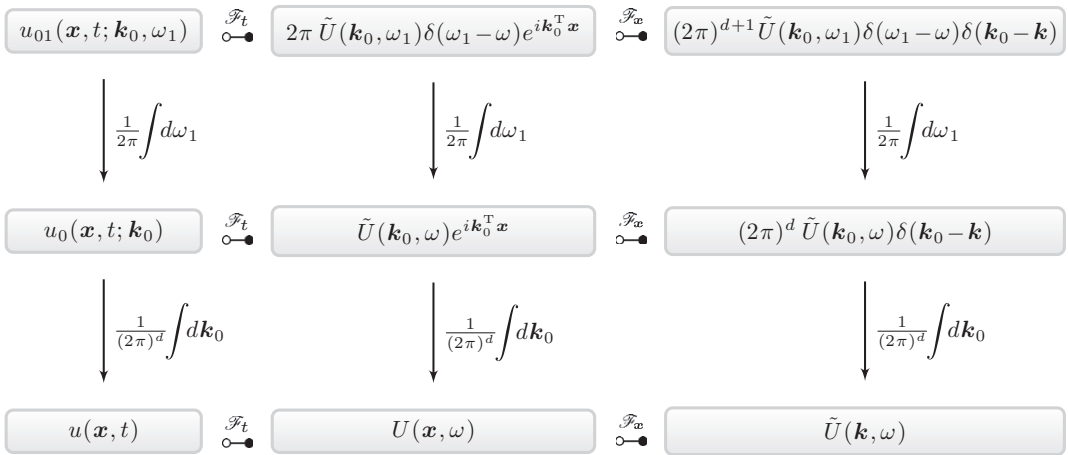


FIGURE 32.9

Representation of the Fourier transformation in time and space as integrals of multidimensional phasors with respect to angular frequency and wave vector.

Note that the Fourier transform $U(\boldsymbol{x},\omega)$ with respect to time is designated by a capital letter and the Fourier transforms $\tilde{u}(\boldsymbol{k},t)$ and $\tilde{U}(\boldsymbol{k},\omega)$ with respect to space by a tilde. These designations are inherited from the complex amplitudes of the respective phasors (see (32.32) and (32.38)).

The transition from the multidimensional phasor $u_{01}(\boldsymbol{x},t;\boldsymbol{k}_0,\omega_1)$ to the general signal $u(\boldsymbol{x},t)$ and its Fourier transformation in time and space $\tilde{U}(\boldsymbol{k},\omega)$ according to (32.37) through (32.42) is shown in Figure 32.9.

The Fourier transformations in time and space can also be used for a concise formulation of the relations (32.33) and (32.41)

$$\mathscr{F}_t\{e^{i\omega_1 t}\} = 2\pi\,\delta(\omega_1-\omega), \tag{32.45}$$

$$\mathscr{F}_{\boldsymbol{x}}\{e^{i\boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{x}}\} = (2\pi)^d\delta(\boldsymbol{k}_0-\boldsymbol{k}). \tag{32.46}$$

They lead to the Fourier transforms in the upper triangle of Figure 32.9.

4.32.3.3.4 *Fourier transformation of plane waves*

The introduction of the Fourier transformations in time and space so far is valid for general functions $u(\boldsymbol{x},t)$ in the sense that they need not represent solutions of the wave equation. Imposing the requirement that $u(\boldsymbol{x},t)$ describes an acoustic wave poses also restrictions on its Fourier transforms. These are derived here for the various cases shown in Figure 32.9. To distinguish general functions $u(\boldsymbol{x},t)$ of time and space from solutions of the wave equation the latter are denoted by $p(\boldsymbol{x},t)$ resembling sound *pressure*.

In (32.29) the complex-valued monofrequent solution of the wave equation has been written in the form

$$p_{01}(\boldsymbol{x},t;\boldsymbol{n}_0,\omega_1) = \tilde{P}(\boldsymbol{n}_0,\omega_1)e^{i\omega_1(t+\frac{1}{c}\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x})} = \tilde{P}(\boldsymbol{n}_0,\omega_1)e^{i\omega_1(t+t_0(\boldsymbol{x}))}. \tag{32.47}$$

The time delay

$$t_0(\boldsymbol{x}) = \frac{1}{c}\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x} \tag{32.48}$$

is the time that a plane wave with the speed c takes to travel from the origin of the spatial coordinate system ($\boldsymbol{x} = \boldsymbol{0}$ and $t_0 = 0$) to the location $\boldsymbol{x}$. This time may also be negative, depending on the position of $\boldsymbol{x}$ relative to the origin and on the direction of the normal vector $\boldsymbol{n}_0$.

The superposition for all possible angular frequencies ω_1 leads to

$$\begin{aligned} p_0(\boldsymbol{x}, t; \boldsymbol{n}_0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} p_{01}(\boldsymbol{x}, t; \boldsymbol{n}_0, \omega_1)\, d\omega_1 &&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{P}(\boldsymbol{n}_0, \omega_1) e^{i\omega_1(t+t_0(\boldsymbol{x}))}\, d\omega_1 \\ &= \mathscr{F}_t^{-1}\left\{\tilde{P}(\boldsymbol{n}_0, \omega) e^{i\omega t_0(\boldsymbol{x})}\right\} &&= \tilde{p}(\boldsymbol{n}_0, t + t_0(\boldsymbol{x})). \end{aligned} \tag{32.49}$$

The function $\tilde{p}(\boldsymbol{n}_0, t)$ describes the waveform of the sound pressure measured at the origin e.g., with a microphone, i.e.,

$$p_{\mathrm{mic}}(t) = p_0(\boldsymbol{0}, t; \boldsymbol{n}_0) = \tilde{p}(\boldsymbol{n}_0, t).$$

The Fourier transform of $p_0(\boldsymbol{x}, t; \boldsymbol{n}_0)$ with respect to time can be read directly from (32.49) as

$$\mathscr{F}_t\{p_0(\boldsymbol{x}, t; \boldsymbol{n}_0)\} = P_0(\boldsymbol{x}, \omega; \boldsymbol{n}_0) = \tilde{P}(\boldsymbol{n}_0, \omega) e^{i\omega t_0(\boldsymbol{x})} \tag{32.50}$$

and Fourier transformation with respect to space results in (see (32.46))

$$\mathscr{F}_{\boldsymbol{x}}\{P_0(\boldsymbol{x}, \omega; \boldsymbol{n}_0)\} = \tilde{P}(\boldsymbol{k}, \omega; \boldsymbol{n}_0) = (2\pi)^d \tilde{P}(\boldsymbol{n}_0, \omega)\delta\left(\frac{\omega}{c}\boldsymbol{n}_0 - \boldsymbol{k}\right). \tag{32.51}$$

The Dirac function in (32.51) restricts the support of the space-time spectrum $\tilde{P}(\boldsymbol{k}, \omega)$ to those values of ω for which the argument of the Dirac function is zero. This statement is again the dispersion relation (32.26), now valid for all space-time frequencies simultaneously.

The relations between the monofrequent plane wave $p_{01}(\boldsymbol{x}, t; \boldsymbol{n}_0, \omega_1)$, its broadband version $p_0(\boldsymbol{x}, t; \boldsymbol{n}_0)$ and their respective Fourier transforms is shown in Figure 32.10.

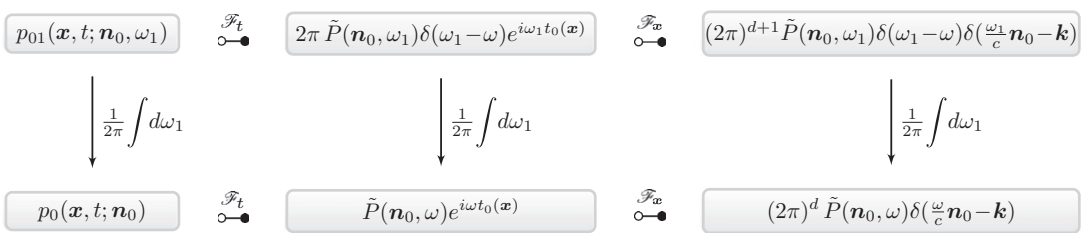


FIGURE 32.10

Representation of a monofrequent plane wave as a phasor and its integration with respect to angular frequency.

Comparing Figures 32.7 and 32.10 shows that the integration with respect to the wave vector in Figure 32.7 (bottom row) has not been performed for the plane wave in Figure 32.10. The reason lies in the dispersion relation which reduces the number of free parameters in the wave vector by one.

In one spatial dimension, the wave vector $\boldsymbol{k}_0$ is equal to the wave number k_0, which can take any real value. The dispersion relation reduces this freedom to only two possible values ± 1 for which the magnitude of the wave number is unity.

In two spatial dimensions, there are two free components in $\boldsymbol{k}_0^{\mathrm{T}} = [k_{0x}\ k_{0y}]$. The dispersion relation reduces the number of free parameters to one since $\boldsymbol{n}_0^{\mathrm{T}} = [\cos\varphi\ \ \sin\varphi]$ depends only on the angle φ which describes the possible angles of incidence. Thus the unit vector $\boldsymbol{n}_0$ describes a circle of radius 1 with the angle φ as parameter.

In three spatial dimensions, the wave vector $\boldsymbol{k}_0$ has three independent components, while the normal vector $\boldsymbol{n}_0$ describes the surface of a unit sphere with two independent parameters, the azimuth and the zenith angle.

The one-dimensional case is discussed below. The two- and three-dimensional cases require the formulation in polar and spherical coordinates, respectively. They are presented in Section 4.32.3.4 and Section 4.32.3.5.

4.32.3.3.5 Plane waves in one spatial coordinate

In one space dimension there are only two possible values for the unit vector $n_0 = \pm 1$. The time delay t_0 from (32.48) has the values

$$t_0(x) = \pm\frac{x}{c} \tag{32.52}$$

and (32.50) has the form

$$P_0(x, \omega; \pm 1) = \tilde{P}(\pm 1, \omega) e^{\pm i\omega x/c}. \tag{32.53}$$

The superposition for all possible values (compare (32.40)) is restricted here to two terms

$$P(x, \omega) = \tilde{P}(+1, \omega) e^{i\omega x/c} + \tilde{P}(-1, \omega) e^{-i\omega x/c}. \tag{32.54}$$

For a shorter notation in the time domain, introduce

$$\tilde{p}_+(t) = \mathscr{F}_t^{-1}\{\tilde{P}(+1, \omega)\}, \quad \tilde{p}_-(t) = \mathscr{F}_t^{-1}\{\tilde{P}(-1, \omega)\}. \tag{32.55}$$

Then the superposition of both components in the space-time domain has the form

$$p(x, t) = \tilde{p}_+\left(t + \frac{x}{c}\right) + \tilde{p}_-\left(t - \frac{x}{c}\right). \tag{32.56}$$

This representation is known as the *D'Alembert-solution* of the wave equation or the *traveling wave solution*. Inspection of (32.56) shows that $\tilde{p}_-\left(t - \frac{x}{c}\right)$ is a wave traveling in the direction of the positive x-axis, since $\tilde{p}_-\left(t - \frac{x}{c}\right) = \text{const}$ for all values of t which increase proportional to x. By the same argument, $\tilde{p}_+\left(t + \frac{x}{c}\right)$ is a wave traveling in the direction of the negative x-axis.

The corresponding relation in the wavenumber-frequency domain follows from (32.53) by Fourier transformation with respect to space as

$$\tilde{P}(k, \omega) = 2\pi\, \tilde{P}(\pm 1, \omega)\delta\left(\pm\frac{\omega}{c} - k\right). \tag{32.57}$$

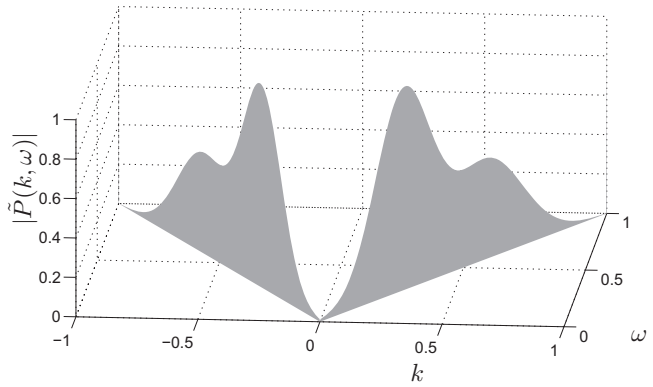


FIGURE 32.11

Frequency-domain representation of $|\tilde{P}(k,\omega)|$ in the k-ω-plane for one space dimension. The k- and ω axes are labeled in multiples of π. The waves traveling in both directions are assumed to be the same, i.e., $\tilde{P}(+1,\omega) = \tilde{P}(-1,\omega)$.

Figure 32.11 illustrates the behavior of $\tilde{P}(k,\omega)$ for one dimension in space. The spectrum $\tilde{P}(k,\omega)$ takes the values $P(\pm 1,\omega)$ only for $\omega = \pm kc$ (s. (32.26)) and is zero elsewhere.

4.32.3.4 Circular harmonics

The plane wave solution of the wave equation has been introduced in Section 4.32.3.2 and further discussed in Section 4.32.3.3 because it has a simple representation in Cartesian coordinates, see e.g., (32.29). General sound fields are however much more complicated than a monofrequent wave from a fixed direction. Nevertheless the plane wave solution is a valuable building block for the description of general sound fields. In particular, a *general sound field* can be described as a superposition of plane waves from a continuum of directions [22]. This superposition is shown here for plane waves from all directions in the horizontal plane. Section 4.32.3.5 extends this idea to plane waves from all directions in the three-dimensional space.

In analogy to Figure 32.9 it is important to distinguish

monofrequent plane waves: waves with a plane wave front indicated by its normal vector $\boldsymbol{n}_0$ and with a monofrequent time characteristic with the fixed angular frequency ω_1 (see also the top row of Figure 32.10),
broadband plane waves: plane waves with a time characteristic that is composed of (possibly infinitely many) monofrequent components (see also the bottom row of Figure 32.10),
general sound fields: sound fields that are composed of broadband plane waves from (possibly infinitely many) different directions.

The role of the plane wave for the description of general sound fields is somewhat similar to the role of sinusoids or complex exponentials for the description of general sound spectra. Although the

spectra of speech and music signals may be very complex, the Fourier transformation allows to write any time signal as a superposition of sinusoids. In the same way, a sound field can be expressed as a superposition of plane waves.

4.32.3.4.1 Monofrequent plane wave in polar coordinates

The superposition of plane waves from all directions in a horizontal plane is best accomplished in polar coordinates introduced in Section 4.32.2.2. First consider the phasor from (32.29) where the notation has been adapted to the description with polar coordinates

$$p_{01}(\boldsymbol{x}, t; \varphi_0, \omega_1) = \tilde{P}(\varphi_0, \omega_1) e^{i\omega_1 (t + t_0(\boldsymbol{x}))}. \tag{32.58}$$

The argument of the complex amplitude $\tilde{P}(\varphi_0, \omega_1)$ indicates that its value depends not only on the angular frequency ω_1 but also on the direction φ_0 of the plane wave. The vector of space variables $\boldsymbol{x}$ and the normal vector $\boldsymbol{n}_0$ are given by

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = r \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \quad \text{and} \quad \boldsymbol{n}_0 = \begin{bmatrix} \cos\varphi_0 \\ \sin\varphi_0 \end{bmatrix}. \tag{32.59}$$

Note that the vector of space variables $\boldsymbol{x}$ may vary with the distance r and angle α, while the wave vector $\boldsymbol{n}_0$ of a plane wave is fixed for a given direction φ_0.

The scalar product $\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x}$ of $\boldsymbol{x}$ and $\boldsymbol{n}_0$ can be expressed like any scalar product by the lengths of these vectors and the angle γ which they include

$$\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x} = |\boldsymbol{n}_0||\boldsymbol{x}| \cos\gamma = r\cos\gamma \quad \text{with} \quad \gamma = \alpha - \varphi_0. \tag{32.60}$$

This general relation follows here from (32.59) with the trigonometric relation for the difference of two angles

$$\boldsymbol{n}_0^{\mathrm{T}}\boldsymbol{x} = r(\cos\alpha\cos\varphi_0 + \sin\alpha\sin\varphi_0) = r\cos(\alpha - \varphi_0). \tag{32.61}$$

The plane wave from (32.58) can now be expressed as

$$p_{01}(r, \alpha, t; \varphi_0, \omega_1) = \tilde{P}(\varphi_0, \omega_1) e^{i\omega_1 t} e^{i\frac{\omega}{c} r\cos(\alpha - \varphi_0)}, \tag{32.62}$$

where the dependence on the polar coordinates is now written explicitly by including radius r and angle α in the list of parameters. To simplify the notation, no separate designation is introduced; i.e., the somewhat loosely formulated identity $p_{01}(r, \alpha, t; \varphi_0, \omega_1) = p_{01}(\boldsymbol{x}, t; \varphi_0, \omega_1)$ is assumed.

The plane wave representation in polar coordinates is obviously a periodic function with respect to the angle α due to the periodicity of the cosine

$$p_{01}(r, \alpha, t; \varphi_0, \omega_1) = p_{01}(r, \alpha + 2\pi, t; \varphi_0, \omega_1). \tag{32.63}$$

Its expansion into orthogonal functions is given by a Fourier series as shown next.

4.32.3.4.2 Expansion of monofrequent plane waves into orthogonal functions

The Fourier coefficients of the monofrequent plane wave in (32.62) follow from an expansion of the complex exponential term. It can be obtained from the following definition of the Bessel functions of the first kind [23][1]

$$i^n J_n(kr) = \frac{1}{2\pi}\int_0^{2\pi} e^{i(kr\cos\gamma - n\gamma)} d\gamma, \quad n \in \mathbb{Z}, \tag{32.64}$$

which represents the Fourier series coefficients of the function $\exp(ikr\cos\gamma)$. Its Fourier series is then given by

$$e^{ikr\cos(\alpha-\varphi_0)} = \sum_{n=-\infty}^{\infty} i^n e^{in(\alpha-\varphi_0)} J_n(kr). \tag{32.65}$$

In the literature on acoustics, this relation is called the *Jacobi-Anger expansion* [22,23].

The Fourier series expansion of (32.62) follows now immediately as

$$p_{01}(r,\alpha,t;\varphi_0,\omega_1) = \tilde{P}(\varphi_0,\omega_1) \sum_{n=-\infty}^{\infty} i^n e^{i\omega_1 t} e^{in\alpha} J_n\left(\frac{\omega_1}{c} r\right) e^{-in\varphi_0}. \tag{32.66}$$

The importance of this representation lies in the fact that its components

$$i^n e^{i\omega_1 t} \cdot e^{in\alpha} \cdot J_n\left(\frac{\omega_1}{c} r\right)$$

are separated in terms that depend on time t, angle α, and radius r. This is not the case in (32.62) where both r and α appear in the exponent of the complex exponential function.

The monofrequent plane wave has been treated in great detail here since it is the building block for the representation of broadband plane waves and general sound fields.

4.32.3.4.3 Expansion of broadband plane waves into orthogonal functions

A *broadband plane wave* $p_0(r,\alpha,t;\varphi_0)$ is composed of monofrequent plane waves with different angular frequencies ω_1 but identical direction φ_0. This composition is formulated in mathematical terms by an integration of the monofrequent plane wave (32.62) with respect to ω_1. Through the term $\exp(i\omega_1 t)$, the integration takes the form of an inverse Fourier transformation

$$\begin{aligned} p_0(r,\alpha,t;\varphi_0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} p_{01}(r,\alpha,t;\omega_1,\varphi_0) d\omega_1 \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{P}(\varphi_0,\omega_1) e^{i\frac{\omega_1}{c} r\cos(\alpha-\varphi_0)} e^{i\omega_1 t} d\omega_1 \\ &= \mathscr{F}_t^{-1}\left\{\tilde{P}(\varphi_0,\omega_1) e^{i\frac{\omega_1}{c} r\cos(\alpha-\varphi_0)}\right\}. \end{aligned} \tag{32.67}$$

This relation is expressed shorter by the Fourier transform of $p_0(r,\alpha,t;\varphi_0)$ as

$$P_0(r,\alpha,\omega;\varphi_0) = \mathscr{F}_t\{p_0(r,\alpha,t;\varphi_0)\} = \tilde{P}(\varphi_0,\omega) e^{i\frac{\omega}{c} r\cos(\alpha-\varphi_0)}. \tag{32.68}$$

[1] The mathematical literature presents various versions of this integral relation which differ in the sign of the summation index n or in the choice of the trigonometric function sin or cos. The equivalence of these versions can be shown by substitution.

Then, with (32.65) the Fourier transform $P_0(r, \alpha, \omega; \varphi_0)$ of the plane wave can be represented by a series expansion

$$P_0(r, \alpha, \omega; \varphi_0) = \sum_{n=-\infty}^{\infty} \tilde{P}(\varphi_0, \omega) i^n e^{in(\alpha-\varphi_0)} J_n\left(\frac{\omega}{c} r\right). \tag{32.69}$$

This series represents a plane wave from the direction φ_0 with the arbitrary spectrum $\tilde{P}(\varphi_0, \omega)$.

4.32.3.4.4 Expansion of general sound fields into orthogonal functions

Finally, a *general sound field* $p(r, \alpha, t)$ can be obtained by composing broadband plane waves from all possible directions $0 \le \varphi_0 < 2\pi$. The composition is formulated as an integration of the Fourier transform of the broadband plane wave (32.69) with respect to φ_0. By extracting the exponential term $e^{-in\varphi_0}$ and applying (32.69) to the remaining integral, it can be written as the calculation of a Fourier coefficient

$$P(r, \alpha, \omega) = \frac{1}{2\pi} \int_0^{2\pi} P_0(r, \alpha, \omega; \varphi_0) d\varphi_0 = \sum_{n=-\infty}^{\infty} \breve{P}_n(\omega) i^n e^{in\alpha} J_n\left(\frac{\omega}{c} r\right), \tag{32.70}$$

where the complex amplitude $\tilde{P}(\varphi_0, \omega)$ is represented by its Fourier coefficients $\breve{P}_n(\omega)$

$$\breve{P}_n(\omega) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{P}(\varphi_0, \omega) e^{-in\varphi_0} d\varphi_0, \tag{32.71}$$

as

$$\tilde{P}(\varphi_0, \omega) = \sum_{\ell=-\infty}^{\infty} \breve{P}_\ell(\omega) e^{i\ell\varphi_0}. \tag{32.72}$$

Finally the Fourier transform $P(r, \alpha, \omega)$ of a general sound field can be expressed as a Fourier series

$$P(r, \alpha, \omega) = \sum_{n=-\infty}^{\infty} \mathring{P}_n(r, \omega) e^{in\alpha} \tag{32.73}$$

with the Fourier series coefficients

$$\mathring{P}_n(r, \omega) = \breve{P}_n(\omega) i^n J_n\left(\frac{\omega}{c} r\right). \tag{32.74}$$

These Fourier series coefficients characterize the angular structure of the sound field, respectively of its Fourier transform. Higher orders n correspond to a finer angular structure. Note that the radius dependent part $J_n\left(\omega r/c\right)$ is independent of the specific wave field.

4.32.3.4.5 Summary of the representations of a general sound field

The transition from a monofrequent plane wave in polar coordinates to a broadband plane wave and to a general sound field is shown in Figure 32.12. It corresponds to the center column of Figure 32.10 with the spatial dependency expressed in polar coordinates as derived in this section.

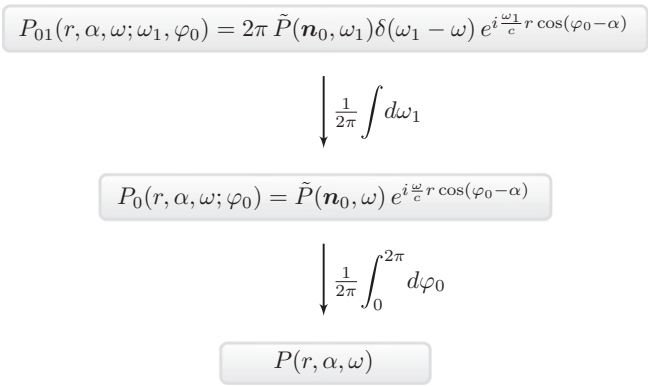


FIGURE 32.12

Representation of the Fourier transform of a monofrequent plane wave in polar coordinates and its integration with respect to angular frequency ω_1 and direction of arrival φ_0.

The relations for the Fourier transform of the general sound field $P(r, \alpha, \omega)$ from (32.69) to (32.74) are compiled in (32.75) for easier reference

$$\begin{aligned} P(r, \alpha, \omega) &= \sum_{n=-\infty}^{\infty} \mathring{P}_n(r, \omega) e^{in\alpha} \\ &= \sum_{n=-\infty}^{\infty} \breve{P}_n(\omega) i^n e^{in\alpha} J_n\left(\frac{\omega}{c} r\right), \\ \tilde{P}(\varphi_0, \omega) &= \sum_{\ell=-\infty}^{\infty} \breve{P}_\ell(\omega) e^{i\ell\varphi_0}. \end{aligned} \tag{32.75}$$

4.32.3.4.6 Orthogonality properties of the complex exponential functions

The relations in this section are based on Fourier series expansions and have exploited the orthogonality properties of the complex exponential functions. Similar properties of more complicated functions are required for three-dimensional sound fields in spherical coordinates in Section 4.32.3.5. To highlight the similarities between these two different systems of orthogonal functions, orthogonality properties of the complex exponential functions are compiled here. They serve as a review on the material on circular harmonics and as well as for later reference from Section 4.32.3.5.

For the sake of a uniform notation, the designation $V_n(\varphi)$ for the complex exponential functions is introduced as

$$V_n(\varphi) = e^{in\varphi} \tag{32.76}$$

with the index n, $n \in \mathbb{Z}$ and the angle φ.

Orthogonality relation with respect to the angle: The integration with respect to the angle φ over one period of length 2π is equivalent to the integration around a circle. The associated orthogonality property is expressed with the Kronecker symbol $\delta_{\mu n}$ as

$$\frac{1}{2\pi}\int_0^{2\pi} V_\mu^*(\varphi)V_n(\varphi)d\varphi = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-\mu)\varphi}d\varphi = \delta_{\mu n}. \tag{32.77}$$

It is easily proven by evaluating the integral separately for the cases $n = \mu$ and $n \neq \mu$.

Closure equation with respect to the index: The functions $V_n(\varphi)$ form a complete set of orthogonal functions and therefore satisfy a closure relation (see [23]). It can be derived by summation with respect to the index n. Applying the orthogonality property (32.77) gives

$$\frac{1}{2\pi}\sum_{n=-\infty}^{\infty} V_n^*(\vartheta)V_n(\varphi) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{in(\vartheta-\varphi)} = \delta(\vartheta-\varphi) \tag{32.78}$$

with the delta function $\delta(\varphi)$.

Compilation of Fourier series expansions: Table 32.2 compiles the Fourier series expansions used in Section 4.32.3.4. The first one is the closure relation from (32.78). The factor $1/(2\pi)$ can be interpreted as a series of Fourier coefficients that are constant with respect to n; the associated function is the delta impulse, see (32.79). When the Fourier series coefficients are expressed by Bessel functions as $i^n J_n(kr)$, then the Fourier series expansion of a plane wave results in (32.80). Multiplying these Bessel coefficients by an arbitrary sequence of coefficients $\check{P}_n(\omega)$ gives the Fourier series expansion of a general sound field $P(r, \varphi-\vartheta, \omega)$ in (32.81).

4.32.3.5 Spherical harmonics

Building a three-dimensional sound field from plane wave components follows the same pattern as described above for components from the horizontal plane. The main difference is obviously that now

Table 32.2 Compilation of Fourier Series Expansions Used in Section 4.32.3.4

$$\frac{1}{2\pi}\sum_{n=-\infty}^{\infty} V_n^*(\vartheta)V_n(\varphi) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{in(\vartheta-\varphi)} = \delta(\vartheta-\varphi) \tag{32.79}$$

$$\sum_{n=-\infty}^{\infty}\left(i^n J_n\left(\frac{\omega}{c}r\right)\right)V_n^*(\vartheta)V_n(\varphi) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\left(i^n J_n\left(\frac{\omega}{c}r\right)\right)e^{in(\vartheta-\varphi)}$$
$$= \exp\left(i\frac{\omega}{c}r\cos(\vartheta-\varphi)\right) \tag{32.80}$$

$$\sum_{n=-\infty}^{\infty}\check{P}_n(\omega)i^n J_n\left(\frac{\omega}{c}r\right)V_n^*(\vartheta)V_n(\varphi) = P(r, \varphi-\vartheta, \omega) \tag{32.81}$$

plane waves not only from all around the horizontal plane but also from above and below are involved. This means in mathematical terms that the integration around a circle that frequently arises in Section 4.32.3.4 is here replaced by an integration on a sphere. The presentation for three-dimensional sound fields therefore repeats the same steps as for planar sound fields in Section 4.32.3.4.

4.32.3.5.1 Monofrequent plane waves in spherical coordinates

The superposition of plane waves from all directions in the three-dimensional space requires spherical coordinates as introduced in Section 4.32.2.2. At first, the definition of the phasor from (32.29) is updated such that the argument of the complex amplitude $\tilde{P}(\varphi_0, \theta_0, \omega_1)$ includes also the zenith angle θ_0

$$p_{01}(\boldsymbol{x}, t; \varphi_0, \theta_0, \omega_1) = \tilde{P}(\varphi_0, \theta_0, \omega_1) e^{i\omega_1 (t + t_0(\boldsymbol{x}))}. \tag{32.82}$$

The time delay $t_0(\boldsymbol{x})$ is given by (32.48), but the vector of space variables $\boldsymbol{x}$ and the normal vector $\boldsymbol{n}_0$ are defined in three spatial spherical coordinates according to (32.6) by

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} \cos\alpha \sin\beta \\ \sin\alpha \sin\beta \\ \cos\beta \end{bmatrix}, \quad \boldsymbol{n}_0 = \begin{bmatrix} \cos\varphi_0 \sin\theta_0 \\ \sin\varphi_0 \sin\theta_0 \\ \cos\theta_0 \end{bmatrix}. \tag{32.83}$$

The scalar product $\boldsymbol{n}_0^{\mathrm{T}} \boldsymbol{x}$ follows now from (32.83) by manipulation with standard trigonometric relations

$$\begin{aligned} \boldsymbol{n}_0^{\mathrm{T}} \boldsymbol{x} &= r(\cos\varphi_0 \sin\theta_0 \cos\alpha \sin\beta + \sin\varphi_0 \sin\theta_0 \sin\alpha \sin\beta + \cos\theta_0 \cos\beta) \\ &= r(\sin\theta_0 \sin\beta(\cos\varphi_0 \cos\alpha + \sin\varphi_0 \sin\alpha) + \cos\theta_0 \cos\beta) \\ &= r(\sin\theta_0 \sin\beta \cos(\varphi_0 - \alpha) + \cos\theta_0 \cos\beta) = r\cos\gamma. \end{aligned} \tag{32.84}$$

The angle γ is again the angle included by the vectors $\boldsymbol{x}$ and $\boldsymbol{n}_0$, now in three-dimensional space, see Figure 32.13.

The plane wave from (32.82) can be expressed similar to (32.62) as

$$p_{01}(\boldsymbol{x}, t; \varphi_0, \theta_0, \omega_1) = \tilde{P}(\varphi_0, \theta_0, \omega_1) e^{i\omega_1 t} e^{i\frac{\omega_1}{c} r\cos\gamma}, \tag{32.85}$$

with the angle γ defined from (32.84). The dependence on the spherical coordinates is written explicitly by including radius r and the angles α and β in the list of parameters as

$$p_{01}(r, \alpha, \beta, t; \varphi_0, \theta_0, \omega_1) = \tilde{P}(\varphi_0, \theta_0, \omega_1) e^{i\omega_1 t} e^{i\frac{\omega_1}{c} r\cos\gamma}. \tag{32.86}$$

This relation looks similar to the corresponding one in (32.62), but the main difference is that the angles are now defined with respect to a sphere, see Figure 32.13. As a consequence, the expansion into basis functions has to take the resulting spherical symmetry into account.

4.32.3.5.2 Definition of the spherical harmonic functions

The spherical harmonic functions or short spherical harmonics are an orthogonal basis for functions defined on a sphere. Their definition and the notation varies slightly in the literature on mathematics,

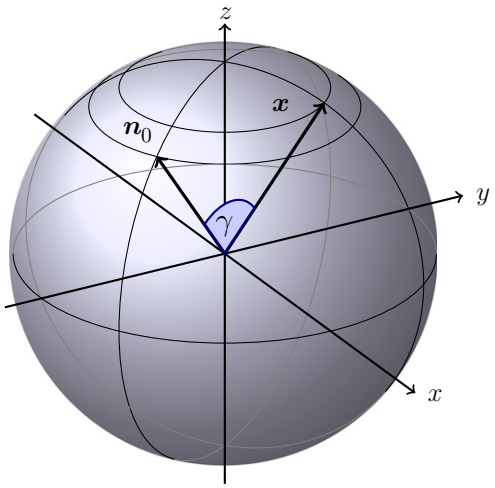


FIGURE 32.13

Spherical coordinate system with direction $\boldsymbol{n}_0$ of a plane wave.

acoustics, quantum physics, and other fields of science. The definition and the description of their properties used here follow [24].

The spherical harmonic functions $Y_n^m(\theta, \varphi)$ may be defined as

$$Y_n^m(\theta, \varphi) = C(m, n) P_n^{|m|}(\cos\theta) e^{im\varphi}, \tag{32.87}$$

The set of functions $P_n^m(\cos\theta)$ are the *associated Legendre functions* of nth degree and mth order. They are defined in terms of the *Legendre polynomials* $P_n(x)$ by

$$P_n^m(x) = (-1)^m (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x), \quad \forall n, m \geq 0, \tag{32.88}$$

$$P_n^{-m}(x) = \frac{(n-m)!}{(n+m)!} P_n^m(x). \tag{32.89}$$

The Legendre polynomials themselves are determined by their generating function

$$(1 - 2xy + y^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} P_n(x) y^n \quad |y| < 1. \tag{32.90}$$

Thus the index n indicates the degree of the Legendre polynomial $P_n(x)$ while the index m indicates the order of the differentiation in (32.88) for the definition of the associated Legendre function $P_n^m(x)$.

The factor $C(m, n)$ is chosen such that the spherical harmonic functions form an orthonormal basis with respect to integration on a sphere (see (32.92)), e.g., as

$$C(m, n) = (-1)^m \sqrt{\frac{2n + 1}{4\pi} \frac{(n - |m|)!}{(n + |m|)!}}. \tag{32.91}$$

Other definitions of the spherical harmonic functions are also in use, which differ from (32.87) and (32.91) mainly with respect to the factor $(-1)^m$. The choice of definition is a matter of taste and convention. The present definition has found to be the most flexible one.

A note on the notation is required here. The designations $P_n(x)$ for the Legendre polynomials and $P_n^m(x)$ for the associated Legendre functions have been chosen here because they are commonly used in the mathematical literature. They must not be confused with the Fourier coefficients $\mathring{P}_n(r,\omega)$ or $\breve{P}_n(\omega)$ e.g., in (32.74). This ambiguity in the notation is permissible here since the further presentation uses almost exclusively the spherical harmonic functions $Y_n^m(\theta,\varphi)$.

4.32.3.5.3 Orthogonality relations for spherical harmonic functions

The properties of the Legendre polynomials are covered in the mathematical literature on orthogonal polynomials. The extension to the associated Legendre functions and to the spherical harmonic functions are found e.g., in [22,23]. Therefore the derivation of the spherical harmonic functions is not repeated here. Some clues on their role as solution of the acoustic wave equation are given in Section 4.32.3.5.

In the context of sound field reproduction, mainly the orthogonality properties of the spherical harmonics are of importance. They are compiled below following the presentation in [23] and in parallel to the orthogonality properties of the complex exponentials in Section 4.32.3.4.

Orthogonality with respect to the azimuth and zenith angle: An integration on a sphere is accomplished by an integration on a circle in the horizontal plane (azimuth angle φ, compare (32.77)) and an integration of the zenith angle θ. It establishes the following orthogonality relation

$$\int_0^{2\pi}\int_0^{\pi} Y_{n_1}^{m_1*}(\theta,\varphi)\,Y_{n_2}^{m_2}(\theta,\varphi)\sin\theta\,d\theta\,d\varphi = \delta_{n_1n_2}\delta_{m_1m_2}. \tag{32.92}$$

Closure equation with respect to the indices n and m: A double summation of the spherical harmonics with respect to the degree n and mth order leads to the closure equation

$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n} Y_n^{m*}(\theta,\varphi)\,Y_n^m(\beta,\alpha) = \delta(\varphi-\alpha)\delta(\cos\theta-\cos\beta) = \delta(\varphi-\alpha)\delta(\theta-\beta)\frac{1}{\sin\theta}. \tag{32.93}$$

Orthogonal expansion of a function defined on a sphere: A function $f(\theta,\varphi)$ of the azimuth angle φ and the zenith angle θ is called a function which is defined on a sphere. The orthogonality relation (32.92) allows to expand such functions into spherical harmonics as

$$f(\theta,\varphi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} f_{mn}\,Y_n^m(\theta,\varphi). \tag{32.94}$$

The expansion coefficients f_{mn} are derived by application of (32.92).

Expansion of a plane wave into spherical waves: An example for the expansion of a function defined on a sphere is the expansion of a plane wave into spherical harmonics

$$4\pi\sum_{n=0}^{\infty}\sum_{m=-n}^{n} i^n\,j_n(kr)\,Y_n^{m*}(\theta,\varphi)\,Y_n^m(\beta,\alpha) = e^{ikr\cos\gamma}, \tag{32.95}$$

where $\gamma = \gamma(\alpha, \beta, \varphi, \theta)$ is related to $\alpha, \beta, \varphi, \theta$ by (compare (32.84))

$$\cos\gamma = \sin\theta \sin\beta \cos(\varphi - \alpha) + \cos\theta \cos\beta. \tag{32.96}$$

The expansion coefficients include the spherical Bessel function $j_n(kr)$ which are described in mathematical texts on special functions e.g., [23].

4.32.3.5.4 Expansion of broadband plane waves into orthogonal functions

The orthogonality relations for spherical harmonic functions allow to expand the monofrequent plane wave from (32.82) and its generalizations into orthogonal functions. The presentation follows closely the circular case from Section 4.32.3.4.

At first, the monofrequent plane wave (32.82) is integrated with respect to its angular frequency ω_1 to obtain a broadband plane wave (see (32.67))

$$p_0(r, \alpha, \beta, t; \varphi_0, \theta_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} p_{01}(r, \alpha, \beta, t; \omega, \varphi_0, \theta_0) d\omega_1. \tag{32.97}$$

Similar to (32.68) follows its Fourier transform with respect to time

$$P_0(r, \alpha, \beta, \omega; \varphi_0, \theta_0) = \mathscr{F}_t\{p_0(r, \alpha, \beta, t; \varphi_0, \theta_0)\} = \tilde{P}(\varphi_0, \theta_0, \omega) e^{i\frac{\omega}{c} r \cos\gamma}, \tag{32.98}$$

where the angle γ is defined on a sphere as shown in Figure 32.13. The expansion into spherical harmonics is given by (32.95) as

$$P_0(r, \alpha, \beta, \omega; \varphi_0, \theta_0) = 4\pi \tilde{P}(\varphi_0, \theta_0, \omega) \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n j_n\left(\frac{\omega}{c} r\right) Y_n^{m*}(\theta_0, \varphi_0) Y_n^m(\beta, \alpha). \tag{32.99}$$

4.32.3.5.5 Expansion of general sound fields into orthogonal functions

The corresponding expansion for general sound fields in a source-free domain follows by superposition of plane wave contributions (32.99) from all directions φ_0 and θ_0

$$P(r, \alpha, \beta, \omega) = \int_0^{2\pi} \int_0^{\pi} P_0(r, \alpha, \beta, \omega; \varphi_0, \theta_0) \sin\theta \, d\theta_0 \, d\varphi_0. \tag{32.100}$$

Applying the spherical integration with respect to φ_0 and θ_0 to the series expansion (32.99) gives

$$P(r, \alpha, \beta, \omega) = 4\pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n j_n\left(\frac{\omega}{c} r\right) \breve{P}_n^m(\omega) Y_n^m(\beta, \alpha) \tag{32.101}$$

with the coefficients

$$\breve{P}_n^m(\omega) = \int_0^{2\pi} \int_0^{\pi} \tilde{P}(\varphi_0, \theta_0, \omega) Y_n^{m*}(\theta, \varphi) \sin\theta \, d\theta \, d\varphi \tag{32.102}$$

of the expansion

$$\tilde{P}(\varphi_0, \theta_0, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \check{P}_n^m(\omega) Y_n^m(\theta, \varphi). \tag{32.103}$$

Similar to (32.73) and (32.74), a general sound field in the three-dimensional space is given by the expansion

$$P(r, \alpha, \beta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \mathring{P}_n^m(r, \omega) Y_n^m(\beta, \alpha) \tag{32.104}$$

with the expansion coefficients

$$\mathring{P}_n^m(r, \omega) = 4\pi i^n j_n\left(\frac{\omega}{c} r\right) \check{P}_n^m(\omega). \tag{32.105}$$

Note that these expansion coefficients are again separable with respect to the radius r and the mode numbers n, m.

4.32.3.5.6 Relation between spherical harmonics expansions and Fourier series

Since the expansion (32.101) converge uniquely and uniformly above a certain threshold, the order of summation may be exchanged [24]. If the spherical harmonics $Y_n^m(\beta, \alpha)$ are then expressed by their explicit formulation (32.87), the Fourier series that is inherent to (32.101) is revealed. It is given by

$$P(\boldsymbol{x}, \omega) = \sum_{m=-\infty}^{\infty} \mathring{P}_m(r, \beta, \omega) e^{im\alpha} \tag{32.106}$$

with the Fourier coefficients

$$\mathring{P}_m(r, \beta, \omega) = \sum_{n=|m|}^{\infty} 4\pi i^n \check{P}_n^m(\omega) j_n\left(\frac{\omega}{c} r\right) C(m, n) P_n^{|m|}(\cos\beta). \tag{32.107}$$

The normalization coefficients $C(m, n)$ are defined by (32.91). Note that the Fourier coefficients $\mathring{P}_m(r, \beta, \omega)$ of the spherical harmonics in (32.107) correspond to the Fourier coefficients $\mathring{P}_n(r, \omega)$ for the circular case in (32.73). On the other hand $P_n^{|m|}(\cos\beta)$ are the associated Legendre Functions from (32.87).

4.32.3.5.7 Summary of the representations of a general sound field

The transition from a monofrequent plane wave in spherical coordinates to a broadband plane wave and to a general sound field is shown in Figure 32.14 in a similar way as for polar coordinates in Figure 32.12.

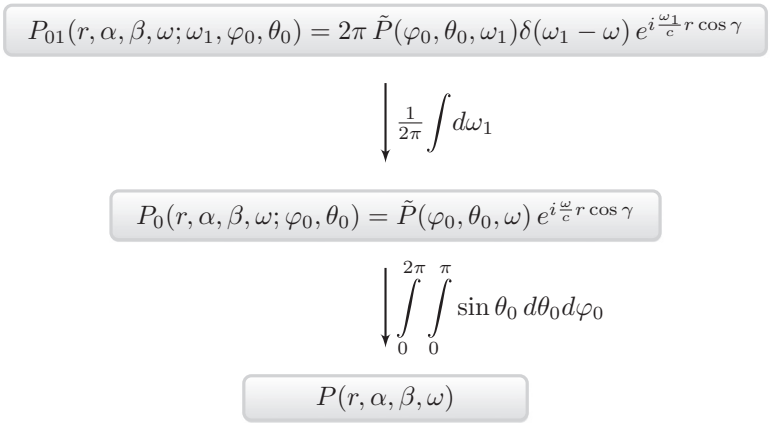


FIGURE 32.14

Representation of the Fourier transform of a monofrequent plane wave in spherical coordinates and its integration with respect to angular frequency ω_1 and direction of arrival φ_0, θ_0.

The relations for the Fourier transform of the general sound field $P(r, \alpha, \omega)$ from (32.100) to (32.107) are compiled in (32.108)

$$\begin{aligned} P(r,\alpha,\beta,\omega) &= \sum_{m=-\infty}^{\infty} \mathring{P}_m(r,\beta,\omega)e^{im\alpha} \\ &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \mathring{P}_n^m(r,\omega)Y_n^m(\beta,\alpha) \\ &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \breve{P}_n^m(\omega)4\pi i^n Y_n^m(\beta,\alpha)\,j_n\left(\frac{\omega}{c}r\right), \\ \mathring{P}_n^m(r,\omega) &= 4\pi i^n\,j_n\left(\frac{\omega}{c}r\right)\breve{P}_n^m(\omega). \end{aligned} \tag{32.108}$$

4.32.3.5.8 Derivation of the spherical harmonic functions

The derivation of the spherical harmonic functions $Y_n^m(\theta, \varphi)$ is found in many textbooks on acoustics [16,22,24] or on mathematical methods in physics [23,25]. The essential steps are

1. Express the Laplace operator in the acoustic wave equation in spherical coordinates.
2. Solve the wave equation by the method of separation of variables to obtain separate ordinary differential equations with respect to time, radius, azimuth and zenith angle.
3. Obtain the solution of the differential equation with respect to time in the form of complex exponentials for time and angular velocity (t, ω).
4. Obtain the solution of the differential equation with respect to the radius in the form of spherical Bessel functions with respect to radius and wave number (r, k).

5. Obtain the solution of the differential equation with respect to the azimuth angle in the form of complex exponentials for azimuth angle and mode number (φ, n).
6. Obtain the solution of the differential equation with respect to the zenith angle in the form of associated Legendre functions for zenith angle and mode numbers (θ, n, m).
7. Form the spherical harmonic functions as the product of the complex exponentials for azimuth angle and mode number and the associated Legendre functions.

The complete process is somewhat tedious and involves various special functions from higher mathematics. Since it is well covered in the standard literature as referenced above, no derivation is given here.

4.32.4 Response to sound sources

The response of a sound field to an acoustic point source is described by the so-called *Green's function*. This section derives some of its basic properties from the acoustic wave equation. Then the Green's function is used to obtain the sound field within a certain spatial region when interior sources or boundary values are given. The latter case serves to introduce the Kirchhoff-Helmholtz integral equation for later use in Section 4.32.5. Finally the Green's function for wave propagation in the free-field is derived. The representation follows classic texts like [15–17,25] and has been adapted from [26].

4.32.4.1 The inhomogenous wave equation

From the homogeneous acoustic wave Eq. (32.3) follows the inhomogeneous wave equation by considering a source term $p_0(\boldsymbol{x}, t)$

$$\Delta p(\boldsymbol{x}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} p(\boldsymbol{x}, t) = p_0(\boldsymbol{x}, t). \tag{32.109}$$

Fourier transformation $\mathscr{F}_t$ according to (32.35) leads to the frequency domain version of the acoustic wave Eq. (32.109), the so-called *Helmholtz equation*

$$\Delta P(\boldsymbol{x}, t) + \left(\frac{\omega}{c}\right)^2 P(\boldsymbol{x}, \omega) = P_0(\boldsymbol{x}, \omega) \quad \boldsymbol{x} \in V. \tag{32.110}$$

It is valid within a spatial region V which may be bounded by the walls of an enclosure. However, V may also be an arbitrary volume in the free field and not directly related to the walls of an acoustic environment.

For later reference the operations on the left hand side of (32.110) can be abbreviated by the linear operator

$$L\{P(\boldsymbol{x}, \omega)\} = \Delta P(\boldsymbol{x}, t) + \left(\frac{\omega}{c}\right)^2 P(\boldsymbol{x}, \omega). \tag{32.111}$$

This notation allows a very concise notation for the Helmholtz equation as

$$L\{P(\boldsymbol{x}, \omega)\} = P_0(\boldsymbol{x}, \omega) \quad \boldsymbol{x} \in V. \tag{32.112}$$

4.32.4.2 Green's function

The effect of a source distribution $P_0(\boldsymbol{x}, \omega)$ on the complete sound field in V is described by the Green's function. This section presents its properties and emphasizes its role for the calculation of sound fields.

4.32.4.2.1 Properties

The Green's function describes the effect of a point source at the location $\boldsymbol{\xi}$ on the sound field $P(\boldsymbol{x}, \omega)$ at the location $\boldsymbol{x}$. This effect depends also on the spectrum of the source signal, therefore the Green's function $G(\boldsymbol{x}|\boldsymbol{\xi}, \omega)$ has two space variables and one frequency variable. To highlight the close relationship of the two space variables, they are separated in formulas by a | rather than by a comma.

The Green's function can be regarded as an equivalent to the impulse response of a one-dimensional system, because the impulse response $h(t, \tau)$ describes the effect of the input signal at time τ on the output signal at time t. For time-invariant systems, the impulse response depends only on the difference $t - \tau$ and is written as $h(t)$.

The propagation of sound in enclosures depends on the distance of sources to the walls. Therefore the signal at a receiver varies also when the sources and the receiver move synchronously within an enclosure. Thus sound propagation in the presence of reflecting surfaces is a shift-variant process. Furthermore there is no preferred direction of sound propagation comparable to the flow of time. This means that the Green's function is not one-sided with respect to space. In summary, contrary to the impulse response of linear and time-invariant systems, the Green's function for the propagation of sound is not one-sided and in general also not shift-invariant. For one-dimensional time-dependent systems, one-sided impulse responses are closely connected to causality. For multidimensional systems the issue of causality is more involved [27].

4.32.4.2.2 The Green's function defined by a differential equation

The Green's function allows to write the sound pressure as a response to the source distribution $P_0(\boldsymbol{x}, \omega)$ by integrating all source locations $\boldsymbol{\xi} \in V$

$$P(\boldsymbol{x}, \omega) = \int_V G(\boldsymbol{x}|\boldsymbol{\xi}, \omega) P_0(\boldsymbol{\xi}, \omega) dV. \tag{32.113}$$

The integration is performed with respect to the volume V which encloses all sources with amplitude $P_0(\boldsymbol{x}, \omega)$, and all locations where the sound pressure $P(\boldsymbol{x}, \omega)$ is of interest. Typically this volume is defined by the walls of a room. Only in reflection-free environments does shift-invariance hold and the integral (32.113) turns into a convolution.

To obtain an equation for the determination of the Green's function, apply the differential operator L from (32.111) to (32.113)

$$L\{P(\boldsymbol{x}, \omega)\} = \int_V L\{G(\boldsymbol{x}|\boldsymbol{\xi}, \omega)\} P_0(\boldsymbol{\xi}, \omega) dV = P_0(\boldsymbol{x}, \omega). \tag{32.114}$$

For the right equality to be valid, the Green's function must satisfy the Helmholtz equation with a spatial delta-impulse as inhomogenity

$$L\{G(\boldsymbol{x}|\boldsymbol{\xi}, \omega)\} = \Delta G(\boldsymbol{x}|\boldsymbol{\xi}, \omega) + \beta^2 G(\boldsymbol{x}|\boldsymbol{\xi}, \omega) = \delta(\boldsymbol{x} - \boldsymbol{\xi}). \tag{32.115}$$

In principle, the Green's function $G(\boldsymbol{x}|\boldsymbol{\xi}, \omega)$ can be obtained from the differential equation (32.115) with the appropriate boundary conditions corresponding the acoustic environment. However, for simple cases (e.g., free field) the Green's function can also be calculated without solving a boundary-value problem (see Section 4.32.4.4).

4.32.4.3 Calculation of the response to interior and exterior sources

When the Green's function is known, it can be used to calculate the response to sources within V using (32.113). However, sources outside of V are not considered by (32.113). But exactly those sources are of interest for the synthesis of sound fields.

This section discusses the complete sound field resulting from sources both within and outside of V. Although the general idea is covered in texts on mathematical physics like [25] a concise self-contained derivation is presented here. Its essential component is the Gauss integral theorem as a generalization of the integration by parts.

4.32.4.3.1 *Solution of the wave equation*

The derivation starts with a so-called divergence expression [25] which contains the Green's function, the unknown sound pressure, and their gradients. To simplify the notation, arguments may be omitted, i.e.,

$$P(\boldsymbol{x}, \omega) = P(\boldsymbol{x}) = P \quad \text{and} \quad G(\boldsymbol{x}|\boldsymbol{\xi}, \omega) = G(\boldsymbol{x}|\boldsymbol{\xi}) = G.$$

Now consider the divergence of $\nabla G \cdot P - G \cdot \nabla P$. Inserting $\pm(\frac{\omega}{c})^2 GP$ yields an expression which contains the differential operator L of the wave equation

$$\begin{aligned} \nabla(\nabla G \cdot P - G \cdot \nabla P) &= \left(\Delta G + \left(\frac{\omega}{c}\right)^2 G\right) \cdot P - G \cdot \left(\Delta P + \left(\frac{\omega}{c}\right)^2 P\right) \\ &= L\{G\} \cdot P - G \cdot L\{P\}. \end{aligned} \tag{32.116}$$

Integration with respect to the volume V on both sides and application of the Gauss integral theorem gives

$$\oint_{\partial V} (\nabla G \cdot P - G \cdot \nabla P) d\boldsymbol{A} = \int_V L\{G\} \cdot P \, dV - \int_V G \cdot L\{P\} dV. \tag{32.117}$$

The surface of the volume V is denoted by ∂V and $\boldsymbol{n}$ is the corresponding unit vector orthogonal to the surface ∂V. The oriented surface element $d\boldsymbol{A}(\boldsymbol{\xi})$ is given in terms of the scalar surface element $dA(\boldsymbol{\xi})$ as $d\boldsymbol{A} = \boldsymbol{n}\, dA$. The dependence on $\boldsymbol{\xi}$ is sometimes omitted in the notation.

Solving for the term with $L\{G\}$ gives

$$\int_V L\{G\} \cdot P \, dV = \int_V G \cdot L\{P\} dV + \oint_{\partial V} (\nabla G \cdot P - G \cdot \nabla P) d\boldsymbol{A}. \tag{32.118}$$

So far the derivation is valid for almost arbitrary functions P and G. Now the special properties of P as solution of the Helmholtz equation according to (32.112) and of G als Green's function according to (32.115) are used. Then the term on the left-hand-side of (32.118) becomes

$$\int_V L\{G(\boldsymbol{x}|\boldsymbol{\xi})\} \cdot P(\boldsymbol{\xi}) dV = \int_V \delta(\boldsymbol{x} - \boldsymbol{\xi}) P(\boldsymbol{\xi}) dV = P(\boldsymbol{x}). \tag{32.119}$$

such that (32.118) turns into

$$P(\boldsymbol{x}) = \int_V G(\boldsymbol{x}|\boldsymbol{\xi}) P_0(\boldsymbol{\xi}) dV + \oint_{\partial V} (\nabla G \cdot P - G \cdot \nabla P) d\boldsymbol{A}. \tag{32.120}$$

Thus the solution $P(\boldsymbol{x}, \omega)$ of the wave equation consists of two components. The first one represents the response to the sources $P_0(\boldsymbol{x}, \omega)$ within V, while the second one considers the behavior of the sound pressure P and of the Green's function G on the boundary. Both components are now discussed in detail.

4.32.4.3.2 *Discussion of the two components of the solution*

Sources within V: When $P(\boldsymbol{x}, \omega)$ and $G(\boldsymbol{x}|\boldsymbol{\xi}, \omega)$ satisfy the same homogenous boundary conditions on ∂V, then the boundary integral in (32.120) vanishes and only the response to the sources within V remains

$$P(\boldsymbol{x}, \omega) = \int_V G(\boldsymbol{x}|\boldsymbol{\xi}, \omega) P_0(\boldsymbol{\xi}, \omega) dV. \tag{32.121}$$

Typical homogeneous boundary conditions arise from the nature of the volume V. If its surface ∂V coincides with the walls of an enclosure then the boundary conditions may be

- $P = 0$ and $G = 0$ (pressure release surface, "soft wall"),
- $\nabla P = 0$ and $\nabla G = 0$ (ideally reflecting surface, "hard wall"),
- $P + \ell_0 \boldsymbol{n}^{\mathrm{T}} \nabla P = 0$ and $G + \ell_0 \boldsymbol{n}^{\mathrm{T}} \nabla G = 0$
  (impedance boundary conditions with suitable wall factor ℓ_0).

On the other hand, the volume V may lie completely in the free field such that free field boundary conditions apply. They correspond to the impedance boundary conditions above with the wall factor replaced by the free field impedance.

Boundary values: Now assume that there are no interior sources $P_0(\boldsymbol{x}, \omega)$, but the sound pressure or its gradient do not vanish at the boundary. Then the sound pressure within V is only determined by the behavior on the boundary

$$P(\boldsymbol{x}, \omega) = \oint_{\partial V} \left(\nabla G(\boldsymbol{x}|\boldsymbol{x}_0, \omega) P(\boldsymbol{x}_0, \omega) - G(\boldsymbol{x}|\boldsymbol{x}_0, \omega) \nabla P(\boldsymbol{x}_0, \omega) \right) d\boldsymbol{A}(\boldsymbol{x}_0). \tag{32.122}$$

Note that ∂V denotes the boundary, $\boldsymbol{x}_0$ is a location on the boundary and $d\boldsymbol{A}(\boldsymbol{x}_0)$ an infinitesimal oriented surface element. Without interior sources, the sound pressure at the boundary can only be caused by sound sources outside of V. The term (32.122) thus reflects the response to exterior sources. It also represents one of several possible forms of the *Kirchhoff-Helmholtz integral equation* which is discussed in more detail in Section 4.32.5.1.

4.32.4.4 Calculation of the Green's function

The determination of the Green's function requires the solution of a boundary value problem consisting of the differential equation (32.115) and the corresponding boundary conditions. These boundary conditions are usually given by the surfaces of the enclosure and are hard to describe analytically.

The most simple case is free field propagation. In reproduction rooms where systems for sound field synthesis are installed this case is approximated by preparing the surfaces to be absorbing.

This section presents the Green's functions for the free field case. At first one-dimensional spatial propagation is considered as a preparation for the three-dimensional case.

4.32.4.4.1 One spatial dimension

In one spatial dimension, the vector $\boldsymbol{x}$ of spatial variables becomes a scalar x and the volume V consists of an interval on the x-axis. For free field propagation, i.e., without reflections at the boundaries of the interval, V covers the complete x-axis. The Laplace operator is simply the second derivative with respect to the space variable $\Delta = \text{div grad} = \partial^2/\partial x^2$.

A monopole source at the location x_0 with a signal $p_0(t)$ is then described by

$$p_0(x,t) = p_0(t)\delta(x - x_0), \tag{32.123}$$

or after Fourier transform with respect to time by

$$P_0(x,\omega) = P_0(\omega)\delta(x - x_0). \tag{32.124}$$

Since the response to this source propagates with the sound speed c way from the source, its Green's function is given by the shift operator

$$G(x|\xi,\omega) = \exp\left(-i\frac{\omega}{c}r(x,\xi)\right). \tag{32.125}$$

Here $r(x,\xi) = |x - \xi|$ is the distance between the location x where the sound pressure is observed and the location of the source ξ. This notation is chosen with respect to the following extension to three dimensions.

The suitability of this Green's function is confirmed by inserting (32.124) into (32.121)

$$\begin{aligned} P(x,\omega) &= \int_V G(x|\xi,\omega)P_0(\omega)\delta(\xi - x_0)\mathrm{d}\xi \\ &= P_0(\omega)G(x|x_0,\omega) = P_0(\omega)\exp\left(-i\frac{\omega}{c}r(x,\xi)\right) \end{aligned} \tag{32.126}$$

and subsequent inverse Fourier transform into the time domain

$$p(x,t) = p_0\left(t - \frac{r(x,\xi)}{c}\right). \tag{32.127}$$

The Green's function (32.125) produces a sound pressure distribution $p(x,t)$ which corresponds to the source signal $p_0(t)$ including a time shift, which results from the distance r to the source and speed of sound c. This is exactly what is expected from wave propagation.

4.32.4.4.2 Three spatial dimensions

The determination of the Green's function of a monopole source in three spatial dimensions can be inferred from the result for one spatial dimension. To this end, a function is required which describes

the sound field of a monopole according to the differential Eq. (32.115). Since the sound waves from a monopole source propagate in a spherical fashion, it is of advantage to use spherical coordinates. Then the Cartesian coordinates (x, y, z) are replaced by the radius, the azimuth and the zenith (or elevation) angles (r, φ, δ). In these coordinates the Laplace operator is given by

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\delta}\frac{\partial}{\partial\delta}\left(\sin\delta\frac{\partial}{\partial\delta}\right) + \frac{1}{r^2\sin^2\delta}\frac{\partial^2}{\partial\varphi^2}. \tag{32.128}$$

Due to the point symmetry of the solution all derivatives with respect to the angles φ and δ are zero and the Laplace operator contains only derivatives with respect to the radius r. Its effect on the sound pressure p takes different forms as

$$\Delta p = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}p\right) = \frac{\partial^2}{\partial r^2}p + \frac{2}{r}\frac{\partial}{\partial r}p = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\,p). \tag{32.129}$$

The wave equation (32.111) is now written with (32.128) and after multiplication with r as

$$\frac{\partial^2}{\partial r^2}(r\,p) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(r\,p) = 0. \tag{32.130}$$

This equation is a wave equation with a scalar space variable r for the unknown $r\,p(r,t)$.

The Green's function for the scalar, i.e., spatially one-dimensional case is already known from (32.125). The relation between the Green's function for the one-dimensional case (designated here by G_{1D}) and the Green's function for the three-dimensional case (designated by G_{3D}) follows as

$$G_{1D}(x|\xi,\omega) = r\,G_{3D}(\boldsymbol{x}|\boldsymbol{\xi},\omega) \tag{32.131}$$

with $r = |\boldsymbol{x} - \boldsymbol{\xi}|$. Inserting (32.125) gives

$$G_{3D}(\boldsymbol{x}|\boldsymbol{\xi},\omega) = \frac{1}{r}\exp\left(-i\frac{\omega}{c}r\right) = \frac{\exp\left(-i\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{\xi}|\right)}{|\boldsymbol{x}-\boldsymbol{\xi}|}. \tag{32.132}$$

The Green's function for the three-dimensional case $G_{3D}(\boldsymbol{x}|\boldsymbol{\xi},\omega)$ describes a spherical wave that propagates from the location $\boldsymbol{\xi}$ with the speed c. Its amplitude decreases with increasing distance r.

4.32.5 Physical foundations of sound field synthesis

This section discusses some of the physical foundations of sound field synthesis. They are exploited for the specific synthesis methods discussed in the subsequent sections.

4.32.5.1 The Kirchhoff-Helmholtz integral equation

Sound field synthesis aims at synthesizing a desired sound field within an extended area V by sources located on the boundary ∂V. In this context the desired sound field is assumed to originate from a *virtual source* and the sources on ∂V are termed as *secondary sources*. Figure 32.15 illustrates the

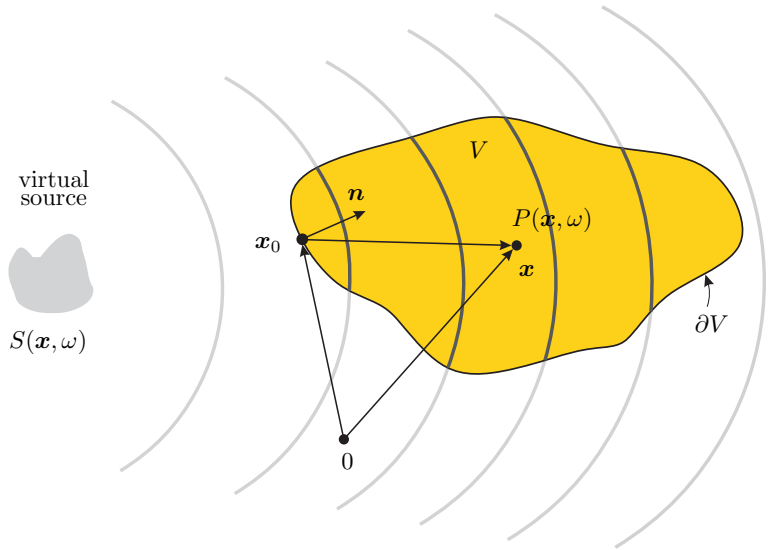


FIGURE 32.15

Illustration of the geometry used to discuss the physical foundations of sound field synthesis. The Kirchhoff-Helmholtz integral equation states that the sound pressure distribution within the area V is uniquely determined by the pressure and the gradient of the pressure of the virtual source $S(\boldsymbol{x}, \omega)$ on the boundary ∂V.

situation. The sound pressure $P(\boldsymbol{x}, \omega)$ within V can be calculated by interpreting the secondary sources on ∂V as inhomogeneous boundary condition. The solution of the homogeneous wave equation for inhomogeneous boundary conditions is provided by (32.122), which is also known as the *Kirchhoff-Helmholtz* integral equation.

Here, the integration of a gradient with respect to an oriented surface element $d\boldsymbol{A}$ can also be expressed by the directional gradient $\partial/\partial\boldsymbol{n}$ as $\nabla P \, d\boldsymbol{A} = \partial P/\partial\boldsymbol{n} \, dA$ and similarly for ∇G. The Kirchhoff-Helmholtz integral equation (32.122) reads then

$$P(\boldsymbol{x}, \omega) = \oint_{\partial V} \left(\frac{\partial}{\partial \boldsymbol{n}} G(\boldsymbol{x}|\boldsymbol{x}_0, \omega) P(\boldsymbol{x}_0, \omega) - G(\boldsymbol{x}|\boldsymbol{x}_0, \omega) \frac{\partial}{\partial \boldsymbol{n}} P(\boldsymbol{x}_0, \omega) \right) dA(\boldsymbol{x}_0). \tag{32.133}$$

The Green's function $G(\boldsymbol{x}|\boldsymbol{x}_0, \omega)$ has to fulfill the homogeneous boundary conditions imposed on ∂V. For sound field synthesis free-field propagation within V is typically assumed, hence that V is free of any objects and that the boundary ∂V does not restrict propagation. The Green's function is then given as the free-field solution of the wave equation and is referred to as *free-field Green's function* $G_0(\boldsymbol{x} - \boldsymbol{x}_0, \omega)$. It can be interpreted as the spatio-temporal transfer function of an acoustic monopole placed at location $\boldsymbol{x}_0$ and its directional gradient as the spatio-temporal transfer function of an acoustic dipole at location $\boldsymbol{x}_0$, whose main axis is parallel to $\boldsymbol{n}$ [22].

Equation (32.133) states that the pressure $P(\boldsymbol{x}, \omega)$ inside V is uniquely determined by the pressure $P(\boldsymbol{x}_0, \omega)$ and its directional gradient on the boundary ∂V. If the Green's function is realized by a *continuous* distribution of appropriately driven monopole and dipole sources that are placed on the

boundary ∂V, the sound field within V can be fully controlled within the volume V. In potential theory the continuous distribution of secondary monopole/dipole sources is also termed as *single/double layer potential* [28].

For sound field synthesis it is desired to synthesize the pressure field $S(\boldsymbol{x}, \omega)$ of the virtual source inside the area V. Concluding the considerations given so far, this can be achieved by a continuous distribution of secondary monopole and dipole sources located on the boundary ∂V of the listening area V, which are driven by the directional gradient and the pressure of the sound field $S(\boldsymbol{x}, \omega)$ of the virtual source, respectively. For the application of the introduced principle, all source contributions of $S(\boldsymbol{x}, \omega)$ are assumed to lie outside of V. The scattering occurring due to listeners inside the listening area does not affect the synthesis in an unfavorable way because the scattered sound field does not depend on the process generating the incident sound field. Hence, the scattering is equal either if the incident sound field is emerging from a source with the characteristics of the virtual source or if the sound field of the virtual source is synthesized by the secondary sources [29]. A listener thus experiences the same scattering in a synthetic sound field like in the corresponding natural one.

4.32.5.2 Monopole only synthesis

It is desirable for a practical implementation to discard one of the two types of secondary sources stated by the Kirchhoff-Helmholtz integral (32.133). Typically the dipole sources are removed, since monopole sources can be realized reasonably well by (commercially available) loudspeakers with closed cabinets. The sound field for monopole only synthesis is expressed by the *synthesis equation*

$$P(\boldsymbol{x}, \omega) = \oint_{\partial V} D(\boldsymbol{x}_0, \omega) G(\boldsymbol{x}|\boldsymbol{x}_0, \omega) dA(\boldsymbol{x}_0), \tag{32.134}$$

where $D(\boldsymbol{x}_0, \omega)$ denotes the strength of the secondary source at position $\boldsymbol{x}_0$, which is denoted as secondary source *driving signal*. Assuming again free-field propagation, $G(\boldsymbol{x}|\boldsymbol{x}_0, \omega)$ in (32.134) can be specialized to the free-field Green's function $G_0(\boldsymbol{x} - \boldsymbol{x}_0, \omega)$. For sound field synthesis, the synthesized pressure $P(\boldsymbol{x}, \omega)$ should be equal to the pressure field of the virtual source $S(\boldsymbol{x}, \omega)$ within the listening area V.

A variety of techniques have been proposed in the past decades to obtain a monopole-only formulation for sound field synthesis. Here, we only discuss those techniques that have led to the well-known approaches that will be outlined later in this chapter. In particular

1. modification of the Green's function employed in the Kirchhoff-Helmholtz integral,
2. the simple source approach, and
3. explicit solution of the single layer potential integral equation.

These three techniques are discussed briefly in the following subsections.

4.32.5.2.1 *Neumann Green's function*

The first addend in the Kirchhoff-Helmholtz integral (32.133) can be suppressed by choosing a Neumann Green's function $G_\mathrm{N}(\boldsymbol{x}|\boldsymbol{x}_0, \omega)$ with

$$\frac{\partial}{\partial \boldsymbol{n}} G_\mathrm{N}(\boldsymbol{x}|\boldsymbol{x}_0, \omega)\Big|_{\boldsymbol{x}_0 \in \partial V} = 0. \tag{32.135}$$

Under this condition the Kirchhoff-Helmholtz integral (32.133) simplifies to

$$P(\boldsymbol{x},\omega) = -\oint_{\partial V} G_N(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}} S(\boldsymbol{x}_0,\omega)\,dA(\boldsymbol{x}_0). \tag{32.136}$$

The explicit form of the Neumann Green's function depends on the geometry of the boundary ∂V. A closed form solution can only be found for rather simple geometries like spheres or planar boundaries [22]. The physical boundary condition (32.135) imposed onto the Neumann Green's function models the boundary ∂V as acoustically rigid. For frequencies that are not equal to the resonance frequencies of the rigid cavity V, the synthesized sound field is equal to the virtual source within V due to the uniqueness of (32.136). Hence, the driving signal is given as the directional gradient of the pressure of the virtual source. A major problem of this approach is that the Neumann Green's function has to be realized by physically existing secondary sources. Other acoustic sources than monopoles or dipoles are typically not available in practice, which renders this approach unfeasible. However, it will be shown later that a sensible approximation of (32.136) forms the basis of WFS.

4.32.5.2.2 Simple source approach and equivalent scattering problem

The second technique, the simple source approach, is based on constructing an acoustic scenario that results in a single layer potential formulation. One way of doing so is to follow the procedure discussed in [22] by constructing two equivalent but spatially disjunct problems. Besides the interior Kirchhoff-Helmholtz integral, given by (32.133), an equivalent exterior Kirchhoff-Helmholtz integral is formulated with the same boundary ∂V but outward pointing normal vector. Note that in this case the exterior and interior regions swap places. It is further assumed that the pressure is continuous and the directional gradient is discontinuous when approaching the boundary ∂V from both sides. These assumptions represent the presence of a secondary source layer. Subtracting the resulting interior from the exterior problem formulation derives

$$P(\boldsymbol{x},\omega) = \oint_{\partial V} D(\boldsymbol{x}_0,\omega) G_0(\boldsymbol{x}-\boldsymbol{x}_0,\omega)\,dA(\boldsymbol{x}_0), \tag{32.137}$$

where $D(\boldsymbol{x}_0,\omega)$ denotes the driving signal of the secondary sources. The continuity conditions for the pressure and its gradient on the boundary ∂V can be interpreted in terms of an equivalent scattering problem [30]. Here, V is replaced by a sound soft object with pressure release boundaries that scatters the field of the virtual source. The driving signal for the simple source approach is then given as

$$D(\boldsymbol{x}_0,\omega) = \frac{\partial}{\partial \boldsymbol{n}} S(\boldsymbol{x}_0,\omega) + \frac{\partial}{\partial \boldsymbol{n}} P_{\mathrm{S}}(\boldsymbol{x}_0,\omega), \tag{32.138}$$

where $P_{\mathrm{S}}(\boldsymbol{x},\omega)$ denotes the pressure of the scattered field in the exterior region. The field in the interior region V matches the field of the virtual source $S(\boldsymbol{x},\omega)$. The insights provided by the simple source approach link the results from acoustic scattering theory to SFS.

4.32.5.2.3 Explicit solution

Equation (32.134) constitutes an integral equation, which can be solved explicitly with respect to the driving signal $D(\boldsymbol{x}_0,\omega)$. According to operator theory [31–33], the integral in (32.134) can be understood as a (compact) Fredholm operator of index zero.

As stated by Fredholm's theory [25], a solution can be found when the secondary source distribution is simply connected and encloses the target volume. The general solution is found by expanding the operator and the virtual sound field into a series of orthogonal basis functions and a comparison of coefficients. It is known from operator theory that the solution is not unique at the eigenfrequencies of the interior homogeneous Dirichlet problem and might be ill-conditioned in practice. Theoretically suitable basis functions can be found for arbitrary simply connected domains V with a smooth boundary. In practice analytic basis functions and solutions are only available for regular geometries like spheres, cylinders and spheroids [15].

The circumstance that the secondary source distribution is required to enclose the target volume constitutes an essential restriction. As will be shown in Sections 4.32.6.3 and 4.32.7, the Fredholm solution can also be applied to non-enclosing geometries of the secondary source distributions. Though, such non-enclosing secondary source distributions exhibit limitations.

4.32.5.3 Three-dimensional synthesis

The particular form of the free-field Green's function and hence the secondary sources depends on the dimensionality of the problem. For a continuous distribution of secondary sources on a surface ∂V surrounding the listening volume V, the three-dimensional free-field Green's function (32.132) is the appropriate choice. It can be approximated reasonably well by loudspeakers in a practical realization. This scenario is termed as *three-dimensional synthesis*.

It has been shown [34–36] that a three-dimensional synthesis can be perfect whereby certain restrictions can apply that are dependent of the geometry of the secondary source distribution.

4.32.5.4 2.5-dimensional synthesis

In many situations the synthesis in a plane only is suitable. This constitutes in principle a two-dimensional scenario. From a physical point of view, the natural choice for the characteristics of the secondary sources used for two-dimensional synthesis would be the elementary solution of the wave equation in two dimensions. The resulting transfer function is given by the two-dimensional free-field Green's function, which can be interpreted as the field produced by a line source [22]. Loudspeakers exhibiting the properties of acoustic line sources are not practical. Using point sources as secondary sources for the synthesis in a plane results in a dimensionality mismatch, therefore such methods are often termed as *2.5-dimensional synthesis*. Ideally the ears of the listeners should be in the same plane like the secondary sources, the *target plane*. However it is well known from WFS and HOA, that even then 2.5-dimensional synthesis techniques suffer from artifacts [35,37]. Most prominent are amplitude deviations with respect to the sound field of the virtual source. Similar artifacts will also be present in other sound field synthesis approaches that aim at correct synthesis in a plane using secondary point sources. These limitations are often discarded in the design of numerical sound field synthesis approaches, a circumstance that can lead to excess regularization since the desired result is physically impossible.

4.32.5.4.1 Model- versus data-based rendering

In general, sound field synthesis can be performed in either a *model-based* or a *data-based* fashion [38]. With model-based objects, all spatial information such as the location of an object (e.g., a sound source) or its radiation properties are described by physical models. A given virtual sound source may

be defined as omnidirectional and being located at a given position that is specified using an appropriate coordinate system. The associated audio signal is then the "input signal" to this source, e.g., a human voice or the performance of a musical instrument captured with a single microphone. Another model-based object could be the virtual venue, the boundary properties of which may be described by an appropriate physical model.

The audio signals associated to data-based objects on the other hand do contain spatial information. Examples are the signals of microphone arrangements that are composed of more than one microphone, e.g., the main microphones of a Stereophonic recording or a spherical or other microphone array. In the case of data-based rendering, a given sound field synthesis system has to determine the loudspeaker driving signals such that the spatial information contained in the input signals is preserved in the presentation. Of course, both model-based as well data-based objects can be apparent in the same scene. A typical scenario is synthesizing a virtual sound sources of given scene model based and then adding reverberation obtained from microphone array measurements [39].

Note that the terms *model-based* and *data-based* auralization initially referred to auralization based on either physical room models or databases of measured room impulse responses [40]. Here, the broader use as explained above is preferred.

We focus in the following on model-based rendering using spatial models for the virtual source. Often applied models in this context are plane waves, point sources or sources with a prescribed complex directivity. For the driving signals derived in the following sections we will consider the synthesis of a plane wave. This special case is suitable to illustrate the basic properties, since other source types can be expressed as a superposition of plane waves [22].

4.32.6 Near-field Compensated Higher Order Ambisonics (NFC-HOA)

Ambisonics is a collective term for a variety of sound field synthesis approaches applied mainly to circular or spherical loudspeaker distributions. Data-based rendering is the traditional rendering technique used in the context of Ambisonics. The expansion coefficients of the desired sound field are extracted from microphone array recordings, transmitted and then used for derivation of the driving signal, e.g., [12,41,42]. This procedure is often referred to as *en- and decoding*.

For ease of illustration we skip the en- and decoding procedure and consider the direct derivation of driving functions for the synthesis of virtual sources following the model-based synthesis paradigm. We review the concept of NFC-HOA based on the analytic formulation presented in [35,43].

4.32.6.1 Outline

Near-field Compensated Higher Order Ambisonics (NFC-HOA), which is also known as *Ambisonics with distance coding*, and related techniques [12,32,35,44,45] base on the explicit solution of the monopole only synthesis Eq. (32.134) by means of decomposing the quantities involved in (32.134)—$P(\boldsymbol{x}, \omega)$, $D(\boldsymbol{x}_0, \omega)$, and $G(\boldsymbol{x}|\boldsymbol{x}_0, \omega)$—into orthogonal basis functions (Section 4.32.5.2). For the considered circular and spherical secondary source distributions, these orthogonal basis functions are given by the surface spherical harmonics and circular harmonics respectively (refer to Sections 4.32.3.5 and 4.32.3.4). Exploitation of the orthogonality of the basis functions leads to a comparison of coefficients that allows for determining the driving function.

The synthesis equation (32.134) can be interpreted as a generalized spatial convolution along the secondary source contour. By identifying the appropriate convolution theorem, the spatial convolution in (32.134) is turned into a scalar multiplication of given expansion coefficients. It is then straightforward to calculate the driving signal, if the expansion coefficients of the desired sound field and the spatio-temporal transfer function of the secondary sources are known. As mentioned in Section 4.32.5.2, the above described procedure theoretically provides also the solution for geometries other than spherical or circular. However, the applicability is restricted due to the complexity involved.

Common variants of Ambisonics and HOA can be derived from the general theory discussed above by varying the model used for the virtual source or the secondary sources. In traditional Ambisonics it is typically assumed that the secondary sources and the virtual source can be modeled as plane waves in the center of the secondary source arrangement [46]. This results in driving functions that are simple amplitude panning laws (refer to Section 4.32.1.2).

4.32.6.2 Spherical secondary source distributions

The synthesis equation (32.134) for an acoustically transparent spherical secondary source distribution with radius R centered around the coordinate origin is given by [30,35,47]

$$P(\boldsymbol{x},\omega) = \int_{S_R^2} D\left(\boldsymbol{x}_0,\omega\right) G\left(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{x}_0)\boldsymbol{\eta}_2,\omega\right) R^2 d\boldsymbol{x}_0. \tag{32.139}$$

$\boldsymbol{\eta}_2 = [0\ 0\ R]^T$ denotes the north pole of the spherical surface S_R^2 and $\boldsymbol{x}_0 = R\,[\cos\alpha_0 \sin\beta_0\ \sin\alpha_0 \sin\beta_0\ \cos\beta_0]^T$ a location on S_R^2. $\boldsymbol{g}(\boldsymbol{x}_0)$ is a rotation matrix the explicit expression of which is waived here for convenience. Refer to the corresponding rotation in the treatment of circular secondary source distributions in Section 4.32.6.3 for an explicit example of such a rotation matrix.

$G\left(\boldsymbol{x}, \boldsymbol{\eta}_2, \omega\right)$ denotes the spatio-temporal transfer function of the secondary source located at $\boldsymbol{\eta}_2 = [0\ 0\ R]^T$. The factor R^2 arises in (32.139) due to the fact that S_R^2 is of radius R and not 1. Refer to Figure 32.16 for an illustration of the setup.

Note that (32.139) implies that the spatio-temporal transfer function of the secondary sources is invariant with respect to rotation around the coordinate origin. In simple words, all secondary sources need to exhibit similar radiation properties and need to be oriented appropriately. For the considered free-field conditions, this circumstance does not constitute an essential restriction.

Following the procedure outlined in Section 4.32.5.2.3 requires that $P(\boldsymbol{x},\omega)$, $D\left(\boldsymbol{x},\omega\right)$, and $G\left(\boldsymbol{x},\omega\right)$ are expanded into appropriate orthogonal basis functions in order to allow for a comparison of the coefficients of the according decomposition. For the geometry under consideration these orthogonal basis functions are given by the surface spherical harmonics discussed in Section 4.32.3.5. This procedure can indeed be straightforwardly applied yielding the desired result. However, we will apply the equivalent procedure from [35], which will later be shown to be applicable also for non-enclosing distributions of secondary sources.

Equation (32.139) can be interpreted as a convolution along the surface of a sphere in which case the convolution theorem [48,49]

$$\mathring{P}_n^m(r,\omega) = R^2 \sqrt{\frac{4\pi}{2n+1}}\, \mathring{D}_n^m(\omega) \cdot \mathring{G}_n^0(r,\omega) \tag{32.140}$$

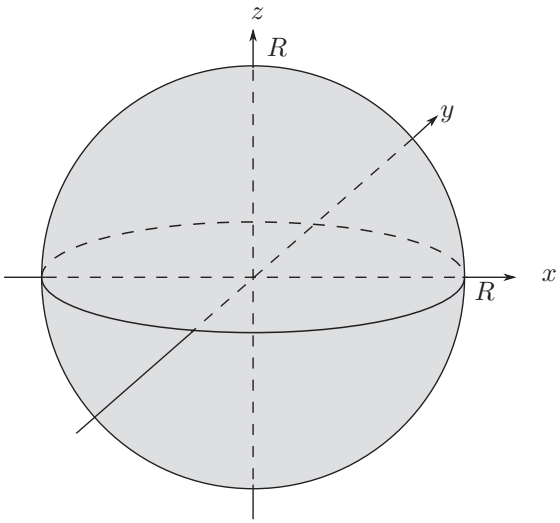


FIGURE 32.16

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

applies, which relates the spherical harmonics expansion coefficients (refer to (32.101)) of the involved quantities via a scalar multiplication. Note that $\mathring{G}_n^0(r,\omega)$ represents the expansion coefficients of $G\left(\boldsymbol{x},\boldsymbol{\eta}_2,\omega\right)$, i.e., of the spatio-temporal transfer function of the secondary source located at the north pole of the sphere.

The asymmetry of the convolution theorem (32.140), $\mathring{P}_n^m(r,\omega)$ vs. $\mathring{G}_n^0(r,\omega)$ is a consequence of the definition of (32.139) as left convolution. An according convolution theorem for right convolutions exists [49].

Rearranging (32.140) yields

$$\mathring{D}_n^m(\omega) = \frac{1}{R^2}\sqrt{\frac{2n+1}{4\pi}}\frac{\mathring{P}_n^m(r,\omega)}{\mathring{G}_n^0(r,\omega)}. \tag{32.141}$$

When introducing the explicit expressions for the coefficients $\mathring{P}_n^m(r,\omega)$ and $\mathring{G}_n^0(r,\omega)$ given by (32.105) into (32.141),

$$\mathring{D}_n^m(\omega) = \frac{1}{R^2}\sqrt{\frac{2n+1}{4\pi}}\frac{\breve{P}_n^m(\omega)\cdot j_n\left(\frac{\omega}{c}r\right)}{\breve{G}_n^0(\omega)\cdot j_n\left(\frac{\omega}{c}r\right)}, \tag{32.142}$$

it can be seen that the parameter r appears both in the numerator as well as in the denominator in (32.142) in the spherical Bessel function $j_n\left(\frac{\omega}{c}r\right)$. It can be shown that the spherical Bessel functions cancel out except for specific situations [35]. It can indeed happen that (32.142) can be undefined for $j_n\left(\frac{\omega}{c}r\right) = 0$ and $\frac{\omega}{c}r \neq 0$. These cases represent resonances of the spherical cavity that can not be controlled by the secondary source distribution. This lack of controllability has not been reported to be a restriction is practice.

It may be assumed that all Bessel functions in (32.141) cancel out yielding

$$\mathring{D}_n^m(\omega) = \frac{1}{R^2}\sqrt{\frac{2n+1}{4\pi}}\frac{\breve{P}_n^m(\omega)}{\breve{G}_n^0(\omega)}. \tag{32.143}$$

In order that (32.143) holds, $\breve{G}_n^0(\omega)$ may not exhibit zeros. This requirement is fulfilled for the three-dimensional free-field Green's function [35].

The secondary source driving function $D(\alpha, \beta, \omega)$ can be composed from its coefficients $\mathring{D}_n^m(\omega)$ using (32.104) to be [30,35,47]

$$D(\alpha, \beta, \omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \underbrace{\frac{1}{R^2}\sqrt{\frac{2n+1}{4\pi}}\frac{\breve{P}_n^m(\omega)}{\breve{G}_n^0(\omega)}}_{=\mathring{D}_n^m(\omega)} Y_n^m(\beta, \alpha). \tag{32.144}$$

In practical applications, the summation in (32.144) can not be performed over an infinite number of addends but has to be truncated. The choice of the summation limits primarily has impact on the artifacts arising in practice. This circumstance will be discussed more in detail in Section 4.32.6.4 conjunction with circular NFC-HOA.

4.32.6.3 Circular secondary source distributions

The specialization of (32.134) to a circular distribution of secondary sources in the horizontal plane and centered around the coordinate origin is given by

$$P(\boldsymbol{x}, \omega) = \int_0^{2\pi} D\left(\boldsymbol{x}_0, \omega\right) G\left(\boldsymbol{x}, \boldsymbol{g}(\alpha_0)\boldsymbol{\eta}_1, \omega\right) R\, d\alpha_0, \tag{32.145}$$

where R denotes the radius of, and $\boldsymbol{x}_0 = R[\cos\alpha_0 \;\; \sin\alpha_0 0]^T$ a location on, the circular secondary source distribution S_R^1. $\boldsymbol{\eta}_1 = [R\; 0\; 0]^T$ denotes that point on the distribution where $\alpha_0 = 0$. $\boldsymbol{g}(\alpha_0)$ is a rotation matrix given by

$$\boldsymbol{g}(\alpha_0) = \begin{bmatrix} \cos\alpha_0 & -\sin\alpha_0 & 0 \\ \sin\alpha_0 & \cos\alpha_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$G(\boldsymbol{x}, \boldsymbol{\eta}_1, \omega)$ in (32.145) denotes the spatio-temporal transfer function of the secondary source located at $\boldsymbol{\eta}_1$. Refer to Figure 32.17 for a sketch of the setup.

Equation (32.145) can be interpreted as a convolution along a circle and thus the convolution theorem [22]

$$\mathring{P}_m(r, \omega) = 2\pi R\; \mathring{D}_m(\omega)\mathring{G}_m(r, \omega) \tag{32.146}$$

holds, which relates the Fourier series expansion coefficients (i.e., the circular harmonics coefficients, refer to (32.70)) of the involved quantities (Section 4.32.3.4). Note that the relation between the Fourier series expansion coefficients and the spherical harmonics expansion coefficients is given by (32.107). $\mathring{G}_m(r, \omega)$ represents the expansion coefficients of $G\left(\boldsymbol{x}, \boldsymbol{\eta}_1, \omega\right)$.

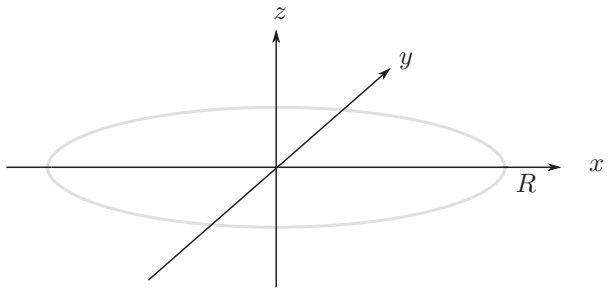


FIGURE 32.17

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

Equation (32.146) can be solved for the coefficients $\mathring{D}_m(\omega)$ of the driving function $D(\alpha_0, \omega)$. The latter can then be composed from its coefficients via (32.73). Unlike with spherical secondary source distributions, the driving function $D(\alpha_0, \omega)$ is generally dependent on the radius r. As a consequence, the synthesized sound field will only be correct on a circle around the center of the coordinate origin. Deviations arise at all other locations. This is a typical case of 2.5-dimensional synthesis as described in Section 4.32.5.4.

Referencing the driving function to the origin of the coordinate system has shown to be most convenient. After some mathematical manipulation, the driving function $D(\alpha, \omega)$ can finally be shown to be [35]

$$D(\alpha, \omega) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi R} \frac{\breve{P}^m_{|m|}(\omega)}{\breve{G}^m_{|m|}(\omega)} e^{im\alpha}. \tag{32.147}$$

Note that $\breve{P}^m_{|m|}(\omega)$ and $\breve{G}^m_{|m|}(\omega)$ denote the spherical harmonics expansion coefficients $\breve{P}^m_n(\omega)$ and $\breve{G}^m_n(\omega)$ of $P(\,\cdot\,)$ and $G(\,\cdot\,)$ for $n = |m|$.

The coefficients $\breve{P}^m_{|m|}(\omega)$ for a plane wave with incidence angle $(\theta_{\mathrm{pw}}, \varphi_{\mathrm{pw}})$ can be deduced from (32.99) to be

$$\breve{P}^m_{|m|}(\omega) = Y^{m*}_{|m|}(\theta_{\mathrm{pw}}, \varphi_{\mathrm{pw}}). \tag{32.148}$$

For omnidirectional secondary sources

$$\breve{G}^m_{|m|}(\omega) = \frac{1}{4\pi} i^{-|m|-1} \frac{\omega}{c} h^{(2)}_{|m|}\left(\frac{\omega}{c} R\right) Y^m_{|m|}\left(\frac{\pi}{2}, 0\right), \tag{32.149}$$

where $h^{(2)}_{|m|}$ denotes the $|m|$th-order spherical Hankel function of second kind [22].

Assuming that a desired plane wave $P(\boldsymbol{x}, \omega)$ propagates in the horizontal plane (i.e., $\theta_{\mathrm{pw}} = \pi/2$) and carries the signal $\hat{S}(\omega)$, the driving function can be determined to be [50]

$$D_{\mathrm{pw}}(\alpha_0, \omega) = \hat{S}(\omega) \sum_{m=-\infty}^{\infty} \frac{2i^{|m|+1}}{R \frac{\omega}{c} h^{(2)}_{|m|}\left(\frac{\omega}{c} R\right)} e^{im(\alpha_0 - \varphi_{\mathrm{pw}})}, \tag{32.150}$$

by exploiting the fact that the associated Legendre functions $P_n^m(\cdot)$, which are contained in the spherical harmonics $Y_n^m(\cdot)$ (see (32.87)), never vanish when their order m equals their degree n [24, Eq. (2.1.50)] so that they cancel out.

4.32.6.4 Spatial sampling and application example

Continuous secondary source distributions as discussed so far can not be implemented in practice but discrete setups of a finite number of loudspeakers have to be used. In the following, we briefly discuss the effects of this spatial sampling of the secondary source distribution. Detailed discussions can be found in [2,51,52].

It can be shown that an equiangular sampling leads to repetitions of the Fourier coefficients of the driving function as [52]

$$\mathring{D}_{m,\mathrm{S}}(\omega) = \sum_{\mu=-\infty}^{\infty} \mathring{D}_{m+\mu L}(\omega), \tag{32.151}$$

where L denotes the number of secondary sources employed and $\mathring{D}_{m,\mathrm{S}}(\omega)$ denotes the Fourier series coefficients of the sampled driving function.

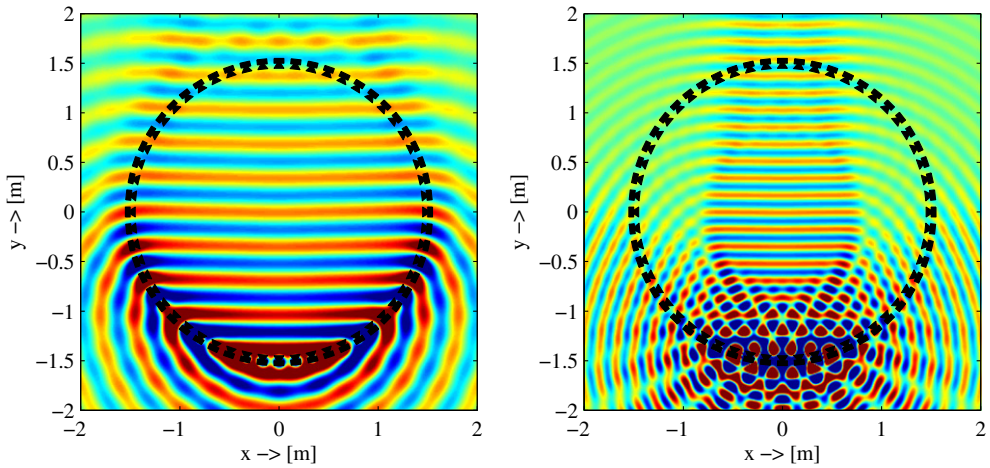
At first stage, the continuous driving function is not bandlimited with respect to the expansion order m, i.e., according to (32.147) the summation over the coefficients $\mathring{D}_m(\omega)$ has to be performed from $m = -\infty$ to $m = \infty$. As a consequence, the repetitions that are apparent in (32.151) overlap and interfere.

In order to avoid such overlap the continuous driving function is typically spatially bandlimited, i.e., the summation in (32.147) is calculated for $m = -L/2+1$ to $m = L/2-1$ for even L and accordingly for odd L. Obviously, such a spatial band limitation causes a loss of information. This is illustrated in Figure 32.18, which shows a virtual plane wave impinging from $\varphi_{\mathrm{pw}} = -\pi/2$ synthesized by a circular distribution of $L = 56$ monopole secondary sources. The geometry is chosen equal to a system built by the authors. At low frequencies, neither the bandwidth limitation nor the spectral repetitions impair the synthesized sound field as depicted in Figure 32.18a. The undesired amplitude decay, which is characteristic for 2.5-dimensional synthesis, is apparent (Section 4.32.5.4).

At higher frequencies though, the imposed bandwidth limitation causes a concentration of the energy of the desired component of the sound field around the center of the secondary source distribution. This is evident in Figure 32.18b. Outside this artifact-free zone, artifacts arise, which are a consequence of the spectral repetitions. The size of the artifact-free zone decreases linearly with increasing frequency.

Of course, it is also possible to choose a spatial bandlimit such that is significantly higher than the one chosen in the Ambisonics context. This does indeed essentially change the properties of the arising artifacts. It has been shown in [52] that a large spatial bandwidth leads to artifacts that are similar to those arising in Wave Field Synthesis as discussed in Section 4.32.8.

As apparent from Figure 32.18b, considerable artifacts arise even at moderate frequencies. The perceptual consequences of these artifacts have not been investigated in detail so far. However, typical loudspeaker setups used for NFC-HOA have shown to provide improved properties for a central listening position when compared to traditional stereophonic techniques [53].



(a) $f_{\mathrm{pw}} = 1$ kHz **(b)** $f_{\mathrm{pw}} = 2$ kHz

FIGURE 32.18

Sound field $P(\boldsymbol{x}, \omega)$ synthesized by a circular distribution of $L = 56$ secondary monopole sources and of radius $R = 1.50$ m when driven with 2.5-dimensional NFC-HOA. The virtual sound field is a monochromatic plane wave of frequency f_{pw} and with incidence angle $\varphi_{\mathrm{pw}} = -\pi/2$. Click Video clips 1 and 2 to see the animation.

4.32.6.5 Extensions to basic principle

The above presented approach has been extended in various ways to enable the synthesis of complex virtual sound fields such as that of focused sources [54–56] or of sound sources with complex radiation properties [57–60] or to enable the employment of non-spherical enclosing secondary source distributions [33] and non-omnidirectional secondary sources [61,62] and active listening room compensation [63,64].

4.32.7 Spectral division method (SDM)

The Spectral Division Method employs the explicit solution to the synthesis equation (32.134) for planar and linear distributions of secondary sources. The following section provides a brief overview of the theory as presented in [34,65,66].

4.32.7.1 Planar secondary source distributions

Consider the synthesis equation (32.134) and assume a secondary source distribution ∂V that consists of a disk V_0 and a hemisphere V_{hemi} of radius r_{hemi} as depicted in Figure 32.19 [22]. As $r_{\mathrm{hemi}} \to \infty$, the disk V_0 turns into an infinite plane and the volume under consideration turns into a half-space. The latter

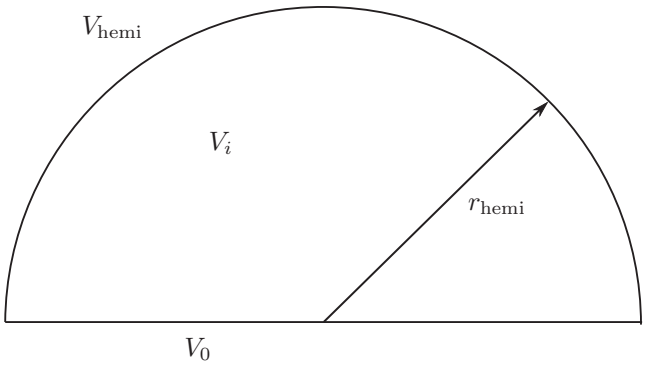


FIGURE 32.19

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

is referred to as *target half-space*. Additionally, the Sommerfeld radiation condition[2] is invoked, i.e., it is assumed that there are no contributions to the desired sound field to be synthesized that originate from infinity so that only the planar part of the boundary needs to be considered. We additionally allow for the synthesis of plane waves that propagate into the target half-space. The Rayleigh integral can be used to prove that such plane waves can indeed be synthesized by the considered secondary source distribution [66].

As a consequence, arbitrary sound fields that are source-free in the target half-space and that satisfy the Sommerfeld radiation condition (as well as plane waves) may now be described by an integration over the infinite plane V_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 Figure 32.20.

The formulation of the synthesis equation (32.134) for an infinite uniform planar secondary source distribution is then given by [34,65,66]

$$P(\boldsymbol{x},\omega) = \iint_{-\infty}^{\infty} D(\boldsymbol{x}_0,\omega)\cdot G(\boldsymbol{x}-\boldsymbol{x}_0,\omega)dx_0dz_0. \tag{32.152}$$

with $\boldsymbol{x}_0 = [x_0\,0\,z_0]^T$. $G(\boldsymbol{x},\omega)$ denotes the spatial transfer function of a secondary source located in the origin of the coordinate system. The term $G(\boldsymbol{x}-\boldsymbol{x}_0,\omega)$ in (32.152) implies that the spatio-temporal transfer function of the secondary sources is invariant with respect to translation along the secondary source contour [22]. In other words, all secondary sources exhibit equal radiation properties and are orientated accordingly. Note the resemblance of (32.152) to the first Rayleigh integral [22].

[2]The Sommerfeld radiation condition can be interpreted as a boundary condition at infinity. It assures that no energy originates from infinity.

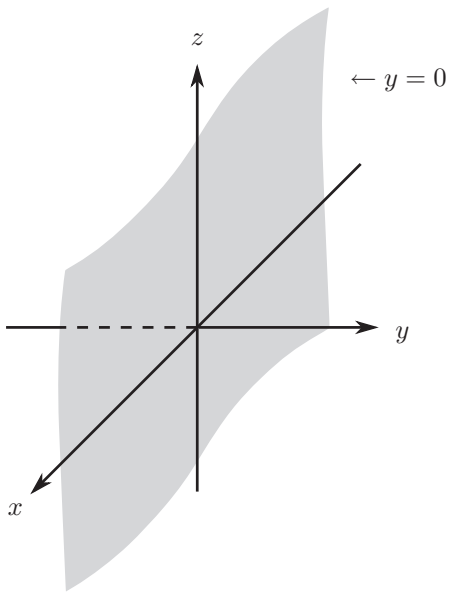


FIGURE 32.20

Illustration of the setup of a planar secondary source situated along the x-z-plane. The secondary source distribution is indicated by the gray 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.

Equation (32.152) essentially constitutes a two-dimensional convolution along the spatial dimensions x and z respectively. This fact is revealed when (32.152) is rewritten as [34,66]

$$\begin{aligned} P(\boldsymbol{x}, \omega) &= \iint_{-\infty}^{\infty} D\left([x_0\ 0\ z_0]^T, \omega\right) G\left([x\ y\ z]^T - [x_0\ 0\ z_0]^T, \omega\right) dx_0 dz_0 \\ &= \iint_{-\infty}^{\infty} D(x_0, 0, z_0, \omega) G(x - x_0, y, z - z_0, \omega) dx_0 dz_0 \\ &= D\left(\boldsymbol{x}\,|_{y=0}, \omega\right) *_x *_z G(\boldsymbol{x}, \omega), \end{aligned} \tag{32.153}$$

where the asterisk $*_i$ denotes convolution with respect to the indexed spatial dimension [67]. Thus, the convolution theorem [67]

$$\tilde{P}\left(k_x, y, k_z, \omega\right) = \tilde{D}\left(k_x, k_z, \omega\right) \cdot \tilde{G}\left(k_x, y, k_z, \omega\right) \tag{32.154}$$

holds, which relates the involved quantities in wavenumber domain with respect to k_x and k_z. Refer to Section 4.32.3.3 for a discussion of the Fourier transform with respect to space.

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

$$\tilde{D}\left(k_x, k_z, \omega\right) = \frac{\tilde{P}\left(k_x, y, k_z, \omega\right)}{\tilde{G}\left(k_x, y, k_z, \omega\right)}. \tag{32.155}$$

In order that (32.155) holds, $\tilde{G}(k_x, y, k_z, \omega)$ may not exhibit zeros. This is indeed fulfilled e.g., for monopole secondary sources [34].

Applying an inverse spatial Fourier transform with respect to k_x and k_z on (32.155) yields [34,65,66]

$$D(x, z, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{\tilde{P}\left(k_x, y, k_z, \omega\right)}{\tilde{G}\left(k_x, y, k_z, \omega\right)} e^{i\left(k_x x + k_z z\right)} dk_x dk_z. \tag{32.156}$$

From (32.155) it is obvious that the driving function is essentially yielded by a division in the spatial frequency domain. The presented approach is therefore referred to as *Spectral Division Method* (SDM).

Equation (32.156) 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 (32.156). It is shown in [2,34,66] that y does indeed cancel out provided that the target half-space is free of virtual sound sources. $D(x, z, \omega)$ is thus independent of the location of the receiver.

4.32.7.2 Linear secondary source distributions

Despite the simple driving function for planar secondary source distributions, 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 [68]. Assuming a linear secondary source distribution of infinite length, the situation may be interpreted as a reduced formulation of the setup treated in Section 4.32.7.1. For convenience, the secondary source distribution is assumed to be along the x-axis (thus $\boldsymbol{x}_0 = [x_0\ 0\ 0]^T$, refer to Figure 32.21).

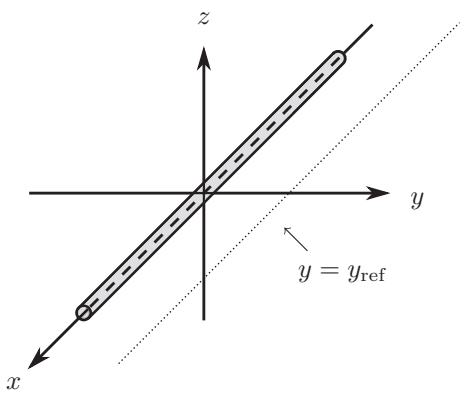


FIGURE 32.21

Illustration of the setup of a linear secondary source distribution situated along the x-axis. The secondary source distribution is indicated by the gray 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 line indicates the reference line (see text).

The specialization of the synthesis equation (32.134) to such a linear distribution of secondary sources along the x-axis is given by

$$P(\boldsymbol{x}, \omega) = \int_{-\infty}^{\infty} D(\boldsymbol{x}_0, \omega) G(\boldsymbol{x} - \boldsymbol{x}_0, \omega) dx_0, \tag{32.157}$$

where $\boldsymbol{x}_0 = [x_0 \; 0 \; 0]^T$ and $G(\boldsymbol{x}, \omega)$ denotes the spatio-temporal transfer function of that secondary source located in the coordinate origin. Equation (32.157) implies again that the spatio-temporal transfer function of the secondary sources is invariant with respect to translation along the secondary source contour.

For the sake of simplicity it is assumed that the listeners' ears are located in that half of the x-y-plane that contains the positive part of the y-axis ($z = 0$, $y > 0$). Refer to [34] for a generalization. Equation (32.157) essentially constitutes a spatial convolution of the driving function $D(\boldsymbol{x}, \omega)$ with the spatial transfer function $G(\boldsymbol{x}, \omega)$ of the secondary source at the coordinate origin, whereby the convolution takes place along the x-axis. Again, the convolution theorem (32.154) of the Fourier transformation can be applied, though in the present case exclusively with respect to k_x. Explicitly,

$$\tilde{P}(k_x, y, \omega) = \tilde{D}(k_x, \omega) \tilde{G}(k_x, y, \omega), \tag{32.158}$$

where k_x denotes the wavenumber in x-direction. Equation (32.158) can now be solved straightforwardly with respect to the driving function $\tilde{D}(k_x, \omega)$ in wavenumber domain. $D(\boldsymbol{x}, \omega)$ is then obtained by applying an inverse spatial Fourier transformation (32.44) given by [34,65,66]

$$D(x, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{P}(k_x, y, z, \omega)}{\tilde{G}(k_x, y, z, \omega)} e^{ik_x x} dk_x. \tag{32.159}$$

In order for (32.159) to hold, $\tilde{G}(k_x, y, z, \omega)$ may not exhibit zeros. This is indeed fulfilled e.g., for monopole secondary sources as can be deduced from (32.162).

As discussed in detail in [34], the driving function (32.159) has to be referenced to a line parallel to the x-axis, which is then the only location where the synthesized sound field is exact. As with circular distributions of secondary sources (Section 4.32.6.3), the present situation constitutes 2.5-dimensional synthesis as discussed in Section 4.32.5.4.

Since we are aiming at the synthesis in the horizontal plane, we reference the driving function (32.159) to $z = 0$ and $y = y_{\text{ref}}$ as indicated in Figure 32.21. The referenced driving function $D(x, \omega)$ is then given by

$$D(x, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{P}(k_x, y_{\text{ref}}, 0, \omega)}{\tilde{G}(k_x, y_{\text{ref}}, 0, \omega)} e^{ik_x x} dk_x. \tag{32.160}$$

When choosing a plane wave with propagation direction $\left(\theta_{\text{pw}} = \frac{\pi}{2}, \varphi_{\text{pw}}\right)$ carrying the signal $\hat{S}(\omega)$ as desired sound field, then $\tilde{P}(k_x, y_{\text{ref}}, 0, \omega)$ can be obtained from (32.46) as

$$\tilde{P}(k_x, y_{\text{ref}}, 0, \omega) = \hat{S}(\omega) 2\pi \delta(k_x - k_{\text{pw},x}) e^{ik_{\text{pw},y} y_{\text{ref}}}. \tag{32.161}$$

$\tilde{G}(k_x, y_{\text{ref}}, 0, \omega)$ for omnidirectional secondary sources is given by [34]

$$\tilde{G}(k_x, y, z, \omega) = -\frac{i}{4} H_0^{(2)}\left(\sqrt{\left(\frac{\omega}{c}\right)^2 - {k_x}^2} \sqrt{y^2 + z^2}\right) \forall 0 \leq |k_x| < \left|\frac{\omega}{c}\right|, \tag{32.162}$$

where $H_0^{(2)}(\,\cdot\,)$ denotes the zeroth-order Hankel function of second kind [69].

Inserting (32.161) and (32.162) into (32.160) yields [34]

$$D_{\mathrm{pw}}(x_0, \omega) = \hat{S}(\omega) \frac{4i e^{i \frac{\omega}{c} y_{\mathrm{ref}} \sin \varphi_{\mathrm{pw}}}}{H_0^{(2)}(\frac{\omega}{c} y_{\mathrm{ref}} \sin \varphi_{\mathrm{pw}})} e^{i \frac{\omega}{c} x_0 \cos \varphi_{\mathrm{pw}}}. \tag{32.163}$$

4.32.7.3 Spatial sampling and application example

As for NFC-HOA, we briefly discuss the effects of spatial sampling of the secondary source distribution. An equidistant sampling with distance Δx between the secondary sources can be modeled by multiplying the driving function (32.163) with a series of spatial Dirac pulses. This results in spectral repetitions in the spatio-temporal frequency domain [70]

$$\tilde{D}_S(k_x, \omega) = 2\pi \sum_{\mu=-\infty}^{\infty} \tilde{D}\left(k_x - \frac{2\pi}{\Delta x}\mu, \omega\right). \tag{32.164}$$

Typical driving functions (e.g., for plane or spherical waves) are not bandlimited with respect to the spatial frequency k_x. The spectral repetitions indicated in (32.164) can therefore overlap and leak into the baseband. The latter is constituted by the continuous driving function, i.e., (32.164) evaluated for $\mu = 0$.

Reformulating (32.158) considering the sampled driving function $\tilde{D}_S(k_x, \omega)$ reads

$$\tilde{P}_S(k_x, y, \omega) = \tilde{D}_S(k_x, \omega) \tilde{G}(k_x, y, \omega). \tag{32.165}$$

It is evident from (32.165), the synthesized sound field $\tilde{P}_S(k_x, y, \omega)$ is given by the driving function $\tilde{D}_S(k_x, \omega)$ weighted by the secondary sources' transfer function $\tilde{G}(k_x, y, \omega)$. Using a simplified model, it may be assumed that $\tilde{G}(k_x, y, \omega)$ is spatially lowpass [51]. This means that it does not alter $\tilde{D}_S(k_x, \omega)$ when $|k_x|$ is smaller but heavily attenuates $\tilde{D}_S(k_x, \omega)$ for large $|k_x|$.

At low angular frequencies ω (and thus at low time frequencies f), the spectral repetitions indicated in (32.164) to not corrupt the baseband and they are attenuated by the spatial lowpass property of $\tilde{G}(k_x, y, \omega)$ [51]. As a consequence, the synthesized sound field is not impaired at low time frequencies f. Refer to Figure 32.22a for an example.

At higher frequencies f, the spectral repetitions due to spatial sampling do corrupt the baseband and therefore also the synthesized sound field. The latter is then composed of the desired sound field superposed by additional undesired wave fronts [51,70]. Refer to Figure 32.22b for an example. For practical setups, the frequency above which considerable artifacts arise is relatively low. When the synthesis of a plane wave is considered, the highest frequency f that can be synthesized without or with only moderate corruption can be calculated via [37]

$$f = \frac{c}{\Delta x (1 + |\cos \varphi_{\mathrm{pw}}|)}. \tag{32.166}$$

For the setup depicted in Figure 32.22, this frequency is approximately 1000 Hz.

Recall that the audible bandwidth typically exceeds 16 kHz. A significant part of the spectrum is thus corrupted. Though, the human auditory system does not seem to be very sensitive to spatial aliasing artifacts for stationary scenarios [71]. The scenario illustrated in Figure 32.22 results in only minor perceptual impairment even when the entire audible frequency range is employed.

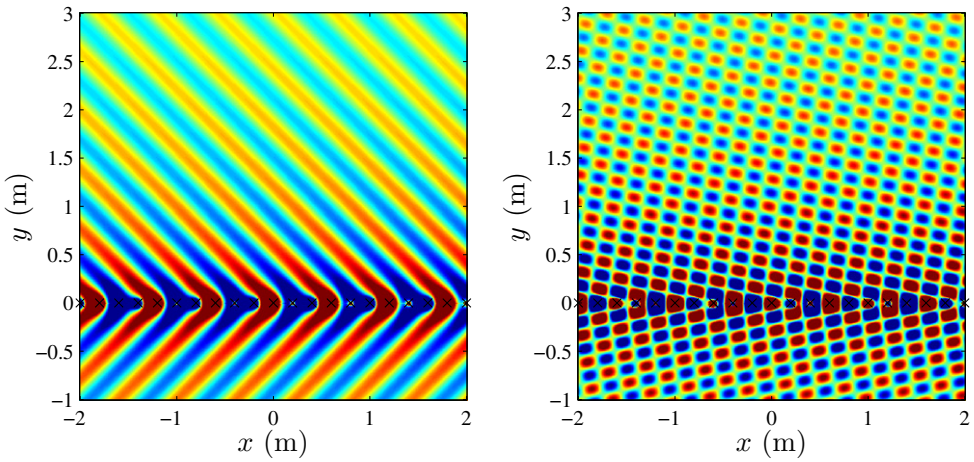


FIGURE 32.22

Snapshot of the sound field $P(\boldsymbol{x}, \omega)$ synthesized by a linear distribution of secondary monopole sources with a spacing of $\Delta x = 0.2$ m driven with SDM. The virtual sound field is a monochromatic plane wave of frequency f_{pw} impinging from direction $\varphi_{\mathrm{pw}} = -(3/4)\pi$ referenced to $y_{ref} = 1$ m. Click Video clips 3 and 4 to see the animation.

4.32.7.4 Approximate solution for non-planar and non-linear secondary source distributions

As pointed out in [30], it is helpful to interpret sound field synthesis 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 [9]. The surface of the scattering object is divided into a region that 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* [9]. The convexity is required in order to avoid re-entry of the scattered sound field.

For such small wave lengths, any arbitrary simply connected convex enclosing secondary source distribution may also be assumed to be locally plane. Consequently, when the driving function (32.156) for planar secondary source distributions is applied in such a scenario, a high-frequency approximation of the driving function is obtained when only those secondary sources are driven that are located in

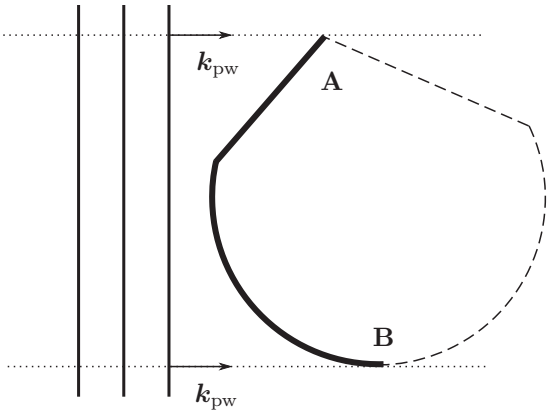


FIGURE 32.23

Secondary source selection for a virtual plane wave with propagation direction $\boldsymbol{k}_{pw}$. Thick solid lines indicate the area that is illuminated by the virtual sound field. 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 $\boldsymbol{k}_{pw}$ and pass the secondary source distribution in a tangent-like manner. In case $\boldsymbol{A}$ tapering may be applied, in case $\boldsymbol{B}$ not.

that region that is illuminated by the virtual sound field. A similar argumentation may be deployed in conjunction with linear secondary source distributions.

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. A further analysis of this inaccuracy may be found in [51,72].

The illuminated area can be straightforwardly determined via geometrical considerations as indicated in Figure 32.23. For a virtual plane wave, the illuminated area is bounded by two lines parallel to the propagation vector $\boldsymbol{k}_{pw}$ of the plane wave passing the secondary source distribution in a tangent-like manner.

If the proper tangent on the boundary of the illuminated area is not parallel to $\boldsymbol{k}_{pw}$ or is not defined (like the boundary of a planar distribution of finite size), a degenerated problem is considered (case $\boldsymbol{A}$ in Figure 32.23). 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. Tapering is a well-established technique in the context of WFS [73].

It has been shown that the illuminated area does not need to be smooth. Corners are also possible with only little additional error introduced [37]. This is also evident from Figure 32.24a, which shows the synthesis of a virtual plane wave by a rectangular distribution of monopoles driven with SDM. The secondary source driving function was deduced from (32.163) via an appropriate translation and rotation of the coordinate system. Since this scenario constitutes 2.5-dimensional synthesis, the synthesized sound field exhibits an undesired amplitude decay.

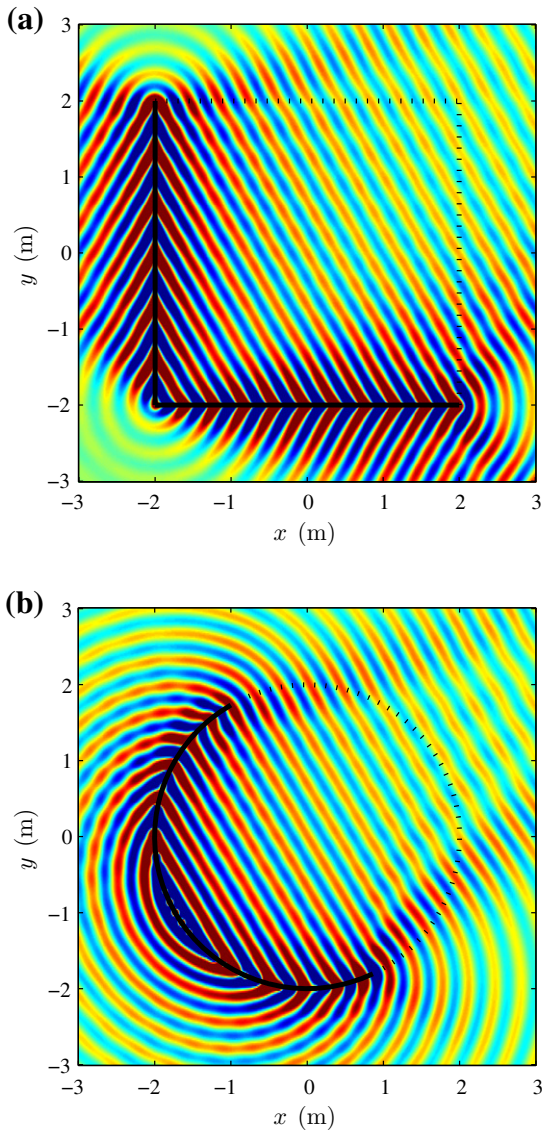


FIGURE 32.24

A cross-section through the horizontal plane of the sound pressure $P_{\text{pw}}(\boldsymbol{x}, \omega)$ synthesized by different secondary monopole distribution driven with SDM synthesizing a virtual plane wave of $f_{\text{pw}} = 1000$ Hz and unit amplitude with incidence angle $\varphi_{\text{pw}} = \frac{7}{6}\pi$ referenced to $y_{\text{ref}} = 1.0$ m. Solid lines indicate the illuminated area; dotted lines indicate the shadowed area. (a) Rectangular secondary distribution and (b) circular secondary source distribution. Click Video clips 5 and 6 to see the animation.

Figure 32.24b depicts the synthesis of a virtual plane wave by a circular distribution of monopoles driven with SDM. This circular distribution may be interpreted as a combination of linear sections of infinitesimal length. Due to the different geometry of the secondary source contour, the amplitude decay of the synthesized sound field is slightly different to the one in Figure 32.24a.

4.32.8 Wave Field Synthesis (WFS)

Wave Field Synthesis constitutes an approximation of the approach to monopole only synthesis that is based on the application of a Neumann Green's function. The physical foundation of this approach is given in Section 4.32.5.2. This section outlines its application in the context of WFS and presents application examples.

4.32.8.1 Outline

The concept of WFS has initially been developed for linear distributions of secondary sources [74]. It bases on a sensible approximation of (32.136). Similar arguments as given in Section 4.32.7 for the SDM, can be used to derive a representation of (32.136) for planar geometries. The required Neumann Green's function $G_{\mathrm{N}}(\boldsymbol{x}|\boldsymbol{x}_0, \omega)$ for this specialized geometry is given by two times the free-field Green's function $G_{\mathrm{N}}(\boldsymbol{x}|\boldsymbol{x}_0, \omega) = 2G_0(\boldsymbol{x} - \boldsymbol{x}_0, \omega)$. The resulting integral is known as Rayleigh's first integral equation [22]. The initial concept of WFS is implicitly based on a specialization of Rayleigh's first integral formula to 2.5-dimensional synthesis. This specialization is derived by degenerating the planar secondary source distribution to a linear distribution using a stationary phase approximation [37,73,74].

WFS has later on been generalized to arbitrary convex secondary source distributions, which may even only partly enclose the receiver area [36,73]. This generalization can be deduced also from (32.136) as is illustrated in the following. The Neumann Green's function in (32.136) has to fulfill homogeneous Neumann boundary conditions imposed on ∂V. Consequently, reflections are included that are caused by the rigid boundary. Since (32.136) provides a unique solution (discarding the eigenfrequencies of the rigid cavity), these reflections will be compensated for by the driving function. As mentioned above, secondary sources with the characteristics of a Neumann Green's function are generally not available for complex geometries. Hence, it is desirable to use secondary sources with the characteristics of a free-field Green's function. WFS is based on an approximation of (32.136) by

1. replacing the Neumann Green's function by the free-field Green's function,
2. limiting the integration path, and
3. prescribing a convex secondary source distribution.

In many practical situations the free-field Green's function can be seen as a far-field/high-frequency approximation of the Neumann Green's function when considering only the interior V of the secondary source contour [72]. Hence, the Neumann Green's function used in (32.136) can be realized by secondary sources exhibiting monopole characteristics. However, as discussed above, the driving function inherently copes with the reflections caused by the Neumann Green's function. These reflections are not present when using monopole sources and consequently do not need to be compensated. This can be accounted for by not those driving secondary sources that compensate for these reflections, by taking

care that only those secondary sources are active that are located in the area that is illuminated by the virtual sound field. This can be formulated by introducing the window function $a(\boldsymbol{x}_0)$ into (32.136)

$$P(\boldsymbol{x}, \omega) = - \oint_{\partial V} \underbrace{2a(\boldsymbol{x}_0) \frac{\partial}{\partial \boldsymbol{n}} S(\boldsymbol{x}_0, \omega)}_{D(\boldsymbol{x}_0, \omega)} G_0(\boldsymbol{x} - \boldsymbol{x}_0, \omega) dA(\boldsymbol{x}_0). \tag{32.167}$$

An alternative motivation for $a(\boldsymbol{x}_0)$ is outlined in Section 4.32.7.4.

The geometry of the boundary ∂V has to be restricted to convex contours in order to avoid that contributions reenter the listening area V. Equation (32.167) constitutes an approximation of (32.136), which has been shown to be of special interest for sound reproduction. Equation (32.167) states that the driving function for WFS is given as

$$D(\boldsymbol{x}_0, \omega) = 2a(\boldsymbol{x}_0) \frac{\partial}{\partial \boldsymbol{n}} S(\boldsymbol{x}_0, \omega). \tag{32.168}$$

The window function for selection of the active secondary sources for a plane wave as virtual source is given as [75]

$$a_{\mathrm{pw}}(\boldsymbol{x}_0) = \begin{cases} 1, \text{ if } \langle \boldsymbol{n}_{\mathrm{pw}}, \boldsymbol{n}(\boldsymbol{x}_0) \rangle > 0, \\ 0, \text{ otherwise.} \end{cases} \tag{32.169}$$

Equation (32.168) is valid for arbitrary convex secondary source contours ∂V. It depends only on local quantities. This is contrary to NFC-HOA and SDM where the driving functions are restricted to a particular geometry and their dependence is non-local. The next section illustrates the practical application of WFS to three- and 2.5-dimensional synthesis.

4.32.8.2 Three-dimensional synthesis

The Green's function used in the synthesis equation (32.167) determines the characteristics of the secondary sources. The specific form of the free-field Green's function depends on the dimensionality of the problem. The three-dimensional free-field Green's function is given by (32.132) which can be interpreted as the field of a point source with monopole characteristics located at the position $\boldsymbol{x}_0$.

Three-dimensional WFS can be realized by surrounding the listening volume V by a continuous distribution of point sources placed on the boundary ∂V. These secondary sources are driven by the secondary source driving function (32.168). The driving function is given by the directional gradient of the virtual source's sound field and the window function $a(\boldsymbol{x}_0)$. Hence, the explicit form of the driving function depends on the virtual source and the geometry of the secondary source distribution. The driving function for a plane wave carrying the signal $\hat{S}_{\mathrm{pw}}(\omega)$ is determined by the direction gradient of the sound field of a plane wave (32.29) and the window function (32.169) as

$$D_{\mathrm{pw,3D}}(\boldsymbol{x}_0, \omega) = 2a_{\mathrm{pw}}(\boldsymbol{x}_0) \frac{\boldsymbol{n}_{\mathrm{pw}}^T \boldsymbol{n}(\boldsymbol{x}_0)}{c} i\omega \hat{S}_{\mathrm{pw}}(\omega) e^{i \frac{\omega}{c} \boldsymbol{n}_{\mathrm{pw}}^T \boldsymbol{x}_0}. \tag{32.170}$$

A time-domain version of the driving function (32.170) is useful to derive an efficient implementation of WFS. Inverse Fourier transformation of (32.170) yields

$$d_{\mathrm{pw,3D}}(\boldsymbol{x}_0, t) = 2a_{\mathrm{pw}}(\boldsymbol{x}_0) \frac{\boldsymbol{n}_{\mathrm{pw}}^T \boldsymbol{n}(\boldsymbol{x}_0)}{c} \frac{d}{dt} \hat{s}_{\mathrm{pw}} \left(t - \frac{\boldsymbol{n}_{\mathrm{pw}}^T \boldsymbol{x}_0}{c} \right), \tag{32.171}$$

where the differentiation theorem of the Fourier transformation was used. Equation (32.171) states that the driving signal for a plane wave can be computed efficiently in the time-domain by weighting the derivative of the time-shifted source signal $\hat{s}_{\mathrm{pw}}(t)$. However, the differentiation of the virtual source signal may also be performed by filtering the signal by a filter with $i\omega$-characteristic.

A planar secondary source distribution is the basic building block of a cuboid shaped reproduction system. A planar distribution will be discussed in detail in the sequel. The closed contour integral (32.136) over the surface ∂V can be degenerated to an integral over an infinite plane. In brief, this degeneration is achieved by splitting the closed contour ∂V into a planar boundary and a half-sphere. The integration over the half-sphere can be omitted by applying the Sommerfeld radiation condition [22]. This procedure was outlined in Section 4.32.7.1 in conjunction with SDM.

It will be assumed in the following, without loss of generality, that the secondary source distribution is located on the xz-plane at $y = 0$ as depicted in Figure 32.20. Other cases can be regarded as simple translation or rotation of this special case.

The synthesized sound field for a planar distribution of secondary point sources on the xz-plane is given as

$$P(\boldsymbol{x}, \omega) = -\iint_{-\infty}^{\infty} D_{3\mathrm{D}}(\boldsymbol{x}_0, \omega) G_{0,3\mathrm{D}}(\boldsymbol{x}|\boldsymbol{x}_0, \omega) dx_0 dz_0, \tag{32.172}$$

with $\boldsymbol{x}_0 = [x_0\ 0\ z_0]^T$. Equation (32.172) is known as the first Rayleigh integral. The synthesized sound field $P(\boldsymbol{x}, \omega)$ will be mirrored at the secondary source distribution as a consequence of the Neumann boundary condition (32.135). Hence, the reproduced wave field is only correct in one of the two half-volumes separated by the secondary source distribution. The direction of the normal vector $\boldsymbol{n}$ specifies the considered half-volume.

The synthesized sound field for a planar continuous distribution of infinite size will exactly match the wave field of the virtual source within the listening area. This can be proven by inserting the driving functions into the synthesis equation (32.172). Artifacts will occur for other geometries of the secondary source distribution. This is due to the fact that the derived Neumann Green's function only fulfills the required Neumann boundary condition exactly in this special case.

4.32.8.3 2.5-Dimensional synthesis

Typical realizations of WFS use secondary source distributions that are located on the boundary ∂V of a planar listening area V. As already discussed in Section 4.32.5.4, this constitutes a 2.5-dimensional scenario when point sources are used a secondary sources. Since WFS is not based on an explicit solution of the underlying mathematical formulation this fact has to be taken into account explicitly.

Traditionally a stationary phase approximation has been applied to (32.167) in order to derive the driving function for 2.5-dimensional reproduction [37]. This procedure results in a spectral correction and a listener dependent amplitude correction. A more instructive approach is followed here. A large-argument approximation of the Hankel function is used to derive the following far-field approximation $\left(\frac{\omega}{c}\,|\boldsymbol{x} - \boldsymbol{x}_0| \gg 1\right)$ of the two-dimensional Green's function

$$G_{2\mathrm{D}}(\boldsymbol{x} - \boldsymbol{x}_0, \omega) \approx \sqrt{\frac{2\pi\,|\boldsymbol{x} - \boldsymbol{x}_0|}{i\,\frac{\omega}{c}}}\,\frac{1}{4\pi}\,\frac{e^{-i\,\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x}_0|}}{|\boldsymbol{x} - \boldsymbol{x}_0|}. \tag{32.173}$$

It can be concluded from (32.173) that a line source can be approximated in the far-field by a point source the spectrum and amplitude of which are corrected. The amplitude correction depends on the observation point $\boldsymbol{x}$. Hence, this correction holds strictly only for one reference point $\boldsymbol{x}_{\text{ref}}$. Calculating the directional gradient of a plane wave and introducing (32.173) into (32.168) results in the driving signal for the 2.5-dimensional synthesis of a plane wave

$$D_{\text{pw,2.5D}}(\boldsymbol{x}_0, \omega) = 2 w_{\text{pw}}(\boldsymbol{x}_0) \hat{S}_{\text{pw}}(\omega) \sqrt{i \frac{\omega}{c}} c_{\text{2.5D}}(\boldsymbol{x}_0) \boldsymbol{n}_{\text{pw}}^T \boldsymbol{n}(\boldsymbol{x}_0) e^{-i \frac{\omega}{c} \boldsymbol{n}_{\text{pw}}^T \boldsymbol{x}_0}, \tag{32.174}$$

where $\boldsymbol{n}_{\text{pw}} = [\cos\theta_{\text{pw}} \;\; \sin\theta_{\text{pw}}]^T$ denotes the normal vector of the plane wave and $c_{\text{2.5D}}(\boldsymbol{x}_0)$ a geometry dependent amplitude factor for 2.5-dimensional reproduction. The amplitude correction may be derived from (32.174). However, it has been shown [37] that the amplitude can be corrected also for a reference line that is parallel to the secondary source distribution.

4.32.8.4 Spatial sampling and application example

As with any other method, practical implementations of WFS systems will not consist of a continuous secondary source distribution but of a limited number of secondary sources placed at discrete positions. Two types of artifacts may emerge from spatial truncation and discretization: (1) truncation and (2) spatial aliasing artifacts. Truncation artifacts can be improved by applying a spatial window function (tapering window) to the driving function [37].

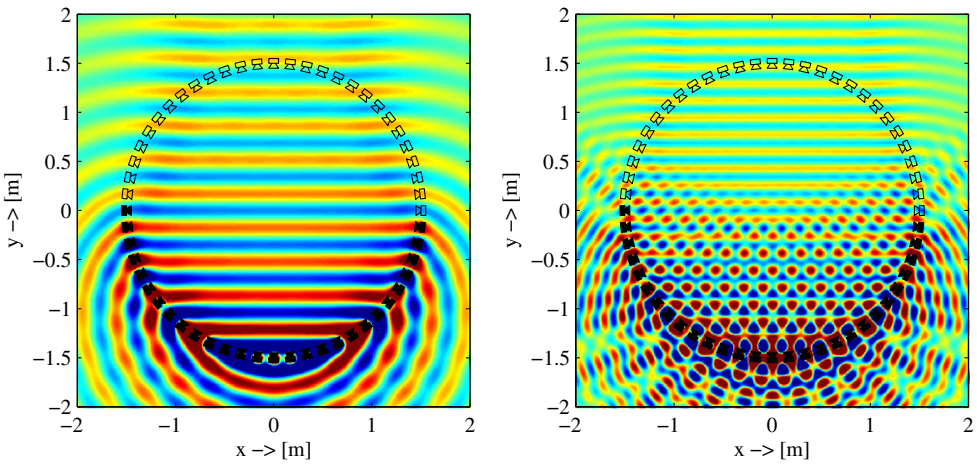
The spatial sampling can be modeled in a similar manner as shown for NFC-HOA and the SDM in Sections 4.32.6.4 and 4.32.7.3, respectively. An detailed analysis of the spatial sampling artifacts for linear and circular WFS systems can be found e.g., in [70].

The same circular geometry as in Section 4.32.6.4 for NFC-HOA is used as application example in order to facilitate the comparison of both approaches. As for NFC-HOA and SDM this constitutes a 2.5-dimensional scenario. Figure 32.25 shows the synthesized sound field for two different frequencies of the considered virtual plane wave. For 1 kHz, the synthesized sound field shows no obvious artifacts. However, when comparing Figure 32.25a with Figure 32.18a some inaccuracies can be observed close to the secondary sources for WFS. This is due to the approximations applied for the derivation of the driving function in WFS. Figure 32.25b shows the situation for 2 kHz. In comparison with Figure 32.18b it is clearly visible that WFS does not exhibit a pronounced area where the synthesis is more accurate. Sampling artifacts are rather evenly distributed over the receiver area, especially at very high frequencies. The amplitude decay in the synthesized plane wave, due to the 2.5-dimensional approach, is clearly visible in Figure 32.25a.

4.32.8.5 Extensions to basic principle

Similarly to NFC-HOA, the basic WFS principle has been extended to the synthesis of various virtual source types such as focused sources [37], sources with complex radiation properties [76], or moving sources [77–79], as well as compensation for loudspeaker directivity [80], auralization of microphone array recordings [39,81], and active listening room compensation [64,82].

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/B978-0-12-396501-1.00032-7.



(a) $f_{\mathrm{pw}} = 1$ kHz **(b)** $f_{\mathrm{pw}} = 2$ kHz

FIGURE 32.25

Sound field $P(\boldsymbol{x}, \omega)$ synthesized by a circular distribution of $L = 56$ secondary monopole sources and of radius $R = 1.50$ m when driven with by 2.5-dimensional WFS. The virtual sound field is a monochromatic plane wave of frequency f_{pw} and with incidence angle $\varphi_{\mathrm{pw}} = -\pi/2$. Filled loudspeaker symbols indicate active secondary sources, hollow loudspeaker symbols indicate in active ones. Video clips 7 and 8 to see the animation.

Acknowledgments

The authors gratefully acknowledge the help of Peter Steffen in proof reading the manuscript and the contribution of numerous figures by Paolo Annibale. The comments by Franz Zotter were helpful in consolidating the final version and are much appreciated.

Relevant Theory: Signal Processing Theory

See Vol. 1, Chapter 2 Continuous-Time Signals and Systems
See Vol. 1, Chapter 3 Discrete-Time Signals and Systems

References

[1] Jens Blauert, Rudolf Rabenstein, Providing surround sound with loudspeakers: a synopsis of current methods, Arch. Acoust. 36 (1) (2012) 1–14.
[2] Jens Ahrens, Analytic Methods of Sound Field Synthesis, Springer, Berlin, 2012.
[3] J. Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization, MIT Press, 1996.
[4] W.G. Gardner, 3-D audio using loudspeakers, PhD Thesis, Massachusetts Institute of Technology, 1997.

[5] V. Pulkki, Spatial sound generation and perception by amplitude panning techniques, PhD Thesis, Helsinki University of Technology, 2001.
[6] R.J. Alexander, Michael Gerzon: Beyond Psychoacoustics, Dora Media Productions, 2008.
[7] G. Theile, On the localisation in the superimposed soundfield, PhD Thesis, Technische Universität Berlin, 1980.
[8] M. Jessel, Acoustique Théorique—Propagation et Holophonie [Theoretical acoustics—propagation and holophony], Masson et Cie, Paris, 1973 (text in French).
[9] D. Colton, R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, second ed., Springer, Berlin, 1998.
[10] O. Kirkeby, P.A. Nelson, Reproduction of plane wave sound fields, J. Acoust. Soc. Am. 94 (5) (1993) 2992–3000.
[11] J. Hannemann, K.D. Donohue, Virtual sound source rendering using a multipole-expansion and method-of-moments approach, J. Audio Eng. Soc. (JAES) 56 (6) (2008) 473–481.
[12] J. Daniel, Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia, PhD Thesis, Université Paris 6, 2000.
[13] F.M. Fazi, P.A. Nelson, The ill-conditioning problem in sound field reconstruction, in: 123rd AES Convention, New York, USA, Audio, Engineering Society (AES), 2007.
[14] D.T. Blackstock, Fundamentals of Physical Acoustics, John Wiley & Sons, 2000.
[15] Philip M. Morse, Herman Feshbach, Theoretical Acoustics, McGraw-Hill Science/Engineering/Math, 1953.
[16] Philip M. Morse, Theoretical Acoustics, Princeton University Press, Uno Ingard, 1987.
[17] A.D. Pierce, Acoustics an Introduction to its Physical Principles and Applications, Acoustical Society of America, 1991.
[18] L.J. Ziomek, Fundamentals of Acoustic Field Theory and Space Time Signal Processing, CRC Press, Boca Raton, 1995.
[19] Jens Blauert, Ning Xiang, Acoustics for Engineers, Springer-Verlag, Berlin, 2009.
[20] Yang-Hann Kim, Sound Propagation, an Impedance Based Approach, Wiley & Sons (Asia) Pte Ltd, Singapore, 2010.
[21] Fridolin P. Mechel (Ed.), Formulas of Acoustics, Springer, Berlin, 2002.
[22] E.G. Williams, Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography, Academic Press, 1999.
[23] George B. Arfken, Hans J. Weber, Mathematical Methods for Physicists, Academic Press, Amsterdam, Weber, 2001.
[24] N.A. Gumerov, R. Duraiswami, Fast Multipole Methods for the Helmholtz Equation in Three Dimensions, Elsevier, Amsterdam, 2004.
[25] Richard Courant, David Hilbert, Methods of Mathematical Physics, Wiley-VCH, Weinheim, Germany, 2009.
[26] Rudolf Rabenstein, Jens Blauert, in: Schallfeldsynthese mit Lautsprechern II—Signalverarbeitung, ITG-Fachtagung Sprachkommunikation, Bochum, 2010.
[27] Alfred Fettweis, Sankar Basu, Multidimensional causality and passivity of linear and nonlinear systems arising from physics, Multidim. Sys. Signal Process. 22 (1) (2011) 5–25.
[28] D.L. Colton, R. Kress, Integral Equation Methods in Scattering Theory, Wiley, New York, 1983.
[29] J. Ahrens, S. Spors, On the scattering of synthetic sound fields, in: 130th AES Convention Audio Engineering Society (AES), May 2011.
[30] F.M. Fazi, P.A. Nelson, R. Potthast, Analogies and differences between three methods for sound field reproduction, in: Ambisonics Symposium, Graz, Austria, June 2009.
[31] J. Giroire, Integral equation methods for the Helmholtz equation, Integr. Equat. Oper. Theory 5 (1) (1982) 506–517.

[32] F.M. Fazi, P.A. Nelson, J.E.N. Christensen, J. Seo, Surround system based on three dimensional sound field reconstruction, in: 125th AES Convention Audio, Engineering Society (AES), San Fransisco, USA, 2008.
[33] S. Spors, J. Ahrens, Towards a theory for arbitrarily shaped sound field reproduction systems, J. Acoust. Soc. Am. 123 (5) (2008) 3930.
[34] J. Ahrens, S. Spors, Sound field reproduction using planar and linear arrays of loudspeakers, IEEE Trans. Audio Speech Lang. Process. 18 (8) (2010) 2038–2050. <http://dx.doi.org/10.1109/TASL.2010.2041106>.
[35] J. Ahrens, S. Spors, An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions, Acta Acust. unit. Acustica 94 (6) (2008) 988–999.
[36] S. Spors, R. Rabenstein, J. Ahrens, The theory of wave field synthesis revisited, in: 24th AES Convention Audio Engineering Society (AES), May 2008.
[37] E.N.G. Verheijen, Sound reproduction by wave field synthesis, PhD Thesis, Delft University of Technology, 1997.
[38] R. Rabenstein, S. Spors, Multichannel sound field reproduction, in: J. Benesty, M. Sondhi, Y. Huang (Eds.), Springer Handbook on Speech Processing and Speech Communication, Springer, Berlin, 2007, pp. 1095–1114.
[39] E. Hulsebos, Auralization using wave field synthesis, PhD Thesis, Delft University of Technology, 2004.
[40] U. Horbach, M. Boone, Practical implementation of data-based wave field reproduction system, in: 108th Convention of the AES, France, Paris, February 2000.
[41] S. Moreau, J. Daniel, S. Bertet, 3D sound field recording with higher order ambisonics—objective measurements and validation of a 4th order spherical microphone, in: 120th Convention of the AES, France, Paris, May 2006.
[42] M.A. Gerzon, Width-heigth sound reproduction, J. Audio Eng. Soc. (JAES) 21 (1973) 2–10.
[43] Y.J. Wu, T. Abhayapala, Soundfield reproduction using theoretical continuous loudspeaker, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, USA, 2008.
[44] M.A. Poletti, Three-dimensional surround sound systems based on spherical harmonics, J. AES 53 (11) (2005) 1004–1025.
[45] Franz Zotter, Hannes Pomberger, Markus Noisternig, Energy-preserving ambisonic decoding, Acta Acust. Unit. Acust. 98 (2012) 37–47.
[46] J.S. Bamford, S. Vanderkooy, Ambisonics sound for us, in: 99th AES Convention Audio Engineering Society (AES), October 1995.
[47] F. Zotter, H. Pomberger, M. Frank, An alternative ambisonics formulation: modal source strength matching and the effect of spatial aliasing, in: 126th Convention of the AES, Munich, Germany, May 2009.
[48] J. Ahrens, S. Spors, A modal analysis of spatial discretization of spherical loudspeaker distributions used for sound field synthesis, IEEE Trans. Audio Speech Lang. Process. 20 (9) (2012) 2564–2574.
[49] J.R. Driscoll, D.M. Healy, Computing Fourier transforms and convolutions on the 2-sphere, Adv. Appl. Math. 15 (2) (1994) 202–250.
[50] J. Ahrens, S. Spors, Analytical driving functions for higher order ambisonics, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, Nevada, March/April 2008.
[51] J. Ahrens, The single-layer potential approach applied to sound field synthesis including cases of non-enclosing distributions of secondary sources, Technische Universität Berlin, Doctoral Dissertation, 2010.
[52] S. Spors, J. Ahrens, A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling, in: 125th AES Convention Audio Engineering Society (AES), October 2008.
[53] S.S. Bertet, Formats audio 3D hiérarchiques: caractérisation objective et perceptive des systèmes ambisonics d'Ordres supérieurs, PhD Thesis, Institut National des Sciences Appliquees de Lyon, 2009.
[54] J. Ahrens, S. Spors, Focusing of virtual sound sources in higher order ambisonics, in: 124th Convention of the AES, Amsterdam, The Netherlands, May 2008, p. paper 7378.

[55] D. Menzies, Calculation of near-field head related transfer functions using point source representations, in: Ambisonics Symposium, Graz, Austria, June 2009, pp. 23–28.
[56] F. Fazi, Sound field reproduction, PhD Thesis, University of Southampton, 2010.
[57] D. Menzies, Ambisonic synthesis of complex sources, JAES 55 (10) (2007) 864–876.
[58] J. Ahrens, S. Spors, Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics, in: 123rd Convention of the AES, New York, NY, October 2007.
[59] M.A. Poletti, F.M. Fazi, P.A. Nelson, Sound reproduction systems using variable-directivity loudspeakers, J. Acoust. Soc. Am. 129 (3) (2011) 1429–1438.
[60] M. Poletti, F.M. Fazi, P.A. Nelson, Sound-field reproduction systems using fixed-directivity loudspeakers, J. Acoust. Soc. Am. 127 (6) (2010) 3590–3601.
[61] J. Ahrens, S. Spors, An analytical approach to 2.5D sound field reproduction employing circular distributions of non-omnidirectional loudspeakers, in: 17th European Signal Processing Conference (EUSIPCO), Glasgow, Scotland, August 2009, pp. 814–818.
[62] J. Ahrens, S. Spors, An analytical approach to 3D sound field reproduction employing spherical distributions of non-omnidirectional loudspeakers, in: IEEE International Symposium on Control, Communication and Signal Processing (ISCCSP), Limassol, Cyprus, March 2010, pp. 1–5.
[63] T. Betlehem, T.D. Abhayapala, Theory and design of sound field reproduction in reverberant rooms, JASA 117 (4) (2005) 2100–2111.
[64] S. Spors, R. Rabenstein, H. Buchner, W. Herbordt, Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering, JASA 122 (1) (2007) 354–369.
[65] J. Ahrens, S. Spors, Reproduction of a plane-wave sound field using planar and linear arrays of loudspeakers, in: Third IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing, St. Julians, Malta, March 2008.
[66] J. Ahrens, S. Spors, Applying the ambisonics approach on planar and linear arrays of loudspeakers, in: 2nd International Symposium on Ambisonics and Spherical Acoustics, Paris, France, May 2010.
[67] B. Girod, R. Rabenstein, A. Stenger, Signals and Systems, John Wiley & Sons, Chichester, UK, 2001.
[68] D. de Vries, Wave Field Synthesis, AES Monograph, AES, New York, 2009.
[69] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover Publications, 1972.
[70] S. Spors, R. Rabenstein, Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis, in: 120th AES Convention Audio, Engineering Society (AES), Paris, France, May 2006.
[71] H. Wittek, Perceptual differences between wavefield synthesis and stereophony, PhD Thesis, University of Surrey, 2007.
[72] J. Ahrens, S. Spors, On the secondary source type mismatch in wave field synthesis employing circular distributions of loudspeakers, in: 127th AES Convention Audio, Engineering Society (AES), New York, USA, October 2009.
[73] E.W. Start, Direct sound enhancement by wave field synthesis, PhD Thesis, Delft University of Technology, 1997.
[74] A.J. Berkhout, A holographic approach to acoustic control, J. Audio Eng. Soc. 36 (1988) 977–995.
[75] S. Spors, Extension of an analytic secondary source selection criterion for wave field synthesis, in: 123th AES Convention Audio, Engineering Society (AES), New York, USA, October 2007.
[76] E. Corteel, Synthesis of directional sources using wave field synthesis, possibilities and limitations, EURASIP J. Adv. Signal Process. Article ID 90509 (2007).
[77] A. Franck, A. Gräfe, T. Korn, M. Strauß, Reproduction of moving virtual sound sources by wave field synthesis: an analysis of artifacts, in: 32nd International Conference of the AES, Hillerød, Denmark, September 2007.

[78] Andreas Franck, Efficient Algorithms for Arbitrary Sample Rate Conversion with Application to Wave Field Synthesis, Universitätsverlag Ilmenau, Ilmenau, 2012.
[79] J. Ahrens, S. Spors, Reproduction of moving virtual sound sources with special attention to the Doppler effect, in: 124th Convention of the AES, The Netherlands, Amsterdam, May 2008.
[80] D. de Vries, Sound reinforcement by wavefield synthesis: adaptation of the synthesis operator to the loudspeaker directivity characteristics, JAES 44 (12) (1996) 1120–1131.
[81] M. Cobos, S. Spors, J. Ahrens, On the use of small microphone arrays for wave field synthesis auralization, in: 45th Conference of the AES, Helsinki, Finland, March 2012.
[82] P.-A. Gauthier, A. Berry, Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: theory, JASA 119 (5) (2006) 2721–2737.